EUROREGIO BNAM2022

Joint Acoustics Conference

9th-11th of May, Aalborg, Denmark

Conference Proceedings

Editors: Flemming Christensen, Rodrigo Ordoñez

ISBN: 978-87-995400-5-1 ISSN: 2245-4365

Welcome by European Acoustics Association (EAA)

On behalf of the EAA Board, it is a great pleasure to welcome all participants to EUROREGIO 2022, the first face-to-face EAA European Acoustics Conference after the pandemic lockdown in 2020 and 2021 and an excellent opportunity to reconnecting during fruitful presentations and discussions. The congress is merged with the bi-annual Baltic-Nordic Acoustic Meetings being held in the Nordic countries since 1954 and thus BNAM 2022 being the 35th since the start. The EUROREGIO concept is much younger with the conference in 2022 being the 5th since the first being held in 2010 in Ljubljana/Slovenia.

We appreciate very much the great efforts done by the Danish Acoustical Society (DAS) and Aalborg University in organizing the congress, and not least for keeping the "physical" setup as a strong goal and preparing major parts of the conference, while many people still considered the pandemic situation and development too uncertain for a face-to-face conference without virtual attendance in parallel. – We also give our acknowledgments to the Nordic Acoustics Association (NAA) for their institutional support and cooperation.

In society and research, the United Nations Sustainable Development Goals (UN SDGs) have got a rapidly developing attention, and in line with this strong trend and necessity, the conference organizers have encouraged focus on the UN SDGs throughout EUROREGIO 2022.

In the weekend preceding the conference, an EAA Training School is held in the Campus of Aalborg University with distinguished experts in the Acoustics field as lecturers. The School is sponsored by EAA and is part of the Young Forum activities, which also comprise participation in the conference, and invitation to the Young Acousticians Network event on Monday evening (sponsored by DAS).

At last, we want to draw your attention to the option extending your conference paper to a full research article in the EAA journal Acta Acustica, now full open access.

We wish you all, organizers, participants, students, sponsors and exhibitors an excellent and fruitful EUROREGIO 2022.

On behalf of the EAA Board
Birgit Rasmussen
EAA Vice-President

BNAM and EuroRegio 2022 in Aalborg Denmark

Dear Colleagues in Europe

On behalf of the Nordic Acoustic Association (NAA) I wish you very welcome to BNAM and EuroRegio in Ålborg in May 2022. You will be well greeted by a great city, and I sincerely hope you will enjoy your stay and earn new, valuable knowledge.

I also extend my heartfelt thanks to Lars Bramsløw and Dorte Hammershøi and all their colleagues for their effort in preparing the conference. In agreement with the NAA, the organizing committee has found leading researchers for keynote talks, supporting the UN sustainable development goals.

The NAA was established in the year 1954. The main driving force and the first president of the NAA was the Professor Fritz Ingerslev from Denmark, who saw the value in Nordic collaboration and coordination on acoustics. The biannual Nordic Acoustics Meetings, later Baltic-Nordic Acoustics Meetings, have been held since that same year 1954 on a rotating basis among the member countries. This time, the NAA welcomes joining forces with the European Acoustics Association (EAA).

Coming from the small nation of Icelanders, it was of great value to me to have good and clever neighbours. After my study at the University in Iceland, I studied acoustics first at DTU in Copenhagen and later at the TU in Berlin.

This is the first BNAM conference after the covid pandemic. What we have learned through the crisis is how to convene an e-conference. It is not the same as a conference in the flesh, but still of value. While they can improve the sustainability of conferences and spread the word wider, meeting in person cannot easily be surpassed. So, we are looking forward to once again having a traditional physical conference!

A very clever teacher I had in Berlin once said: “We can divide the world in three parts: Material, Energy, and Information. Albert Einstein taught us the connection between the first two. The Information is however unique: It does not diminish by being spread.”

I wish you a very valuable conference,

Ólafur Hjálmarsson
NAA President

Welcome to the Joint

Euroregio / Baltic-Nordic Acoustics Meeting 2022

Our proposal for hosting the Euroregio 2022 as part of our host obligation for BNAM 2022, was motivated by the purpose of the Euroregio conferences; to move around in Europe and embrace the regional conferences. The BNAM conferences have a solid circulation record, and it is always a pleasure to host it in Denmark, where it originally started.

Little did we know of the events that would follow, and the uncertainty that should follow the project for the coming two years. We watched national and international conferences turn digital, and we got used to prepared videos or live presentations with a variety of challenges. The flexibility of viewing presentations offline proved some upsides as well, when everything had to be managed from the same home environment. But even the most professional attempts to make interaction in digital environments would exclude many whatever the options offered.

The idea of conferences is to interact with peers, formally and informally, and to create networks that would support professional life in science, whether academic or industrial. This is the one thing, which the majority of digital conferences have failed to do.

While we wait for new conference formats to emerge, which will eventually combine the best from both worlds, we should celebrate the opportunities that allows us natural interaction with all its pros and cons.

It is therefore with great pleasure that we welcome delegates for the ER BNAM 2022 conference in Aalborg, where we will enjoy the scientific talks, investigate opportunities in the exhibition, and introduce juniors to the real life experience of attending a physical conferences, with all its opportunities for professional interaction.

Dorte Hammershøi
Professor, Aalborg University
Chair

Lars Bramsløw
President
Danish Acoustical Society

Welcome Exhibitors and Sponsors

On behalf of the EAA, NAA and the Danish Acoustical Society, the organizers are pleased that so many companies have shown interest in the conference and exhibition that takes place in Aalborg 9th-11th May 2022.

We are proud that the sponsorship options have been received well, and that so many different organizations are supporting the Euroregio/BNAM 2022 as either Platinum, Gold or silver sponsors.

The exhibition has since long time been fully booked with a variety of exciting companies, and this demonstrates that physical meetings are important for all of us.

Thanks to all sponsors and exhibitors. We look forward to spending these days with you in the technical sessions and as a part of the social program.

Welcome to Aalborg.

On behalf of the EAA, NAA and the Danish Acoustical Society

Dorte Hammershøi, Chair, Prof. Aalborg University

Lars Bramsløv, President of the Danish Acoustical Society

The organizing committee

- Dorte Hammershøi, Professor PhD, Aalborg University (chair)
- Lars Bramsløw, Senior scientist, Eriksholm Research Center and president of the Danish Acoustical Society (co-chair)
- Flemming Christensen, Associate Professor PhD, Aalborg University (co-chair)
- Claus Møller Petersen, Senior Consultant, SWECO, Denmark
- Rodrigo Ordoñez, Associate Professor PhD, Aalborg University
- Anders Olsen, Vibratec, Denmark (Sponsors and Exhibition management)
- Dorthe Sparre, Administrative secretary, Aalborg University
- Sreeram Kaithali Narayanan, PhD student, Aalborg University (YAN liason)
- Henriette Bugge Mortensen, Senior consultant, Destination Nord & Partners A/S (registration, hotels and tours)

Liasons of the Nordic Acoustic Association (NAA)

- Ólafur Hjálmarsson, NAA president
- Finnur Pind, Treble Technologies, Iceland
- Hans Bodén, Professor, KTH Royal Institute of Technology, Sweden
- Henrik Möller, senior consultant, AKUKON, Finland
- Mari Alvik Hagen, Senior consultant, Rambøll, Norway

Liasons of the European Acoustics Association (EAA)

- Birgit Rasmussen, Senior scientist, Aalborg University (EAA vice-president, conferences)
- Elie Abi Raad, PhD student RWTH Aachen University (YAN chairman)
- Karin Loh, PhD student, RWTH Aachen University (YAN vice-chairman)

Scientific committee

- Lars Bramsløw, Eriksholm Research Centre, Snekkersten, Denmark
- Dorte Hammershøi, Aalborg University
- Flemming Christensen, Aalborg University, Denmark
- Rodrigo Ordoñez, Aalborg University, Denmark
- Piotr Majdak, Acoustics Research Institute, Austrian Academy of Sciences, Austria
- Douglas Manvell, DMdB, Denmark
- Teresa Carrascal García, Quality in Construction Unit, Eduardo Torroja Institute for Construction Science. IETcc–CSIC, Spain
- Michael Haverkamp, Germany
- Christian Sejer Pedersen, Aalborg University, Denmark
- Claus Møller Petersen, Sweco, Denmark
- Alain Bradette, Sweco, Norway
- Tobias Neher, University of Southern Denmark, Denmark
- Louena Shtrepi, Politecnico di Torino, Italy

TABLE OF CONTENTS

Session R015 - Noise assessment and road traffic noise 2

Session R041 - Sound quality

Session R018 - Technical audiology

Session R016 - Noise assessment and road traffic noise 3

Session R019 - Objective and perceptual aspects of measured and simulated alternatives to conventional spaces

Session R020 - Current topics in hearing-aid processing and fitting 1

Session K006 - Keynote 2

Session R017 - Noise assessment and road traffic noise 4

Session P002 - Poster session

Session R025 - Room and Concert hall acoustics 4

Session R028 - Building acoustics 2

Session R031 - Numerical acoustics 2

Session R032 - Subjective experience of soundscapes

Session R047 - Closing ceremony

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Silent City - Europe's First Living Lab for Traffic Noise
Paper for Euroregio BNAM 2022

Sif Enevold[1,]

[1]Gate 21, Albertslund, Denmark

sif.enevold@gate21.dk

Abstract
Around Copenhagen, in the most noise emission plagued areas in Denmark, 11 municipalities and The Capital Region of Denmark have joined forces to improve citizens' health and quality of life. Traffic noise is one of the largest environmental problems in our society. The partnership encourages exchange of experiences across different local policies and geographical spaces. Since the start in 2015, we have helped to put traffic noise on the agenda, and on that basis implemented several concrete initiatives, including a public consultation at the Danish parliament, preparation of a white paper, targeted press efforts and concrete demonstration projects in a Living Lab. The partnership will help to put the participating municipalities and Denmark on the world map within urban noise control.

Keywords: Noise abatement, Municipality, Partnership, Policy, Living Lab

1 Together We are Stronger, and Wiser

Around Copenhagen, in the most noise emission plagued areas in Denmark, 11 municipalities and the Capital Region of Denmark, have joined forces in the partnership “Silent City”, to improve citizens' health and quality of life. Silent City works with making knowledge available and supporting new solutions on traffic noise abatement towards lower traffic noise levels and higher quality of life for citizens.

1.1 The Core Value

The Silent City partnership works on several levels with political influence, opinion-making, competence development and the establishment of a Living Lab. The core value lies in working together across thematic and organisationeal boundaries in a quadruple helix collaboration where public sector, companies, research institutions and citizens work together to reduce traffic noise and improve quality of life. Together we are stronger, and wiser.

The partnership started out in 2015 as an association of the municipalities immediately west and south of Copenhagen, with a common challenge of noise from the state highways. Since then, the partnership has expanded to include municipalities north and northwest of the capital, also struggling with heavy traffic noise from the state roads.

The partnership builds on Gate 21 renowned capacity to work with local municipalities in green innovation and green growth, combining the broad competence in understanding green transition and local policies, with the efforts to combat traffic noise.

1.2 The Living Lab – In Real Scale, In Real Life.

The partnership test and develop innovative solutions for combating traffic noise on a 1:1 scale in natural urban environments. In Silent City Living Lab we demonstrate technologies in real scale, in real life.

The Living Lab is an obvious focal point for testing new solutions and practical knowledge sharing amongst municipalities and private actors. With urbanisation, increased mobility, and other human activity it is not always technically or economically feasible to remedy the noise. In that case, there should be more precise guidelines for subsidies for noise insulation, replacement, and expropriation at new facilities where the noise cannot be reduced sufficiently.

The complex problem of traffic noise gives partners the opportunity to combine several agendas where Denmark is strong: sound and acoustics, public transport solutions and mobility, environment and climate, urban development and quality of life, sustainable construction, and infrastructure. The Living Lab gives companies a reference to show how the technology works for future costumers and collaborators. It gives the public sector a picture of future possibilities beyond existing off-the-shelf items. The main criteria for a solution to be part of the Living Lab is in other words, the need of test and demonstration in a specific context to explore the effect, and the solution not being a commonly used solution on the market.

1.3 A Sound Transition

Noise does not respect municipal boundaries. The dialogue between municipalities, region, and state on how we combat traffic noise is urgent and important. In Denmark, over 1.3 million people are estimated to be affected by traffic noise [1]. WHO has declared noise as one of the biggest environmental risks to health in Europe. The negative effects of traffic noise on people and society also speak its clear language; many healthy living and production years are lost as a result of high noise levels, and socio-economic calculations estimate that traffic noise costs society many billions of kroner annually. In the Capital Region alone, the costs due to traffic noise amounted to approximately DKK 2.4 billion in 2014 [2].

The Silent City Whitpaper ”Noise Emission demand Action” from 2020 point out that EU has succeeded in putting focus on the challenges of traffic noise, among other things through a directive on noise mapping. But the directive has not been followed up with binding EU legislation that can concretely affect traffic noise levels in Europe. Many of the technologies are investment-intensive and extensive, innovation efforts in this area are limited, decisions are long-term and often it is not the one that pays the bill who benefits from the investment. The opportunities for the noise plagued citizens is also limited.

Fortunately, tools are already in place to limit nuisances and health effects from traffic noise. In Denmark, we have over the past 25 years implemented several noise reduction initiatives along the country's busy roads. Regardless of the scope of these efforts, there are still more than 720,000 homes affected by traffic noise above the Danish Environmental Protection Agency's guideline limit value [3]. At the same time, there is a sharp increase in traffic volumes on all highways in the Silent City parnter municiaplities A condition that will further increase the noise level if no effort is made. The municipalities are experiencing increasing pressure from citizens, citizen groups and landowners' associations, who complain about the noise, lack of night sleep and the opportunity to stay outside their homes. The task is therefore far from solved. There is an urgent need for a sound transition.

1.4 Location, location, location

The Silent City municipalities are placed on the top of national traffic noiseload statistics. The traffic noise is expected to rise in the coming years, due to increased traffic on the highways and expansions of the road network. In the municipalities, there is an ongoing need for competence building and knowledge sharing in order to qualify the local effort. At the same time, there is a need for the issue to be taken seriously from the national level.

The eleven municipalities; Albertslund, Brøndby, Hvidovre, Gentofte, Gladsaxe, Glostrup, Høje-Taastrup, Ishøj, Køge, Lyngby-Taarbæk, and Vallensbæk are challenged by traffic noise from four highways and a number of major traffic roads and rail links.

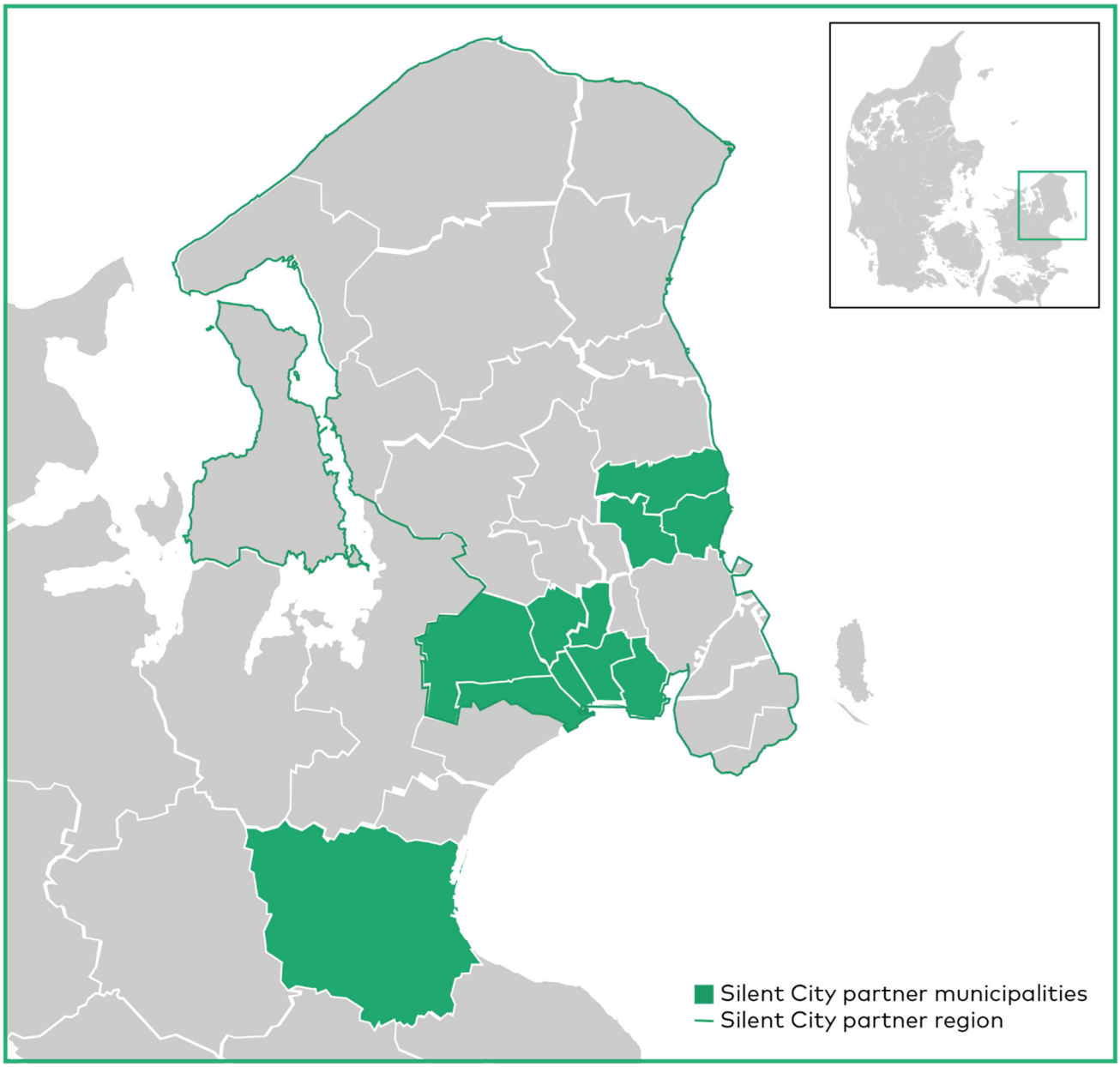


Figure 1: Silent city consists of 11 municipalities and one region. This includes Albertslund, Brøndby, Hvidovre, Gentofte, Gladsaxe, Glostrup, Høje-Taastrup, Ishøj, Køge, Lyngby-Taarbæk, and Vallensbæk municipaities and the Capital Region.

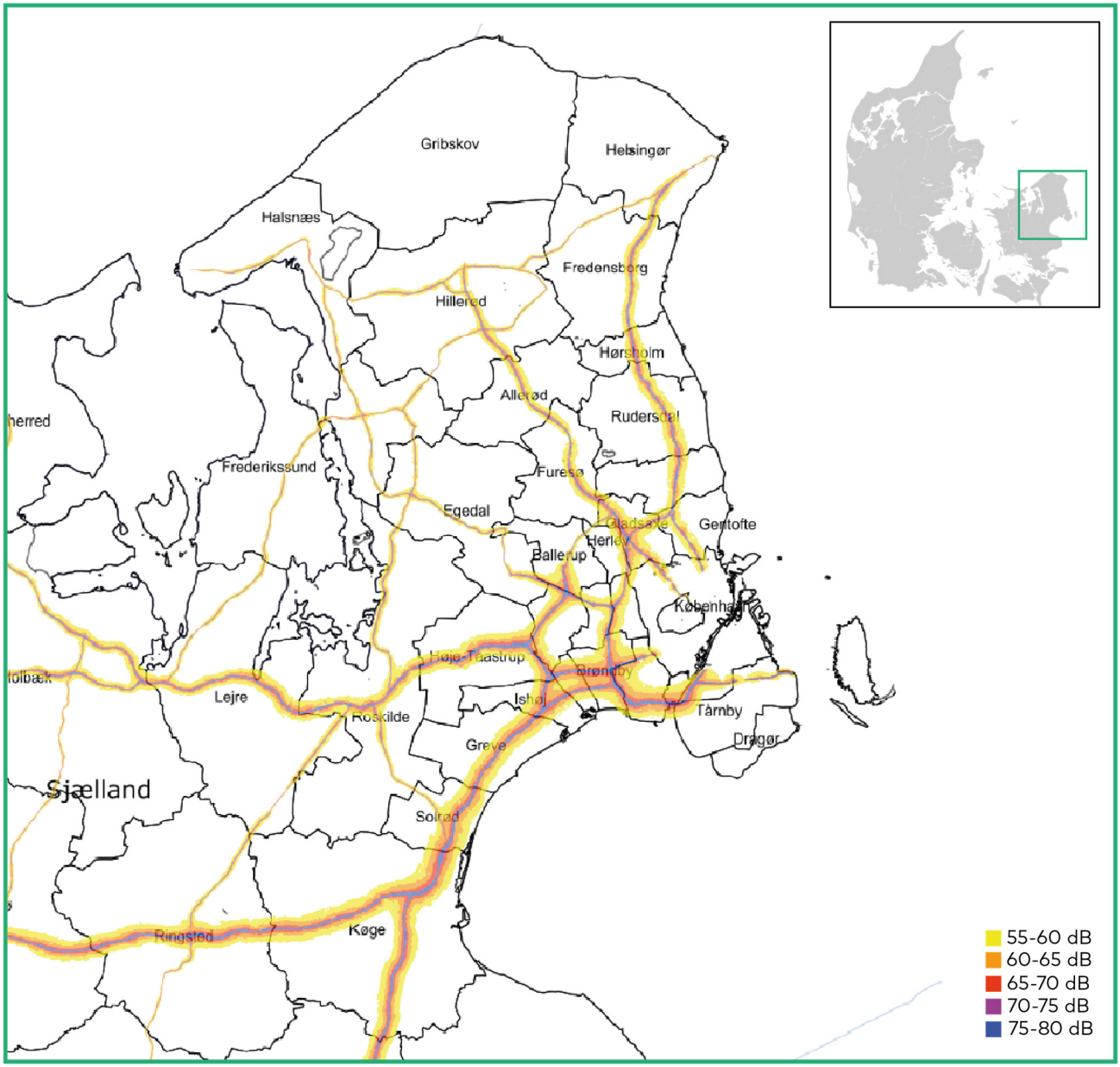


Figure 2: Noise loads from the state motorways in the Silent City geography [4]. The eleven municipalities are challenged by traffic noise from four highways and a number of major traffic roads and rail links.

2 Policy Instruments on The Noise Agenda

The Silent City partnership have made considerable developments on political awareness raising and joint development on policy measures such as urban development with noise considerations, citizens dialogue and strategic knowledge sharing between local governments. Here we give some examples on how our partnership have worked with policy instruments on the noise agenda with three examples, national infrastructure plan, speed limitations and political open fora.

2.1 Joint Statement to the National Infrastructure Plan

In April 2021, a historically large and comprehensive plan for Danish infrastructure of more than DKK 160 billion DKK was announced by the Danish government [5]. The plan includes massive investments in reducing congestion on Danish roads, a significant boost in public transport and initiatives to support the green transformation of the transport sector. Several elements in the plan have a positive effect on traffic noise and about DKK 3 billion will contribute to the solution. This allocation is a very good start; however, the scale of the problem is greater and will require further financing to solve.

The Silent City municipalities made a joint statement for the national plan, showing a direction for cooperation on noise abatement. The main points in the joint statement are twofold; invest where it helps the most citizens, and to give priority to help existing noise plagued citizens. The Danish noise map show a concentration of more than 50 percent of homes that are noisy due to noise from state roads in the metropolitan area, so at least 50 percent of the noise reduction funds around should be set aside in the capital region.

Priority must be given to helping the existing noise plagued citizens, rather than new infrastructure projects triggering investment in noise reduction actions. Proposed highway projects will not only have consequences in new areas, but also have derivatives. Include the proposed noise surcharge principle for expansion of the highway network to include existing highways.

2.2 Capacity Building on Speed Regulations and Traffic Adjustments

The faster a vehicle drive, the noisier it is. Speed regulations should be used as a tool to reduce traffic noise in residential areas. Silent City have created collaborations and knowledge exchange on municipalities ability and jurisdiction to create speed regulations and traffic adjustments to combat traffic noise.

In 2019, the municipalities of Hvidovre, Vallensbæk and Brøndby made a joint request to the national authority for a speed reduction to 90 km per hour on the highway passage in their municipalities. For state roads, municipalities have limited legal or sufficient financial means to act. However, following a new regulation in December 2017, a municipality may request speed reduction to the national authority in special cases, even as low as 60 km per hour solely for the sake of reducing the noise level [6].

The joint statement from the three municipalities is justified because, nowhere in Denmark, do so many people live so close to very busy highways. A total of 24,000 homes in the 3 municipalities are affected by noise above the recommended limit value of 58 dB, corresponding to more than 50% of the municipalities' homes. 60% of noise-affected homes are above the limit value solely due to noise from highways. The Danish Road Directorate's own noise survey shows that the three municipalities are in the top when it comes to the proportion of citizens who are affected by noise from the state's highways [7].

On their own roads, a municipality may request the police to reduce the speed limit in noisy residential areas. Silent City partnership has worked on sharing knowledge between municipalities on how to make traffic analysis with simulations on speed regulations and traffic adjustments. A traffic analysis can be used to show where the greatest noise reduction can be achieved and how many homes it will affected. Subsequently, the traffic analysis can be used to prioritize the speed reductions that provide the greatest value for as many people as possible.

Municipalities have gained knowledge on parameters important to speed limitation, for example travel times might be counted as a socio-economic cost, and how to avoid unwanted redistribution of traffic. The partnership has also raised awareness of other traffic regulations, for example green waves, with a more even flow in traffic, to avoid the noise from braking and acceleration. And by narrowing the roadway, introducing road bumps or other interventions in the physical infrastructure, you reduce the average speed of traffic.

2.3 Joint Understanding of Noise Mapping

Every fifth year, the European noise directive stipulate an update of the mapping of noise from major roads and in major urban areas. The next mapping is ongoing and must be submitted to the European Union in summer 2022. This year's mapping will, however, be slightly different from previous mappings, as EU has introduced a new common calculation model, CNOSSOS.

The implementation of the new CNOSSOS model is positive in the sense that all member states will use the same methods and will have comparable mappings. However, the model previously used in Denmark, Nord2000, is more advanced in aspects like propagation and meteorological conditions, which will result in generally lower noise levels for CNOSSOS than for Nord2000 calculations [8]. Therefore, Denmark will continue to use the Nord2000 model parallel to the new EU model.

For the Danish municipalities, the transition, and the double mapping, will undoubtedly cause some challenges. This is both in terms of the mapping, which serves as the underlying elements for the subsequent action plans, but especially also for the communication towards citizens. We will use the cooperation between the municipalities in the Silent City partnership and work together to ensure a common and well-founded understanding of the results from the new model. The work will aim at raising the technical knowledge of city planners, as well as creating suitable communication strategies to inform citizens on the consequences of the new mapping.

2.4 The Creation of a Policy Arena

With a joint political mandate from several local administrations and the Capital Region, Silent City have managed to help creating a momentum for political awareness of traffic noise emissions being one of the largest environmental problems in our society. The partnership has been able to raise the issue broader and on a more strategic level, than a single noise plagued municipality would be able to do alone. Under a joint flag and joint understanding of the problem, mayors of Silent City municipalities and the Capital Region have managed to create a conference in the Danish parliament, as well as more local events directed towards citizens.

In May 2016, the Silent City partnership, in collaboration with other strategic partners, launched the first whitepaper for traffic noise in Denmark and opened the doors to the conference 'Traffic noise: An overlooked societal challenge'. Here, media, politicians, professionals, and citizens were invited to a joint debate on how we can bring traffic noise down in Denmark. Local and parliamentary politicians and the leading Danish companies in the field got together and listen to the latest international research on health and socio-economic consequences of traffic noise and the perspectives and opportunities and challenges with solutions existing on the market.

Silent City have also created several citizens focused events. For example, the events "It's Still Noisy" in October 2018 and "Traffic Noise Demands More Action!" in September 2021, attracted several hundred of citizens and national and local media. In an open forum, citizens were invited to a one-on-one with their local politicians, listen to experts talking about noise and health, and get a firsthand idea on how to mitigate noise in their surroundings. Additionally, these events have sent a strong signal to the Danish parliament that traffic noise is an issue that affects and worries many citizens in the Greater Copenhagen area.

Figure 3: Panel of mayors debate the future of collaborative noise abatement and noise reduction with citizens at the event "Traffic Noise Demands More Action!".

3 Show – Don't Tell – Real Scale Living Lab

In the Silent City Living Lab, municipalities allow concrete demonstration in the real traffic conditions. In triple helix collaborations between cities, technology owners and knowledge institutions, the demonstrations give the public sector and the citizens a picture of future possibilities beyond existing off-the-shelf items. In the following part, we give you some examples of how demonstrations projects in Silent City comes into play, the silent bench, noise absorbing guardrails, and real time monitoring of noise. From demonstration to market is still a bumpy path, and not all solutions will succeed to market. But with real life measurements of effect, citizen evaluations, communication, and on-site visitor service, the Living Lab facilitate innovation and open possibilities for new solutions.

3.1 A "Silent Bench" Can Halve the Perceived Noise

South of Copenhagen, at a recreational area with a football golf course situated next to a major railway and highway corridor, you will find the Silent Bench. The bench consists of a 2-meter high, slightly angled, and curved noise screen and is a result of a collaboration between technology experts FORCE Technology, developing engineers G9 Landskab, and Vallensbæk Municipality. On the inside, the bench is lined with perforated sheets with sound-absorbing material. Thus, the back of the bench acts as a shield for both sound and wind, while the sound that may still creep over or around the screen can be absorbed in the porous plates on the inside. It is important to consider wind direction and potentially reflective surfaces when choosing the location.

The open landscape location provides a good protection against the noise. It also provides a visual shielding of the traffic, which can have an impact on the perceived noise level. In denser urban spaces, the reflection of sound between the hard surfaces will change the measured effect, and one will not be able to expect the same high degree of attenuation. In some noise spectra, the noise level inside the bench is reduced by 13-14 dB, this equates to the noise being more than halved [9].The reduction of noise is highest in the most intrusive and sharp sound levels.

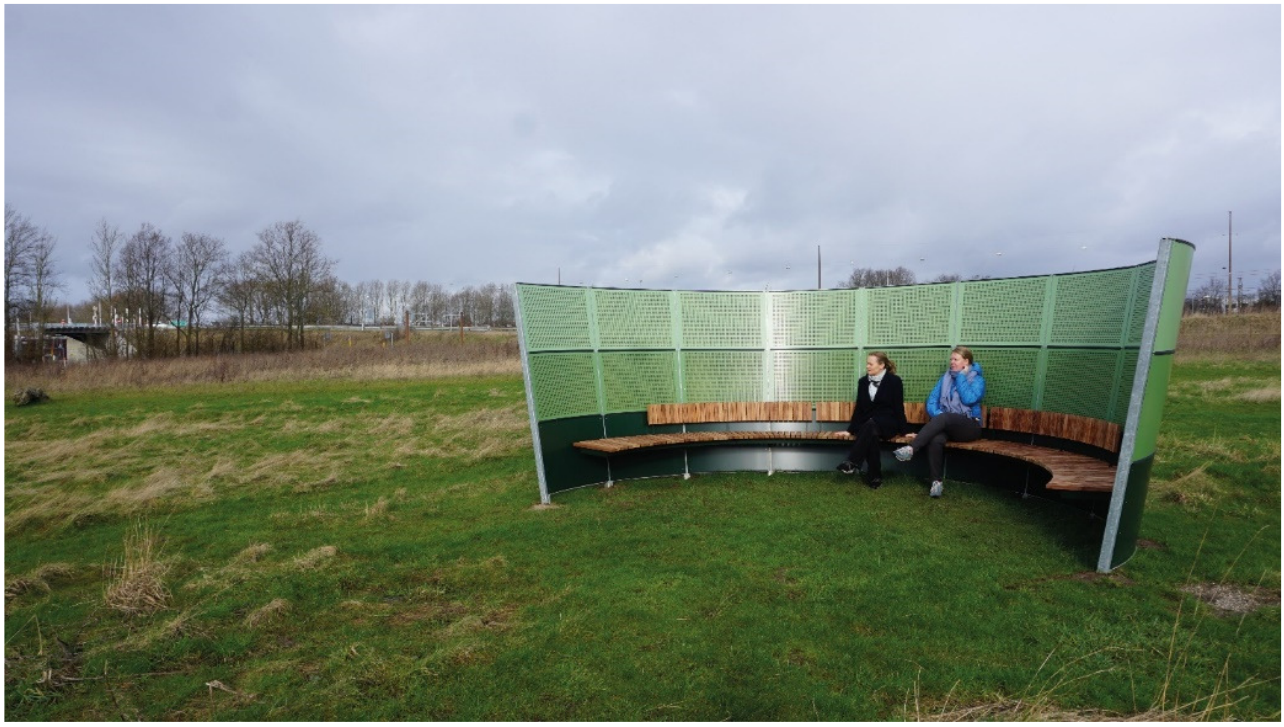


Figure 4: The "Silent Bench" at the football golf course in Vallensbæk

3.2 From Pilot to Product on Noise Absorbing Guardrails

On Vejlegårdsvej in Vallensbæk and on Ishøj Stationsvej you will find the country's lowest noise screens with reduction effect. Here the company NAG 1 ApS, as part of Silent City Living Lab, and with support from Danish start-up funds, has tested noise absorbing guardrails. The Silent City partnership helped the company from pilot, to launching of final product, by allocating demonstration site and reference surroundings with citizens involvement. The guardrails are now a patented CE-standard crash barrier. The technology can be integrated in existing guardrails and the low noise screen has good effect on this stretch because it is installed directly at the wayside, close to the source and especially attenuates that part of the noise coming from the tire-to-road contact.

A survey shows that 67 percent of the residents next to Vejlegårdsvej believe that noise reduction protection has had a good effect. They experience a moderate noise reduction that corresponds to the expected power of between 3 and 5 decibels for homes closest to the road. 3 to 5 decibels are experienced as an audible change. Only 10 percent of residents says they experience no - or very little - noise reduction [8].

3.3 Real Time Noise Measurements Improve Decision Making

Silent City have demonstrated a dynamic noise chart using real-time noise measurements in an area. The technology is developed by the company SoundEar and the GTS institute FORCE Technology. A set of microphones connected to a computer show both the area's current and past noise levels, which are presented in a dynamic noise map. The map shows day by day, hour by hour, how noise development is in different areas.

Official noise mapping in Denmark shows average values over a year. These maps can be good to point out which areas are exposed to traffic noise; how many are exposed to noise and help to prioritize the noise reduction effort. But they do not say much about the current situation, and they cannot be used as a basis for understanding daily fluctuations or the perceived nuisance of noise. As a municipality it can be challenging to deal with complaints from citizens about elevated noise levels if the noise mapping does not indicate averages above the limit values.

The dynamic noise chart has embedded meteorological data and can take height of wind direction and strength into account. The real-time measurements and the dynamic noise maps can give noise-disturbed citizens a better insight into the noise that they experience. The program can also show where and when noise limits are exceeded, and the map can calculate the proportion of highly disturbed citizens in the area.

The map can also be used to calculate the effect of different noise reduction measures, such as noise barriers, noise walls or reduced speed in a particular area. This can provide better decision-making basis for new noise reduction measures and give citizens a better incentive to seek municipal noise pools or co-finance the measures.

4 Conclusions

Silent City is a collaboration between some of the most noise plagued municipalities in the capital region of Denmark. The goal is to reduce traffic noise, increase traffic noise abatement and promote the quality of life and health of citizens. Through the Silent City Living Lab and cooperation with companies and knowledge institutions, the partnership takes part in the necessary development of instruments needed for the traffic noise abatement. The partnership shows the benefit of a joint effort to raise awareness on one of the largest environmental problems in our society. Since 2015, the collaboration has strengthened political cooperation between municipalities, and to speak with a strengthened voice towards the national level. Silent City have demonstrated state of the art technology and methods to target traffic noise and disseminate the latest knowledge about traffic noise strengthen dialogue with citizens on joint initiatives.

Acknowledgements

References

[1] Gate 21, Rambøll, FORCE Technology, Trafikstøj kræver handling, Fakta udfordringer og løsninger; Hvidbog, 2020, s.11. *www.gate21.dk*

[2] Region Hovedstaden, Hovedrapport - vurdering af luft og støj kortlægning og virkemiddelkatalog, april 2018

[3] Miljøstyrelsen, National kortlægning af boliger belastet af vejstøj i 2012. *Arbejdsrapport fra Miljøstyrelsen nr. 5*, 2013

[4] Miljøstyrelsen, Støj-Danmarkskortet, *www.mst.dk*

[5] Transportministeriet, Danmark fremad – Infrastrukturplan 2035, april 2021

[6] Transport- og Boligministeriet, Bekendtgørelse nr. 1486, *Retsinformation.dk*

[7] Vejdirektoratet, Støjhandlingsplan for statens veje 2018-2023, Rapport 593, *www.vejdirektoratet.dk*

[8] Force Technology, Støjkortlægning med CNOSSOS - EU's fælles beregningsmetode for vejstøj, *Trafik og veje*, marts 2021

[9] Force Technology, Stillebænk – en oase fra støjen, *Trafik og veje*, april 2020

[10] Gate 21, Styrk indsatsen mod trafikstøj – Inspiration til handling, 2021, *www.gate21.dk*

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Health effects of environmental noise in a wind power area

Valtteri Hongisto,* Jenni Radun, Henna Maula, Pekka Saarinen, Jukka Keränen, Reijo Alakoivu

Turku University of Applied Sciences, Turku, Finland.

*valtteri.hongisto@turkuamk.fi

Abstract

Our purpose was to determine how wind turbine noise level and road traffic noise level are associated with symptoms in a wind turbine area. We applied a case-control design, which included a wind power area and a control area farther away not exposed to wind turbine noise nor visibility of turbines. Noise levels were modeled to the yard of all potential respondents. Road traffic and wind turbine noise level varied between 32–64 dB and 17–39 dB L_{Aeq}. Altogether 684 residents responded to a living environment questionnaire measuring various factors. Here, we focus on stress, symptoms, health, and annoyance of environmental noise. No associations with wind turbine noise level were found for symptoms, diseases, nor stress. Instead, prevalence of several symptoms and prevalence of heart disease were higher if road traffic noise level was higher. Wind turbine noise level varied between 17–39 dB and did not exceed the Finnish regulated value of 40 dB L_{Aeq} in any residential yard of the wind turbine area. Therefore, the findings can be applied to all areas where the regulated value is applied. Our study emphasizes that road traffic noise is a larger problem than wind turbine noise even in wind power areas, not to mention areas where the road traffic noise is higher than in our study.

Keywords: road traffic noise; wind turbine noise; health effects; noise annoyance; epidemiology

1 Introduction

Most of the residential surveys investigating the effects of wind turbine (WT) noise until 2015 were conducted in wind power areas where a large proportion of residents close to the WTs were exposed to A-weighted sound pressure levels (SPLs) higher than 40 dB L_{Aeq} [1, 2]. Those studies were utilized as scientific evidence in the development of Finnish WT noise regulations which were published in 2015 [3]. These regulations give the upper limit of WT noise both for daytime ($L_{Aeq,07–22}$ =45 dB) and nighttime ($L_{Aeq,07–22}$ =40 dB). In practice, the nighttime regulation is applied also during daytime since energy companies do not want to invest on WTs which cannot produce full power during the nighttime. Since the complaints about WT noise have not ended after 2015, it is important to investigate the effects of WTs in areas, which fulfill the new regulation. This will provide evidence-based information about the health effects of WTs under current regulations. This is especially important in areas with large WTs, since they induce more concerns and fears than smaller WTs. The concern about the infrasound exposure is also an important driver to investigate the health status of residents living close to WT areas fulfilling regulations.

Because onshore WTs are usually erected close to main roads, WT noise is not the only form of environmental noise in WT areas. Previous studies have not shown any other health effects of WTs than noise annoyance. Instead, several different health effects have been found to be associated with road traffic (RT) noise [4]. From public health point of view, it is relevant to investigate the health effects of both WT and RT noise in parallel to provide a holistic understanding of the health effects of environmental noise in the living environment.

Focusing solely on the health effects of WT noise is not reasonable, if it is self-evident that the residents are also exposed to other forms of environmental noise. There are very few studies which have investigated the health effects of both WT and RT noise in parallel in the same area.

Our purpose was to determine, how RT and WT noise are associated with noise annoyance, symptoms, stress, and diseases. The full study has been published [5] and this study gives a summary of it.

2 Materials and methods

The study applied *case-control* design, where the exposure variables (independent, objectively measurable variable) were the SPLs of RT noise and WT noise in the yard and the response variables (dependent variables) were the subjective responses to the questionnaire. The study was accepted by the ethic committee of Turku University of Applied Sciences.

We wanted to study a couple of WT areas, where the Finnish WT noise regulation [3] is fulfilled but, at the same time, the areas should have high population density. We could not apply random sampling over all Finnish WT areas since it is necessary to conduct the noise emission measurements of the WTs and they are very time-consuming. The best choice in Finland was Hamina city, which has three nearby WT areas:

- Harbor involves two 3 MW WTs erected in 2015
- Summa involves three 3 MW WTs erected in 2010
- Mäkelänkangas involves four 2 MW WTs erected in 2012.

The control area was a suburb in Kotka city 6.8–8.0 km west from the eastmost WT of Mäkelänkangas. The control area resembled Hamina w.r.t. building base, socioeconomics, and seaside environment.

The permanent residential houses were identified with map services. Their building IDs were used to obtain the basic information of residents. We asked individual address of a single adult from each household. In the case of two or more adults in the same address, randomization was made w.r.t. age and gender.

All households were invited to respond, which located closer than 2,8 km from the WTs. The living environment questionnaire was mailed to 2560 households in the WT area and 498 households in the control area, altogether to 3058 households.

Questionnaire survey was conducted in autumn 2018. Response time was 4 weeks after which a reminder was sent to all households. We received 684 responses, 563 from WT area and 121 from control area. Response rate was 22.4%. Questionnaire was mailed in paper form. Web option was available in Finnish, Swedish, and English. Low response rate could be explained by the length of the questionnaire and the fact that the filled questionnaire had to be returned to the nearest ordinary mailbox.

The questionnaire was masked: it was not possible to see, that our purpose was to study WT and RT noise effects on human. Here, we focus on a minor proportion of the questionnaire items:

- Annoyance of WT and RT noise. Noise annoyance was measured using an 11-step response scale (0 Not at all, 10 Extremely). The responses were dichotomized so that people who responded 5 or more were rated to be annoyed (%A).
- Prevalence of non-specific symptoms during last 12 months (migraine or headache including nausea, vomiting, and sensitivity to light and sound; dizziness; ringing, whistling or other sounds in your ears that have no actual source, e.g., tinnitus; impaired hearing; blocked ears or a sense of pressure in your ears; rash or itchy skin; back pain or backache; regular stomach problems; blurred vision; tachycardia or heart palpitations; problems in concentrating or remembering things; panic attacks or similar sensations)
- Prevalence of diseases during last 12 months (chronic pain; asthma; joint inflammation; cancer; depression; elevated blood pressure; bronchitis, pulmonary emphysema, or chronic obstructive pulmonary disease; diabetes; heart disease; sleep problems, including sleep apnea and insomnia; restless legs syndrome).

For each respondent, the following independent variables were determined:

- distance to the nearest WT (0.9–2.7 km),
- A-weighted equivalent SPL, when all WTs are producing maximum electricity ($L_{Aeq,WT}$), and
- A-weighted equivalent SPL of RT noise during daytime hours 07–22 ($L_{Aeq,07\text{-}22,RT}$).

Both L_{Aeq}'s were determined in the respondent's yard at 4 m height. It should be noted that $L_{Aeq,07\text{-}22,RT}$ represents well the RT noise exposure every day throughout the year. On the other hand, $L_{Aeq,WT}$ is only valid during a windy condition (wind speed larger than 12 m/s at hub height). Such conditions occur less than 10% of the year. Because such definition is used in legislation, we had to adopt it and we did not conduct our analyses using the annual WT noise level. Annual level would probably be more than 5 dB smaller.

SPLs in the yards were simulated using CadnaA -software using the national topographic maps. $L_{Aeq,WT}$ was determined using a national method [6]. The prediction accuracy of this method has been found to be very good [1]. Because the WTs were not new and the noise emission data was vague, we measured the sound power level of the WT's in each three WT area using another national method [7], which is principally in agreement with IEC 61400-11. $L_{Aeq,07\text{-}22,RT}$ was determined using Nordic model [8]. It takes into account the traffic amount, share of heavy vehicles, road surface type, and traffic speed. Traffic numbers were obtained from a national authority.

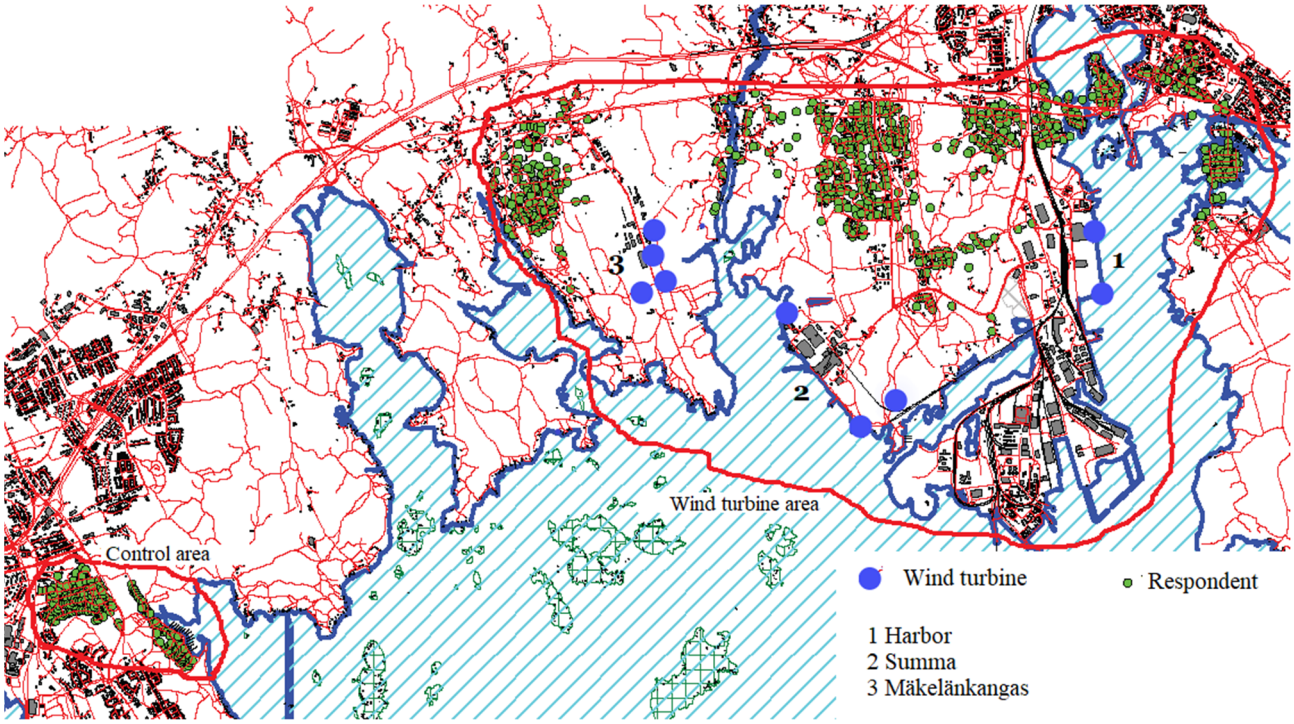


Figure 1: A map of the study areas. Control area is on the left and the WT area is on the right. Both areas are surrounded by red line. Hamina city center is locating in the right top corner.

3 Analyses and results

Results were analyzed using three methods. Method 1 compared the prevalence of stress, symptoms, and diseases in WT area and the control group. Method 2 assessed the association of stress, symptom, and disease prevalences with WT and RT sound levels. Method 3 assessed the noise annoyance related to WTs and RT in different noise level categories.

Method 1. Respondents were divided to four WT noise level categories (groups) having approximately similar sizes: group 17–25 dB (122 respondents), group 25–30 dB (282 respondents), group 30–40 dB (159

respondents) and control group with no audible WT noise exposure (121 respondents). Three first groups locating in the WT area were compared to the control group using binary logistic regression. The model involved also age, gender, and RT noise level ($L_{Aeq,07\text{-}22,RT}$). The outcome of the analysis was that no significant differences were observed between the control group and three other groups w.r.t. the prevalence of stress, symptoms, nor diseases (p>0.05). In conclucion, exposure to WT noise within 17–40 dB $L_{Aeq,WT}$ was not associated with higher prevalence of stress, symptoms, or diseases.

Method 2. This method did not use the groups but continuous noise levels and the focus was only in the WT area. The association of continuous sounds levels ($L_{Aeq,07\text{-}22,RT}$ and $L_{Aeq,WT}$) and prevalence of stress, symptoms, and diseases was analyzed using binary logistic regression. The model included also age and gender. WT noise level, $L_{Aeq,WT}$, was not associated with the prevalence of stress, symptoms, nor diseases. Instead, higher RT noise level, $L_{Aeq,07\text{-}22,RT}$, was significantly associated with higher prevalence of heart disease, and five symptoms. Figure 2 reports the risk coefficients in percentages. Figure 3 clarifies the meaning of the risk coefficients at different RT noise levels.

Method 3. Respondents in the WT area were divided in the three noise categories (groups) according to their WT noise exposure as described above. All respondents were also divided in the five noise categories according to their RT noise exposure: 32–40 dB, 40–45 dB, 45–50 dB, 50–55 dB, and 55–64 dB. The relationship between annoyance (%A) and noise level is show in Figure 4.

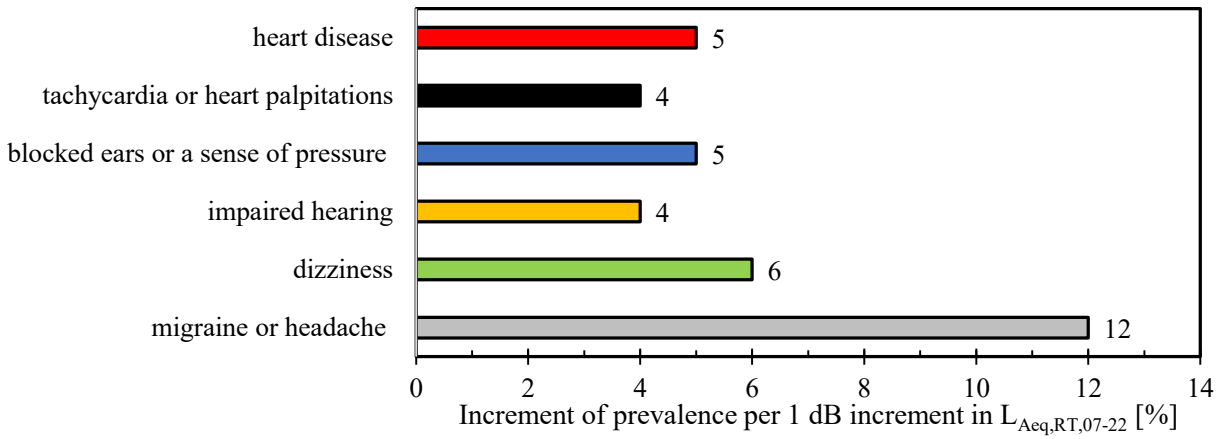


Figure 2: The increment of prevalence for heart disease and five symptoms per 1 dB increment in RT noise level, $L_{Aeq,07\text{-}22,RT}$. The x-axis corresponds to 100·(Exp(B) – 1). $L_{Aeq,07\text{-}22,RT}$ represents the value in the respondents' yard.

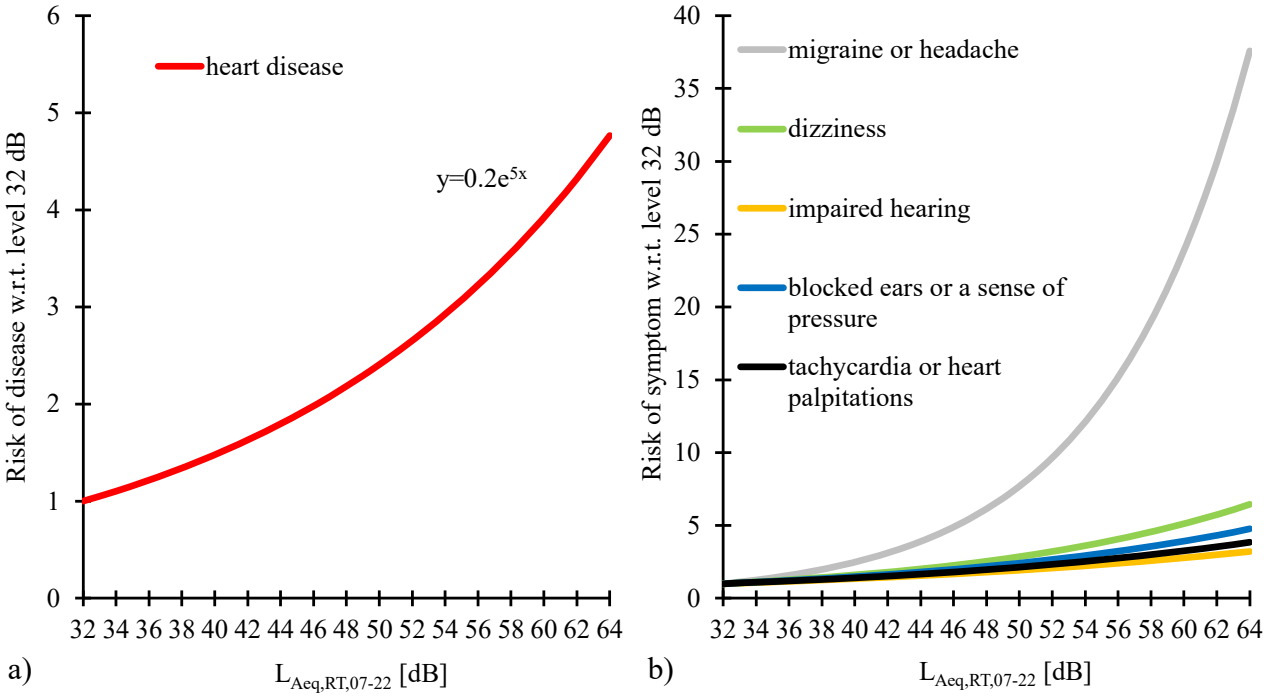


Figure 3: a) Prevalence of heart disease was significantly higher, when the A-weighted SPL of RT noise, $L_{Aeq,RT,07–22}$, was higher. Prevalence increased by 5 %, when the level increased by 1 dB. b) Prevalence of five non-specific symptoms was significantly higher, when $L_{Aeq,RT,07–22}$ was higher. $L_{Aeq,07-22,RT}$ represent the value in the respondents' yard.

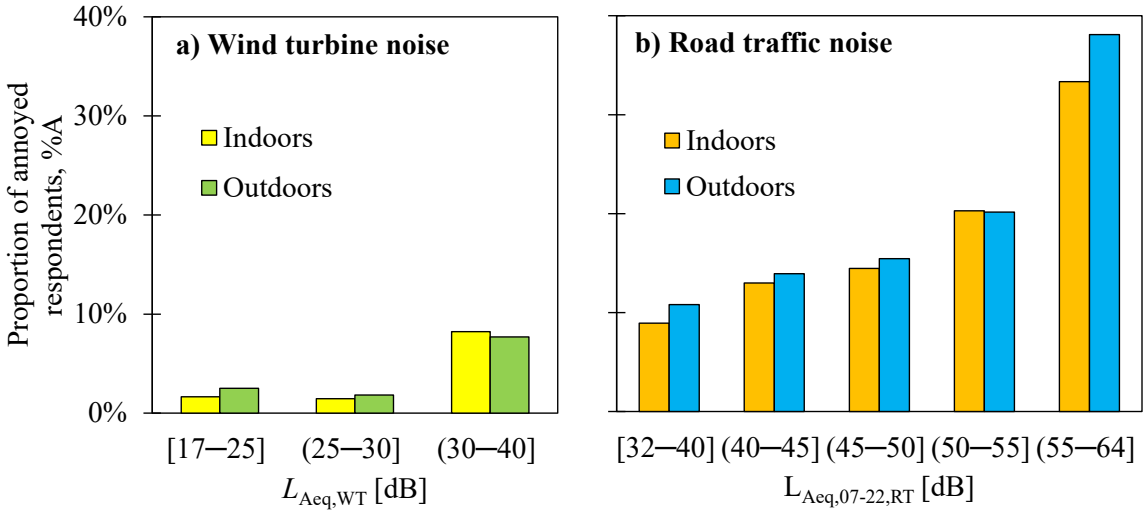


Figure 4: (a) Proportion of respondents reporting WT noise annoyance both indoors and outdoors was higher when the WT noise level, $L_{Aeq,WT}$, was higher. (b) Proportion of respondents reporting RT noise annoyance both indoors and outdoors was higher when the RT noise level, $L_{Aeq,RT,07–22}$, was higher. %A describes the probability, that the response in 11-step noise annoyance response scale from 0 to 10 was five or higher. Both $L_{Aeq,WT}$ and $L_{Aeq,07-22,RT}$ represent the value in the respondents' yard.

4 Discussion

Altogether 2560 residents were invited to respond in the WT area but only 563 responded, giving a response rate of 22 %. Number of non-respondents was 1997. Respondents and non-respondents did not differ from each other w.r.t. WT noise level, distance to the nearest WT, nor RT noise level. We found that respondents were older (mean age 63 y) than non-respondents (56 y). We do not have reasons to believe that the results would not apply to the whole population in the studied WT area.

Results with method 1 agree with previous literature: the only undisputed health effect of WT noise under 40 dB is noise annoyance [1, 9]. Based on our result, it is improbable that WT noise exposure under 40 dB would cause elevated prevalence of stress, symptoms, or diseases in any WT area.

Results with method 2 agreed with the results obtained with method 1 regarding WT noise. Although the RT noise exposure was not very high in the studied areas, we found that higher RT noise level was associated with higher prevalence of five symptoms and heart disease, although the RT noise levels were not excessively high in any yard (32–64 dB). The finding related to heart disease agrees with previous literature [4]. The finding is very important from public health point of view: it suggests that RT noise was a larger problem in this WT area than WT noise. Because WT areas are erected close to roads, it is probable that similar findings would be made also in other WT areas where the RT noise levels reach 64 dB. Current Finnish WT noise regulation [3] seems to be sufficiently tight, because lawfully implemented WT areas do not have an effect on residents' stress, symptoms, nor diseases. Instead, the regulation of RT noise [10] was violated in many yards of the WT area. It states that the RT noise level shall not exceed 55 dB daytime ($L_{Aeq,07\text{-}22,RT}$) and 50 dB nighttime ($L_{Aeq,22–07,RT}$). It is possible that our results would change, if we excluded yards, where 55 dB was exceeded.

Results with method 3 showed an expected result that annoyance increased with increasing noise level. The same finding was made for both investigated environmental noise types. This agrees with previous literature [1,11,12]. An important new finding was that the proportion of annoyed respondents was approximately the same both for RT noise and WT noise, when the RT noise level and WT noise level was within 30–40 dB. This is in conflict with Janssen *et al.* [12] suggesting that the annoyance due to WT noise at this noise level range was much larger than annoyance due to RT noise. Janssen *et al.* collected the exposure–response relationships of different environmental noise types from different studies. Such comparisons are biased because annoyance depends on the study area [11] and annoyance reporting also depends on the subjective metrics (precise question, response scale). We measured both WT noise annoyance and RT noise annoyance in the same area using identical subjective metrics.

Our study has large value from public health perspective, because the study was conducted in a WT area, where the regulated SPL of WT noise (40 dB) was not exceeded in any yard. This justifies the assessment of the operability of the WT noise regulation [3]. Based on our survey, there is no health-based reason to tighten the current Finnish WT noise regulations [3]. Based on the results of Method 3, it is even possible to question why the noise regulations of WT noise [3] are tighter than the noise regulations of RT noise [10]. Further research is needed about the comparison of exposure–response relationships between different environmental noise types. Because all health effects were associated with RT noise, it is more relevant to focus on RT noise control in the future.

Acknowledgements

The study was part of Anojanssi –project funded by Business Finland (60 %, Tekes grant 828/31/2015), Turku University of Applied Sciences, Ministry of the Environment, Ministry of Social Affairs and Health, several companies and several associations.

References

[1] Hongisto, V., Keränen, J., Oliva, D. (2017). Indoor noise annoyance due to 3–5 MW wind turbines - an exposure-response relationship. The Journal of the Acoustical Society of America 142(4) 2185–2196. Online at: http://dx.doi.org/10.1121/1.5006903.

[2] Hongisto, V., Suokas, M., Varjo, J., Yli-Kätkä, V.-M. (2015). Tuulivoimalamelun häiritsevyys kahdella tuulivoima-alueella, Ympäristö ja Terveys -lehti, 6 54–59 (In Finnish). Available at: https://tuulivoimayhdistys.fi/media/794-hongisto_ym_2015_ymparisto_ja_terveys.pdf.

[3] Ministry of the Environment (2015), Government Decree 1107/2015, "Government Decree on the Outdoor Noise Level Guide Values caused by wind turbines," 31st August 2015, Helsinki, Finland. Available in English at: http://www.finlex.fi/en/laki/kaannokset/2015/en20151107.pdf .

[4] van Kempen, E., Casas, M., Pershagen, G., Foraster, M. (2018). WHO environmental noise guidelines for the European region: a systematic review on environmental noise and cardiovascular and metabolic effects: a summary. Int. J. Environ. Res. Publ. Health 15 379 (59 p). Available at: https://doi.org/10.3390/ijerph15020379.

[5] Radun, J., Maula, H., Saarinen, P., Keränen, J., Alakoivu, R., Hongisto, V. (2021). Health effects of wind turbine and road traffic noise on people living near wind turbines. Renew. Sust. Energ. Rev. 157 112040 (13 p). Available at: https://www.sciencedirect.com/science/article/pii/S1364032121013022.

[6] Ministry of Environment (2014). Modelling of wind turbine noise. Environmental Administration Guidelines 2|2014, Helsinki, Finland (In Finnish). Available at: https://julkaisut.valtioneuvosto.fi/handle/10138/42937.

[7] Ministry of Environment (2014). Using measurements to verify noise emissions from wind turbines. Environmental Administration Guidelines 3|2014, Helsinki, Finland (In Finnish). Available at: https://julkaisut.valtioneuvosto.fi/handle/10138/42938.

[8] Nielsen, H. (1996). Road Traffic Noise – Nordic Prediction Method. Nordic Council of Ministers, Copenhagen, Denmark.

[9] Freiberg, A., Schefter, C., Girbig, M., Murta, V.C., Seidler, A. (2019). Health effects of wind turbines on humans in residential settings: Results of a scoping review. Environ. Res. 169 446–63.

[10] Finnish Ministry of the Environment, Government Decision on the Noise Level Guide Values (993/1992), Helsinki, Finland.

[11] Radun, J., Hongisto, V., & Suokas, M. (2019). Variables associated with wind turbine noise annoyance and sleep disturbance. Building and Environment 150 339–348.

[12] S.A. Janssen, S.A., Vos, H., Eisses, A.R., Pedersen, E. (2011). A comparison between exposure-response relationships for wind turbine annoyance and annoyance due to other noise sources. J. Acoust. Soc. Am. 130 3746–3753.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Uncertainty estimation in environmental road traffic noise measurements using ISO 1996-2:2017

Simona Domazetovska[1,*], Maja Anachkova[1], Viktor Gavriloski[1], Vasko Changoski[1]

[1]Faculty of Mechanical Engineering, University Ss Cyril and Methodius in Skopje, North Macedonia

*simona.domazetovska@mf.edu.mk

Abstract
When measuring the road traffic noise, the environmental noise levels are quantitatively described using the equivalent noise pressure level parameter (L_{eq}). The value of L_{eq} based on the measurements done with sound level meter would probably differ from the true one due to the sources that can cause measurement uncertainty. According to this, the ISO 1996-2:2017 standard proposes calculation of the standard measurement uncertainty, so that the results from the measurements could be more accurate. This paper proposes a guideline on estimating the measurement uncertainty in compliance with the ISO standard for short-term measurements of road traffic noise. To verify the proposed method, twenty short-term measurements are done in same spot in morning and afternoon period from Monday to Friday for two weeks; and afterwards, the measurement uncertainty is calculated for each measurement. To confirm the accuracy of the results, additional acoustic modeling was made. By comparing the results for the L_{eq} parameter from the measurements and from the acoustic noise maps in accordance with the calculated measurement uncertainty, it could be confirmed the results accuracy. In conclusion, this paper presents a reflection on why estimation of the uncertainty of the measurement is essential in environmental noise analysis.

Keywords: uncertainty estimation, environmental noise, noise measurements, acoustic noise maps

1 Introduction

Environmental noise is defined as any unwanted or harmful outdoor sound created by human activity, such as noise emitted from transport, traffic and industrial activity. Noise pollution is known to be one of the main reasons for environmental pollution in the urban areas, causing harmful effects on the quality of life of the population that is directly exposed to noise. The noise pollution is known to be mainly caused by noise sources in the heavy urban traffic and is followed with increasing number of complaints from the public [1]. According to this, significant health problem can be noticed followed with loss of hearing, sleep distribution, reduced productivity, and traffic accidents [2]. The traffic noise is major environmental noise source, and by the analysis made in [3], it is confirmed to be unpleasant and most influential sound source that causes urban noise pollution.
In order to properly determine the environmental noise pollution assessment and to design effective noise control, noise measurements are required. According to the research in [4], the noise measurement is an important diagnostic tool in noise control technology and noise pollution assessment. To quantitively describe the environmental noise, the equivalent sound pressure level parameter L_{eq} is widely used, using the A weighted curve. The measured value of L_{eq} based on the sound pressure level measurements by sound level meter will probably differ from the true one because of the measurement uncertainty due to the sources that disturb the measurement operation, especially when measuring the outdoor noise pollution. The result from a measurement in the environment is only an approximation or estimate of the 'true' value, and thus, is only complete when accompanied by a statement of the uncertainty. The uncertainty leads in defining measurement accuracy, where the term 'accuracy' shows how a measurement result is close to the 'true' or accepted value.

Guidelines on estimating the measurement uncertainty are given by the ISO Guide to Uncertainty of Measurements [5] and the ISO 1996-2:2018 standard [6]. According to the standard, the measurement uncertainty for outdoor noise measurements is caused by several parameters: type of the sound level meter, the residual sound, type of source, meteorological conditions, and source location. These guidelines are followed by a lot of researchers, leading in producing significant contribution in the field of environmental noise measurement uncertainty [7,8,9].
On the other hand, the noise level can be also calculated by creating acoustic noise maps. In addition to this, by using the number of vehicles from the traffic flow in chosen urban area as input in noise mapping software, the researchers in [10] use noise-mapping techniques as a cartographic representation of the noise level in a defined area and period. By comparing the noise level from the measurement and the predictive noise maps, the results in [11] show similar noise levels, indicating the traffic as main noise source.
It can be concluded that the noise level results can be validated by comparing the results for the L_{eq} parameter estimated using different methodologies. Followed by this, the calculation of the noise uncertainty can help in evaluating the results within the uncertainty range, and from here, verify their accuracy.
According to the previous work, this paper aims to evaluate the accuracy of the equivalent noise pressure level parameter using two methodologies: noise measurements using hand-held analyzer from 1st class using the ISO 1996-2:2018 standard; and acoustic modelling by creating predictive noise maps that use the vehicles as an input sound source. To verify the proposed method, twenty short-term measurements are done in same spot in morning and afternoon period from Monday to Friday for two weeks; and afterwards, the measurement uncertainty is calculated for each measurement. Also, the L_{eq} is estimated from the predictive noise maps that are generated using the traffic flow as line source for the same time period as the measurements.

2 Estimation of the equivalent noise pressure level (L_{eq}) from road traffic noise

2.1 Measurement methodology in chosen urban area

Twenty short-term measurements of the traffic noise were carried out at one measurement point in the central part in the city of Skopje known to have constant noise pollution as shown. Figure 1 shows the measurement location and setting.
The measurements were done in two weeks period from Monday to Friday in the first week of June and the first week of September.

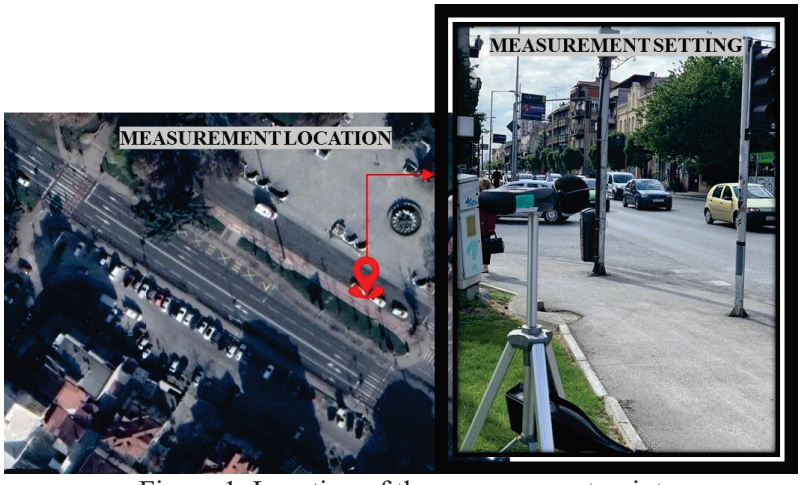


Figure 1. Location of the measurement point

The applied methodology for the noise measurements is based on the European Noise Directive and the international standards transposed by the national by-laws. For the noise measurements, 1st class hand-held analyzer Brüel & Kjær 2250 was used, using the A-weighting filter. Before the measurements, the hand-held analyzer was calibrated at 94 dB and 114 dB. Each measurement lasted 10 minutes, which is long enough to identify the road traffic noise sources. The hand-held analyzer was set at 1.5 meters height and 3 meters horizontal distance from the edge of the road. The atmospheric impact, such as temperature, humidity and wind speed have also been noted for every conducted measurement.

2.2 Modelling of acoustic predictive noise maps

Using the French Method for Road Traffic Noise Prediction (NMBP routes 96) which is transposed in the national by-laws, the traffic flow is considered as main parameter when creating acoustic prediction model. The acoustic modeling was made using the IMMI software. When creating the predictive noise maps, three operational phases must be considered: input of the topography and geometry, the road surface topology, the average vehicle speed, and the traffic flow as line noise source. The flowchart indicating the operational phases is shown on figure 2.
For the traffic flow input, the number of vehicles for the crossroad on the selected location was provided for the same period as the measurements were done from the Traffic Management and Control Center. After statistically processing and analyzing the traffic flow data, the number and type of the vehicles were set as an input noise source.

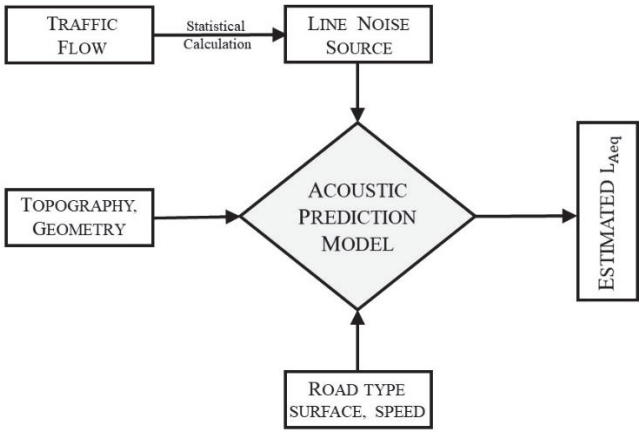


Figure 2. Flowchart of the methodology for creation of acoustic noise maps

By inserting the above-mentioned parameters, the IMMI software calculates the noise maps dispersion models, from where the equivalent sound level parameter L_{eq} can be obtained.

3 Measurement uncertainty estimation

Measurements of environmental noise are complex to perform because of the great number of variables that have to be considered. As each measurement occasion is subject to current source and meteorological conditions which cannot be controlled by the operator, it is often not possible to control the resulting uncertainty of the measurements. Instead, the uncertainty is determined after the measurements based on an analysis of the acoustic measurements and collected data on source operating conditions and on meteorological parameters important for the sound propagation.

The measurement uncertainty of sound pressure level depends on several sources: sound source, measurement time interval, meteorological conditions, distance from the source and the measurement method and instrumentation. According to ISO 1996-2 and the ISO/IEC Guide 98-3 (GUM), there are several methods for noise level estimation:

- Modelling approach by identifying and quantifying all major sources of uncertainty by using the uncertainty budget;
- Inter-laboratory approach by carrying out a round-robin test to determine standard deviation of reproducibility of the measurement method;
- Hybrid approach by combining the modelling and the inter-laboratory approach.

The modelling method consisted of determining the budget uncertainty is preferred and highly used method, especially when having short-term measurement. That is why, this method will be described and analyzed in the following paragraph, and later will be used for determining the measurement uncertainty of the applied measurements. According to the modelling method, the measurement uncertainty will be explained, focusing on the general model and the environmental noise measurement uncertainty.

3.1 General model

According to the ISO/IEC Guide 98-3, the mathematical general model for uncertainty estimation needs to identify and determine the measurement uncertainty of each source that causes uncertainty. The value of the measurement L is function of the measured parameters x_i that have influence on the uncertainty is determined:

$$L = f(x_1, x_2, x_3, \dots, x_j) \tag{1}$$

If each of the parameters x_j has standard uncertainty u_j, the combined standard uncertainty will be:

$$u(L) = \sqrt{\sum_{1}^{n} (c_j u_j)^2} \tag{2}$$

The parameters x_j are independents, and the c_j is sensitivity coefficient that is calculated as a function of the measured parameter:

$$c_j = \frac{df}{dx_j} \tag{3}$$

The overall measurement uncertainty is shown as expanded uncertainty, that is product of the combined standard uncertainty (u) and the numerical coverage factor (k):

$$U = ku \tag{4}$$

The coverage factor is a numerical factor used as a multiplier of the combined standard uncertainty to obtain an expanded uncertainty. The coverage factor is stated so that the standard uncertainty of the measured quantity can be used in calculating the combined standard uncertainty of other measurement results that may depend on that quantity. The value of the coverage factor is chosen based on the level of confidence (confidence level) required of the interval. The coverage factor is based on Gauss distribution. The mean $k = 1$ corresponds to 68%, the $k = 2$ to 95%, and $k = 3$ to 99.7% confidence interval.
Usually, the used coverage factor is 2 ($k = 2$), which confirms the result with 95% accuracy. The final given value is the corrected result with the expanded measurement uncertainty in the following form:

$$L \pm 2u \tag{5}$$

3.2 Budget uncertainty for outdoor environmental noise measurements

Determining the measurement uncertainty of environmental noise in road traffic measurements is complex operation to determine the function f as influence of the measured parameters. The estimated value during the specified conditions followed with the measurement uncertainty based on the measurement budget is calculated with the following equation:

$$L = L' + 10\lg\left(1 - 10^{-0,1(L'-L_{res})}\right) dB + \delta_{sou} + \delta_{met} + \delta_{loc} \quad (6)$$

Where:
L is the estimated value during the specified conditions expressed in decibels (dB);
L' is the measured value including residual sound;
L_{res} is the residual sound;
δ_{sou} is an input quantity that shows the uncertainty due to deviations from the expected operating conditions of the source;
δ_{met} is an input quantity to allow for any uncertainty due to meteorological conditions deviating from the assumed meteorological conditions;
δ_{loc} is an input quantity to allow for any uncertainty due to the selection of receiver location.

The equation (6) shows that each source of measurement uncertainty is function of several sources of uncertainty and in addition, it will be applied for the short-term measurements that were carried out for the purpose of this paper. The measurement uncertainty that depends on the sources δ_{sou} and the meteorological characteristics δ_{met} are determined directly from the measurements' conditions. The measured sound pressure level L' and the residual sound level L_{res} depend on the measurement uncertainty arising from the measuring instrument (δ_{slm}).
Table 1 shows the budget of the measurement uncertainty when measuring the sound pressure level, including all of the sources that can cause uncertainty, which are discussed in detail in section 3.3. According to the ISO 1996-2, the estimated value in dB from the source and meteorological conditions is considered to be 0, while the measurement uncertainties for both parameters vary, and their calculation is showed in section 3.3.

Table 1. Budget of measurement uncertainty

Value	Estimated value (dB)	Standard measurement uncertainty y, *uj* dB	Magnitude of sensitivity coefficient, *cj*	Determined measurement uncertainty, $c_j u_j$, dB
$L' + \delta_{slm}$	L'	$u(L')$	$\frac{1}{1-10^{-0,1(L'-L_{res})}}$	
δ_{sou}	0	u_{sou}	1	
δ_{met}	0	u_{met}	1	
δ_{loc}	0.0 – 6.0	u_{loc}	1	
$L_{res} + \delta_{res}$	L_{res}	u_{res}	$\frac{10^{-0,1(L'-L_{res})}}{1-10^{-0,1(L'-L_{res})}}$	
Combined measurement uncertainty			$u(L) = \sqrt{\sum_{1}^{n}(c_j u_j)^2}$	
Expanded measurement uncertainty (95% confidence level k=2), $2u$				
Final value				$L \pm 2u$

Depending on the sources that cause the measurement uncertainty in the measurement of environmental noise, the combined measurement uncertainty is calculated according to the formula:

$$u(L) = \sqrt{(c_{slm}u_{slm})^2 + (c_{sou}u_{sou})^2 + (c_{met}u_{met})^2 + (c_{loc}u_{loc})^2 + (c_{res}u_{res})^2} \quad (7)$$

3.3 Sources of measurement uncertainty

The measurement uncertainty can be caused by:

1. Standard measurement uncertainty from the sound level meter ($L' + \delta_{slm}$)

According to the standards IEC 61672 and ISO 1996-2:2018 [6, 12], for sound level meter that refers to a class 1, the standard uncertainty should be taken as 0.5 dB.

2. Standard measurement uncertainty from the source (δ_{sou})

For road traffic noise measurement, the standard measurement uncertainty depends on the number of vehicles n and the coefficient that depends on the type of vehicles C:

$$u_{sou} \cong \frac{C}{\sqrt{n}}\, dB \tag{8}$$

3. Standard measurement uncertainty from the meteorological conditions (δ_{met})

According to the meteorological conditions if the weather is favorable and stable then the default uncertainty will be:

$$u_{met,fav} = 2\, dB \tag{9}$$

This applies only when the horizontal distance between the noise source and the receiver is less than 400 meters ($D < 400\ m$).

4. Standard measurement uncertainty from the microphone location (δ_{loc})

Depending on the location of the microphone, the noise level can be adjusted between 0 and 6 dB. If the microphone is set on a distance lower than 0.5 meters of reflecting surface, the noise level should be corrected for 6 dB, while if the distance is between 0.5 to 2 meters, correction of 3 dB should be applied. For measurements in free field with more than 2 meters distance from reflecting surface, no correction is applied (the value should be set to 0). Because the measurements in this study are done in free field, this value is set to be 0.

5. Standard measurement uncertainty from the residual sound ($L_{res} + \delta_{res}$)

The residual sound while measuring road traffic noise can be calculated by comparing measured equivalent sound level with the residual sound which can be considered as the indicator L_{95} which represent the noise level present in 95% of the time.

4 Results

4.1 Short-term measurements

According to the proposed methodology in section 2, short-term noise measurements are done in the chosen location in urban area, and afterwards, the measurement uncertainty is calculated according to the methodology shown in section 3. Table 2 shows the calculated L_{eq} for each measurement with their measurement uncertainty. For the measurements number which is shown in the format MX-Y-Z, the M stands for measurement, the X is the number of the day of the measurement (ex. 1 for Monday, 2 for Tuesday etc.), the Y represents the week (1 is for the first week of June, while 2 is for the first week of September), and Z represents the time when the measurement is done (1 for the morning measurements and 2 for the afternoon measurements).

For example, for measurement M1-1-1, the calculated equivalent noise level is 67.28 with measurement uncertainty of $\pm$4.26. The measured L_{eq} is 67.33 dB, while the residual sound is 57.57 dB. By applying the formula (6), the calculated L_{eq} is measured to be 67.28 dB. Next, the measurement uncertainty is calculated. The measurement uncertainty for the sound level meters is $u(L') = 0.5$, while the magnitude of sensitivity coefficient is calculated due to the formula shown on table 1 and its value is 1.12. The measurement uncertainty from the source is calculated from formula 8 and its value is estimated to be 0.48, having $n = 432$ number of

vehicles with $C = 10$ coefficient for mixed types of vehicles, while the magnitude of the coefficient is set to be 1. Next, the measurement uncertainty from the meteorological condition is shown, which is set to be 2 as favorable meteorological conditions were followed during the measurements. As stated earlier, the microphone location was in free field, and according to the standard [6], the measurement uncertainty of the location is set to be 0, with magnitude of sensitivity coefficient 1. At the end, the measurement uncertainty from the residual sound has value 0, because according to the standard, if the values between the Leq and Lres differ more than 5 dB, then the uncertainty from the residual sound would be close to 0. As for the magnitude of the sensitivity coefficient, according to the formula shown on table 1, this value is 0.08. After evaluating all the measurement uncertainty parameters, the equation (7) is applied, from where the combined measurement uncertainty is calculated to be 2.13 dB. As the confidence level of the result is 95%, the expanded measurement uncertainty multiplied with 2, this leading to a value of 4.26 dB. From here, it could be stated out that the equivalent noise level for the measurement M1-1-1 is 67.28 ± 4.26 dB. This methodology is applied to all measurements, and the results could be seen on table 2.

Table 2. Calculated equivalent noise level and measurement uncertainty

Morning measurements (09:00 – 09:10) From Monday to Friday in the first week of June			
Measurement number	Date	Calculated L_{eq} (dB(A))	Measurement uncertainty (dB)
M1-1-1	31.05	67.28	±4.26
M2-1-1	01.06	67.73	±4.25
M3-1-1	02.06	71.8	±4.25
M4-1-1	03.06	72.23	±4.27
M5-1-1	04.06	71.56	±4.25

Afternoon measurements (14:00 – 14:10) From Monday to Friday in the first week of June			
Measurement number	Date	Calculated L_{eq} (dB(A))	Measurement uncertainty (dB)
M1-1-2	31.05	67.37	±4.24
M2-1-2	01.06	69.72	±4.24
M3-1-2	02.06	72.89	±4.26
M4-1-2	03.06	73.65	±4.26
M5-1-2	04.06	67.29	±4.28

Morning measurements (09:00 – 09:10) From Monday to Friday in the first week of September			
Measurement number	Date	Calculated L_{eq} (dB(A))	Measurement uncertainty (dB)
M1-2-1	30.08	73.18	±4.25
M2-2-1	31.08	74.92	±4.24
M3-2-1	01.09	75.33	±4.25
M4-2-1	02.09	73.3	±4.25
M5-2-1	03.09	76.7	±4.26

Afternoon measurements (14:00 – 14:10) From Monday to Friday in the first week of September			
Measurement number	Date	Calculated L_{eq} (dB(A))	Measurement uncertainty (dB)
M1-2-1	30.08	73.27	±4.23
M2-2-1	31.08	73.94	±4.26
M3-2-1	01.09	74.28	±4.24
M4-2-1	02.09	72.47	±4.25
M5-2-1	03.09	73.36	±4.25

As it can be noticed from the results, for the measurements in the first week of June, the noise level is in the range of $67.28 - 73.65$ dB(A). The expanded measurement uncertainty is calculated due to the proposed uncertainty budget, and it is in range of $4.24 - 4.28$ dB(A). This means that the result of the calculated L_{eq} could vary between the calculated measurement uncertainty for each measurement.
From the results from the first week of September higher values of L_{eq} could be noticed. According to the number of vehicles and the traffic in this week, there could be noticed higher number of vehicles. Due to this, we can find relation why the noise level is higher.

4.2 Acoustic predictive noise maps

In relation with the measured noise levels in the proposed time intervals, the number of vehicles was provided by the State Traffic and Control Management Center. In order to make comparison between the measured value of the L_{eq} and the predicted value of L_{eq} from the acoustic prediction maps generated in the IMMI software, 20 different maps were generated. Figure 3 shows one representative predictive noise map, while table 3 shows the results of the calculated L_{eq} from the noise maps for each measurement with no uncertainty estimate.

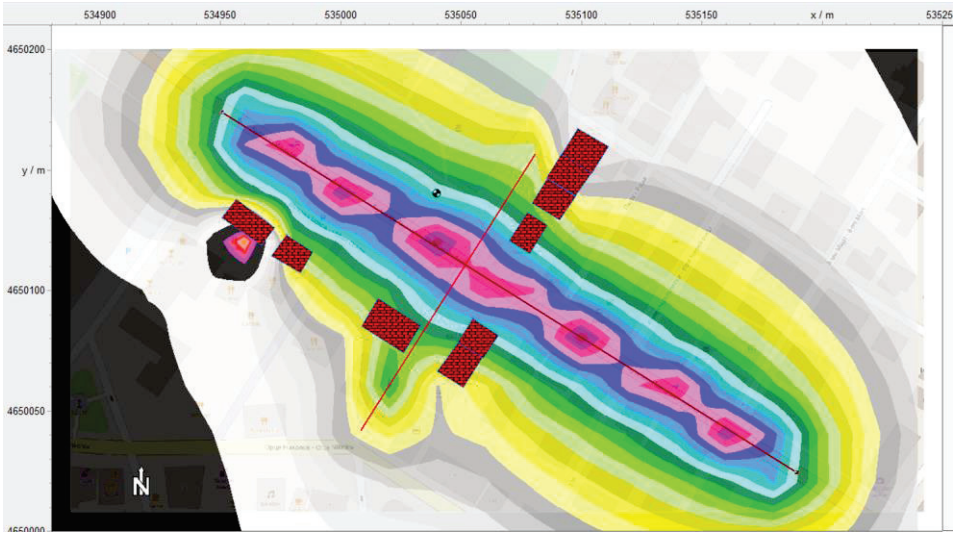


Figure 3. Representative predictive noise map generated from the IMMI software

Table 3. Results from the predicted equivalent noise level L_{eq}

Results from IMMI software (09:00 – 09:10) From Monday to Friday in the first week of June		
Measurement number	Date	Calculated L_{eq} (dB(A))
M1-1-1	31.05	65.43
M2-1-1	01.06	66.4
M3-1-1	02.06	69.78
M4-1-1	03.06	70.54
M5-1-1	04.06	70.77

Results from IMMI software (14:00 – 14:10) From Monday to Friday in the first week of June		
Measurement number	Date	Calculated L_{eq} (dB(A))
M1-1-2	31.05	66.37
M2-1-2	01.06	70.12
M3-1-2	02.06	70.32
M4-1-2	03.06	71.62
M5-1-2	04.06	65.9

Results from IMMI software (09:00 – 09:10) From Monday to Friday in the first week of September		
Measurement number	Date	Calculated L_{eq} (dB(A))
M1-2-1	30.08	70.28
M2-2-1	31.08	74.12
M3-2-1	01.09	73.78
M4-2-1	02.09	73.24
M5-2-1	03.09	72.52

Results from IMMI software (14:00 – 14:10) From Monday to Friday in the first week of September		
Measurement number	Date	Calculated L_{eq} (dB(A))
M1-2-1	30.08	69.7
M2-2-1	31.08	71.27
M3-2-1	01.09	72.18
M4-2-1	02.09	70.98
M5-2-1	03.09	74.16

4.3 Comparison between the results

When comparing the results from calculated L_{eq} generated from the measurements shown on table 2 with the predicted L_{eq} from the IMMI software shown on figure 3, there could be noticed deviation between the results. The results for the L_{eq} parameter from the measurements and from the acoustic noise maps were compared in accordance with the calculated measurement uncertainty. On figure 4, the measurement uncertainty is shown as a range, while the value shows the difference between the measured and the predicted L_{eq} in dB(A). From the comparison, it could be stated out that the differences between the results are within the limits of the measurement uncertainty, which confirms that the used methodologies are correctly proposed and used.

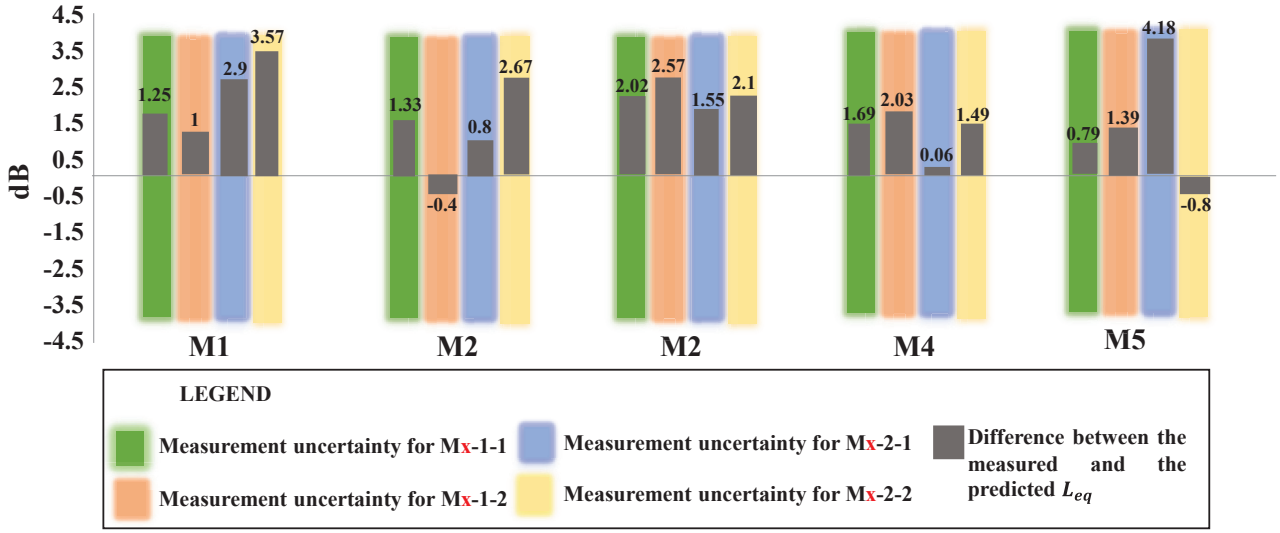


Figure 4. Difference between the measured and the predicted L_{eq}

In 90% of the measurements, the values from the calculated equivalent noise level from the acoustic maps are lower than the measured ones. This is expected, as for the acoustic noise maps the vehicles are the only noise source, while when doing the measurements, there are more parameters causing noise. That is why, when measuring the noise level with hand-held analyzer, in order to get the real value, the measurement uncertainty should be calculated.

5 Conclusions

The outdoor measurements of the environmental noise pollution can be complex to perform as there are several parameters that cause measurement uncertainty. By using the ISO 1998-2:2018 standard, the sources of uncertainty were defined and calculated, providing the equivalent noise level parameter with 95% confidence interval. The results were further compared by estimating the L_{eq} parameter from the predictive noise maps that use the traffic flow as an input parameter. Given the measurement accuracy, by comparing the results for the L_{eq} parameter, the accuracy of the proposed methodologies and the results was confirmed.

References

[1] Singh, N., & Davar, S. C. (2004). Noise pollution-sources, effects and control. *Journal of Human Ecology*, *16*(3), 181 Singh, N. and Davar, S.C., 2004. Noise pollution-sources, effects and control. *Journal of Human Ecology*, *16*(3), pp.181-187.

[2] Nazneen, S., Raza, A. and Khan, S., 2020. Assessment of noise pollution and associated subjective health complaints and psychological symptoms: analysis through structure equation model. *Environmental Science and Pollution Research*, *27*(17), pp.21570-21580.

[3] Yang, W., He, J., He, C. and Cai, M., 2020. Evaluation of urban traffic noise pollution based on noise maps. *Transportation Research Part D: Transport and Environment*, *87*, p.102516.

[4] Cvetković, D., Praščević, M. and Mihajlov, D., 2011. Estimation of uncertainty in environmental noise measurement. *measurement*, *100*, p.3.

[5] Iso, I. and OIML, B., 1995. Guide to the Expression of Uncertainty in Measurement. *Geneva, Switzerland*, *122*, pp.16-17.

[6] ISO 1996-2:2018 Acoustics – Description, Measurement and assessment of environmental noise – Part 2: Determination of sound pressure levels

[7] Sonia Alves, David Waddington. Estimation of Uncertainty Using revised draft ISO 1996-2. Forum Acusticum 2014

[8] Lindsay Hannah, Wyatt Page and Stuart McLaren. An introductory guide to uncertainty in acoustic measurements. New Zeland acoustics

[9] Manvell, D., 2020. Uncertainty in International Acoustics Standards. In *Uncertainty in Acoustics: Measurement, Prediction and Assessment* (pp. 431-443). CRC Press.

[10] Kim, P., Ryu, H., Jeon, J.J. and Chang, S.I., 2021. Statistical road-traffic noise mapping based on elementary urban forms in two cities of South Korea. *Sustainability*, *13*(4), p.2365.

[11] Domazetovska, S., Anachkova, M., Gavriloski, V., Sokolikj, A. and Stojkovska, S., 2021, August. The impact of COVID-19 lockdown on the noise pollution: case study in the city of Skopje. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (Vol. 263, No. 4, pp. 2087-2094). Institute of Noise Control Engineering.

[12] IEC-International Electrotechnical Commission, 2013. IEC 61672-1.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Statistical analysis of urban noise measurement data: case study for the city of Skopje

Maja Anachkova[1,*], Simona Domazetovska[2], Filip Nikolovski[3], Viktor Gavriloski[4]

[1,2,3,4]Faculty of Mechanical Engineering in Skopje, North Macedonia

*maja.anachkova@mf.edu.mk

Abstract

Noise from the road transport, particularly from vehicles in urban city areas largely accounts for the general noise level and annoyance of the citizens. The numerous volumes of motor vehicles flow can be treated statistically, which can establish a deeper insight into the contribution of the road noise to the prevalent noise pollution and its' characteristics. According to ISO 362 and ISO 1996:2, the environmental noise level from traffic is highly dependent on the vehicle category regarding the factor of contribution to the overall urban noise level. The purpose of this study is to analyze the dependence between the number and types of vehicles and measured standardized parameters (L_{eq}, L_{AF} and L_{95}) for noise level assessment by implementing a statistical model analysis of the collected results. The number and the type of the vehicles is obtained from the States' traffic management and control center for a chosen road in the center of the city, whereas noise level measurements have been conducted with a Bruel&Kjaer sound level meter by using a standardized noise level measurement methodology procedure for the selected period on the given location. This study provides a detailed statistical approach of the collected noise and traffic volume data to obtain conclusions and prediction models for further management of the noise pollution problem in the city.

Keywords: road traffic noise, statistical analysis, urban noise level measurement

1 Introduction

In the past decade, the noise from high density traffic networks in urban cities including residential areas has been rapidly increasing and evidently contributes to citizens' annoyance **[1]**. Consequently, methods for evaluation and assessment of the urban road-traffic noise need to be established as a part of a general sustainable management plan for its' prediction and control. Common method to facilitate systematic management of the road traffic noise are the noise maps of cities introduced in relation to the Environmental noise directive of the European Parliament and Council (Directive 2002/49/EC). Noise maps contribute to understanding the noise situation in the city, including the location of hot spots and the specific noise level of a specific noise source to which the population is exposed. Nevertheless, when a noise map is not available for an existing city, a statistical relationship between the urban form indicators and parameters and the road-traffic noise can be practical and efficient**.**

Many researches have been conducted in order to explore the traffic and road noise impact on the noise pollution in urban cities and some of them are reviewed in this section. Khan, J. et.al. in **[2]** have presented a review paper where current methods of modelling and exposure assessment techniques for road traffic noise pollution are analyzed. In **[3]**, Akgüngör et.al. have conducted a study to explore traffic-based noise pollution in the city of Kırıkkale in Turkey whereby using a multi regression model, the effect of traffic density and composition of the traffic noise level in the city has been investigated. Pobedin et. al in **[4]** have performed a range of experimental research of the noise emissions of the passenger car individual elements by creating a mathematical model for computation of the expected noise in passenger car cabin and a method of probabilistic

computational evaluation of car noise. The system discussed in paper **[5]** by Zambon et. al. has applications in terms of traffic noise prediction in large urban environments and obtains an accurate description of traffic noise by conducting measurements of road noise from monitoring stations positioned over a zone of interest. In **[6]**, the authors have provided a comparative analysis between the measured and predicted noise levels obtained by a traffic noise propagation model combined with an urban traffic noise map. Ramirez et. al. in **[7]** elaborate the development and evaluation of a stochastic dynamic traffic noise prediction model for the city of Bogotá established on noise functions for vehicle classes, their speed, and different types of vehicles. In **[8],** the same relation is investigated by means of numerical calculations for two cities, Amsterdam and Rotterdam. Researchers in **[9]** describe the roadside traffic noise assessment conducted in dense built-up urban areas in Hong Kong by using multiple regression to provide a set of empirical formulas for predicting road traffic noise level. Measurement data included statistical analysis including the number of light vehicles, the number of heavy vehicles, the total traffic flow, and the time period on the given location is presented. In Section 3, the noise mapping for the discussed area is presented. In Section 4 the results obtained are analyzed and explained, and in Section 5 the conclusions from the research are provided.

2 Methodology

The purpose of this research is to provide a statistical analysis between actual measurements of noise levels in a chosen area in the city center and the number of vehicles circulating through the streets which limit that area. For this purpose, it is necessary to analyze the correlation of these results in order to gain a better insight into the relationship between the noise level and the number of vehicles, as well as the type of vehicles that make up the traffic in that area. The methodology for providing the results needed to conduct this analysis primarily involves realistic noise level measurements at the selected location to assess the noise level. This was done through short-term measurements at a point near an intersection in the city center which was previously determined to be one of the busiest intersections in traffic. The traffic at that intersection which is formed by two streets is two-way, on one street there are 4 traffic lanes, for each two, and the second street with which it is intersected has one traffic lane for each direction. Namely, the measurements were conducted at the same measuring point during two working weeks in the two most frequent months of the year, June and September. The acoustic measurements were carried out two times a day, at the morning and at the evening, during a 10-minute sampling interval. The 20 acoustic measurements have been conducted by using a Bruel&Kjaer 2250 Class 1 sound level meter, mounted at the point of measurement, at a height of 1.5 meters in the direction of a sound source 3.5 meters from the most exposed façade.

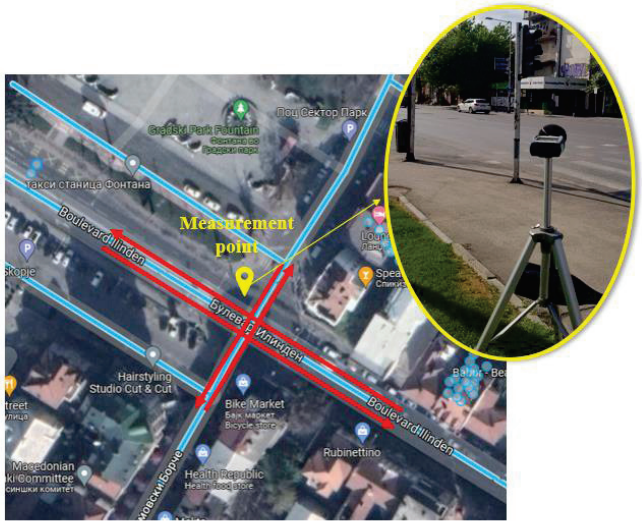


Figure 1. Measurement location (crossroad) in the city centre

At each measuring point the noise level is represented by the L_{eq}, L_{AF} and L_{95} parameter in accordance with recommendations explained in **[1].**

As the traffic jams at this intersection significantly affect the total noise level at the intersection, it was necessary to provide data on the number of vehicles passing at that intersection. For this purpose, data were obtained from the Traffic Control and Management Center, which is a state institution obliged to monitor traffic density across city roads through a network of sensors set on every crossroad in the city. The data were obtained in the form of the number of vehicles that passed on both streets in one hour, and then they were standardized to obtain the total number of vehicles that passed at the intersection for a period of 10 minutes while the measurement was performed. On the other hand, another aspect related to the overall noise level on the certain location is that different types of vehicles such as passenger cars, trucks and motorcycles contribute differently to the overall noise level and cause different noise levels **[15].** Therefore, during the 10-minute acoustic measurement, the number of heavy and medium vehicles, as well as motorcycles was manually counted and carefully noted. These numbers were used as an input parameter in the statistical analysis that was conducted to provide a correlation between the total sound level on the location and the number of vehicles, but also a relation between the types of vehicles and their contribution to the overall noise level.

3 Acoustic noise maps

The data for the number of vehicles and the type of vehicles for the 20 measurements on the crossroad were used as an input in the noise mapping software IMMI in order to create an acoustic predictive noise map for the area. This software package is used for professional and research purposes and is used for static mapping and prediction of environmental noise. For the calculations, IMMI uses the national guidelines and standards and this map was calculated according to the French National Method for Calculating Noise Indicators for Street Traffic [NMPB Routes], that classifies the equivalent noise level for the day time (from 7am-19 pm), evening time (from 19pm-23pm) and night time (from 23pm-7 am).

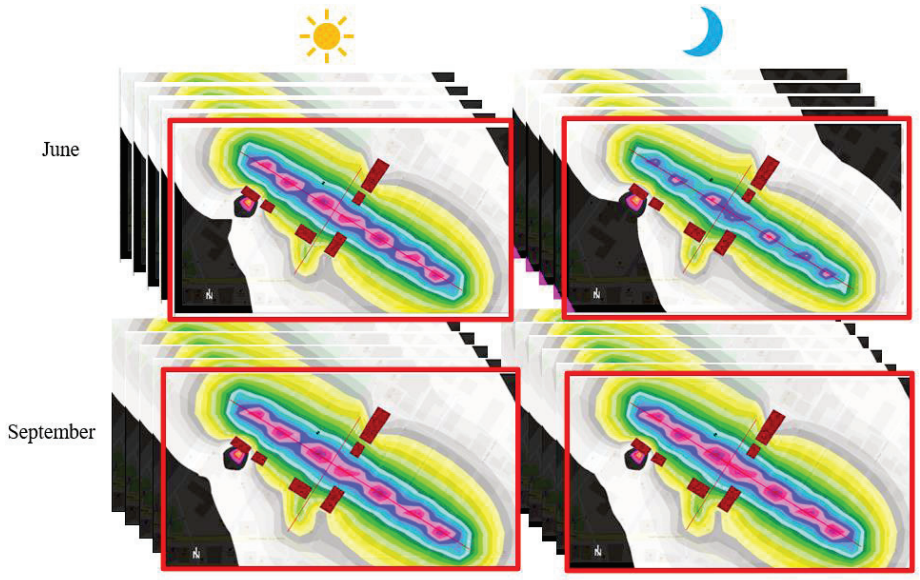


Figure 2. Four representative noise level dispersion maps for morning and evening period of two days during the measurement week

The correct input of the sound sources such as the total traffic flow (number of vehicles) and the types of vehicles for the specific streets and periods of the day is an important aspect in the process of creating the maps. In this case, additional interventions before inserting the traffic noise sources should be made in terms of defining the height of the objects as a part of the infrastructure around the area. Namely, for proper

calculation of the sound propagation, the heights of the buildings are manually inserted into the model after conducting previous calculations of the individual objects heights in the area. The number of vehicles, average speed and types of vehicles characterizes the urban traffic noise and are required to be input in the software model. In this map, the number of vehicles for the two streets forming the crossroad, differing the types of vehuicles as heavy, medium and motorcycles over the two periods of the day (day and evening) The average traffic speed on the boulevards is 50 km/h. Having these parameters, the program calculates the noise maps dispersion models and the obtained results for the equivalent noise level L_{eq} for the day and evening scenarios expressed in L_d and L_e are presented in Fig.2. The calculation is visualized using the vehicle flow on the two streets modeled in the software as line sources of noise.

4 Results from the statistical analysis and discussion

In order to obtain conclusions about the impact of the traffic volume and different types of vehicles on the total noise level at the selected location, the linear correlation statistical method was performed on the obtained data. It was noticed that data for the number of medium and heavy vehicles contain matching extremes (Fig.3, left), but the same data are not extremes for the number of motorcycles. Therefore, the data analysis approach was decided to be looked at from three aspects: with extremes, without extremes and ordinally (considering only the ranking of values, not the measured / enumerated values). The correlation between the number of vehicles by type (medium, heavy, motorcycles), the number of vehicles on both streets and the total number of vehicles with the variables that emit noise levels (L_{eq}, L_{AF}, L_{A95}) was compared. The aim was to determine how the noise level depends on the number and type of vehicles.

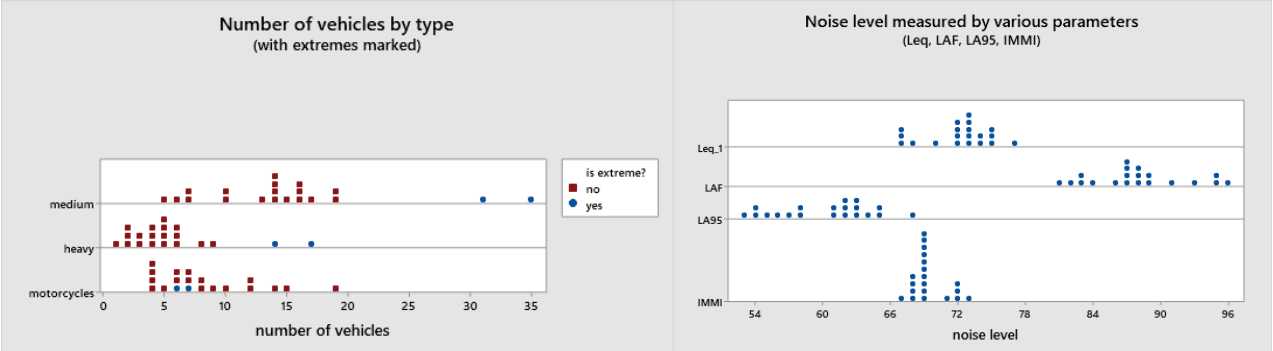


Figure 3. Review of number of vehicles by type (left) and values for the measured noise parameters (right)

After the initial overview of the data, three scatter plots were created. Each of them shows the number of vehicles on the x-axis and the corresponding noise measure (L_{eq}, L_{AF}, L_{95}) on the y-axis. Additionally, the vehicle type (with a different symbol and color for each type) was marked and added a simple linear model for

each type to see the trend (trendline). This approach and these elements were chosen because it is easy to see the trend of the noise measure according to the number of vehicles.

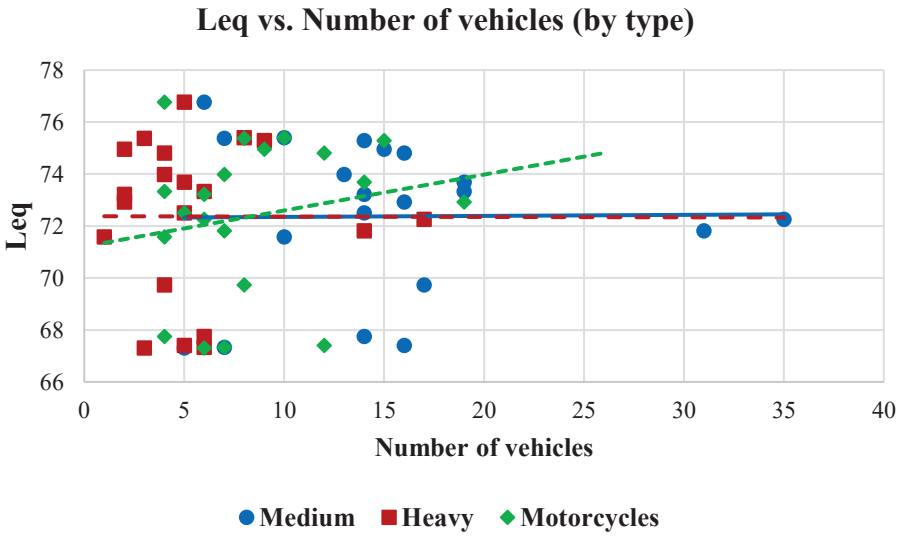


Figure 4. Correlation between the number of all three types of vehicles and L_{eq} parameter

From the initial analysis, it is obvious that the more vehicles were included in the traffic, the higher noise level was measured (positive correlation). The dependence is moderate in all cases. The volume of traffic on the main boulevard Ilinden is most influential when it comes to noise level, and the impact of the intersecting street is smaller. On the other hand, medium and heavy vehicles are weakly correlated with the noise level, except when considering the normal (ordinal) values for heavy vehicles. In this case the correlation is negative and moderate, which means that in the days ranked as days with higher volume of heavy vehicles are in relation to days ranked as days with lower noise levels (Fig. 4).

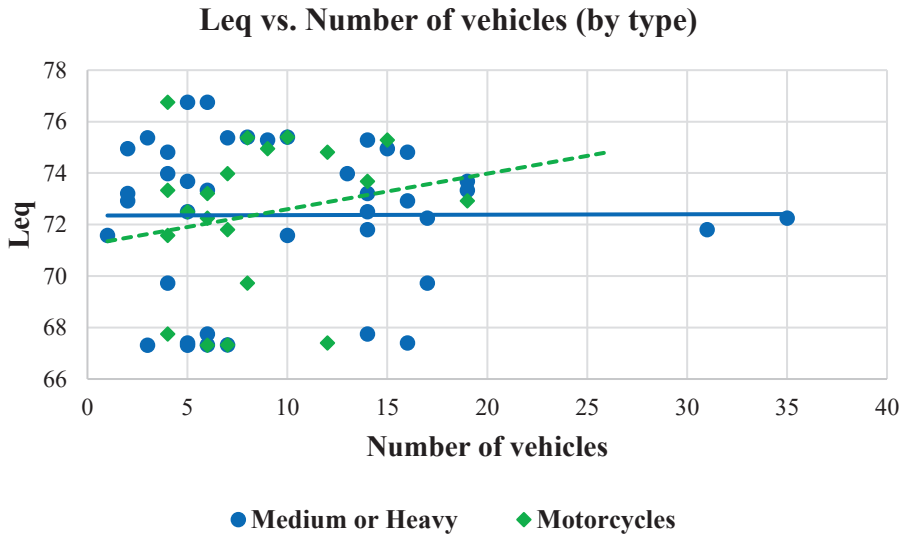


Figure 5. Correlation between the number of medium and heavy vehicles and motorcycles on L_{eq} parameter

For data with extremes, there is a moderate positive correlation between the number of vehicles by type and noise level. This is an indication that there is a dependence between noise level and vehicle type, but this dependence is moderate. By rejecting the extremes or looking at the ranked data, the correlation between medium and heavy vehicles on the one hand and the noise level on the other is significantly reduced. It is interesting that there is a reverse trend in motorcycles where the correlation is increasing (Fig.5). The number

of motorcycles is a stable indicator of the noise level (there is a moderate positive correlation with the level in all cases). The volume of traffic on Ilinden and the total volume of traffic are a good indicator of the noise level (positive, fairly strong correlation). According to the type of vehicles, a modest correlation with the noise level can be noted.

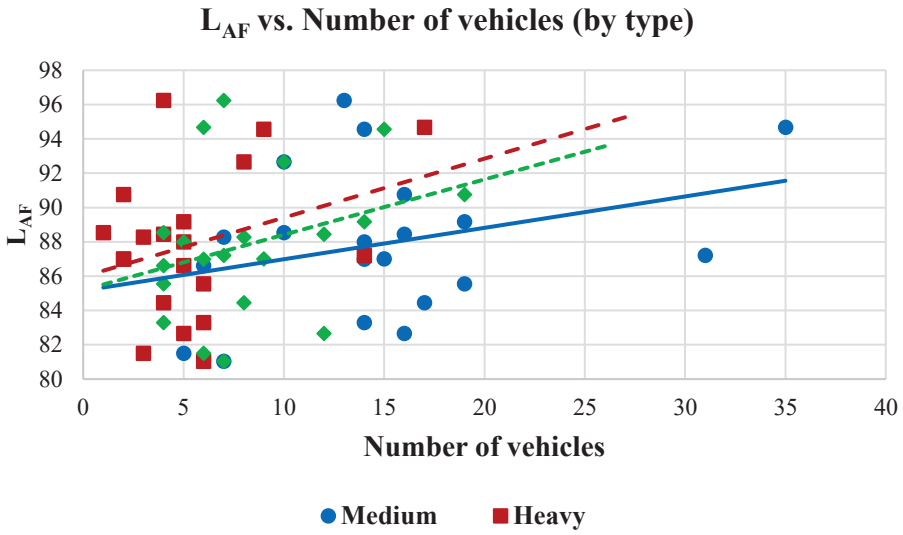


Figure 6. Correlation between the number of medium and heavy vehicles and motorcycles on L_{AF} parameter

What is interesting, even unexpected is the negative dependence in some cases: a larger number of vehicles are associated with a lower noise level. In heavy and medium vehicles, huge changes occur in the results based on the approach of analysis. The volume of traffic on Ilinden and the total volume of traffic are a good indicator of the noise level (positive, fairly strong correlation). The dependence between vehicle type and noise level is quite modest, even weak (except for medium vehicles where the correlation is moderate and negative).

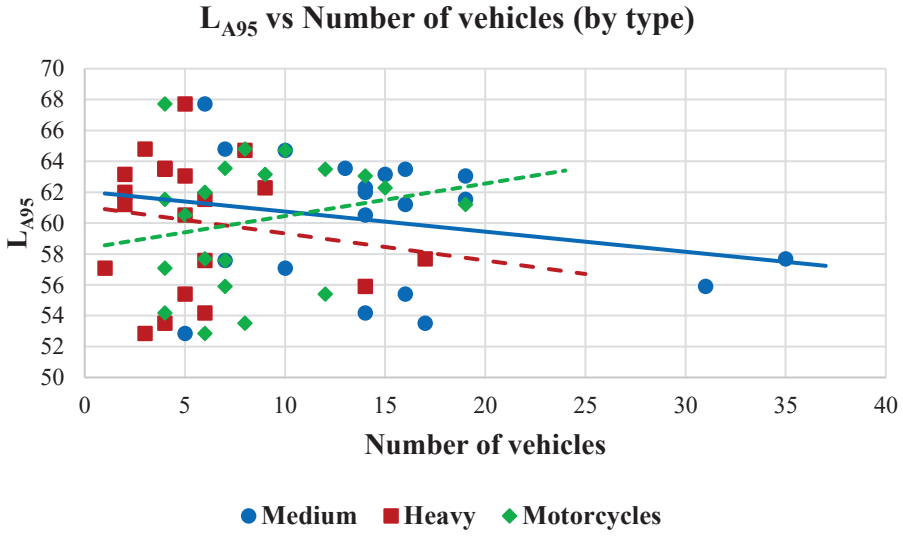


Figure 7. Correlation between the number of medium and heavy vehicles and motorcycles on L_{A95} parameter

From the statistical analysis can be concluded that the number of motorcycles is the most stable indicator of the noise level which has been noticed in all combination analyzes. On the other hand, it was concluded that it is preferable to work with L_{AF} (Fig. 6) or L_{A95} (Fig. 7) indicators for noise level. In some cases, there is a negative correlation between the number of vehicles of a certain type and the noise level (for example medium

vehicles and L_{A95} as shown in Fig.7). This is an indication that on average a larger number of medium vehicles means less noise, meaning the dominant noise is caused by another type of vehicle.

5 Conclusions

This paper presents the results from a conducted statistical analysis approach of the impact of traffic density, as well as the type of vehicles that make up the traffic on the overall noise level measured through three parameters (L_{eq}, L_{AF}, L_{A95}) at a selected crossroad in the center of Skopje. To obtain the input parameters for the noise level, 20 short 10-minute noise measurements were made at the selected location, and afterwards, data for the number of vehicles that passed through the intersection in those periods, as well as the type of vehicles that contributed to the traffic was obtained.

From the exposed detailed analysis, can be concluded that the heavy traffic at the selected location has a large, dominant impact on the noise level at that location. At the same time, it can be concluded that the different types of vehicles that were considered in this study have different contributions to the overall noise level. It was noticed that the most stable indicator of increased noise level are the motorcycles, and the parameters L_{AF} and L_{A95} are the most stable parameters of increased noise level, because in general they have shown a positive correlation with the total number of vehicles.

In order to provide more detailed results from the application of this statistical method for assessing the impact of traffic noise on the total noise level, it is desirable to provide additional more frequent and longer measurements at the selected location to be able to draw stronger conclusions in future. It is also necessary to take into account additional factors of influence such as the age of the vehicles, the type of engine and the speed and also, incidental short-term vehicle noises, such as whistling or siren, should also be considered. All this will add input to the models for statistical analysis and will deepen the analysis, as well as the conclusions from it.

For future work, it is proposed and planned to provide more detailed data through conducting more frequent and longer measurements and considering additional parameters for the traffic contents. Furthemore, creation of more detailed maps for noise level dispersion to provide a predictive method and monitor the level of road traffic noise in the city should not be mistreated.

References

[1] EU Directive. Directive 2002/49/EC of the European parliament and the Council of 25 June 2002 relating to the assessment and management of environmental noise. Off. J. Eur. Commun. 2002, 12, L189.

[2] Khan, J., Ketzel, M., Kakosimos, K., Sørensen, M. and Jensen, S.S., 2018. Road traffic air and noise pollution exposure assessment–A review of tools and techniques. *Science of The Total Environment*, *634*, pp.661-676.

[3] Akgüngör P.A. & Demirel, A. (2008) Investigating urban traffic based noise pollution in the city of Kirikkale, Turkey, Transport, 23:3, 273-278, DOI: 10.3846/1648-4142.2008.23.273-278

[4] Pobedin, A.V., Shekhovtsov, V.V. and Dolotov, A.A., 2017. Computational probabilistic evaluation of passenger cars noise level. *Procedia Engineering*, *206*, pp.1558-1563.

[5] Zambon, G., Roman, H.E., Smiraglia, M. and Benocci, R., 2018. Monitoring and prediction of traffic noise in large urban areas. *Applied Sciences*, *8*(2), p.251.

[6] Di, H., Liu, X., Zhang, J., Tong, Z., Ji, M., Li, F., Feng, T. and Ma, Q., 2018. Estimation of the quality of an urban acoustic environment based on traffic noise evaluation models. *Applied Acoustics*, *141*, pp.115-124.

[7] Ramírez, A. and Domínguez, E., 2013. Modeling urban traffic noise with stochastic and deterministic traffic models. *Applied Acoustics*, *74*(4), pp.614-621.

[8] Salomons, E.M. and Pont, M.B., 2012. Urban traffic noise and the relation to urban density, form, and traffic elasticity. *Landscape and Urban Planning*, *108*(1), pp.2-16.

[9] To, W.M., Ip, R.C., Lam, G.C. and Yau, C.T., 2002. A multiple regression model for urban traffic noise in Hong Kong. *The Journal of the Acoustical Society of America*, *112*(2), pp.551-556.

[10] Ryu, H., Park, I.K., Chun, B.S. and Chang, S.I., 2017. Spatial statistical analysis of the effects of urban form indicators on road-traffic noise exposure of a city in South Korea. *Applied acoustics*, *115*, pp.93-100.

[11] Seong, J.C., Park, T.H., Ko, J.H., Chang, S.I., Kim, M., Holt, J.B. and Mehdi, M.R., 2011. Modeling of road traffic noise and estimated human exposure in Fulton County, Georgia, USA. *Environment international*, *37*(8), pp.1336-1341.

[12] Say, N., Okten. O., Aysu, A., Yalcinkaya, M.N., 2017. Seasonal Change and Spatial Distribution of Traffic Noise in Seferihisar, Turkey. ARCHIVES OF ACOUSTICS Vol. 42, No. 4, pp. 631–642.

[13] Gaja, E., Gimenez, A., Sancho, S., Reig, A., 2002. Sampling techniques for the estimation of the annual equivalent noise level under urban traffic conditions. Applied Acoustics 64 (2003) 43–53

[14] Przysucha, B., Szeląg, A., Rosa, W. and Kania, K., 2021. Representativeness of the measurement sample in the determination of the long-term noise indicators of traffic noise. *Measurement*, *180*, p.109550.

[15] Garcia, A. and Faus, L.J., 1991. Statistical analysis of noise levels in urban areas. *Applied Acoustics*, *34*(4), pp.227-247.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Assessment of the proposed changes regarding heavy vehicles in the statistical pass-by draft standard ISO/DIS 11819-1 (2021)

Michiel Geluykens[1,*], Ablenya Grangeiro de Barros[1], Luc Goubert[2], Cedric Vuye[1]

[1]EMIB Research Group, Faculty of Applied Engineering, University of Antwerp, 2020 Antwerp, Belgium.

[2]Belgian Road Research Centre, Woluwedal 42, 1200 Brussels.

*michiel.geluykens@uantwerpen.be

Abstract
Exposure to excessive (road) traffic noise can result in sleep deprivation, increased stress, and even heart diseases. Accordingly, accurately determining the acoustic quality of different pavement types is crucial for implementing noise-reducing measures such as low-noise pavements. In this regard, the ISO 11819-1:1997 standard, which describes the Statistical Pass-By method for assessing the acoustic performance of road pavements, is currently under revision. The proposed changes include combining two vehicles' categories, currently defined as dual-axles (HD) and multiple-axles (HM) heavy vehicles, into a single category: heavy vehicles (H), by adding 2.7 dB to the maximum sound pressure level (L_{Amax}) of HD vehicles. Additionally, a new method is presented to define the noise versus speed relation of this new caterory. In this work, these changes were evaluated based on a large dataset collected by the authors over 8 months, with an abundance of heavy vehicles (2820) while also accounting for passenger cars (2896). The average difference in L_{Amax} between HD and HM vehicles ranged between 2.1 – 2.3 dB, meaning that the draft standard may overestimate this difference. Furthermore, while the draft standard prescribes a coefficient of 25 for the L_{Amax} variation with speed for the road surfaces here investigated, we observed a lower speed dependency, with a coefficient of 20.5 on average. Lastly, we proposed and investigated the introduction of a new vehicle category containing delivery vans. It appears that vans can form a vehicle category with consistent noise behaviour and could thereby aid in describing the pavements' acoustic performance, ultimately contributing to reducing the measurement duration. The noise produced by vans presents a similar spectra shape and coincident noise levels as passager cars at 1000 to 2000 Hz third-octave bands.

Keywords: Road traffic noise, SPB, heavy vehicles, ISO 11819-1

1 Introduction

Road traffic noise is the largest contributor to environmental noise in urban areas. Excessive road traffic noise exposure is linked to non-auditory adverse health outcomes such as the incidence of ischemic heart diseases, cortical awakenings and self-reported sleep disturbance, among others [1][2][1]. In short, it decreases one's quality of life, and furthermore, the Environmental Noise Guidelines for the European Region indicate that over one million healthy years of life are lost yearly due to traffic-related environmental noise in Western Europe. With the increase in traffic and densely built-up areas, more and more people are exposed to (road) traffic noise [3]. In order to abate this problem by various means, there must first be a solid base for assessing road traffic noise and its influencing factors.
Several standardized methods already exist for this purpose, for example, UNECE R.117 contains requirements for the determination of sound emissions of tyres, and uses the same coast-by measurement principle as in ISO 13325:2019. ISO 11819-2:2017 describes the Close ProXimity method (CPX), the On-Board Sound Intensity method is described in AASHTO T 360-16:2016, and ISO 11819-1 defines the

Statistical Pass-By method (SPB), which all serve to evaluate the the acoustical quality of the pavement. In more detail, in ISO 11819-1 the maximum A-weighted noise levels' L_{Amax}' of passing vehicles, also called 'pass-bys' are captured together with the respective vehicle speed at the roadside under free-flow traffic conditions [4]. The currently active standard, ISO 11819-1:1997, is under revision (ISO/DIS 11819-1, 2021) [5].

The overall noise emitted by a vehicle passage is the combination of the contributions from several noise mechanisms, commonly categorised under power unit noise, aerodynamic noise and tyre/road noise [6]. Power-unit noise is the mechanical and combustion noise resulting from vehicle propulsion; aerodynamic noise originates from the turbulent airflow around the body (at higher speeds), and lastly, the tyre/road noise is generated by the interaction between the tyre and the pavement. Aerodynamic noise is negligible concerning noise emissions to the environment, and tyre/road noise is dominant over power unit noise already at moderate speeds (about 30 km/h for cars and 50 – 60 km/h for heavy vehicles). For this reason, the pavement plays a critical role in the generation of road traffic noise. Tyre/road noise itself is a convoluted combination of various generation and amplification mechanisms that result in noise in different, but often overlapping, frequency ranges[7]. However, the noise generation differs amongst vehicles on the roads, and moreover, it is also dependent on temperature and vehicle speed [8]. This complicates using free-flowing traffic conditions for the purpose of acoustic characterisation of pavements, as in ISO 11819-1. To some extent, the noise emissions are similar for certain vehicle types. By defining vehicle categories containing specific types of vehicles, it is assumed in ISO 11819-1 that the noise behaviour of the vehicles in each category is sufficiently similar to describe the pavement's acoustic performance. The allocation of vehicles to the categories is made visually, therefore the category descriptions are based on vehicle appearance. The standard describes the following categories; Passenger cars (P), Dual-axled Heavy vehicles (HD) and Heavy vehicles with more than two axles (HM) [4]. These categories are then separately analysed further; first, the L_{Amax} is corrected for temperature. Next, in order to exclude the impact of speed on the measurement, a linear regression is applied to the temperature corrected L_{Amax} and logarithm of the speed of the vehicle pass-bys in the same vehicle category. Finally, the formula of the resulting regression line, see Eq. 1, is used to calculate the L_{Amax} at a reference speed, which represents the acoustic performance of the pavement for that specific category. This value can be used to compare pavements directly, or to calculate the Statistical Pass-By Index (SPBI), a value where the contribution of the different vehicle categories is weighted.

$$L_{\mathrm{Amax}} = \mathrm{A} + \mathrm{B} * \log(v) \tag{1}$$

To ensure the accuracy of the method, sufficient data is needed to construct the linear regression model. For this reason, the current ISO 11819-1 sets a minimum number of 100 pass-bys for passenger cars and 30 for both HD and HM. Additionally, the vehicle categories of HD and HM combined, also called 'Heavy vehicles', should contain at least 80 pass-bys. However, reaching these minima often proves difficult, specifically regarding the HD and HM categories. While there usually is an abundance of passenger car vehicles, the number of valid HD and HM pass-bys is much lower, and often one of the categories is almost absent which extends the duration of the measurement [9]. Lowering the minima therefore would be of great practical and economic interest. For this reason, methodological changes are included in a new draft version of the ISO standard. Most notably, the minimum required number of heavy vehicles is reduced to 40, and the individual minimum for these categories is abolished [5].

However, only reducing the minimum number of vehicles would compromise the statistical accuracy of the method. For this reason, the draft standard proposes a new approach for processing the data for these categories. Firstly, the data of HD and HM are combined into a single category after the HD pass-bys are 'normalised' to HM pass-bys by adding 2.7 dB to their L_{Amax} levels. Still, the spread of pass-by speeds within the heavy vehicles category is generally lower compared to the passenger car category, and simultaneously, the variation in L_{Amax} is larger, which results in large uncertainties when applying the linear regression method [9]. To solve this issue, Eq. 1 relating L_{Amax} and the logarithm of speed is constructed not using the linear regression analysis. Instead the slope, also referred to as speed coefficient 'B', is assumed to be a predefined semi-generic value selected based on the pavement type. The intercept 'A' is calculated as the arithmetic average of the noise levels corrected to a reference speed using the speed coefficient. Lastly, another change included in the draft standard is a more strict description of vehicles for the HD category with a newly imposed lower limit on Gross Vehicle Mass (GVM) of 8 tons. This aims to increase the compatibility of HD and HM categories for their combined analysis by reducing the variation among the vehicles, and thus the sound

emission profiles, within this category. However, this lower limit also excludes more vehicles from the measurement. In this study, a large dataset of SPB measurements was collected according to the ISO/DIS 11819-1 draft standard and utilized to evaluate the proposed changes compared to the original standard ISO 11819-1:1997. Furthermore, the possibility of also considering the pass-bys of vans, currently excluded from the data analysis, is explored.

2 Materials and methods

2.1 Statistical Pass-By (SPB)

The influence of the pavement on traffic noise is evaluated through wayside noise measurements according to ISO/DIS 11819-1. The overall vehicle noise is measured, which is more representative of the real-world impact of traffic noise compared to tyre/road noise captured in the CPX-method. To measure the noise levels, a sonometer setup is placed at 7.5 m from the centre of the measured traffic lane at a height of 1.2 or 4.0 m, or both. In this study, only a single microphone at 1.2 m was used, connected to a sonometer type 2260 by Brüel and Kjaer, see Figure 1. Since the pass-by noise is greatly impacted by vehicle speed, it is captured by a tachometer also placed at the roadside; a model KR-10 SP was used in this study. Moreover, the speed of the pass-bys needs to be checked actively as only vehicles approaching at a constant speed should be measured; accelerating/decelerating causes additional engine or breaking noise, which is not desired when assessing the pavements acoustic performance. The sonometer and tachometer are connected to a data acquisition system plugged into a laptop, where the real-time noise level and speed can be evaluated, and the pass-by data can be saved. In order to capture the acoustic spectra in third-octave bands, a second sonometer, an NTi XL2, was set up in parallel for the second half of the measurement campaign. The two devices were individually calibrated and their accuracy was checked through a round-robin test at each measurement site. Lastly, also temperature impacts the noise levels. In more detail, higher temperatures result in lower noise levels, regarding passenger cars, this effect is as large as 1dB per 10 °C [8]. For this reason, the air temperature was captured every 15 min. with an accuracy of 0.1 °C and the measured L_{Amax} were temperature corrected according to ISO/DTS 13471-2:2021. In a previous work, we established that this temperature correction is appropriate for this dataset [8].

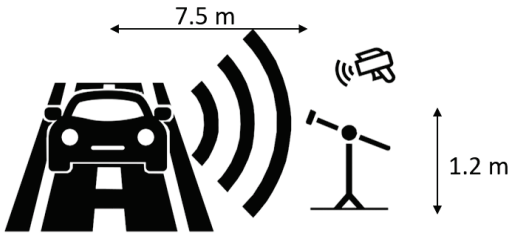


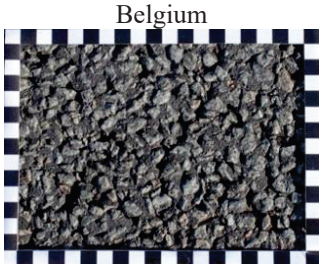
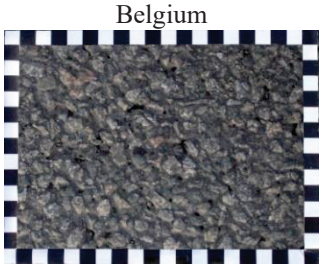
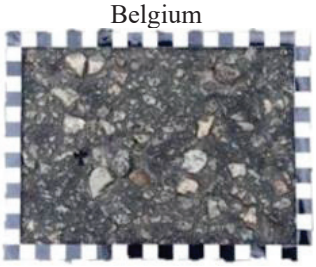
Figure 1: SPB method setup.

2.2 Locations

In the SPB method, the overall vehicle noise is captured at the roadside, therefore, acoustic reflections of surfaces in the vicinity of the microphone can affect the measured noise levels. In this regard, the acoustic behaviour of the ground surface between the measured lane and microphone position is of great importance. For this reason, the standard sets requirements regarding the acoustic properties of this surface area and other objects that disturb the acoustic free-field around the microphone [4], [5]. In this study, SPB measurements were repeatedly performed at three locations, which were carefully selected to fulfil these requirements. However, a few minor deviations were unavoidable. Furthermore, as the goal of this study was to evaluate the changes made in the draft standard, and the majority of these changes are regarding the heavy vehicle categories, the locations for this study were specifically selected for their large share of heavy vehicles in the traffic volume. Due to safety concerns, the maximum speed limit at the selected locations was 70 km/h. However, the average speeds during the measurements were remarkably lower, especially for the heavy vehicle

categories. Lastly, the pavement either consisted of an AC or SMA type. All available data about the measurement sites is presented in Table 1.

Table 1: Overview of measurement locations.

	Location 1			**Location 2**			**Location 3**		
Address	Keetberglaan, Zwijndrecht, Belgium			Krijgsbaan, Zwijndrecht, Belgium			Stuivenbergvaart, Mechelen, Belgium		
	 Scale: 1 square = 1 cm			 Scale: 1 square = 1 cm			 Scale: 1 square = 1 cm		
Asphalt type	Stone Mastic Asphalt (SMA-10)			Stone Mastic Asphalt (SMA-10)			Asphalt Concrete (AC)		
Max. aggregate size	10 mm			10 mm			14 mm		
Mean Profile Depth	1.2 mm			1.0 mm			1.8 mm		
Surface between lane and mic.	Asphalt parking lane			Cobblestone parking lane Concrete bike lane			Concrete block pave bike lane and sidewalk		
Remarks regarding acoustically free field	• Very thin bushes on the opposite side of the road • Small ditch behind the microphone, banking was covered by grass • Highway near, background noise levels were checked			• Vegetation and houses across the street at approximately 20 and 30 m from the microphone position			• Highway near, background noise levels were checked • small ditch behind the microphone position		
Speed limit	70 km/h			70 km/h			70 km/h		
Vehicle category	P	HD	HM	P	HD	HM	P	HD	HM
Avg. speed	63 km/h	55 km/h	53 km/h	51 km/h	46 km/h	45 km/h	51 km/h	47 km/h	46 km/h
N° pass-bys	773	101	1855	898	137	384	1225	57	286
Measurement count	7 days			5 days			6 days		

3 Results and discussion

3.1 Difference in L_{Amax} of vehicle groups HD and HM

To increase the practicality of the SPB measurements, the draft standard proposes to reduce the minimum number of measured vehicles by combining the data of HD and HM categories into a single category named 'Heavy vehicles' (H). However, as the noise-speed behaviour of these categories differs, the data of the HD category is first normalised by adding 2.7 dB to the L_{Amax} value of each pass-by. To evaluate this proposed change, the data of the individual measurement days was combined per location to achieve a robust dataset, from which the difference in L_{Amax} from categories HM and HD was calculated and compared to the proposed 2.7 dB.

First, in Figure 2, the L_{Amax} values of both HD, without the 2.7 dB correction, and HM pass-bys are plotted versus the logarithm of their speed for each location, and a linear regression analysis is perfomed for each vehicle category. From the graphs in Figure 2, the difference in noise behaviour of the HD and HM vehicle categories is directly clear.

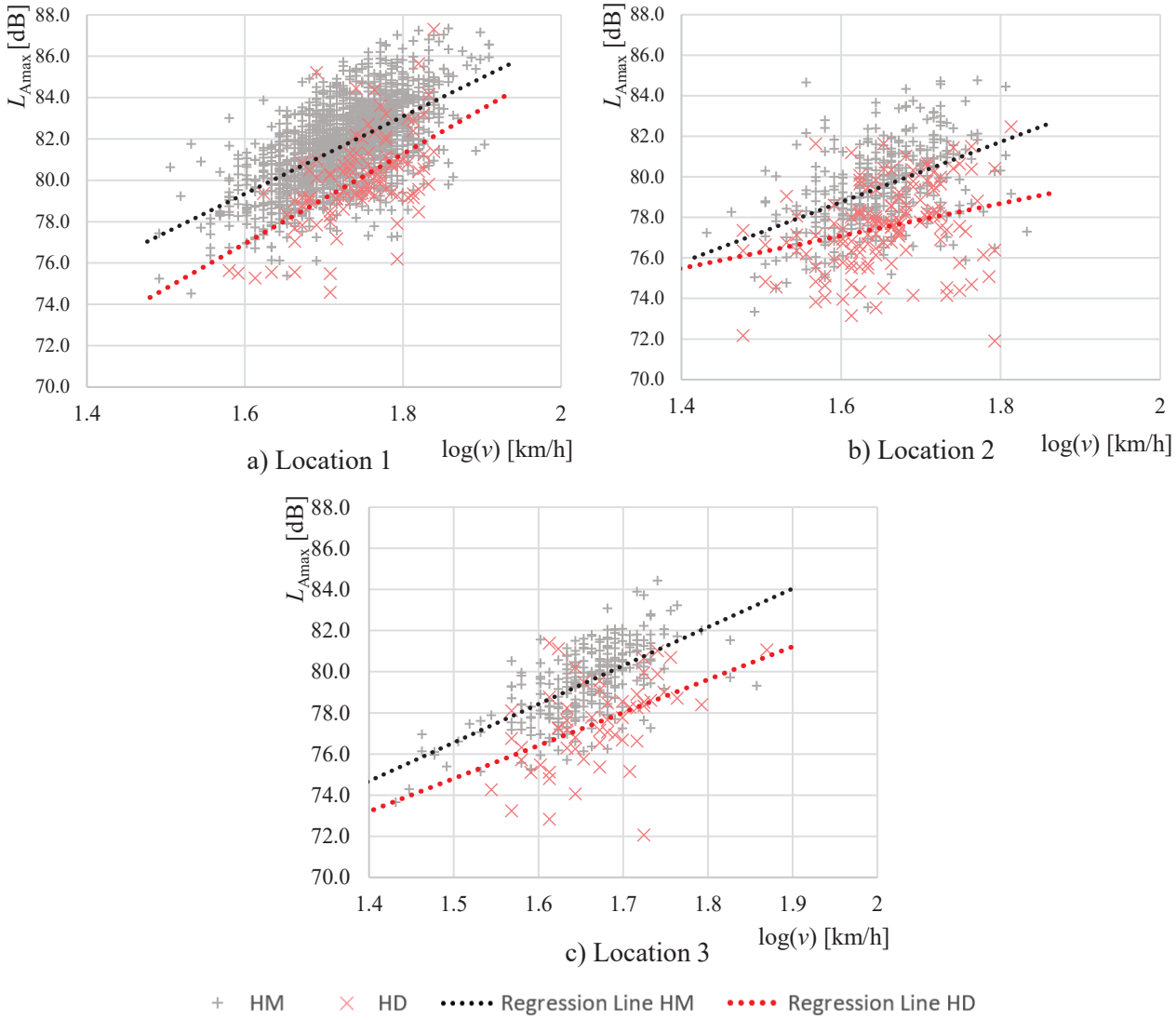


Figure 2: Regression models for HD and HM vehicle categories per location.

Due to the large amount of data of this study, the regression models are assumed to be an accurate representation of the noise behaviour of the vehicle categories. Therefore, in order to evaluate the proposed +2.7 dB in a quantitative manner, the difference between the vehicle categories' C' was calculated as the difference between the regression lines at the reference speed using Eq. 2:

$$C = (A_{HM} - A_{HD}) + (B_{HM} - B_{HD}).\log v_{ref} \quad (2)$$

where $A_{HD/HM}$ and $B_{HD/HM}$ are the regression line's intercept and slope, respectively, on the HD or HM data, and v_{ref} is the reference speed at 50 km/h (log(50) ≈ 1.7). The C values for each location are presented in Table 2. The values for C at 50 km/h range from 2.1 to 2.3 dB with an average value of 2.2 dB.

Table 2: C values per location.

Location 1	Location 2	Location 3	Average
2.1	2.3	2.3	2.2

Figure 3 also displays the C values over the whole speed range of the datasets. The C values are not constant across the speed range as the regression lines of $L_{A,max}$ vs. log(v) are not parallel. There is a decreasing trend for Location 1, meanwhile an increasing trend is found at Locations 2 & 3.

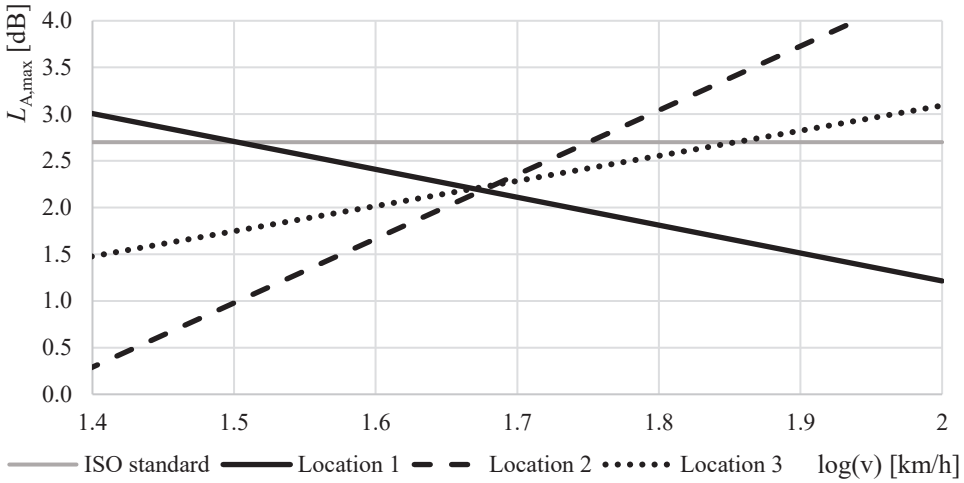


Figure 3: C values per location over the speed range.

The overall average of the C values is 2.2 dB, which is 0.5 dB lower compared to the 2.7 dB proposed in the draft standard. The data used to establish this value has not been published yet. Regarding other literature, [10] reported an average difference of 2.5 dB between the vehicle categories based on a compilation of measurements from different locations. However, there was no detailed information on the pavement type or speed range at the measurement sites. Furthermore, since the report dates from 2015, it did not impose the 8 tonnes lower limit for the HD category imposed by the new draft standard, which affects the selection of vehicles for the HD category. In another study from the UK, HD and HM SPB values were also reported from a series of measurements [11]; the difference between the categories was on average 2.4 dB. However, in that study, the vast majority of the sites were specifically selected because they were paved with low-noise pavements, which acoustic behaviour is very different from the dense asphalt surface considered in this study, and the study was performed at highway speeds. In short, the values reported in the literature were closer to the findings of this study. On the other hand, differences in measurement sites and traffic make the interpretation ambiguous. An overestimation of C by the standard, as the results from this study indicate, would lead to an overestimation of $L_{SPB:H,v}$ at locations where HD vehicles dominate the traffic. However, as the difference is only 0.5 dB, the impact would be limited. Yet, it is still important to further evaluate the correctness of this value, not only for dense asphalt surfaces considered in this study but also for other pavement types. Furthermore, it is necessary to evaluate the speed dependency of the C value in more detail.

3.2 Evaluation of speed coefficients for H

Generally, the H category presents a much larger range of L_{Amax} while simultaneously covering a much smaller speed range. This spread of data results in large uncertainties on the L_{Amax} – log(v) relation in the linear regression model. To solve this problem, an alternative method for processing the data using semi-generic speed coefficients is proposed in the draft standard. There are three coefficients for three surface types: dense asphalt, open asphalt and concrete pavements, and these semi-generic speed coefficients are the equivalent of the slope' B' of the regression line in the L_{Amax} – log(v) regression model. As this study consists of a remarkably large number of HM pass-bys, the dataset allows to perform the regression analysis on the pass-bys of the individual measurement days. The proposed speed coefficient for dense asphalt is then compared to the equivalent slope of the regression lines in this study. With regard to the previous section, where it was found that the average SPL difference between HD and HM was smaller compared to what is proposed in the standard, only the data of HM has been considered in this section.
The B values from this study range from 14.9 to 27.7 dB/log(km/h) with a mean value of 20.5 dB/log(km/h), one outlier at -2.7 was removed, see Table 3. A one sample t-test was performed to check if the mean B value of this study is different from the provided coefficient of 25 in the standard. The result is strongly significant with a p-value <0.001. This indicates that at these specific measurement sites the proposed value is not applicable. Nevertheless, the effect of the difference in speed coefficients on the subsequent calculations is negligible. As indicated in the draft itself, a variation of +/-5 in B results in 0.25 dB over 10 km/h.

Table 3: B and R^2 values from the linear regression on the data of individual measurement days.

Meas. day	B	R^2	Location	Meas. day	B	R^2	Location
1	14.9	0.26	1	10	15.5	0.27	1
2	25.6	0.55	1	11	20.0	0.51	3
3	20.3	0.46	3	12	26.6	0.71	3
4	19.8	0.37	3	13	20.5	0.40	2
5	17.5	0.35	3	14	16.8	0.41	3
6	23.0	0.47	2	15	27.7	0.61	2
7	21.0	0.43	2	16	19.4	0.34	1
8	-2.7*	0.01	2	17	21.2	0.45	1
9	19.0	0.34	1	18	20.0	0.42	1

*Removed from the one sample t-test as an outlier

3.3 Vans as an alternative category

Vans are currently not included in any of the vehicle categories, neither in the original ISO 11819-1:1997, nor the ISO/DIS 11819-1:2021 draft. This is because the purpose of the standard is to evaluate the impact on the acoustic performance of the pavement; as the impact is different for various vehicles, with light and heavy vehicles at the extreme ends, the standard primarily focusses on these two vehicles types. Furthermore, the impact on noise behaviours of cars and trucks is more relevant since the first makes up the majority of the traffic composition, while the latter results in the highest L_{Amax} values. Meanwhile, 'medium' vehicles, such as vans, are not considered to be of interest. Nonetheless, in urban landscapes, the presence of vans is becoming increasingly more significant due to a broad range of activities such as postal, courier and delivery services, construction, etc. Vans are visually different from passenger cars in appearance due to their larger size and different build, and moreover, a van category 'V'could also be easily distinguished from the HD category by vehicle length, height and tyre size. Images of typical vehicles assigned to these categories are presented in Figure 4.

Figure 4: Images of typical vehicles assigned to the P, V and HD categories in this study.

Interestingly, a linear regression on L_{Amax} and the logarithm of speed on the data of each individual measurement day presented on average higher R^2 values for vans compared to the HD and HM vehicle categories, see Table 4. In this application, the R^2 indicates how much variation of L_{Amax} is explained by the variation in speed, thus the higher R^2 values indicate that the regression model for the V category better describes the acoustic behaviour of the pavement compared to the HD and HM categories. While it is true that the linear regression method will no longer be applied to the HD and HM vehicle categories in the ISO/DIS 1119-1:2021 draft, the R^2 values remain a good indicator for the uncertainty when using the proposed alternative method. For this reason, considering vans as an alternative vehicle category may yield valuable information that potentially could aid in reducing the measurement length. The noise behaviour of this alternative vehicle category in relation to the other vehicle categories will be explored in the next section.

Table 4: R^2 values per vehicle category.

	P	HD	HM	V
Mean	0.507	0.298	0.410	0.443
Stdev	0.164	0.245	0.150	0.224
N° pass-bys	2896	295	2525	682

3.4 Comparison of noise behaviour of different vehicle categories

Figure 5 displays the regression lines resulting from the L_{Amax} – log(v) relation for the four different vehicle categories and three measurement locations.

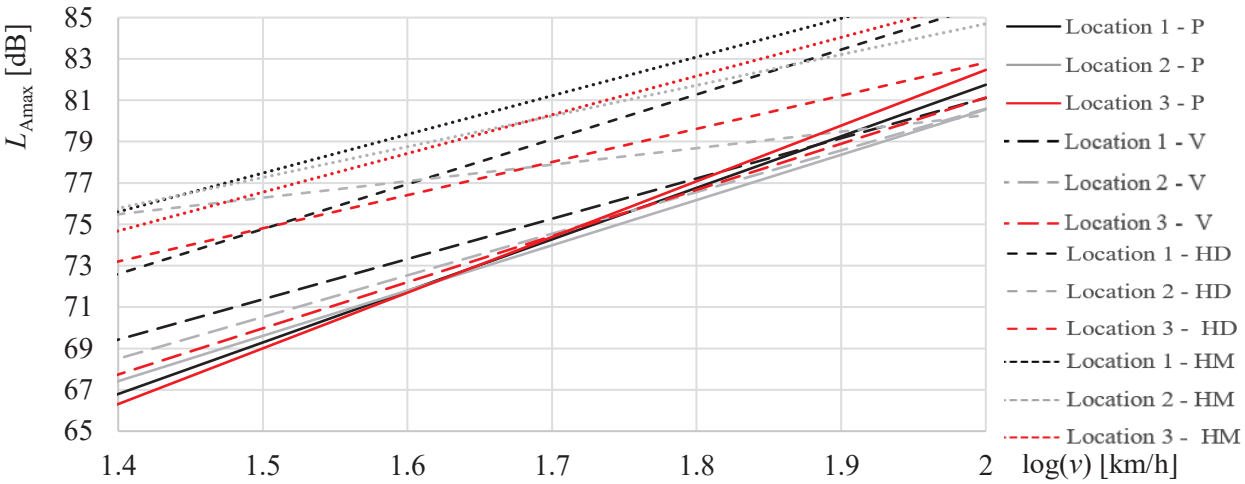


Figure 5: Regression lines for each vehicle class at all locations.

The P category presents the most inclined L_{Amax} vs log(v) relation; thus, for low speeds (below 45 km/h), their noise levels are considerably lower than at higher speeds compared to the other vehicle categories. At low speeds, power-unit noise outweighs tyre/road noise, while at higher speeds, tyre/road noise is dominant in contributing to the overall vehicle noise. In this sense, it can be interpreted that the share of power-unit noise over vehicle noise of P pass-bys, even at low speeds, is consistently smaller than for other vehicle categories. This behaviour is expected due to their smaller engine and other differences in the intake system/exhaust/muffling design.

Even though the regression lines for P pass-bys are slightly steeper than for vans, Figure 5 shows that the two categories produce similar noise levels, mostly at higher speeds. Vans, especially those based on car platforms, may only require passenger rated tyres (tyre class C1, as per the ECE Regulation 117 [12]). Larger commercial vans tend to demand specific van tyres (C2) to cope with the payload of the vehicle, but these tyres have similarities with passenger car tyres, such as tread depth, that may result in a tyre/road noise generation behaviour closer to the C1 than the tyres generally used on heavy vehicles (C3).

The 2.7 dB addition in HD pass-bys noise levels results in a vertical shift of the regression lines in Figure 5. As discussed in Section 3.1, even though this addition will bring the HD pass-bys L_{Amax} to similar noise levels as HM, the L_{Amax} vs log v relation slopes are consistently different, emphasising that a speed-depencendy of this value should be investigated.

To gain more insight in the acoustic behaviour of the different vehicle categories, Figure 6 depicts the average L_{Amax} spectrum from the three locations for the vehicle categories in third-octave bands between 315 Hz and 5000 Hz. To eliminate the speed dependency, and allow a direct comparison among vehicle categories, the noise levels ($L_{Amax,i}$) for each individual pass-by v_i were normalised to a reference speed of 50 km/h for each vehicle category and location, using the slope B of the regression lines presented in Figure5, before calculating the average noise levels, following Eq.(3).

$$L_{Amax,i,speednormalised} = L_{Amax,i} - B * \log\left(\frac{v_i}{v_{ref}}\right) \quad (3)$$

V and P pass-bys present a similar behaviour regarding the value of peak L_{Amax} ranging around the 800-1000 Hz, Additionally, their spectra almost overlap between 1000 and 2500 Hz third-octave bands. This could be attributed to the fact that air-pumping mechanisms [6], induced by displacement of air in/out the tyre tread under its volume changes upon contact with the pavement, are the main tyre/road noise generation mechanism at these frequencies and are similar for vans and passenger car tyres due to similarities in construction.

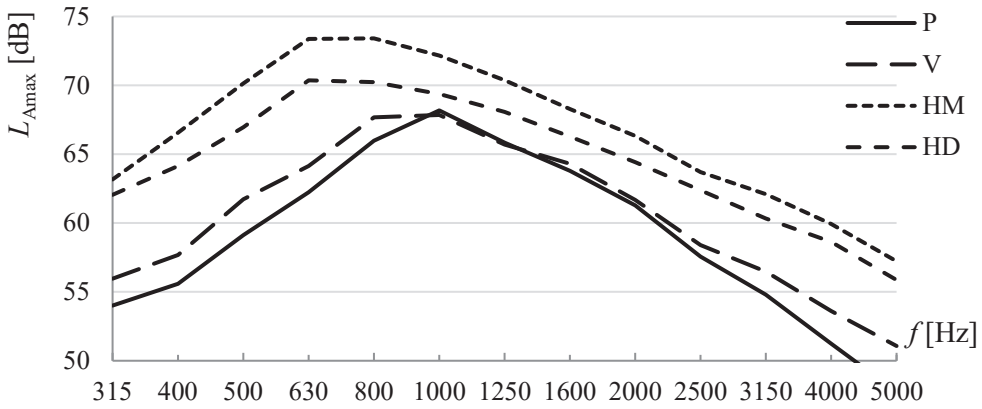


Figure 6: Average spectrum of LAmax, corrected for speed, per vehicle category P (n=616), V(n=166), HD (n=74), and HM (n=905).

HD and HM present similarly shaped spectra, shifted in amplitude by 2.1±0.7 dB on average. These spectra show that these two vehicle groups are impacted by the different tyre/road noise generating mechanisms in a similar manner and, in this sense, may produce a consistent single group by combining HM pass-bys with the normalised HD, as proposed in the ISO/DIS 11819-1 draft standard.
Peaks in noise levels for HM and HD occur at lower frequencies than P and V. A possible explanation for this can be a lower crossover frequency from the dominance of impact mechanisms to air displacement mechanisms for truck tyres (around 500 Hz), compared to car tyres (800-1200 Hz) [6].

4 Conclusions

In this study, a large dataset collected over 18 measurement days at three locations with a large share of heavy vehicles was used to evaluate the changes in ISO/DIS 11819-1 draft standard compared to the original currently active version ISO 11819-1:1997. The following was concluded:

- This study reports a difference in maximum overall noise levels of vehicle pass-bys 'L_{Amax}' between heavy vehicles with two axles (HD), and more than two axles (HM) of 2.2 dB. This is 0.5 dB lower compared to the 2.7 dB proposed in the draft standard. An overestimation of this difference C by the standard, as the results from this study indicate, would lead to overestimation of $L_{SPB:H,v}$, used to describe the acoustical performance of the pavement, at locations where HD vehicles dominate the traffic, although a difference of 0.5 dB entails only a limited impact.
- The slope of the regression line B for the HM vehicle category of all individual measurement days in this study was on average 20.5 dB/log(km/h). A one-sample t-test showed that this result is significantly different from the value of 25, which is proposed in the draft standard as an equivalent value to replace the regression analysis by an alternative data processing method.
- The spectra of the HD and HM vehicle categories indicate that they are similarly impacted by the different noise generating mechanisms, thus producing a consistent vehicle group after the $L_{A,max}$ normalisation for HD vehicles.
- By presenting good R^2 values in the linear regression analysis, vans appear to be an appealing group as an additional vehicle category with relatively consistent noise behaviour and could thereby reduce the measurement duration. Moreover, a deeper look showed that vans present strong similarities with passenger cars (P): regarding L_{Amax}, the two categories produce similar noise levels, mostly at higher speeds where tyre/road noise is dominant. Furthermore, their noise spectra almost overlap from 1000 to 2500 Hz with only limited differences outside this frequency range. A possible explanation is the similarity of the vehicle's tyres resulting in similar tyre/road noise behaviour.

Acknowledgements

The authors would like to thank the assistance of Anneleen Bergiers (BRRC), for providing relevant data for the drafting of this paper. Furthermore, this work could not have been completed without the technical support from the technicians and researchers from BBRC who provided the equipment and necessary explanations for the SPB measurements. Lastly, we thank the former master's students Dario Regnier and Nathan Muts for their assistance during the SPB measurements.

References

[1] Elise Van Kempen, Maribel Casas, Göran Pershagen, and Maria Foraster. WHO Environmental Noise Guidelines for the European Region: A Systematic Review on Environmental Noise and Cardiovascular and Metabolic Effects: A Summary. *International Journal of Environmental Research and Public Health,* 15(2): 379, 2018. doi.org: 10.3390/ijerph15020379

[2] Mathias Basner and Sarah McGuire. WHO Environmental Noise Guidelines for the European Region: A Systematic Review on Environmental Noise and Effects on Sleep. *International Journal of Environmental Research and Public Health,* 15(3): 519, 2018. doi.org: 10.3390/ijerph15030519

[3] WHO. Environmental Noise Guidelines for the European Region. *WHO*, 2018. Available online: https://www.euro.who.int/en/publications/abstracts/environmental-noise-guidelines-for-the-european-region-2018

[4] ISO 11819-1:1997. Acoustics—Measurement of the Influence of Road Surfaces on Traffic Noise—Part 1: Statistical Pass-By Method. The International Organization for Standardization, 1997.

[5] ISO/DIS 11819-1. Acoustics—Measurement of the Influence of Road Surfaces on Traffic Noise—Part 1: Statistical Pass-By Method. The International Organization for Standardization, 2021.

[6] Ulf Sandberg and Jerzy A. Ejsmont. Tyre/Road Noise Reference Book. *Informex*, 2002; ISBN 9789163126109.

[7] Senling Ling, Fan Yu, Daquan Sun, Guoqiang Sun, and Lei Xu. A comprehensive review of tire-pavement noise: Generation mechanism, measurement methods, and quiet asphalt pavement. *J. Clean. Prod*, 287: 125056, 2021. doi: 10.1016/j.jclepro.2020.125056.

[8] Michiel Geluykens, Ablenya G. de Barros, Luc Goubert, and Cedric Vuye. Empirical Study on Temperature Influence on Noise Measurements with the Statistical Pass-By Method. *Sustainability,* 14(4): 2099, 2022. doi: 10.3390/su14042099.

[9] Luc Goubert and Ulf Sandberg. The balance between practicality and uncertainty in the SPB method: How many heavy vehicles do we need to measure? *Proceedings of InterNoise 2020*, 3982-4994, 23–26 August 2020, Seoul (Korea).

[10] Jørgen Kragh, Fabienne Anfosso-Lèdèe, Wolfram Bartolomaeus, Marek Zöller, Truls Berge, Anneleen Bergiers, Matthew Muirhead, and Reinhard Wehr. Deliverable D2.3: Analysis of methods for measuring noise properties of road pavement. *Rosanne project*, 2015. Available online: https://www.rosanne-project.eu/ajax/DownloadHandlerFM.php/downloadFile?i d=11821

[11] Matthew Muirhead, L. Morris, and R. E. Stait. The performance of quieter surfaces over time. *Transport Research Laboratory report*, 2010.

[12] Regulation No 117 of the Economic Commission for Europe of the United Nations (UNECE)—Uniform provisions concerning the approval of tyres with regard to rolling sound emissions and/or to adhesion on wet surfaces and/or to rolling resistance [2016/1350]. Off. J. Eur. Union 2016, L 218, 1–106.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Relationship between tyre-road noise and temperature under noncontrolled traffic flow conditions

Manuel Sánchez-Fernández [1, 2], David Montes González [1,*], Juan Miguel Barrigón Morillas [1], Pedro Atanasio Moraga [1], Guillermo Rey Gozalo [1], Rosendo Vílchez Gómez [1], Alicia Bachiller León [1]

[1] Universidad de Extremadura, INTERRA, Lambda, Departamento de Física Aplicada, Avda. de la Universidad s/n, 10003 Cáceres, Spain

[2] Universidad de Extremadura, INTERRA, NEXUS, Avda. de la Universidad s/n, 10003 Cáceres, Spain

*davidmg@unex.es

Abstract
The influence of temperature on tire-road noise level is a topic on which many investigations have been carried out on controlled road traffic flow conditions based on international standards. Its practical application under uncontrolled vehicle flow conditions may be of interest for the assessment and management of road traffic noise through strategic noise maps. This paper presents the results of an investigation carried out by means of in situ measurements of road traffic noise levels and environmental temperature and a discussion is conducted with respect to those findings available in the scientific literature. A negative and highly significant dependence relationship of the road traffic noise level with both pavement and air temperature was found on the two days of measurement, with similar explanations for variability in the range 59-75% for both temperatures but with differences of around 12-14% between the two sets of measurements. When the slope coefficient is compared, increases of approximately 40% for pavement temperature and 36% for air temperature are found from day 2 to day 1. While the values of the coefficient of variation of noise level with temperature range from -0.05 to -0.07 dBA/◦C for pavement temperature, they are in the range of -0.14 to -0.19 dBA/◦C for air temperature. These values are similar to those published in the scientific literature for pavement temperature, but higher than those reported for air temperature under controlled traffic flow conditions. The followed methodology could be applicable in different conditions.

Keywords: road traffic noise, temperature correction, strategic noise map, on-site measurements.

1 Introduction

Environmental noise from road infrastructure is a topic widely addressed in the scientific literature due to its impact on the surrounding environment, including both people and wildlife [1,2]. The noise level generated by vehicle traffic depends on different aspects related to road and vehicle characteristics as well as weather conditions.

In connection with the characteristics of this type of infrastructure, Freitas et al. [3] point out the relevance of road surface conservation. Based on a study carried out in Guimarães (Portugal), they found that pavement maintenance in the early stages of problem development is, especially on low-speed roads, very important to reduce the environmental noise. The type of pavement is another relevant point in relation to tyre/road noise. For example, from a study carried out in Brazil, Mendes Knabben et al. [4] indicated that the porous friction layer (PFL) was the surface with the lowest noise index (CPXI) when compared to other pavements. But the characteristics of the vehicles driving on the road are also important in the generation of noise. The role of the kind of tyres on the sound pressure level generated on low-noise road surfaces was showed by Licitra et al.

[5], while other studies have shown the influence of factors such as the speed and type of vehicle on road traffic noise [6].

Concerning the meteorological conditions, different authors conducted investigations following international standards such as ISO 11819-1, ISO 11819-2 and ISO/PAS 11819-4 [7,8,9] that found linear relationships between the temperature and the tyre/road noise with coefficients varying between −0.03 and −0.11 dBA/ºC [10,11,12,13,14,15]. Jaben [16] conducted a study on a road with non-controlled traffic circulation, obtaining values of the temperature correction between -0.03 and -0.12 dBA/ºC for light vehicles depending on the speed (50-140 km/h) and -0.04 dB/ºC for middleweight trucks in the range 70-100 km/h. A coefficient for temperature correction is proposed in the ISO/TS 13471-1 standard [17] for two specific tires under controlled traffic conditions, that ranges between −0.04 and −0.11 dBA/ºC depending on the pavement and the vehicles speed. But other values independent of speed are also proposed in this standard ranging between −0.05 and −0.10 dBA/ºC. CNOSSOS-EU method [18] also suggest a correction to the sound power emitted by road traffic to show its reduction as temperature rises: −0.08 dB/°C for light vehicles (category 1) and −0.04 dB/°C for heavy vehicles (categories 2 and 3). If the tyre/road noise is analysed in frequency bands, CNOSSOS-EU method [18] and the ISO/TS 13471-1 standard [17] point out that the same correction should be applied in all bands. In this line, some investigations reported a higher linear relationship between temperature and tyre/road noise in the ranges of 31.5–630 Hz and 1.6–5 kHz [10,11,13].

Considering that this dependence of tyre/road noise on temperature may have a practical application under uncontrolled vehicle flow conditions in the assessment and management of road traffic noise, a paper is presented where an experimental test of the relationships between temperature and road traffic noise on a main road under free traffic flow conditions.

2 Methods

To conduct the present study, in situ measurements were carried out on the section between Cáceres and Malpartida de Cáceres of the national road N-521 in Extremadura (Spain). As can be seen in Figure 1, this is a two-way road with one lane of traffic for each direction. The pavement is seven years old and can be considered as porous asphalt [17] and NL01 class [18]. More specific details of the pavement composition can be found in the paper by Sanchez et al. [19]. The area around the measurement point is flat and can be considered as acoustically absorbent (Figure 1).

The official data from the Ministry of Transport of the Spanish National Government for vehicle flow on the stretch of road under study over the last ten years indicate that the average daily traffic (ADT) is 7,658 vehicles/day, where 95.1% are light vehicles and 4.9% are heavy vehicles. Estimates made from this data indicate that this section of the road can be considered as a major road [20] and that a flow of 600 vehicles per hour during the day can be predicted.

Figure 1. Location of the measurement site (by Google Maps)

The aim of this study was to analyse the dependence of road traffic noise on temperature under actual traffic conditions, which means considering simultaneously a large variability in terms of vehicle model and tyre type and maintenance condition in the experimental test. For this purpose, taking into account the ADT value, a measurement period of 10 minutes was selected to ensure a vehicle flow of at least 100 vehicles. A class 1 sound level meter/analyser was used during the measurements, whose microphone was placed 15 m from the centre of the road [21] and 1.5 m above the ground [22]. No obstacles were present between the microphone and the road that could generate acoustic shielding effects [23] and no reflective surfaces were in the vicinity that could cause reflections of sound [24,25]. Eighteen 10-minute measurements were carried out in each of the two campaigns, recording the equivalent sound pressure level in broadband ($L_{eq,A}$) and 1/3 octave bands (L_{Xeq}). A verification of the calibration of the sound level meter/analyser was performed before and after of each series of measurements.

Simultaneously, the air and pavement temperature, the relative humidity and the wind speed were recorded for each measurement. In the case of air temperature, it was monitored at a height of 1.5 m above the ground [8,10,11,13] avoiding direct exposure of the thermohygrometer to sunlight, while the pavement temperature was taken with a thermal camera on the nearest lane. The wind speed was zero or negligible during the tests. Vehicle flow and class were also monitored according to the vehicle categories indicated in the CNOSSOS-EU method (European Directive, 2015). A speed camera with a 100 km/h limit was in the vicinity of the monitoring point and some measurements showed that vehicles drive on this stretch at an average speed slightly below the limit (most of them in the range of ± 5 km/h). In this concern, Institute for Vehicle Technology estimated a maximum variation of about 1 dBA for variations in speed of ±5 km/h [6].
A period of 10 minutes was considered to ensure the passage of at least 100 vehicles in each measurement, but it is assumed that the number of vehicles will not be exactly the same in all measurements. A normalisation with respect to a reference flow of 780 vehicles/hour (in situ monitored average) was therefore proposed to deal with this aspect [19]. Given that the sound power emitted by categories of vehicles 2, 3 and 4 [18] can be considered different from that of category 1 (light vehicles) [6], a second normalisation was employed by considering the equivalence between the noise level emitted by category 1 vehicles and the rest of the vehicle categories [19]. The coefficients for vehicle categories 2, 3 and 4 in this normalisation were obtained by Sandberg [21].

Since the total equivalent number of light vehicles (category 1) is again likely not to be the same for each measurement, an additional normalisation was performed to study the relationship between measured noise level and temperature. The equivalent vehicle value V_{eq1} derived from the second normalisation was again normalised considering the 930 category 1 vehicles/hour obtained as the average total equivalent category 1 vehicle flow in the measurements. Then, the category 1 equivalent normalised sound pressure level L_{Neq1} was derived from Equation (1) [19]:

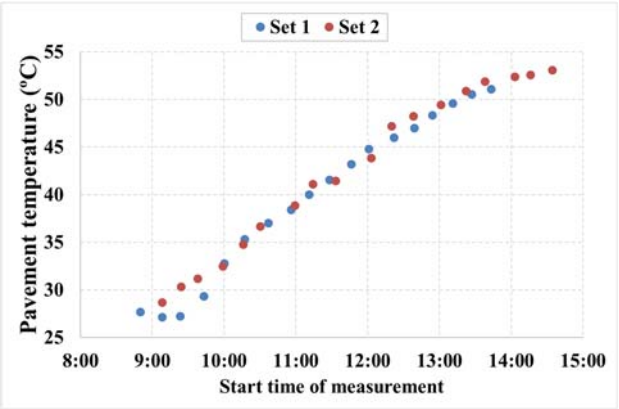
$$L_{Neq1} = L_0 - 10 * log_{10}\left(\frac{V_{eq1}}{930}\right) \quad (1)$$

where L_0 is the recorded sound pressure level; V_{eq1} is the total number of equivalent vehicles in category 1.

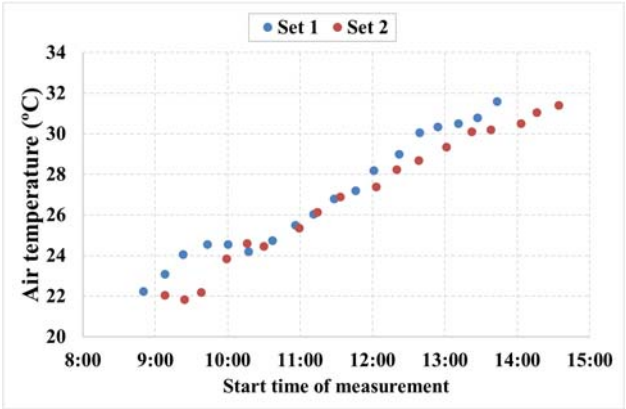
3 Results and discussion

3.1 General conditions

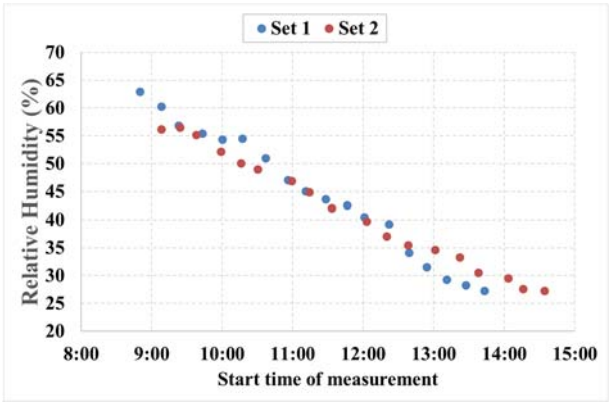
Two sets of 18 on-site measurements were made on consecutive days and at similar hours to try to reduce the variability related to traffic flow and weather conditions. Figure 2 shows the evolution of air (Figure 2a) and pavement temperature (Figure 2b) in relation to the start time of each of the measurements for the two sets. The time-related evolution of the pavement temperature in the two sets is very similar, with some differences in the temperatures recorded in the early morning (between 9:00 and 10:00 a.m.). In the case of air temperature, the differences observed in the early hours are greater than those for pavement temperature. Moreover, from 12.00 onwards, the temperatures recorded in the two surveys also show differences of up to 1ºC. These slight differences between the two sets of measurements in the early morning and from 12:00 onwards are also reflected for relative humidity (Figure 2c). Regarding the percentage of heavy vehicles obtained in each measurement (Figure 2d), it is generally in the range of approximately 2% to 8% and is similar for both sets. No trend is observed in the distribution of the percentage of heavy vehicles in relation to the time of day.



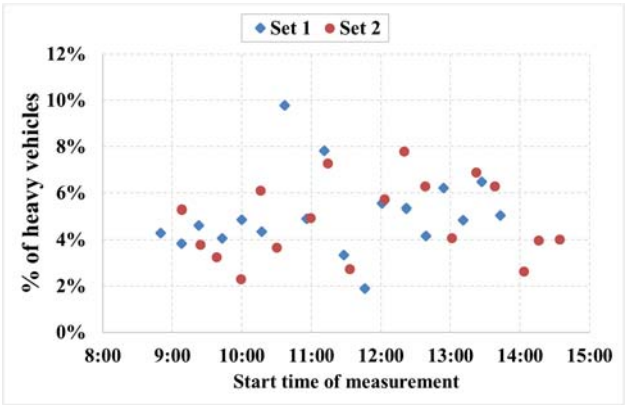
a)



b)



c)



d)

Figure 2. Time evolution during measurement sets 1 and 2 of: a) pavement temperature, b) air temperature, c) relative humidity and d) percentage of heavy vehicles.

More detailed information on the maximum, minimum and average values of the measurement conditions regarding temperature, relative humidity and percentage of heavy vehicles for each set of measurements is shown in Table 1. As can be seen, there are no relevant differences in terms of averages in the measurement conditions of both sets. One aspect that is worth to be pointed out is that, although the average air temperature coincides on both measurement days, the average pavement temperature is approximately 2.6 ºC higher in measurement set 2. The average percentage of heavy vehicles recorded in both measurement campaigns was similar and around 5%, although a maximum value of 9.7% was recorded in set 1, higher than the maximum value in set 2.

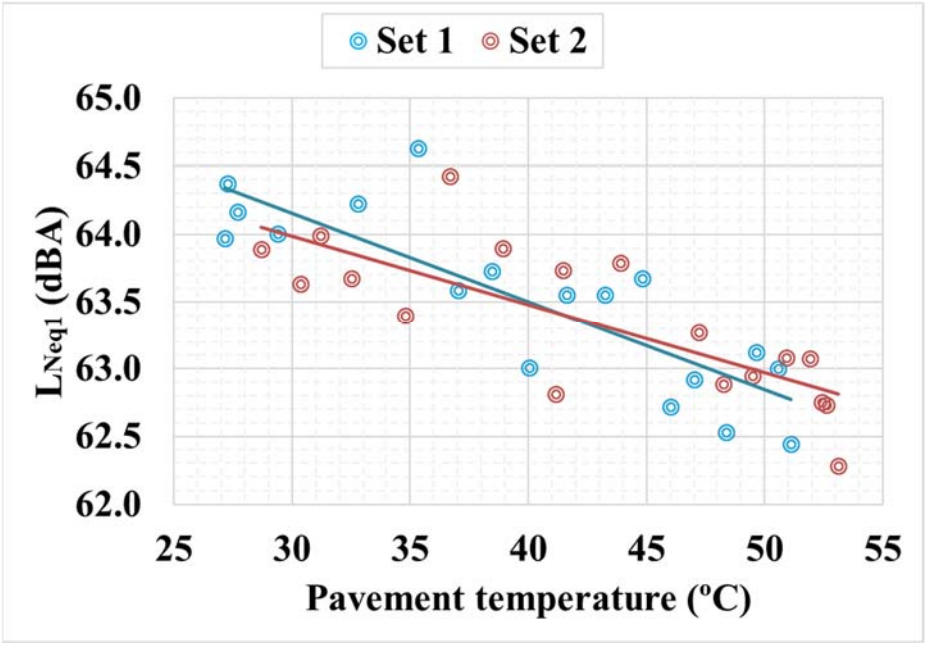
Table 1. Summary of temperature and vehicle data

		Pavement Temperature (ºC)	Air Temperature (ºC)	Relative humidity (%)	% of heavy vehicles
Set 1	Maximum	51.1	31.6	63	9.7
	Minimum	27.2	22.3	27	1.9
	Average	39.9	26.9	45	5.1
Set 2	Maximum	53.1	31.4	57	7.8
	Minimum	28.7	21.9	27	2.3
	Average	42.5	26.9	42	4.8

3.2 Broadband analysis

As mentioned in the methodology section, taking into account the variability in the total number of vehicles and in the percentage of heavy vehicles in each measurement, some normalisations were applied to obtain the category 1 equivalent normalised sound pressure level L_{Neq1}. This allowed to more accurately study the effect of temperature on the sound pressure level from road traffic. Figure 3 shows a diagram of the relationship between the category 1 equivalent normalised sound pressure level L_{Neq1} and the pavement and air temperature for each of the measurements sets 1 and 2, as well as their corresponding linear regression. As can be seen, a negative dependence relationship of the road traffic noise level with both pavement and air temperature is obtained on the two days of measurement, which means that the measured noise level decreases with increasing temperature.

a)



b)

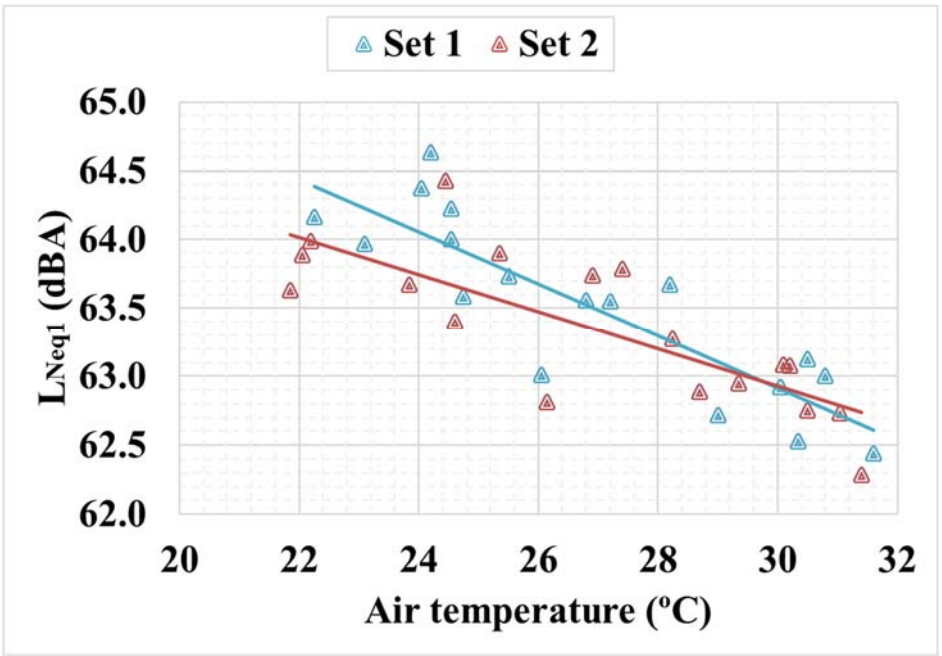


Figure 3. Variation of the category 1 equivalent normalised sound pressure level (L_{Neq1}) with pavement (a) and air temperature (b).

To analyse the results in more detail, Table 2 shows the values of the linear regression parameters as the slope (βi) and its standard error, the coefficient of determination (R^2), the intercept coefficient and the significance, for each set of measurements and for each of the temperatures considered. First, it seems interesting to highlight that all the relationships shown are highly significant (>99.9%), which implies that there is a very low probability that these relationships occur randomly. If the explanation of the variability of the noise level from temperature is studied, it is quite similar for both pavement temperature (59-71%) and air temperature (61-75%). It is also interesting to note an increase in the explanation of variability of 12% for pavement temperature and 14% for air temperature from day 2 to day 1, even though all relationships are highly significant and have a similar intercept coefficient. When comparing the slope of the regression lines, increases of approximately 40% for pavement temperature and 36% for air temperature are found from set 2 to set 1. A larger number of

measurement series would be useful to analyse in depth variations of this nature under similar measurement conditions. It is worth pointing out that while the values of the coefficient of variation of noise level with temperature range from -0.05 to -0.07 dBA/◦C for pavement temperature, they are in the range of -0.14 to -0.19 dBA/◦C for air temperature. This difference between the two temperatures is mainly due to the fact that, for the same range of noise level variation in both cases, the range of variation of the pavement temperature (approximately 24 ºC) is considerably larger than in the case of the air temperature (approximately 9.5 ºC). The findings of these experimental tests for the coefficient of variation of road traffic noise level with pavement temperature are similar to those obtained in previous investigations [10, 11] for tyre/road noise level under controlled traffic flow conditions using a Controlled Pass-By (CPB) method [26] and a Close-Proximity method (CPX) [8] respectively. However, the slope values found in the case of air temperature are higher than those reported so far in the scientific literature under controlled conditions. In this regard, values between approximately -0.08 and -0.11 were determined in previous studies [10,13,27,28] under controlled traffic conditions following procedures such as Statistical Pass-By method (SPB) [7], Close-Proximity method [8] and On-Board Sound Intensity (OBSI) Method [29].

Table 2. Linear regression parameters between the the category 1 equivalent normalised sound pressure level (L_{Neq1}) and the air and pavement temperatures.

	Independent variable	βi (dBA/◦C)	Standard error βi (dBA/◦C)	R^2	Intercept coefficient (dBA)	Sig.
Set 1	T_P	-0.07	0.01	0.71	66.12	< 0.001
Set 2	T_P	-0.05	0.01	0.59	65.50	< 0.001
Set 1	T_A	-0.19	0.03	0.75	68.64	< 0.001
Set 2	T_A	-0.14	0.03	0.61	67.02	< 0.001

4 Conclusions

This paper presents an experimental test of the relationships between temperature and road traffic noise on a main road under free traffic flow conditions. Two sets of 18 in situ measurements were performed at the same point on two successive days to carry out a comparison of the results, considering also the findings reported in the scientific literature which have mainly been performed under controlled conditions. A normalisation was applied to consider the variability in the total number of vehicles and in the percentage of heavy vehicles in each measurement.

A negative and highly significant dependence relationship of the road traffic noise level with both pavement and air temperature was found on the two days of measurement. Similar explanations for variability in the range 59-75% were reported for both temperatures, but with differences of up to 14% between sets of measurements. If the slope coefficient is compared, increases of up to 40% are found from set 2 to set 1. A larger number of measurement series would be useful to analyse in depth variations of this nature under similar measurement conditions.

While the values of the coefficient of variation of noise level with temperature range from -0.05 to -0.07 dBA/◦C for pavement temperature, they are in the range of -0.14 to -0.19 dBA/◦C for air temperature. These values are similar to those published in the scientific literature for pavement temperature, but higher than those reported for air temperature under controlled traffic flow conditions.

Acknowledgements

This project was co-financed by European Regional Development Fund (ERDF) and Junta de Extremadura, Consejería de Economía, Ciencia y Agenda Digital (IB18050 and GR21061).
This work was co-financed by Consejería de Educacion y Empleo of Junta de Extremadura, the Extremadura Public Employment Service (SEXPE) and European Union (FSE) through grants to promote the hiring of research support personnel in the Autonomous Community of Extremadura (PAI20/40).
This work was also supported by Consejería de Economía, Ciencia y Agenda Digital of Junta de Extremadura and European Regional Development Fund (ERDF) through grants for the financing of industrial research and experimental development projects (IDA-19-0022-3).
This work was also supported by Consejería de Economía, Ciencia y Agenda Digital of Junta de Extremadura through grants for attracting and returning research talent to R&D&I centres belonging to the Extremadura Science, Technology and Innovation System (TA18019), where University of Extremadura was the beneficiary entity.

Unión Europea
Fondo Europeo de Desarrollo Regional
Una manera de hacer Europa
JUNTA DE EXTREMADURA
Consejería de Economía, Ciencia y Agenda Digital

References

[1] Khomenko, S., Cirach, M., Barrera-Gómez, J., Pereira-Barboza, E., Iungman, T., Mueller, N., Foraster, M., Tonne, C., Thondoo, M., Jephcote, C., Gulliver, J., Woodcock, J., Nieuwenhuijsen, M. Impact of road traffic noise on annoyance and preventable mortality in European cities: A health impact assessment. Environment International 162, 107160, 2022. https://doi.org/10.1016/j.envint.2022.107160.

[2] Ortiz-Urbina, E., Diaz-Balteiro, L., Iglesias-Merchan, C. Influence of anthropogenic noise for predicting cinereous vulture nest distribution. Sustainability (Switzerland) 12(2), 503, 2020. https://doi.org/10.3390/su12020503.

[3] Freitas, E., Silva, L., Vuye, C. The influence of pavement degradation on population exposure to road traffic noise. Coatings 9(5),298, 2019. https://doi.org/10.3390/coatings9050298.

[4] Knabben, R.M., Trichês, G., Vergara, E.F., Gerges, S.N.Y., Keulen, W. van, 2019. Characterization of tire-road noise from Brazilian roads using the CPX trailer method. Applied Acoustics 151, 206–214. https://doi.org/10.1016/j.apacoust.2019.03.013.

[5] Licitra, G., Teti, L., Cerchiai, M., Bianco, F., 2017. The influence of tyres on the use of the CPX method for evaluating the effectiveness of a noise mitigation action based on low-noise road surfaces. Transp. Res. Part D Transp. Environ. 55, 217–226. https://doi.org/https://doi.org/10.1016/j.trd.2017.07.002.

[6] Institute for Vehicle Technology. Investigations on Noise Emission of Motor Vehicles in Road Traffic. Wuerselen, Germany, 2005.

[7] ISO 11819-1, 1997. Acoustics — Measurement of the influence of road surfaces on traffic noise — Part 1: Statistical Pass-By method.

[8] ISO 11819-2, 2017. Acoustics — Measurement of the influence of road surfaces on traffic noise — Part 2: The close-proximity method.

[9] ISO/PAS 11819-4, 2013. Acoustics — Method for measuring the influence of road surfaces on traffic noise — Part 4: SPB method using backing board.

[10] Anfosso-LédéE, F., Pichaud, Y. Temperature effect on tyre-road noise. Appl. Acoust. 68, 1–16, 2007. https://doi.org/10.1016/j.apacoust.2006.06.001.

[11] Bueno, M., Luong, J., Viñuela, U., Terán, F., Paje, S.E. Pavement temperature influence on close proximity tire/road noise. Appl. Acoust. 72, 829–835, 2011. https://doi.org/10.1016/j.apacoust.2011.05.005.

[12] Bühlmann, E., Sandberg, U., Mioduszewski, P. Speed dependency of temperature effects on road traffic noise. International Congress and Exposition on Noise Control Engineering (Inter-Noise), 2015.

[13] Bühlmann, E., Ziegler, T., 2011. Temperature effects on tyre/road noise measurements. International Congress and Exposition on Noise Control Engineering (Inter-Noise), pp. 557–564, 2011.

[14] Sandberg, U. Standardized corrections for temperature influence on tire/road noise. International Congress and Exposition on Noise Control Engineering (Inter-Noise), 2015.

[15] Kneib, G., Belcher, D., Beckenbauer, T., Beyeler, H.P. Continuous road traffic noise monitoring and aging of asphalt surfaces. International Congress and Exposition on Noise Control Engineering (Inter-Noise), pp. 7122–7133, 2016.

[16] Jabben, J. Temperature effects on road traffic noise measurements, J. Basic Appl. Phys., 2, 43-46, 2013.

[17] ISO/TS 13471-1, 2017. Acoustics — Temperature influence on tyre/road noise measurement — Part 1: Correction for temperature when testing with the CPX method.

[18] European Directive, 2015. Commission Directive (EU) 2015/996 of 19 May 2015 establishing common noise assessment methods according to Directive 2002/49/EC. European Union.

[19] Sánchez-Fernández, M., Morillas, J.M.B., González, D.M., Gozalo, G.R., 2021. Relationship between temperature and road traffic noise under actual conditions of continuous vehicle flow. Transportation Research Part D: Transport and Environment 100, 103056. https://doi.org/10.1016/j.trd.2021.103056.

[20] European Directive, 2002. Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002 Relating to the Assessment and Management of Environmental Noise (END).

[21] Sandberg, U. Vehicle categories for description of noise source. HARMONOISE, 20003.

[22] Montes González, D., Barrigón Morillas, J.M., Rey Gozalo, G., 2015. The influence of microphone location on the results of urban noise measurements. Applied Acoustics 90, 64–73. https://doi.org/10.1016/j.apacoust.2014.11.001.

[23] Montes González, D., Barrigón Morillas, J.M., Rey Gozalo, G., Godinho, L. Evaluation of exposure to road traffic noise: Effects of microphone height and urban configuration. Environ. Res. 191, 110055, 2020. https://doi.org/10.1016/j.envres.2020.110055.

[24] Zagubień, A., Wolniewicz, K., 2021. Impact of measuring microphone location on the result of environmental noise assessment. Applied Acoustics 172, 107662. https://doi.org/10.1016/j.apacoust.2020.107662.

[25] Montes González, D., Barrigón Morillas, J.M., Rey Gozalo, G., Atanasio Moraga, P. Microphone position and noise exposure assessment of building façades. Appl. Acoust. 160, 107157, 2020. https://doi.org/10.1016/j.apacoust.2019.107157.

[26] AFNOR S 31-119, 1993 – Acoustique: Caractérisation in situ des qualités acoustiques des revêtements de chaussées – Mesurages acoustiques au passage.

[27] Liao, G., Heitzman, M., West, R., Wang, S., Ai, C. Temperature Effects on the Correlations between Tire-Pavement Noises and Pavement Surface Characteristics. International Symposium on Frontiers of Road and Airport Engineering (IFRAE), pp. 219–232, 2015. DOI: 10.1061/9780784414255.021.

[28] Yuan, M., Ni, D., Zhang, D., Wei, X., Wang, Z., Wang, C. Effect of temperature on road pass-by noise of light vehicle. IOP Conference Series: Earth and Environmental Science. IOP Publishing, p. 22003, 2019.

[29] AASHTO. Measurement of Tire/Pavement Noise Using the On-Board Sound Intensity (OBSI) Method. AASHTO TP 76-08, 2008.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Solving structure borne noise problems from car wash facility

Jorge Torres[1*], Anders Buen[2]

[1,2] Brekke Strand Akustikk AS, Oslo, Norway.

*jto@brekkestrand.no, anb@brekkestrand.no

Abstract

A car wash facility is a complex system with many noise sources. When the car wash is on the ground floor of a residential building, there are frequently reported noise complaints. The number of complaints usually increases when the car wash service is open 24 hours, and it is constantly in use during the day.

This paper shows how to predict propagation of structure borne noise and airborne noise inside a residential building. It is described different measures, not only to solve structure borne noise problems from a car wash facility, but also to reduce airborne noise. It is presented a case study were structure borne noise was the main noise source. In Norway, noise limits for car wash facility installed in the same building should not exceed a sound level of 25 dB inside the residential apartment ($L_{p,A,T} \leq 25$ dB). This limit is quite difficult to achieve in some situations.

Keywords: car wash facility, structure borne noise, air borne noise, building acoustic.

1 Introduction

A car wash is a facility used to clean the exterior of vehicles. Car washes can be self-service (where the vehicle's owner self-wash his car), full-service (with employees who wash the vehicle), and fully automated (usually connected to a gas station).

Car washes facilities are usually questioned not only for discharge of "dangerous" substances, but also for noise pollution [1]. Noise emissions associated with the car wash would not be a problem when the car washes are situated outside the populated residential areas (for instance adjacent to motorways). However, car washes are often situated in residential areas. Of course, depending on where the car wash is situated, it won't be a problem. For instance, cities that already have a great deal of noise generated by traffic in the area. On the other hand, suburban areas, could have zoning regulations regarding noise pollution. As such, a new car wash must be planned with noise reduction before their permits are approved. For example, Banayan [2] and Kierzkowski and Law [3] presents a case study where a car wash does not comply with city regulations (noise limits during daytime on outdoors residentials areas). The authors investigated several features to reduce noise from the proposed car wash.

This work deals with the noise problems generated from an automated car wash situated on the ground floor in a residential building. The car wash is located on an urban area where any noise coming from the car wash was not problematic since the street noise is already over the noise regulations for car wash facilities in Norway. However, due to car wash facility is in the building, it was found several noise sources associated with their operation. The noise measurements revealed that noise transmitted to neighbours in the same building have two components: airborne noise and structure borne noise.

2 Noise levels regulations in Norway

In Norway the noise regulations are defined according to the national standard NS 8175, *acoustic conditions in buildings* [4]. The guideline indicates limit values for indoors and outdoors that are sufficient to meet the requirements in the building technical regulations. Current noise limit values are shown in Table 1.

Table 1: Sound levels limit, indoors and outdoors, from technical installations in the same building. Excerpt from NS 8175, sound rating C.

Type of user area	Sound limit in NS 8175, sound rating C
In living rooms and bedrooms from technical installations in business and service activities in the same building	$L_{p,A,T} \leq 25$ dB $L_{p,AF,max} \leq 27$ dB
Sound level in outdoor living area and outside window from technical installations in the same building and in another building	Night, 23-07, $L_{p,AF,max} \leq 35$ dB Evening, 19-23, $L_{p,AF,max} \leq 40$ dB Day, 07-19, $L_{p,AF,max} \leq 45$ dB

3 Case study

In one urban area of Oslo city, it is installed an automated car wash in a residential building. The existing site currently has a gas station and convenience store on premises. The gas station and convenience store are opened 24 hours a day, but the car wash is to operate from 07:00 a.m. and 11:00 p.m.

Figure 1 shows a noise map from the area. During the day was not complaints from adjacent neighbours in the area, due to any noise coming from the car wash facility was masked by street noise. However, neighbours on the plan 1 complaints from noise that come from car wash facility. Initial measurements showed noise levels of $L_{p,A,T}$ = 30-36 dB, in the apartment on plan 1. During measurements, it was possible to hear/follow the washing process. Noise levels were measured for three different types of cars, and results variated 2-3 dB.

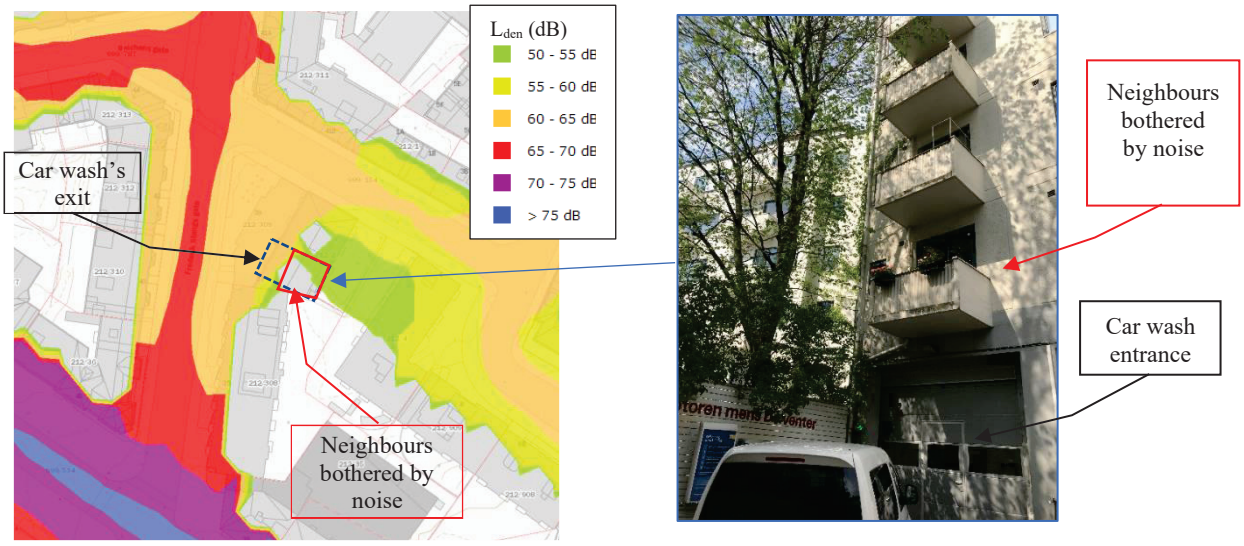


Figure 1 - Left: noise map (excerpt from https://miljoatlas.miljodirektoratet.no/). Right: picture entrance of car wash.

3.1 Construction detail

Walls and floor in the building are made of concrete and brick (thickness is unknown). The car wash is encased with sandwich element of Promisol 100 mm. The sandwich elements are embedded in the foundation plate. The apartment on the plan 2 is a small apartment. The bedroom and living room are located over the technical room which belong to car wash, see Figure 2. There are two high pressure pumps in the technical room. Those are used to remove chemical with high pressure water. The pumps run at a speed of 2930 rpm (48.8 Hz).

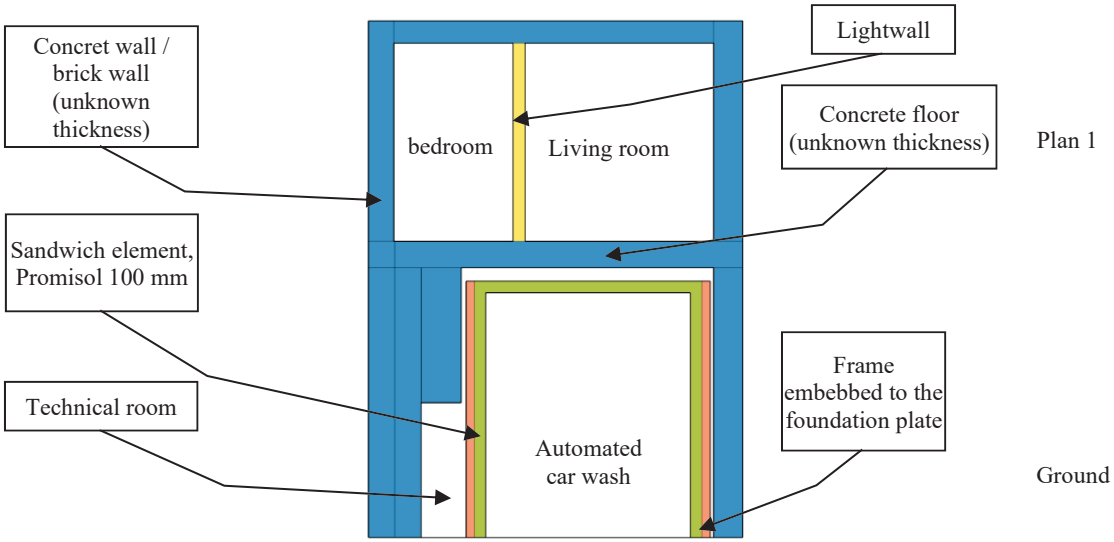


Figure 2: Sectional drawing of the car wash and the apartment which is bothered by noise.

4 Sound and vibration measurements

To map transmission paths, it was performed following measurements:

- Noise and vibration measurements in the car wash, while a car is being washed.
- Noise and vibration measurements in the apartment on plan 1, while washing a car.
- Vertical airborne sound insulation between the car wash and the apartment on plan 1.
- Horizontal airborne sound insulation of the door to the car wash.
- Horizontal airborne sound insulation of building facade.
- Noise level on the facade while washing a car.

Measurements were performed with sound level meter Nor140, with microphone Nor1225 and accelerometer Nor1270.

4.1 Noise and vibration levels

The car wash has different washing programs. Noise measurements were performed for washing program XL which has extra prewash and extra polishing. This washing program has the following processes:

1) Degreasing.
2) Wheel washing + polishing.
3) High pressure flushing.
4) Extra flushing (hood, roof, and boot, without high pressure).
5) Washing with brushes (three brushes, 2 on the side door / frame, and one on the middle) + wheel wash.
6) New round with 3 brushes + extra polishing at the same time.
7) Drying.

Noise measurements were performed in accordance with ISO 16032: 2004 [5]. Simultaneous measurements of vibration levels were also performed on the wall in the apartment and technical room / car wash.

Figures below show results from noise and vibration measurements in the car wash and in the apartment before mitigation measures. A reference line at 25 dB and 80 dB is shown in orange in all figures. Figure 3 shows noise measurements in the car wash and the apartment. The noise level in the apartment (blue curve) increases from $L_{p,A,T}$ = 22-23 dB (background noise) after the washing process started, and it is approximately constant around 35 dB through alt washing process. Lower values were registered in the last part of the washing (drying of the car), but the noise level is also above the requirement.

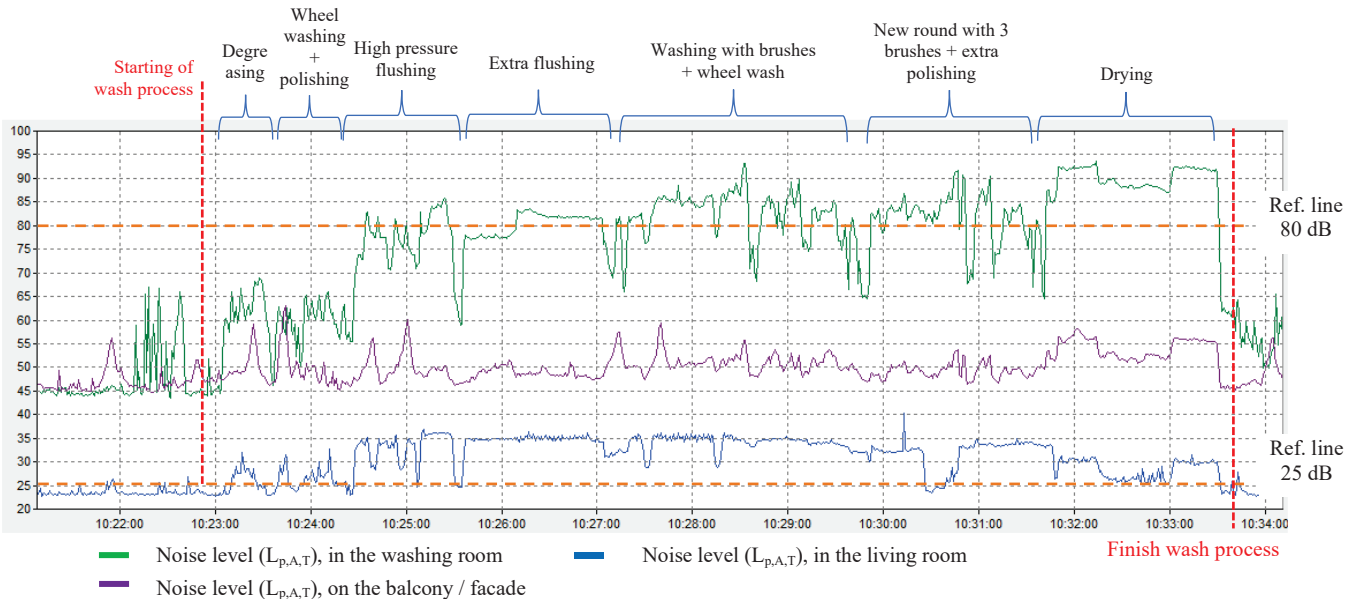


Figure 3: Noise measurements in the car wash and in the apartment, living room.

Figure 4 shows vibration and noise measurements in the apartment while a car is being washed. The vibration level in the apartment (light blue curve) increases when pumps in technical room are activated (brown curve). The measurements show that the noise is mainly transmitted as structure borne noise.

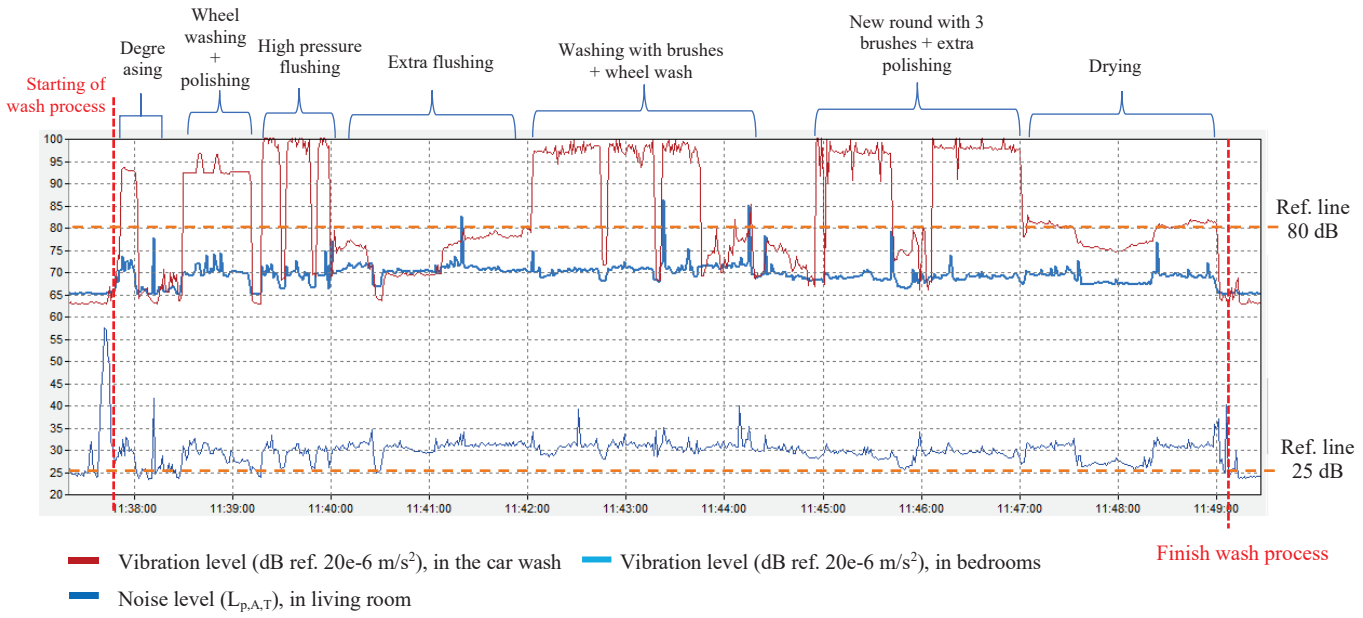


Figure 4: Noise and vibration measurements in the technical room and in the apartment, bedroom.

Figure 5 shows the frequency spectrum from vibration measurements. It was measured on the concrete walls in the technical room (dot red line), in the bedroom (yellow and green line) and in the living room (dash blue line). The results show that fundamental frequency from high pressure pumps (48.8 Hz) is propagated via concrete walls and radiated from the floor and walls into the apartment. The spectrum from noise measurements in the living room shows harmonics components at 100 Hz and 150 Hz when high pressure pumps are operative, it means, the sound energy is structure borne noise transmitted to the apartment.

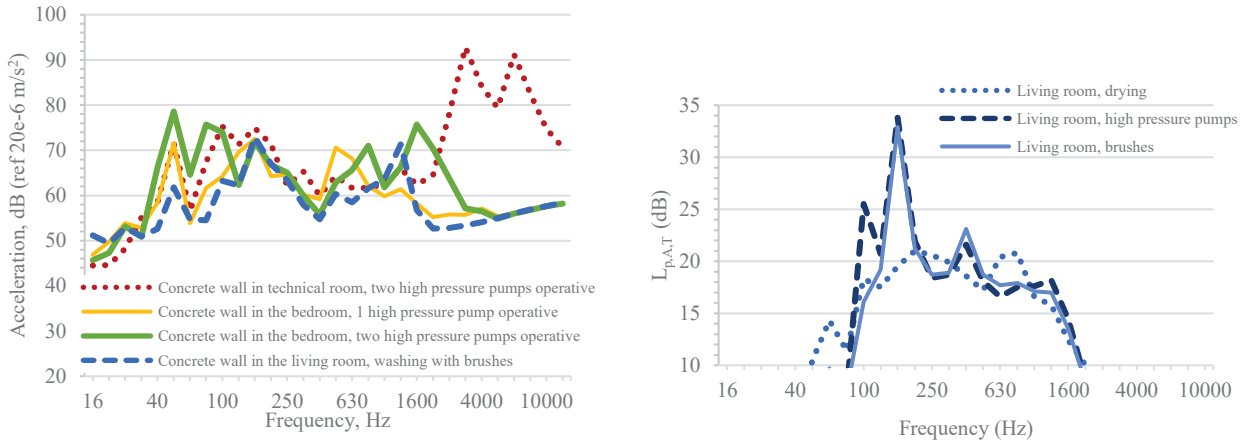


Figure 5: Spectrum from vibration (left) and noise measurements (right).

On the other hand, it was also measured the vertical airborne sound reduction from washing room to the apartment. The result from measurements was $R'_w = 65$ dB and it is not enough to reduce the airborne noise from the washing room. Noise level in the washing room was measured to $L_{p,A,T} = 83 - 93$ dB, see Figure 3. This means that part of noise measured in the apartment is airborne.

5 Recommended mitigation

5.1 Car wash gate

The car wash gate is a rolling gate, model Crawford OH 1042P. This gate is made with steel and aluminium, and it has a thickness of 42 mm filled with CFC polyurethane foam. It was done sound measurements according to standard ISO 16283-1:2014 [6]. Sound measurements show that gate has an airborne sound reduction of $R'_w = 23$ dB. This agrees with producent datasheet. A reduction of 23 dB is not enough to attenuate noise from the car wash facility. The results also show weakness over 400 Hz in the sound reduction. During the inspection, it was observed sound leaks around the gate. It was sealed spaces around the gate with elastic rubber.

5.2 Ventilation facade

The car wash entrance is situated in an alley, and it was built a noise wall around the entrance. After several noise measurements it was discovered that noise level is “trapped” in the alley, and it is transmitted via the facade. Noise level on the facade increases to $L_{P,A,T} = 50 - 57$ dB during drying prossces, see Figure 3.

The facade is built with concrete / brick, and it has windows 6-16-4 (Sound isolation $R_w = 37$ dB). It was measured airborne sound reduction of $R'_w = 30$ dB according to standard ISO 16283-1:2014. The measurements show weakness between 500 – 2000 Hz. It was concluded that some of the noise level measured in the apartment comes via the facade.

In this residential building, the fresh air was supplied to the bedroom with not damping air vents on the façade, it means existing vents did not attenuate the sound. Several tests were performed, and the results show that air vents on the facade contribute to noise measured in the apartment. It was recommended to replace the existing air vents with a sound-absorbing air vents, type TL80DE from Fresh ventilation.

5.3 Technical room

The high-pressure pumps were screwed to the foundation plate with neoprene vibration isolators. This method does not provide sufficient vibration isolation of low-frequency noise. In addition, some water pipe from pumps were fastened with plastic supports and other were fastened with rubber clamps, but it was rigid connected the concrete walls. Vibration transmission from water pipes to the apartment was identified.

High pressure pumps were vibration isolated with polyurethane material, type Sylomer SR220 with thickness 12.5 mm. Regarding pipes, some were fastened with anti-vibration clamps and other were changed with hosepipes and it was disconnected from the building walls.

5.4 Washing room

The car wash has two vertical washing brushes, one horizontal washing brush, and two-wheel brushes on the sides. Figure 6 shows the brushes with yellow arrows. Each brush is connected to a motor that run at 1400 rpm (23.3 Hz).

The space between sandwich element and concrete floor was approximately 20 mm. Such as, it was not possible to improve vertical sound reduction. Sound computation was done with the software for room acoustics, ODEON. As mitigation measure, the ceiling and approximately 50 % of the wall inside washing room were covered with absorber plate, sound absorption class A, type Loke^{TM} strong 40 mm.

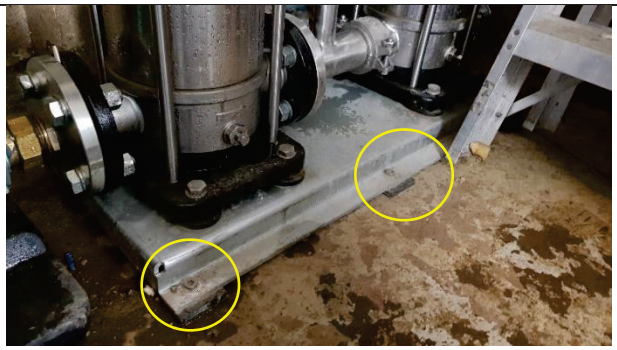
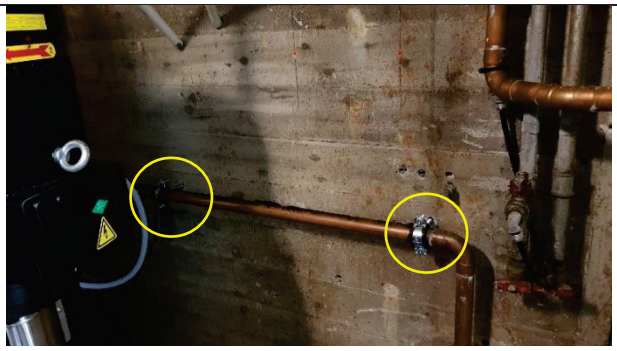
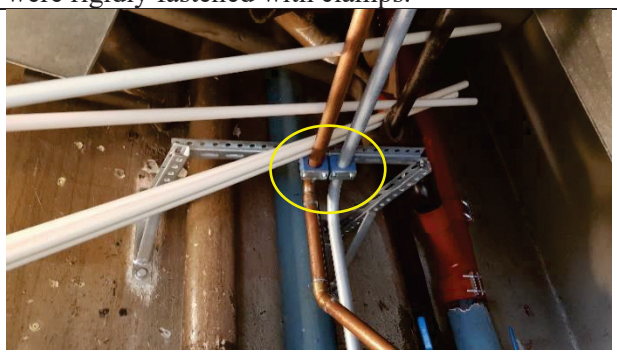
Figure 6: Washing room

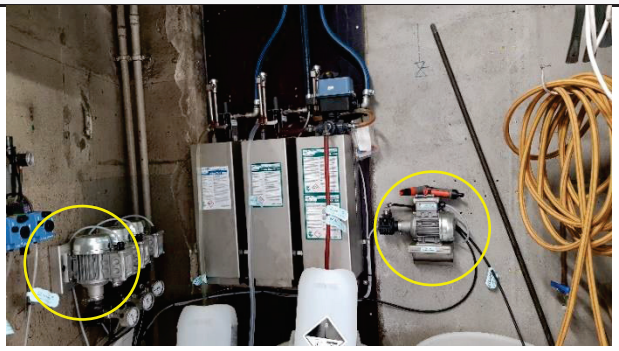
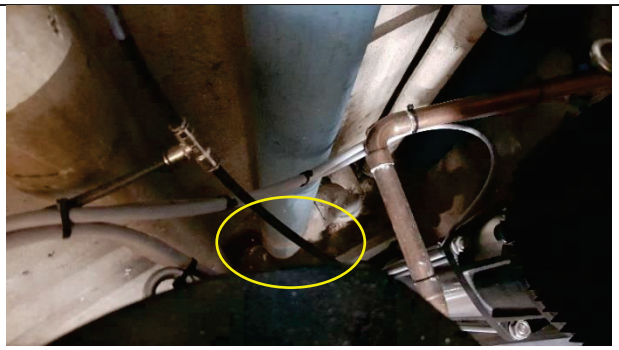
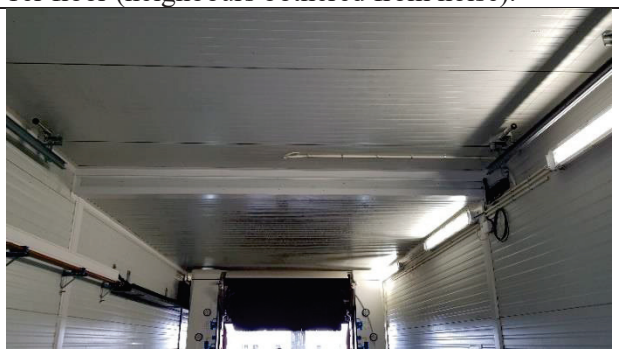
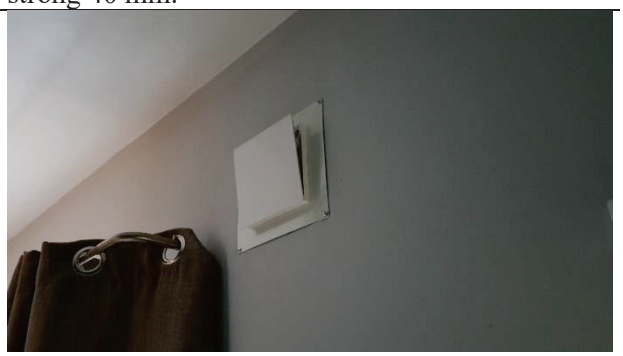
5.5 Exhaust ventilation duct

There was an exhaust ventilation duct located close to the sandwich elements in the technical room. This was used to extract air in the ground before the car wash was installed. It was checked whether the duct was connected to the vents in the bedroom in plan 1. After some studies, it was found that the air extraction in the building was changed to mechanical from the roof. Thus, the ventilation duct was cut and sealed with acoustic panels.

Table 2 shows a resume of mitigation measures.

Table 2: Resume of mitigation measures

#	Situation BEFORE measures	Situation AFTER measures
1		
	The car wash gate had sound leaks around.	Sound leaks were sealed with elastic rubber, but there is still a small leak on the corner.
2		
	High-pressure pumps were screwed to the foundation plate with neoprene vibration isolators.	High pressure pumps were vibration isolated with polyurethane material, type Sylomer SR 220, thickness 12.5 mm.
3		
	Water pipes connected to high-pressure pumps were rigidly fastened with clamps.	The copper tubing was remplaced with hosepipes and it was disconnected from the building walls.
4		
	Water pipes connected to high-pressure pumps were attached with plastic.	The copper tubing was vibration-insulated with rubber.

#	Situation BEFORE measures	Situation AFTER measures
5	 Small pumps for flushing chemicals were not vibration insulated.	 Small pumps for flushing chemicals were vibration-insulated with rubber material.
6	 Several drainpipes from the building were embedded to the technical room and go up to the 1er floor (neighbours bothered from noise).	No measures were taken on drainage pipes due to high-pressure pumps and their pipes were isolated.
7	 The car wash is built with sandwich element of Promisol 100 mm.	 Ceiling and walls in the washing room and technical room were covered with sound absorbing plate, sound absorption class A, type Loke™ strong 40 mm.
8	 Two fresh air vents in the bedroom on the façade were not soundproofed.	 It was changed ventilation air vents with soundproofed vents, type FRESH TL80DE.

#	Situation BEFORE measures	Situation AFTER measures
9	 There was a ventilation duct in the technical room which was connected to exhaust air system in the bedroom (plan 1).	 The ventilation duct was cut and sealed with acoustic panels.
10	There was also a duct for the heat exchanger in the technical room. This was not in use.	 The heat exchanger was cut and sealed with acoustic panels.

6 Results - noise levels

The results from the sound measurements are summarized in Table 3. The table presents the noise level for different washing prosses. The results in the table are corrected for background noise according to measurement standard ISO 16032:2004. The background noise level was measured to $L_{p,A,T}$ = 23-24 dB.

Table 3: Results for measurements in the apartment, plan 1 (corrected for background noise according to ISO 16032).

Washing prosses	Sound limit, $L_{p,A,T}$	Initial noise levels $L_{p,A,T}$	Noise levels, $L_{p,A,T}$, after mitigation measures	Comments
Degreasing	≤ 25 dB	29 dB	25 dB	The noise level for five washing processes satisfies the sound requirement after mitigation measure. The noise from drying and wheel washing is barely audible, and it has a deviation of 2 dB.
Wheel washing + polishing		28 dB	27 dB	
High pressure flushing		30 dB	25 dB	
Extra flushing		31 dB	24 dB	
Washing with brushes + wheel wash		31 dB	24 dB	
New round with 3 brushes + extra polishing		28 dB	25 dB	
Drying		30 dB	27 dB	

Before mitigation measures, the noise level in the washing room was measured to $L_{p,A,T}$ = 83–93 dB, and after measures it was measure to $L_{p,A,T}$ = 80–85 dB.

7 Conclusions

Measurement results show that implemented measures had a good effect. Noise level in apartment on plan 1 was reduced 1-7 dB.

Noise level for almost all washing processes is below the noise requirement now. There are still sound leaks on the corner of the gate, and this may be the reason why the noise level is 2 dB above the requirement.

The noise levels were compared with the relevant Room Criteria curves (RC curves) [7] to detect any interfering components. The measurement results show that the noise from the car wash does not contain disturbing sound components after mitigation measures.

Before measures, noise was measured when washing two cars, a small car, and a van. The noise level for the van was somewhat lower for several washing processes.

Acknowledgements

Acknowledgements to Brekke Strand Akustikk to support in this work.

References

[1] Dorothee Brantz og Avi Sharma. Urban Resilience in a Global Context. *Urbanstudies*, Berlin, 2020.

[2] Chantly Banayan. Operations noise study for a proposed automatic car wash in the Sherma Oaks Distric of the Los Angeles. *Advanced Engineering Acoustics*. Revision 1, 2018.

[3] Marek Kierzkowski and Harvey Law. Car wash noise and EPA regulation. *Acoustic 2017*. Australia.

[4] Committee for Acoustic NS 8175. *Acoustic conditions in buildings, Sound classification of various types of buildings*. Norwegian Standard, 2012.

[5] Technical Committee ISO/TC 43. *Acoustics - Measurement of sound pressure level from service equipment in buildings - Engineering method.* European Standard ISO 16032, September 2004.

[6] Technical Committee ISO/TC 43. *Acoustic- Field measurement of sound insulation in buildings and of building elements, part 1: Airborne sound isolation*. European Standard ISO 16283-1, February 2014.

[7] ANSI S12.2:1995. *Criteria for evaluating room noise.* Acoustical Society of America.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Subjective and objective evaluation of the tone-to-tone timbre variability of historical flute designs.

Michael Haverkamp

Independent Researcher, Köln, Germany.

michael.haverkamp@netcologne.de

Abstract
During the seventeenth and eighteenth century, the variability of timbre of music instruments in terms of a 'tone colour' was essential to enhance the emotional involvement of the audience, the naturalness of expression and the language-likeness of instrumental sounds. An expression of multi-sensory content by 'painting with music' was popular. Timbre is the essential base of the colourfulness of music. Given that flutes of that era show a more simple design than the modern flute, specific fingering patterns are thus needed for the full chromatic scale, which causes a characteristic variation of timbres from tone to tone. For the first time, this study compares the natural timbre variability of recorder and Baroque transverse flute (traverso) with the modern Boehm flute. Possibilities of subjective comparison are discussed. Perceived timbre is seen from a phenomenological point of view. It combines a pitch-dependent portion with effects caused by the distribution of partial tones and formants. For the subjective approach, the pitch dependent portion of timbre perception is removed. This enables subjective evaluation of the influence of partial tones on timbre perception. Objective acoustic analysis of spectra clarifies the associative content with a view to similarities with other instruments. Colour scales are used for a brief overview on the tone-to-tone variability of an instrument. Such colour scales prove to be beneficial to qualitatively document the associative content and its variability in the sense of tone colour. The results indicate that the total sound of each instrument is based on a variety of spectral properties/timbres rather than unique properties for all tones. Surprisingly, this also applies to the modern flute that has been developed for similarity of perceived tone quality throughout the scale. Each flute type shows variable references to sounds of other instruments. Timbre variation is not an error state, but it essentially contributes to the music aesthetics of the Renaissance and Baroque era.

Keywords: flute, recorder, timbre, tone-to-tone variability, colour scale, HIP

1 Introduction

Throughout the last 200 years, improvements of woodwind instruments have been based on acoustic understanding and systematic calculations. These efforts have improved intonation for equally-tempered scale as well as the tonal and dynamic ranges of instruments. Moreover, the aim has been to achieve a substantial uniformity of timbre throughout the whole tone range, as well as improving the possibilities of players' creative influence on the music parameters.
However, in contrast to the development of increasingly powerful instruments, the popularity of historic instruments has grown in approximately the last 50 years. Numerous concerts have been performed, and a nearly endless number of recordings have been published. One element of this trend is the rediscovery of historical ways of performance, phrasing and articulation in terms of 'historically-informed performance' (HIP). Another essential contributor is the reproduction of historical sounds by means of period instruments. One argument for the usage or reconstruction of such historic instruments is the awareness of the aesthetic correlations between the composer's intent, the listener's taste and the tone qualities, which depend on the

instrumental capabilities. In other words, it is understood that the historic timbre and dynamics is an intrinsic part of any contemporary composition. It is neither simply a result of the limited skills of the instrument makers, nor caused by a lack of knowledge on the acoustics of tone generation and the perception of sound.
For this reason, it is meaningful to compare the timbre of historic instruments and appropriate replica with modern types, whereby historical flutes hold particular interest. The flutes family divides into two groups with different excitation principles: the transverse flute and the recorder. Recorders of the Renaissance and Baroque eras show different timbre due to major geometric differences [1]. The design of the transverse flute changed twice during three centuries: from roughly cylindric to conical (tapered towards the foot) and back to a cylindric bore [2]. These design changes caused essential modifications of tone properties [3, 4]. In comparison with the modern Boehm flute, it is supposed that the timbre behaviour of historical transverse flutes and recorders is characterised by a higher variability from tone to tone.
Prior to analysing this feature, it would be beneficial to consider the aesthetic preconditions of the respective eras. Due to space restrictions, this shall be a topic of a separate paper.

2 Timbre and other qualia of instrumental sound

Since the beginning of research on music perception, the question on the nature of timbre has always been a core theme. During recent years, this discussion has gained intense actuality [5]. However, the description of timbre is a complicated task. In contrast to pitch, it is neither a single feature that can precisely be described by a single quantity – such as the frequency in Hz – nor does it find its unambiguous place in a graphical representation, like a specific height of a note within staves. If we imagine a sound event that is clearly distinguished from the acoustic background as an entity, it has a specific Gestalt in our perception [6]. However, due to the qualia problem of perception, we are unable to precisely communicate this ‘sound object’ to other subjects. Timbre is not caught by a single quantity or simple terms, because it is an essential portion of the whole Gestalt. Many quantities as spectral properties relevant for timbre have already been evaluated, such as spectral centroid, attack time, brightness, and others [see e.g. 7]. An evaluation of auditory descriptors of a Baroque transverse flute and various Boehm flutes is provided online by the University of Vienna [8].
A first approach for understanding the relation between physical spectra and perceived timbre of sound has been developed by Hermann von Helmholtz, who stated that the instrumental timbre depends on the number and magnitude of the partial tones [9, p. 113f]. This finding remains a good base to search for differences in the steady-state appearance of tones, especially in case of instruments with a small number of partials and high fundamental. It is also a good starting point for estimating the associative content of sound and similarities between instrumental timbres. In case of flutes – which typically feature a relatively small number of partials – it is beneficial to focus on the approach by Helmholtz by first comparing the steady-state spectra themselves. The proportion of the magnitude of single partials towards each other is then crucial for the perceived timbre, because the onset and release phase of the signal are intentionally ignored. It is self-evident that a focus on the partials of each tone is beneficial in the case when various tones of an instrument show different spectra. Therefore, seeking uniform timbre behaviour throughout the whole tonal range of an instrument would require spectra for all tones with a comparable relation between the main harmonics. Similarity thus means the same distribution of magnitudes of partials relative to the given fundamental frequency.
Beside the relations among the partial tones, the absolute frequency of the fundamental influences the perception of timbre. Pitch is not an attribute of sound that can be heard separately, whereas timbre includes the main perceivable aspects of the auditory Gestalt of a sound. Pitch is simply a descriptor that can be evaluated from an auditory perception. Pitch and timbre are thus not simply orthogonal dimensions of a perceived sound. In order to understand the influence of the partial tones on the perceived timbre itself, it is meaningful to standardise all tones of an instrument to one frequency, whereby the spectral differences are clearly perceivable. However, the timbre of a tone shifted in frequency is different from the real timbre composed from partial tone distribution and real fundamental frequency. At flute instruments with tone ranges equal or above c’/C4 – like those investigated here – pitch is always related to a measurable

fundamental frequency. Perception of a pure residual pitch does not occur because tones with completely suppressed fundamental do not exist.
Investigations of a variety of instruments have shown that formant-like parts of spectra can determine the perception of timbre. The formant-driven concept of instrumental timbre has been developed by Karl Erich Schumann [10]. Schumann's formant theory implies that fixed formants induce the timbre that is typical for a specific instrument and its auditory interactions with other instruments. Formants thus represent the musical timbre of many instruments and of the human voice. The spectral analysis of flutes presented herein, however, has proven that in this case formants do not play a major role for the perception of timbre. It is thus appropriate to focus on Helmholtz's perspective. With view to a pragmatic approach on perception of timbre of musical tones, it is evident that

- Tones of similar pitch change timbre with their spectral content, i.e. the pattern of partial tones.
- Tones of similar spectral content change timbre with pitch, represented by the fundamental frequency.

Even if a pure sinusoidal tone is shifted in frequency, listeners will report that high pitch will sound "somewhat" different than low pitch of the tone. For a demonstration of the tone-to-tone variability of instrumental timbres, it is thus beneficial to separate both influencing factors from each other: the spectral content and pitch. If pitch of a variety of instrumental tones is standardised to a single value, the spectral influence on timbre variability becomes clearly perceivable, although it includes a modification of the real timbre. Figure 1 visualises the effect in case of instruments with uniform spectrum such as a simple electronic keyboard compared to historic and modern flutes. This concept is here applied for subjective and analytic comparison of complete tone scales of flutes.

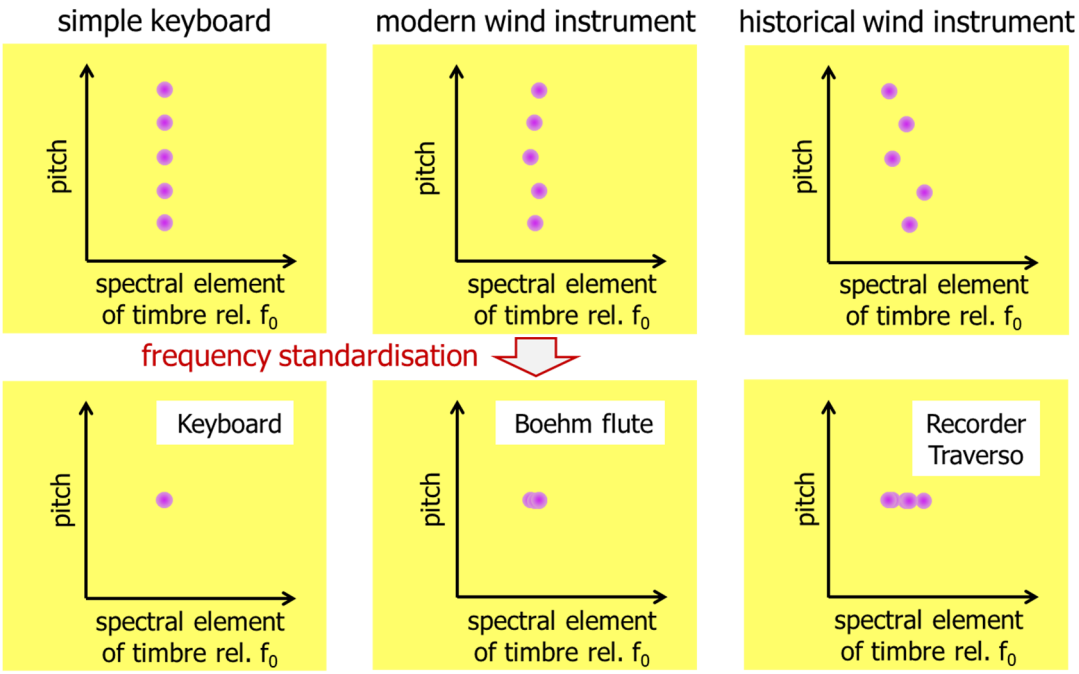


Figure 1: Effect of frequency standardisation for subjective comparison of spectral contributions to timbre.

The subjective method is suitable for demonstrating and comparing timbre variations of woodwind instruments without major influence of formants. The analysis of steady-state spectra includes estimating spectral similarities with other instruments and helps to establish colour scales for visualisation of timbre variability.

3 Instruments used

Both subjective and objective methods have been proven with various instruments. Due to the fact that original builds were not available, replica were analysed which are used for soloistic demands of HIP. Naturally, these designs imply some adaptation on recent requirements of playing techniques, such as double holes. Examples discussed herein are (Figure 2):

- Baroque F-recorder after Pierre Jaillard Bressan (1663-1731), boxwood, a'/A4=415Hz
- Baroque traverso after Carlo Palanca (1688-1783), blackwood, a'/A4=440Hz
- Boehm flute, solid silver, ring keys, solid silver mouthpiece, with golden lip plate, a'/A4=440Hz

Additionally, flute and recorder stops of a simple electronic keyboard (a'/A4=440Hz) have been used as a reference for a complete lack of timbre variability.

Figure 2: Instruments for which results of timbre variability are herein discussed.

4 Subjective comparisons of timbre

Various examples of flute tones can be found online. A comparison of single tones for transverse flutes is documented in [11]. Sound examples and a set of descriptors can be picked here [8]. Both documentations, however, do not include recorder sounds.
In this study, the pitch-dependent part of timbre is separated from the portion that depends on the spectral pattern itself. This approach allows an uncomplicated overview of the overall variability of timbres of various instruments. It is proposed to shift all tones to the same fundamental frequency by means of a time-domain operation such as stretching or squeezing in order to avoid signal distortion. Thus, a *timbre/spectral profile* of all chromatic tones is achieved. In this first approach, standard pitch (415, 440, or 466 Hz) is used as the reference. Sound recording was conducted in front of the player at a distance of 1m, and a height of 1m. The complicated directivity of sound radiation influences the spectral properties of sound and the perceived timbre. Furthermore, directivity differs from tone to tone due to the variation of fingering patterns. However, a definite direction of recording provides a situation similar to the fixed position of the listener of music performances.
It should again be emphasised that the result comprises tones with modified pitch, except the single tone, which already shows the reference pitch. Nonetheless, the timbre profile makes it much easier to gain an idea of the spectral changes within the chromatic scale. Specifically, this applies if the subjective impression is compared with the objective analysis, as described below. It is reasonable that the proposed method may cause some confusion if the instrumental sound is dominated by fixed formants. An additional analysis has clarified that formants do not play a major role for timbre of the given instruments.

Besides timbre changes, the auditory variability of tones played with period instruments extends to further parameters:

- The loudness that can be achieved with one tone differs due to the limited dynamic range of these instruments. This also influences the possibility to perform crescendo and decrescendo.
- The temporal response of the instrument differs with different tones. It thus influences articulation/attack phase, the tone release phase, and the potential of vibrato, which is performed by means of either airstream or finger movement.
- The potential precision of intonation varies from tone to tone.

These influences have been eliminated. Players were instructed to avoid vibrato and loudness changes during playing. A "natural" instrumental tone was required without application of particular playing techniques. Tone cut-out was used to eliminate onset and decay, which were substituted by standard envelope curves to avoid any clicks. All signals have been normalised in magnitude. The COVID-19 pandemic impeded the use of one identical equipment and recording room for all players. Recordings were thus conducted with stereo recording equipment used by the musician, such as Zoom H2 handy recorder or comparable devices. Signal processing in Wave-file format was performed using REAPER (cockos), final processing by WavePad software (NCH). The measurement location in relation to the instrument was standardised as mentioned above. The aim of this project is to demonstrate the principle of timbre variations on period woodwind instruments, which can be achieved without perfect standardisation of measurement equipment and environment. In case a detailed comparison of instruments with small differences is intended, a more precise proceeding would be required. An essential advantage of the proposed timbre profile is given by the possibility to simply arrange test setups for specific questions. Tones can be sorted for a chromatic scale, or other series. Specific selections can be made, such as arrangements of tones achieved with simple fingering compared with those which require cross-fingering. As result of the procedure, a set of standardised signals for the full chromatic scale is available for all instruments.

5 Objective comparison of spectral properties

The adjustment of all tone recordings to the standard pitch of the respective instrument and an equal magnitude enables a simple visual comparison of spectra along the whole scale. Given that the timbre of most tones of the investigated flutes is defined by a relative small number of partial tones, this provides an elementary overview of timbre variations. In a further step, the timbre variation is described by comparison of partial tone patterns of different instruments, like clarinets, string instruments and organ pipes. This provides a first glance at the associative potential of timbre, which refers to sounds of other instruments, and which varies from tone to tone in case of historical instruments. Finally, the spectra are sorted with a view to similarities with other instruments. This provides the possibility to use colour codes for a summary of timbre variations. In figure 3, the pitch range of the Baroque recorder built after Bressan is analysed for two octaves from f′ - f‴ (F4 – F6). Alternative fingering for dis‴ is used, noted as es‴. The strong variation of timbres from tone to tone can be understood with a view to the various types of spectra that occur. In case of this Late Baroque recorder, four principle types can be distinguished. Some tones show a dominant fundamental frequency f1 (Figure 4, top left). The fundamental f1 is present at each tone of the recorder, but usually supported by a variety of harmonics that clearly differentiate the flute timbre from a pure sinusoidal tone. The colour code is used for colour scaling of timbre classes as described below. Some tones show a wide spectrum with seven or more strong partials (Figure 4, bottom left). This behaviour generates a smooth, string-like timbre. As an example, the spectrum of a violin tone is characterised by a high number of harmonics with slightly decreasing magnitude towards the higher partials. String-like timbres of organ pipes are intended to simulate such instruments, like the orchestra-viola [12, p. 252]. Typical examples here are observed at fis'/F#4 and c''/C5. The most typical behaviour of a Baroque recorder with its distinct conical bore is characterised by protruding uneven-numbered harmonics f1, f3, f5 and f7 (Figure 4, top right). While f3 is the second partial that occurs, the behaviour has been named "duodecimal" [13]. This spectral feature is

well known from organ pipes with a closed tube, named ”gedackt” or “gedeckt” (German for “covered”) [14, p. 252]. This is also typical for the lowest register of the clarinet. Helmholtz already knew this behaviour and its effect on the perceived timbre [9, p. 180-1].

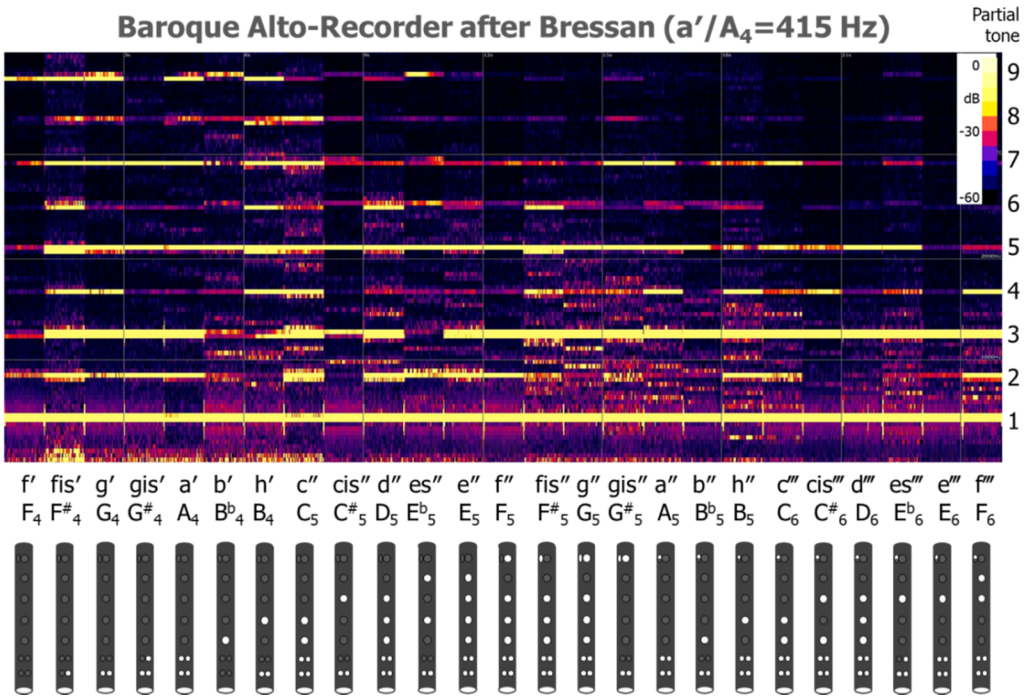


Figure 3: Spectral profile of an F-alto recorder of the Late Baroque period. Spectra normalised and shifted to 415Hz. Hole pattern added to distinguish between simple- and cross-fingering.

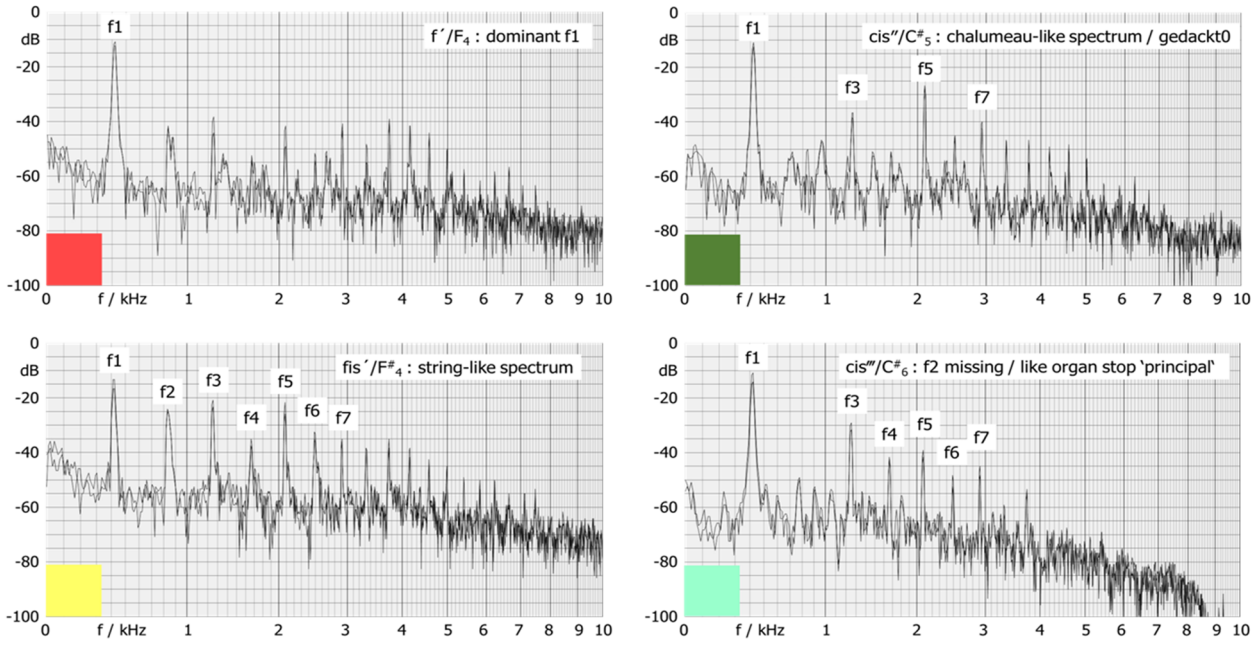


Figure 4: Typical spectra of an F-alto recorder of the Late Baroque period.
The audio signal is normalised and shifted to 415Hz.

This spectral behaviour is herein named “chalumeau-like” with reference to the so-called woodwind instrument that was quite popular during the Baroque era. Its design is similar to the recorder, but makes use of a single reed like the clarinet instead of a windway with labium. The chalumeau-like behaviour typically occurs at many tones of recorders. It is the spectrum most frequently found at the Baroque instrument investigated here. This finding confirms that it is a determining feature of Baroque recorders. At some tones, the spectrum shows a homogenous group of partials, but without a distinct f2 (Figure 4, bottom right). This type is known from the open organ pipes of the principal stop [14, p. 252]. The principal is the most bright register of the organ. It was the only register of the organ used until the Renaissance era. The principal stop manifests the typical sound usually associated with the acoustic organ.

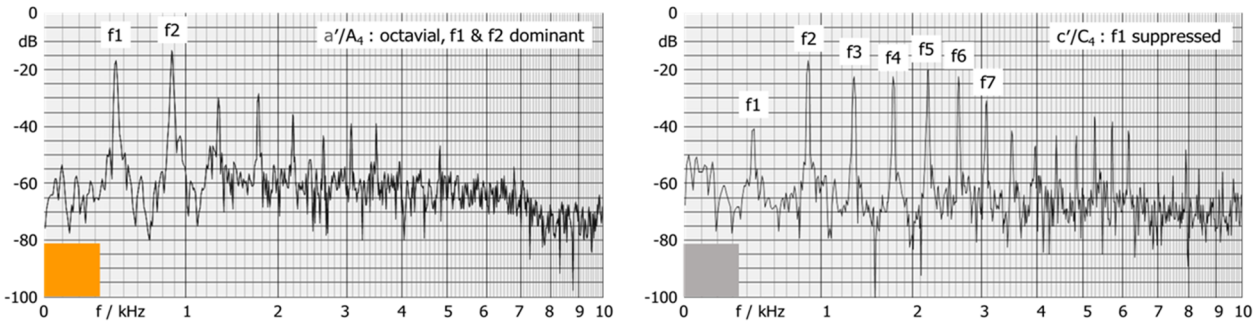


Figure 5: Typical spectra of transverse flutes such as Baroque traverso and Boehm flute.
The audio signal is normalised and shifted to 440Hz.

A special behaviour of transverse flutes is the strong contribution of the second partial tone f2. This is mainly observed in the lower octave. At some tones, f2 can exceed the magnitude of f1 (Figure 5 left). Such “octavial” behaviour is very typical for transverse flutes. It usually does not occur at recorders [13]. By contrast, the chalumeau-like duodecimal behaviour observed at recorders is an exception in spectra of transverse flutes. Nonetheless, it can occur at single tones. In case of the observed Palanca-flute, this was found in the spectrum at g’’/G5 and gis’’/G#5. Pure chalumeau-like spectra have not been identified there. Two exceptions occurred at dis’’/D#5 and d’’/D5 with slightly pronunciation of the first odd partials. Such type of spectrum mixed from string- and chalumeau-like contributions is marked with a light green colour (spectrum not shown here). Furthermore, inharmonic partials occur in some cases, but these tones may have a minor influence on timbre due to their low magnitudes. With respect to the dominant f2 at many tones of the transverse flute, and the overall typical octavial characteristics – in contrast to frequent duodecimal spectra of recorders – it can be concluded that these are spectral features which clearly distinguish transverse flutes from recorders. These spectral characteristics audibly distinguish the steady-state timbre of both types of flutes, although this does not apply to each single tone. Furthermore, it explains why transverse flutes do not need an octave hole/flap, because enhancement of f2 by means of the embouchure allows smooth switching to the upper octave. This opportunity does not exist with the typical chalumeau-like spectra of a recorder. At the Boehm-flutes investigated, string-like and octavial spectra occur most frequently. A broad spectrum of harmonics can be generated at specific tones of the Boehm flutes. Such spectrum is often further characterised by an extenuated f1, as found in the lower register (Figure 5 right).

6 Visualisation of timbre variation as colour scales

In order to provide a brief overview of the timbre variation of each flute, the application of a colour code has been examined for this study. Each aforementioned type of spectrum is characterised by a specific colour. In the past, various trials have been made to attribute colours to tone pitch, with a focus on either the auditory perception or physical properties of sound and light [15]. Many people, however, intuitively connect colours to timbres, as a result of fundamental studies on synaesthetic perception [16, 17]. However, in this study the

colours are intentionally chosen and thus do not refer to a subjective synaesthetic perception. The specific colour refers to the types of physical spectra as described above, by means of the following attributions:

dominant f1 | string-like | chalumeau-like | principal-like
weak f1 | octavial | chalumeau with f2

The most characteristic spectra found at the investigated recorders are shown in the upper line, including the frequent chalumeau-like behaviour. Additional variants found at transverse and Boehm flutes are included in the lower line. As stated above, the octavial type is typical for transverse flutes. The colour scales for the instruments exemplified here are depicted in figure 6. Mixed spectra are indicated by increased brightness of the colour code. As an example, a string-like spectrum with a slight duodecimal contribution is indicated by a light green colour.

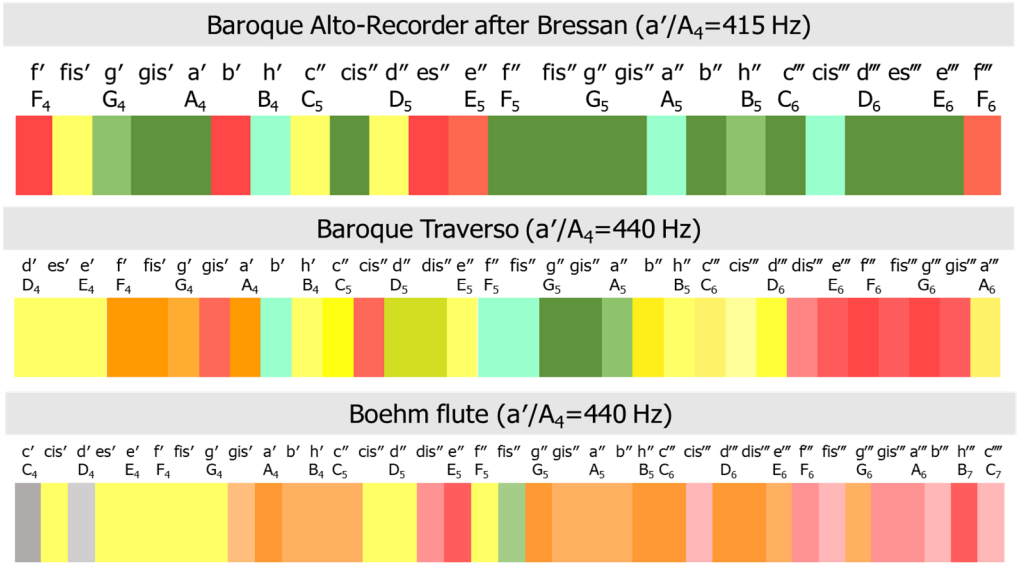


Figure 6: Colour scales related to the timbre profiles of the described Baroque recorder and the traverso compared to the Boehm flute as described above.

The chalumeau-like spectrum occurs at recorders of both Renaissance and Baroque era. Enhanced studies verify that this sound feature is the most distinct of Baroque recorders [18]. Spectra with dominant f1, string-like and principal-like behaviour are also found, but scattered over the whole tone scale. The colour scales thus illustrate the strong variability of timbres from tone to tone, literally the *colourfulness* of flute sounds.
The majority of tones of the investigated traverso shows octavial or string-like properties. Dominance of f1 is typical for the third octave. Chalumeau-like spectra are seldom but evident at g"-a" (G5-A5). At d"-dis" (D5-D#5), such spectra show an additional f2. Some tones have spectra similar to principal organ pipes. Therefore, when compared with the results of recorders, the tone-to-tone variation of timbres is not significantly reduced at the transverse flute. On the other hand, it is not more pronounced. This finding is contrary to Linde's statement that the timbre variability is more pronounced for the Baroque transverse flute rather than a recorder [19, p. 34].
However, the Boehm flute spectra appear to be more equalised but by no means homogeneous between all tones, as aspired by Boehm [20]. String-like and octavial tones as well as those with protruding f1 characterise the timbre profile overall. Few exceptions are the chalumeau-like characteristics, only at fis"/F#5, and spectra with extenuated f1. As mentioned above, a comparison of different instrument designs to the full extend would need a high number of instruments, players, and operating modes such as variation of the embouchure.
This first approach was intended to derive new possibilities for describing the timbre variability of woodwind instruments. Colour scales are beneficial to gain an overall impression of the timbre behaviour of

such instruments throughout the tone scale. They help to compare the spectral variations and the main tendency of all flutes. Furthermore, they refer to comparable timbres of other instruments. In case of the appearance of distinct types of spectra, the attribution to colours is easy, although the attribution of the various mixed types of spectra is a challenge. It is still uncertain where the borderline has to be set for clear differentiations, such as transition of a spectrum with dominance of f1 to a spectrum of further harmonics, or the limits at which a chalumeau-like spectrum morphs to a string-like type. The aim of colour selection was to provide an intuitive overview of the highly complicated pattern of timbre variations of flutes, and explore the possibilities and limits of such approach. It is still a challenge to establish colour scales that are strictly consistent for both the subjective (auditory) *and* the objective (physical/acoustical) appearance of spectral variations. Nonetheless, in combination with the subjective impression of the standardised signals, the colour scales displayed in figure 6 provide a strong impression of the relevance of timbre variations and help to draw conclusions regarding the principal differences in sounds of various instrumental designs.

7 Conclusions

The standardisation of recorded tones regarding similar magnitude and pitch is helpful to gain understanding of the variability of timbres from tone to tone. This is important due to the fact that the associative/iconic content of timbre cannot be quantified. Furthermore, the method is beneficial for subjective evaluations of specific questions, such as concerning the effect of cross-fingerings or features of instrument design. A colour scale is useful to gain a quick visual overview. It is based on the distribution of partial tones, thus providing a reference to comparable instrumental sounds. This colour scale may need some more systematic elaboration on the assignment of specific colours to specific spectra features. It is essential to provide objective criteria for classification of the spectra. A further method development shall include appropriate auditory descriptors, like the spectral centroid. The brightness of the colours used can then be aligned with analysis results of spectral brightness. Continuative investigation need to include more instruments, dynamic variations and various players.

All flutes investigated in this study show a specific variability of timbres throughout the scale. This applies to the early designs from Renaissance and Baroque, but surprisingly also to the modern Boehm flute. It is essential to understand the formation of sound perception of each instrument based on its variety of spectral properties and thus timbres, not just based on a selection of single spectra that are generalised. A simplified description of the timbre of historical flute designs that tends to embrace all tones with few words is not beneficial considering the high variability of timbres. By means of the colour scale, it is possible to perceive the multiplicity of timbres upon one glance. It avoids rash jumping to a generalisation on “how the instrument sounds”. At the investigated flutes with high fundamental frequencies and rather few partial tones, formants with fixed frequencies across the tone scale did not show any significant contribution.

The tone-to-tone variability of sound of the Boehm flute is not as uniform as its acoustic properties might suggest. Transverse flutes and recorders can be distinguished by octavial versus duodecimal (chalumeau-like) spectra, although this does not apply to each single tone. The recorder is a unique instrument with unique properties and expressions. Even besides temporal effects like the onset behaviour, simply by analysing the steady-state condition it shows its specific timbre characteristics. Further clarification is needed with view to the relation between variability of timbre and music aesthetics of the respective era and to its significance for HIP.

Acknowledgements

Many thanks to Eva Kuen for her expertise, helpful suggestions and proof reading. I also want to address my gratitude to the musicians who provided the sound recordings, especially to Eva Kuen, Alexandra Kollo, and Anja Lautermann.

References

[1] Adrian Brown: Database. https://adrianbrown.org/renaissance-recorder-database/ (5.3.2022)

[2] Gustav Scheck. Die Flöte und ihre Musik. Mainz: B. Schott's Söhne, 1975

[3] Ardal Powell. The flute. The Yale Musical Instrument Series. Yale University Press, 2002

[4] Dorothea Seel. Der Diskurs um den Klang der Flöte im 19. Jahrhundert. Augsburg: Wißner, 2020

[5] Kai Siedenburg, Charalampos Saitis, Stephen McAdams, Arthur N. Popper, and Richard R. Fay (eds.). Timbre: acoustics, perception, and cognition. Heidelberg: Springer Nature, 2019

[6] Albert S. Bregman. Auditory scene analysis. Cambridge, MA: MIT Press, 1990

[7] Groffroy Peeters, Bruno L. Giordano, Patrick Susini, Nicolas Misdariis, and Stephen McAdams. The Timbre Toolbox: Extracting audio descriptors from musical signals. *JASA,* 130, 2902 (2011), DOI: 10.1121/1.3642604

[8] Christoph Reuter and Julie Delisle. Sound radiation and other features of flute instruments, 2019, https://www.univie.ac.at/muwidb/flutes/ (8. 3. 2022)

[9] Hermann (von) Helmholtz. Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik. Braunschweig, 1863

[10] Karl Erich Schumann. Zur Physik der Vokalklangfarben. In: Heinrich Hüschen (ed.). Musicae Scientiae Collectanea, Köln: Arno-Volk-Verlag, 1973

[11] Joe Wolfe. Flute acoustics. https://newt.phys.unsw.edu.au/music/flute/ (8.3.2022)

[12] Christoph Reuter. Die auditive Diskrimination von Orchesterinstrumenten. Frankfurt/Main: Peter Lang, 1996

[13] Thomas Lerch. Der Klang historischer Blockflöten. Ein Berechnungsverfahren zur Klangvorhersage. In: Günther Wagner (ed.). Jahrbuch des Staatlichen Instituts für Musikforschung Preußischer Kulturbesitz, Stuttgart: J. B. Metzler, 1997, 146-68

[14] Donald E. Hall. Musical acoustics. Pacific Grove (Ca.): Brooks/Cole Publishing Company, 1991

[15] Jörg Jewanski. Ist C = Rot? Eine Kultur- und Wissenschaftsgeschichte zum Problem der wechselseitigen Beziehung zwischen Ton und Farbe. Sinzig: studio, 1999

[16] Julia Simner and Edward M. Hubbard (eds.): The Oxford handbook of synesthesia. Oxford: University Press, 2013

[17] Jörg Jewanski, Rustem Sakhabiev, and Christoph Reuter. About blue flutes and red trumpets. Features and tendencies of timbre-color mappings. In: Anton Sidoroff-Dorso and Sean A. Day (eds.). Synaesthesia Conference Moscow 2019, in print

[18] Michael Haverkamp. Klangfarben-Variabilität von Blockflöten aus Renaissance und Barock. *Annual Conference DAGA 2022*, Stuttgart. Berlin: DEGA, 2022

[19] Hans-Martin Linde. Handbuch des Blockflötenspiels. Engl. edition: The recorder-players handbook. Mainz: B. Schott's Söhne, ^{2}1984

[20] Theobald Boehm. Die Flöte und das Flötenspiel, 1871. Engl. edition: The flute and flute-playing. Translation and edition by D. C. Miller. New York: Dover, 1964

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Interdependencies of Humming, Rumbling and Booming

Florian Doleschal, Gloria-Tabea Badel, Jesko Verhey

Department of Experimental Audiology, Otto von Guericke University Magdeburg

*florian.doleschal@med.ovgu.de

Abstract
Humming, rumbling and booming are sensations that require the presence of low-frequency tonal components in the noise. If it only contains a low-frequency tone the sound is perceived as humming, if this component is modulated it is perceived as rumbling and if more tonal components with higher frequencies are present that are related to the low-frequency tonal component (e.g. are harmonics), the sound is perceived as booming. Sounds eliciting these sensations are often observed in vehicle acoustics when downspeeding the engine, which is a strategy to reduce fuel consumptions in vehicles with combustion engines. The present study investigates how these sensations interact using original sounds from the vehicle interior and sounds where parts of the spectrum are altered to change the magnitude of the sensations. The results show that the three sensations correlate in some aspects. Furthermore, dominance effects have been observed, so that measures to manipulate a certain sensation do not take effect if another sensation is much more pronounced.
Keywords: Rumbling, Humming, Booming, Interdependencies

1 Introduction

Running combustion engines at lower speeds is a common approach to improve the fuel efficiency of passenger vehicles. At these low engine speeds, low-frequency sensations, such as humming, rumbling and booming, are often associated with the vehicle interior sound and can severely reduce its pleasantness [1]. While the sensation of humming is elicited by the appearance of tonal components in a frequency range below 100 Hz, the sensation of rumbling involves additional amplitude modulations. The sensation of booming is an interaction of low- and medium-frequency components, where the modulation frequency of the medium-frequency components is the same as the audio frequency of the low-frequency component [1]. Even though, the sensations involve some common signal properties and therefore imply certain interdependencies, the terms humming, rumbling and booming are commonly used interchangeably with different meanings [1, 2, 3]. Therefore, an experiment was set up, where the three sensations were evaluated for the same set of stimuli, which involve both original recordings and recordings with spectro-temporal manipulations [4]. The results show, to which extent the sensations are interdependent of each other and if certain sensations are dominant.

2 Experiment

2.1 Introduction phase

Many listeners are not familiar with the exact meaning of the terms humming, rumbling and booming in the context of vehicle sounds. Therefore, an introduction phase was included in the experiment, where both

artificial and recorded vehicle interior sounds were presented to the listeners to familiarize them with the sensations. The synthetic sounds (see Figure 1) contain key features of the sensations, which have been determined by a previously conducted experiment with experts in vehicle acoustics. The humming introduction sound consists of a single 50-Hz sinusoid. In contrast, the rumbling sound is an amplitude modulated 40 Hz sound with a modulation frequency of 8 Hz. The resulting sound therefore consists of three sinusoids, a 40-Hz sinusoid and two sinusoids with frequencies of 32 and 48 Hz. As mentioned before, booming is an interaction of low- and medium-frequency components. Therefore, the introduction sound for the sensation booming consists of a low-frequency sinusoid at a frequency of 38 Hz and higher frequency components with frequencies of 152 Hz, 304 Hz, 456 Hz and 608 Hz, which are modulated at a rate of 38 Hz.

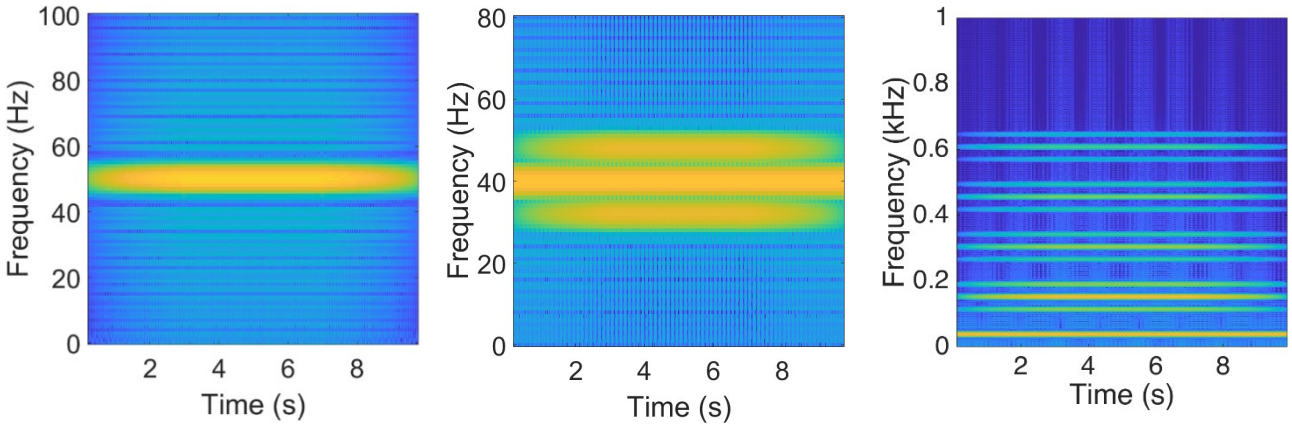


Figure 1: Spectrograms of the introduction sounds for the sensations humming (left panel), rumbling (center panel) and booming (right panel).

In addition to the synthetic introduction sounds, for each sensation, one recorded sound, which strongly elicited the respective sensation and one sound, which did not elicit any of the three sensations (see Figure 2), were presented. The presented sounds have been determined by experts in the field of vehicle acoustics and were not part of the stimuli set of the main experiment. The sound, whose spectrogram is displayed in the first row of Figure 2, strongly elicited the humming sensation, due to the dominant low frequency components at frequencies of 50 Hz and 100 Hz. The sound of the second row contained a low-frequency component at a frequency of about 30 Hz with audible amplitude modulations, eliciting a rumbling sensation. The third sound involves a low-frequency component and several modulated higher-frequency components, eliciting a strong booming. The last sound was of a noisy character and did not elicit any of the three sensations. Therefore, it was presented as a counterexample of a sound eliciting the respective sensation.

All sounds were switched on and off using raised-cosine ramps with a duration of 10 ms. The listeners were allowed to listen to both the artificial introduction sounds and the recorded introduction sounds until they were familiar with the sensation. Then, they started the main experiment.

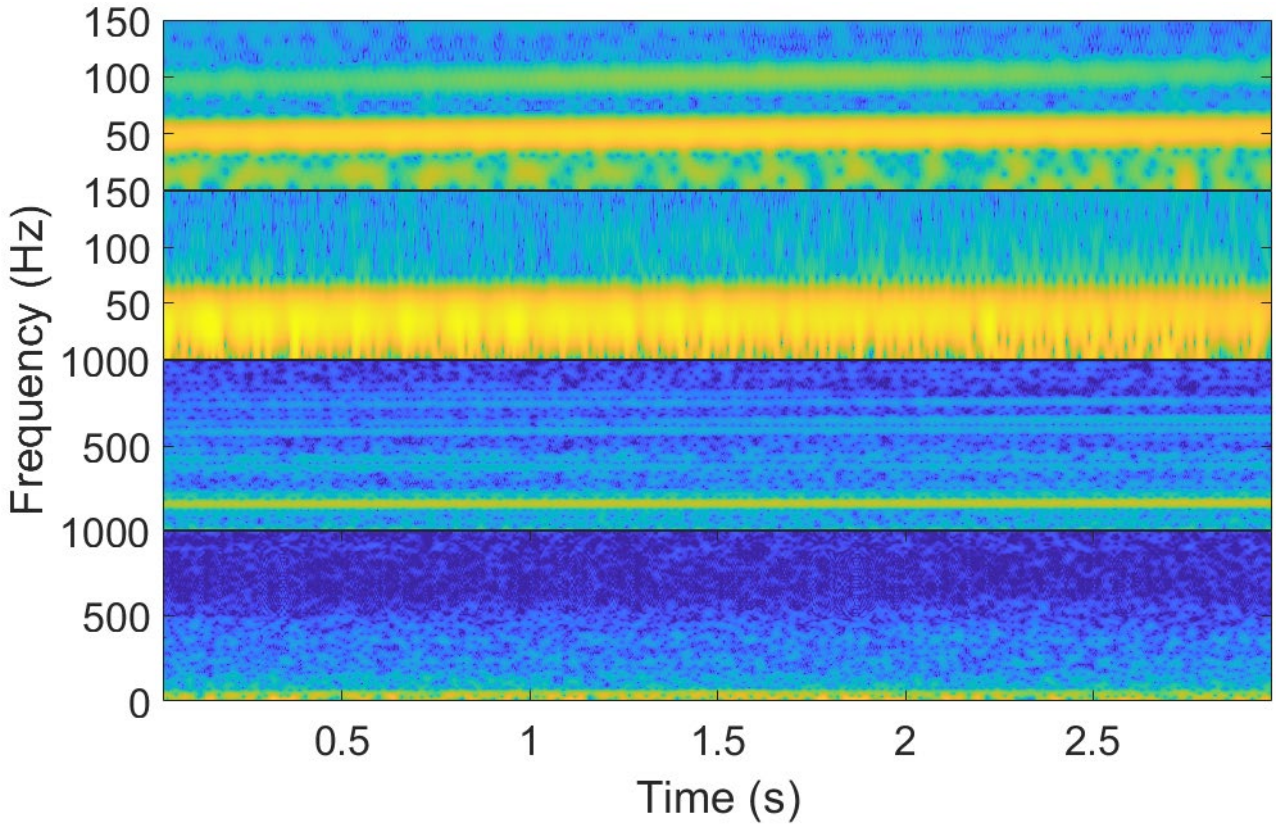


Figure 2: Spectrograms of the recorded introduction sounds for the sensations humming (first row), rumbling (second row), booming (third row) and a sound, which did not elicit any of the three sensations (fourth row).

2.2 Experimental procedure

After the listeners completed the introduction phase, the main experiment started. Before the first run of each sensation, the listeners listened to all sounds of the experiment one after another to become familiar with the dynamic range of the sensation. Afterwards, each sound was played individually in a random order. The task was to rate the respective sensation of each sound on a 9-point Likert scale (see Figure 3). The study was carried out in German, so that both the originally used German terms and the English translations are provided. As an example, for the sensation humming ("Brummen"), the extreme value 1 was labelled with "nicht brummend" ("not humming") and the extreme value 9 with "extrem brummend" ("extremely humming"). For further orientation, the scale ticks 3, 5 and 7 were labelled with "wenig brummend" ("little humming"), "mittel brummend" ("medium humming") and "deutlich brummend" ("clearly humming"). The scale ticks were adapted from the method of categorical loudness scaling [5]. In contrast to the introduction phase, the listeners were not allowed to listen to the sounds repeatedly. For the other two sensations rumbling and booming, the terms "wummernd" and "dröhnend" were used instead of "brummend". Each listener rated all sensations for all sounds three times. The results of the three runs per sensation were averaged and the overall mean values and the interindividual standard errors were calculated.

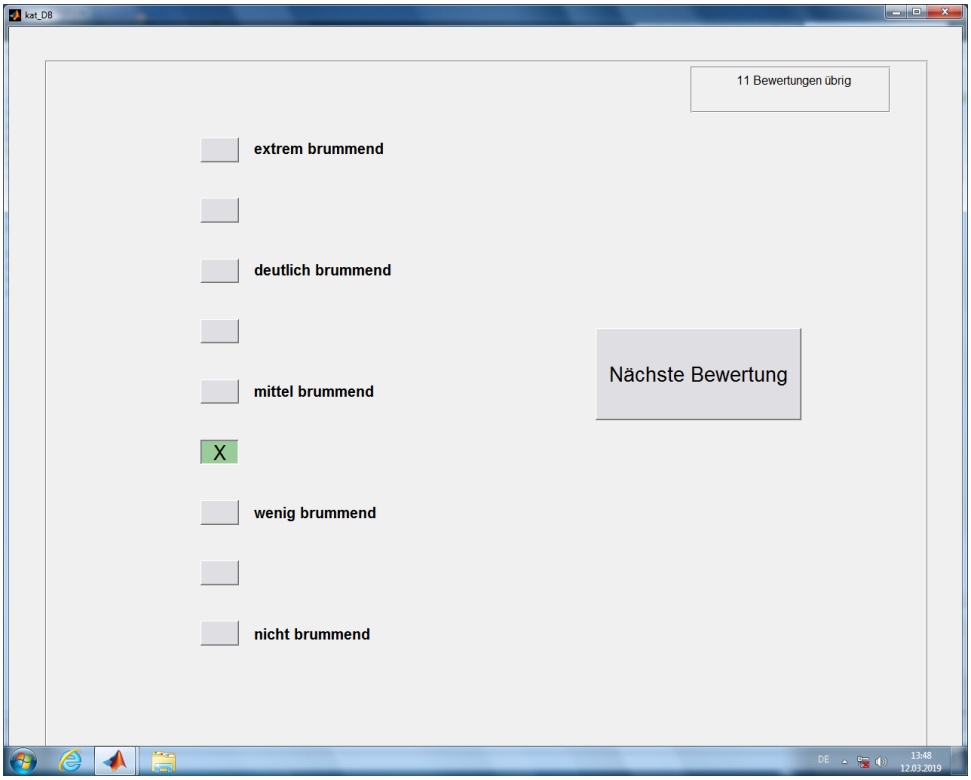


Figure 3: User interface for the categorical rating of the sensation humming (“Brummen”).

2.3 Apparatus

The listeners were seated in a sound-attenuating booth. The stimuli were converted from digital to analog signals using the sound card RME Fireface UC and presented binaurally via Sennheiser HD 650 headphones. The listeners rated each sensation using a touchscreen in front of them.

2.4 Stimuli

Table 1: Variations of the five original sounds of the experiment. The first column shows the sound number of the respective original sound. The second column shows the altered frequency bands. The third column shows the level variation steps of the frequency bands shown in the second column.

Sound number	Frequency bands	Level variations
19	30-42 Hz (1. C.)	-20 dB
	28-34 Hz & 40-46 Hz (SB)	-20 dB
42	35-45 Hz (1. C.)	+5 dB, +10 dB
49	36-44 Hz (1. C.)	-20 dB, -5 dB
63	34-42 Hz (1. C.)	-5 dB
	72-80 Hz (2. C.)	-5 dB
69	30-40 Hz (1. C.)	+5 dB, +10 dB, +15 dB
	30-40 Hz & 68-74 Hz (1. & 2. C.)	+5 dB, +10 dB, +15 dB

For the study, five original sounds were used, which were binaurally recorded at the driver’s seat during slow run-up driving conditions, which are labelled as *Original*. From these original sounds, spectro-temporal variations were adapted by manipulating the level of certain sound components using the Sound Engineering

Tool of the HEAD Acoustics ArtemiS Suite 9.0. The variations were subcategorized into variations of the side bands of a low-frequency component (*SB*), variations of a low-frequency component itself (*1. C.*), variations of a low-frequency component, whose frequency is higher than the one of the first component (*2. C.*) and variations of both the first and second component (*1. C. & 2. C.*). The component levels were either decreased by 5 or 20 dB or increased by 5, 10 or 15 dB. The adjustment steps were selected in a way that the resulting variations sounded considerably different from each other and that the sound pressure levels did not exceed the maximum level according to the vote of the ethical committee. In total, ten variations (see Table 1) of the sounds were created in addition to the five original sounds, so that in total, 15 stimuli were evaluated in the experiment.

2.5 Listeners

Fourteen normal-hearing listeners participated in the experiment. None of them had any hearing difficulties and their audiometric thresholds were 20 dB HL or below for the standard audiometric frequencies between 125 and 8000 Hz.

3 Results and Discussion

Figure 4 shows the mean values and standard errors for the three sensations rumbling, humming and booming. Mean rumbling values range from 2.4 to 7.1, mean humming values from 2.4 to 6.7 and mean booming values from 2.3 to 7.4.

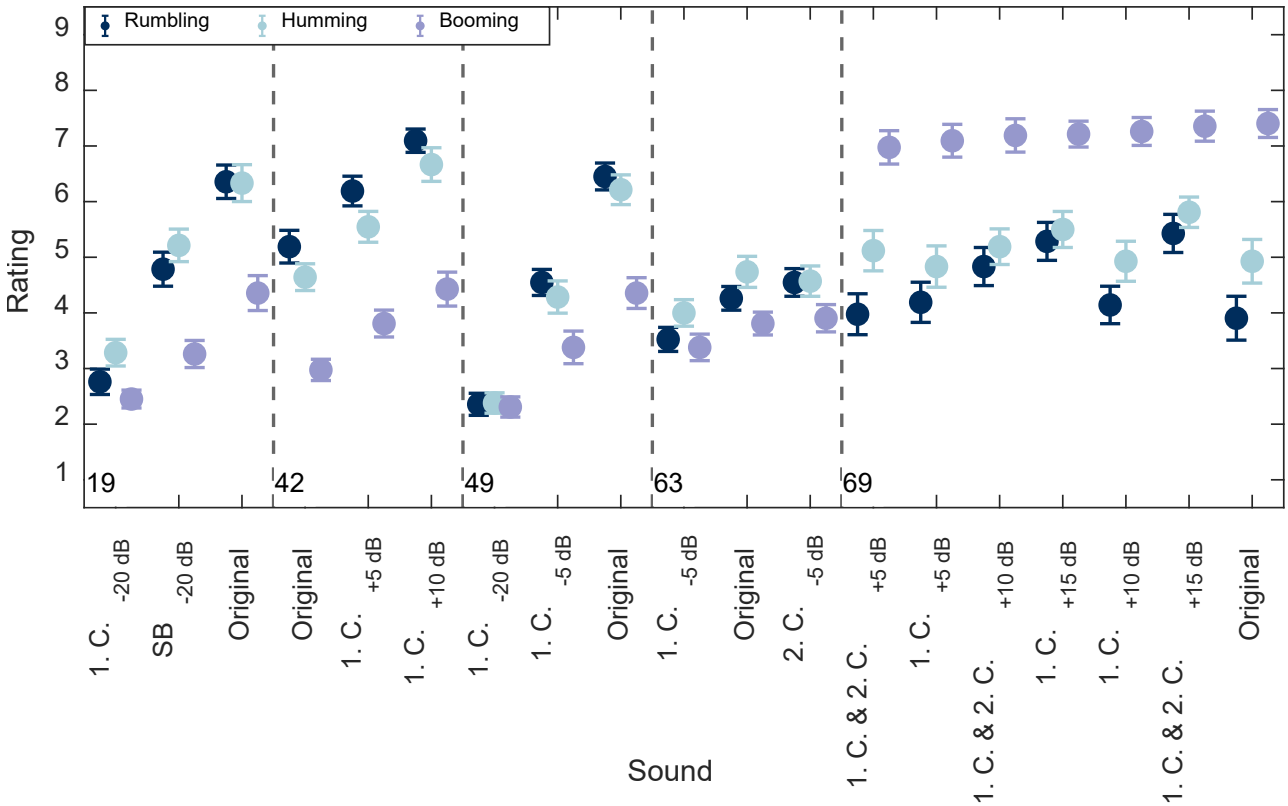


Figure 4: Mean values and interindividual standard errors for the three sensations rumbling (dark blue), humming (light blue) and booming (purple). The numbers in the bottom-left corner of each subpanel refer to the original sounds of Table 1. The labels of the abscissa indicate the respective variation of each sound. The ordinate shows the magnitude of the three sensations, ranging from 1 to 9.

For sound 19, the attenuation of the first sound component by 20 dB resulted in a stronger decrease of the sensations rumbling and humming than the attenuation of the side bands. The influence of on the perceived booming is even smaller, presumably because only low-frequency components were altered. Similar observation can be made for the results of sounds 42 and 49. For sound 42, the level increase of the first component results in a slightly stronger increase of the sensations rumbling and humming than of the sensation booming. This difference is more pronounced for sound 49. While the *Original* sound elicited relatively strong rumbling (6.5) and humming (6.2) sensations, the perceived booming was lower (4.4). A level reduction of the first component by 20 dB resulted in nearly equal values for the three sensations. The variations of sound 63 and sound 69 only revealed small differences for the three sensations. While for sound 63, the small changes could be explained by the relatively small changes of only 5 dB for both the first and the second component, it does not explain the results of sound 69. For sound 69, the increases of the first, the second and both components by up to 15 dB resulted in only relatively small changes in the perceived rumbling and humming. For the *Original* sound all its variations, the booming was similar and much higher than for the other two sensations. This indicates that the sensation of booming dominates the sensations of humming and rumbling for this sound. The listeners hardly seem to notice the changes in the low frequency range when rating the sensation. The more or less constant high booming magnitude is presumably a result from the medium-frequency components which are the same for all these sounds.

When analyzing the interdependencies between the sensations, the rank correlation coefficient according to Kendall [6] of 0.66 reveals that the sensations rumbling and humming were correlated. This correlation may be explained on the basis of the common spectral characteristics that elicit these sensations. Both sensations have in common that they solely depend on the appearance of low-frequency components. However, this study indicates that a strong booming sensation could result in only small changes in humming and rumbling when altering the relevant frequency bands. The rank correlation coefficient of 0.41 between the sensations humming and booming was slightly smaller. The correlation may be explained by the common contribution of an unmodulated low-frequency component to these sensations. However, booming further depends on the appearance of higher-frequency components, which were not altered in this study, hence the smaller correlation than for rumbling and humming. The results for sound 69 and its variations indicate that for prominent higher frequency components, changes of the low-frequency component have a negligible effect on booming. With a value of 0.17, the rank correlation between the sensations rumbling and booming is the lowest correlation between the sensations. This low correlation may be explained by differences in the key features of the sound for the two sensations. Even though rumbling and humming involve low-frequency components and audible modulations, not only the components that are modulated differ but also the modulation frequencies of the sounds eliciting a strong booming are commonly higher than those eliciting a strong rumbling [1].

4 Conclusions

The study shows, that the sensations humming, rumbling and booming are partially correlated, presumably due to common signal properties. All the three sensations depend on low-frequency components, while rumbling additionally involves amplitude modulations and booming is an interaction of a low-frequency component and higher frequency components. The study revealed a certain dominance effect of booming over the two other sensation when booming is large. Even though the terms humming, rumbling and booming are often used interchangeably in the literature, listeners can distinguish them when they are clearly defined prior to the experiment. The sensations should be clearly distinguished, because the sensory impressions and the impact on the pleasantness are considerably different. Overall, a clear definition of the sensations may help to better characterize sound perception and to conduct the appropriate countermeasures to improve sound quality.

Acknowledgements

This report is the scientific result of a research project undertaken by the FVV (The Research Association for Combustion Engines e. V.) and performed by the Otto von Guericke University Magdeburg under the direction of Prof. Dr. Jesko Verhey and the Acoustics Group at Carl von Ossietzky University Oldenburg under the direction of Prof. Dr. Steven van de Par. The FVV would like to thank the professors van de Par and Verhey and their scientific research assistants Dr. Arne Oetjen (Oldenburg) and Gloria-Tabea Badel (Magdeburg) for the implementation of the project. The project was conducted by an expert group led by Dr. Harald Stoffels (Ford Werke GmbH) and Dr. Sebastian Lucas (Volkswagen AG). We gratefully acknowledge the support received from the chairmen and from all members of the project user committee. The research project was self-financed (FVV funding no. 1304) by the FVV (Research Association for Combustion Engines e. V.).

References

[1] K. Genuit, B. Schulte-Fortkamp, A. Fiebig and M. Haverkamp, "Bewertung von Fahrzeuggeräuschen," in *Sound-Engineering im Automobilbereich*, Berlin Heidelberg, Springer, 2010, pp. 109-182.

[2] S. Hatano and T. Hashimoto, "Booming index as a measure for evaluating booming sensation," *Proceedings of the 29th International Congress and Exhibition on Noise Control Engineering 27-30 August 2000, Nice, FRANCE,* 2000.

[3] H. Lee and S. Lee, "Objective evaluation of interior noise booming in a passenger car based on sound metrics and artificial neural networks," *Applied Ergonomics 40 (5),* 2009.

[4] G.-T. Badel, F. Doleschal and J. Verhey, "Spektro-temporale Geräuschmanipulationen als Grundlage zur Erforschung der Empfindungsgröße Wummern," in *Fortschritte der Akustik - DAGA 2020*, Hannover, Deutsche Gesellschaft für Akustik e. V., 2020, pp. 313-314.

[5] T. Brand and V. Hohmann, "An adaptive procedure for categorical loudness scaling," *Journal of the Acoustical Society of America 112 (4),* pp. 1597-1604, 2002.

[6] M. Kendall, "A New Measure of Rank Correlation," *Biometrika 30 (1/2),* pp. 81-93, 1938.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Experimental evaluation of earplug behavior in front of high-level impulse noises.

Cyril Blondé-Weinmann[1,2,*], Thomas Joubaud[2], Pascal Hamery[2], Sébastien De Mezzo[2], Véronique Zimpfer[2], Sébastien Roth[1].

[1]Laboratoire Interdisciplinaire Carnot de Bourgogne, UMR 6303 CNRS/Univ. Bourgogne Franche-Comté, France.

[2]Acoustics and Soldier Protection, French-German Research Institut of Saint-Louis, 68300 Saint-Louis, France.

[*]cyril.blonde-weinmann@isl.eu

Abstract

High-level impulse noises such as weapon noises or mine charges can cause irreversible damage to the hearing system. Hearing protectors are used to reduce airborne sound propagation. However, the resulting attenuation of protectors worn singly or in combination is limited and may be insufficient in some extreme situations. One reason for these limitations is the behavior of the protectors under high stresses. For example, high-level impulse noises could induce a slight earplug movement in the ear canal. In order to quantify this effect, a new experimental set-up with an artificial simplified ear canal is developed. Thus the earplug movements are measured with a laser Doppler vibrometer and a high-speed camera under various configurations (high-impulse stimulation levels, ear canal lengths, earplug adjustments, and positions). These investigations highlight relative displacements that can exceed one millimeter and modify the final earplug position in the ear canal. The observed effects could be responsible for an alteration of the protection efficiency. New hypotheses are formulated to limit these phenomena and improve hearing protectors.

Keywords: high-level impulse noise, hearing protection, earplug.

1 Introduction

Prolonged exposure to high-level noises can irreversibly damage the sensory cilia of the organ of Corti and result in varying degrees of hearing loss. Alone or in combination with continuous high-level noise, impulses can cause more injuries to the auditory system [1]. There are many sources of impulse noises, such as gunfires, explosions, and landmines. These are defined by their maximum peak pressure and A-duration (positive phase duration), two characteristics that reflect their danger. Indeed, the greater the amplitude, the more harmful the consequences on the hearing system. In addition, the impulse duration is also decisive since it can reach the inner ear by supplanting the stapedius reflex [2]. It is conventional to use hearing protectors to occlude the ear canal external meatus to attenuate the direct air propagations. However, this protection is partial since it provides only a limited attenuation depending on the type of occlusion and the stimulation characteristics [3]. The addition of a second protector does not always improve the protection [4] which gradually erodes the idea that waves propagate only through air paths. Therefore, imperfect ear canal occlusion is not the only source of hearing protection limitation. The tissue conduction by the skin [5], the bones [6], or the cartilages [7] are, for example, as many paths taken by the waves. It then activates mechanisms of hearing at the origin of an auditory perception [8] by loading the hearing protection environment. In addition, some of the hearing protectors' limitations could result directly from the protection behavior itself. Indeed, depending on the impulse characteristics, passive protectors can adopt structural non-linearities, which make their behaviors difficult to

apprehend [9]. For instance, in the case of high-level impulses, it has been demonstrated that earmuffs act as acoustic antennas by capturing the waves and retransmitting them to the external ear tissues [10]. This could partly explain the insufficiency of the earmuffs worn alone for the protection against impulse noises [11]. Similarly, earplugs could also carry consequences, perhaps even more critical because of their positioning close to the eardrum. Indeed, the induced ear canal walls displacement resulting from earplugs [12] are in favor of waves transmissions through the secondary acoustic paths. This paper will study the earplug behavior using a simplified Artificial Ear Canal (AEC) and a laser Doppler vibrometer (LDV). Although this measuring device can provide accurate results, the context of impulse noise makes the operations difficult because of the diverse signal alterations. The protocol presented in this study lists these alterations and proposes an adapted methodology. A comparison with a high-speed camera (HSC) is performed to compare the results. Besides, the various situations considered allowed to evaluate the influence of the plug's initial position, its adjustment, the length of the ear canal, and the incident impulse level.

2 Materials and methods

2.1. High-level impulse noise generation

Impulse generation using loudspeakers does not achieve the levels encountered by military personnel, which may reach a 190 dB peak. In this study, the impulse waves were generated using explosive charges. Indeed, the use of explosive detonation has the advantage of generating impulses whose characteristics (peak level and A-duration) will depend on the weight of the charge and the distance. It allows an adjustment to obtain the required amplitude. The earplug's behavior has been studied in the case of impulses of 172 and 176 dB peaks. In Table 1, the different explosive charges (mass and type), as well as the distances for the impulse waves generation, are described for these two levels. The resulting temporal pressure evolution after the impulse is shown for both loads in Figure 1.

Table 1: Required type and mass of explosive charge and distance between the explosive charge and the measurement point.

Peak pressure level [dB]	Charge masse [g]	Charge type [-]	Charge distance [m]
172	70	$C4$	7.0
176	220	$C4$	7.0

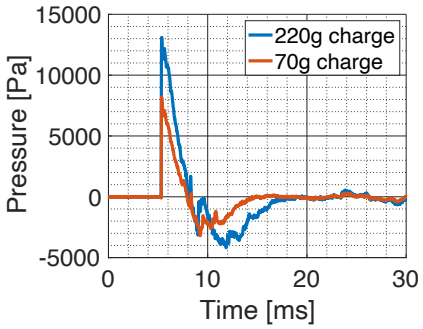
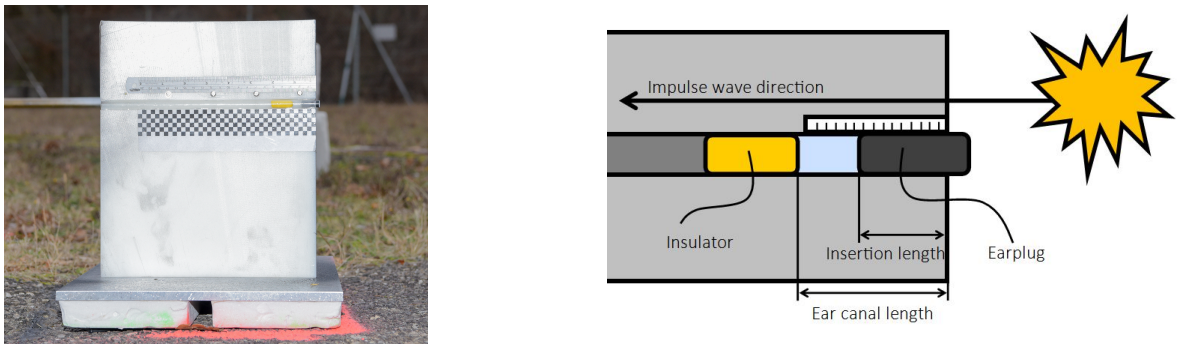


Figure 1: Pressure variations at 7.0m for a 70g and 220g C4 charges.

2.2. Earplug displacement visualization in simplified ear canal

It would be difficult to visualize the longitudinal earplug movement in the ear canal of an artificial head (acoustic test fixture). For this reason, an experimental device was built to make the plug movement visible. A transparent artificial ear canal of 24.7 cm in length and 8.05 mm in diameter is used to allow measurements under various conditions of lengths. In particular, it was desired to evaluate the displacement of the plug without pressure constraints resulting from a closed air volume while limiting the propagation of the wave on the opposite side of the occlusion. The device is presented in Figure 2a. The simplified cylindrical geometry of this artificial ear canal cannot be held as an exact representation of reality. However, numerical simulations in the frequency

domain have shown a limited impact for the 1.5-4 kHz frequency range. Outside this range, the consequences are negligible [13]. As the frequency spectrums of the impulses used in this study are mostly low frequencies (under 1 kHz), this simplification can bring first relevant elements compared to a human ear canal. This geometry also respects the intention to recreate an axisymmetric numerical model of the experiment. The material used to manufacture the canal is plexiglas®. Its thickness is 2.2 mm. The use of rigid material is justified by the will to limit secondary acoustic conductions, in particular, the canal deformations following the shock wave, although they are inevitable for very high-level impulses. The canal is inserted in polyethylene support to limit lateral stresses from the impulse side. The insulation with the vibrations propagated by the ground is made with a layer of polystyrene. A part of the artificial ear canal is exposed to measure the earplug displacements visually. The differences with an artificial head are related to the coupling between the plug and the canal due to the materials and possible effects of the pinna, not transcribed here. A 3M E-A-R Classic acoustic foam plug is used as an insulator to limit the ear canal length. The displacement is measured in the direction of the propagating impulse wave. These designations are presented schematically in Figure 2b.



(a) AEC photography. (b) AEC schematic representation.

Figure 2: AEC and corresponding study lengths designations.

2.3. Earplug displacement measurement

Two measuring devices capture the longitudinal earplug's motions. A high-speed camera PHOTRON SAZ with a 50 kHz sampling frequency is positioned face to the uncovered lateral side of the artificial ear canal. An OPTOMET SWIR laser Doppler vibrometer with a 51.2 kHz sampling frequency points at the base of an Acrylonitrile Butadiene Styrene (ABS) earplug made in our laboratory and records the motion velocity of the earplug in the AEC axis with a dynamic range of 245 mm/s. The measurement of the impulse noise with the LDV requires to take into account the superposition of various signals:

- the perturbation of the impulse wavefront entering the laser measuring field with the duration of the initial impulse;
- the useful signal, i.e., the actual displacement of the plug measured by the vibrometer, also approximatively of the initial impulse duration;
- the disturbance of the LDV body by the impulse that can last longer because of the induced mechanical vibrations and the duration of the dispersion;
- The involuntary reflections of the environment (as ground reflection for instance) that have a limited impact on the measurement of the plug displacement due to energy dissipation.

To avoid a superposition of these signals, it is essential to distribute sensors, source, and target so that each perturbation intervenes separately in time. Therefore, the propagation time and duration of each perturbation

must be taken into account. A measurement of the perturbations resulting from the disposition shown in Figure 3 is given in Figure 4. For example, the beginning of perturbation 1 corresponds to the propagation of distance d_1, which value is given in Table 2.

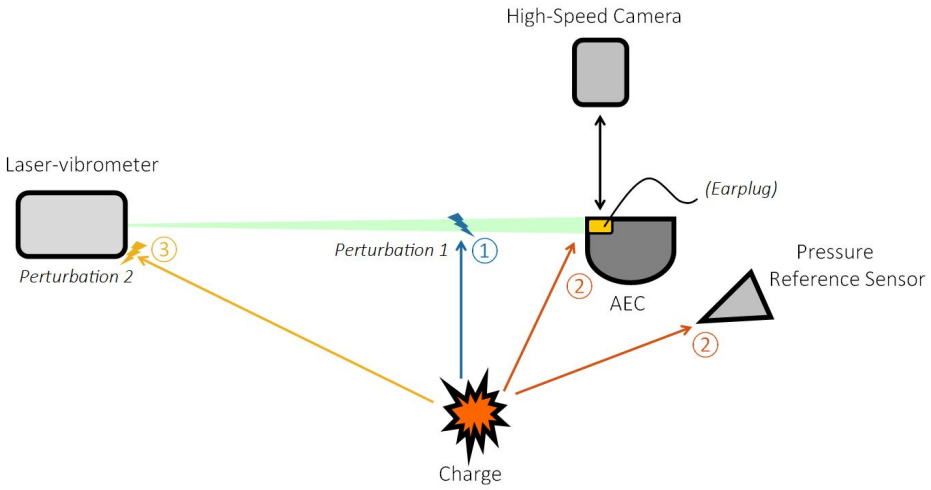


Figure 3: Positioning of the source, ear canal and sensors. The pressure reference sensor was manufactured in our laboratory from a KISTLER-6031 quartz sensor.

The distances between the sensors considered and the signal obtained with the laser vibrometer are explained respectively on the Figure 4 and the Table 2.

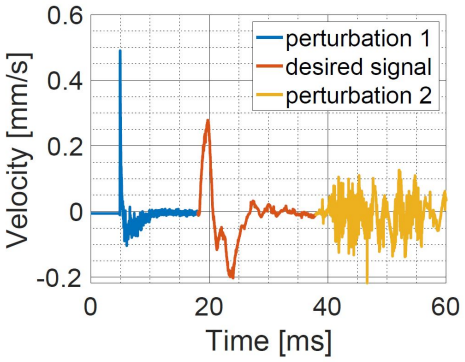


Figure 4: Velocity measured with the LDV with the distances between the sensors considered in the Table 2 allowing for a disambiguation of the perturbations and desired signals.

Table 2: Distance and approximate time from which the perturbation would be visible on the measurement signal of the laser vibrometer.

Distance name [-]	Distance value [m]	Approx. prop. duration [ms]
d_1	2.0	6
d_2	7.0	20
d_3	14.1	41
d_4	20.7	–

2.4. Configurations studied

(a) Comparative study between the laser Doppler vibrometer and the high-speed camera tracking

First, a comparative study between the displacement of an 8.05 mm diameter ABS plastic earplug measured with the LDV and with the HSC was carried out in order to compare the obtained results. The LDV allows a temporal evaluation of the earplug's position. In contrast, the HSC snapshots evaluation aims to evaluate the plug's extreme positions (initial position, maximum insertion, maximum extraction, and final position). This is due to the displacement measurement accuracy, which is much coarser with the HSC than with the LDV. To do this, four iterations of the shooting protocol for an open ear canal and four iterations for a 22.5 cm length ear canal were performed. It was assumed that the longer canal involved higher pressure variations and would less

damped plug displacements. Thus, it would allow better visualization of the displacements on the HSC whose accuracy is lower than the LDV and about 0.2 mm.

b) Behavior of the plug under different conditions

To visualize the effects of various mechanical and physical conditions on the earplug's movement, three sets of measurements were performed:

- the initial position influence of an adjusted plug protector in the case of an open artificial ear canal was studied. Two measurements were performed: a measurement with a plug totally inserted (inserted condition) and a measurement mostly extracted from the canal (partially inserted condition).
- Another series of measurements were realized with a 3 cm length artificial ear canal with two different earplug diameters. One measurement was performed with an adjusted earplug of 8.05 mm diameter and a second with a non-adjusted plug of 8.00 mm diameter.
- Last but not least, the incidence of the impulse peak level on the earplug behavior was also evaluated. For an 8.05 mm diameter earplug and a 3 cm length artificial ear canal, 172 dB-peak, and 176 dB-peak levels were studied.

3 Results

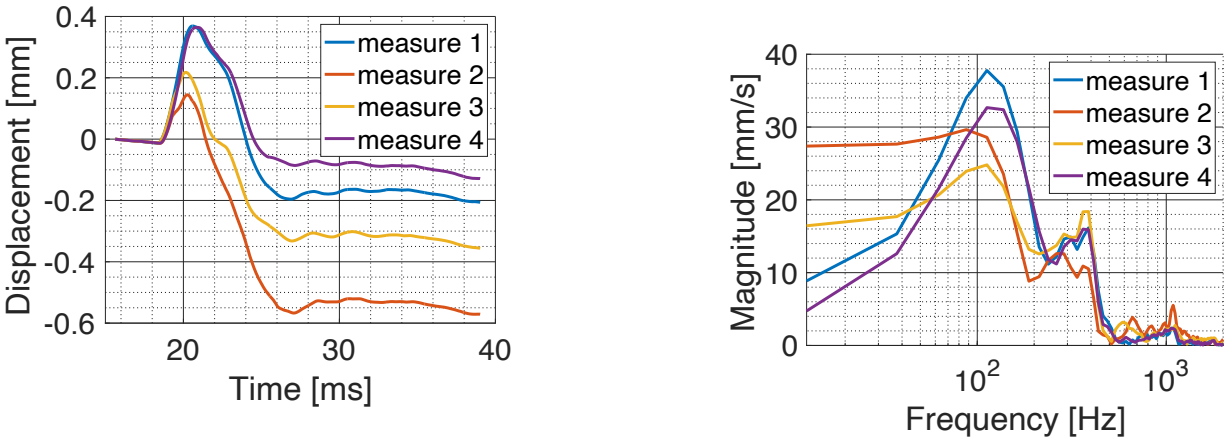
3.1. Comparative study between the laser vibrometer and the fast camera tracking

a) Open artificial ear canal

The time displacement and spectral velocity of an 8.05 mm diameter ABS earplug for four 172dB-peak charge iterations with an artificial open ear canal are presented in Figure 5. The earplug's positions (maximal insertions and extractions as well as final positions) determined with the HSC for each iteration and its comparison with the LDV measurement are listed in Table 3. As the initial positions of the earplug were different for each iteration, they were also reported in the same table. It can be observed that the behavior of the earplug respects a similar evolution for the four measurements. However, significant variations of the maximum insertion (extreme values measured 0.14 mm and 0.37 mm) and the final relative position (extreme values measured -0.09 mm and -0.53 mm) appear. Nevertheless, the variations between the maximum insertion and maximum extraction displacement (extremum distance) remain globally constant for all measurements. The spectral representations in Figure 5b highlight a displacement velocity concentrated mainly at low frequencies with two remarkable spectral densities at about 0.1 kHz and between 0.3 and 0.4 kHz corresponding to the resonant frequency of the open canal that appears theoretically at 345 Hz.

Table 3: Comparison between the HSC relative displacements and LDV measured relative displacements for the open ear canal configuration. Positive values correspond to a displacement in the propagation direction of the impulse wave.

Measurement	Measurement 1		Measurement 2		Measurement 3		Measurement 4	
Initial insertion [mm]	17.4		17.2		19.5		17.0	
Sensor	HSC	LV	HSC	LV	HSC	LV	HSC	LV
Relative maximal insertion [mm]	0.4	0.37	0.1	0.14	0.2	0.22	0.4	0.36
Relative maximal extraction [mm]	−0.2	−0.20	−0.5	−0.57	−0.3	−0.33	−0.1	−0.09
Relative final position [mm]	−0.2	−0.16	−0.5	−0.53	−0.3	−0.32	−0.1	−0.09
Extremum distance [mm]	0.6	0.57	0.6	0.71	0.5	0.55	0.5	0.45



(a) Time displacement

(b) Spectral velocity

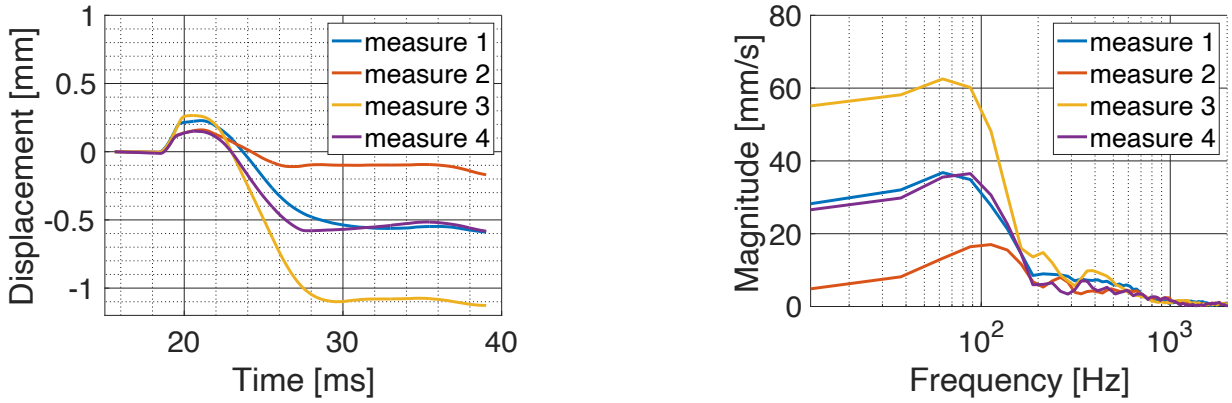
Figure 5: LDV measurement for an artificial open ear canal, 8.05 mm diameter ABS earplug, and 172 dB-peak pressure impulse level stimulation. Positive values correspond to a displacement in the propagation direction of the impulse wave. HSC values are listed in Table 3.

b) 22.5 cm length closed artificial ear canal

The time displacement and spectral velocity of an 8.05 mm diameter ABS earplug for four 172dB-peak charge iterations with an artificial 22.5 cm length closed ear canal are presented in Figure 6. The ear canal is closed with a foam earplug on the lateral side in this configuration. The positions of the earplug determined with the HSC for each iteration and its comparison with the LDV measurement are listed in Table 4. Again, the earplug movement evolution remains the same for all measurements, but very significant variations appear for the same configuration. Suppose the relative insertion has a certain reproducibility from one measurement to another (extreme values measured of 0.27 and 0.15 mm). In that case, especially the final position differs significantly (extreme values measured of -0.10 and 1.08 mm). These final position variations corroborate the spectral behavior of the displacement velocity at low frequencies, as presented in Figure 6b. Then, the extreme distances in this configuration are not reproducible. Besides, no correlation between the earplug's initial position and the earplug's final position seems to be sketched. It is also remarkable that occlusion on the opposite side of the artificial canal leads to suppressing the spectral components above 0.2 kHz. Comparisons between the values measured with the LDV and the HSC allow finding consistent results between the two methods. The differences between the two sensors are contained within the measurement uncertainties of the HSC ($\pm$ 0.2 mm).

Table 4: Comparison between the HSC relative displacements and LDV measured relative displacements for the 22.5 cm length closed ear canal configuration. Positive values correspond to a displacement in the propagation direction of the impulse wave.

Measurement	Measurement 1		Measurement 2		Measurement 3		Measurement 4	
Initial insertion [mm]	15.7		16.1		15.1		17.2	
Sensor	HSC	LV	HSC	LV	HSC	LV	HSC	LV
Relative maximal insertion [mm]	0.3	0.23	0.2	0.16	0.3	0.27	0.2	0.15
Relative maximal extraction [mm]	-0.5	-0.56	-0.1	-0.11	-0.9	-1.10	-0.6	-0.58
Relative final position [mm]	-0.5	-0.55	-0.1	-0.10	-0.9	-1.08	-0.5	-0.52
Extremum distance [mm]	0.8	0.79	0.3	0.27	1.2	1.37	0.8	0.73



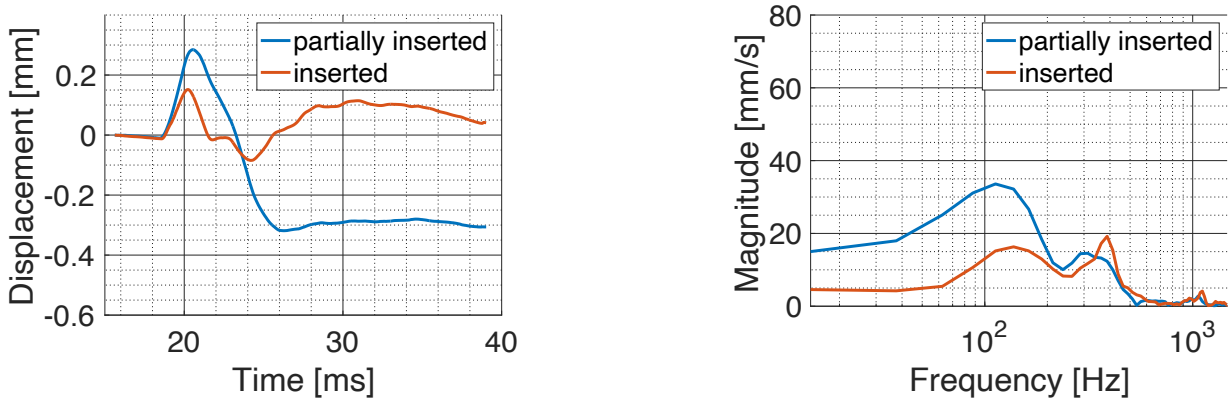
(a) Time displacement (b) Spectral velocity

Figure 6: LDV measurement for an artificial 22.5 cm length closed ear canal, 8.05 mm diameter ABS plug, and 172 dB-peak pressure impulse level stimulation. Positive values correspond to a displacement in the propagation direction of the impulse wave. HSC values are listed in Table 4.

3.2. Behavior of the earplug under different conditions

a) Influence of earplug initial position

The time displacement and spectral velocity of an 8.05 mm diameter ABS earplug for four 172dB-peak charge iterations with an artificial open ear canal are presented in Figures 7. The "inserted" condition corresponds to a total insertion of the earplug, i.e., 2 cm: only the lateral flat face of the earplug is exposed to the impulse wave. The "partially inserted" position refers to the insertion of two-thirds, i.e., 1.2 cm. The behavior is very different for the two initial positions. The "partially inserted" condition leads to an almost two-times higher insertion. Above all, the final position of the plug is approximately -0.3 mm extracted from the initial position. The "inserted" condition performs an in-out-in oscillation before reaching a final position slightly more inserted than the initial one (less than +0.1 mm). The spectral velocity represented in Figure 7b is also very different, with much lower low-frequency components for the "inserted" condition and a remarkable spectral peak at 0.4 kHz more accentuated than for the "partially inserted" condition.

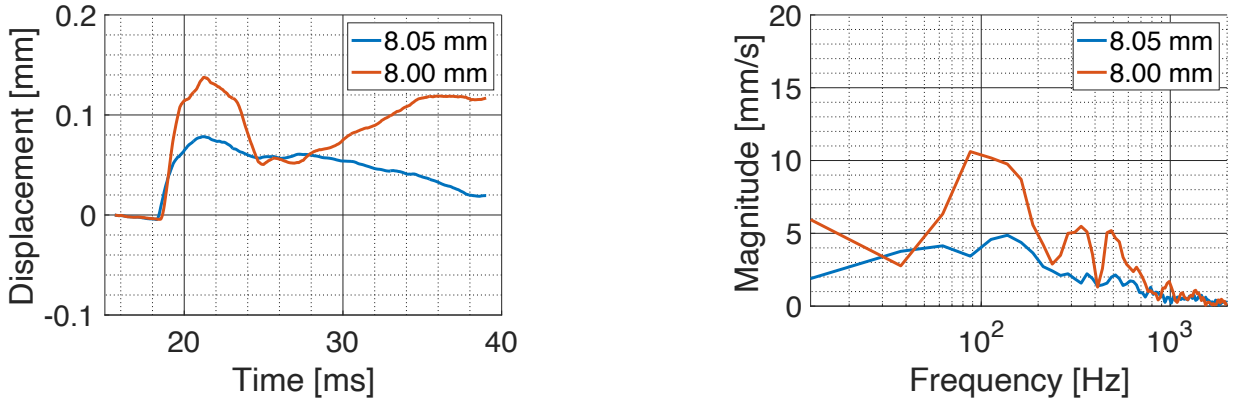


(a) Time displacement (b) Spectral velocity

Figure 7: LDV measurement for an open ear canal, 8.05 mm diameter ABS earplug, and 172 dB-peak pressure impulse level stimulation for the two initial plug position conditions. Positive values correspond to a displacement in the propagation direction of the impulse wave.

b) Influence of earplug diameter

The time displacement and spectral velocity for two ABS earplug diameters (8.00 and 8.05 mm) on 172dB-peak charge stress with an artificial 3 cm length closed ear canal are presented in Figure 8. The 8.00 mm diameter plug that can move without constraint (except the gravity forces) in the artificial canal has a first insertion amplitude following the attack of the wavefront that is nearly 60% greater than that of the adjusted plug. Nevertheless, it is mainly the final position of the two earplugs that distinguishes them: the 8.05 mm diameter earplug tends to find a final position close to the initial position with insertion of 0.02 mm. In contrast, the 8.00 mm diameter earplug oscillates and tends towards an inserted position of 0.12 mm. The 8.00 mm diameter earplug has a higher spectral peak between 0.1 and 0.2 kHz than the 8.05 mm diameter plug, as well as additional peaks at 0.3 kHz and 0.5 kHz, as visible in Figure 8b.



(a) Time displacement

(b) Spectrale velocity

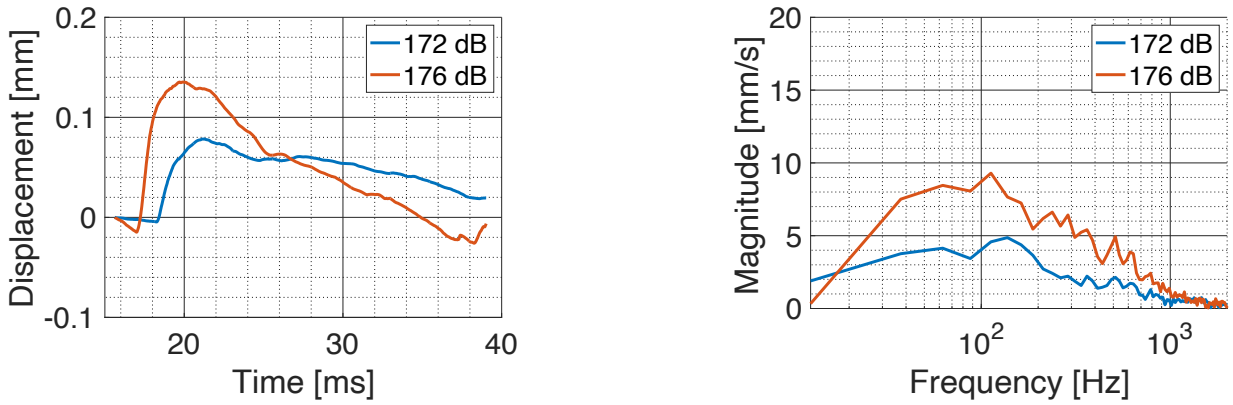
Figure 8: LDV measurement for a 3-cm length closed ear canal, two diameters ABS earplugs and 172 dB-peak pressure impulse level stimulation. Positive values correspond to a displacement in the propagation direction of the impulse wave.

c) Influence of the impulse peak level

The time displacement and spectral velocity for an 8.05 mm diameter ABS earplug under two different charge stresses (172 dB-peak and 176 dB-peak) with an artificial 3 cm length closed ear canal are presented in Figure 9. A 4 dB higher charge leads to 1.75 times larger insertion displacement. This displacement seems to constrain the earplug more. It tends to return to its initial position quicker, leading to a consequent extraction displacement for the 176 dB than the 172 dB charge (0.02 mm extraction displacement for 176 dB and 0.02 mm insertion displacement for 172 dB). The spectral velocity represented in Figure 9b does not show the appearance of significant peaks between the two charges. In general, the magnitudes between the two conditions appear only amplified.

4 Discussion

A new measurement protocol was used to evaluate the behavior of an earplug activated by a high-level impulse under different configurations. The results obtained with an HSC were compared to those acquired with an LDV. This comparison highlighted a good correspondence between the two methods. The advantage of LDV measurement is that it enables a more precise and accurate measurement than an HSC. This significant result makes it possible to consider measurements with the LDV when visibility required for the use of the HSC is not possible. However, it is necessary to judiciously place the sensors, target, and source on the experiment scheme



(a) Time displacement

(b) Spectrale velocity

Figure 9: LDV measurement for a 3 cm length closed ear canal, 8.05 mm diameter ABS earplug and two dB-peak stimulation condition. Positive values correspond to a displacement in the propagation direction of the impulse wave.

to avoid the superposition of undesired disturbances on the desired signal. Besides, these comparisons have also highlighted a significant disparity between the measurements under the same configuration. If the extreme position displacements are small in an open canal, this situation is not encountered in the case of a 22.5 cm occluded canal. The initial position of the plug could have explained these variations. However, no correlation could be found for minor variations of this position. From a spectral point of view, these differences are also illustrated by low-frequency spectral components of different magnitudes for the displacement speed. These differences must be considered in the interpretation of the subsequent tests. For initial positions of the plug fully inserted and two-thirds inserted, the plug behaves significantly differently. A fully inserted plug appears to possess a more inserted final position, while a two-thirds partially inserted plug possesses a more extracted final position. This could be due to two combined reasons. The first could be the result of the mechanical impact of the wavefront. This part depends only on the depression degree of the plug. The second could result from the plug and ear canal surface conditions. Thus, the coupling between the plug and the ear canal appears to depend on the mechanical properties of the materials involved in the interaction. Further studies on the materials and the impact of mechanical properties on the coupling would quantify these effects and limit the harmful consequences. In addition, it might be possible to design a geometry that limits the movement of the plug, in particular the insertion movement, which could result in an increase in pressure at the eardrum and a depression that would decrease the insertion of the plug. Incorporating fixation at the scaphoid fossa or tragus might be a practical option. The final position of the plug seems to be an important consideration: the outward displacement of the plug that occurs with a single impulse could be repeated with each impulse. It could be dangerous because of the effectiveness loss of the protector with a shallower insertion. This seems even more true when the plug is poorly adjusted, and the impulse is higher. Indeed, the earplug fitting is naturally a sensitive issue. A tight plug leads to less plug movement. However, this is multifactorial data with implications on both comfort and wave retransmission to the ear canal walls by solid conduction. Thus, a single consideration of the displacement is insufficient, and a global approach must be taken. These characteristics should be studied in more detail using a real geometry of an artificial canal (for example, with an acoustic test fixture, with a more realistic ear canal). Then a measurement of the pressure alterations behind the protection would quantify the dangers for the auditory system and highlight the main weakness of the protections.

5 Conclusion

A new measurement protocol using an LDV has made it possible to assess the displacement of a plug in an AEC during high-level impulses up to a 176 dB peak. After being validated by comparisons with an HSC, the experimental evaluation highlighted significant variations from one measurement to another and seemed to designate the plug-canal coupling as essential data in the protector movement. The impulse level, the plug adjustment, and the plug's initial position are all parameters that can negatively influence these movements. Further research is needed to quantify these earplug movements' impact on eardrum pressure. In parallel, modeling works are carried out to transcribe the observable phenomena and explain their origins and mechanical interactions.

References

[1] A Dancer, K Buck, P Hamery, G Parmentier, et al. Hearing protection in the military environment. Noise and Health, 2(5):1, 1999.

[2] Prem G Nair, N Shivashankar, B Indira Devi, SG Srikanth, V Shanmugham, KS Gayathri, et al. Acoustic reflex decay and acoustic reflex latency threshold test findings in patients with cerebellopontine angle tumors: Correlation with tumor type, size, and extent. Amrita Journal of Medicine, 16(4):164, 2020.

[3] William J Murphy, Cameron J Fackler, Elliott H Berger, Peter B Shaw, and Mike Stergar. Measurement of impulse peak insertion loss from two acoustic test fixtures and four hearing protector conditions with an acoustic shock tube. Noise & health, 17(78):364, 2015.

[4] Yu Luan, Olivier Doutres, Hugues Nélisse, and Franck Sgard. Experimental study of earplug noise reduction of a double hearing protector on an acoustic test fixture. Applied Acoustics, 176:107856, may 2021. doi: https://doi.org/10.1016/j.apacoust.2020.107856.

[5] Haim Sohmer. Soft tissue conduction: Review, mechanisms, and implications. Trends in Hearing, 21: 233121651773408, oct 2017. doi: 10.1177/2331216517734087.

[6] G. Von Bekesy. Zur theorie des horens bei der schallaufnahme durch knochenleitung. Annalen der Physik, pages 111--136, 1932.

[7] H Hosoi. Receiver. Japanese Patent Application, (166644), 2004.

[8] Cyril Blondé-Weinmann, Thomas Joubaud, Véronique Zimpfer, Pascal Hamery, and Sébastien Roth. Characterization of cartilage implication in protected hearing perception during direct vibro-acoustic stimulation at various locations. Applied Acoustics, 179:108074, aug 2021. doi: https://doi.org/10.1016/j.apacoust.2021.108074.

[9] Pascal Hamery, Véronique Zimpfer, Karl Buck, and Sébastien De Mezzo. Very high level impulse noises and hearing protection. Euronoise, 2015.

[10] Cyril Blonde-Weinmann; Thomas Joubaud; Véronique Zimpfer; Pascal Hamery; Sébastien Roth. Involvement of the outer ear's cartilage and soft-tissues in hearing protection limitation during high-level impulse noise exposition. In ICSV, 2021.

[11] J Starck, E Toppila, H Laitinen, G Suvorov, V Haritonov, and T Grishina. The attenuation of hearing protectors against high-level industrial impulse noise; comparison of predicted and in situ results. Applied Acoustics, 63(1):1--8, jan 2002. doi: 10.1016/S0003-682X(01)00025-1.

[12] S Benacchio, O Doutres, A Le Troter, A Varoquaux, E Wagnac, Virginie Callot, and F Sgard. Estimation of the ear canal displacement field due to in-ear device insertion using a registration method on a human-like artificial ear. Hearing research, 365:16--27, 2018.

[13] Takuji Koike, Hiroshi Wada, and Toshimitsu Kobayashi. Modeling of the human middle ear using the finite-element method. The Journal of the Acoustical Society of America, 111(3):1306--1317, mar 2002. doi: http://dx.doi.org/10.1121/1.1451073.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Perception of Virtual Reality Based Audiovisual Paradigm for People with Hearing Impairment

Kang Sun[1,2,*], Niels H. Pontoppidan[1], Dorothea Wendt[1,3], Lars Bramsløw[1]

[1]Eriksholm Research Centre, Oticon A/S, Snekkersten, Denmark.

[2]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark.

[3]Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark.

*knsu@eriksholm.com

Abstract
Integrating new and emerging technologies, such as virtual reality (VR), with established test methods to improve the ecological validity is gradually used by practitioners in some fields (e.g., soundscape research) but remains limited use in hearing science. In this paper, an audiovisual setup was introduced to create an augmented speech-in-noise test for people with hearing impairment (HI). The environment containing four competing talkers was recorded with 360-degree video and ambisonics audio. In the scene recreating the environment, the video was displayed in the VR headset, while the audio was presented to the participants via a circular loudspeaker array. Furthermore, a (silent and fixed) physical avatar (manikin) was included in the video recording as a placeholder for the audio stream of the target speech, which was added into sound presented over the loudspeaker array. This setup was used to test the effect of different hearing aid (HA) settings under the same condition for each participant. In total, 27 HI participants were tested. VR has been known to cause motion discomfort, which is referred as VR sickness or cybersickness nowadays. The simulator sickness questionnaire (SSQ, [11]) was used to quantify the sickness measurement. The data showed a low degree of Nausea and Disorientation, but scattered Oculomotor responses. Moreover, a general questionnaire assessing scene recreation, test method and outcome expectation was administrated. In general, the audiovisual system received high appraisal in realism; the augmented speech-in-noise test method was well accepted; participants highly agreed that the difference between programs could be distinguished. However, the sense of physical immersion decreased due to the weak binding between the avatar and the target speech. Furthermore, when comparing the three components (Nausea, Disorientation, Oculomotor) in SSQ and items in the general questionnaire, the Oculomotor was found significantly correlated to the perceived binding of target speech and the avatar. Specifically, participants who were less convinced that the speech came from the avatar, also rated the Oculomotor higher.
Keywords: virtual reality, hearing impairment, perception, audiovisual

1 Introduction

Conventionally, research fields involve human perception of sounds, such as soundscape, adopt laboratory tests in audio-only presentation. However, driven by the notion that many established methodologies in laboratory lack sufficient realism to produce adequately findings in real life, adopting a holistic manner in delivering multisensory information in laboratory studies to improve their ecological validity gets increasingly popular [1]. For example, researchers in soundscape address the role of visual component via the audiovisual interaction (e.g., [2]) in recent years. Until now, audio-only presentation has dominated in laboratory tests in

hearing science. Though being a relatively new concept, ecological validity, in hearing science, refers to the degree to which research findings reflect real-life hearing-related function, activity, or participation [3].

New and emerging technologies – virtual reality (VR) – was integrated with established test method and used in soundscape studies (e.g., [4,5]). The technology is often well-perceived as it increases the immersion of human participation. In hearing science, effort has been taken recently to use VR as a part of rehabilitation tool for young persons with hearing difficulty [6,7], with a particular focus on stimulated environments [8]. However, the reproduction of real-life recorded environments in lab were not particularly interested due to the increased difficulty of augmentation to meet the study designs. VR has been known to cause motion discomfort, which is referred as VR sickness or cybersickness nowadays. One of the major concerns of applying VR on people with hearing impairment (HI) lies on that older (65+ yrs) adults reported more sickness than younger adults in simulated environments [9]. Furthermore, hearing loss was found associated with poorer spatial awareness [10], which might bring further discomfort in VR environments. Nevertheless, limited attention was paid on investigating the perception of HI on the VR technology, particularly in recorded real-life environments reproduction.

In this paper, we present a VR based audiovisual paradigm with augmented real-life recording, in which HI could perform a speech-in-noise test in a close-to-real-life ambience. To obtain a sketch of the paradigm, participants were recruited and instructed to evaluate a set of hearing aid (HA) programs. We used a well-known stimulator sickness questionnaire (SSQ, [11]) to evaluate the sickness symptoms of HI experiencing our setup. Furthermore, a comprehensive questionnaire was used to assess the reproduction quality, test method and outcome expectation.

2 Method

2.1 Audiovisual scene recording

The initial audiovisual scene recording was administrated in a canteen. Four actors (two male and two female) were recruited to sit in prescribed positions to engage a paired conversation (i.e., the two actors on the same side to the avatar had a conversation, Figure 1). The actors were provided a list of topics for non-invasive daily conversation such as hobby, vacation, food. The actors were also instructed to maintain the consistency of the conversation, that is to avoid rapid volume change or silence from both pairs. Furthermore, an avatar, represented by a B&K 4128 head and torso simulator (HATS, Brüel & Kjær, Denmark), was placed in the recording as a placeholder for the target speech in the lab reproduction. The HATS manikin only served as an avatar and had no acoustic purpose. All actors and the avatar were positioned 1.4m from the recording position, as shown in Figure 1.

At the recording position, a GoPro Fusion 360 camera was used for 360-degree video (5.2k, 4992 x 2496 resolution, 30 fps) and synchronized 4-channel first-order (B-format) ambisonics recording. The camera was set at the eye level of the avatar, and both the camera and the avatar were aligned to the centre of the table. Moreover, an additional microphone linking to a B&K 2250 sound level meter (Brüel & Kjær, Denmark) was hanged on the ceiling right above the recording position. The total duration of the audiovisual recording was 60 min. The sound level meter performed 5-min segment recordings continually and collected 12 recordings.

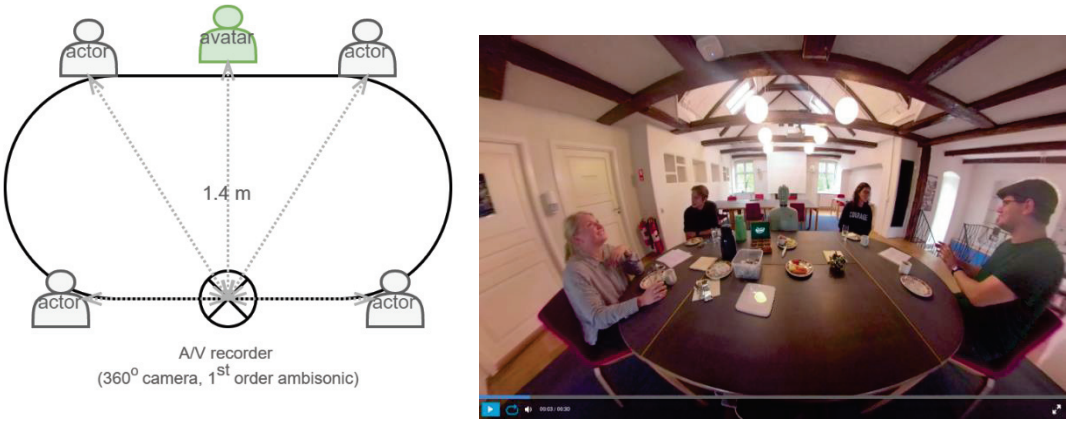


Figure 1 – Sketch of the audiovisual scene recording. (Left: top view of the relative positions of the actors, the avatar and the A/V recorder, all actors and the avatar were 1.4m from the A/V recorder, sketch is not true to scale; Right: the recorded view from the point of the A/V recorder, i.e., the participant view in the test.)

2.2 Audiovisual scene reproduction and augmentation

The reproduction of the audiovisual scene was administrated in an anechoic room. This study used Oculus Go, a stand-alone device, as the head-mounted display (HMD) for visual recreation. Though Oculus Go provided integrated headphones and even possibility to customize headphones, this option was omitted considering the test participants in this study all wore a pair of HAs. Instead, a loudspeaker array containing 24 loudspeakers forming a circle of 1.4m radius was used for sound reproduction. All loudspeakers were horizontally aligned and evenly distributed at an interval of 15 degrees, as shown in Figure 2.

Recorded footage were processed with GoPro Fusion studio. As a raw rendering, the audiovisual document was saved in a format of H.264 video codec, 5.2k resolution, and 360 audio (ambix) from Fusion Studio. This rendering was then split into a 4-channel B-formatted audio in ambix convention and a muted video document (no audio). A decoding procedure using Matlab AmbiToolbox transformed 4-channel audio into 24-channel audio while maintaining spatial characteristics. The decoding exported audio matching the loudspeaker positions as per se, and reconstructed the sound field at the centre of the array. The 24-channel audio was played via Reaper (v6.13, suggested by [12]) in this study. As the audio was presented via the loudspeakers and the video via the HMD, a time alignment procedure was implemented. First, a single channel linear time sync audio (LTC) was merged with the muted video. The merged document was saved in the HMD and played via the Skybox VR player. Then while the LTC-merged video was played, the coded signal LTC indicated a LTC reader software (Reaper) the timestamp of the audio (via an audio cable, indicated in Figure 2-Right), and thus synchronized the decoded 24-channel audio. The playback was calibrated at the array centre (i.e., participant's seat), using the sound level meter recordings as reference. To maintain the equivalence of sound level across the whole experiment, the overall average sound level was adjusted (the adjustment was limited) and fixed at 65 dB ($L_{eq,5min}$ continually throughout the 60 min recording).

Since the essential form of this study was a speech-in-noise test, the audiovisual scene augmentation used the standard speech material archive Danish HINT (hearing in noise test, [13]) sentences as target speech and the recorded scene as the mask (i.e., noise). A separate HINT audio was merged to the front loudspeaker relative to the participant's seat (the loudspeaker marked as T in Figure 2-Left, overlapped with the avatar). During the test, it was presented as that the actors were having conversations as the mask, while the avatar was delivery HINT sentences, from the participant's view. The HINT audio was also calibrated at the array centre, and a graphic user interface (developed in Matlab) was provided to the experimenter to manually adjust the speech level, and hence signal-noise ratio (SNR), during the test.

Figure 2 – Sketch of the recreation system. (Left: top view of the loudspeaker array of 24 loudspeakers in a circle of 1.4m radius, the VR position indicates the seat for the participants, the shaded portrait indicates the relative positions of the recorded avatar and actors reflected in the loudspeaker array; Right: an illustration of a test participant during the test, an audio cable was connected to the VR headset, which sent timestamp to Reaper for synchronization.)

2.3 Questionnaire

The simulator sickness questionnaire (SSQ) from [11] was used. SSQ contained 16 questions and derived into three categories: Nausea, Oculomotor, and Disorientation. The score scale, the calculation means, and the coefficients were true to [11] (details attached in Appendix). Moreover, a general questionnaire containing four aspects was designed and administered. As shown in Table 1, Q1-Q3 described the audiovisual scene recreation from perspectives including realism (Q3), immersion (Q2) and the binding between the speech and the avatar (Q1). Q2 and Q3 were adopted from [4,14]. Q4 and Q5 concerned the test method focusing on the perception of the augmentation (Q4) and the test objective (Q5). Q4 was adopted from [15]. Q6-Q8 expressed the outcome expectation and the repeatability of the test. Lastly, Q9 and Q10 asked the representativity of the reproduced scenario (Q9) and the general attitude in coping with it (Q10). All questions were answered in a 5-point likert-scale (from Strongly Disagree to Strongly Agree), started by asking "To which degree you agree with the following statement".

Table 1: Overview of the general questionnaire.

Category	No.	Questionnaire
Scene recreation	1	I sense the speech is from the avatar.
	2	I feel physically immersed in the presented scene.
	3	I feel the presented scene is real.
Test method	4	I can accept to talk to a chatbot like the avatar.
	5	I sense the difference of different programs in hearing during the experiment.
Outcome expectation	6	I believe this format of experiment will bring me better hearing aid fitting.
	7	I can describe my hearing experience to an audiologist better in this test format.
	8	I wish to experience more of this type of test scenarios.
Others	9	I encounter the type of listening environment in this experiment a lot in life.
	10	I try to avoid the presented situation in this experiment in my life.

2.4 Participants and Procedure

The protocol of the study was approved by the Capital Region of Danish Committee System on Health Research Ethics (Hovedstaden, reg. nr.: H-20068237). Twenty-seven HI (10 female, 17 male; $mean_{age}$=72, $median_{age}$=74, SD_{age}=8.6, age range: 51-87 yrs.) with mild to moderate hearing loss were recruited as test participants. All participants were provided with a consent form before the experiment and a signature was obtained from each participant.

The participants were instructed to evaluate a set of HA programs under the same condition. As stated before, the mask (noise) was set to 65 dB. First, in the training phase, an adaptive procedure was applied to estimate the signal-to-noise ratio (SNR), at which participants reached 50% of speech intelligibility (SNR50) with the same reference program of the HAs. Afterwards, in the test phase, each participant was tested at fixed SNR, i.e. their individual SNR50, and thus a fixed sound level of the target speech (HINT). Across all participants, the average SNR was -0.19 dB (SD=2.33, range: -5~+5). During the test, the participants sat comfortably in the chair in the centre of the loudspeaker array. The chair was height adjustable so that the loudspeaker array and the ears of the participants were set at the same level. As the sanity check of the spatial alignment of the VR view and the loudspeakers, the participants were asked to point out the avatar in the VR view at the beginning of wearing the HMD, before the loudspeakers were activated (i.e., in silence). The physical pointing position was required to be at the front loudspeaker (T). After all the test, the questionnaires were presented to the participants via a tablet.

3 Results and Analysis

3.1 Perception of the audiovisual paradigm

In [11], three principal factors were extracted from 16 symptoms in the SSQ, namely Nausea (nausea, stomach awareness, increased salivation, burping), Oculomotor (eyestrain, difficulty focusing, blurred vision, headache), and Disorientation (dizziness, vertigo). Detailed list is attached in the Appendix. Each of the three factors was used as the basis for an SSQ subscale, and together they formed an SSQ profile for an assessed item. The proportional result of the SSQ profile of the audiovisual scene was shown in Figure 3. The mean proportional score in all factors was lowest in Nausea (mean=6.70%, SD=6.12%), followed by Disorientation (mean=6.88%, SD=9.55%) and highest in Oculomotor (mean=15.87%, SD=12.30%). In similar studies using HMD and adopting the SSQ, [16] involved younger normal hearing (NH) participants to perform a visual task, which suggested a comparable Nausea (mean proportional score 18%) response but a much higher Oculomotor (39%) and Disorientation (42%) response than the current study. [17] also involved younger NH participants to experience the VR scene with a visual discrimination task, which reported no absolute score in any SSQ factors but a deviation under different test conditions. However, six out of twenty participants withdrew the test, which indicated a high degree of discomfort.

In this study, the older HI shown limited symptoms of Nausea and Disorientation after the audiovisual paradigm, which was comparable to (and even better than) the younger NH peers in other studies. Earlier research [9,18] suggested more sickness expected in older participants in complex environments. However, a correlation analysis on age and the SSQ items (including three factors and the total score) in the current study shown no significant correlations (all p>.05). The current study used a static recording of an indoor, familiar, and real-life environment [19], which could contribute to a low degree of cybersickness.

Nevertheless, a sphere space was presented to the participants and 26 (out of 27) participants experienced a pair of VR glasses for the first time. Adaptation was expected, which might explain the slight increase in Oculomotor response (compared to Nausea and Disorientation). Furthermore, a hard task was involved in the test, which might increase the Oculomotor response, despite it being an audio task. Unlike visual tasks in which participants paid high attention to the visual presentations in VR (which might evoke considerable head rotation and body turn), an audio task did not require the participants to observe the visual in detail, which might explain the relatively lower SSQ response in this study than in [16] or potentially in [17]. In addition, the high standard deviation (SD) in Oculomotor (and Disorientation) response indicated a (relatively) large individual difference in this study. Nevertheless, the low SSQ profile in this study suggested that the audiovisual paradigm was not invasive and cybersickness friendly to HI. In fact, after the test, no participant reported any discomfort at any given time.

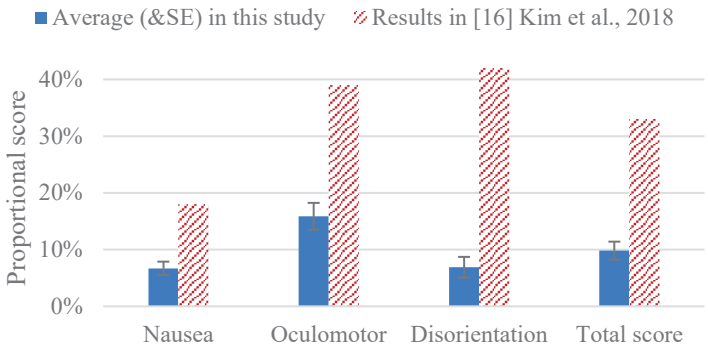


Figure 3 – SSQ profile (average score and standard error) of the audiovisual system in the test (All scores were converted to proportions to the full scale of each SSQ factors, and therefore the scale in this figure was unified. The corresponding interpretation of the degree of the symptoms: 0% – not at all, 33.33% – slightly, 66.67% – moderately, 100% – very. The slashed bars were results adopted from [16]).

Furthermore, Figure 4 illustrated the response of the general questionnaire. The participants perceived the reproduction as real (Q3, 81.47% agree or strongly agree). However, this appreciation decreased in immersion (Q2) and the binding between the avatar and the target speech (Q1). Particularly, 37.04% of participants were negative in answering Q1, while 33.33% of participants answered neutral in Q2. The results of immersion and realism were comparable to [14], which also used real-life recording reproduction but with no augmentation. Even though in this study an augmentation was attempted, the avatar still lacked features related to the audio and visual binding, such as lip movement, facial expression, and body language.

Moreover, the objective of the test was clear to the participants (Q5, 88.88% agree or strongly agree), suggesting such a paradigm was suitable for a speech-in-noise test format. A reasonably high degree of accepting the intention of augmentation (Q4, 66.66% agree or strongly agree) was achieved. Combining results in Q1, the participants understood and accepted the augmentation attempt, however, required a better tailored implementation. In addition, the repeatability of the test was high (Q8, 85.19% agree or strongly agree). The participants expected better hearing aid fitting after such type of test (Q6, 66.67% agree or strongly agree). However, describing their hearing difficulties remained a problem (Q7, 66.67% neither, disagree or strongly disagree), despite the improvement of ecological validity in this study. Lastly, though nearly 60% participants agreed that they met challenging scenarios as such in their lives (Q9). However, the attitude in coping with the challenge was almost evenly distributed, that is 44.44% of the participants wanted to face this type of scenarios while 40.74% decided to avoid them.

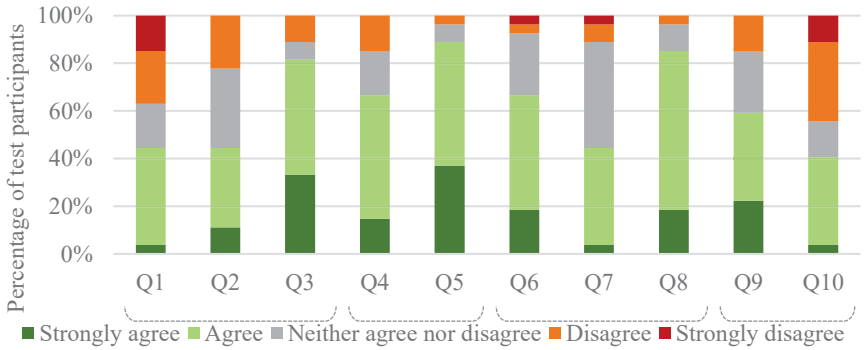


Figure 4 – The answer distribution of the questions in the general questionnaire. (Q1-Q10 refer to the question numbers in Table 1; parentheses group questions into four categories according to Table 1.)

The response of the general questionnaire was transferred into a scale comparable to the SSQ, in which "Strongly Disagree" was set to 0 and "Strongly Agree" was set to 4. Then a Person's correlation test was applied among the SSQ factors and questionnaire items (Table 2).

On one hand, all factors in the SSQ were significantly correlated with each other, which suggested that participants who rated higher score in one factor, also rated higher in the other two factors and therefore in Total score. That is, despite the three SSQ factors being distinct [11], participants had stronger cybersickness felt discomfort from all dimensions. A significant individual difference in VR sickness measurement could be expected. On the other hand, among the items in the general questionnaire, Q3 was significantly correlated to Q6, Q8, and Q9. It indicated that the more real the participants found the reproduction (Q3), the better outcome from the test (Q6) they expected, the better willingness to conduct more of such tests (Q8) they shown, and the more of similar environments in own lives as what shown in the reproduction (Q9) they experienced. Furthermore, Q4 was highly correlated to Q6 and Q7. Namely that the participants who had a good acceptance of the virtualized scene, also expected more positive outcomes from such test. Moreover, the three questions (Q6-Q8) in outcome expectation were all significantly correlated, which suggested that the refreshing technology and yet narrative context in this study could stimulate higher outcome expectation, better difficulty description and higher willingness of repeatability, at least for some participants. Lastly, Q6 was significantly correlated to Q10, which indicated that the participants who expected better HA fitting after such test also were more willingly to take the challenge of similar difficult communication environments in their lives.

Between the SSQ factors and the general questionnaire, Q1 was found significantly correlated to Oculomotor and Total score in the SSQ. That is, the participants who were less convinced of the binding between the avatar and the target speech, also rated the Oculomotor and the Total score higher. One potential explanation might be that the participants who were less convinced by the A/V augmentation, would visually focus less on the avatar but more on the ambience in VR, and thus triggered a higher Oculomotor response. However, this test lacked recordings of the head movement of the participants, this interpretation thus cannot be examined.

Table 2: Person's correlations between SSQ items and questionnaire items.

		SSQ				Scene recreation			Test method		Outcome expectation			Others	
	Var.	N[1]	O[2]	D[3]	TS[4]	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
	N[1]	—													
	O[2]	0.696 ***	—												
	D[3]	0.462 *	0.736 ***	—											
	TS[4]	0.772 ***	0.954 ***	0.867 ***	—										
Scene recreation	Q1	-0.187	-0.523 **	-0.261	-0.407 *	—									
	Q2	0.041	0.015	-0.176	-0.05	-0.157	—								
	Q3	0.019	0.073	0.171	0.107	-0.102	0.114	—							
Test method	Q4	-0.075	-0.286	-0.164	-0.224	0.304	-0.044	0.371	—						
	Q5	0.139	0.156	0.033	0.125	-0.291	0.213	0.26	0.167	—					
Outcome expectation	Q6	-0.067	-0.212	0.042	-0.105	0.23	-0.028	0.445 *	0.472 *	0.247	—				
	Q7	-0.059	-0.316	-0.294	-0.285	0.323	3.130 e -20	0.131	0.402 *	0.123	0.702 ***	—			
	Q8	-0.216	-0.237	-0.028	-0.182	0.19	0.118	0.422 *	0.615	0.151	0.719 ***	0.476 *	—		
Others	Q9	-0.333	-0.278	-0.056	-0.242	0.054	0.28	0.423 *	0.334	0.154	0.19	0.185	0.17	—	
	Q10	0.235	0.253	0.185	0.255	0.053	-0.035	0.004	-0.145	0.03	-0.38*	-0.36	-0.294	-0.033	—

N[1]: Nausea, O[2]: Oculomotor, D[3]: Disorientation, TS[4]: Total score;
***: p<.001, **: p<.01, *: p<.05;
Q1-Q10 refer to the question numbers in Table 1.

3.2 Limitation

In this study, the major concern from the participants' feedback of this audiovisual paradigm was the binding between the fixed avatar and the target speech, i.e., the implementation of the augmentation. The visual placeholder (avatar) lacking convincing features such as lip movement and facial expression might contribute to the weak binding perception. Furthermore, the room acoustic feature (e.g., reverberation time) was not applied for the target speech material, which could introduce a decrease of realism and immersion for the participants. Last but not the least, the purpose of this paradigm was to integrate VR technology with established test method to improve the ecological validity of the test. In this test, speech intelligibility was assessed via word score of the target speech, which was far from daily conversations. Meanwhile the ambience and the noise mask were a real-life recording. This controversial combination might prohibit the participants revivificating their hearing conditions in real life.

4 Conclusion

This paper reported an audiovisual paradigm, in which an augmented real-life recording was presented to perform a speech-in-noise test. The cybersickness perception of this paradigm was measured by the SSQ and a general questionnaire. The result indicated that it was a non-invasive and cybersickness friendly setup for HI, who experienced high immersion and shared good appreciation in the study. Furthermore, it provided an important enhancement for realistic and ecological HA assessment in laboratory.

Acknowledgements

This research received funding from Innovation Foundation Denmark and William Demant Foundation. These organizations are gratefully acknowledged. The authors would like to thank all the HI participating in this study, and the four actors in the recording. The authors would also like to thank Sergio Aguirre for helping the experiment setup.

References

[1] Aletta, F., Kang, J. and Axelsson, Ö., 2016. Soundscape descriptors and a conceptual framework for developing predictive soundscape models. Landscape and Urban Planning, 149, pp.65-74. doi: 10.1016/j.landurbplan.2016.02.001.

[2] Preis, A., Kociński, J., Hafke-Dys, H. and Wrzosek, M., 2015. Audio-visual interactions in environment assessment. Science of the Total Environment, 523, pp.191-200. doi: 10.1016/j.scitotenv.2015.03.128.

[3] Keidser, G., Naylor, G., Brungart, D.S., Caduff, A., Campos, J., Carlile, S., Carpenter, M.G., Grimm, G., Hohmann, V., Holube, I. and Launer, S., 2020. The quest for ecological validity in hearing science: What it is, why it matters, and how to advance it. Ear and hearing, 41(Suppl 1), p.5S. doi: 10.1097/AUD.0000000000000944.

[4] Sanchez, G.M.E., Van Renterghem, T., Sun, K., De Coensel, B. and Botteldooren, D., 2017. Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space. Landscape and Urban Planning, 167, pp.98-107. doi: 10.1016/j.landurbplan.2017.05.018.

[5] Sun, K., De Coensel, B., Filipan, K., Aletta, F., Van Renterghem, T., De Pessemier, T., Joseph, W. and Botteldooren, D., 2019. Classification of soundscapes of urban public open spaces. Landscape and urban planning, 189, pp.139-155. doi: 10.1016/j.landurbplan.2019.04.016.

[6] Zirzow, N.K., 2015. Signing avatars: Using virtual reality to support students with hearing loss. Rural Special Education Quarterly, 34(3), pp.33-36. doi: 10.1177/875687051503400307.

[7] Lau, S.T., Pichora-Fuller, M.K., Li, K.Z., Singh, G. and Campos, J.L., 2016. Effects of hearing loss on dual-task performance in an audiovisual virtual reality simulation of listening while walking. Journal of the American Academy of Audiology, 27(07), pp.567-587. doi: 10.3766/jaaa.15115.

[8] van de Par, S., Ewert, S.D., Hladek, L., Kirsch, C., Schütze, J., Llorca-Bofí, J., Grimm, G., Hendrikse, M.M., Kollmeier, B. and Seeber, B.U., 2021. Auditory-visual scenes for hearing research. arXiv preprint arXiv:2 doi: 10.48550/arXiv.2111.01237. 111.01237.

[9] Keshavarz, B., Ramkhalawansingh, R., Haycock, B., Shahab, S. and Campos, J.L., 2018. Comparing simulator sickness in younger and older adults during simulated driving under different multisensory conditions. Transportation research part F: traffic psychology and behaviour, 54, pp.47-62. doi: 10.1016/j.trf.2018.01.007.

[10] Jiam, N.T.L., Li, C. and Agrawal, Y., 2016. Hearing loss and falls: A systematic review and meta-analysis. The Laryngoscope, 126(11), pp.2587-2596. doi: 10.1002/lary.25927.

[11] Kennedy, R.S., Lane, N.E., Berbaum, K.S. and Lilienthal, M.G. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. The international journal of aviation psychology, 3(3), pp.203-220. doi: 10.1207/s15327108ijap0303_3.

[12] Spatial workstation. [Online]. Available: https://facebook360.fb.com/spatial-workstation/.

[13] Nielsen, J.B. and Dau, T., 2011. The Danish hearing in noise test. International journal of audiology, 50(3), pp.202-208. doi: 10.3109/14992027.2010.524254.

[14] Sun, K., Botteldooren, D. and De Coensel, B., 2018, December. Realism and immersion in the reproduction of audio-visual recordings for urban soundscape evaluation. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 258, No. 4, pp. 3432-3441). Institute of Noise Control Engineering.

[15] Zamora, J., 2017, October. I'm sorry, Dave, I'm afraid I can't do that: Chatbot perception and expectations. In Proceedings of the 5th international conference on human agent interaction (pp. 253-260). doi: 10.1145/3125739.3125766.

[16] Kim, H.K., Park, J., Choi, Y. and Choe, M. 2018. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. Applied ergonomics, 69, pp.66-73. doi: 10.1016/j.apergo.2017.12.016.

[17] Carnegie, K. and Rhee, T., 2015. Reducing visual discomfort with HMDs using dynamic depth of field. IEEE computer graphics and applications, 35(5), pp.34-41. doi: 10.1109/MCG.2015.98.

[18] Jang, D.P., Kim, I.Y., Nam, S.W., Wiederhold, B.K., Wiederhold, M.D. and Kim, S.I., 2002. Analysis of physiological response to two virtual environments: driving and flying simulation. CyberPsychology & Behavior, 5(1), pp.11-18. doi: 10.1089/109493102753685845.

[19] Chang, E., Kim, H.T. and Yoo, B., 2020. Virtual reality sickness: a review of causes and measurements. International Journal of Human–Computer Interaction, 36(17), pp.1658-1682. doi: 10.1080/10447318.2020.1778351.

Appendix

Table A summarized all items included in the SSQ test. All symptoms were followed by a 4-point likert-scale from 0 (not at all) to 3 (very) to indicate the degree of each symptom. The three major components (Nausea, Oculomotor and Disorientation) included a different set of symptoms in the list. To calculate the three components, the sum was first calculated by adding symptom scores and then multiplied by a coefficient.

Table A: Computation of SSQ scores (adopted from [11]).

SSQ symptom[a]	Nausea	Oculomotor	Disorientation
1. General discomfort	|	|	
2. Fatigue		|	
3. Headache		|	
4. Eyestrain		|	
5. Difficulty focusing		|	|
6. increased salivation	|		
7. Sweating	|		
8. Nausea	|		|
9. Difficulty concentrating	|	|	
10. Fullness of head			|
11. Blurred vision		|	|
12. Dizzy (eye open)			|
13. Dizzy (eye closed)			|
14. Vertigo			|
15. Stomach awareness	|		
16. Burping	|		
Total[b]	[1]	[2]	[3]

Score
Nausea = [1] * 9.54
Oculomotor = [2] * 7.58
Disorientation = [3] * 13.92
Total score = ([1]+[2]+[3]) * 3.74

[a]scored (label): 0 (not at all), 1 (slightly), 2 (moderately), 3 (very);
[b]sum obtained by adding symptom scores.

Figure A- Left reported the boxplot of the SSQ profile (in original score) of the system in the test. The average score in all factors was lowest in Nausea (mean=13.43, SD=12.25), followed by Disorientation (mean=20.11, SD=27.93) and highest in Oculomotor (mean=25.27, SD=19.57). However, due to the difference between the coefficient, all factors in SSQ had a different scale. Figure A-Right further explained the correspondingly scale in each factor with markers at "not at all", "slightly", "moderately" and "very". Furthermore, the average scores for each factor were also plotted (marked as "average" in Figure A-Right).

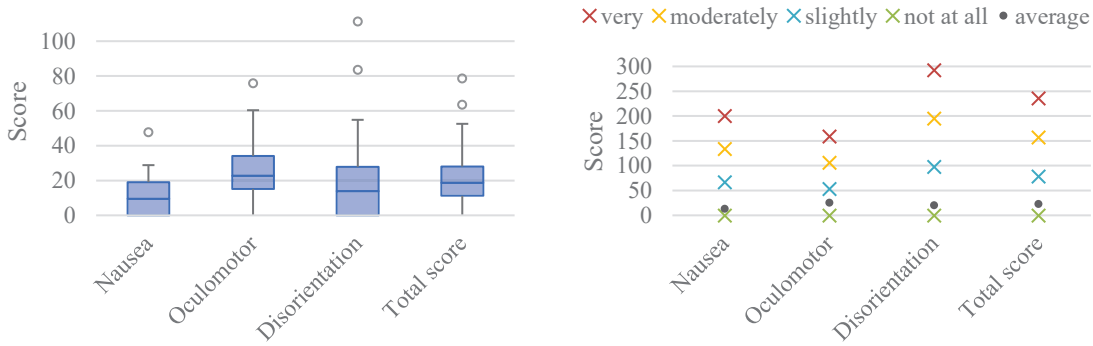


Figure A – Left: SSQ profile of the audiovisual system in the test (original score); Right: overview of SSQ subscale and corresponding verbal ticks.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Perceptual evaluation of auralizations from a wave-based method in a virtual environment

Tanmayee Pathre[1,*], **Maarten Hornikx**[1], **Armin Kohlrausch**[2]

[1]Department of the Built Environment, Eindhoven University of Technology, Eindhoven, Netherlands.

[2]Human Technology Interaction Group, Eindhoven University of Technology, Eindhoven, Netherlands.

[*]t.u.pathre@tue.nl

Abstract

In recent years, Virtual Reality (VR) has become a powerful tool in studies involving perceptual evaluation of auralization methods. The aim of this study was to investigate the extent to which binaural auralizations from a wave-based method are perceptually close to those from measurements for two different room acoustical scenarios. The signals used for auralization are simulated using the Discontinuous Galerkin method. The simulations encompass geometric details of the room, material properties of the objects in the room and receiver directivity including head orientation. The measured Binaural Room Impulse Responses (BRIRs) were limited to an upper frequency of 2.5 kHz to match the bandwidth of the simulations. The reverberation time results show a fairly good agreement between the measured and simulated data, thereby indicating that the wave-based method can be held as a reference for achieving perceptually realistic auralizations. To evaluate this, a listening experiment was carried out in an interactive VR environment employing a 3D model of the room integrating a dynamic convolution framework with headphone reproduction. The BRIRs were convolved in real-time with two types of source signal – percussion and female voice speech. Twelve normal hearing participants performed a rating task for judging differences between measured and simulated BRIRs for four perceptual attributes - reverberance, clarity, bassiness and externalization. The results indicate that the mean difference between measurement and simulation is statistically significant for reverberance, clarity and bassiness but not for externalization. Additionally, this study highlights a state-of-the-art experimental framework which can be used for perceptual evaluation studies independent of the computational method used to derive the auralizations.

Keywords: perceptual evaluation, wave-based method, auralization, virtual reality, perceptual attribute.

1 Introduction

Several computational acoustics methods are applied for room acoustic modelling and auralization purposes. Amongst them, the wave-based methods are known to accurately model low-frequencies and physical phenomena such as reflection, diffraction and scattering. Although computationally expensive at high frequencies, the time-domain wave-based methods have been successfully applied for room acoustic modelling [1], [2], [3]. The room acoustic simulation when rendered for auditory reproduction serves as an input for subjective evaluation studies. For an auralization to be realistic, the room material properties, directivity of source, geometrical details of the room and receiver directivity needs to be taken into account. Several studies on perceptual evaluation of room acoustic modelling methods and various auralization systems provide an account of the needed requirements aiming to achieve realistic auralizations [4], [5] [6]. Lately, Virtual Reality (VR) has widely been implemented in the domain of building design and building acoustics [7]. The incorporation of VR is growing in addressing questions such as accuracy of sound localization in reverberant environments [8] and sound effects of building characteristics on cognitive performance [9].

The current study implements VR in a listening test for subjective evaluation of the auralizations derived from a wave-based method. Our goal is to investigate the degree to which the auralizations come close with the measurements for the four perceptual attributes - reverberance, clarity, bassiness and externalization [10]. The choice of perceptual attributes was primarily based on the computational method used to derive the simulation and what answers it can provide to potentially improve the simulation.

This paper starts by providing an overview of the approach adopted for the measurements and wave-based simulations and the objective results [11]. Next, an elaborate account on the listening experiment design is presented. It informs the reader about the VR system integrated with a real-time convolution framework and design of the 3D user interactive environment. Finally, the results from the listening experiment are presented.

2 Background

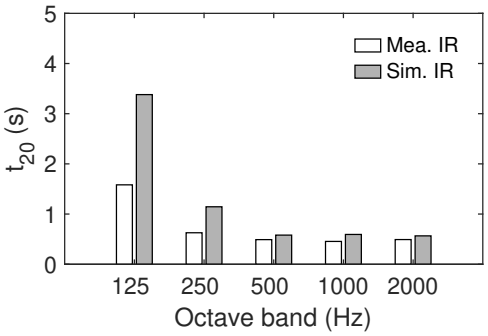
2.1. Measurements

A small room (87.63 m^3) in a laboratory was acoustically treated to achieve two different room acoustical scenarios (see Figure 2). The room received acoustical treatment with carpet and acoustic porous panels. This well-treated environment was referred to as Scenario 1. The second scenario received a fairly less amount of acoustical treatment and was referred to as Scenario 2. For each scenario, BRIR measurements were carried out for a source-receiver distance of 1.5m and 3.5m. The BRIR measurements were carried out using B&K TYPE 4125, an omnidirectional source with an exponential sine sweep of 175s and the B&K 4128 Head-and-Torso Simulator (HATS) as a receiver. The measurements were performed for 72 different head orientations in the horizontal plane with a resolution of 5°. The surface impedances of the acoustic panels and the carpet tiles and were measured both with a pressure-velocity (PU) probe and an impedance tube.

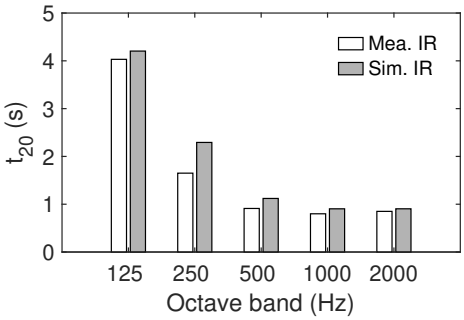
2.2. Wave-based simulations

The simulations use the frequency dependent time-domain boundary conditions using the time domain Discontinuous Galerkin (DG) method. They encompass geometric details of the room, surface impedances of the porous panel and carpet, and the reflection coefficient of the hard concrete walls. The sound source was simulated as an omnidirectional Gaussian pulse. In order to simulate the head-rotations and binaural aspects of human hearing, the simulations were rendered to be reproduced binaurally over headphones. A dual concentric sphere of receiver points was placed in the simulation domain. Spherical harmonics along with the Head Related Transfer Functions (HRTFs) from a database of measured BRIRs [12] were applied to simulate the same set of head-rotations as in the measurements [13]. For the purpose of the listening experiment, an improved set of simulations was re-computed compared to the ones mentioned in [11]. The new calculations were carried out for all four room acoustic scenarios with an upper frequency bound set to 2.5 kHz . Within this new set of simulations, the receiver grid for spherical harmonics reconstruction was corrected. The reflection coefficient of the hard concrete walls was adjusted from 0.991 to 0.996. An optimization algorithm was used for angle adjustment in the horizontal plane between source and receiver for a better match with the measurements. The reader can refer to [11] for further details on measurements and simulations and the methods adopted for their post-processing.

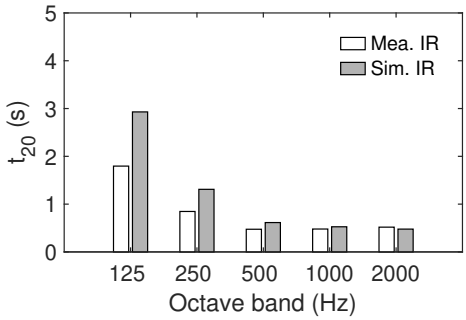
The reverberation time results are presented in Figure 1. It can be observed from Figure 1 that overall, for all four scenarios the simulations are in fairly good agreement with the measurements. A closer look at the results tells us that the simulations are in a much better agreement with the measurements for Scenario 2 as compared to Scenario 1 for both source-receiver distances 1.5m and 3.5m.



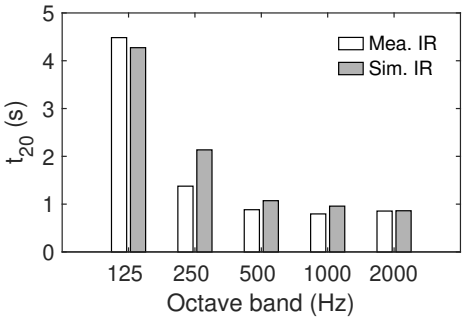
(a) Scenario 1 1.5m



(b) Scenario 2 1.5m



(c) Scenario 1 3.5m



(d) Scenario 1 3.5m

Figure 1: Reverberation Times T_{20} of IRs for the four room acoustic scenarios

3 Listening Test Experimental Design

Virtual Reality Dynamic Convolution Framework

A virtual reality integrating a dynamic convolution framework forms the cornerstone of the listening experiment setup. Unity Game Engine [14] and Max/MSP (Cycling'74) [15] are the two major building blocks of the system design. A new object, called dynamic convolution object was developed with the convolution approach presented in [16]. The object was programmed for Max/MSP to perform a real-time convolution on the CUDA. The source code for the dynamic convolution object was written in C and compiled using Visual Studio 2017. The object was implemented as a dynamic library incorporated in Max/MSP for general usage and specifically connected to the Unity Game Engine to feed the spatial parameter (head movements from the user). The dynamic convolution object takes pre-computed BRIRs for left and right channels separately to be convolved with source signals.

Unity Game Engine provides a platform for functioning of VR. When the user/s when mounts the VR Head Mounted Display (HMD), their head-movements/rotations are represented as angles within Unity Game Engine [14]. The angles (real valued numbers) are converted into bytes and sent over UDP (User Datagram Protocol) to Max/MSP. In Max/MSP, these bytes are translated into real numbers which correspond to the angles or head-orientation of the user. The Max patch is equipped with functions to evaluate the angles and match it with the associated BRIR filter number. This filter convolved with the source signal in real-time on CUDA. A change in head-orientation of the user, would update the BRIR filter number allowing convolution with a different filter. The convolved audio signal is then reproduced over the headphones to be heard by the user.

3.1. 3D Interactive Environment Design

Four 3D models of the room, each corresponding to a room acoustical scenarios were created in SketchUp [17] as shown in Figure 2.

(a) Scenario 1 1.5m

(b) Scenario 2 1.5m

(c) Scenario 1 3.5m

(d) Scenario 2 3.5m

Figure 2: 3D room models of four room acoustic scenarios created in SketchUp.

The assignment of materials and textures to various room objects was also performed in SketchUp. The 3D models of the room were imported to Unity (Version 2019.3.15f) in .fbx format. Visual rendering and offline lighting was done in Unity to achieve realistic and immersive scenarios. The 3D models of all the four room acoustic scenarios were rotated by 270°in the Unity Game Engine as SketchUp and Unity follow different co-ordinate systems. The material and textures of all the 3D room model components such as tables, chairs, wall panels, carpet were further upgraded using built-in high definition post-processing features in Unity to obtain visually high quality materials. An example of the 3D room model in the Unity Game Engine is provided in Figure 3. The experimental interface was designed using the User Interface (UI) components from Unity as shown in Figure 4.

The UI components included buttons for stimuli presentation, recording the response of the participant, loading source signals and saving the response of the participants. Four sliders each corresponding to the four perceptual attributes were implemented as a bipolar continuous rating scale with anchors -50 to +50 at either ends. However, only the scale end labels were visible to the participants as shown in Figure 4. An additional slider, with a unipolar continuous scale ranging between 0 to 100 was implemented for the overall difference question. The behaviour of UI components was controlled by means of C# scripts compiled in Visual Studio Environment.

Figure 3: 3D model of Scenario 1 1.5m as seen in Unity Game Engine. The camera in the image (indicated by ”L” in red) is the position of the listener inside the virtual room when VR HMD is mounted.

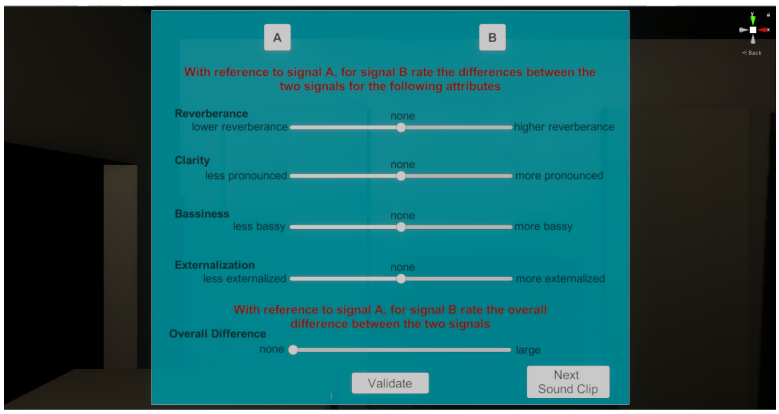


Figure 4: Experimental Interface as seen by the participants while wearing VR HMD

The C# scripts handled the following processes for Max/MSP and Unity respectively:

- Sending commands to Max/MSP for loading new sets of measured and simulated BRIRs, cleaning the GPU off the previously loaded BRIR set, randomising the order of source signals and loading a source signal from the list of source signals, playing the convolved measured and the convolved simulated signal.
- Loading a random 3D model of the room such that there was no mismatch between the BRIRs loaded into the memory and the corresponding 3D room model.

4 Materials and Method

4.1. Participants

Twelve participants (5 Female, 7 Male) took part in this study. The age of the participants ranged between 20-32 years. All participants reported that they were normal hearing listeners.

4.2. Source Signals and Audio Apparatus

Two types of source signals were used - percussion instruments and speech sentences by a female talker. Anechoic recordings of two different percussion instruments, Turkish drums, Darbuka and Bendir played with the same rhythm "dum-tek-tek" [18] and four anechoic recordings of speech sentences by a female talker [19] were used in the experiment. The duration of the signals was 3-6s. The source signals were re-sampled at 44.1 kHz. The audio playback was handled by Max/MSP. The output signal was routed through the Audient iD14 audio interface to Sennheiser HD 800 S headphones.

4.3. Procedure

The experiment was divided into four steps. Figure 5 shows the steps of listening experiment procedure.

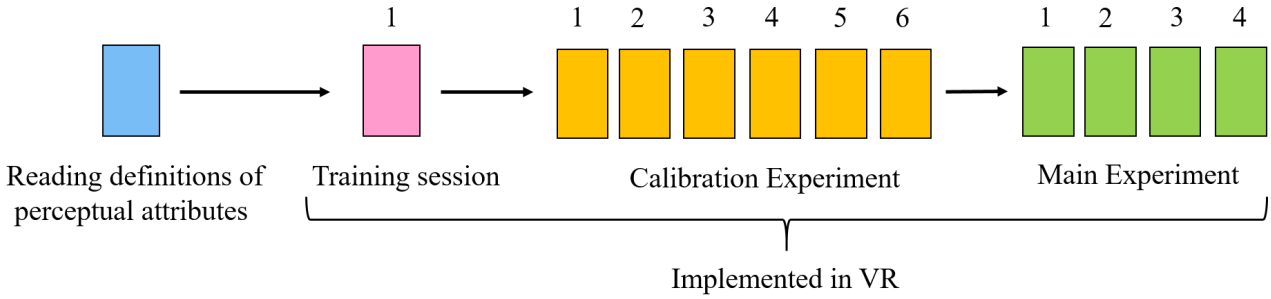


Figure 5: Graphical representation of the listening experiment procedure. The listening experiment consists of four steps. The numbers on top of the coloured blocks indicate the number of sessions in each experiment. The first step involves participants reading the definitions of the four perceptual attributes. The second step was one training session in VR. The third step was a calibration experiment composed of 6 sessions and the last experiment was the main experiment having 4 sessions.

The listening experiment was carried out in a double-walled sound insulated listening booth. The participants were first informed about the general nature of the experiment involving sound quality evaluation in Virtual Reality for four perceptual attributes - reverberance, clarity, bassiness and externalization. Before they began the experiment, they read the definitions of the four attributes. The definitions of the attributes were modified compared to mentioned in [10] to be understood by naive listeners. In case the definition of the attributes was not clear to the participants, the experimenter elaborately explained the meaning of the attributes through examples. After the participants read the definition of the attributes, a short training session was administered before the commencement of the experiment. The purpose of the training session was to familiarise the participants with the task involved in the experiment and to get them accustomed to the VR environment. Oculus Rift VR was used during this study. They first mounted the Head Mounted Display (HMD) and then the headphones. Throughout the experiment, the primary button of the Oculus Rift right hand touch controller was used by the participants to provide their response.

After the training session, a calibration experiment was conducted prior to the main experiment. The motivation for the design of calibration experiment was to understand how the participants map the meaning of perceptual attributes onto the rating scale used in the experiment. The calibration experiment was an audio only session implemented in VR. The 3D models of the rooms were eliminated. However, for the sake of brevity the data from calibration experiment has not been presented and discussed in this paper.

The experiment consisted of four sessions. Within each session, participants were presented with 6 source signals. The order of the source signals and experiment sessions was randomised throughout the experiment. Each session in the experiment represented one audio-visual condition corresponding to one of the four room acoustic scenarios as shown in the Table 1.

Table 1: Table showing the audio-visual sessions implemented in the listening experiment

Sessions	Signal A (Convolved measured signal)	Signal B (Convolved simulated signal)
1	Scenario 1 – 1.5m	Scenario 1 – 1.5m
2	Scenario 1 – 3.5m	Scenario 1 – 3.5m
3	Scenario 2 – 1.5m	Scenario 2 – 1.5m
4	Scenario 2 – 3.5m	Scenario 2 – 3.5m

The participants began the experiment by pressing 'Start Experiment' button on the interface. A 3D model of one of the four room acoustic scenarios was generated in the VR headset such that the participant was placed in the virtual room. Simultaneously, the measured and simulated BRIRs corresponding to the room acoustic scenario were loaded into the computer's memory. The process of loading BRIRs took around 110 seconds. During this time, the participants were prompted to wait by a display message on the interface until the files were loaded in the memory. They were asked to have a look at the virtual room in which they were placed or to take a break by taking off the VR HMD. Once the files were loaded, a 'start' button appeared on the interface. The participants started the session by pressing the button.

The method adopted in the listening experiment was an A/B testing method in which Signal A was the reference signal, calculated by convolving the source signal with the measured BRIRs and signal B was calculated by convolving the source signal with the simulated BRIRs as mentioned in Table 1. The participants were not informed about the nature of the signals A and B. They were instructed to provide their response for signal B with signal A as the reference. Their task was to rate the differences between the two signals A and B for the four perceptual attributes and an additional question regarding the overall difference between the two signals.

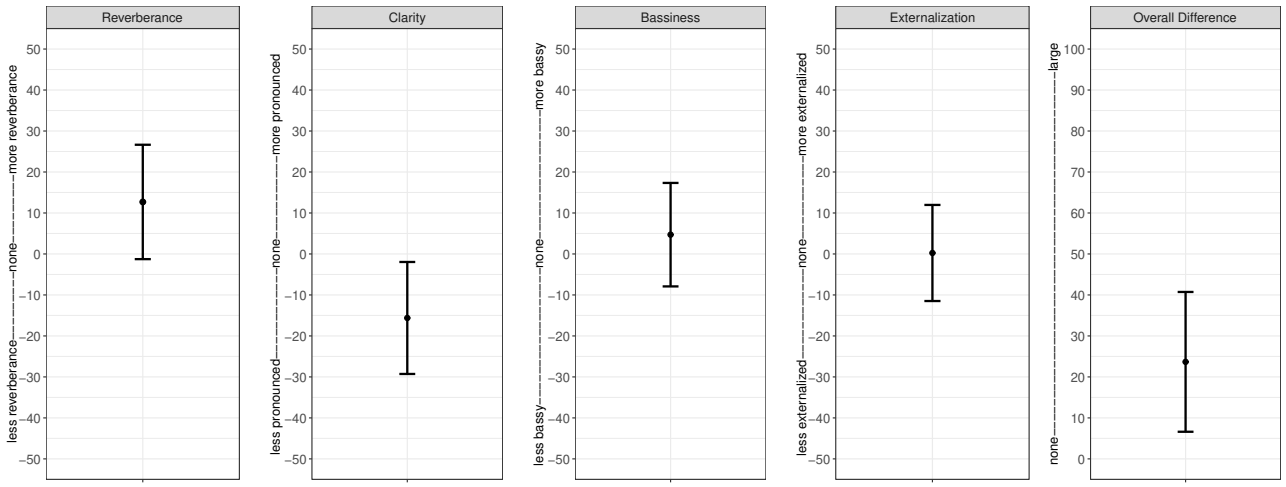
The participants listened to signals A and B as many times as they would like. They provided their response by dragging the slider for each of the attributes. The participants pressed the Validate button on the interface to save their response. After which they would press the Next Sound signal button on the interface. This button would load a different source signal to be convolved with the BRIRs.

After the completion of each session, the participants pressed the Next session button. By pressing next session, the participants were placed in another virtual room (3D model) and the BRIR set corresponding to the visual condition was loaded into the system. This process happened until the participants completed all four sessions. All the participants performed the rating task for all source signals and for all four sessions in the experiment only once. In total there were, 4 room acoustic conditions x 6 source signals = 24 observations per participant. The approximate duration of the experiment was 30 minutes.

5 Results

For each of the four perceptual attributes, a direct difference between the two signals, A (measured) and B (simulated) was obtained from the rating task. Figure 6 shows the judged difference between the measurement and simulation for each of the four perceptual attributes. Each sub-figure provides the mean response combined over all four scenarios for each perceptual attribute. The Y-axis of each sub-figure shows the scale labels for the four attributes as implemented in the listening experiment. A positive value on the Y-axis represents a higher rating for the simulation than for the measurement. It can be observed from Figure 6a that the perceived reverberance was judged higher for the simulation than for the measurement. From Figure 6b, it can be seen that the clarity in the stimuli was perceived to be less clear or less pronounced for the simulation than for the measurement. The perceived bassiness was judged higher for the simulation than for the measurement as shown in Figure 6c. For the externalization attribute based on the VR experience, the simulation was rated to be slightly more externalized than the measurement as seen in Figure 6d. Lastly, from Figure 6e the overall

difference between simulation and measurement was not perceived as large but a perceptual difference between the two stimuli was still observed.



(a) (b) (c) (d) (e)

Figure 6: Mean of the difference rating between measurements and simulations for each of the four perceptual attributes ($Mean \pm s.d.$).(a) Reverberance (13.35 ± 13.96), (b) Clarity (-15.62 ± 13.66), (c) Bassiness (4.71 ± 12.62) (d) Externalization (0.24 ± 11.72) and (e) Overall Difference (23.76 ± 17.05). Error bars indicate standard deviation.

Statistical analysis was performed using R Statistical Software. One sample Wilcoxon signed rank test, a non-parametric test, was performed as the responses for each of the attributes was found to be not normally distributed. The analysis revealed that the mean difference between measurement and simulation was found statistically significant for reverberance ($V = 27695$, $p < 0.001$), clarity ($V = 974$, $p < 0.001$), bassiness ($V = 2468$, $p < 0.001$) but not for externalization ($V = 10456$, $p = 0.7216$). The overall difference between the simulation and measurement was found to be statistically significant ($V = 40755$, $p < 0.001$).

6 Discussion

The VR integrated dynamic convolution framework explained in section 3 provides the listener with a 3D simulated audio-visual environment. The dynamic convolution framework is independent of the number of BRIR filters allowing to feed BRIRs of a desired resolution. For this study, we used 72 BRIR filter pairs (360°/5°), 5°resolution for measurements and simulations. A finer resolution between the BRIR filter pairs might prove useful in perceptual sensitivity studies.

For attribute reverberance, the perceptual judgement results shown in Figure 6a are in fairly well agreement with the objective results for T_{20} Figure 1.

For the attributes clarity, the results shown in Figure 6b are in line with the inverse relationship between reverberation time and clarity [20]. This opposite trend can also be observed in Figure 6a and Figure 6b showing that the reverberant aspect of the simulation resulted in lack of its perceived clarity.

The bassiness attribute was judged higher for the simulation than for the measurement (Figure 6c). The T_{20} results from Figure 1 for octave bands 125 Hz and 250 Hz are in well agreement with perceptual results for all room acoustic conditions except for 125 Hz band of Scenario 2 3.5m (see Figure 1d where there is more energy in the 125 Hz band for the measured IR). The inherent bassy trait of the percussion instruments, Bendir and Darbuka, is possibly responsible for more perceived bassiness in the simulation than the measurement.

Externalization rating were not found to be statistically significant from zero. This finding tells us that for one of

the perceptual attributes the simulation was perceptually perceived close to the measurement (see Figure 6d). A review study on sound externalization provides evidence that the phenomenon is influenced by various cues such as binaural cues (Interaural Time Differences (ITDs) and Interaural Level Differences (ILDs)), reverberation in the stimuli, visual cues and head-movements [21]. The measured and simulated stimuli used in this study both encompass identical source-directivity information and were reverberant in nature. They were dichotically presented to the participants in a VR environment capturing natural head movements of the listener. The fact that in both the measured and the simulated stimuli the cues required for sound externalization were preserved, may explain that the simulations were judged to be perceptually close to the measurements for the externalization attribute.

Lastly, the overall difference was found statistically significant (see Figure 6e). The overall difference refers to any perceptually noticeable differences including the four perceptual attributes that could be involved in rating the difference between the simulated vs. the measured signals.

7 Conclusions and Further Work

In this study, a listening experiment was conducted in VR for subjective evaluation of the auralizations from a wave-based method. The findings from the listening experiment show that the wave-based simulations were found perceptually indistinguishable from the measurements for the externalization attribute. The simulations and measurements were found to be perceptually distinguishable from the measurements for the attributes reverberance, clarity and bassiness, the results from the listening experiment are in agreement with the objective results of reverberation time T_{20}. Future work focuses on examining results from the calibration experiment to understand how listeners used the scale implemented in the listening experiment. The experimental design presented in this study will be extended for perceptual evaluation of auralizations using computationally efficient room acoustic modelling methods. Lastly, investigations on material properties of objects used to create the two room acoustic scenarios are still ongoing to obtain a better match between measurements and simulations.

Acknowledgements

We would like to thank fellow researchers of the Building Acoustics Group Wouter Wittebol and Baltazar Briere de la Hosseraye for providing simulation data and objective results. Alessia Milo for trouble-shooting complexities during the design of dynamic convolution framework.

References

[1] Brian Hamilton and Stefan Bilbao. FDTD methods for 3-d room acoustics simulation with high-order accuracy in space and time. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25 (11):2112–2124, 2017. doi: 10.1109/TASLP.2017.2744799.

[2] Jelle Van Mourik and Damian Murphy. Explicit higher-order fdtd schemes for 3d room acoustic simulation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(12):2003–2011, 2014.

[3] Huiqing Wang, Indra Sihar, Raúl Pagán Muñoz, and Maarten Hornikx. Room acoustics modelling in the time-domain with the nodal discontinuous galerkin method. *The Journal of the Acoustical Society of America*, 145(4):2650–2663, 2019.

[4] Matthias Blau, Armin Budnik, Mina Fallahi, Henning Steffens, Stephan D Ewert, and Steven Van de Par. Toward realistic binaural auralizations–perceptual comparison between measurement and simulation-based auralizations and the real room for a classroom scenario. *Acta Acustica*, 5:8, 2021.

[5] Fabian Brinkmann, Lukas Aspöck, David Ackermann, Steffen Lepa, Michael Vorländer, and Stefan Weinzierl. A round robin on room acoustical simulation and auralization. *The Journal of the Acoustical Society of America*, 145(4):2746–2760, 2019.

[6] Barteld NJ Postma and Brian FG Katz. Perceptive and objective evaluation of calibrated room acoustic simulation auralizations. *The Journal of the Acoustical Society of America*, 140(6):4326–4337, 2016.

[7] Finnur Kári Pind Jörgensson, Cheol-Ho Jeong, Hermes Sampedro Llopis, Kacper Kosikowski, and Jakob Strømann-Andersen. Acoustic virtual reality – methods and challenges. In *Proceedings of BNAM 2018*, 2018. Baltic-Nordic Acoustics Meeting 2018, BNAM 2018 ; Conference date: 15-04-2018 Through 18-04-2018.

[8] Hermes Sampedro Llopis, Finnur Pind, and Cheol-Ho Jeong. Development of an auditory virtual reality system based on pre-computed b-format impulse responses for building design evaluation. *Building and Environment*, 169:106553, 2020.

[9] Imran Muhammad, Michael Vorländer, and Sabine J Schlittmeier. Audio-video virtual reality environments in building acoustics: An exemplary study reproducing performance results and subjective ratings of a laboratory listening experiment. *The Journal of the Acoustical Society of America*, 146(3):EL310–EL316, 2019.

[10] Alexander Lindau, Vera Erbes, Steffen Lepa, Hans-Joachim Maempel, Fabian Brinkman, and Stefan Weinzierl. A spatial audio quality inventory (saqi). *Acta Acustica united with Acustica*, 100(5):984–994, 2014.

[11] Fotis Georgiou, Baltazar Briere de la Hosseraye, Maarten Hornikx, and Philip W. Robinson. Design and simulation of a benchmark room for room acoustic auralizations. Proceedings of the International Congress on Acoustics, pages 723–730, September 2019. doi: 10.18154/RWTH-CONV-239684.

[12] Ramona Bomhardt, Matias de la Fuente Klein, and Janina Fels. A high-resolution head-related transfer function and three-dimensional ear model database. In *Proceedings of Meetings on Acoustics 172ASA*, volume 29, page 050002. Acoustical Society of America, 2016.

[13] Maarten Hornikx. Reconstruction of binaural room impulse responses using spherical harmonics. In *23rd International Congress on Acoustics, integrating 4th EAA Euroregio 2019 (ICA2019)*, 2019.

[14] Unity game engine. https://unity3d.com/unity/whats-new/2019.3.15, 2019.

[15] Max/msp/jitter. https://cycling74.com/, 2019.

[16] Øyvind Brandtsegg, Sigurd Saue, and Victor Lazzarini. Live convolution with time-varying filters. *Applied Sciences*, 8(1):103, 2018.

[17] Sketchup pro 2020. https://www.sketchup.com/node/4446, 2020.

[18] Sound examples of percussion detection. http://users.spa.aalto.fi/ajylha/percussion/.

[19] P Demonte. Harvard speech corpus—audio recording 2019. *University of Salford Collection*, 2019.

[20] Gilbert A Soulodre and John S Bradley. Subjective evaluation of new room acoustic measures. *The Journal of the Acoustical Society of America*, 98(1):294–301, 1995.

[21] Virginia Best, Robert Baumgartner, Mathieu Lavandier, Piotr Majdak, and Norbert Kopčo. Sound externalization: A review of recent research. *Trends in Hearing*, 24:2331216520948390, 2020.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Effect of Wireless Transmission Errors on Sound Zone Performance at Low Frequencies

Christian Sejer Pedersen[1,*], Martin Bo Møller[2], Jan Østergaard[1]

[1]Department of Electronic Systems, Aalborg University, Aalborg, Denmark.

[2]Bang & Olufsen A/S, Struer, Denmark.

*cp@es.aau.dk

Abstract
Sound zones can provide different audio content to several persons in the same room by creating separate zones of individual audio content with minimal interference between the zones. Optimal strategies for creating sound zones depend on the wavelength of the sound. For low frequencies it is advantageous to distribute a number of woofers in the room and use sound field control based on measured transfer functions. For such a distribution of woofers it is desirable to stream the audio signals wirelessly to the woofers, but wireless streaming to multiple receivers can be challenging since transmission artefacts may occur. As the sound field control depends on the superposition of the sound from all the woofers, any errors in the reproduced sound will not only affect the audio quality, but also have a negative effect on the sound zone generation. This study investigates relevant types of degradation that transmission errors can have on the sound zone performance for a specific sound zone setup with eight woofers. Errors such as transmission delays and data dropouts are simulated. It is demonstrated that all these transmission artefacts lead to decreased contrast between the sound zones which is mainly caused by increased leakage to the dark zone. A mitigation strategy where the affected woofer is excluded from the sound zone control shows promising results with slightly lower contrast than the error-free system, but with much better performance than the system affected by transmission errors.

Keywords: sound zones, sound field control, wireless transmission, low frequencies

1 Introduction

Sound zones can be used to provide different audio content to several persons in the same room [1]. It uses multiple loudspeakers and sound field control to create a bright zone, where a specific audio content is wanted and a dark zone, where that specific audio content is unwanted i.e. it should preferable be inaudible. This is often practically impossible so some leakage from the bright zone to the dark zone is expected. The leakage between the zones is quantified by the contrast, which is the level difference between the bright and dark zone. By creating bright and dark zones for the different audio content the personal sound zones can be superimposed on each other.

There are different solutions for creating sound zones and the most suitable solution depend on the wavelength of the reproduced sound. For high frequencies the directivity of drivers is suggested [2], while beamforming may be suitable at mid frequencies. At low frequencies the omnidirectional nature of low-frequency drivers and low absorption at room boundaries means that a suitable method is to distribute several woofers in the room and use sound field control to minimize the mean square error between desired sound pressure and the actual sound pressure in the two zones (also referred to as pressure matching [1]). The sound field control method is based on a feedforward control scheme, that assumes that all physical characteristics remain linear and time invariant. Any violation of these assumptions can lead to deviation between the target and the actual

sound fields in the bright and the dark zone and therefore lower contrast. For example, physically moving one or more woofers would cause considerable performance loss [3] and it is important to keep the woofers in their linear amplitude range [4]. Even physical characteristics like temperature changes can potentially affect the contrast performance [5].

As the woofers are distributed in the room it is most convenient to transfer audio signals wirelessly – especially for domestic use, where cabling may give practical challenges. But low latency streaming of audio or video over a wireless network to multiple receivers can be challenging [6]. Due to multipath and Doppler effects, the channel conditions become time varying and the instantaneous transmission rates, jitter, and packet loss probabilities, can be hard to determine [7]. It is generally necessary to use buffering [8] [9] [10] and forward error control strategies [11] to ensure smooth playback at the receivers. In [12] it was furthermore shown that that a joint design of the buffer size and forward error correction codes can lead to a reduced delay in real-time streaming applications. A complementary approach to forward error correction is packet loss concealment [13] [14]. At the expense of a longer delay, interleaving techniques can be advantageous in order to deal with packet burst errors [15]. It is not the purpose of this study to simulate these complex errors, but rather to quantify the effect on the sound zones in simple scenarios of missing one woofer channel, having one woofer channel delayed, and having one woofer channel losing a packet of audio (data dropout) at different error rates. How these simple scenarios reflect real transmission errors depend very much on the system architecture and the transmission scheme and codec used.

This study uses a specific laboratory setup with eight woofers which is used in a larger research project [16] and simulate simple transmission errors to one of the woofers in order to evaluate the degradation in the performance of the sound zones. If the sound zone system detects transmission errors on a channel, then the mitigation solution of removing this channel from the sound zone generation is evaluated.

2 Method

2.1 Laboratory setup

The sound zone setup is placed in a laboratory room (dimensions 7.0 m x 8.12 m and a height of 3.0 m) at Aalborg University. The room has heavily damped walls and ceiling and a wooden floor with a carpet (reverberation time T20 ranging from 0.6 s at 50 Hz to 0.2 s at 400 Hz). The sound zone setup consists of eight woofers distributed in the room and a soundbar for the mid to high frequencies placed in front of the sound zones at a distance of 2.911 m as shown in Figure 1. The two sound zones are at a left (Zone A) and right (Zone B) seat of a sofa.

The woofers are 10’’ woofers in custom made MDF cabinets and are designed to be used for frequencies below 500 Hz enabling different cross-over frequencies depending on the setup. The connections to each woofer are cabled, but for future implementations closer to a commercial system the transmission is expected to be using a wireless system of transferring the audio to the woofers.

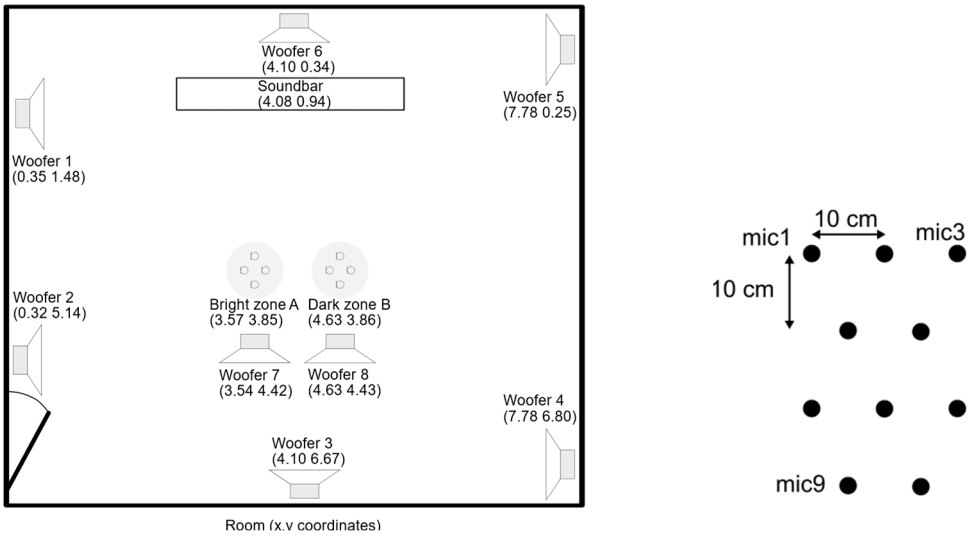


Figure 1: Diagram of the room seen from above. The two zones A and B and the eight woofers 1-8 are shown. The center coordinates in meters are shown in brackets. Right figure shows microphone positions for the transfer function measurements seen from above. Microphone 1-3 are facing the sound bar and two heights (1.051 m and 1.164) are used.

2.2 Sound zone control method

The sound zone control method used in this study is a time domain method that attempts to match a reference target in the bright and the dark zone while minimizing the mean square error between the reference target and the actual sound [17]. Additionally, the method shapes the envelope of the filters in order to reduce pre- and post-ringing as this is expected to produce better audio quality [18]. It is specifically the method proposed by [18] that is used in this study.

The block diagram including sound zone control filters, w_k, wireless transmission, t_k, transfer functions between woofers and microphone positions, h_k, and their resulting pressures in the microphone positions in the bright zone, p_b, and dark zone, p_d, can be seen in Figure 2. For the laboratory setup the impulse responses of woofer 7 with a modelling delay, $w_k^{(eq)}$, is the reference target function for Zone A due to its close proximity to this zone. The reference target for the dark zone is zeroes i.e., no sound is wanted in the dark zone.

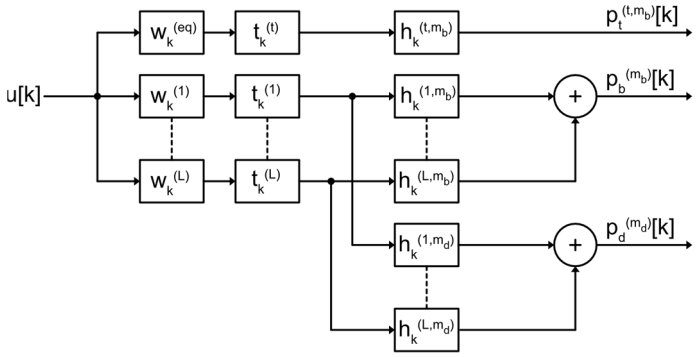


Figure 2: Block diagram of the signal flow from input u[k] to target pressure in the bright zone $p_t^{(t,m_b)}[k]$, reproduced bright zone pressure $p_b^{(m_b)}[k]$ and dark zone pressure $p_d^{(m_d)}[k]$. w_k are the sound zone filters, h_k are the transfer functions from each woofer to each microphone positions in the bright zone, m_d, and microphone positions in the dark zone m_d. The diagram is based on [19] with the addition of the wireless transmission blocks t_k for each woofer 1 to L.

The frequency range of the sound field control focus on the frequencies from 35 to 500 Hz, which is the range of the woofers. The sampling frequency is chosen to 1200 Hz in order to have the most efficient setup for the used frequency range and an integer relation to the 48 kHz used for the playback system and a FIR filter length of 100 taps is chosen to have adequate performance while getting efficient real time processing for the sound zone playback system.

2.3 Woofer transfer function measurements

Transfer functions are measured from each woofer to each 20 microphone positions in the bright and dark zones using the logarithmic swept sine technique [20]. The microphones positions shown in right Figure 1 are ten positions in two different heights of 1.051 m and 1.164 m. This particular layout was chosen in order to have a reasonable approximation of the continuous sound field below 450 Hz over a range of seated ear positions in the zones. The measurements are performed at 48 kHz sampling rate and down-sampled to 1.2 kHz in order to reduce the required processing power both with regard to filter calculation and the actual processing of the audio playback. The complete set of transfer functions are measured twice – one set of for sound zone filter generation and another set for evaluating the performance of the sound field control in order to prevent what is referred to as the “inverse crime” [18].

2.4 Simulation of transmission errors

The transmission errors are simulated both directly with the transfer functions, but also with audio signals processed through the sound zone filters for preparing the audio streams for each woofer and then introducing the errors in t_k on blocks of audio streams to the 20 microphones positions in the bright and dark zone using the scheme shown in Figure 2. The source signal for the plots is gaussian white noise, while for informal listening the source signal is a mono signal from different music signals with bass content.

2.4.1 Simulation of one woofer not playing

The scenario of having one woofer not playing i.e. if it is not turned on or the wireless connection to it is lost is simulated by removing the contribution from that woofer in the summation of the transfer functions (see Figure 2) by effectively setting the impulse response for the affected channel to all microphone positions to zero. For time-domain testing with audio signals the audio stream for the affected channel is omitted in the superposition at the microphone positions.

2.4.2 Simulation of delay in one woofer signal

The scenario of having a synchronization issue resulting in a delay on one woofer is simulated by shifting the contribution from that woofer by integer samples in the summation of the transfer functions (see Figure 2). As the sampling frequency in the simulations is 1200 Hz each sample shift corresponds to a delay of approx. 0.83 ms. For time-domain testing with audio signals the audio stream for the affected channel is delayed by integer samples in t_k .

2.4.3 Simulation of random packet loss on one woofer channel

The scenario of packet loss of audio on one woofer channel is done by preparing the source signal streams and then by random (from uniform probability distribution, independent from previous and future packets) introducing an empty block (packet loss) resulting in data dropout at a predefined error rate in the blocks of streams for the selected woofer channel. For the chosen sampling frequency of 1200 Hz, the block size has been chosen to 24 samples which correspond to 20 ms of signal and simulations of 5%, 10% and 15% packet loss are evaluated.

2.5 Mitigation by removing affected woofer channel if error is detected

If the transmission to a woofer is affected by errors, and there is some way of detecting this, then it is possible to mitigating the error. If a stable delay is detected, then the sound zone system should mitigate this by getting all woofer signals in sync, and the effect of the error will disappear. But packet loss or a missing woofer needs a more drastic solution of removing the affected woofer from the sound zone generation. This requires a new set of control filters based on the seven other woofers. It is expected that this will reduce the performance as compared to an error-free eight woofer solution as there are fewer woofers to control the sound, but it is expected that this solution will perform better than no mitigation.

2.6 Method of evaluating of sound zone performance and transmission error effects

The evaluation is based on the method described in [21] and the performance metrics used in this study is contrast and normalized mean square error as defined in the following.

For the time domain the contrast is following the notation of [19] and is defined as:

$$Contrast[k] = 10log_{10}\left(\frac{M_b^{-1}\sum_{m=1}^{M_b}\left|{p_b}^{(M_b)}[k]\right|^2}{M_d^{-1}\sum_{m=1}^{M_d}|{p_d}^{(M_d)}[k]|^2}\right), \quad (1)$$

where ${p_b}^{(M_b)}[k]$ is the sound pressure at the M_b microphone control points in the bright zone at time sample k and ${p_d}^{(M_d)}[k]$ is the sound pressure at the M_d microphone control points in the dark zone at time sample k. In the frequency domain the contrast is following the same notation but with the time sample k replaced by the angular frequency ω.

The normalized mean square error follows the notation from [19]:

$$nmse[k] = \left(\frac{\|p_b[k] - p_t[k]\|_2^2}{K^{-1}\sum_{k=1}^{K}\|p_t[k]\|_2^2}\right) \quad (2)$$

where $p_t[k]$ is the target pressure in the bright zone and K is the number of time-steps used to evaluate the normalized mean square error.

In addition to the quantitative results also informal listening was performed on one center microphone signal in the bright and dark zone in order to assess the qualitative effects of transmission errors on the sound quality.

3 Results

3.1 Evaluation of one woofer missing

The effect of a missing woofer in the sound zone generation can be seen in left Figure 3. This shows that the loss in contrast is depending much on which woofer is affected and especially woofer 7 and 8, which are close to the sound zones, are important for maintaining contrast between the zones.

It can be seen by comparing right top and bottom plots in Figure 3 that the loss in contrast is mainly due to the increased level in the dark zone as the level in the bright zone is much less affected by the missing woofer 7. This illustrates how sensitive the destructive interference in the dark zone is to violations of the linear time invariance assumptions.

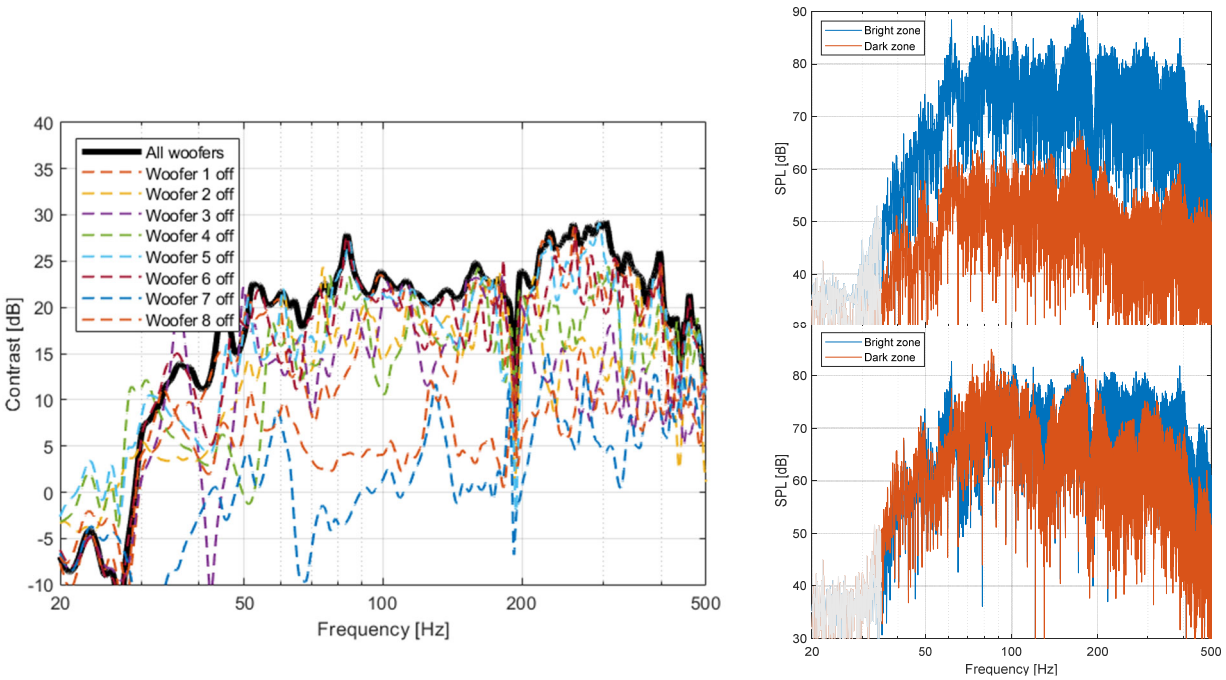


Figure 3: Left: frequency domain mean contrast (Equation (1)) between Zone A and Zone B when all woofers are playing (no error) and the effect of individual woofers not playing. Right: Frequency response of a 10 s gaussian white noise signal in the bright and dark zone; top figure is when there is no error and bottom figure is when woofer 7 is not playing.

3.2 Evaluation of one woofer channel delayed

The effect of having one woofer channel delayed at different integer sample delays for either woofer 5 (best case scenario) and woofer 7 (worst case scenario) can be seen in Figure 4. As expected, the degradation in contrast performance increases with increased delay and as can be expected from the results in Figure 3 it also depends on which woofer is affected.

Left Figure 5 shows the mean time domain contrast depending on delay and affected woofer. In general, the mean contrast drops from zero to two samples delay and is then more or less stable with larger delays, but depends on which woofer is affected.

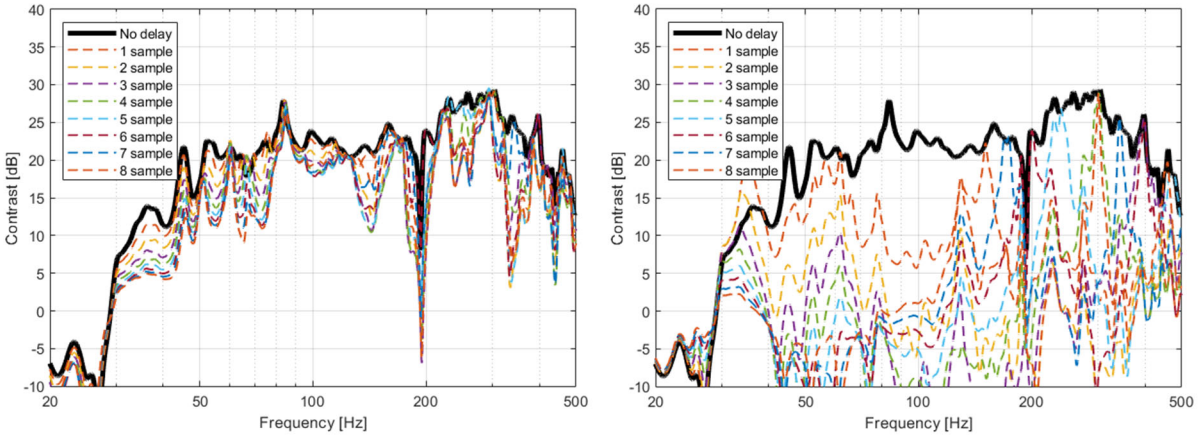


Figure 4: Left is the effect of delay of woofer 5 (far from the sound zones), and right is the effect of delay of woofer 7 on the frequency domain contrast (Equation (1)).

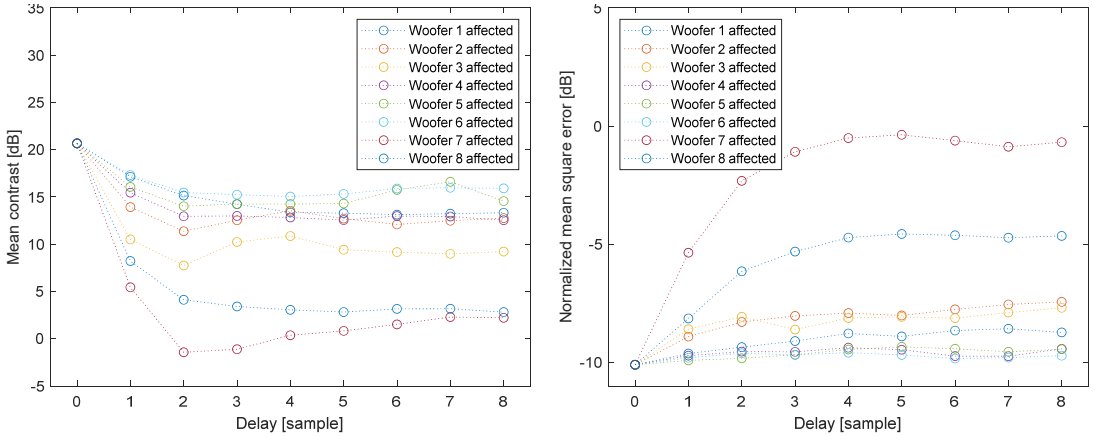


Figure 5: Left shows mean time domain contrast (Equation (1)) and right shows normalized mean square error (Equation (2)) depending on integer sample delay on specific woofer signal (10 s gaussian white noise at 1200 Hz sampling frequency).

The degradation in normalized mean square error between the reference target audio and the resulting audio in different situations of a delayed woofer signal can be seen in right Figure 5. This shows that the normalized mean square error increases with increased delay until approx. four samples delay. As expected, delay in woofer 7 affect the performance in the bright zone the most, as it is closest to the zone and furthermore, it is the transfer functions for this woofer that is used as the desired reference in the sound field control.

3.3 Evaluation of packet loss at different error rates based on gaussian white noise signal

Mean time domain contrast under different situation can be seen in left Figure 6. It is seen that increased packet loss results in decreased mean contrast. The degradation in normalized mean square error between the reference target audio and the resulting audio in different situations can be seen in right Figure 6. It is clear that especially woofer 7 and to some degree woofer 8 are important for reproducing the desired signal in the bright zone. If other woofers are affected the error is minimal.

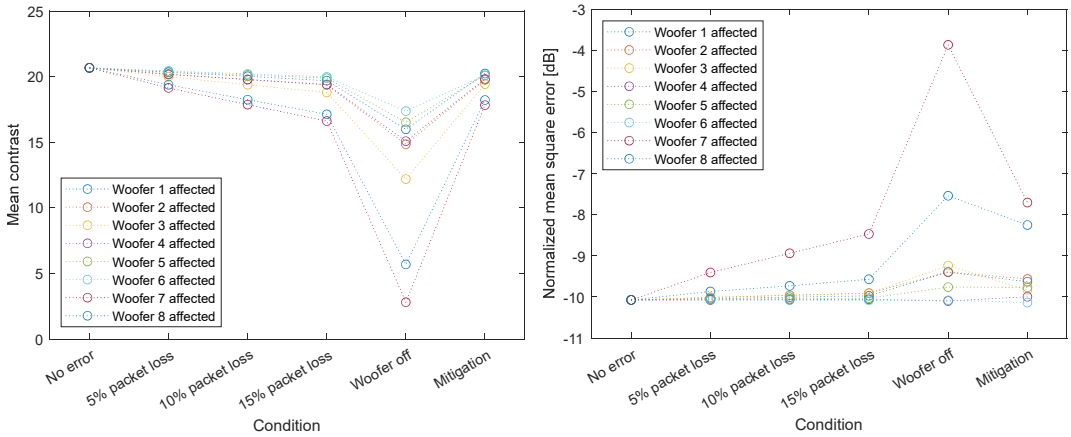


Figure 6: Left shows mean time domain contrast (Equation (1)) and right shows time domain normalized mean square error (Equation (2)) for 10 s gaussian white noise depending on error and affected woofer.

3.4 Evaluation of mitigation based on omitting affected woofer

Sound zone control filters based on the seven remaining woofers (as described in section 2.5) are calculated for each combination and their frequency response in the bright zone the contrast performance compared to

the eight woofers situation can be seen in Figure 7 and the mean time domain contrast and normalized mean square error is also shown in Figure 6. As expected, the overall contrast is a little lower, but it depends very much on which woofer is off. Woofer 7 and 8 are the most important woofers in the setup due to their close proximity to the sound zones, and missing one of these affects the performance the most. As the reference target for the bright zone is the response of woofer 7 it is evident that the deviation in the frequency response in the target zone is also highest if woofer 7 is missing.

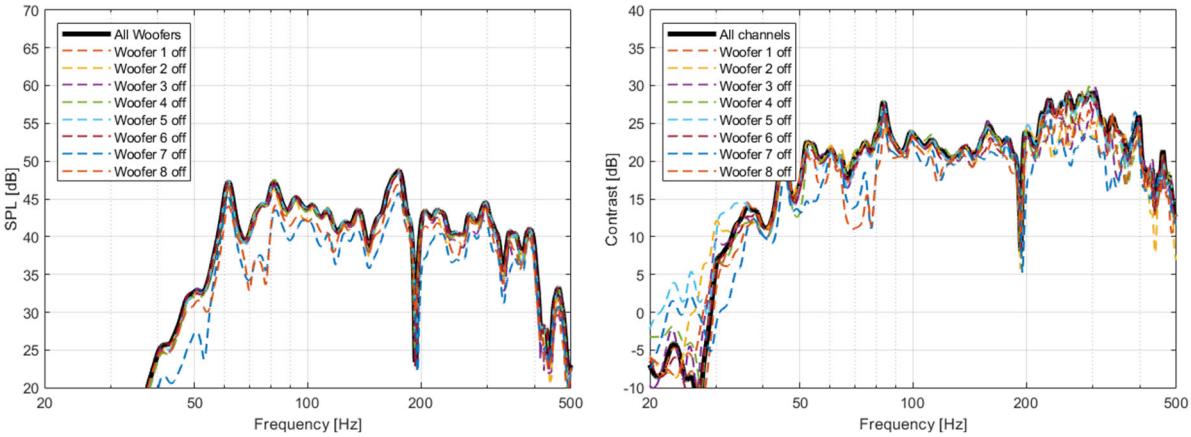


Figure 7: Left is the frequency response in bright zone and right is the frequency domain contrast (Equation (1)) for the mitigation by removing the affected woofer and using new filters.

3.5 Informal listening to simulated sound in bright and dark zone

From informal listening to the simulated sound in both the bright and the dark zone it is clear that the effect of having a woofer off gives increased leakage to the dark zone as the level is much higher. Packet loss gives audible transients, so it is expected to be more "distracting" than what can be expected by just looking at the mean contrast (Figure 6). Here the mitigation solution provides an audio experience similar to the "no error" scenario, but with a slight increased leakage in the dark zone.

4 Discussion

The transmission error manifest itself especially as noticeable leakage to the dark zone as destructive interference is very sensitive to deviations from the "perfect" signal. The reduction in contrast due to transmissions errors is frequency dependent due to the fact that the different woofers are contributing differently at different frequency ranges, i.e. one woofer may be more important in a specific frequency range. Therefore, the contrast performance of the sound zones is depends very much on which woofer is affected as could be expected from the investigated layout of the woofer placement. The woofers closer to the sound zones are more important for the sound zone performance. This opens the question if another reference target could give better performance. Possibly a reference target with a smoother frequency characteristic would give a more "transparent" sound reproduction inside the target zone.

The mitigation solution requires that the transmission error is detected, but the mitigation is effective in maintaining acceptable performance of the sound zones. For a specific setup the performance of the mitigation depends on how many woofers that are running error-free as the performance gets lower with the lower number of woofers available for the sound zone generation. In order to take any situation into account it is necessary to calculate sound zone control filters for all possible permutations of working woofers i.e., there could be any number and combination of woofers with transmission errors.

The setup investigated here consists of eight woofers, but for setups with less woofers, e.g. four woofers it is expected that the performance degradation due to transmission errors will be more severe as there are fewer woofers contributing appropriately to the sound zone generation. As a consequence, it is expected that a mitigation solution with only three woofers will show a larger performance degradation as compared to this study. Accordingly, setups with more than eight woofers will probably be less affected than what is seen here and the mitigation is expected to have less performance penalty.

As the transmission errors mainly causes increased leakage to the dark zone, the main consequence is that person(s) in the sound zones will be more distracted from the audio material from the other zone [22]. How distracting the error is in an actual setup with two sound zones with independent audio material depends very much on the audio material as this may provide masking if it is in the same frequency range [22]. But it is also seen that the distraction model proposed by [22] fails to predict the distraction when the music characteristics cannot be well described by the energy-based features used in the model. Informal listening revealed that especially random packet loss errors are audible due to transients. This opens up for future research in the audibility of transmission errors and mitigation solutions.

5 Conclusion

For feedforward sound zone systems any violation of the assumed linear time invariance will affect the performance. Missing the sound from a woofer or getting errors in the sound from one woofer gives significant increases in leakage to the dark zone seen as decreased contrast ranging from 3.3 to 17.9 dB depending on the affected woofer . This increased leakage can effectively be mitigated by removing this channel from the sound zone control system and by using another set of control filters where this woofer is excluded from the filter calculation. The mitigation only decreases the contrast 0.5 to 2.9 dB depending on the woofer. However, this requires some kind of error checking system in order to identify which woofer is affected and by which type of errors and a large set of precalculated filters in order to effectively deal with the possible combinations of errors. The results shown in this study are for a specific setup, but similar effects are expected to be general also for other sound zone setups based on distributed woofers.

Acknowledgements

This work is partly funded by the Innovation Fund Denmark (IFD) under File No 9069-00038B in the project: Interactive Sound Zones for Better Living (ISOBEL) [16].

References

[1] T. Betlehem, W. Zhang, M. A. Poletti and T. D. Abhayapala, "Personal Sound Zones: Delivering interface-free audio to multiple listeners," *IEEE SIGNAL PROCESSING MAGAZINE,* pp. 81-91, March 2015.

[2] W. F. Druyvesteyn and J. Garas, "Personal Sound," *J. Audio Eng. Soc,* vol. 45, no. 9, pp. 685-701, 1997.

[3] P. Coleman, P. J. B. Jackson, M. Olik, M. Møller, M. Olsen and M. A. Pedersen, "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array," *J. Acoust. Soc. Am.,* vol. 135, no. 4, p. 1929–1940, April 2014.

[4] X. Ma, P. J. Hegarty, J. A. Pedersen and J. J. Larsen, "Impact of loudspeaker nonlinear distortion on personal sound zones," *J. Acoust. Soc. Am.,* vol. 143, no. 1, p. 51–59, January 2018.

[5] M. Olsen and M. B. Møller, "Sound zones: on the effect of ambient temperature variations in feed-forward systems," in *AES 142th Convention*, Berlin, Germany, 2017.

[6] S. i. Hahm, P. Kang, H. Bang and H. j. Yeon, "Dynamic media buffer control scheme for seamless streaming in wireless local area networks," in *IEEE Wireless Communications and Networking Conference Workshops*, 2016.

[7] Rappaport, T. S., Wireless communications: principles and practice, vol. 2, New Jersey: Prentice Hall PTR, 1996.

[8] T. Kim and M. H. Ammar, "Receiver buffer requirement for video streaming over TCP," in *Electronic Imaging 2006*, San Jose, California, United States, 2006.

[9] K. J. Ma, M. L. A. Huang and R. Bartos, "Video rate adaptation in mobile devices via http progressive download of stitched media files," *Communications Letters, IEEE,* vol. 15, no. 3, p. 320–322, 2011.

[10] T. Kim, N. Avadhanam and S. Subramanian, "Dimensioning receiver buffer requirement for unidirectional vbr video streaming over TCP," in *Image Processing, 2006 IEEE International Conference on*, 2006.

[11] J. Hansen, J. Østergaard, J. Kudahl and J. H. Madsen, "Sequential Use of Block Codes and Convolutional Codes in a Real-Time Multi-Hop Network," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, 2018.

[12] J. H. Sørensen, P. Popovski and J. Østergaard, "Delay Minimization in Real-Time Communications With Joint Buffering and Coding," *IEEE Communications Letters,* vol. 21, no. 1, pp. 52-55, Jan 2017.

[13] C. Perkins, O. Hodson and V. Hardman, "A survey of packet loss recovery techniques for streaming audio," *IEEE Netw,* vol. 12, no. 5, pp. 40-48, 1998.

[14] X. Lu, H. He and H. Tan, "A low complexity packet loss recovery method for audio transmission," in *Proc. 2nd Int. Conf. on Computer Science and Electronics Engineering (ICCSEE 2013)*, 2013.

[15] A. Botta and A. Pescapé, "IP packet interleaving for UDP Bursty losses," *J. Syst. Softw.,* vol. 109, p. 177–191, 2015.

[16] ISOBEL, "Interactive Sound Zones for Better Living," 2022. [Online]. Available: https://isobel.dk/.

[17] M. F. S. Gálvez, S. J. Elliott and J. Cheer, "Time Domain Optimization of Filters Used in a Loudspeaker Array for Personal Audio," *IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING,* vol. 23, no. 11, pp. 1869-1878, 2015.

[18] M. B. Møller and M. Olsen, "Sound zones: on envelope shaping of fir filters," in *ISCV24*, London, 2017.

[19] M. B. Møller and J. Østergaard, "A Moving Horizon Framework for Sound Zones," *IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING,* vol. 28, pp. 256-265, 2020.

[20] S. Müller and P. Massarani, "Transfer-Function Measurement with Sweeps," *Journal of Audio Engineering Society,* vol. 49, no. 6, pp. 443-471, 2001.

[21] M. B. Møller and M. Olsen, "Sound Zones: On Performance Prediction of Contrast Control Methods," in *Proc. Audio Engineering Society Conference: 2016 AES International Conference on Sound Field Control*, Guildford, United Kingdom, 2016.

[22] J. Francombe, R. Mason, M. Dewhirst and S. Bech, "A Model of Distraction in an Audio-on-Audio Interference Situation with Music Program Material," *J. Audio Eng. Soc.,* vol. 63, no. 1/2, pp. 63-77, January/February 2015.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

The Effect of Fixed-point Arithmetic on Low Frequency Sound Zone Control

Peter Koch[1,*], **Jan Østergaard**[1]

[1]Department of Electronic Systems, Aalborg University, Aalborg, Denmark.

[*]pk@es.aau.dk

Abstract

By use of several loudspeakers it is possible to create constructive and destructive acoustic interference leading to spatial sound zones with high (bright zone) and low (dark zone) sound pressure levels, respectively. The loudspeaker control signals are commonly created by filtering the audio signals using Finite Impulse Response (FIR) control filters. We are interested in using finite precision multiply-accumulate arithmetic in the FIR filter operations in order to reduce the complexity of a potential hardware implementation. In a simulation study based on a setup with 8 loudspeakers and for the frequency range 50 – 600 Hz, we demonstrate that using only 8 and 12 bits signed fixed-point arithmetic for the multiplier and adder, respectively, reduces the mean acoustic contrast ratio between the bright and the dark zones by only 0.1 dB as compared to when using 32 bits arithmetic. Reducing the word length of the adder below 12, significantly increases the sound pressure in the dark zone.

Keywords: sound zones, fixed point arithmetic, FIR filtering, variable word length simulation.

1 Introduction

Control of personal or individual sound zones refers to a specific problem within sound field control, where one is interested in generating individual listening zones for separate listeners in some enclosed space, for example, a room [1] or a car cabin, [2]. It is common to consider the control of two sound zones in a room, and where the sound field outside the zones is not controlled, [3]. In one of the zones, one would like to reproduce a specific sound signal with a high quality and at a high sound pressure level. This zone is usually referred to as the *bright* zone. In the other sound zone, one would like to reduce the sound pressure level as much as possible so that it is harder to hear the audio signal. This zone is usually referred to as the *dark* zone. By combining a pair of bright and dark zones, it is then possible to achieve different audio content in different spatial locations in the same room.

To control the low frequencies in the sound zones, it is advantageous to make use of knowledge of the acoustical properties of the enclosed space – such as the room transfer functions (RTFs), [4]. For the mid-frequencies, conventional beam forming techniques are often sufficient, whereas for the highest frequencies, the directivity of the loudspeakers can be exploited, [4].

The design of sound zones is currently an active area of research, see for example, [5, 6, 7, 8, 5, 9, 10]. To take into account the RTFs of the enclosed space, it is often necessary to use long FIR filters having several hundreds taps. Moreover, to achieve a high degree of acoustic separation (contrast) between the zones, it is generally necessary to use several loudspeakers. Since the individual loudspeakers requires different filtered signals, the complexity in terms of hardware requirements, execution time, and energy consumption of the filter operations can be quite significant. To reduce this complexity it is possible to use finite word length arithmetic for the implementations of the FIR filter operations, [11].

The aim of this paper is to demonstrate in a simulation study the effect that finite precision arithmetic has on the resulting acoustic separation between the bright and the dark zone as well as on the audio quality in the bright zone. We will use a conventional FIR filtering approach to control the bright as well as the dark zone. We will be using fixed-point implementations of the multiplier and the adder used for the FIR filter operations, and we will focus on the low frequency region from 50 – 600 Hz. The RTFs used in the study to evaluate the performance are real measurements of a room equipped with 8 loudspeakers. Our experiments reveal that in particular it is the suppression of sound in the dark zone that is destroyed when the accuracy of the arithmetic operations becomes too low. We also observe that a high degree of acoustic contrast and sound quality (in a squared error sense) is possible using only 8 bits for the multiplier and 12 bits for the adder in the FIR filtering operations.

2 Background on FIR filtering based sound zones

Let us assume the availability of a set of $L > 0$ loudspeakers, which are arbitrarily distributed within a room. We adopt the notation and setup from [10], and assume that the sound pressure is known in M_b and M_d points in the bright and dark zones, respectively. For example, this knowledge can be obtained by placing microphones in the sound zones of the room. The sound pressure $p_b^{(m)}$ at the m-th point in the bright region is given by the combined output of all L loudspeakers convolved with the their room impulse responses. Specifically, let $u[k]$ be the k-th sample of the single audio signal to be filtered and played out by the loudspeakers. Moreover, let $\bar{w}^{(l)} \in \mathbb{R}^{N_w}, \ell = 1, \ldots, L$, be the impulse response of the control filter for the ℓ-th loudspeaker, where N_w denotes the length in samples of the impulse response. We limit our attention to FIR filters, and N_w is therefore finite. Finally, let $\bar{h}_b^{(m,\ell)} \in \mathbb{R}^{N_h}$ be the room impulse response (RIR) between the ℓ-th loudspeaker and the m-th point in the bright zone, and where we limit the response to N_h samples. We define $\bar{h}_d^{(m,\ell)}$ in a similar manner for the dark zone. Using this notation, we can express the sound pressures $p_b^{(m)}[k]$ and $p_d^{(m)}[k]$ at time k as follows, [10]:

$$p_b^{(m)}[k] = \sum_{\ell=1}^{L} \left(\bar{h}_b^{(m,\ell)} \star \bar{w}^{(\ell)} \star u \right)[k] = \sum_{\ell=1}^{L} \sum_{j=0}^{N_w-1} \sum_{i=0}^{N_h-1} \bar{h}_b^{(m,\ell)}[i] \bar{w}^{(\ell)}[j] u[k-i-j], \tag{1}$$

$$p_d^{(m)}[k] = \sum_{\ell=1}^{L} \left(\bar{h}_d^{(m,\ell)} \star \bar{w}^{(\ell)} \star u \right)[k] = \sum_{\ell=1}^{L} \sum_{j=0}^{N_w-1} \sum_{i=0}^{N_h-1} \bar{h}_d^{(m,\ell)}[i] \bar{w}^{(\ell)}[j] u[k-i-j], \tag{2}$$

where $\star$ denotes linear convolution. We assume the filters and the RIRs to be time invariant.

The average accumulated squared sound pressure level (or sound *energy*) P_{bright} and P_{dark} in the bright and dark zones, respectively, are then given by:

$$P_{\text{bright}} \triangleq \frac{1}{N_u M_b} \sum_{m=1}^{M_b} \sum_{k=0}^{N_u-1} |p_b^{(m)}[k]|^2, \quad P_{\text{dark}} \triangleq \frac{1}{N_u M_d} \sum_{m=1}^{M_d} \sum_{k=0}^{N_u-1} |p_d^{(m)}[k]|^2, \tag{3}$$

where N_u denotes the length of the time-domain audio signal $\{u[k]\}$. A similar notation applies for the energy P_{dark} in the dark zone. We are now in a position to introduce the mean acoustic contrast ratio (expressed in dB) and the normalized mean squared error (MSE), which are defined as follows:

$$C = 10 \log_{10} \left(\frac{P_{\text{bright}}}{P_{\text{dark}}} \right) \text{[dB]}, \quad Q = \frac{\sum_{m=1}^{M_b} \sum_{k=0}^{N_u-1} |p_b^{(m)}[k] - \tilde{p}_b^{(m)}[k]|^2}{\sum_{m=1}^{M_b} \sum_{k=0}^{N_u-1} |\tilde{p}_b^{(m)}[k]|^2}. \tag{4}$$

where $\tilde{p}_b^{(m)}[k]$ is a specific desired target pressure level at time k at the m-th position in the bright zone. We will be using the contrast ratio and the normalized MSE to quantify the performance of the system, when we

are changing the precision of the arithmetic operations involving the convolutions between the control filters and the audio signals. The contrast ratio quantifies the amount of "separation" between the two zones, and the normalized MSE quantifies the quality of the resulting audio signal in the bright zone.

3 Finite word length FIR filtering in sound zones

The convolutions described by Equations (1) and (2) can be separated into two stages. In the first stage, the audio signal $\{u[k]\}$ is convolved with the control filters $\{\bar{w}^{(\ell)}\}$, and in the second stage the *filtered* audio signal is played out and will thereby be convolved with the impulse responses $\{\bar{h}_b^{(m,\ell)}\}$ and $\{\bar{h}_d^{(m,\ell)}\}$ of the room. In this work, we will focus on the first stage, where we will modify the filter coefficients $\{\bar{w}^{(\ell)}\}$ as well as the convolution operator $\star$ in order to model the effect of using finite precision arithmetic. The filtering operations in the second stage are not affected by the modifications in the first stage.

3.1. Quantization of filter coefficients

Operating in a finite word length environment where signed arithmetic operations are needed, several potential number representations may be considered, ranging from the simple "signed magnitude" notation to advanced representations such as redundant binary number systems. The number representation being used highly impacts the implementation cost as well as the execution time of the overall application. For instance, the advantage of a redundant number system is that it can perform addition in a constant time independent of the word length, the drawback being a significant hardware overhead demanded by multiple bits per digit as well as input/output converters needed for interfacing against a traditional binary number representation.

Since our primary aim in this work is to investigate the numerical robustness of sound zones operated in a finite word length environment, our experiments will be conducted using a d-bit 2's complement fixed-point number representation which easily handles signed arithmetic operations in the dynamic range $[-1; 1[$, but which at the same time may be far from optimal in terms of implementation cost and execution time. We will address specific hardware implementation issues in a future work.

Preparing the control filters for fixed-point execution, we initially quantize the N_w coefficients of each of the L filters. To do this, we scale all coefficients by the same factor so that all coefficients are within $[-1, 1]$. Specifically, let $\bar{w}^{(\ell)} \in \mathbb{R}^{N_w}$ denote the impulse response of the ℓ-th filter. Then let $c_\ell = \max_{\ell,j} |\bar{w}^{(\ell)}[j]|$, where $\bar{w}^{(\ell)}[j]$ denotes the j-th element of the vector $\bar{w}^{(\ell)}$, and $|\cdot|$ denotes the absolute value operator. We form the normalized filters $\bar{w}^{(\ell)} c_\ell^{-1}$, and next quantize these filters to word length $d > 0$ using the following operations:

$$\hat{w}^{(\ell)}[j] = \left\lfloor \frac{\bar{w}^{(\ell)}[j]}{c_\ell} 2^{d-1}(1-\epsilon) \right\rfloor 2^{-(d-1)} \quad j = 0, \ldots, N_w - 1, \tag{5}$$

for $\ell = 1, \ldots, L$, where $\lfloor \cdot \rfloor$ denotes rounding towards the nearest integer from below, e.g., $\lfloor -0.6 \rfloor = -1$, and $\lfloor 0.6 \rfloor = 0$. Multiplying by $(1-\epsilon)$, where $0 < \epsilon \ll 1$ is a small positive constant, guarantees that all the filter coefficients are in the range $[-2^{d-1}, 2^{d-1} - 1]$ before being quantized to nearest integer. For the special case where $d = 0$, we simply replace the filters by a unit impulse, i.e., $\hat{w}^{(\ell)} = [1, 0, \ldots, 0]^T, \forall \ell$. Thus, in this case we do not control the sound field in the sound zones but simply play out the audio without any filtering taking place (except the convolutions of the audio with the RIRs).

3.2. 2's complement fixed-point arithmetic

In order to conduct the control-filter computations, i.e., multiply-accumulate operations reflecting as accurately as possible a real-time hardware execution, we initially design bit-true 2's complement Matlab-based multiplier and adder simulation models. We therefore briefly introduce the underlying mathematical fixed-point operations applied in these models which are next implemented with appropriate input/output converters such that their

input operands and the resulting product/sum can be represented as floating point numbers, [12].

Radix-4 multiplication Given two d-bit numbers X and Y being the multiplicand and the multiplier, respectively. Expressing Y as a 2's complement number, which represents the individual filter coefficients $\hat{w}^{(\ell)}[j]$, we use a notation where the Most Significant Bit (MSB) is indexed as 0. Normally, the MSB is indexed $d-1$, but in this FIR filter context where we scale the input signal (X) and the coefficients (Y) to the dynamic range $[-1;1[$, the 0-indexing of MSB is a notation which conveniently is used to express Y in binary notation, and thus the product P as:

$$Y = -y_0 + \sum_{j=1}^{d-1} y_j \cdot 2^{-j}, \quad P = Y \cdot X = -y_0 \cdot X + \sum_{j=1}^{d-1} (y_j \cdot X) \cdot 2^{-j}, \tag{6}$$

where the fixed point is located immediately to the right of the sign bit y_0.

From Equation (6) we see that the product consist of d partial products which are individually left-shifted and added, starting from the Least Significant Bit (LSB) end. In order for this to work, appropriate sign-extension has to be enforced prior to the addition. Since X and Y both have format Q1.d-1, the product is format Q2.2d-2 which due to two identical sign bits (in case of no overflow) is easily adjusted into format Q1.2d-1 by a logical left shift. Now, using the 2's complement notation and the assumption that d is an even number (identical arguments can be derived for d odd), the multiplier is parted into two sums represented by the odd and the even indices, respectively;

$$Y = -y_0 + \sum_{j=1,odd}^{d-1} y_j \cdot 2^{-j} + \sum_{j=2,even}^{d-2} y_j \cdot 2^{-j} \tag{7}$$

Adding and subtracting the "even indexed sum" on the right-hand side of Equation (7), it can be rewritten as

$$Y = -y_0 + \sum_{j=1,odd}^{d-1} y_j \cdot 2^{-j} + \sum_{j=2,even}^{d-2} y_j \cdot 2^{-j+1} - 2 \cdot \sum_{j=2,even}^{d-2} y_j \cdot 2^{-j-1}. \tag{8}$$

If Y is appended at the LSB-end with a bit y_d, which is identically equal to 0 and therefore does not alter the numerical value of Y, the two "even indexed sums" can now be reformulated in terms of two identical but "odd indexed sums";

$$\sum_{j=2,even}^{d-2} y_j \cdot 2^{-j+1} = \sum_{j=1,odd}^{d-1} y_{j+1} \cdot 2^{-j} \quad \text{and} \quad -2 \cdot \sum_{j=2,even}^{d-2} y_j \cdot 2^{-j-1} = -2 \cdot \sum_{j=1,odd}^{d-1} y_{j-1} \cdot 2^{-j} + y_0. \tag{9}$$

The multiplier Y, and hence the product P can therefore be written as:

$$Y = \sum_{j=1,odd}^{d-1} (y_j + y_{j+1} - 2 \cdot y_{j-1}) \cdot 2^{-j}, \quad P = \sum_{j=1,odd}^{d-1} (z_j \cdot X) \cdot 2^{-j} \tag{10}$$

where

$$z_j = y_j + y_{j+1} - 2 \cdot y_{j-1}; z_j \in \{0, \pm 1, \pm 2\}. \tag{11}$$

From Equations (10) and (11) it is concluded that P is the sum of $d/2$ left-shifted and sign-extended partial products (PP) which can take on the values $\{0, \pm X, \pm 2X\}$ depending on the bit pattern of three consecutive bits of the multiplier Y, starting with y_0 at the MSB-end.

In the general case, where d can be even or odd, we obtain:

$$Y = \sum_{j=0}^{\lceil \frac{d}{2} \rceil - 1} (y_{2j+1} + y_{2j+2} - 2 \cdot y_{2j}) \cdot 2^{-(2j+1)}, \quad P = \frac{1}{2} \cdot \sum_{j=0}^{\lceil \frac{d}{2} \rceil - 1} X \cdot z_j \cdot 4^{-j}. \tag{12}$$

From Equation (12) it is easily noticed that the PPs are individually shifted two bit positions against each other, i.e., Radix 4, and similarly that the final product is obtained only after a 1-bit right shift of the sum of the $\lceil \frac{d}{2} \rceil$ PPs which are all represented as sign-extended 2's complement numbers. In our model, the PPs are calculated sequentially (despite that a parallel computation is possible in a dedicated hardware configuration) and next converted into a floating point representation before they are consecutively added. Since the PPs initially are represented as fixed-point numbers, conducting the additions using floating point arithmetic significantly simplifies the simulation model but does not alter the accuracy, i.e., the resulting floating point product is generated with an accuracy equivalent to a 2's complement number in Q1.2d-1 format. The word length d is an adjustable parameter in our model enabling experimentation with varying accuracy of the calculated products.

2's complement addition One of the most unique features of 2's complement numbers, as compared to other more straight forward number representations like for instance the signed magnitude representation, is the possibility to perform signed addition (and thus also subtraction) directly on the two input operands. Due to this, contrary to multiplication, there is no need to distinguish between algorithms for signed and unsigned addition.

Consequently, a traditional d-bit Ripple Carry Adder (RCA) or a Carry Look-ahead Adder (CLA), eventually extended with overflow detection, can be opted for in a signed signal processing context like the one addressed in this work. While both of these two adder concepts perform a numerically exact computation (for a given word length d), the CLA introduces a mechanism which in a parallel manner pre-calculates the carry at several selected bit positions throughout the total word length, thus reducing the worst case propagation delay, i.e., the overall execution time, as compared to an RCA of same word length. The overhead however, being a significantly higher complexity as well as an irregular circuit layout.

Since in this work we want to prepare for the least complex hardware topology, the obvious choice for a 2's complement adder simulation model therefore is the d-bit RCA which generates the sum of two Q1.d-1 numbers by performing bit-wise iterative addition of the operands, starting from the LSB-end. At the same time, the RCA performs addition of carry information from the lower consecutive bit position, using a 3-2 adder compressor at each bit position. In this work, the carry into the LSB is defined identically equal to 0.

Our RCA-based simulation model takes as input two floating point operands which in the FIR filter context are the individual filter tap products and the accumulated sum, respectively. Using floating point number representation for the inputs makes it easy to interface against the products generated by the multiplier. Furthermore, from a numerical perspective the product accumulation can be conducted in any precision, i.e., single-, double- or overflow precision, since internally our model implements a d-bit RCA where the word length is an adjustable parameter which allows experimentation with varying addition accuracy.

For all FIR control filters we conduct appropriate numerical scaling of the input signal in order to avoid an overflow condition at the output variable, and similarly we perform online check for any overflow internally in the structure. For this reason no guard bits are needed in our adder model. The adder thus produces a sum which initially is derived in Q1.d-1 format, and next converted to and presented at the output as a floating point number with an identical accuracy.

4 Simulation study

In this section, we consider an experimental setup having $L = 8$ loudspeakers and two sound zones, a bright and a dark zone. We let the audio signal $\{u[k]\}$ be a low-frequency 50 – 600 Hz band-limited white Gaussian noise signal sampled at 1.2 kHz. The audio signal is filtered by the L FIR control filters $\{\bar{w}^{(\ell)}\}, \ell = 1, \ldots, L$, before being played out. To design the control filters, we use the design method presented in [13]. We use real measured RIRs $\{\bar{h}_b^{(m,\ell)}\}$ and $\{\bar{h}_d^{(m,\ell)}\}$ for the design of the control filters and when simulating the resulting performance. The room was of size 7.00x8.12x3.00 meters. Each control filter has length $N_w = 100$. An example of one of the filter impulse responses is shown in Figure (1) (left), and similarly an example of one the

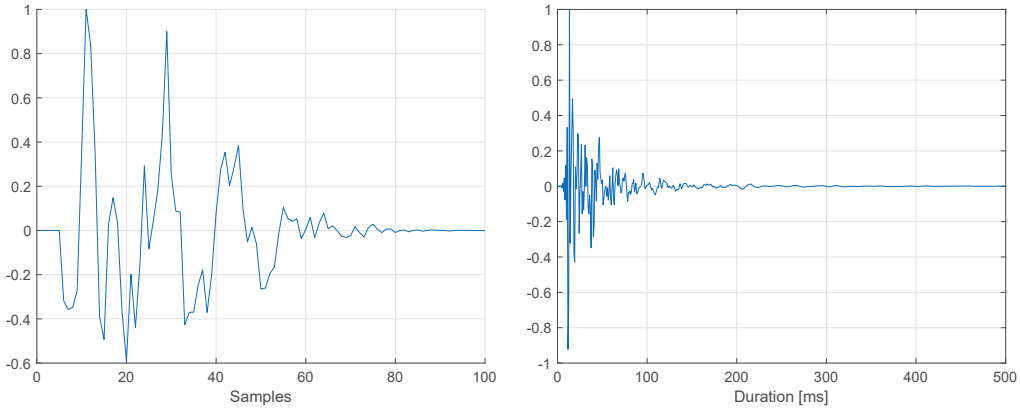


Figure 1: Left: Example of one of the control filters. Right: Example of one of the RIRs used when evaluating the resulting performance of the sound zones.

Table 1: Contrast ratio in dB as a function of the word length of the adder, and for fixed 8 and 16 bits multipliers. The contrast ratio when using 32 bits adder and multiplier is 23.09 dB.

Multiplier / Adder	6 bits	8 bits	10 bits	12 bits	14 bits
16 bits	15.08 dB	20.56 dB	22.72 dB	23.06 dB	23.09 dB
8 bits	14.95 dB	20.67 dB	22.72 dB	22.98 dB	22.98 dB

RIRs is shown in Figure (1) (right).

For reference purposes, we initially calculate i) the normalized MSE, and ii) the contrast ratio using a 32 bit word length, both for the multiplier and for the adder in the control filters. We obtain reference values equal to -9.89 dB and 23.09 dB, respectively, for the two performance metrics.

We next demonstrate the effect of replacing the 32-bit arithmetic operations by shorter word length multiplications and additions. Figure (2) (left) shows the normalized MSE as a function of the adder word length for different multiplier word lengths, and similarly Figure (2) (right) shows the corresponding acoustic contrast ratios. Table 1 shows the contrast ratio as a function of the adder word length, and for a fixed 8 and 16 bits multiplier.

The results shown in Figure (2) represents an average performance covering all frequencies. To better illustrate the impact of finite-precision arithmetic on the resulting sound zones, we therefore introduce Figures (3) and (4) which show the resulting Power Spectral Densities (PSD) for the bright and dark zones, respectively.

These experiments illustrate several interesting performance features of sound zones operated in a reduced numerical accuracy environment. First and foremost we observe a significant difference among the bright and the dark zone. The bright zone is mostly unaffected by a reduction in the multiplication accuracy given a fixed reference adder word length, Figure (3) (right), the exception being the top-most 50 Hz of the frequency range where a 4-bit multiplier word length leads to an approximately 15 dB degradation of the sound field as compared to longer multiplier word lengths.

Maintaining the reference multiplier word length for a reduced adder accuracy, we observe a somewhat more sensitive bright zone in the upper 100 Hz of the investigated frequency range when the adder word length is decreased below 10 bits, Figure (3) (left). Particularly, despite that the adder word length seems to have no or very little impact in the center part of the frequency band, an up to 20 dB degradation is discovered at the band edges when applying a 4-bit adder.

For both of the above discussed situations we explain the increased sensitivity in the upper part of the frequency band by the following considerations. All control filters are implemented as ordinary transversal filters, i.e.,

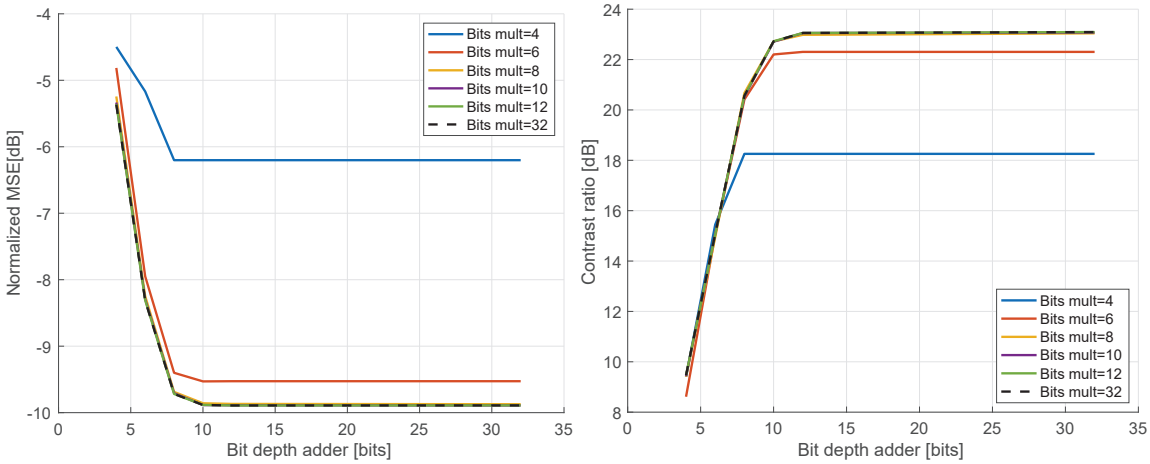


Figure 2: Left: Normalized MSE as a function of the adder word length for given multiplier word lengths. Right: Acoustic Contrast Ratio as a function of the adder word length for given multiplier word lengths.

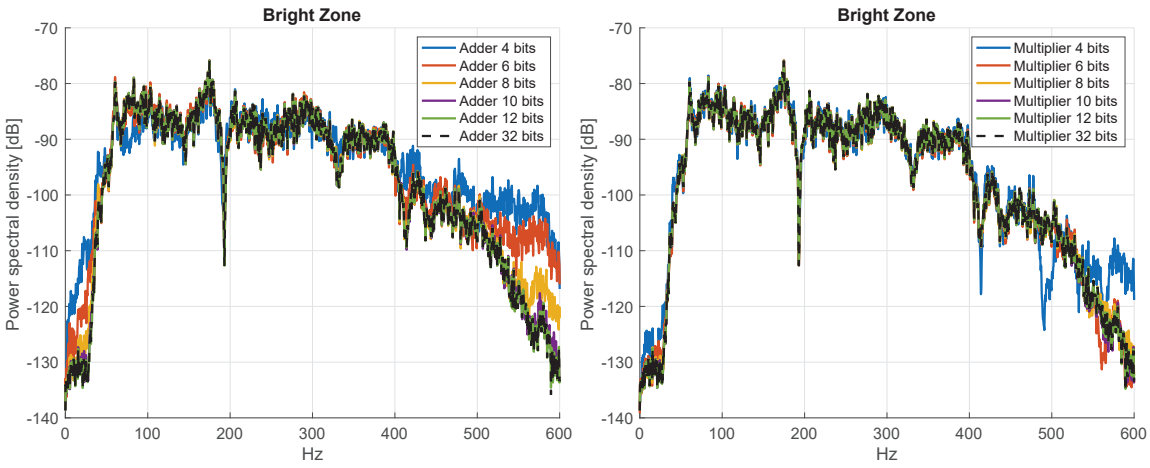


Figure 3: Power spectral density of the bright zone. Left: 32-bit reference multiplier and shorter word length adder. Right: 32 bit reference adder and shorter word length multiplier.

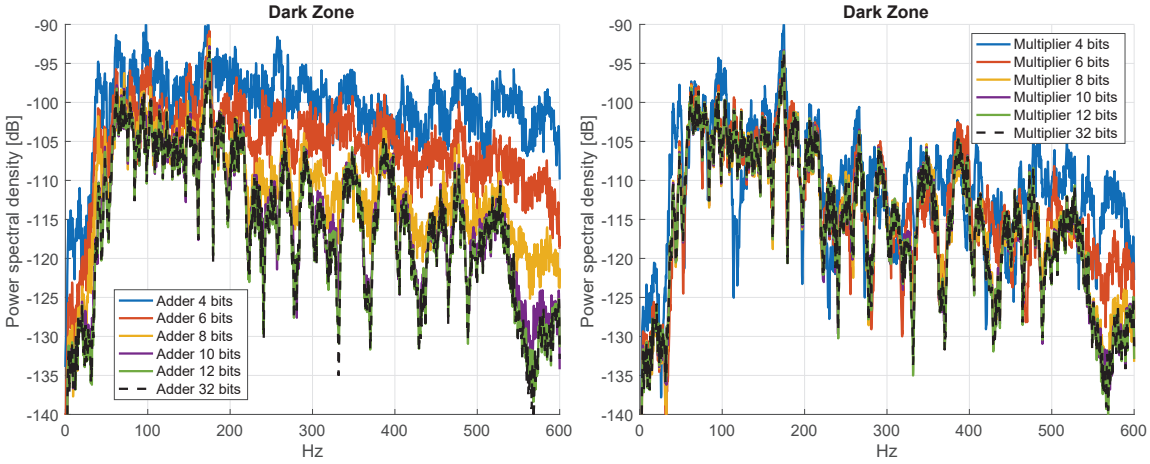


Figure 4: Power spectral density of the dark zone. Left: 32-bit reference multiplier and shorter word length adder. Right: 32 bit reference adder and shorter word length multiplier.

the tapped delay line is storing the input samples $u[k-j]$, $j = 0, ..., N_w - 1$, represented as 2's complement numbers, all scaled to comply with the dynamic range $[-1; 1[$. If the frequency of the input signal is increased (for a fixed sample frequency), the probability for changing the MSB part of the word, which is a representation of the sign, is similarly increased among consecutive input samples. Consequently, for a bit j in the MSB part of the word, the temporal correlation among consecutive samples, given as

$$\rho_j \propto E[u_j[k]u_j[k-1]] \tag{13}$$

therefore decreases, which indicates an increased transition probability, [14].

After all N_w samples in the delay line are multiplied with their corresponding constant filter coefficients, the individual products from the FIR taps similarly show decreased temporal correlation for the bits at the MSB end when the input signal frequency becomes higher. When the word length in the adder chain is reduced, a proportionally larger part of the operand bits are used for sign information, and thus fewer bits are available to represent fractional accuracy. In a "high frequency" scenario, where the sign information becomes increasingly more important, the accuracy of the product summation therefore suffers from proportionally fewer fractional bits, thus impacting the overall numerical quality of the sum. This set of arguments also explains why the sound zone application is less sensitive towards a reduction in the multiplier word length as compared to the adder word length.

Concerning the dark zone, we generally found a much more pronounced dependency of the arithmetic accuracy. We explain this mainly by the fact that creating a dark zone is significantly more complicated as compared to generating a bright zone. In the bright zone, the constructive interference increases the sound pressure level by a certain factor depending upon the number of sources. On the other hand, in the dark zone, ideally there is zero sound pressure due to a complete cancellation of all the direct and reflected sound paths from the L loudspeakers. This is of course not possible in practice and our simulations also illustrate that the sound pressure level is not zero in the dark zone. For the dark zone scenario we therefore also discover exactly the same phenomenons in the "high frequency" regions as discussed above for the bright zone. Moreover, there are several other important observations associated with the dark zone.

Consider first the fixed reference adder scenario, Figure (4) (right). Here we observe an interesting behaviour, namely a somewhat inconsistent PSD relation between the reference multiplier and the shorter word length multipliers. Normally, one would think that a word length reduction leads to a consistent performance degradation, but for several frequencies in the mid-range region we see the opposite effect, particularly for 6- and 4-bit multipliers. One possible explanation might be that the white Gaussian noise input signal, due to its random nature combined with a word length dependent accuracy reduction in the FIR filter tap products, which next are to be added, enables an enhanced "self cancellation" effect of the direct and reflected sound contributions at some arbitrary frequencies – in several cases up to 15 dB better than the reference. Overall however, the PSD obtained by the 32-bit reference and the lower word length multipliers agrees, although with a generally higher fluctuation throughout the entire frequency range as compared to the bright zone scenario.

Additionally, the experiments indicate that significant degradation, in comparison to the 32-bit reference multiplier, only occurs for a multiplier word length less than 8 bit which is therefore considered as the lower bound word length for the multiplier. This observation complies with result also illustrated in Figure (2).

Finally, for the fixed reference multiplier scenario, Figure (4) (left), we observe similar behaviours as already discussed. One exception though being the pronounced full-range performance degradation discovered for a continuous adder word length reduction. Not only does the PSD degrade with up to 30 dB in the high end of the frequency band, but throughout the entire frequency band the degradation, except for a very narrow band around 180 Hz, never falls below 10 dB (using a 4-bit adder). Similarly, we observe a very distinct, almost linear performance degradation when the adder word length is decreased in 2-bits steps. Deviation from the 32-bit reference adder becomes pronounced for a 10-bit and lower word length adder, and similarly more pronounced for increased frequency. These findings clearly indicate that the dark zone is specifically sensitive towards the adder word length which we explain by the same set of arguments as introduced for the bright zone with fixed reference multiplier.

5 Conclusions

In order to prepare recently developed Sound Zone Control algorithms for implementation in a real-time fixed-point reconfigurable hardware environment with the possibility for individual word length selection of the arithmetic units – and thus potential power-, time-, and area savings – we investigate the sound zone performance as related to the necessary multiplier- and adder word lengths. Using 2's complement based Ripple Carry Addition and Radix-4 Multiplication, we demonstrate in terms of i) Normalized Mean Square Error, ii) Acoustic Contrast Ratio, and iii) Power Spectral Density, in the bright as well as in the dark sound zone, that the multiplier- and adder word lengths can be reduced to 8 bit and 12 bit, respectively, when compared against a 32-bit reference word length.

For these specific word lengths, we conclude that it is possible to maintain a reduction in the ratio of sound pressure levels between the bright and the dark zones of at most 0.1 dB which is considered sufficiently small in order not to disturb the overall sound perception individually in the two zones.

Furthermore, our studies have clearly demonstrated that the dark zone is difficult to construct when being operated in a reduced word length scenario. Maintaining a destructive interference, in a given acoustic/physical environment, requires a certain amount of arithmetic operations, i.e., a necessary order of the control filters being executed with a sufficient numerical accuracy. We demonstrate that in particular a reduced adder word length impacts the possibility to maintain a high fidelity dark zone, which on the other hand is significantly less sensitive to modification of the multiplier word length.

The constructive interference to be established in the bright zone is also significantly less sensitive to word length minimization, although we for this zone conclude that a reduction of the adder word length has a significantly more negative impact on the overall performance as compared to a similar reduction in the multiplier word length.

In terms of the frequency related performance, we have demonstrated that the dark zone performance is impacted essentially in the complete 50 – 600 Hz frequency band which has been subject for our investigation. Although the bright zone performance is also frequency dependent for varying word lengths, we conclude that this can be observed only to a less extent as related to PSD degradation and frequency range. Despite that a shortened adder word length reduces the bright zone performance, we found that this occurs for very limited bandwidths, primarily in the upper part of the frequency band.

Our work has shaded some previously unknown light onto the way Sound Zone Control is influenced by and potentially could be operated optimally in a reconfigurable hardware fixed-point environment. Our results are therefore of particular importance when it comes to practical realization of this emerging audio technology. Despite our many new discoveries, there are still various unsolved problems and questions needed to be answered. In our future work we therefore focus on topics related to the mapping of the control filters onto a real-time hardware platforms. In particular, we will address how essential design metrics such as execution time and power consumption potentially can be reduced when the fixed-point multipliers and adders are replaced by arithmetic units which perform their calculation approximately, [15]. Our work has shown that the sound zone application allows substantial word length reductions, as compared to a 32-bit reference, and therefore an interesting study is to investigate how arithmetic units, which introduce a certain amount of approximation for a given word length, can be operated in sound zones in order to minimize time- and power overhead in a real-time dedicated hardware architecture.

References

[1] Stephen J. Elliott, Jordan Cheer, Harry Murfet, and Keith R. Holland. Minimally radiating sources for personal audio. *The Journal of the Acoustical Society of America*, 128, 2010. ISSN 0001-4966. doi: 10.1121/1.3479758.

[2] Jordan Cheer, Stephen J. Elliott, and Marcos F.Simón Gálvez. Design and implementation of a car cabin personal audio system. *AES: Journal of the Audio Engineering Society*, 61, 2013. ISSN 15494950.

[3] Joung-Woo Choi and Yang-Hann Kim. Generation of an acoustically bright zone with an illuminated region using multiple sources. *The Journal of the Acoustical Society of America*, 111, 2002. ISSN 0001-4966. doi: 10.1121/1.1456926.

[4] W.F Druvesteyn and J. Garas. Personal sound. *Journal of Audio Engineering*, 45, 1997.

[5] A. Canclini, D. Markovic, M. Schneider, F. Antonacci, E. A.P. Habets, A. Walther, and A. Sarti. A weighted least squares beam shaping technique for sound field control. volume 2018-April, 2018. doi: 10.1109/ICASSP.2018.8461292.

[6] Michael Buerger, Christian Hofmann, Cornelius Frankenbach, and Walter Kellermann. Multizone sound reproduction in reverberant environments using an iterative least-squares filter design method with a spatiotemporal weighting function. volume 2017-October, 2017. doi: 10.1109/WASPAA.2017.8170005.

[7] Yefeng Cai, Ming Wu, Li Liu, and Jun Yang. Time-domain acoustic contrast control design with response differential constraint in personal audio systems. *The Journal of the Acoustical Society of America*, 135, 2014. ISSN 0001-4966. doi: 10.1121/1.4874236.

[8] Yefeng Cai, Ming Wu, and Jun Yang. Design of a time-domain acoustic contrast control for broadband input signals in personal audio systems. 2013. doi: 10.1109/ICASSP.2013.6637665.

[9] Daan H.M. Schellekens, Martin B. Moller, and Martin Olsen. Time domain acoustic contrast control implementation of sound zones for low-frequency input signals. volume 2016-May, 2016. doi: 10.1109/ICASSP.2016.7471698.

[10] Martin Bo Møller and Jan Østergaard. A moving horizon framework for sound zones. *IEEE/ACM Transactions on Audio Speech and Language Processing*, 28, 2020. ISSN 23299304. doi: 10.1109/TASLP.2019.2951995.

[11] M. Aktan, A. Yurdakul, and G. Dundar. An algorithm for the design of low-power hardware-efficient fir filters. *IEEE Trans. Circuits Syst.*, 55:1536 – 1545, 2014.

[12] Behrooz Parhami. *Computer Arithmetic, Algorithms and Hardware Design*. Oxford University Press, 2000.

[13] Martin Bo Møller and Martin Olsen. Sound zones: On envelope shaping of fir filters. *24th International Congress on Sound and Vibration, ICSV 2017*, 2017.

[14] S. Ramprasad, N. R. Shanbhag, and N. Hajj. Analytical estimation of signal transition activity from word-level statistics. *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, 16:718 – 733, 1997.

[15] H. Jiang, F. J. H. Santiago, H. Mo, L. Liu, and J. Han. Approximate arithmetic circuits: A survey, characterization, and recent applications. *Proceedings of The IEEE*, 108:2108 – 2135, 2020.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

A comparison of audio-to-tactile conversion algorithms for melody recognition

Razvan Paisa[1,*], Jesper Andersen[2], Niels Christian Nilsson[1], Stefania Serafin[1]

[1]Multisensory Experience Lab, Aalborg University, Copenhagen, Denmark

[2]The Royal Danish Academy of Music, Copenhagen, Denmark

[*]rpa@create.aau.dk

Abstract

Besides language, music is one of the two major acoustic channels for expression of human nature and it is ubiquitous to all cultures. Due to a tight correlation between auditory and haptic stimuli, more and more attention is focused on the importance of the latter sensation in a musical context [1]. For the hearing impaired especially, tactile feedback has been investigated extensively for it's musical applications and hearing assistive devices, as early as 1983 [2]. This study compares three common audio-to-haptic signal processing algorithms designed for full range vibrotactile transducers used for tactile augmentation of music. The focus is on melody discrimination over three instruments: double bass, digital subtractive synthesizer with a sawtooth oscillator and trumpet. The transducer used is a high fidelity Tactuator BM1C, enclosed in a custom anatomical handheld case, inspired by an orthopedic *resting hand splint*. An evaluation was conducted on 34 participants and used a within-group design with three alternative forced choice task assessing the participants ability to match melodies to tactile stimuli. The results indicate that no algorithm performs better than others, which is in line with the literature regarding the overall poor frequency discrimination of the skin. Nevertheless, post experiment interviews suggest that some participants perceived multiple frequencies simultaneously, on different areas of their hand, similar to auditory polyphony.

Keywords: Vibrotactile music, Vibrotactile discrimination, Vibrotactile display, Hearing impaired music, Tactile music perception

1 Introduction

In recent years haptic feedback has received increasing attention from the sound and music community, mainly because of the strong connection between the auditory and haptic experiences. This has given birth to *musical haptics* field of research [1]. The mechanism linking auditory and tactile sensations is called multisensory integration, pioneered by Barry Stein and Alex Meredith. It describes how humans form coherent, valid, and robust perception of reality, by processing sensory stimuli from various modalities [3]. The classical rules for multisensory integration demand that enhancement occurs only for stimuli that are temporally coincident and propose that enhancement is strongest for those stimuli that individually are least effective[3]. For this integration to occur, the input from various sensors must eventually converge on the same neurons. In the specific case of auditory-somatosensory stimuli, recent studies demonstrate that multisensory integration can in fact occur at very early stages of cognition, resulting in supra-additive integration of touch and hearing [4, 5]. This translates to a robust synergy between the two sensory apparatuses, than can be exploited to synthesize experiences impossible to achieve by unisensory means. Furthermore, research within auditory-tactile interactions has shown that tactile stimulus can influence auditory stimulus and vice-versa [6, 7, 8]. It can therefore be observed that

auditory and haptic stimuli are capable of modifying or altering the perception of each other when presented in unison [9].

This study is the first in a project that has as long term goal to help partially impaired hearing individuals and cochlear implant users to appreciate music. With that in mind, the aim of this particular study is to compare three signal processing methods that convert full spectrum music into vibrotactile haptic feedback suitable for the proprieties of skin receptors, namely the Meissner's corpuscles and Pacinian ones, while preserving the melodic information encoded in the original signal. The three processing methods were chosen from existing literature [10, 11, 12, 13, 14, 15]. The experiment revolved around a handheld device designed to be comfortable to hold for longer periods of time, and capable of reproducing full spectrum audio signal. The hand was identified as the most sensitive body region for touch, due to a very high density of receptors[16, 17].

This paper describes the device built for the study followed by a detailed presentation of each signal processing technique used to convert music to vibrotactile stimuli. Subsequently, it is presented the experimental study evaluating user performance when tasked to match the haptic stimuli to a coherent auditory one. The aim of the study was twofold: (1) to evaluate the three signal processing methods in terms of their ability to convey the melodic structure existing in the original signal.(2) To evaluate the proposed hardware in terms of it's ergonomics and ease of use, as well as it's ability to produce a satisfying haptic experience. Specifically it was considered relevant to determine if a satisfactory experience can be elicited with a single, high-fidelity actuator.

2 Background

Live concerts, especially amplified ones, as well as movie scores are know to create haptic sensations coupled with the sound, conveying valuable information such as articulation and timing. Several studies have tried to replicate and quantify this phenomena with compelling results [18, 17, 19, 20].

Merchel approached the topic from an architectural acoustics point of view, aiming to prove that concert halls with a strong haptic feedback improve the overall quality of the concert experience [19]. His studies proposes several signal processing techniques to be used for the haptic channel, indicating that in music with a rich low end, the audio signal passed through a low pass filter is enough to improve the experience. Furthermore, he suggests that simple sinusoidals with frequencies not related to the audio signal will produce an enhanced listening experience, but the frequency of these haptics oscillators will have an impact in the overall perception [19].

Other group of authors suggested to account for haptic feedback at the composition stage, creating a coherent audio-haptic experience, instead of approaching haptics as an afterthought [21, 20]. Gunther and O'Modhrain coined the term *tactile composition* as a *system that facilitates the composition and permeation of intricate, musically structured spatio-temporal patterns on the surface of the body*, emphasizing the importance of a compositional language for the sense of touch[20]. Their 2001 *Concerts for the skin* experiments surface some important notions like selective haptic attention - the ability to selectively direct attention into different stimuli, if several body areas are actuated at the same time [20]. On top of that they suggest that the music-haptic relationship does not need to produce congruent stimuli at all times, and the composer should engage into a parallel multimodal composition that inter-plays between the two sensory channels.

Listening for pitch is almost always dependant on the frequency of the audio content, while the timbre and amplitude rarely have an impact on pitch perception [22]. In contrast, the perception of frequency from a vibrotactile stimuli is more complicated due to the multi-channel nature of the cutaneous sensing organ - the skin [22]. Moreover, perception of tactile frequency is amplitude and time dependant, and it varies significantly depending on the position on the body. Nevertheless, there is one important similarity between auditory and tactile pitch perception: within certain frequencies, the discrimination fits a critical band model [23]. Specifically, certain frequency ranges are perceived as distinct sensations, indicating that with enough exposure, tactile pitch perception can be interpreted similarly to the auditory one - a fact proven by many hearing impaired people [24]. This does not mean that understanding music through vibrotactile stimuli is equivalent to hearing it, but

the experience, while different, could be just as enjoyable.

Music usually uses a wider frequency spectrum than the skin can provide, and the tactile pitch-amplitude coupling only makes understanding music without hearing it more complicated. Unlike the ear, with its single receptor capable of 20Hz-20000Hz frequency range perception, the skin has multiple types of receptors, each with its own frequency and temporal characteristics. For music perception, two of them prove to be useful: the Meissner corpuscles and the Pacinian receptors [25]. The Meissner corpuscles, also known as Rapid Adapting (RA) receptors, have a very high innervation density and have a limited frequency range of 10Hz-100Hz, with a peak sensitivity around 40Hz. The Pacinian receptors are larger than the RA ones, have a low spatial resolution, and a frequency response between 40Hz and 1000Hz, and are most sensitive around 250Hz [25]. In an attempt to describe the tactile music properties, Erp & Spapé conducted an experiment on the perceptual attributes of vibrotactile melodies [26]. Their results indicate that users can perceive and evaluate multiple characteristics from the tactile stimuli(f.ex. aggressive, soft, alarming, bombastic, etc) and that melodies generally land in one of four clusters, on a two dimensional tempo-intrusiveness map [26]. In a similar fashion, Ternes & MacLean designed a large set of distinguishable tactile rhythms, further highlighting the potential of tactile melodies [27].

3 Implementation

3.1. Hardware

The hardware device is an ovoid shape with the following dimensions: 84mm wide, 58mm tall and 89mm deep and can be seen in figure 1.

Figure 1: Side and front view of the haptic device

The shape was inspired by the resting hand position when fixed with an orthopedic splint. This pose should minimize the strain on the wrist, and allow the fingers to relax in their natural rest position. The initial shape was created using modelling clay, aiming to ensure the finger position is anatomic, each digit having its own socket. The clay artefact was 3d scanned using Autodesk ReCap[1], by analyzing 40 still images of the subject, taken from multiple angles with a Fujifilm X-T1 camera and a Fujinon XF 35mm @ f2.0 lens. The artefact was suspended in midair with fishing line, affording visibility from all angles. The 3D scan resulted in a very high fidelity digital model, but in order to improve topology, a new 3d model was created using the scan as an outline.

[1] https://www.autodesk.com/products/recap/

The final shape was split in half horizontally, to have access inside where the electronics would eventually lie. The two halves were held together by three M3x16 bolts that have been incorporated in the design show in figure 1.

For the actuator, a socked was created on the bottom half of the device, and a 3.5mm female jack opening has been installed on top half to connect the transducer to the amplifier. The jack was oriented towards the left side of the device, allowing for cable connection that should not interfere with the user while holding it. The haptic device halves were fabricated with 2mm wall thickness using an Ultimaker 3 and PLA material.

The device was for left hand only, as it was intended to have the users navigate the experiment's questionnaire with the computer mouse, which is generally used with the right hand. Initial informal tests showed that people unfamiliar with the device tend to hold it in unintended ways, thus for the experiment finger positioning visual signifies have been painted on. When it comes to the transducer, a Tactuator BM1C vibrotactile actuator manufactured by Tactile Labs[2] was used. Haptuator Mk1 and Mk2 were also tried, but the Tactuator BM1C proved to have the highest amplitude in the current setup, and the distortion (if any) was non disruptive, as would be the case with Mk2 and M1 that rattle rather loud when overdriven. All transducers tested offer full spectrum reproduction. The tactuator requires amplification to achieve desirable amplitude therefore a high gain Behringer HA8000 headphone mixing and distribution amplifier was used.

3.2. Tactile signal processing

In an attempt to improve the perception of pitch through tactile sensing, three signal processing methods that convert arbitrary auditory signal into a tactile one were compared. Each of the processing methods was inspired from existing literature, and was re-implemented to exploited one physical or perceptual trait relevant for music listening.

Method 1: Compression of frequency spectrum

The first method focused on compressing the musically relevant frequency spectrum defined between 40Hz and 2093 Hz into a narrower "tactile range" one up to 1046 Hz, to address the Pacinian receptors exclusively, since these are the most sensitive to vibrotactile stimuli[22, 25]. The lower limit represents the crossover between RA receptors and Pacinian receptors, and the high frequency represents the top range of the Pacinian ones. The frequency compression was implemented as seen in Figure 2 as following:

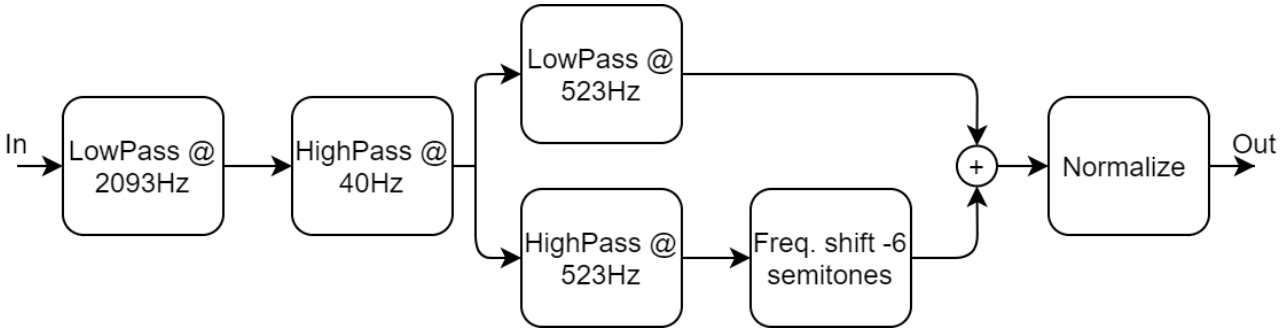


Figure 2: Signal processing for first condition

1. Apply a lowpass FIR filter with 60 dB/octave attenuation at 2093Hz - the corresponding frequency of the fundamental for a C7 note with A4 = 440 Hz tuning. This meant that only the highest octave available on a piano was ignored. Nevertheless, that the majority of instruments, including the human voice, have the high limit considerably lower than C7, Since the upper harmonics are not contributing much to melody perception, the frequency limitation was not consider to be a practical problem [28].

[2] www.tactilelabs.com

2. Apply a highpass FIR filter with 60 dB/octave attenuation filter at 40Hz to limit the actuation of RA receptors.
3. Split the frequency band at 523 Hz (C5 note) in two spectra, using a lowpass and a high pass filter. The lower one (40Hz - 523Hz) will be called Spectrum A, and the higher on B. The C5 note was chosen in relationship to the fundamental frequencies of the melodies used, and described in 3.
4. Pitch shift down spectrum B 6 semitones to shift the high frequency content into the tactile sensible range
5. Add the Spectrum A (original) and Spectrum B (pitch shifted)
6. Normalize to 1 to avoid clipping, and ensure equal amplitude throughout the melodies selection.

Method 2: Sinusoidal oscillators
The second method focused on Pacinian receptors as well, and it used sinusoidal with the frequency equal to the fundamental one of the actual tone, instead of the original signal as suggested by Merchel [10, 13, 14, 19]. This was done with the aim of avoiding higher frequency content from masking or diminishing the fundamental harmonic perception, since tactile spectral masking works similar to auditory one [23]. The sinusoidals were generated using the same MIDI information as the auditory signal, using Xfer Serum[3] wavetable synthesizer with *Basic Shapes* table, on position one and a square envelope(0 attack, max sustain, 0 decay). In order to ensure amplitude coherence between the auditory and tactile stimuli, the contour/envelope was extracted from the original file and applied to the haptic one. The last step was to apply normalization, similar to method 1.

Method 3: Tactile transient reinforcement
The last tactile signal processing tried to make use of both the RA and Pacinian receptors. The tactile signal combined the auditory stimuli with a haptic reinforcement one, aimed at the RA receptors in order to emphasize changes in pitch, practically working as an exciter or *transient emphasizer*. This feedback approach was inspired by the way frets provide guitar players feedback about the note selection, as described it [29]. The haptic signal was created by adding a haptic reinforcement component, to the signal generated with method 2. The haptic reinforcement was generated similarly to the sinusoidal described above, but 3 octaves lower than the auditory signal. This meant that the frequencies lied in the peak frequency response of the RA receptors [25]. An attack-decay (AD) envelope was used for the haptic reinforcement signal, with 10ms attack time, in order to reduce artefacts(clicks), and 500ms logarithmic decay time to avoid temporal masking over the higher frequency signal. The two signals were summed with amplitudes of 0.8 for the haptic reinforcement, and 0.2 for the original, unprocessed signal, followed by normalization. The large difference in volume between the two signals is due to the lower amplitude response of the Tactuator BM1C below 40Hz.

All processing was done in Matlab unless specified otherwise. Highpass and lowpass filters had a steepness of 0.8 (default in Matlab). The amplitude contour was computed as the moving RMS envelope of the unprocessed melody every 5000 samples, in order to avoid artefacts introduced by abrupt changes in loudness. Sampling frequency used was 48kHz, and all files were exported uncompressed (wav).

3.3. Melodies

The 75 musical melodies composed for the project, were all of a duration of three to eight seconds and spread across a randomized selection of different major- and minor keys; figure 3 shows the distribution of notes across all melodic phrases. The melodies were simple and kept in a melodic style easily recognizable for listeners familiar with western music. They all represented a small musical progression with a beginning and an end. Tempo was 120bpm and, rhythmically, there were a mixture of whole- quarter- eighth and sixteenth- notes. For each of the 75 true melodies, two false were added. The false melodies always had the same rhythmic content as the true one, but at least 75 percent of the tones were changed. In the false melodies, the musical progression would not be perceived as natural, since the selection of notes were not following western melodic tradition. Figure 4 shows the average number of semitones deviating from the correct melodies.

[3] https://xferrecords.com/

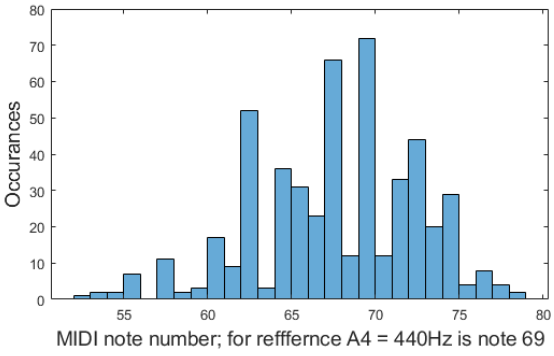


Figure 3: Distribution of notes in the correct melodies

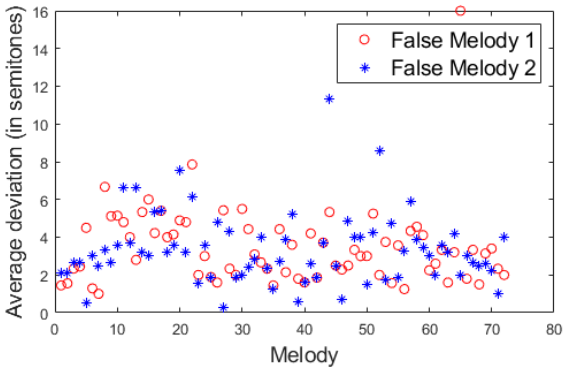


Figure 4: False melodies deviation from the correct ones

4 Evaluation

The aim of the study was to (1) evaluate three signal processing techniques for converting audio material to tactile stimuli, and (2) to evaluate the usability of the haptic device itself. To meet this aim a within-subjects study was performed, comparing four conditions that varied in term of the tactile feedback provided through the haptic device, when listening to melodic phrases.

The hypothesis was: *There is a difference in terms of tactile melody discrimination between an unprocessed signal and a processed one when presenting congruent bi-modal melodic phrases through a single-actuator handheld vibrotactile device and headphones.*

4.1. Task and Stimuli

The task for the participant was to select the haptic feedback that matched the auditory signal played through the headphones. A three alternative forced choice design was used, with only one correct option; for each trial the participants were presented with three types of haptic feedback. The experiment had four conditions, with different singal processing techniques described in section 3: [1]Control condition with no processing, [2]Frequency compression, [3]Sine wave at the fundamental frequency and [4]Tactile reinforcement of transients.

There were 72 trials in total: 18 for each processing technique, plus 18 for unprocessed acting as control condition. Melodies where chosen and presented randomly out of pool of 75 possibilities, distributed equally among the three instruments. A total of 900 possible trials were used for the whole experiment: 75 melodic lines * 4 conditions * 3 instruments, ensuring a high level of validity. For each instrument, condition and melodic line there was one correct haptic stimulus, and two incorrect ones. The order for melodies, order of conditions and choice of instruments were assigned randomly, in real time, for each participant. In order to ensure similar exposure levels for all participants, they were allowed to experience each stimuli/melody combination only once. The experiment took place in the Multisensory Experience Lab at Aalborg University campus in Copenhagen.

4.2. Participants

Participation in the experiment was voluntary and the majority of participants were students of *Sound and Music Computing* and *Medialogy* programs, that are affiliated with the Multisensory Experience Lab, and the Tonmeister program from the Royal Danish Academy of Music. Some participants had musical experience, but this was not a selection criteria. The participants have been encouraged to partake in the experiment at their convenient time and there was no reward for doing it. The data collected has been anonymous, without the possibility of matching answers sets with the participant. There were a total of 34 participants (24 male, 10 female). Although the ultimate goal is to create a device for hearing impaired users, current COVID-19

restrictions prevented us to test on the relevant target group. The experiment was conducted with 3 participants in parallel, in 3 different rooms, that were briefed and debriefed together, by the first author.

4.3. Setup and Equipment

The hardware setup consisted of a Windows computer running the experiment application with a Behringer HA8000 headphone distribution amplifier connected to it. The left audio channel contained the haptic melodies, and the right channel carried the auditory signal. The distribution amplifier routed the auditory signal to both headphones channels. The headphones used were different due to availability, but had similar price and quality level: *Creative Aurvana Live!, AKG K240 and Sennheiser HD240 Pro*. The level balance between the haptic and auditory signal was set by the second author, and calibrated to have a natural balance between the two sensory inputs to assure that not one would overpower the other. First the headphones were adjusted to a comfortable playback-level and then the haptics were added up close to the distortion limit of the transducer. The mix was constant for all participants.

4.4. Procedure

As mentioned, three participants partook in the experiment at the same time. They were welcomed and introduced to the task, emphasizing that it is not required from them to over-analyze the vibrotactile stimuli, but instead they should answer based on their intuition. The participants were then guided to the setup rooms and required to experience Queen's "Don't stop me now", as training and accommodation with the system. The song was chosen due to its popularly, but also because it features many combinations of instruments and intensities: from low intensity voice only, to high intensity full band playing. This should provide the users with most of the possible stimuli expected throughout the experiment. After the accommodation phase was finished, the users were required to click on "Start" button to begin the experiment. Each trial consisted of listening to the same auditory melody three times, with different haptic stimuli for each as described in 2. A visual indicator was signaling what exposure was playing at all times, and the "Select the haptic stimuli that matches best the melodic phrase you heard" message was permanently displayed in the bottom of the page, followed by the trial number. There was a 2 seconds gap between exposures within the same trial. Since all potential melodies were fairly similar in length, the experiment was completed in 18-19 minutes. After all 3 participants in a series finished, they were gathered for a post-experiment debriefing discussing about their experience, comfort, amplitude and potential suggestions. The setup, similar to an ad-hoc focus group, facilitated interesting discussions between participants, as well as between conductors and participants.

4.5. Data Collection

The data collected has been anonymous, without the possibility of matching answers sets with the participant. The following information was logged: trial number(1-72), melody number(1-72), condition(1-4), instrument(1-3), correct answer, user answer, and inevitably, the log file creation time.

5 Results

The data collected from the 34 participants was treated as nominal and was analyzed using Friedman tests. The main test was run to determine if there were significant differences between the three proposed signal processing techniques with respect to number of correct identification of the matching haptic stimuli. In addition to this analysis, the data was analysed on a per-instrument basis, as well as instrument performances against each-other. The number of correct answers were not statistically significantly different among conditions $\chi^2(3) = 0.885, p = .829$. Similarly, no significant differences were found when the instruments were analyzed independently: bass trials ($\chi^2(3) = 2.590, p = .459$), synth trials ($\chi^2(3) = 4.528, p = .210$), and

trumpet trials ($\chi^2(3) = 1.401, p = .701$).

No effect on the number of correct answer was observed, but while inspecting the descriptive statistics, it seemed that one instrument (synthesizer) stands out therefore an exploratory analysis was ran, comparing the sets of trials for each instrument against each other. The results are $\chi^2(2) = 1.746, p = .418$. Furthermore, when looking at the best and worst performing three melodies in terms of correct answers, it was discovered that the ones with multiple short notes had a slightly higher average number of correct answers, while the ones with longer, sustained notes had a lower number of correct answers, even when harmonic content is very similar. The correlation between average note length in the melody and number of correct answers is ρ = -0.19 with p = 0.1. Figure 5 show the best performing melody with 14 correct answers out of 34 (41.1%) and worst performing one with 2 correct answers (5.8%). Lastly, there was no correlation found between the average deviation from correct melody as show in 4 and the number of correct notes: $\rho < 0.07$.

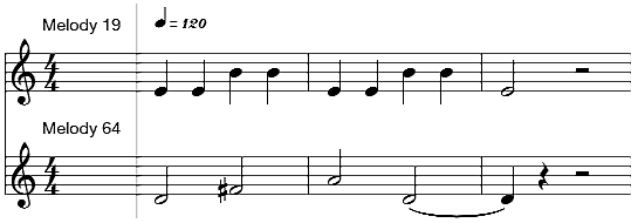


Figure 5: Best(19) and worst(64) performing melodies

5.1. Post-experiment interview

The post experiment interview highlighted some interesting facts regarding the physical design, the experiment as well as potential directions for further experimentation. Several participants remarked that the experiment is too long and repetitive, loosing focus towards the end. Regarding the physical properties, some subjects reported that they experimented with different arm resting positions (arm resting on the knee facing up/down, arm resting on the table, crossed arms) noticing that each position will produce slightly different results. Out of those who mentioned position, there seemed to be a consensus that palm facing up feels the best, with one mention that it felt stronger. Probably the most interesting feedback was that some participants felt different frequencies in different areas of the hand, one participant mentioning that sometimes it could sense two frequencies at the same time. This has been expressed in various forms, some claiming that higher frequencies feel too strong, especially for the fingers, but the lower frequencies feel good.

Regarding the hardware, the feedback has been generally good, but some participants suggested that the device was either too big or too small for their hands. At the same time, few participants reported that it is a slightly uncomfortable to hold for long time while most mentioned it was comfortable.

6 Discussion

The results related to *number of correct answers* suggest that the proposed processing techniques do not result in statistically significant different performances. Furthermore, looking at the nature of the haptic stimuli, it is observed that the *bass, synthesizer and trumpet* perform similarly among the 4 conditions. Even though the performed tests does not permit us to conclude that the results are statistically equivalent, the descriptive statistics does seem to indicate that the effect of processing methods was negligible, and possibly non-existing. This can be seen in figure 6 showing the median for all conditions is 1 and the variance is consistent, regardless of the instrument presented. A potential explanation for this similarity can be found in the fact that a higher number of shorter notes are easier to recognize, opposite to longer, sustain ones as seen in Figure 5. A similar behavior was observed by Tommerdahl et. al in 2005 in his studies on the vibrotactile discrimination capacity of skin for various stimuli lengths, concluding that the cerebral cortex undergoes a profound inhibition withing 1-2 seconds from the start of a 200Hz stimuli [30]. That being said, this study did not investigate the impact of legato or staccato as all harmonic events were rhythmically independent, therefore no conclusion can be reached

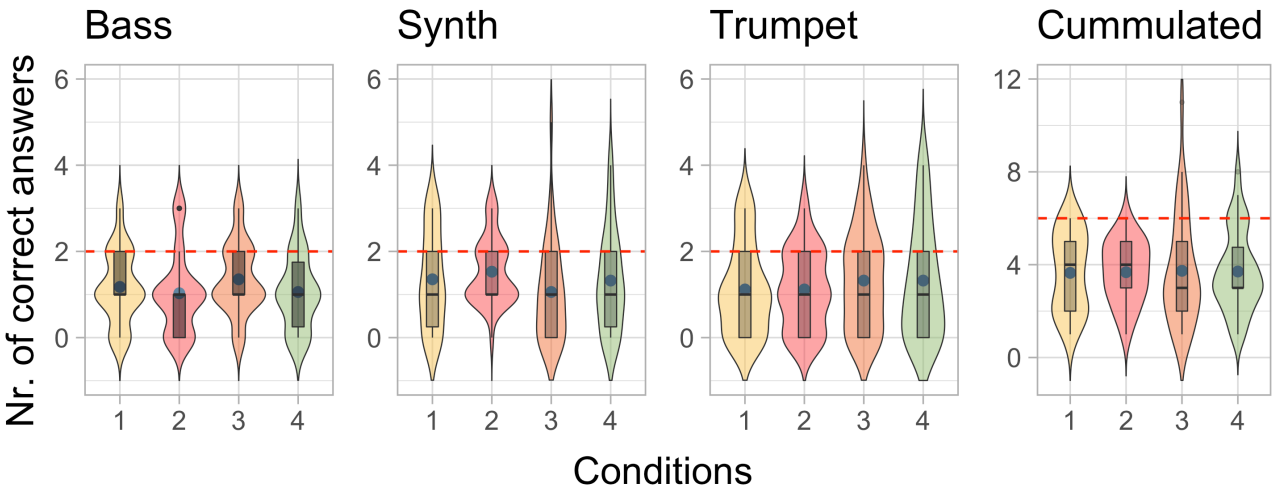


Figure 6: Distribution of number correct answers (max. possible = 6/instrument/condition),for each instrument, for each of the four conditions: 1 = control condition; 2 = frequency compression; 3 = sine wave at the fundamental frequency; 4 = haptic reinforcing of transients. Box plots inside the violin plots represent interquartile range, blue circles represent the mean and red line represents the expected chance level.

on whether the rapid melodic changes or the short duration explain this phenomena. Nonetheless, if condition 1 - control and condition 3 - sinusoidal oscillators, are analysed in isolation, the results do not align with the findings of Merchel, that suggest using sine signals with the frequency matching the fundamental one from the auditory signal produces a better tactile experience - at least in terms of melodic content identification [19].

When it comes to the comfort and performance of the physical device, the results are mixed. Some users claimed it was comfortable and provided appropriately strong vibrotactile stimuli, while other complained that it can be too strong at times or that it becomes uncomfortable to use for longer periods of time. These findings indicate a preferences for individual customization of device size as well as control over the haptic intensity.

Lastly, an interesting phenomena was describe by several users, claiming that different frequencies are felt in different areas of the hand. This is a direction worth exploring further, since it can indicate that single actuator devices can address different areas of the hand, providing an extra dimension for communication.

7 Conclusion

In this paper it was proposed a system that allows musical signals to be converted to vibrotactile stimuli. The system was evaluated in a user experiment exploring impact of three signal processing techniques used for audio to haptic conversion, in terms of user's ability to identify the melodic information. The stimuli used for the experiment was composed of short melodies played on double bass, synthesizer and trumpet. The results indicated that there was no significant difference between the 3 proposed techniques and no processing at all, when it comes to melody identification, underlining the need for new algorithms that can be empirically validated. However, there was an indication that users do perform better at the identification task when the haptic stimuli contains shorter notes, regardless of processing algorithm or instrument played. Finally, it was surfaced that different areas of the hand can sense separate frequencies, but further research needs to be conducted in order to fully understand the phenomena.

Acknowledgements

This work is supported by NordForsk's Nordic University Hub - Nordic Sound and Music Computing Network.

References

[1] Stefano Papetti and Charalampos Saitis. *Musical Haptics: Introduction*. 2018. doi: 10.1007/978-3-319-58316-7_1.

[2] P. L. Brooks and B. J. Frost. Evaluation of a tactile vocoder for word recognition. *The Journal of the Acoustical Society of America*, 74, 1983. ISSN 0001-4966. doi: 10.1121/1.389685.

[3] B.E. Stein, P.C.B.E. Stein, and M.A. Meredith. *The Merging of the Senses*. A Bradford book. ISBN 9780262193313.

[4] John Foxe, Glenn Wylie, Antigona Martinez, Charles Schroeder, Daniel Javitt, David Guilfoyle, Walter Ritter, and Micah Murray. Auditory-somatosensory multisensory processing in auditory association cortex: An fmri study. *Journal of neurophysiology*, 88:540–3, 08 2002. doi: 10.1152/jn.00694.2001.

[5] Christoph Kayser, Christopher Petkov, Mark Augath, and Nikos Logothetis. Integration of touch and sound in auditory cortex. *Neuron*, 48:373–84, 11 2005. doi: 10.1016/j.neuron.2005.09.018.

[6] Ryuta Okazaki, Hiroyuki Kajimoto, and Vincent Hayward. Vibrotactile stimulation can affect auditory loudness: A pilot study. pages 103–108, 06 2012. doi: 10.1007/978-3-642-31404-9_18.

[7] Ryuta Okazaki, Taku Hachisu, Michi Sato, Shogo Fukushima, Vincent Hayward, and Hiroyuki Kajimoto. Judged consonance of tactile and auditory frequencies. pages 663–666, 04 2013. ISBN 978-1-4799-0087-9. doi: 10.1109/WHC.2013.6548487.

[8] Tony Ro, Johanan Hsu, Nafi Yasar, Caitlin Elmore, and Michael Beauchamp. Sound enhances touch perception. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 195:135–43, 04 2009. doi: 10.1007/s00221-009-1759-8.

[9] Gareth Young, David Murphy, and Jeffrey Weeter. Haptics in music: The effects of vibrotactile stimulus in low frequency auditory difference detection tasks. *IEEE Transactions on Haptics*, PP:1–1, 12 2016. doi: 10.1109/TOH.2016.2646370.

[10] S. Shin, C. Oh, and H. Shin. Tactile tone system: A wearable device to assist accuracy of vocal pitch in cochlear implant users. 2020. ISBN 9781450371032. doi: 10.1145/3373625.3418008.

[11] Mark Fletcher, Nour Thini, and Samuel Perry. Enhanced pitch discrimination for cochlear implant users with a new haptic neuroprosthetic. *Scientific Reports*, 10:1–10, 2020. URL https://eprints.soton.ac.uk/441275/.

[12] M.J. Luciá, P. Revuelta, Á. Garciá, B. Ruiz, R. Vergaz, V. Cerdán, and T. Ortiz. Vibrotactile captioning of musical effects in audio-visual media as an alternative for deaf and hard of hearing people: An eeg study. *IEEE Access*, 8:190873–190881, 2020. doi: 10.1109/ACCESS.2020.3032229.

[13] U. Trivedi, R. Alqasemi, and R. Dubey. Wearable musical haptic sleeves for people with hearing impairment. pages 146–151, 2019. ISBN 9781450362320. doi: 10.1145/3316782.3316796.

[14] M.D. Fletcher, S.R. Mills, and T. Goehring. Vibro-tactile enhancement of speech intelligibility in multi-talker noise for simulated cochlear implant listening. *Trends in Hearing*, 22, 2018. doi: 10.1177/2331216518797838.

[15] G.W. Young, D. Murphy, and J. Weeter. Vibrotactile discrimination of pure and complex waveforms. pages 359–362, 2015. ISBN 9780992746629.

[16] ALVAR WILSKA. On the vibrational sensitivity in different regions of the body surface. *Acta Physiologica Scandinavica*, 31, 1954. ISSN 1365201X. doi: 10.1111/j.1748-1716.1954.tb01139.x.

[17] Antonella Mazzoni and Nick Bryan-Kinns. How does it feel like? an exploratory study of a prototype system to convey emotion through haptic wearable devices. 2015. doi: 10.4108/icst.intetain.2015.259625.

[18] Marcello Giordano, John Sullivan, and Marcelo M. Wanderley. *Design of Vibrotactile Feedback and Stimulation for Music Performance*. 2018. doi: 10.1007/978-3-319-58316-7_10.

[19] Sebastian Merchel and M. Ercan Altinsoy. Auditory-tactile music perception. volume 19, 2013. doi: 10.1121/1.4799137.

[20] Eric Gunther and Sile O'Modhrain. Cutaneous grooves: Composing for the sense of touch. *International Journal of Phytoremediation*, 21, 2003. ISSN 15497879. doi: 10.1076/jnmr.32.4.369.18856.

[21] Joanne Armitage and Kia Ng. doi: 10.1007/978-3-319-46282-0_9.

[22] David M. Birnbaum and Marcelo M. Wanderley. A systematic approach to musical vibrotactile feedback. 2007.

[23] JC Makous, RM Friedman, and CJ Vierck. A critical band filter in touch. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 15(4):2808—2818, April 1995. ISSN 0270-6474. doi: 10.1523/jneurosci.15-04-02808.1995.

[24] Lisa Bruns, Dirk Mürbe, and Anja Hahne. Understanding music with cochlear implants. *Scientific Reports*, 6, August 2016.

[25] Ki-Uk Kyung, Minseung Ahn, Dong-Soo Kwon, and Mandayam A. Srinivasan. Perceptual and biomechanical frequency response of human skin implication for design of tactile displays. In *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, WHC '05, page 96–101, USA, 2005. IEEE Computer Society. ISBN 0769523102. doi: 10.1109/WHC.2005.105. URL https://doi.org/10.1109/WHC.2005.105.

[26] Jbf Van Erp and Mma Spapé. Distilling the underlying dimensions of tactile melodies. *Proceedings of Eurohaptics 2003*, 2003.

[27] David Ternes and Karon E. MacLean. Designing large sets of haptic icons with rhythm. volume 5024 LNCS, 2008. doi: 10.1007/978-3-540-69057-3_24.

[28] John M. Eargle. *Frequency Ranges of Musical Instruments and the Human Voice*, pages 324–325. Springer US, Boston, MA, 2002. ISBN 978-1-4615-2027-6. doi: 10.1007/978-1-4615-2027-6_156. URL https://doi.org/10.1007/978-1-4615-2027-6_156.

[29] Håkon Knutzen, Tellef Kvifte, and Marcelo M. Wanderley. Vibrotactile feedback for an open air music controller. volume 8905, 2014. doi: 10.1007/978-3-319-12976-1_3.

[30] M Tommerdahl, KD Hester, ER Felix, M Hollins, OV Favorov, PM Quibrera, and BL Whitsel. Human vibrotactile frequency discriminative capacity after adaptation to 25 hz or 200 hz stimulation. *Brain research*, 1057(1-2), September 2005. ISSN 0006-8993. doi: 10.1016/j.brainres.2005.04.031.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Conception of the listening test procedure for quantifying speech intelligibility in Slovak language – a preliminary study

Dominika Húdoková[1,2*], Vojtech Chmelík[2], Lukáš Zelem[2], Daniel Urbán[2], Monika Rychtáriková[1,2]

[1]Department of Architecture - Campus Brussel and Gent, KU Leuven, Belgium.
[2]Department of Materials Engineering and Physics, Faculty of Civil Engineering, STU Bratislava, Slovakia.

* dominika.hudokova@kuleuven.be / dominika.hudokova@stuba.sk

Abstract
This paper describes the methodology used for the conception of the Slovak sentence listening test, suitable for speech reception thresholds (SRT) experiments in quiet and noise. The procedure on choose of suitable sentences in Slovak language is addressed and method on sorting them according to their intelligibility is described. Results shown in this paper are based on a preliminary headphone-based listening test experiments performed on 39 normal hearing female and male human subjects. The task of each respondent was to repeat presented sentence in noise, with individualized signal to noise ratio, tuned for each test person in a pre-test by means of the method of adjustment. The application of the newly developed well-balanced sets of sentences in Slovak language is twofold. Its direct usage can be seen in audiological experiments with severely hearing-impaired individuals and cochlear implant users with Slovak as native language. Secondly, it can serve as a tool for validation of speech related room acoustic quantities, such as *STI* and D_{50} or U_{50} for Slovak language.

Keywords: speech intelligibility, listening test, Slovak language, methodology, sentences

1 Introduction

The speech intelligibility in architectural context is typically performed using single number quantities such as *STI*, C_{50}, D_{50}, U_{50}, etc. [1, 2]. Although the topic of speech intelligibility in architectural context is not a new field of interest, there are several aspects which have not been tackled or fully explored yet e.g. to what extent the mentioned acoustic quantities represent the speech intelligibility of Slovak language? Slovak language, in general, contains large number of consonants and sibilants, which carry lots of ”information”, but carry less sound energy. Thus, we can set a hypothesis that Slovak language might be more difficult to understand in unfavourable conditions when comparing to e.g. Dutch or English. To answer this question, it is necessary to perform a number of listening tests. Several standardized speech intelligibility tests procedures already exist, that has been designed for different languages. Well known are e.g. listening tests based on meaningful sentences using everyday language [3-8], so called matrix test which are specific by means of fixed order of items: name, verb, numeral, adjective, noun etc. [9]. This particular strict content of sentences produces an unpredictable meaning of sentences. [10, 11].
This paper discusses partial results of a dissertation thesis concerning Speech intelligibility in various architectural environments. The main aim is a detailed discussion on methodology for the design of the sentence listening test for Slovak language.

2 Methods

2.1 Criteria for selection of sentence database

Altogether, 800 ordinary Slovak sentences were suggested. The choice of the sentences was based on the following presumptions: (a) to be short enough to repeat. It means that sentences contain minimum 2 and maximum 12 words. (b) Sentences represent conversational speech. There are no scientific phrases, difficult names or specific brand names. (c) None exclamations or questions. (d) Sentences should not to be too unessential. (e) There are no proverbs, to avoid the conceit. The criteria for designing Slovak sentence test material were inspired by van Wieringen and Wouters [7].

2.2 Psychoacoustic method used for listening test

Psychophysics generally focuses on detecting the relationship between a physical stimulus and its perception and on detecting the phenomenon [12]. In psychoacoustic listening tests, we rely on people's answers. Therefore, statistical analysis is essential in data analysis. Statistically relevant sample of people is one of the starting conditions. The first method used in our experiment is called method of adjustment, see in the Figure 1 below. In this method, subject is asked to control the level of the stimulus and keep it changing until the stimulus is narrowly detectable against the background noise [12].

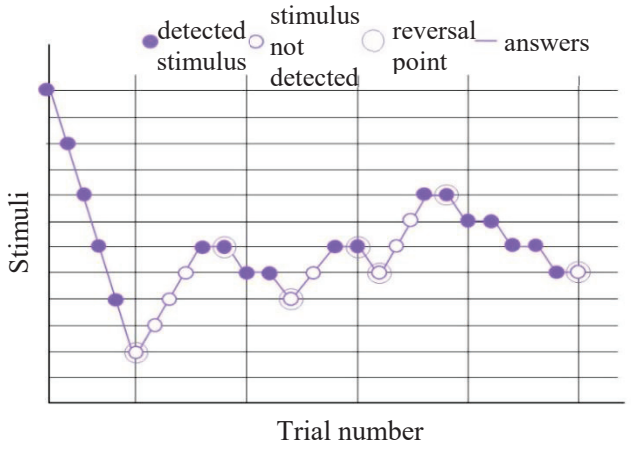


Figure 1: Graph illustration of adjustment method.

3 Listening experiments

3.1 Test subjects

In total, 39 human subjects (20 male and 19 female) participated in the listening test experiment. Their age was between 25 and 61 years. The average age is 33 years, the modus value is 27 and the median value is 28 years. Every individual has declared that he/she did not suffer from any hearing impairment and all of them had normal hearing (<20 dB HL hearing loss range for octave frequencies from 125 Hz to 8000 Hz, measured in the laboratory). Each test subject, that participated in the experiment, received the same explanation from the test operator. They were also explained that the psychoacoustic experiments are not an exam with correct or incorrect answers or a competition on performance and that the most important is to give honest and true answers.

3.2 Stimuli

The recordings of the 800 chosen sentences were performed in the Acoustic lab – the semi-anechoic room at STU Bratislava, Faculty of civil Engineering, Department of Materials Engineering and Physics. They were spoken by female person and recorded at normal speech level, regular speed and spoken in monotonic way without emphasis on any of the words. During the recordings, the speaking person stood in front of microphone in the empty semi-anechoic room alone. The recorded signals were sampled at 44100 Hz (16 bit A/D converter). In editing process, 30 ms cosine window was applied at the beginning and the end of each sentence to avoid click sound.
The whole sentence material was stored in a ".wav" format on disk of the computer. The masking noise necessary for performance of the tests was based on the long-term average speech spectrum of randomly picked 40 sentences (the duration of sample was 2 minutes) of 800 sentence the database. It was a stationary noise with a spectrum of the female voice (used for recordings of the signal in this experiment) thus, a so called spectrally matched masker.

3.3 Set-up

The set up consisted of the computer with home-made, Visual Basic based listening test application, the sound card (Focusrite Scarlett 6i6), two pairs of headphones (one used by test subject and one by operator) and a microphone. The test subject was seated in the semi-anechoic room in the Acoustic lab (Figure 2) and performed listening tests by means of calibrated open headphones (Sennheiser HD650) with a flat frequency response. The headphones were calibrated using analyser Norsonic 140. The sentence material was played from a computer into the headphones via sound card. The test subjects communicated with operator via dynamic microphone, e.g. when repeating the sentences during the listening test, that they have heart via headphones. The test operator was seated in the control room and listened to the responses of people via the second headphones (Figure 2).

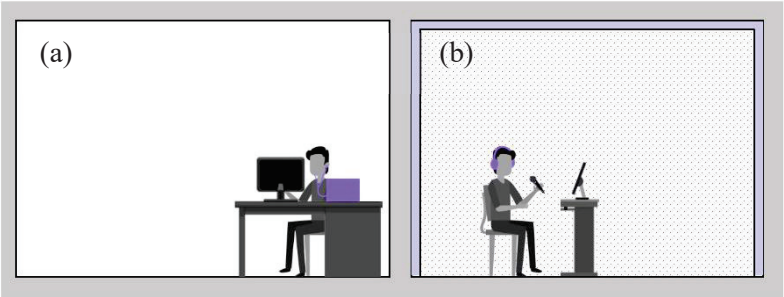


Figure 2. The schematic sketch of test rooms ((a) the control room, (b) the semi-anechoic room).

Test No.1
The first test served for estimation of the approximate SRT (Speech Reception Threshold) necessary for setting up the correct S/N ratio for the Test No.2. The psychoacoustic method used in the test No.1 was adjustment method and consisted of 40 sentences. Each one of the 40 sentences started at S/N ratio = -20 dB. The task of the test subject was to reduce the noise in 1 dB steps until he/she was certain about the meaning of that particular sentence, what means he/she could understand the whole sentence. Each sentence was played in the loop so that the test subject had the opportunity to listen to it as many times as needed. When it was done, the test subject could continue to another sentence. When the Test No.1 was finished, the signal to noise ratio was calculated from the average of all (40) tested sentences.

Test No.2 (main test)
The second test consisted of 760 sentences. These sentences were masked by the noise signal with fixed S/N ratios obtained in the Test No.1, and thus individual for each test subject. The task of a test subject was to listen each sentence in the fixed S/N ratio and to repeat it. Each sentence was played once. The operator seated

in the control room was present all the time and has evaluated answer, i.e. if the sentence was repeated correctly (i.e. each word of the sentence was correctly repeated) and noted the score. The operator could hear each sentence (played to test person) without presence of masking noise. For double check, the sentence which was played to a test person was at the same time also displayed on the screen of the operator. To avoid fatigue and errors in the test, participants were allowed to ask for a break anytime.

4 Results and discussion

The results of the Test No.1 are shown in the Figure 3. They show that the S/N ratio of the most (17) participants was -6 dB. Only two participants have adjusted the S/N ratio to -8 dB.

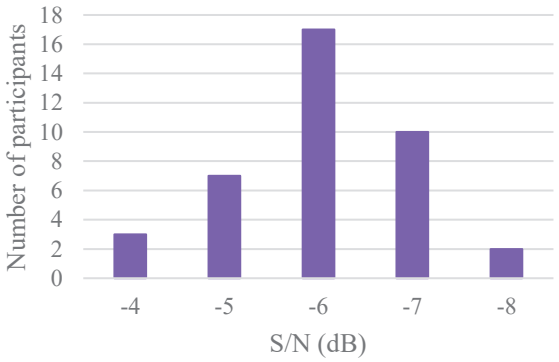


Figure 3: The histogram of S/N (dB).

The results of the Test No.2 were gathered in the matrix, where correct answers were indicated as “1” and incorrect as “0”. Doing so, each sentence got a score, which was later expressed in percentage of correct answers. Sentences were later sorted according to the easiest to most difficult in a form of a histogram. The global result can be seen in the Figure 4.

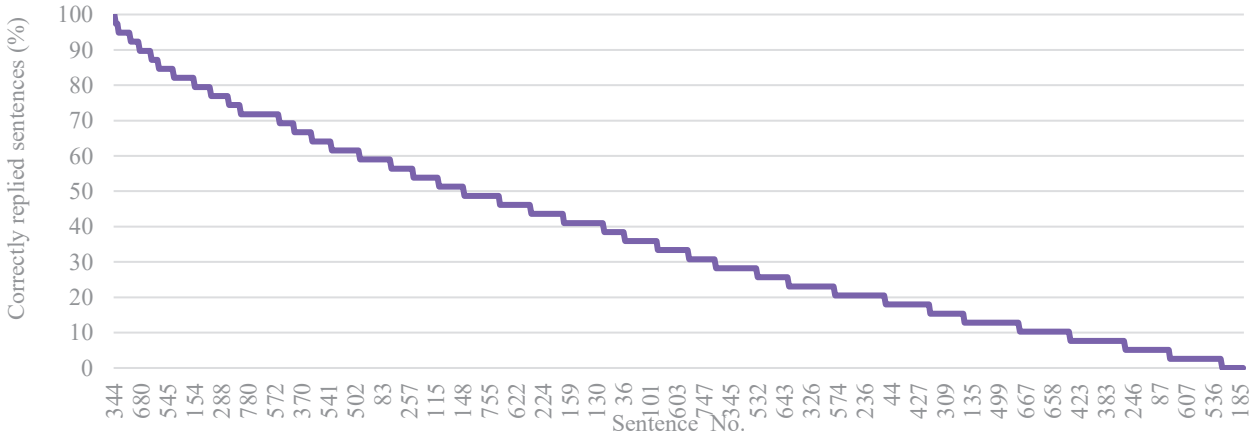


Figure 4: The percentage of the correctly replied 760 sentences. Showing the course of most simple to most difficult sentences.

The percentage of success rate includes whole range from 0% to 100%. We can see by looking at the Figure 4 that the whole sentence material included higher amount of too difficult sentences in comparison with too easy sentences to repeat. This is represented by slowly decreasing slope of presented histogram.
In order to determine the set of equally difficult sentences to understand, the following method was used: (1) the standard deviation of the percentage of correct replied sentences was calculated (STD=26%). (2) the centre of the whole percentage range was indicated (50%). (3) the standard deviation was added to centre of

percentage range as well as subtracted from the centre to obtain both the top and the low border of the range (50%+26%=76% it is the top border; 50%-26%=24% it is the low border). We can say that the equally difficult sentences are between 24% and 76% what means 52% of the sentence material.
If we look into the results deeper we can conclude the following: The levels of signal to noise ratio oscillated between -4 to -8 dB, but the level -6 dB was detected in the most of tested participants (Figure 3). In comparison with van Wieringen and Wouters (S/N = -8 dB; -10 dB) [7] we obtained lower average value of S/N. This can have several reasons, among which one can be caused by the fact that the Slovak language due to its large consonant content, is more difficult to be understand under presence of noise. However, this would need to be confirmed by a dedicated study. The majority of participants were adults between 25 and 35 years. Based on our observation, we can claim that age of participants had not affected the results of signal to noise ratio. (Figure 5). Moreover, the age has not affected the number of correctly repeated sentences either.
The range of correctly repeated sentences over participants was between 71 and 472 (Figure 6).

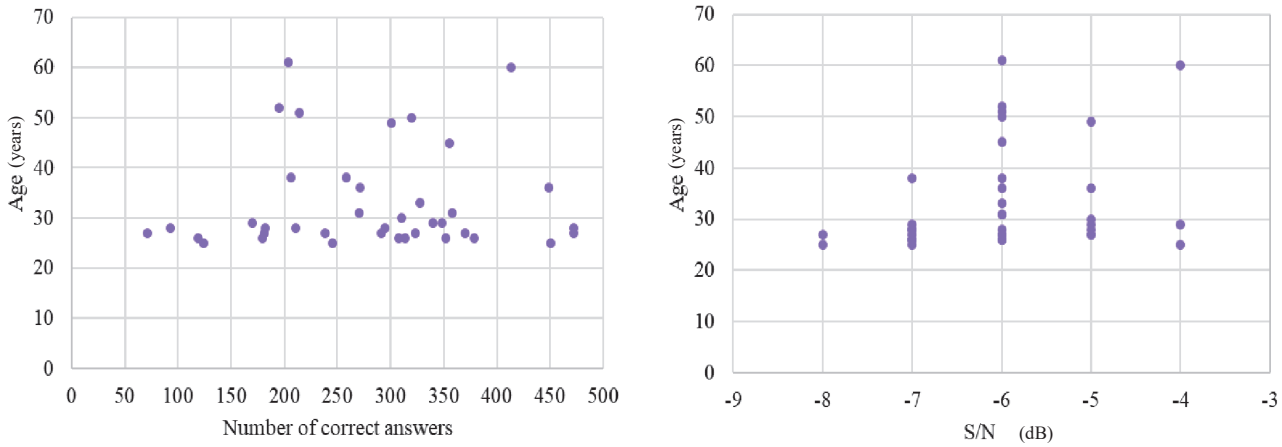


Figure 5: Correctly repeated sentence and S/N ratio in a function of the age (of participants).

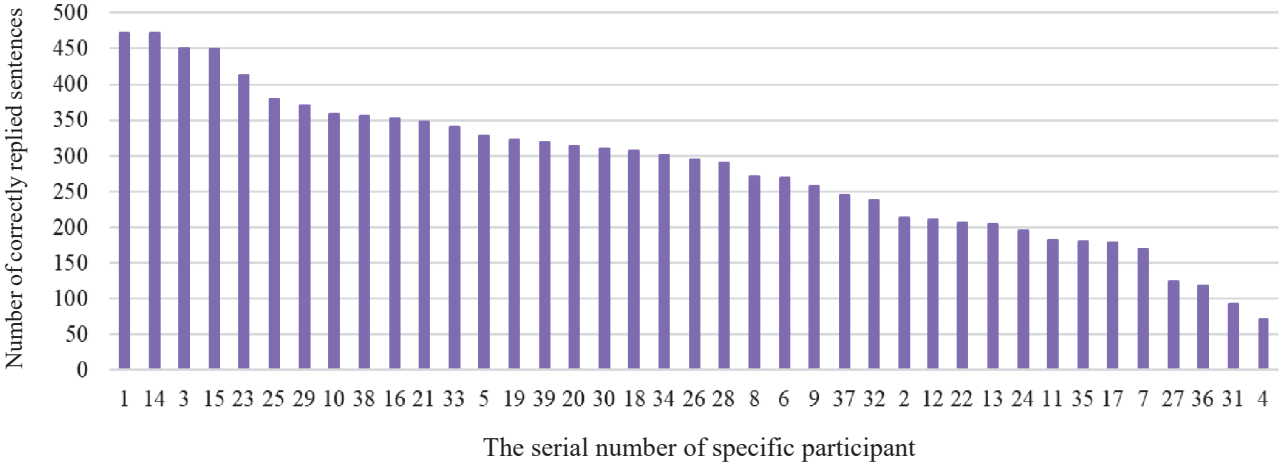


Figure 6: The quantity graph of correctly replied sentences of each test participant.

5 Conclusions

In this paper a sentence material consisting of 800 Slovak sentences was created following the rules based on the work of van Weiringen and Wouters [7]. According to the results based on laboratory listening tests, with 39 test subjects, the ranking of proposed sentences from the easiest to the most difficult has been made. A

deeper look into the sentences and explanation upon their difficulty is given in [14]. Altogether 376 sentences have been indicated as equally difficult.

Acknowledgements

The authors would like to express a gratitude to all participants of listening tests.

References

[1] ISO 3382-1 Acoustics – Measurements of room acoustic parameters, Part 1: Performance Spaces.

[2] IEC 60268 Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index.

[3] Plomp, R. – Mimpen, A. M.: Improving the reliability of testing the speech reception threshold for sentences, *Audiology*. Volume 18, Issue 1, p. 43-52, 1979.

[4] Nilsson, M. – Soli, S. D. – Sullivan, J. A.: Development of the hearing in noise test for the measurement of speech reception thresholds in quite and in noise, *The Journal of the Acoustical Society of America*. Volume 95, Issue 2, 1994.

[5] Versfeld, N. J. – Daalder, L. – Festen, J. M. – Houtgast, T.: Method for the selection of sentence materials for efficient measurement od the speech reception threshold. *The Journal of The Acoustical Society of America*. Volume 107, February 2000.

[6] Wong, L. L. N. – Soli, S. D.: Development of the Cantonese Hearing in noise test (CHINT), *Ear and Hearing*. Volume 26, Issue 3, p. 276-289, 2005.

[7] van Wieringen, A. – Wouters, J.: LIST and LINT: Sentences and numbers for quantifying speech understanding in severely impaired listeners for Flanders and the Netherlands. *International Journal of Audiology*. Volume 47, Issue 6, p. 348 - 355, 2008.

[8] Ozimek, E. – Kutzner, D. – Sek, A. – Wicher, A.: Polish sentence tests for measuring the intelligibility of speech in interfering noise, *Itnernational Journal of Audiology*, Volume 48, Issue 7, p. 433-443, 2009.

[9] Kollmeier, B. – Warzybok, A. – Hochmuth, S. – Zokoll, M. A. – Uslar, V. – Brand, T. – Wagener, K. C.: The multilingual matrix test: Principles, applications, and comparison across languages: A review, *International Journal of Audiology*. Volume 54, Issue 2, 2015.

[10] Hagerman, B.: Sentences for testing speech intelligibility in noise, *Scandinavian Audiology*. Volume 11, Issue 2, p. 79-87, 1982.

[11] Panocová, R. – Gregorová, R.: Designing the Slovak matrix sentence test, *International journal of applied languages and cultural studie*. Volume 2, Issue 2, 2019.

[12] Rychtáriková, M.: Psychoakustické testy v stavebnej akustike. 2015 ISBN: 9788097199364

[13] Gescheider, G.: The Classical Psychophysical Methods. Psychophysics: the fundamentals (3rd ed.). Lawrence Erlbaum Associates. ISBN 978-0-8058-2281-6, 1997.

[14] Hudcovičová, B. – Petrášová, M.: Linguo-Acoustic acpects of speech, *EUROREGIO/BNAM2022.* Aalborg, Denmark, 2022.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Innovation Loam Noise Barrier

Michael Chudalla*, Fabio Strigari

Section Environmental Protection, Immissions, Federal Highway Research Institute, Bergisch Gladbach, Germany.

*chudalla@bast.de

Abstract
Cob (in German ”Wellerlehm”) and stamped loam are traditional earthen construction methods that have been used in Europe for hundreds of years. With a pilot project near Nebelin in the German federal state of Brandenburg, this forgotten construction technology, which can now be considered innovative, is to be applied to a noise barrier at a planned service area on the A14 motorway. The project is called the ”Alhambra of Brandenburg”. The aim is to determine the necessary boundary conditions for the realisation of noise protection walls in earthen construction, which meet both the technical requirements of the German ”ZTV-Lsw” (Additional Technical Contractual Conditions for noise barriers) as well as ecologically sensible criteria. Due to the wall thickness of a loam noise barrier of up to one metre and the density of the building material alone, achieving a sufficient airborne sound insulation does not represent a big problem. However, to improve the absorbing properties, the dense and sound-reflecting surface of the cob must be supplemented by a kind of facing shell. Promising results in this regard are shown by initial measurements in the impedance tube with Heraklith and hemplime, a composite material made of hemp, light wood and lime binder. In further test series, the material compositions of the sample bodies are varied in order to optimise the absorption spectrum to the road traffic spectrum.
It is hoped that the natural and locally occurring materials will contribute to a reduction in energy consumption and, due to the high content of renewable raw materials, a favourable CO_2 balance and a reduction in greenhouse gas emissions. Furthermore, the purely naturally occurring raw materials make the building completely recyclable.

Keywords: innovations, sustainable, noise barrier, new material

1 Introduction

Innovations in noise protection are usually understood to be new, artificially created materials or geometries that follow complex physical effects and require extensive calculations to function correctly. These can be, for example, so-called metamaterials or geometries which cause improvements in the near field due to diffraction, scattering or refraction effects, and thereby leading to level reductions in the far field.
The innovation of this study is a new but historically ancient building material for road noise protection, namely loam. Loam is an ancient building material used all over the world. There are still numerous buildings, some of them historical, which were built using loam. The loam construction method cob was particularly widespread in eastern Germany. The mixture of cob consists mainly of loam and straw, or other long-stemmed plant stalks. Partly due to environmental protection and sustainability considerations, this building material is experiencing a renaissance. Clay is again being used as a building material, not only in residential construction but also in larger industrial buildings. It is therefore not surprising that there are efforts to use it in road construction, too. This will be described in more detail here.

1.1 Sustainable loam construction

The expansion of the German A14 motorway [1] between Karstädt and Wittenberge runs in the immediate vicinity of and partly through the biosphere reserve "Elbe River Landscape". A rest area is to be built near Nebelin, which is to be embedded in the surroundings in a way that is as landscape- and people-friendly as possible [2]. A feasibility study is being conducted to determine whether various structures of the planned rest area can be built using earthen construction methods. For example, the acoustic and non-acoustic properties of loam construction could be investigated on the perimeter walls of the facility in order to assess its suitability for noise protection purposes. For official use, these walls must meet the requirements of the "Additional Technical Contractual Conditions and Guidelines for the Construction of Noise Barriers on Roads", in short ZTV-Lsw 06 [3].
One advantage of this construction method would be that the excavated soil produced on site can be used directly as building material and the erected noise barrier is thus fully recyclable. It can be returned to nature without any reprocessing - possibly directly on site. Loam construction is also reflected in several of the 17 Sustainable Development Goals (SDGs) to which the German government has committed itself [4].

1.2 Loam noise barrier

Depending on the type of loam and the construction method - both influencing the composition of the building material - the produced loam has a density between 1500 and 2200 kg/m³ and is thus in the range of concrete. To ensure stability and take into account initial weathering (i.e. the reduction of the wall thickness towards a balanced condition), loam noise barriers would be much thicker than conventional walls made of concrete, wood, glass or aluminium cassettes. Thus, depending on their height, they could have a thickness of 1 m at their base and still have a thickness of 0.5 m at their upper edge.

2 Acoustic condition of cob

Compliance with the minimum sound insulation of 25 dB required in Germany according to ZTV-Lsw 06 [3] is not a problem due to the density of the loam and the thickness of a loam noise barrier. With regard to its absorption properties, the surface of a wall in conventional stamped or cob construction represents a sound-hard surface, as initial measurements on the impedance measuring tube have shown, i.e. incident sound is largely reflected. In order to improve the sound absorption properties, the use of admixtures, incorporated absorption layers and facing shells is being investigated.

2.1 Improvement of the absorption capacity

The sound-reflecting surface of cob and stamped loam severely restricts its use as a noise barrier. Efforts have therefore been made to establish sound absorbing properties in a loam noise barrier. Initial attempts to improve the absorption properties by increasing the proportion of straw did not lead to any significant improvement. With the following three alternatives (see Figure 1), this goal seems to be achievable. Either an absorbent layer can be incorporated during the construction of the loam noise barrier (Figure 1 left and centre) or it can be attached to the loam noise barrier like a facing layer after its construction (Figure 1 right).

The three samples that instantly provided promising absorption curves in the impedance measuring tube are:

- Loam with reed stalks (see Figure 1, left)
- Cob with incorporated Heraklith (see Figure 1, centre)
- Hemp lime (see Figure 1, right)

The results of these investigations are shown and discussed below.

Figure 1: Test specimen, from left to right: loam with reed stalks, cob with inserted Heraklith slab, hemp lime.

2.2 Specimens for the impedance measuring tube

The samples examined were made directly to size (diameter 10 cm) or taken from larger blocks with a core drill. The height of the samples varied and can be taken from Table 1.

Table 1: Dimensions of the examined samples with a diameter of 10 cm

	Sample height /cm
Loam with reed stalks	10,5
Heraklith 2,5 cm x 5 cm	11
Hemp lime	11, 6.4 and 4

Loam with reed stalks

Pure loam is used here. It has a doughy consistency and, above all, no stones or other additives. Thus, reed stalks can be worked into this mass perpendicular to the surface without damage. As a result, the round openings of the reed stalks, which can vary in diameter, remain on the surface. Due to the open stalks, an open-pored structure of the surface is achieved. Absorption inside the stalks and interference effects due to phase-rotated reflection at the end of the channel can occur here and be responsible for a reduction in reflection.

Cob with Heraklith

Heraklith consists of wood wool and mineral binders. Heraklith is used in residential construction for thermal insulation and is available in panels. The Heraklith used in the sample is 2.5 cm thick and was worked into the cob in a 5 cm deep strip. Due to the open porosity of the material, sound can enter the structure and its energy can be absorbed.

Hemp lime

Hemp lime is made from hemp hurds (crushed soft wood from the inner core of the hemp stalk) and lime. The result is a very porous and open-pored product, which strongly reminds of the open-pored asphalt used in road construction. Sound can easily enter the structure, interfere and be absorbed. Hemp lime is very light and can be incorporated directly when building a loam noise barrier.

3 Absorption measurements in the impedance measuring tube

The measurements on the test specimens were carried out in an impedance measuring tube to accommodate specimens with a diameter of 10 cm and the heights listed in Table 1. In an impedance measuring tube, two

microphones are used to measure the sound absorption capacity of a test specimen, which is exposed to sound from a loudspeaker in the opposite direction to the specimen. The sound absorption capacity can be calculated from the difference between the incident sound energy on the sample and the sound energy reflected by it.
Figure 2 shows the absorption curves of the sample "loam with reed stalks" (blue) and the sample with incorporated Heraklith panels (red) in comparison to the initial absorption capacity of the pure cob sample (purple). Here, the clay with reed stalks shows an absorption maximum of 0.6 at about 800 Hz. The Heraklith plate incorporated into the reflecting cob shows a first low maximum of a bit more than 0.4 at about 350 Hz and a higher and very broad maximum of about 0.65, which extends from 1000 Hz to 1200 Hz and shows an absorption coefficient of at least 0.6 already from 900 Hz and up to 1300 Hz.

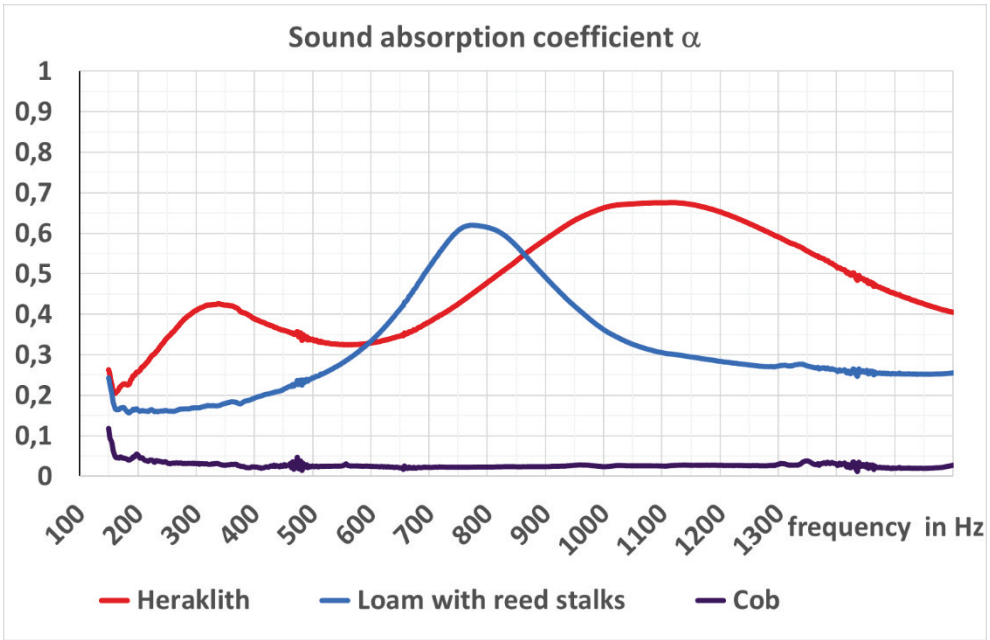


Figure 2: Sound absorption coefficients of cob, cob with Heraklith and loam with reed stalks measured in an impedance measuring tube.

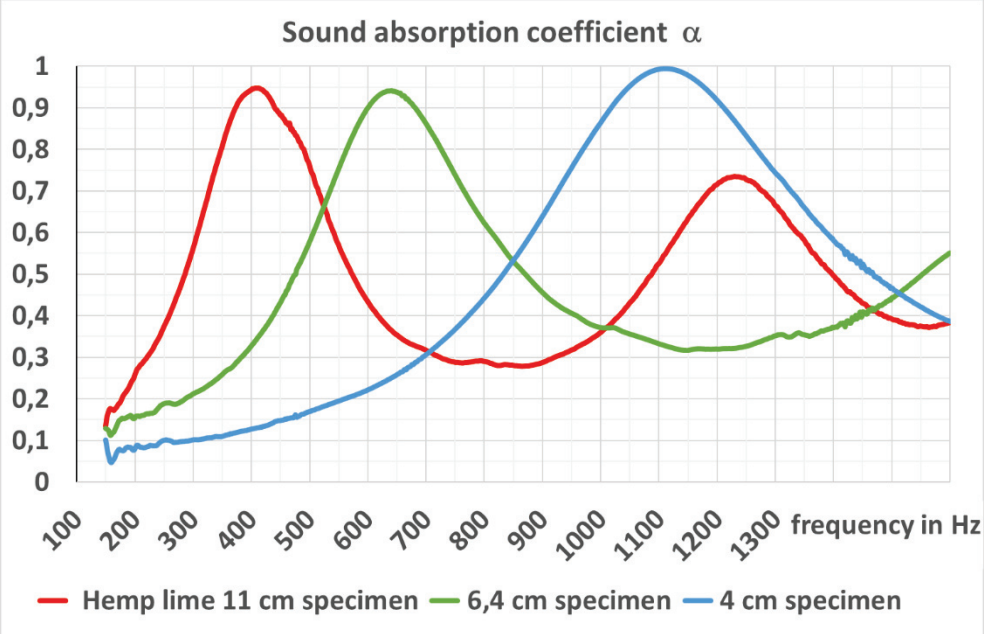


Figure 3: Sound absorption coefficient of hemp lime of different thickness (11 cm, 6.4 cm and 4 cm).

Figure 3 shows the absorption curves of hemp lime specimen. Starting with an 11 cm high sample (red), this was divided into a 6.4 cm (green) and a 4 cm (blue) sample to adjust the absorption maximum. The 11 cm specimen shows its absorption maximum of 0.95 at 300 Hz. For the lower samples, a clear shift of the absorption maximum of also 0.95 for the 6.4 cm sample to about 650 Hz and for the 4 cm sample to about 1100 Hz can be seen.

4 Conclusions

After the initial poor absorption values of samples made of cob or stamped loam, significant improvements could be achieved by incorporating reed stalks, integrating Heraklith and using hemp lime. The findings of the first series of measurements are now being used to produce new test specimens, which are to run through a further optimisation process.
Further improvements in the absorption behaviour are to be achieved through variations in the mixture composition and in the absorption layer thickness. Here, the phenomenological model according to Hamet [5] might be used for optimisation. According to Hamet, the shape factor, porosity, flow resistance and depth of a material determine the position and characteristics of the absorption maxima. The flow resistance, the form factor and the sample thickness have a main influence on the position of the maxima. Halving the sample thickness, for example, shifts a maximum to twice the frequency. The porosity in turn has an influence on the height and width of the maxima. These parameters can be used to optimise the adaptation to the traffic noise spectrum.
The aim is to achieve a highly absorbent surface structure in terms of acoustics and thus create the possibility of sustainable noise protection.

References

[1] DEGES projects, URL: https://www.deges.de/projekte/projekt/a-14-abschnitt-4-wittenberge-bis-as-karstaedt/

[2] Brandenburgs Alhambra. Zentrum für Peripherie, Nebelin und Düsseldorf , URL: https://zentrum-fuer-peripherie.org/en/

[3] Zusätzliche Technische Vertragsbedingungen und Richtlinien für die Ausführung von Lärmschutzwänden an Straßen (ZTV Lsw 21), ISBN: 3-939715-10-7 FGSV-Nr.: 258 (2006)

[4] Federal Government, Global Sustainability Strategy, URL: https://www.bundesregierung.de/breg-de/themen/nachhaltigkeitspolitik/nachhaltigkeitsziele-verstaendlich-erklaert-232174

[5] Hamet, J.F., Modelisation acoustique d'un enrobe drainant, Rapport inrets No 159, Octobre 1992, ISBN 2-85782-360-6

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Measurement of the acoustic effectiveness of diffractors

Fabio Strigari*, Ralf Becker, Wolfram Bartolomaeus

Section Environmental Protection, Immissions, Federal Highway Research Institute, Bergisch Gladbach, Germany.

*strigari@bast.de

Abstract

One possible approach for innovative noise protection from road traffic noise is to use the diffraction principle to cause an upward diffraction of the sound. Technically, this involves periodic grid structures with resonance chambers of different depths. These produce a reduction effect that extends over a broadband frequency range. Diffractors can be embedded in the side space next to the road or mounted on a noise barrier.

In the present study, the acoustic effectiveness of two diffractor types was determined by means of controlled and statistical pass-by measurements. In order to ensure the best possible quantification of the noise reduction potential, the measurements were carried out on two times four microphones, simultaneously at the diffractor and at a reference. An equivalent noise barrier of the same height or the propagation over free field served as reference. The simultaneous measurements at the diffractor and the reference allow a direct comparison of the same emission source with almost identical emission strength and characteristic. The evaluation of the level differences shows a significant and height-dependent additional reduction effect by the diffractors. The analysis of the one-third octave bands illustrates the working principle of the diffractors and their optimisation for the road traffic noise spectrum.

In addition to classifying the acoustic effectiveness of the investigated diffractors, the general suitability of the measurement principle is also discussed. The results moreover provide an important input for possible approaches to consider diffractors in the calculation of sound propagation along roads, as presented at Euroregio/BNAM 2022 in the article "Simulation of the sound field behind diffractors".

Keywords: innovative noise protection, sound diffraction, pass-by measurements

1 Introduction

The possibility to reduce road traffic noise by exploiting the effect of diffraction has already been known for a long time in theory. For instance, in the EU-project HOSANNA (Holistic and sustainable abatement of noise by optimized combinations of natural and artificial means), conducted from 2009 to 2013, periodic grid structures with resonance chambers were discussed.

The working principle of such noise reducing devices is based on the upward diffraction of sound. A resonance of the sound waves occurs in the grooves of the diffractor. These grooves have different depths and are thus tuned to different frequencies, so that in total a broadband effect is created. The resonances create a change in impedance making the horizontally propagating sound waves experience a resistance and deflecting them upwards. The sound pressure level is thereby reduced at shallow angles of propagation, while it is increased at greater angles.

A first successful transfer of technology from theory to practical application is realized by the WHIS® product series of the Dutch company 4Silence, which allows to make use of the diffraction effect alongside roads and railways in situations where other mitigation measures cannot be realised or to reduce the necessary height of a noise barrier.

1.1 Description of the diffractor types under investigation

In the present study two types of diffracting elements were investigated, namely the so-called WHIS®stone and WHIS®wall.

The WHIS®stone is a structured concrete slab that is installed in the banquet next to the road. The dimensions of a single element are specified as 1020 x 980 mm with a height of 168 mm [1]. The weight is about 400 kg. The specified frequency range for noise reduction is 800 – 1200 Hz. The reduction of the sound pressure level for a single row installation is 2.5 dB according to the product description. For a double row arrangement (not examined here), the reduction is given as 4.0 dB.

The WHIS®wall is a combination of a low concrete noise barrier with a steel diffractor on top. Other material combinations are possible. The total structure is 1.11 m high and the maximum depth (given by the diffracting top) is 1.05 m [2]. The weight is about 1000 kg/m. The frequency range for noise reduction is given as 400 – 2000 Hz. According to product description, the achievable reduction of the sound pressure level is of the order of 7 – 9 dB, corresponding to the effect of a 3 m high noise barrier.

1.2 Existing studies

Measurements at the slab diffractor were carried out in 2013 and 2014 in the Netherlands, on the N 413 near Soesterberg and on the N 314 near Hummelo [3]. At distances of 7.5 m and 15 m from the closest lane in four different heights, sequential measurements were taken in the middle of three 100 m long sections (reference section, one and two diffracting rows), using the statistical pass-by (SPB) method. The results show a clear and significant upward bending of the sound. As expected, this leads to a sound reduction at low heights and to a sound increase at the higher microphone positions.

The diffractor wall was also experimentally examined in Soesterberg in 2019 [4]. Two microphone arrays, each consisting of four microphones, were used to determine the effect of the diffractor. The microphones were placed at heights of 1.2 m, 2.0 m, 3.0 m, and 4.0 m. Two measurement sections were considered: a reference section without any protection measure and a section with diffractor. Again, the SPB method was used, but both measurement sections were measured simultaneously (i.e. the pass-by of the same vehicle was recorded at both the reference and the diffractor). Although only a limited number of valid vehicles could be recorded, the results give a clear picture of the diffracting effect. In summary, the noise reductions recorded behind the diffractor (in a distance of 7.5 m) are $\gtrsim$ 9 dB at 1.2 m height, $\approx$ 9 dB at 2.0 m height, $\gtrsim$ 6 dB at 3.0 m height and $\gtrsim$ 3 dB at 4.0 m height.

2 Measurements

Controlled pass-by (CPB) measurements took place at the "Technology Base" in Enschede, Netherlands. The site is located at Twente Airport. Two test tracks were available for the investigations. On one of the test tracks, the WHIS®stone is installed in a single row over a length of approximately 25 m, followed by a longer reference section with grassland. On the other test track, a WHIS®wall (1 m total height) over a length of about 50 m is installed. This is followed by a conventional noise barrier of the same height and on the other end by a longer reference section with grassland. Both test tracks are situated on flat ground and the width of the roadway is 2 x 3.5 m.

The measuring site for the SPB measurements was located along the federal road B 25 near Nördlingen, Germany. The 2 x 3.5 m wide roadway lies at an embankment and has a slight left curve leading away from a roundabout. In the 300 m long section, two 100 m long strips of the slab diffractor (single row arrangement) were installed, separated by a 100 m long area without modification of the banquet (reference section). In the front segment, the diffractor stones were rotated by 180°, in the rear segment the intended orientation (with the drainage facing the road) was used.

2.1 Measurement setup

The following figures show the measurement setup used for the pass-by measurements. For the investigation of slab diffractor, both for the CPB in Enschede and the SPB in Nördlingen, the pass-by level was determined at a distance of 7.5 m from the far lane (microphone positions M1 and M2) and at 7.5 m distance from the near lane (microphone positions M3 and M4) at a height of 1.2 m and 2.4 m above the level of the road (cf. Figure 1). The microphones were placed 15 m in front of and behind the transition between diffractor and reference section. The speed measurement system was located directly at the transition point. In contrast to the common right-hand traffic situation in the case of the SPB measurements, the same direction of travel was chosen for both lanes for the CPB runs (as indicated by the yellow arrows in Figure 1).

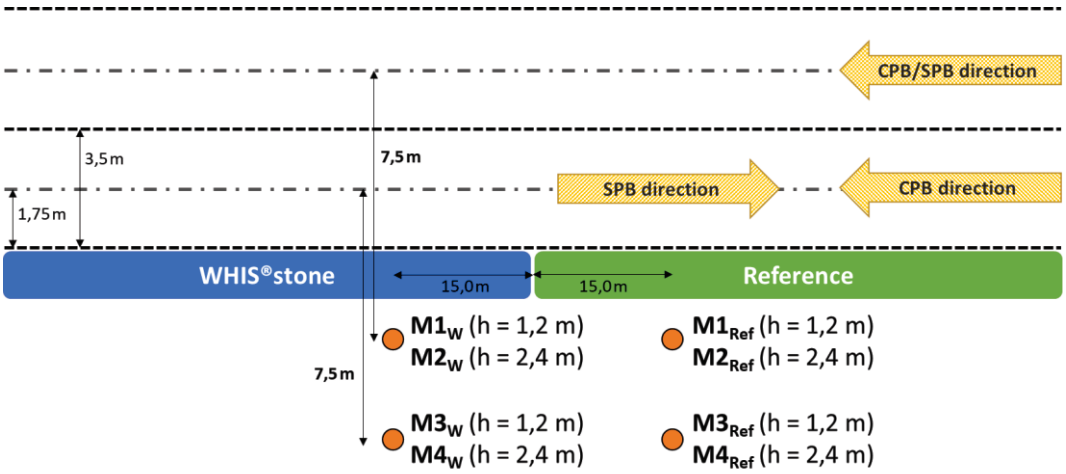


Figure 1: Measurement setup for the pass-by measurements on the slab diffractor (not to scale). Reference was free field sound propagation over grassland (ground reflections are not excluded).

The setup for the measurements at the diffractor wall differed slightly from the situation at the slab diffractor (cf. Figure 2). Here, one microphone was located directly between the first lane and the diffractor/reference (microphone position M1). Behind that, three microphones were situated at a distance of 5 m from the first microphone at heights of 1.2 m, 2 m and 3 m (microphone positions M2, M3 and M4). This choice was motivated, among other things, by the intention to achieve comparability with the measurement results from Ref. [4]. As above, the controlled pass-by always occurred from the same direction.

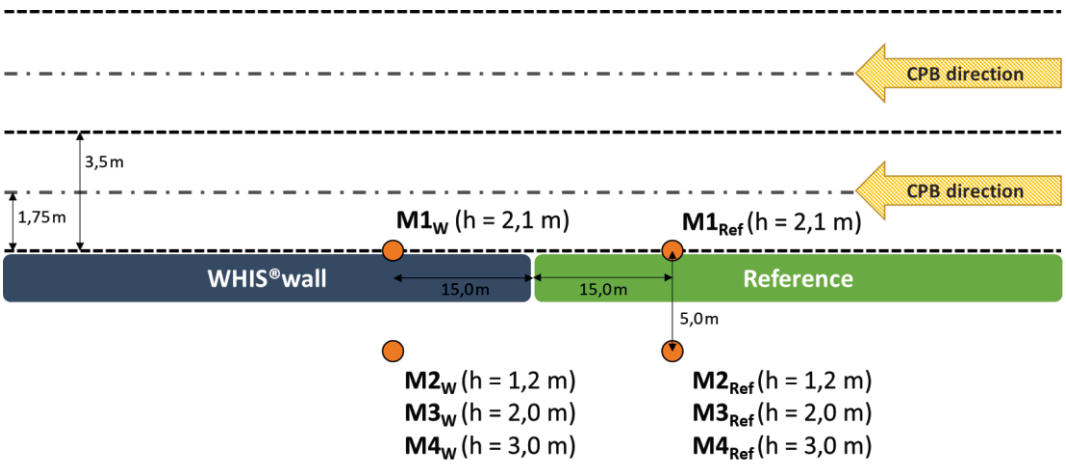


Figure 2: Measurement setup for the pass-by measurements on the diffractor wall (not to scale). Reference was free field sound propagation over grassland (ground reflections not excluded) and a conventional low noise barrier of the same height.

2.2 Measurement procedure

The CPB measurement campaign comprised three comparisons: slab diffractor vs. free field, diffractor wall vs. conventional noise barrier and diffractor wall vs. free field. The expression "free field" here is deliberately used for the reference sound propagation over grassland (ground reflections are not excluded). Each comparison consisted of 18 measurements (9 per lane) at speeds of 60 km/h, 70 km/h and 80 km/h (three measurements each). In all cases, the sound level of the same passing car was recorded simultaneously at both cross sections (i.e. diffractor and reference, distance to each other = 30 m).

Each time, the car was accelerated well outside the measurement cross-section and kept at the desired speed via automatic cruise control, so that at the moment of the pass-by always the same emission conditions prevailed (approximately). The speed and meteorology were also recorded. It was slightly cloudy and the wind speed was less than 5 m/s throughout the day. The air temperature was in the range of 19°C to 21°C.

The SPB measurement campaign was carried out on two days: one day the regular orientation of the diffractor stones was examined, on the second day the rotated version was investigated. Vehicles in both directions of travel were recorded. The free field sound propagation always served as reference. The pass-by sound levels of individual passing cars as well as light and heavy trucks were measured simultaneously at both measurement cross-sections. The data from the speed measurement system and the meteorology were fed into both measurement systems. These data include the passing speeds, the distance of the vehicles to the speed measurement system, wind speed and direction and air temperature.

On the first day, the active measurement time was about 6 h. It was cloudy and the wind speed was less than 5 m/s throughout the day. The air temperature was in the range of 15°C to 18°C. On the second day, the active measurement time was about 4.5 h. It was overcast and the wind speed was partly more than 5 m/s from midday onwards, so that single SPB runs had to be dropped as invalid. The air temperature ranged from 14°C to 19°C.

3 Results

In all pass-by measurements, the A-weighted individual sound levels of the respective measurement cross sections were determined simultaneously at the time of the maximum level at the loudest microphone position. For this point in time, also the third-octave band spectra were recorded and stored for further evaluation.

At the test tracks for the CPB, the signal-to-background ratio met the 10-dB criterion in all runs. The surrounding was very quiet, so that background noise played almost no role. The SPB results were analysed by first assigning the raw data of the respective cross sections to each other via measurement time, speed and distance. Implausible measurements and those that did not fulfil the 6-dB criterion (regarding the acoustic separation of two successive vehicles) were discarded.

To quantify the acoustic effectiveness of the investigated diffractors, the sound level differences $\Delta L1_{max,i}$, $\Delta L2_{max,i}$, $\Delta L3_{max,i}$ und $\Delta L4_{max,i}$ (running index i) between the corresponding microphones on diffractor and reference are calculated. The level differences are also determined for the recorded one-third octave bands (100 Hz to 20 kHz). The arithmetic mean is calculated for the total of n = 18 passings and for the vehicle passing on the near or far lane (n = 9) only. The empirical standard deviation is used as a measure of dispersion for the calculated mean values. The SPB data are analysed statistically by using common methods.

3.1 Slab diffractor

Figure 3(a) and (b) show the measured level differences [slab diffractor – reference (free field)] for the CPB on the near and far lane as box plots. The number of evaluated passings for one lane is n = 9. Each box represents one of the microphone positions, as depicted in Figure 1. Table 1 summarises the corresponding mean values and standard deviations.

The level reduction at position M1 (front-bottom, h = 1.2 m) is significant compared to the free field reference. On average, it amounts to -1.7 dB for the vehicle passing on the near lane and -1.4 dB for the far lane.

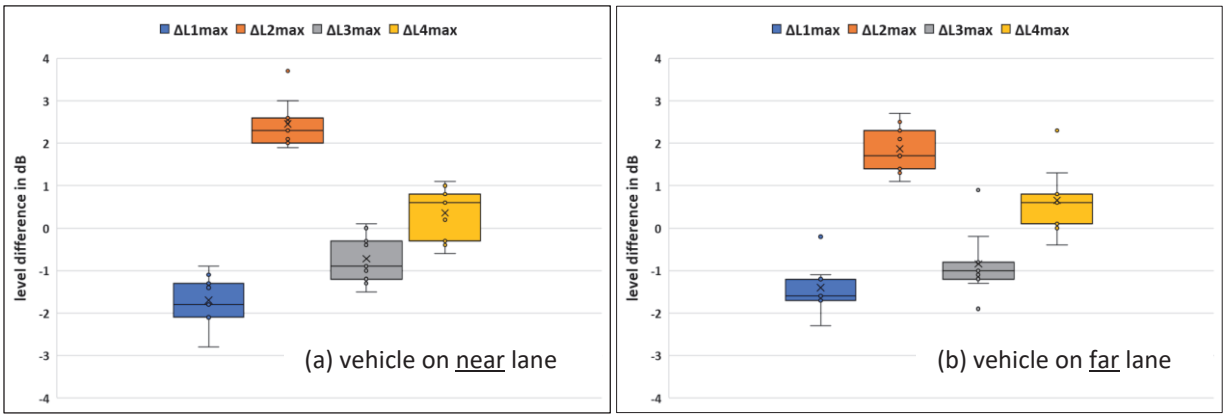


Figure 3: Box plots of differences of maximum CPB sound levels [slab diffractor – reference (free field)] in dB for the vehicle passing (a) on the near lane and (b) on the far lane at the microphone positions M1, M2, M3 and M4

Table 1: Mean values $\Delta\bar{L}i_{max}$ and standard deviation σ of the CPB level differences [slab diffractor – reference (free field)] in dB at the microphone positions M1, M2, M3 and M4; averaged over vehicles passing on the near lane, the far lane and over all lanes.

	$\Delta\bar{L}i_{max}$ **near**	σ	$\Delta\bar{L}i_{max}$ **far**	σ	$\Delta\bar{L}i_{max}$ **all**	σ
M1	-1.7	0.6	-1.4	0.6	-1.6	0.6
M2	2.5	0.6	1.9	0.6	2.2	0.6
M3	-0.7	0.7	-0.8	0.6	-0.8	0.8
M4	0.4	0.7	0.7	0.7	0.5	0.8

At M3 (rear-bottom, h = 1.2 m) a level difference of -0.7 dB (near) and -0.8 dB (far) can be observed. On the contrary, front position M2 at h = 2.4 m shows a significant level increase, namely 2.5 dB and 1.9 dB for the near and far lane, respectively. At the rear-top position M4, the detected level increase is only about 0.5 dB. The measurements generally show a good reproducibility and the standard deviation for all mean values is of a similar order of magnitude (0.6 dB to 0.8 dB).

The measurement results clearly illustrate the working principle of the diffractor. The sound energy is "redistributed" from bottom to top, so that a level reduction occurs for low emission angles, while the sound level increases accordingly for high emission angles. Averaged over all measurements (see Table 1), the maximum reduction effect at the lower height amounts to -1.6 dB.

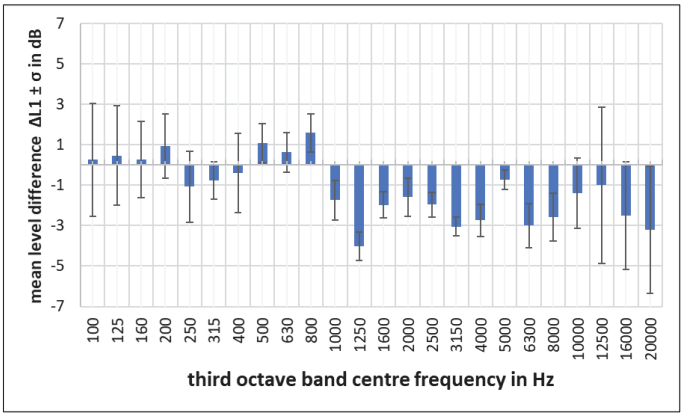


Figure 4: Mean value and standard deviation of spectral differences of maximum CPB sound levels [slab diffractor – reference (free field)] in dB for vehicle passing on the near lane at M1

The frequency dependence of the level differences (third-octave band analysis) is shown in Figure 4 for the front-bottom microphone M1 and passings on the near lane. The reduction effect of the diffractor is mainly effective in the frequency range above 800 Hz. Regarding the distant lane, the reduction already starts above 500 Hz (not shown), but – as expected – somewhat less pronounced. At measurement position M2 (front-top), a corresponding effect becomes evident in the opposite direction: for most frequency bands and especially above 800 Hz, a level increas occurs, i.e. the upward deflection of the sound is also noticeable here.

It is also worth mentioning that significant level reductions were also measured at the upper microphones M2 and M4 in the third-octave bands of 630 Hz and 800 Hz (approximately 2 to 3 dB).

In general, all trends and statements concluded above from the CPB measurements are confirmed by the results of the SPB measurement campaign at Nördlingen. Figure 5 shows the level differences [slab diffractor – reference (free field)] for passenger cars on the near lane as violin diagrams: The black cross indicates the mean value $\Delta\bar{L}i_{max}$, the white circle indicates the median ("P50") of the distribution. The blue box represents the range from the first ("P25") to the third quartile ("P75"), containing 50 % of all data. The red lines draw the range of the 5 % quantile ("P05") to the 95 % quantile ("P95"). 90 % of all data lie in this range. In Figure 6 the same representation is used for the data resulting from heavy trucks passing on the near lane.

The number of valid vehicles was well above 100 in both cases. At the positions M1 and M3, the mean level reduction compared to the reference is 2.3 dB and 5 dB, respectively. The mean level increase at the position M2 amounts to 2.1 dB. At the position M4 the level increase is only 0.3 dB. The significant level differences are lower in absolute value by up to 0.1 dB with respect to the respective mean values (statistical significance level = 0.05).

The statements obtained for passenger cars tend to remain valid for heavy trucks (and for light trucks as well, not shown here). However, the diffractor has a somewhat lower effect on trucks than on passenger cars. The mean level reduction at M1, for example, is only 0.8 dB for heavy trucks. Likewise, the level increase at the position M2 is also lower. For heavy trucks, the value there is 1.3 dB.

Table 2 summarises and compares the SPB results for passenger cars, light and heavy trucks. Both the level reductions and the level increases – resulting from the SPB data – are significantly lower for the far lane. This is also in agreement with the findings from the CPB measurements.

As explained in 2.1 and 2.2, on the second day of the SPB measurement campaign a rotated installation of the slab diffractor was investigated. The diffractor stones were rotated by 180° in this alternative arrangement. The resulting level differences (not shown here) are quite similar. What is striking, however, is the difference at the front-bottom position for passenger cars passing on the near lane. Instead of -2.2 dB in the normal arrangement, with the rotated diffractor stones the measurements yield only -1.0 dB level difference compared to the reference (free field).

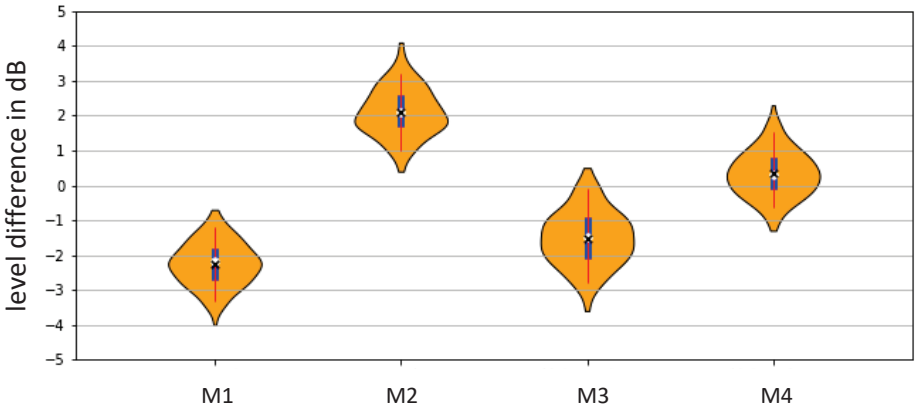


Figure 5: Violin plots of differences of maximum SPB sound levels [slab diffractor – reference (free field)] in dB for passenger cars passing on the near lane at the microphone positions M1, M2, M3 and M4

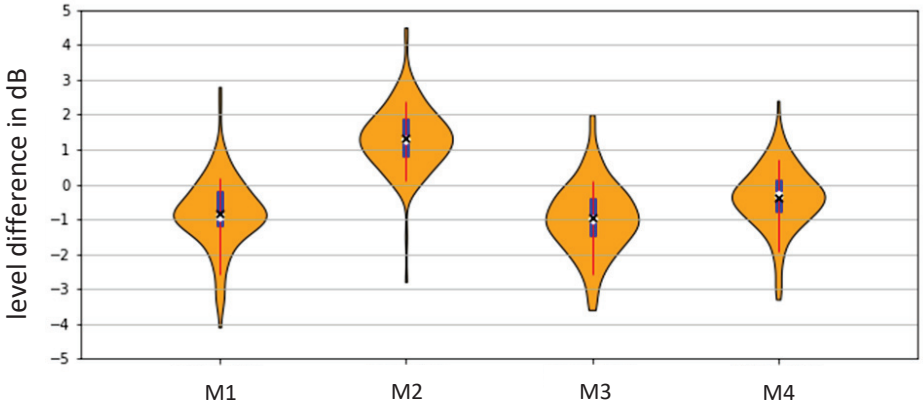


Figure 6: Violin plots of differences of maximum SPB sound levels [slab diffractor – reference (free field)] in dB for heavy trucks passing on the near lane at M1, M2, M3 and M4

Table 2: Significant level differences [slab diffractor – reference (free field)] in dB for passenger cars, light and heavy trucks in the near and far lane at the measurement positions. The last row gives the number N of valid vehicles used in the statistical analysis.

	near lane			far lane		
	passenger cars	light trucks	heavy trucks	passenger cars	light trucks	heavy trucks
M1	-2.2	-0.8	-0.7	-0.7	-0.1	-0.6
M2	2.0	1.4	1.2	1.4	1.2	0.7
M3	-1.5	-0.6	-0.9	-0.2	0.0	-0.5
M4	0.2	0.3	-0.3	0.3	0.2	-0.1
N	158	39	152	46	17	115

3.2 Diffractor wall

Figure 7(a) and (b) show the measured level differences [diffractor wall - reference (noise barrier)] for the CPB on the near and far lane as box plots. The number of evaluated passings for the near lane is n = 10, for the far lane n = 9. Please note that the measurement positions differ from the setup at the slab diffractor (cf. Figure 2). In Table 3 the corresponding mean values and standard deviations can be found.

Microphone M1 is located directly at the first lane in front of the diffractor. Here, the measured level differences vanish within the measurement accuracy, i.e. the emission levels at the diffractor and the reference noise barrier match as expected. At a height of 1.2 m and 2.0 m (M2 and M3), a considerable level reduction can be observed behind the diffractor compared to the reference. Since the total height of the diffractor wall and the reference noise barrier are the same, the measurement reflects the additional reduction effect coming from the diffractor. The values are -5.3 dB and -3.9 dB at M2 and -4.4 dB and -2.1 dB at M3 for the near and far lane, respectively (see also Table 3). Again, the height dependence of the level reduction is clearly recognisable. At h = 3.0 m, the reduction regarding the close lane is only -1.0 dB, and for the distant lane no reduction is measured any more ($\Delta L4_{max}$= 1.2 dB).

These CPB measurements also show a good reproducibility. The standard deviation is slightly higher for distant lane measurements at the two upper microphones M3 and M4 – thus, the degree of scatter of the individual measurements is somewhat greater here.

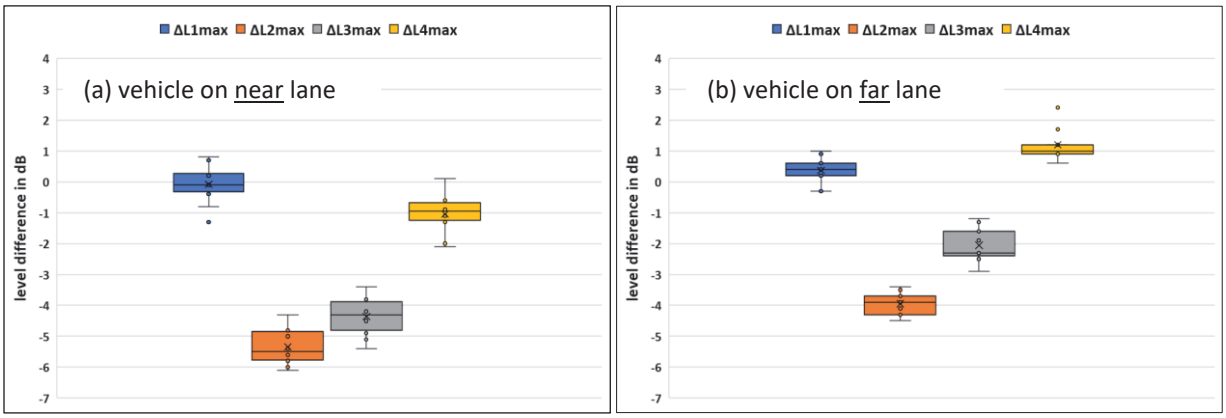


Figure 7: Box plots of differences of maximum CPB sound levels [diffractor wall – reference (noise barrier)] in dB for the vehicle passing (a) on the near lane and (b) on the far lane at the microphone positions M1, M2, M3 and M4

Table 3: Mean values $\Delta\bar{L}i_{max}$ and standard deviation σ of the CPB level differences [diffractor wall – reference (noise barrier)] in dB at the microphone positions M1, M2, M3 and M4; averaged over vehicles passing on the near lane, the far lane and over all lanes.

	$\Delta\bar{L}i_{max}$ **near**	σ	$\Delta\bar{L}i_{max}$ **far**	σ	$\Delta\bar{L}i_{max}$ **all**	σ
M1	-0.1	0.6	0.4	0.5	0.1	0.6
M2	-5.3	0.6	-3.9	0.4	-4.7	0.9
M3	-4.4	0.6	-2.1	1.1	-3.3	1.3
M4	-1.0	0.7	1.2	0.9	0.0	1.3

The third-octave band analysis of the level differences is shown in Figure 8 for M2 (h = 1.2 m) and close passings. The spectral range of the reduction effect of the diffractor is rather broad. Except for 315 Hz and 400 Hz, a level reduction is measurable in all third octave bands. The greatest effect can be seen in the range between 630 Hz and 1250 Hz (more than 7 dB for 630 Hz and 800 Hz). The results for the higher microphone M3 show a similar behaviour, but in a somewhat weaker form (not shown). Here, the largest detected reduction of -5.9 dB is in the 800 Hz band. Especially at the highest microphone M4 for the vehicle passing on the distant lane, there are also frequency-dependent level increases (not shown).

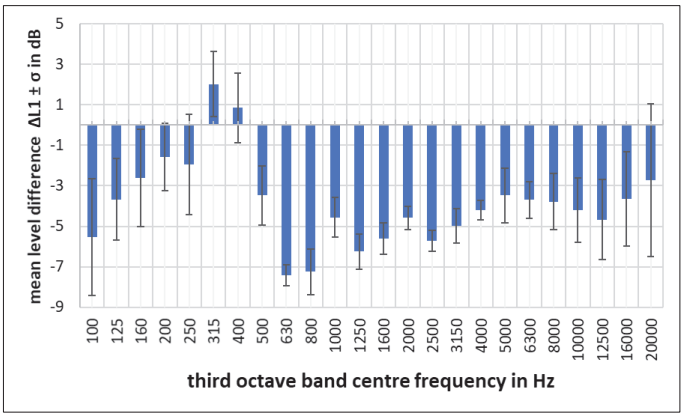


Figure 8: Mean value and standard deviation of spectral differences of maximum CPB sound levels [diffractor wall – reference (noise barrier)] in dB for the vehicle passing on the near lane at the microphone position M2

The results of the CPB at the diffractor wall in comparison to free field as reference are summarised in Table 4. The measured level differences are a measure for the total insertion loss of the diffractor wall (i.e. shielding plus diffraction effect). In agreement to the measurements befor, there is a clear height-dependent level reduction behind the diffractor. The maximum effect at M2 (at a height of 1.2 m) amounts to -15.3 dB for the near lane and -12.3 dB for the far lane. With increasing height (M4), the detected level reduction decreases to 5.1 dB (near) and 0.2 dB (far).

Table 4: Mean values $\Delta\bar{L}i_{max}$ and standard deviation σ of the CPB level differences [diffractor wall – reference (free field)] in dB at the microphone positions M1, M2, M3 and M4; averaged over vehicles passing on the near lane, the far lane and over all lanes.

	$\Delta\bar{L}i_{max}$ **near**	σ	$\Delta\bar{L}i_{max}$ **far**	σ	$\Delta\bar{L}i_{max}$ **all**	σ
M1	0.5	0.5	0.1	0.5	0.3	0.5
M2	-15.3	0.5	-12.3	0.4	-13.8	1.6
M3	-11.8	0.7	-6.1	0.7	-9.0	3.0
M4	-5.1	0.4	0.2	0.8	-2.4	2.8

3.3 SEL analysis

In a further analysis routine, the sound level over time is considered to calculate the sound exposure level (SEL). For this purpose, a fixed aperture angle of 120° is assumed. The SEL is equivalent to the sound energy reaching the respective microphones from this angular range. The corresponding time window is centred at the point in time of the maximum sound level. The windows size is of the order of 0.5 s – 2.3 s, depending on the pass-by velocity and the distance between the microphone and considered lane. This analysis approach is applied to the CPB data only, in order to assess the quality of the maximum sound level analysis.

Since the scope of the present manuscript is limited, only the resulting level differences [slab diffractor – reference (free field)] on the near lane are depicted in Figure 9. Qualitatively, the conclusions about the acoustic effectiveness of the diffractor remain the same as in the analysis via the maximum pass-by levels. However, when comparing to Figure 3, one can see (i) that the range of scatter of the SEL results is smaller, and (ii) that the mean level reduction at the lower microphones and level increase at the higher microphones become less in their absolute value. These two effects are observed also for the CPB measurements at the diffractor wall (not shown) and for passings on the far lane (not shown), whereas here the changes (ii) in the absolute values of the level differences are less prominent than on the near lane.

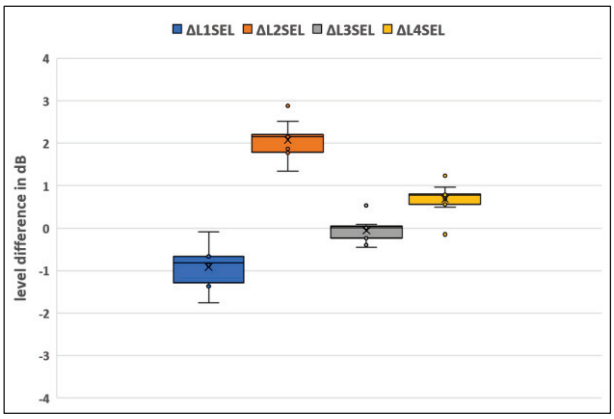


Figure 9: Box plots of differences of SEL sound levels [slab diffractor – reference (free field)] in dB for CPB on the near lane at the microphone positions M1, M2, M3 and M4

4 Conclusions

All CPB/SPB results indicate a significant acoustic effectiveness of the investigated sound diffractors. The reduction effect is generally lower for vehicles passing on the far lane than for a close pass-by. For the slab diffractor the level difference for low emission angles and at the microphones close to the road amounts to about -2 dB. According to the diffraction principle, an increase in the level difference (compared to free field) is detected at higher emission angles. For trucks the reduction effect turns out to be weaker than for cars.

The diffraction effect is also clearly evident in the CPB results for the diffractor wall. The combination of a low noise barrier with an attached diffractor shows a significantly higher acoustic effectiveness compared to the reference noise barrier of the same height. Additional level differences down to -5.3 dB are measured at 5 m from the first lane. Compared to the free field reference, level differences down to -15.3 dB are recorded.

The frequency dependence confirms the behaviour seen in the total level difference. Moreover, the one-third octave band analysis shows that the slab diffractor has its centre of action mainly in the frequency range above 800 Hz (or above 500 Hz for the distant lane). The effective frequency range of the diffractor wall is broader.

The differences between the SEL and maximum pass-by level analysis do not surprise from a physical point of view: The working principle of the diffractor is optimised for perpendicular sound incidence. Sound contributions reaching a receiver point from a certain aperture angle do not experience the same diffraction effect. Therefore, a superposition of sound coming from different directions leads to a weaker total effect. Nevertheless, considering the maximum pass-by levels still has a high validity, as contributions from angles close to perpendicular incidence are energetically dominant.

The measurement principle applied here has proven to be well-suited for the quantification of the acoustic effectiveness of sound diffractors. The simultaneous measurement of two cross sections allows a direct comparison of the same emission source with a well-defined reference and yields more meaningful and relevant results than doing a before-and-after measurement, which typically are prone to weather-related influences and traffic-induced differences.

Based on these promising measurements, the questions arise how the results can be used to draw conclusions about the acoustic effectiveness at greater distances and how the acoustic effect of diffractors can be considered in a more general calculation of sound propagation from roads. A first approach to do so has been realized by modelling the diffractors as a strip with a frequency-dependent acoustic impedance and using an impedance jump model for calculating the sound propagation. This is presented by Bartolomaeus *et al.* in the article "Simulation of the sound field behind diffractors" [5], which is kindly recommended to all interested readers.

Acknowledgements

I would like to thank the company 4Silence for the kind admission at its headquarters on Twente airport and giving us the possibility to carry out measurements on site.

References

[1] WHIS®stone product description. *https://www.4silence.com/whisstone/*, last access: March 07, 2022.

[2] WHIS®wall product description. *https://www.4silence.com/whiswall/*, last access: March 07, 2022.

[3] J. Hooghwerff, W. J. van der Hejiden, H. F. Reinink and Y. H. Wijnant. Whisstone, a sound diffractor: does it really affect traffic noise?. *Proceedings of EuroNoise 2015*, 2015.

[4] Y. H. Wijnant and Z. A. Lok. Metingen Soesterberg. *Measurement protocol University of Twente*, 2019.

[5] W. Bartolomaeus, R. Becker and F. Strigari. Simulation of the sound field behind diffractors. *Proceedings of Euroregio/BNAM 2022*, 2022.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Simulation of the sound field behind diffractors

Wolfram Bartolomaeus*, Ralf Becker, Fabio Strigari

Section Environmental Protection, Immissions, Federal Highway Research Institute, Bergisch Gladbach, Germany.

*Bartolomaeus@bast.de

Abstract

In analogy to optical reflection gratings, acoustic diffractors cause an upward diffraction of the sound due to their periodic structure. They can be recessed in the roadside space or mounted on a noise barrier. The prerequisite for their effectiveness, the coherence of the incident sound, is fulfilled for passing vehicles at small angular differences and not too high frequencies. By using chambers of different depths, a spectral broadband effect of the diffractors is achieved.

Based on the results presented at Euroregio/BNAM 2022 in the article "Measurements of the acoustic effectiveness of diffractors", the question arises how the acoustic effect of diffractors could be taken into account in the calculation of sound propagation from roads. For this purpose, the propagation of traffic noise over flat terrain described by Rasmussen already in 1982 was extended with an impedance jump road - grassland and applied to the situation with diffractors. By comparing the sound signals of pass-by measurements recorded at different positions, it should be possible to model the acoustic effect of an additional strip of diffractors placed next to the road by a corresponding strip with a frequency-dependent acoustic impedance.

The transformation of the results obtained here into the purely energetic approach of ISO 9613-2 for sound propagation outdoors is certainly not easy. An indirect possibility would be the introduction of a frequency-dependent vertical directivity of the sound sources, which takes the effect of the diffuser into account.

Keywords: diffractor, propagation.

1 Introduction

In 2019 we first noticed the product WHIS®stone of the Dutch company 4Silence. The possibility to reduce road traffic noise by a periodic grid structures with resonance chambers was already known, e.g. from the EU-project HOSANNA (Holistic and sustainable abatement of noise by optimized combinations of natural and artificial means) conducted from 2009 to 2013. But with WHIS®stone a product is available to be practically used for noise abatement at minor roads.

In the supplementary regulation for noise in the Environment Act, published in the Government Gazette of the Netherlands on March 26th 2021 [1], the diffractor is already considered. In the appendix IV the acoustical effect of the diffractor is defined and how to integrate such a device in the Dutch calculation method. But how to include a diffractor in a more general noise propagation method like ISO 1913-2 [2]?

2 Theory

The theory outlined here is for the purpose of understanding the underlining principles in modelling the sound propagation above ground. It is far from completeness.

2.1. Impedance

The impedance of a surface like a ground or a wall $\underline{Z}_W$ is defined as quotient of sound pressure $\underline{p}$ and velocity $\underline{v}$ at the surface

$$\underline{Z}_W := \frac{\underline{p}}{\underline{v}}\bigg|_{surface} \tag{1}$$

The acoustical behaviour of ground is characterised best by the normalised admittance $\underline{\beta}$ as the quotient of the characteristic acoustic impedance of the fluid $Z_0 = \rho_0 c_0$ with density ρ_0 and velocity c_0 and the surface impedance $\underline{Z}_W$

$$\underline{\beta} := \frac{Z_0}{\underline{Z}_W} \tag{2}$$

In the sound propagation model NORD2000 [3] the flow resistivity of ground is classified in seven impedance classes, based on a measuring procedure of Nordtest [4].

In Figure 1 real and imaginary part of the normalised admittance are shown for impedance classes A to G of NORD2000 in the frequency range from 50 Hz to 10 kHz, following the Darcy model, described by Taraldsen in [5].

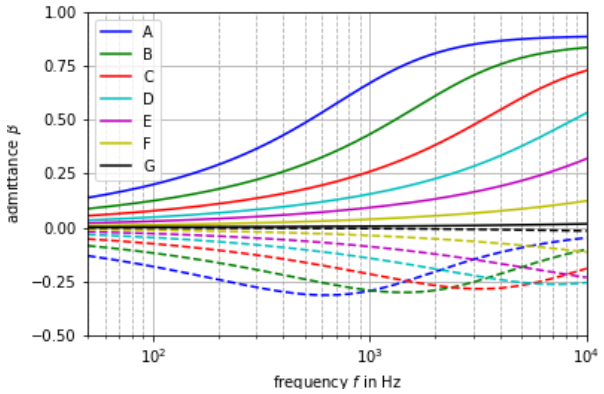


Figure 1: normalised admittance for impedance classes A to G of NORD2000 depending on frequency; solid line: real part, dashed line: imaginary part

The impedance is getting acoustically harder from class A to G. Accordingly, the admittance decreases and the maximum of real and imaginary part of the admittance is moving to higher frequencies. Class "C" is referred as "rough grassland" with flow resistivity of 100 KNs/m^4 while Class "G" corresponds to a hard surface of "dense asphalt" or "concrete" with flow resistivity of 20 000 KNs/m^4.

2.2. Noise propagation above surface

In [6] Nobile the sound propagation over an impedance plane is described. The problem can be solved in cylindrical coordinates. A point source is located in the height of h_Q above a plane with the normalised admittance of $\underline{\beta}$. A receiver is located in a distance of d in the height of h_E above this plane (see Figure 2).

From this, the radial distances between source and receiver, R_1, and between mirror source and receiver, R_2, can be determined.

$$R_1 = \sqrt{(h_E - h_Q)^2 + d^2} \tag{3}$$

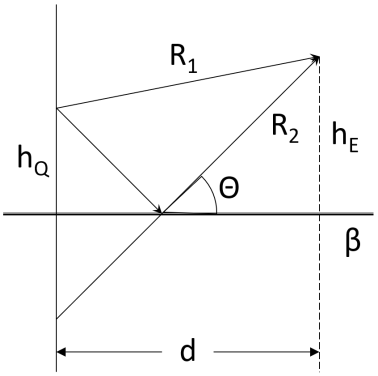


Figure 2: Geometry for the propagation above an impedance plane

$$R_2 = \sqrt{(h_E + h_Q)^2 + d^2} \tag{4}$$

The glancing angle Θ — complementary angle to $\pi/2$ — has the value

$$\Theta = \arctan\left(\frac{d}{h_E + h_Q}\right) \tag{5}$$

The velocity potential with angular wave number k above a local reacting impedance plane (impedance independent of incident angle) is describe by summation of the velocity potentials $\underline{\Psi}_0(R) = nicefrac{e^{ikR}}{R}$ of two spherical sound waves

$$\underline{\Psi}(R_1, R_2) = \underline{\Psi}_0(R_1) + \underline{Q}\underline{\Psi}_0(R_2) = \frac{e^{ikR_1}}{R_1} + \underline{Q}\frac{e^{ikR_2}}{R_2} \tag{6}$$

The spherical reflection factor $\underline{Q}$ is defined by the reflection factor $\underline{R}_p$ of a plane wave and the ground wave function $\underline{F}$

$$\underline{Q} := \underline{R}_p + \left(1 - \underline{R}_p\right)\underline{F} \tag{7}$$

For a sound source near the ground the curvature of the spherical wave has a big influence on reflection. In some cases there is no destructive interference but a ground wave is starting to propagate near the ground.

The ground wave function is defined as

$$\underline{F} := 1 + \mathrm{i}\sqrt{\pi}\underline{w}e^{-\underline{w}^2}\,\underline{\mathrm{erfc}}\left(-\mathrm{i}\underline{w}\right) \tag{8}$$

with the numerical distance

$$\underline{w} := \sqrt{\frac{\mathrm{i}kR_2}{2}}\left(\cos\theta + \underline{\beta}\right) \tag{9}$$

and the complementary error function $\underline{\mathrm{erfc}}$.

The sound pressure of an harmonic point source can be derived from the time deviation of the velocity potential, equation (6), multiplied by the density of the fluid.

In Figure 3 the sound pressure level over an impedance plane of class ”C” of NORD2000 at 1000 Hz is shown for a range of 400 m width and 400 m height.

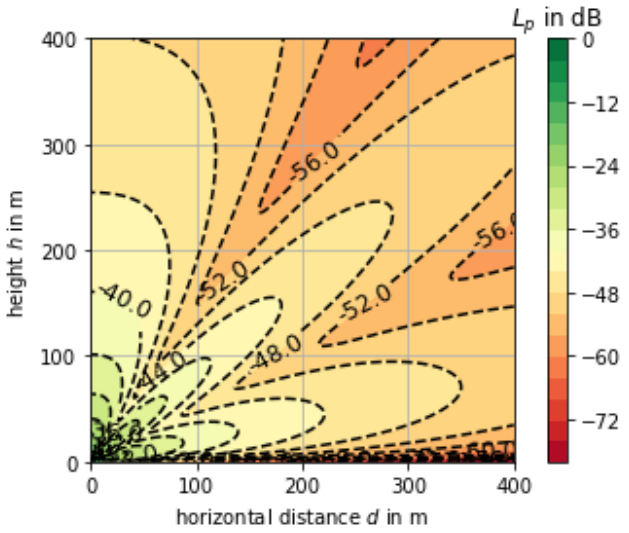


Figure 3: Sound pressure level over an impedance plane of class ”C” at 1000 Hz

The source height was 0.5 m and the source strength was normalised to 0 dB at 1 m distance for free field conditions.

Depending on the impedance of the ground, radial structures of level minima and maxima, originating at the source, occur in the sound field.

2.3. Model with an impedance jump

Rasmussen [7] described a method, developed by de Jong [8] for sound propagation over a plane with an impedance jump. The geometry is displayed in Figure 4.

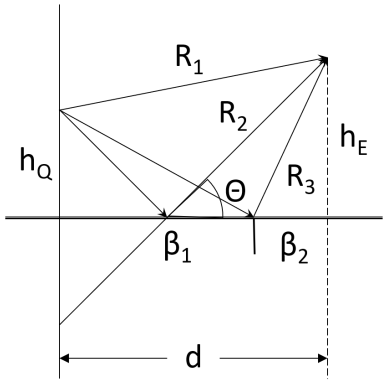


Figure 4: Geometry of sound propagation over an impedance jump

The specular (direct) reflection occurs at the front half plane with admittance β_1.

Beside direct sound and reflection at the impedance plane, corresponding to equation (6) an additional correction term for the velocity potential of the reflection from the impedance jump is introduced

$$\underline{\Psi}_{S\pm} = \frac{e^{-\mathrm{i}\pi/4}}{\sqrt{\pi}} \left[\frac{e^{\mathrm{i}kR_1}}{4\pi R_3} \underline{H}\left(k\sqrt{R_3 - R_1}\right) \pm \frac{e^{\mathrm{i}kR_2}}{4\pi R_3} \underline{H}\left(k\sqrt{R_3 - R_2}\right) \right] \tag{10}$$

The auxiliary function $\underline{H}(x)$ can be expressed by the complex complementary error function

$$\underline{H}(x) = \frac{\sqrt{\pi}}{2} \frac{1+\mathrm{i}}{\sqrt{2}} \underline{\mathrm{erfc}}\left(\frac{1-\mathrm{i}}{\sqrt{2}} x\right) \tag{11}$$

There are two times two cases for the equations. Reflection in the front (1) or in the rear (2) half plane. Impedance of the front half plane is acoustically harder (<) or softer (>) than the rear half plane. For these four cases the equations are

$$\begin{aligned}
1<: \quad \underline{\Psi}_{S,1<} &= \underline{\Psi}_0(R_1) + \underline{Q}_1 \underline{\Psi}_0(R_2) + \left(\underline{Q}_2 - \underline{Q}_1\right) \underline{\Psi}_{S+} \\
1>: \quad \underline{\Psi}_{S,1>} &= \underline{\Psi}_0(R_1) + \underline{Q}_1 \underline{\Psi}_0(R_2) - \left(\underline{Q}_2 - \underline{Q}_1\right) \underline{\Psi}_{S-} \\
2<: \quad \underline{\Psi}_{S,2<} &= \underline{\Psi}_0(R_1) + \underline{Q}_2 \underline{\Psi}_0(R_2) + \left(\underline{Q}_2 - \underline{Q}_1\right) \underline{\Psi}_{S-} \\
2>: \quad \underline{\Psi}_{S,2>} &= \underline{\Psi}_0(R_1) + \underline{Q}_2 \underline{\Psi}_0(R_2) + \left(\underline{Q}_2 - \underline{Q}_1\right) \underline{\Psi}_{S+}
\end{aligned}$$

In Figure 5 the sound pressure level over an impedance plane of class ”G” until 2 m and class ”C” behind at 1000 Hz is shown for a range of 5 to 50 m width and 1 to 5 m height.

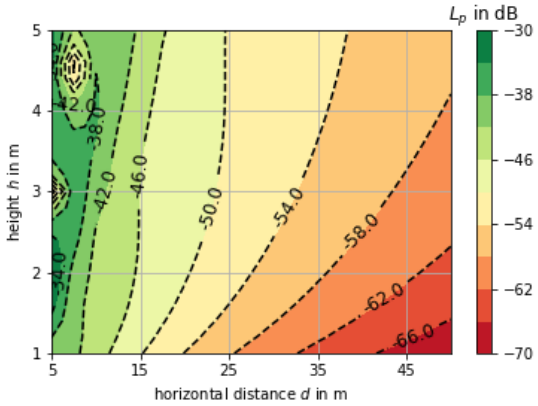


Figure 5: Sound pressure level over an impedance plane of class ”G” until 2 m and class ”C” behind at 1000 Hz

The source height was 0.5 m again and the source strength was normalised to 0 dB at 1 m distance for free field conditions.

At long distance and low height the sound level takes its minimum. At short distance the sound level gets maximum values. This can be assumed as a typical sound field for the radiation of sound from traffic on a road over grassland nearby.

2.4. Modell with two impedance jumps

Hothersall [9] has extended the method of de Jong for a stripe with a different impedance included in the impedance plane. This method can be expanded further for two impedance jumps where all areas are having different impedances. The geometry is shown in Figure 6.

Beside direct sound and specular reflected sound, according equation (6) with a correction for reflection on the first impedance jump according to equation (10) an additional correction term $\underline{\Psi}_{S2\pm}$ for the reflection at the second impedance jump is introduced. For this, in equation (10) the difference between the spherical reflection factors of the two first impedances $\left(\underline{Q}_2 - \underline{Q}_1\right)$ is substituted by the difference of the two last impedances $\left(\underline{Q}_3 - \underline{Q}_2\right)$ and instead of the distance R_{S1} the distance R_{S2} is used.

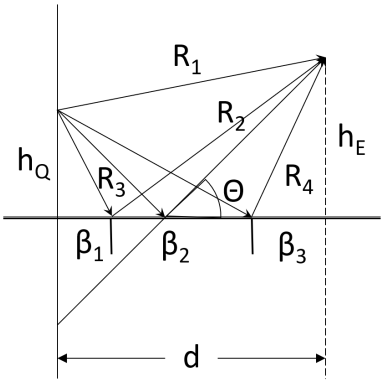


Figure 6: Geometry for sound propagation over an plane with two impedance jumps

Depending on whether the specular reflection occurs in section 1, 2 or 3 and which admittances β_1, β_2 and β_3 in these sections are existent:

$$\begin{aligned}
&\text{a:} \quad |\beta_1| \leq |\beta_2| \leq |\beta_3| \\
&\text{b:} \quad |\beta_1| \leq |\beta_2| \wedge |\beta_2| > |\beta_3| \\
&\text{c:} \quad |\beta_1| > |\beta_2| \wedge |\beta_2| \leq |\beta_3| \\
&\text{d:} \quad |\beta_1| > |\beta_2| > |\beta_3|
\end{aligned}$$

there are twelve different cases:

$$\begin{aligned}
\text{1a:}\ \underline{\Psi}_{S,1<<} &= \underline{\Psi}_0(R_1) + \underline{Q}_1\underline{\Psi}_0(R_2) + (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1+} + (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2+} \\
\text{1b:}\ \underline{\Psi}_{S,1<>} &= \underline{\Psi}_0(R_1) + \underline{Q}_1\underline{\Psi}_0(R_2) + (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1+} - (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2-} \\
\text{1c:}\ \underline{\Psi}_{S,1><} &= \underline{\Psi}_0(R_1) + \underline{Q}_1\underline{\Psi}_0(R_2) - (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1-} + (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2+} \\
\text{1d:}\ \underline{\Psi}_{S,1>>} &= \underline{\Psi}_0(R_1) + \underline{Q}_1\underline{\Psi}_0(R_2) - (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1-} - (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2-} \\
\text{2a:}\ \underline{\Psi}_{S,2<<} &= \underline{\Psi}_0(R_1) + \underline{Q}_2\underline{\Psi}_0(R_2) + (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1-} + (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2+} \\
\text{2b:}\ \underline{\Psi}_{S,2<>} &= \underline{\Psi}_0(R_1) + \underline{Q}_2\underline{\Psi}_0(R_2) + (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1-} - (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2-} \\
\text{2c:}\ \underline{\Psi}_{S,2><} &= \underline{\Psi}_0(R_1) + \underline{Q}_2\underline{\Psi}_0(R_2) + (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1+} + (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2+} \\
\text{2d:}\ \underline{\Psi}_{S,2>>} &= \underline{\Psi}_0(R_1) + \underline{Q}_2\underline{\Psi}_0(R_2) + (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1-} - (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2-} \\
\text{3a:}\ \underline{\Psi}_{S,3<<} &= \underline{\Psi}_0(R_1) + \underline{Q}_3\underline{\Psi}_0(R_2) + (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1-} + (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2-} \\
\text{3b:}\ \underline{\Psi}_{S,3<>} &= \underline{\Psi}_0(R_1) + \underline{Q}_3\underline{\Psi}_0(R_2) + (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1-} + (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2+} \\
\text{3c:}\ \underline{\Psi}_{S,3><} &= \underline{\Psi}_0(R_1) + \underline{Q}_3\underline{\Psi}_0(R_2) + (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1+} + (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2-} \\
\text{3d:}\ \underline{\Psi}_{S,3>>} &= \underline{\Psi}_0(R_1) + \underline{Q}_3\underline{\Psi}_0(R_2) + (\underline{Q}_2 - \underline{Q}_1)\,\underline{\Psi}_{S1+} + (\underline{Q}_3 - \underline{Q}_2)\,\underline{\Psi}_{S2+}
\end{aligned}$$

3 Simulation

The equations of the models without, with one and with two impedance jumps were coded in Python. The verification of the model without a jump was done by comparison with analytical results for simple configurations like e.g. a single sound source on hard ground.

The verification of the computer code for one jump was done by comparison with the transfer grid method developed by Rasmussen [7]. In this method the radiation from the source is received at 20 000 grid points in a plane above the impedance jump. All these points are acting as sources, radiating the sound to the receiver at the other side of the impedance jump.

In Figure 7 the differences in sound level between transfer grid method and de Jong model for one impedance jump at 1 kHz are displayed for distances from 5 to 50 m and for heights from 1 to 5 m. The source height is

positioned 0.5 m above the origin. The impedance of the plane in front of the jump is of class ”G” and behind the jump at a distance of 2 m of class ”C” like in Figure 5.

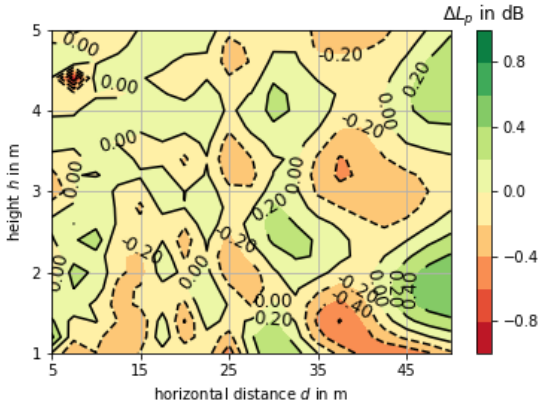


Figure 7: Differences in sound level for the simulation models with one jump at 1 kHz

In the simulation of the transfer grid method 200 points in width and 100 points in height were used. The grid constant was set to 0.2 times the wavelength. Hence the dimensions of the grid plan were about 10 m by 6 m.

The level differences are almost in a range of $\pm$ 0.2 dB except the area near the ground far away from the source and one small spot very high above the source, where strong interferences occur.

The model for two jumps was verified by using the same admittance on both sides of one of the jumps and comparing the results with the ones obtained by the model with one jump.

3.1. Description of the Model

The geometry of the simulation model for comparison with measurements, explained in the next section, is shown in Figure 8.

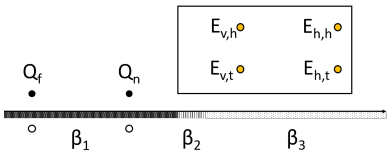


Figure 8: Geometry for sound propagation over a plane with one or two impedance jumps; comparison with measurements

The sound source, a passenger car, was located far ”Q_f” or near ”Q_n” the impedance jump. The admittance area before the jumps is denoted as ”β_1”, between the jumps as ”β_2” and after the jumps as ”β_3”. With ”$E_{v,t}$”, ”$E_{v,h}$”, ”$E_{h,t}$” and ”$E_{h,h}$” the microphone positions ”ahead” (v), ”behind” (h) and ”low” (t), ”high” (h) are marked, respectively.

3.2. Measurements

The acoustic effectiveness of two diffractor types was determined by means of controlled and statistical pass-by measurements. In order to ensure the best possible quantification of the noise reduction potential, the measurements were carried out on two times four microphones, simultaneously at the diffractor and at a reference. For comparison with simulations only the maximum sound levels from controlled pass-by measurements at Twente Airport near Enschede where used, because of the flat terrain there. For details see Strigari [10].

3.3. Comparison with measurements

The admittance for the stripe between the jumps, β_2, is not known. But from measurements with and without the stripe at the four microphone positions, the values for different frequency bands can be determined. For this purpose, simulations with different values for the admittance of the strip were conducted for all third octave bands. To compare the results with the measurements at the four microphone positions, the results of different frequencies were added up energetically.

In Figure 9 the comparison simulations for different admittances at the strip for the third octave band of 1 kHz are shown.

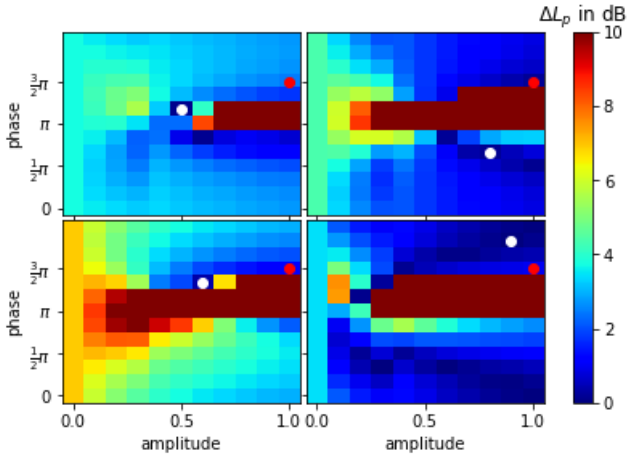


Figure 9: Comparison of four measurements with simulations for different admittances at the strip for the third octave band of 1 kHz

The complex admittance of the strip is parametrised with its two polar coordinates in the complex plane, the phase from 0 to 2π and the amplitude from 0 to 1. The level differences between simulation and measurements for the four microphone positions are shown in Figure 8. The individual minimum is marked by a white dot, the overall minimum for all positions together by a red dot. In this way for all third octave bands the parameters for an optimal fit were calculated. In Figure 10 the obtained parameter are displayed as normalised admittances (cyan lines) together with the ones for impedance classes ”C” and ”G” of NORD2000.

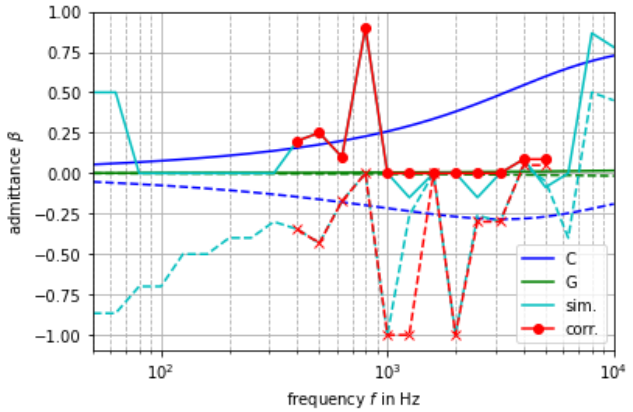


Figure 10: normalised admittance for impedance classes ”C” and ”G” of NORD2000 and for the simulation depending on frequency; solid line: real part, dashed line: imaginary part

There are some strange effects visible due to uncertainties in measurements and fitting process. These where corrected (red lines). Fore some frequencies, the results are even not jet satisfying.

3.4. Long range propagation

With the parameters of the normalised admittance β, obtained within the comparison of measurements and simulations, it is possible to calculate the long-range sound level.

The total level calculation from the third octave sound pressure levels $L_{p,i}$ was done with the normalised traffic noise spectrum [11] $\Delta L_{p,i}$ of third octave bands i:

$$L_t = 10 \lg \left[\sum_i 10^{0.1(L_{p,i} + \Delta L_{p,i})} \right] \tag{12}$$

From the difference of simulation with and without the diffractor the total noise reduction ΔL_p for the whole spectrum can be found. The calculation starts from 400 Hz and some third octave bands, namely 1250, 2500 and 5000 Hz were left out for reasons of faulty simulation results.

In Figure 11 the difference in total level ΔL_p is calculated for distances from 5 to 50 m and for heights from 1 to 5 m. In this case the diffractor had a width of 1 m and was placed 1.65 m from the near lane of the road, positioned at a distance of 0 m with a height of 0.5 m.

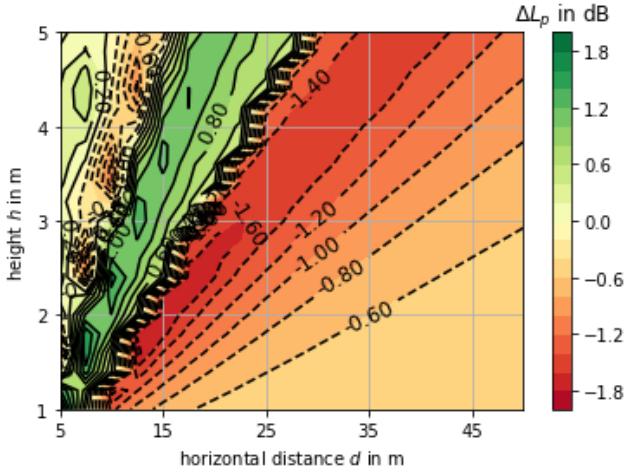


Figure 11: Total sound level difference for a diffractor of 1 m width for the near lane

The noise reduction is up to 2 dB for a sector with an angle of about 7° away from the sound source.

4 Conclusions

It is possible to simulate the noise propagation over impedance jumps with a set of compact mathematical formulations. This can used as a tool to integrate new noise abatement measures with physical frequency-dependant behaviour into engineering models of sound propagation like ISO 9613-2. The way to do this is using analogous frequency-dependant models for sources and barriers.

The noise reduction for WHIS®stone found near the road is conserved at greater distances. The research about acoustical effectiveness will continue, also for the products WHIS®wall and WHIS®top.

Acknowledgements

I would like to thank the company 4Silence for the kind admission at their headquarter on Twente airport and giving us the possibility for measurements at their site.

References

[1] Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. Regeling van de staatssecretaris van infrastructuur en waterstaat en de minister van binnenlandse zaken en koninkrijksrelaties van 19 maart 2021 tot wijziging van de omgevingsregeling vanwege het opnemen van regels met het oog op de beheersing van geluid afkomstig van wegen, spoorwegen en industrieterreinen (aanvullingsregeling geluid omgevingswet), mar 2021. URL https://zoek.officielebekendmakingen.nl/stcrt-2021-15868.html#volgtOpmain_reg0001_d3088e1_cmp_IVA.

[2] ISO 9613-2. Acoustics - attenuation of sound during propagation outdoors - part 2: General method of calculation, December 1996.

[3] B. Plovsing and J. Kragh. Comprehensive outdoor sound propagation model. part 1: Propagation in an atmosphere without significant refraction. Technical report, DELTA, Danish Electronics, Light & Acoustics, Lyngby, 2001.

[4] *NT ACOU 104 - Ground surfaces: Determination of the acoustic impedance*, 1999.

[5] G. Taraldsen. The delany-bazley impedance model and darcy's law. *Acta Acust united Ac*, 91:4150–4165, 2005.

[6] M. A. Nobile and S. I Hayek. Acoustic propagation over an impedance plane. *J. Acoust. Soc. Am.*, 78(4): 1325–1336, 1985.

[7] K. B. Rasmussen. Propagation of road traffic noise over level terrain. *J. Sound Vib.*, 82(1):51–61, 1982.

[8] B. A. de Jong. Reekenmodellen om de invloed te voorspellen van en asfalt-grasovergang en van afscherming van een wal. Technical report, Technical University of Delft, Laboratorium voor Technische Natuurkunde, 1978.

[9] D. C. Hothersal and J. N. B. Harriott. Approximation models for sound propagation above multi-impedance plane boundaries. *J. Acoust. Soc. Am.*, 97(2):918–926, 1985.

[10] F. Strigari, R. Becker, and W. Bartolomaeus. Measurement of the acoustic effectiveness of diffractors. In *Euregio / BNAM2022*, 2022.

[11] EN 1793-3. Road traffic noise reducing devices - test method for determining the acoustic performance - part 3: Normalised traffic noise spectrum, November 1997.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Shore power connection for offshore vessels – Measured noise reduction in port and dock

Bernt Mikal Larsen[1,*]

[1]Department of Acoustics, Multiconsult Norway, Kristiansand, Norway.

*Bernt.Mikal.Larsen@multiconsult.no.

Abstract

The presentation will summarize measured noise reduction for drilling, pipelay and offshore support vessels due to shore power connection. The noise level at office or storage building in distance of 100-250 m from the vessels without significant influence from other background noise has been registered continuously through days and nights. Noise levels at nights without and with shore power connection have been compared. A drilling vessel had L_w 110 dB and a noise reduction due to shore power connection of 18 dB. A pipelay vessel had L_w 105 dB without shore power, and a noise reduction of 15 dB due to the effect of shore power. Both drilling and pipelay vessel seem to have L_w of 90-92 dB with shore power connection. Two different offshore support vessels have also been measured with sound power level L_w of 107 dB and 100 dB. The noise reduction with shore power connection is 13 dB for the first and 6-7 dB for the latter. With shore power connection, both offshore suppert vessels have L_w of 94 dB. The two offshore support vessels show that shore power can give different noise reduction for ships with different design, and that the sound power level with shore power connection seems to be the same within a category of vessels, due to the fact that ship noise level from ventilation and fans seems to be the same. All of the vessels measured have sound power level between 90 and 95 dB with shore power connection.

1 Introduction

Noise from ships in port has been measured at Kongsgård-Vige from drilling and pipelay vessel, and noise from ships in dock has been measured from offshore support vessels at Andøya Industripark in Kristiansand in Norway. This document will compare the measured noise level with and without shore power connection. The method used will be explained, and also the considerations for the estimated noise reduction.

2 Hypothesis and method for estimating noise reduction due to shore power

The basic hypothesis is that shore power connection of a vessel will give a constant noise reduction compared to the situation without shore power. At daytime and evening there will normally be many sources and activities which contributes to a measured noise level. The best method for determining the effect of shore power should therefore be to compare noise levels at nights, with and without shore power. Nights without heavy rain and without strong wind have to be used, for background noise being as low as possible. The method

for determining effect of shore power is a continuous registering of noise level in port/dock area, for both nights with and without shore power connection.

The measurement method was to use one microphone position with instrument Norsonic N140 at port building with low influence from other background noise. If more microphone positions were used, the issue of background noise in different positions had to be carefully analysed. The microphone on port building is placed in height not shielded by the vessel. In noise calculations vessels are normally modelled as point source with omni-directional sound pressure level (same directivity in all directions), therefore it is considered as good enough to have one microphone position (due to omni-directional sound radiation) where it is known that the level from other background noise is significant low. The main purpose is to compare the measured noise reduction at port building, which is explained by shore power when other background noise is low.

The sound power level without shore power is calculated by modelling the vessel as a point source, calibrated against the measured level at night at the port buildings. Compared with literature about sound power from ships and vessels, the estimated L_w from measurements shows very good correspondence with what should be expected. The sound power level with shore power is estimated by the level difference measured at the port building, with and without shore power. Calculations with Cadna using L_w with shore power show very good match with measured noise level at the port building. The Cadna model with different ships/vessels gives noise levels which corresponds very well with short time measurements in different directions within the port area and surrounding neighbourhood.

The two situations are shown in the following figures.

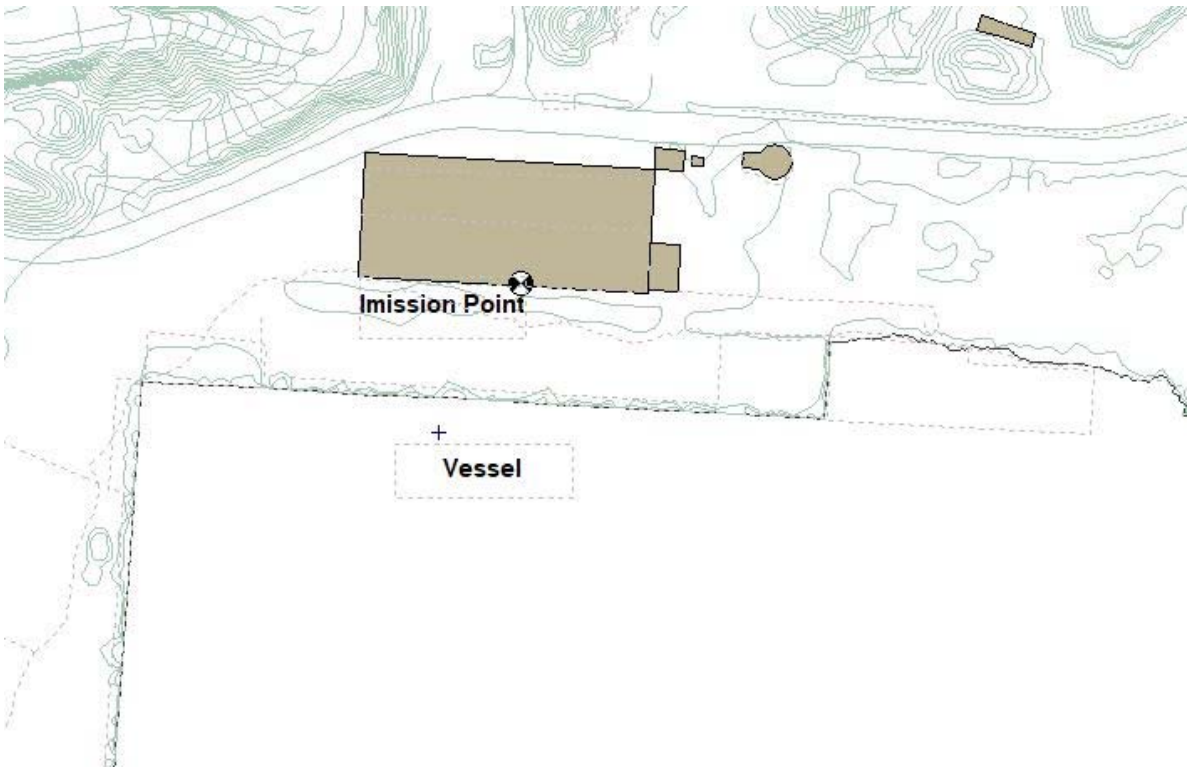


Figure 1a Vessel and imission point for noise registration at Kongsgård-Vige

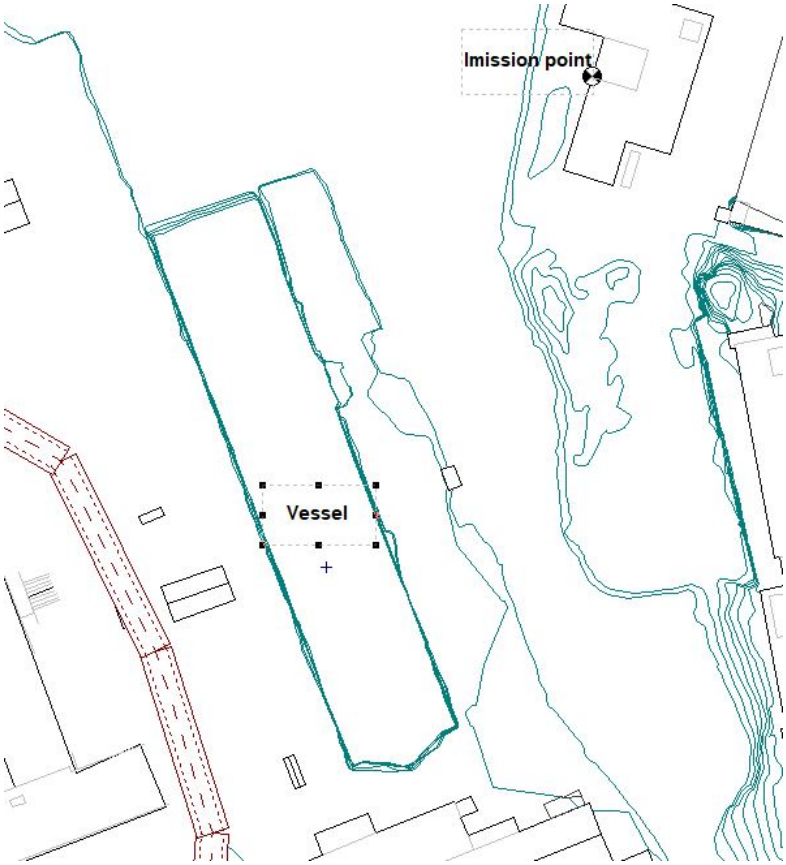


Figure 1b Vessel and imission point for noise registration at Andøya Industripark

3 Drilling and pipelay vessel – noise reduction due to shore power

Measurement of noise from drilling and pipelay vessel was done with continuously registering by a harbor building at Kongsgård-Vige. The noise levels at nights with and without shore power were compared.
The analysis of measurement data gives the following:

1. Drilling Vessel
    - L_w 110 dB without shore power
    - L_w 92 dB with shore power

2. Pipelay vessel
    - L_w 105 dB without shore power
    - L_w 90 dB with shore power

Picture 1 Drilling and pipelay vessel, Kongsgård in Kristiansand

Frequency spectrum shows significant reduction in all frequencies. Normally the annoyance from neighbours is related to low frequencies, where shore power connection shows very good effect. The L_w with shore power connection is dominated by noise from ventilation, fans and other noise sources on the ship and seems to be similar for the two vessels.

The pipelay vessel Lorelay was in port, with and without shore power. In figure 2 the blue curve is noise level without shore power, and the other curves show noise level with shore power connection. By comparing noise level at nights for a period in autumn 2021 and winter 2022 (only with contribution from Lorelay) we can see the noise reduction due to shore power connection. The measurement in January 2022 shows more than 15 dB reduction in a wide frequency range. Some of the measurements (with shore power) will also contain other background noise.

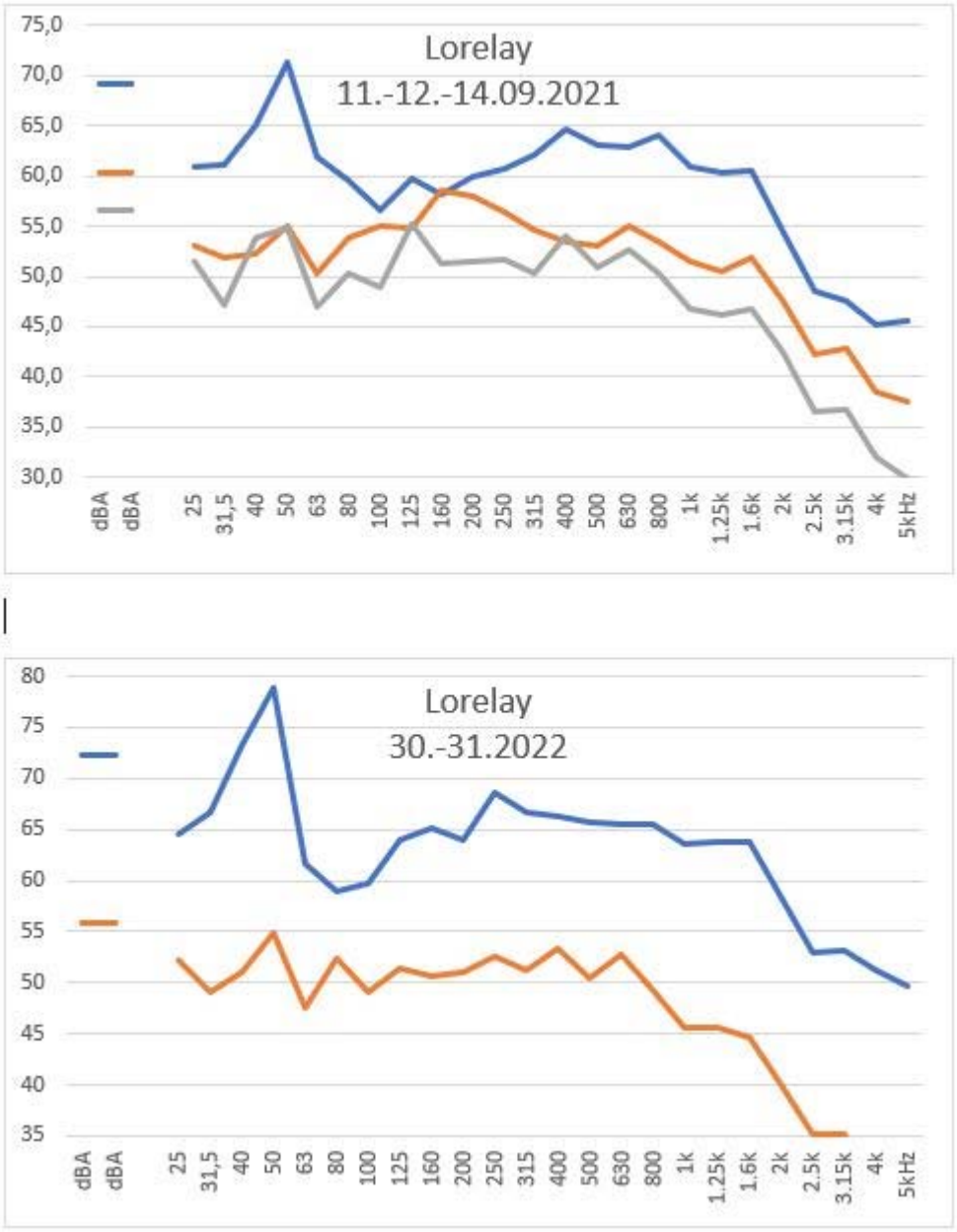


Figure 2 Frequency spectrum for the pipelay vessel Lorelay, with and without shore power, measured at port building

The noise levels from the two periods of Lorelay with shore power are summarized in the tables below.

Table 1 Pipelay vessel Lorelay – Noise level without and with shore power September 2021

Lorelay – September 2021							
Octaveband	**63 Hz**	**125 Hz**	**250 Hz**	**500 Hz**	**1000 Hz**	**2000 Hz**	**4000 Hz**
Without shore power	**72**	**63**	**66**	**68**	**67**	**62**	**51**
Shore power	**57**	**57**	**56**	**58**	**53**	**48**	**39**
Reduction	**15**	**6**	**10**	**10**	**14**	**14**	**22**

Table 2 Pipelay vessel Lorelay – Noise level without and with shore power January 2022

Lorelay – January 2022							
Octaveband	**63 Hz**	**125 Hz**	**250 Hz**	**500 Hz**	**1000 Hz**	**2000 Hz**	**4000 Hz**
Without shore power	**79**	**68**	**72**	**71**	**69**	**65**	**56**
Shore power	**57**	**55**	**56**	**57**	**52**	**46**	**36**
Reduction	**22**	**13**	**16**	**14**	**17**	**19**	**20**

Table 2 shows different and more noisy activity than table 1. The noise reduction due to shore power is good in all frequencies. It is to be noted that the noise level with shore power connection is quite similar in all octave bands for the two periods. This shows that in both cases the noise level is reduced to a similar level which is dominated by ventilation, fans or other noise sources on the vessel. The reason for the different level without shore power in September 2021 and January 2022 is not exactly known, but the sound power level with shore power is approximately the same for both measurements. Some of the nights in September shore power connection gave noise level according to the orange curve, the reason for this is assumed to be that noise sources due to testing or other activity gave higher level with shore power than expected (the grey curve in September). The noise data summarized for shore power in September last year is based on the grey curve.

4 Offshore support vessels – noise reduction due to shore power

Measurement of noise from offshore support vessels was done with continuously registering by an office building at Andøya Industripark. The noise levels at nights with and without shore power were compared. The analysis of measurement data gives the following:

Offshore Support Vessel

- L_w 100-107 dB without shore power
- L_w 94 dB with shore power

Picture 2 Offshore support vessel in dock

The difference in sound power level for the two different offshore support vessels is mainly explained by different design of these two vessels. Registration of noise level at the facade of office building was done for two different vessels in December 2021 and January 2022:

1. MMA Pinnacle
   - Noise level 61-62 dBA at nights 13th to 15th of December
   - Noise level 49-50 dBA at nights 20th to 21st of December, and most quiet periods noise level of 42-43 dBA

2. Normand Jarstein
   - Noise level 54-55 dBA night between 11th and 12th of January
   - Noise level 48-49 dBA night between 12th and 13th of January – with shore power

MMA Offshore Asia has confirmed that some of the nights MMA Pinnacle was in dock were without the harbor generators running. This explains the reduced noise level the night between 20th and 21st of December. It seems that both vessels give noise level of 49-50 dBA when harbor generators are off or the vessel is connected to shore power. The noise level of 48-50 dBA is assumed to be dominated by fans and ventilation on the vessels. The low noise level of 42-43 dBA from MMA Pinnacle in some quiet periods the night between 20th and 21st of December may be explained by that either fans or ventilation were turned off such periods.

The frequency spectrum for Normand Jarstein (with shore power connection in dock) is shown in the table below.

Table 3 Normand Jarstein – Noise level without and with shore power January 2022 measured at office building

Normand Jarstein – January 2022							
Octaveband	63 Hz	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz
Without shore power	59	74	53	52	51	46	40
Shore power	52	51	50	53	51	47	41
Reduction	7	23	3				

From the table we can see that the noise level by the office building is dominated by background noise from 250 Hz due to either weather/wind or ventilation/fans on the vessel. Since the level in the upper frequencies are quite similar with and without shore power, it is highly probable that the noise level from 250 Hz and above is due to a constant source like ventilation or fans. The measurement for Normand Jarstein shows very good effect of shore power in the low frequencies, which normally gives annoyance for the neighbourhood.

Because there is no sure knowledge of what installations were turned off for MMA Pinnacle, noise level from this vessel regarding frequency spectrum is not summarized. In the night with lowest level, it is registered that the noise level in the lowest frequencies (50-200 Hz) is 50-55 dBA as for Normand Jarstein.

Picture 3 Continuous measurement at facade of office building

Picture 3 shows the measurement point for continuous registration at the office building. The frequency spectrum for the situation with and without shore power are based on the levels registered at this building.

5 Conclusions

Shore power connection for drilling and pipelay vessels brings the sound power level L_w down from 105-110 dB to 90-92 dB. For offshore support vessels, the sound power level is 94 dB with shore power connection in dock. In all cases, shore power connection gives significant reduction of low frequency noise. For drilling, pipelay and offshore support vessels, the level measured at port or office building is 50-60 dB in the frequency range from 63 to 1000 Hz with shore power connection. The effect of shore power seems to be much better than predicted in earlier international project [2] and shows similar results as for offshore rigs [1].

Acknowledgements

May I offer my deepest acknowledgements to Kristiansand Port and Andøya Industripark for the possibility to do such measurements, that are very helpful for good estimates of noise reduction due to shore power connection. I want to give special thanks to Odd-Leif Berg, Trond Sikveland and Halvard Aglen in Kristiansand Port and to Svein Erik Halvorsen and Geir Aamo at Andøya Industripark.

References

[1] Bernt Mikal Larsen. “Emitted noise in harbors – Effect of shore power” – BNAM AND INTERNOISE 2018

[2] Rob Witte. “Regulation of noise from moored ships in ports” – EURONOISE 2015

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Applicability of ISO standard 3744 to UA

Julia Treichel[1,*], Jan Foerster[2], Andreas Volkert[3], Teemu Joonas Lieb[3]

[1]German Environment Agency (UBA), Federal Agency, Dessau, Germany,

[2]Technische Universität Berlin (TU Berlin), University, Berlin, Germany,

[3]German Aerospace Center (DLR), Research Institute, Braunschweig, Germany

*julia.treichel@uba.de

Abstract
Unmanned Aircraft Systems (UAS) are used for a variety of purposes. Especially the industrial or professional use of unmanned aircraft (UA) will lead to an increasing number of possible applications. The steadily rising number of UA raises the question of noise impact on the society from these vehicles.

For the first time, an EU regulation provides a uniform noise rating for UA. It involves the introduction of a label for the guaranteed sound power level. This level is to be determined via EN ISO 3744:2010 by means of an enveloping surface method. Manufacturers are required to document the guaranteed sound power level as part of their CE marking. In addition, the EU regulation specifies a maximum permissible sound power level. The permitted level depends on the weight of the UA.

Therefore, the German Environment Agency has started with acoustic investigations of UA. Various small multicopter) were used for the measurements in accordance with the EU regulation. This paper presents the results of the measurements and shows whether the requirements of the EU regulation are complied with. The challenges for users of the applicable measurement standard are also highlighted.

Keywords: UA, UAS, noise, measurements, regulation

1 Introduction

More and more people are using unmanned aircraft (UA), both privately and commercially. While initially the focus was primarily on photography and video, today's applications are much more diverse. UA are used for inspection and maintenance work on e.g. infrastructure, for survey tasks or transporting medical goods. As a result, UA will be increasingly used in the future, not least because longer flight times and larger payloads will become possible.

The steadily increasing number of UA raise the question of expected noise effects. It is foreseeable that more and more people will feel annoyed by the noise of UA flights in the future. However, at present, there are no sound findings on this topic, either nationally or internationally.

The Commission Delegated Regulation 2019/945 [1] includes a noise measurement procedure to label UA similar to the Outdoor Directive 2000/14/EC [2]. This means that the guaranteed sound power level must be displayed on the device or packaging of the UAS. As far as the authors are aware, no practical measurements following the referenced noise measurement procedure has been reported in the literature so far. Therefore, this paper aims at investigating the measurement procedures presented in reference [1] in a practical test setup. Additionally, the paper attempts to identify and highlight practical issues of measurement procedure.

2 Measurement Setup

In its Annex Part 13, reference [1] specifies a noise test specification for recording and calculating the sound power level for UA classes C1-3 and C5-6 in the Open Category. EN ISO 3744:2010 [3] is to be used as the basic standard for the measurement. This standard presents procedures to be used for determining the sound power level of a noise source from sound pressure levels. An enveloping surface enclosing the noise source shall be formed. The measurements are to be carried out in an environment, that approximates acoustic free field conditions in the vicinity of one or more reflecting surfaces.

According to reference [1], the enveloping surface shall be a hemisphere with the microphone arrangements from [3] Annex F. For our measurement we used 12 microphones. The individual microphone positions were calculated as per [3] Annex F. Similarly, the UA is to be measured above a reflective surface, and the rest of the measurement environment must meet the requirements of [3] Annex A.

The measurements took place outdoors at two different airfield sites (see Figure 1 and Figure 2). Runways and taxiways of the airfield were chosen as the reflecting surface in each case. The remaining measurement environment was a spacious lawn without any tree cover or buildings. Further reflections could therefore be excluded.

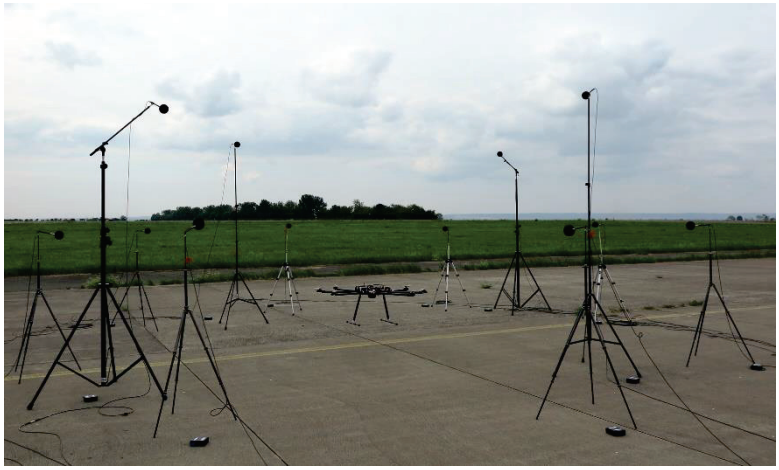


Figure 1: Setup at the National Experimental Test Center for Unmanned Aircraft Systems (DLR) in Cochstedt/Germany

Figure 2: Setup on the model airfield of the model flying club "Hugo Junkers" Dessau-Rodleben/Germany

Based on the sizes of the measured UA (see Table 1), the smallest possible radius of the microphone hemisphere of 4 m was assumed from [3] Annex F and the microphone positions were determined accordingly. The smallest radius was also chosen to make the difference between extraneous noise and UA noise as large as possible.

Table 1: Specification of the used UA models

	UA 1	**UA 2**	**UA 3**	**UA 4**	**UA 5**	**UA 6**	**UA 7**	**UA 8**
weight in g	7000	6200	10000	1320	499	6200	2355	1700
size in cm	162 x 162 x 78	125 x 125 x 58	170 x 170 x 80	38.3 x 38.5 x 24	33 x 38 x 6	96 x 96 x 50	89 x 89 x 22	42 x 42 x 21
construction type	Octo-copter	Octo-copter	Octo-copter	Quadro-copter	Quadro-copter	Hexa-copter	Quadro-copter	Quadro-copter
theoretical class*	C3	C3	C3	C2	C1	C3	C2	C2

*This corresponds to the Open Category classification if all models had been placed on the market after publication and mandatory application of [1]. In fact, they currently correspond to class C5, which is to be labelled, but is not subject to any noise limit value. This is also true for C3.

All models were measured 0.5 m (centre of the UA) above the reflective surface in the hover condition. The definition of the flight altitude for noise measurements of UAs was removed after the amendment of reference [1] by Delegated Regulation 2020/1058 [4]. However, it is still included in DIN EN 4709-001:2021-02 [5]. This draft standard provides technical specifications and test methods to support compliance with [1]. The measurements were repeated at least 5 times for each UA model, while the measurement duration was 5-10 seconds.

3 Measurement Results

All measurement results were provided with the correction parameters K_1 (background noise correction) and K_2 (environmental correction) according to [3]. Besides that, A-weighting was added to the sound level.

Table 2 shows the results of the surface time-averaged sound pressure level ($\overline{L_P}$) according to the following equation (1):

$$\overline{L_P} = \overline{L_P'(ST)} - K_1 - K_2 + A \tag{1}$$

with

$\overline{L_p'(ST)}$ the mean value formed over all microphone positions on the measurement surface, while the noise source under investigation is in operation.

K_1 background noise correction parameter determined according to [3] section 8.2.3

K_2 environmental correction parameter, here $K_2 = 0$, since according to [3] section 4.3.1 K_2 can be neglected for outdoor measurements with a reflecting plane of asphalt or concrete.

A A-weighting

According to reference [1], *"the A-weighted surface time-averaged sound pressure level* [$\overline{L_p}$] *shall be determined at least three times for each UA configuration. If at least two of the determined values do not differ by more than 1 dB, further measurements will not be necessary; otherwise the measurements shall be continued until two values differing by no more than 1 dB are obtained. The surface time-averaged sound pressure level*

[$\overline{L_{px}}$] *to be used for calculating the sound power level of a UA configuration is the arithmetic mean of the two highest values that do not differ by more than 1 dB."*

To comply with the requirements, the values that satisfy the 1 dB criterion and the number of measurements were selected from the measurement results. The values used to determine the arithmetic mean are highlighted with blue background colour in Table 2. In Table 2, "x" denotes inadequate measurements, which were consequently excluded from the validation. This is owed to the fact of sudden gusts of wind, too much background noise or undesired position corrections.

Table 2: Results of the individual measurements and the surface time-averaged sound pressure level.

	UA 1	**UA 2**	**UA 3**	**UA 4**	**UA 5**	**UA 6**	**UA 7**	**UA 8**
measurement	$\overline{L_P}$ in dB(A)							
1	78.5	x	x	65.8	59.2	75.6	x	66.8
2	78.6	77.6	77.9	64.7	x	75.5	67.8	66.8
3	78.6	78.3	x	x	59.9	76.4	66.5	67.0
4	78.7	x	77.8	65.1	60.7	x	66.8	66.6
5	78.9	77.5	77.4	66.1	59.2	75.7	66.4	x
6	78.9	78.1	78.0	-	-	-	67.5	67.3
7	-	77.9	-	-	-	-	-	66.9
8	-	x	-	-	-	-	-	67.4
$\overline{L_{px}}$ in dB(A)	78.6	78.0	77.8	65.5	60.3	76.0	67.3	66.9

The sound power level was determined according to the following equation (2):

$$L_{WA} = \overline{L_{px}} + 10\,lg\,\frac{S}{S_0}\,dB \tag{2}$$

with

S area of the measurement surface, here area of a hemisphere with radius $r = 4$ m

S_0 $S_0 = 1\,m^2$

According to reference [1] Part 14, the guaranteed sound power level must be indicated in a pictogram on the product. The definition of the guaranteed sound power level in this regulation is given as follows:

"[…] *a sound power level determined in accordance with the requirements laid down in Part 13 of the Annex which includes the uncertainties due to production variation and measurement* […]" [1]

Thus, an expanded uncertainty U was determined according to paragraph 9 of reference [3] with equation (3) and (4):

$$U = k\sigma_{tot} \tag{3}$$

$$\sigma_{tot} = \sqrt{\sigma_{R0}^2 + \sigma_{omc}^2} \tag{4}$$

with

k coverage factor, here $k = 2$ because only one model per species was available and no further database is known.

σ_{tot} total standard deviation

σ_{R0} all uncertainties allowed by the measurement standard, except those for the instability of the sound power of the source under investigation (σ_{omc}). In this case 1.5 dB, since no value specific to the product is known or has been determined so far.

σ_{omc} standard deviation for determining variations in operating and installation conditions, determined, here over all valid measurements from Table 2.

The expanded uncertainty U was added to the calculated L_{WA} to obtain the guaranteed sound power level. The results are shown in Table 3.

Table 3: calculated vs. guaranteed vs. maximum sound power level

	UA 1	**UA 2**	**UA 3**	**UA 4**	**UA 5**	**UA 6**	**UA 7**	**UA 8**
calculated L_{WA} in dB (A)	98.6	98.0	97.8	85.5	80.3	96.0	87.3	86.9
U in dB	3	3.1	3	3.3	3.3	3	3.1	3.1
guaranteed L_{WA} in dB (A)*	102	101	101	89	84	99	90	90
max. L_{WA} in dB (A)				88.1	85.0		92.7	90.1

*According to [1] the guaranteed sound power level needs to be rounded to integer values.

Likewise, the theoretical maximum sound power level (max. L_{WA}) was determined according to [1] Part 15. For UA weighing more than 4 kg (classes C3, C5 and C6), unfortunately no values are given in the reference, so only the values of the "smaller/lighter" UA are apparent. For further details refer to sections 5 and 6.

4 Application Issues

The obtained results cannot be regarded as representative/absolutely conforming to standards, since some compromises or assumptions were made during the measurements and evaluation. The reason for this is that too little experience or no testing of the measurement method with UA is available so far and therefore a suitable database is missing.

According to [3], the measurement time should be 20 s, but must be at least 10 s. However, experience from the measurements carried out shows that a duration of 20 s is hardly possible. Even under optimal conditions, the UA drifts away from its initial hover position and corrections have to be made by the UA pilot to remain at this position. Likewise, the GPS signal is usually not accurate enough to maintain the required position (± 0.05 m) [5]. This could affect the measurement significantly. Therefore, for some UA models the measurement time had to be reduced to 5 s. However, this was seen as acceptable for the purpose of the measurements.

In case it is not possible to keep a stable position, the UA shall be mounted on a tripod or should be tethered to limit horizontal movements of the UA according to reference [5]. This is hardly possible for users who cannot intervene in the UA system, e.g. by setting specific rotor speeds. If mounted on a tripod, the UA would immediately apply full torque to the motors/rotors as it tries to lift off to get into hover position. The feedback controllers would try to compensate for the weight of the tripod. This, would distort the measurement results.

Another ambiguity was the specification of the hover position of 0.5 m above ground level. Before the amendment of [1], no further information was given as to where the 0.5 m was to be applied (UA center, rotor plane, foot plane). If the center or rotor plane was assumed, the problem arose, especially with UA over 4 kg and larger dimensions (cf. UA 1-3, Table 1), that they were only allowed to lift off a few centimeters above the ground. Thus, safe hovering was hardly possible. Due to changes in [1], the height requirement has been eliminated. However, this is included in the draft of the concretizing standard [5], where the foot plane is taken as the reference. It is questionable to what extent the resulting different heights of the rotor plane for different UA models will affect the measured values.

5 Evaluation Issues

For the evaluation the correction parameters K_1 and K_2 must be determined. K_1 is determined from the difference between the extraneous noise level and the product noise. Here the problem arises that especially for smaller UA the difference is insufficient. The reason for this is the given radius in [3] Annex F. This Annex was originally developed for the measurement of construction machines. The smallest option of $r = 4$ m is usually too far for UA below 4 kg and therefore causes an invalid K_1. Consequently, the UA noise differs too little from the extraneous noise.

The environmental correction parameter K_2 can be neglected in our case. According to [1] the determination of K_2 for measurements in space is left optional. Annex A of reference [3] gives various possibilities for determining the environmental noise correction.

To determine the uncertainties required for the guaranteed sound power level, a coverage factor k and further uncertainties must be specified. There is currently a great deal of leeway here, since the regulation does not specify any correction values. As far as the authors know, there is a lack of preliminary investigations to determine suitable values. The coverage factor k could be determined through a series of tests by the manufacturer. In each configuration, a certain number of the same model should be tested and indicated in the measurement report. Yet there is no obligation to do so. Section 9.1 of reference [3] indicates that a coverage factor of $k = 1.6$ dB would be more appropriate when comparing the sound power level to a limit value. However, since no data basis exists yet, this assumption is not worth supporting. The same applies to the uncertainty σ_{R0}. For this purpose, large-scale interlaboratory tests would have to be performed by different manufacturers, UA models, and measurement institutes to produce valid results.

Due to the missing clear specification of these two values, the manufacturer is free to correct the measured, guaranteed or maximum sound power levels in a desired direction.

Another discontinuity in [1] is that the term maximum sound power level is not specified. This raises two questions: Should the maximum permissible guaranteed sound power level or the maximum measured sound power level be specified here? Which sound power level should be included in the report?

Furthermore, it is not clear how the indicated maximum sound power level was determined or on what data basis it was based. The authors are not aware that there has been any analysis of existing products in this regard. The Outdoor Directive 2000/14/EC [2] was used as a model for parts 13-15 of reference [1]. However, the noise limits specified there are based on a variety of preliminary studies, the results of which form the basis for the setting [6, 7]. These studies were funded by the Commission for this reason.

6 Conclusions

Although the measured models comply with the defined noise limits (as far as possible), there are still some issues that must be clarified.

In general, it can be emphasized that there are still too few studies dealing with UA noise and its mitigation. The Delegated Regulation 2019/945 [1] can be regarded as a first step towards reducing the physical noise pollution caused by UA. However, it should be optimized in the sections mentioned in this paper and a database should be evaluated. Similarly, it is not yet sufficient to evaluate the extent of noise effects, such as annoyance. This is influenced by various acoustic and non-acoustic factors that need further investigation.

In most cases, the data from literature sources only refer to UA of the multicopter type. A broader data base and a standardized measurement practice that also includes other design configurations would be necessary. Furthermore, the emission model should distinguish different operating states: hovering, take-off, landing, climbing, descending and level flight with "typical" forward speed. In practice, different operating states have different noise characteristics that are not comparable with each other. Hovering is the quietest and most

unrealistic condition and section 6.6 of ISO 3744:2010 specifies that the measured object shall be measured under conditions that are representative of the noisiest mode of operation under normal use.

Just as the development of UA is far from complete, the standardization framework as well as the legal framework must also be further developed, adapted and tested in order to create appropriate specifications for UA operation. Environmental protection and noise abatement (still) plays a subordinate role but should be given greater consideration in order to protect the public from an additional noise source.

Acknowledgements

I would like to thank my co-authors and colleagues at TU Berlin and DLR for the measurement campaigns carried out with us and for providing the UA models. I would also like to thank my former colleague Steffen Körper for his support of my measurement idea and the associated implementation.

References

[1] Commission Delegated Regulation (EU) 2019/945 of 12 March 2019 on unmanned aircraft systems and on third-country operators of unmanned aircraft systems. OJEU L 152 11.6.2019, p. 1.ISSN 1977-0677.

[2] Directive 2000/14/EC of the European Parliament and of the Council of 8 May 2000 on the approximation of the laws of the Member States relating to the noise emission in the environment by equipment for use outdoors. OJEU L 162 03.07.2000, p.1.ISSN 0378-6988.

[3] ISO 3744:2010:2010-10. Acoustics - Determination of sound power levels and sound energy levels of noise sources using sound pressure - Engineering methods for an essentially free field over a reflecting plane.ICS17.140.01.

[4] Commission Delegated Regulation (EU) 2020/1058 of 27 April 2020 amending Delegated Regulation (EU) 2019/945 as regards the introduction of two new unmanned aircraft systems classes. OJEU L 232 20.07.2022, p.1.ISSN 1977-0677

[5] DIN EN 4709-001:2021-02. Aerospace series, Unmanned Aircraft Systems, Part 001: Product requirements and verification; English version prEN 4709-001:2021.ICS 49.020.

[6] ODELIA (OutDoor Equipment Noise LImit Assessment). Final Study Report for European Commission Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs Tender No 414/PP/ENT/119427. link: https://ec.europa.eu/docsroom/documents/18281/attachments/1/translations/

[7] Preparatory study on the environmental performance of residential room conditioning appliances (aircon and ventilation). Draft report of Technical Analysis of existing Products. TREN/D1/40-2005/LOT10/S07.56606. link: https://circabc.europa.eu/sd/a/773b634f-9e34-444c-9406-3693982e00b3/V%20_%20Ventilation%20_%20final%20report.pdf

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Electric Cars - Noise Simulation of AVAS Effects

Sebastian Eggers[1,*], Heinz Steven[2]

[1]LÄRMKONTOR GmbH, Hamburg, Germany

[2]HS Data Analysis and Consultancy, Heinsberg, Germany

*s.eggers@laermkontor.de

Abstract
Besides their environmental impact on e.g. air pollution, electric cars are often expected to bring a noticable reduction in noise emissions. Even for speeds on urban roads, the noise contribution of the tyres (i.e. rolling noise) is dominant. A reduction in urban areas could still be expected as the engine noise is significantly lower. Since mid 2021, the ”Acoustic Vehicle Alert System” (AVAS) is mandatory for new electric vehicles to improve safety for pedestrians and cyclists, creating artificial noises for speeds up to 20 km/h. Overall, the impact of AVAS on urban road noise remains unclear as the emission varies between different cars and can also be in magnitude of combustion engine noise emissions. As part of recent research, a noise simulation was carried out using detailed noise emissions for single cars (ROTRANOMO/TraNECaM). Rolling noise and propulsion noise (noise from the engine, the exhaust and the transmission) are calculated for various traffic situations in second by second steps, depending on the vehicle’s acceleration and speed. The model allows the detailed simulation of AVAS as the emission is only active for low vehicle speeds. Simulations were carried out for different crossing situations with percentages of cars passing at higher speed and stopping. Results will be presented as well for the average noise level as for an assessment of the changing noise level and single noise events.
Keywords: vehicle noise, electric cars, AVAS

1 Introduction

Besides their environmental impact on e.g. air pollution, electric cars are often expected to bring a noticable reduction in noise emissions of road vehicles. A reduction in urban areas (at low vehicle speeds) could still be expected as the noise emission of an electric motor is significantly lower than the noise emission of an internal combustion engine. This advantage could be deteriorated by the ”Acoustic Vehicle Alert System” (AVAS) which has become mandatory since mid 2021for new electric vehicles to improve safety for pedestrians and cyclists, creating artificial noises for speeds greater than 0 km/h up to and inclusive 20 km/h. The requirements and limitations of the AVAS are specified in UN-ECE Regulation 138 [1]. Revision 1 of this Regulation permits the activation of the AVAS beyond 20 km/h (up to 32 km/h, which is equivalent to 20 mph, the upper speed limit of the US regulations).

From previous measurement reports and experience e.g. with TraNECaM [2] (see chapter 2.1) it is known that vehicle emissions from electric and combustion vehicles are almost identical at speeds of 50 km/h and above, as these are dominated by rolling noise and propulsion noise is low in modern passenger cars. In contrast, at speeds around 30 km/h and below, the propulsion noise may dominate for internal combustion engine vehicles. Electric vehicles could therefore be fundamentally quieter in this range than internal combustion engine vehicles - due to the lower propulsion noise. However, the use of AVAS at speeds below 20 km/h adds an additional emission, the effect of which we want to demonstrate in this presentation.

From the emission modelling in TraNECaM/ROTRANOMO, the emissions of rolling noise and propulsion noise will be available on a second-by-second basis for individual vehicle types and vehicle speed traces representing individual traffic situations. The emissions are to be located along the route with second-by-second resolution. Thus, the influence area of an intersection with regard to noise emissions can be identified in noise propagation calculations. This allows the determination of the distance from the intersection to which the influence of an AVAS will have an effect in noise levels and how vehicle emissions will change from vehicles with internal combustion engines to vehicles with electric motors in an intersection location.

In addition, the difference in emissions is determined in comparison with the "free flow" traffic situation. For the area "approaching" the intersection (deceleration), lower emissions may be expected, for the area "leaving" the intersection (acceleration), higher emissions may be expected. By superimposing the traffic (proportion of accelerating and decelerating vehicles), the difference to the "free flow" situation can be assessed.

2 Emission model

For the intended analysis, an emission-sensitive noise calculation is required, in which a separation of rolling noise and propulsion noise is possible. Based on this, the differences in propulsion noise between vehicles with an internal combustion engine and an electric motor (lower engine and no exhaust noise for electric vehicles), but also additional noise emission by artificial sound from AVAS for electric vehicles can be taken into account. The relevant noise emission by AVAS will be determined, and as a result, noise emissions for vehicles with different variants of AVAS emissions will be generated and compared.

2.1 TraNECaM/ RoTraNoMo

For the analysis, the noise model TraNECaM (Traffic Noise Emission Calculation Modell) [2] was used. This model was developed within a research project for the German Federal Environment Agency (UBA) in 1998 to 2000. It allows a more detailed emission calculation than the conventional calculation models and also takes into account the technical progress of motor vehicles. The data basis of the model is considerably broader in terms of road and vehicle categories than that of the conventional models and was originally developed on behalf of the Federal Environment Agency.

It was expanded and updated with financial support from the EU Commission and the Norwegian Pollution Control Authority. In Norway, TraNECaM has been used to quantify nationwide the effects of abatement measures that could be used to achieve policy abatement targets. Up to 2021 the traffic situations stored in TraNECaM were based on the traffic situations from HBEFA 3.1 [3], which is a comparable model for emission factors of air pollutants. Since the traffic situations and the underlying driving cycles were significantly updated in the time period 2018 to 2020, leading to the publication of HBEFA 4.1, in 2021 a complete recalculation of the TraNECaM model was performed using the driving cycles from HBEFA 4.1. This recalculation required an upstream model of TraNECaM, the RoTraNoMo model in analogy to the HBEFA model which requires the PHEM model as upstream model.

The RoTraNoMo model was developed within the project ”Development of a Microscopic Road Traffic Noise Model for the Assessment of Noise Reduction Measures” funded by the European Community under the ‘Competitive and Sustainable Growth’ Programme. The model calculates an instantaneous pass by level for different vehicle subcategories and emission stages on the basis of second by second vehicle speed traces separately for tyre/road and propulsion sound levels. The model was developed in 2003 to 2005 but was updated in later projects in order to cover also the 3 emission stages defined in 2009/661/EU and 540/2014/EU, phase 1 to phase 3.

Based on the experience of previous research projects, the vehicle related noise sources (engine, powertrain, exhaust, intake) are summarised as propulsion noise and modelled as function of engine speed and engine load. A further split into different sources is possible in principle but at present there is no data available for the development of such an amendment. The tyre/road, or rolling, noise component L_{roll} is modelled as a function of the tyre, the road surface and the vehicle speed.

2.2 AVAS

When modelling the noise emissions of an electric vehicle approaching, crossing and leaving an intersection by RoTraNoMo, we found that the influence of AVAS in the intersection area depends very much on the chosen approaches for the emissions of the AVAS.

Regarding the "minimum requirements", we can apply the emissions according to UNECE Regulation 138 "Uniform provisions concerning the approval of Quiet Road Transport Vehicles with regard to their reduced audibility (QRTV)" [1]. However, it is to be expected that in reality vehicles do not comply exactly with these minimum standards (tolerances, liability issues, etc.).

A requirement for the maximum emission level for the AVAS can also be found in Annex VIII of EU Regulation 540/2014 [7] in No. 3.c): "The sound level generated by the AVAS shall not exceed the approximate sound level of a vehicle of the M1 category equipped with an internal combustion engine and operating under the same conditions." This emission is normally lower than the maximum allowed AVAS sound emission according to UNECE Regulation 138 (75 dB(A) at 2 m distance).

Overall, the range (at 10 km/h: 50-75 dB(A) at 2 m) is very high, the result (especially for a detailed consideration in the intersection area) is virtually solely determined by the modelling approach. For a better assessment of real emission levels, a limited literature survey was conducted, supplemented by contacts to car manufacturers.

A study from 2019 [4] conducted measurements on a number of cars and showed that emissions with AVAS differed greatly between manufacturers. The minimum required value is exceeded by 4-13 dB (lowest value: Smart, highest value: Renault Zoe). The AVAS was virtually undetectable in an eGolf. In some cases, significant deviations between the left and right side of up to 7 dB (Renault Zoe) were determined. This is increasingly the case when only one loudspeaker is used to generate sound: To ensure minimum emission in each direction, the emission in one direction may be significantly higher than required. Another earlier publication [5] concluded that in those cases analysed, an all-electric operation without AVAS showed the lowest noise levels. The combustion engine is slightly louder than the electric vehicle without AVAS at lower speeds. However, at low speeds the combustion engines noise level is on average lower than the noise level of a vehicle with AVAS system. In a further study [6], AVAS was not detectable at 20 km/h (Nissan eNV200). The rolling noise had been dominant.

Own investigations with interviews of car manufacturers have shown that, in addition to the pure A-weighted sound level, especially the frequency response requirements at different speeds have a strong impact on the emissions. To ensure the required emissions, a high tolerance of the systems used must be foreseen, which according to oral testimony (at individual frequencies) can be up to 12 dB.

2.3 Emission cycle and noise emission contributions

From the set of road types implemented in the newest version of TraNECaM with the update to road types of HBEFA 4.1, one representative cycle was chosen for a typical intersection. The cycle chosen is used for “distributor roads” with a speed limit of 50 km/h, “saturated” [1] level of service (quality of traffic flow) and 0 %

[1] Level 3 of 5, where 1 is free flow and 5 is high congestion

gradient. .The whole cycle covers about 500 s and a travelled distance of about 3.3 km. The vehicle speed over the time travelled is shown in Figure 1.

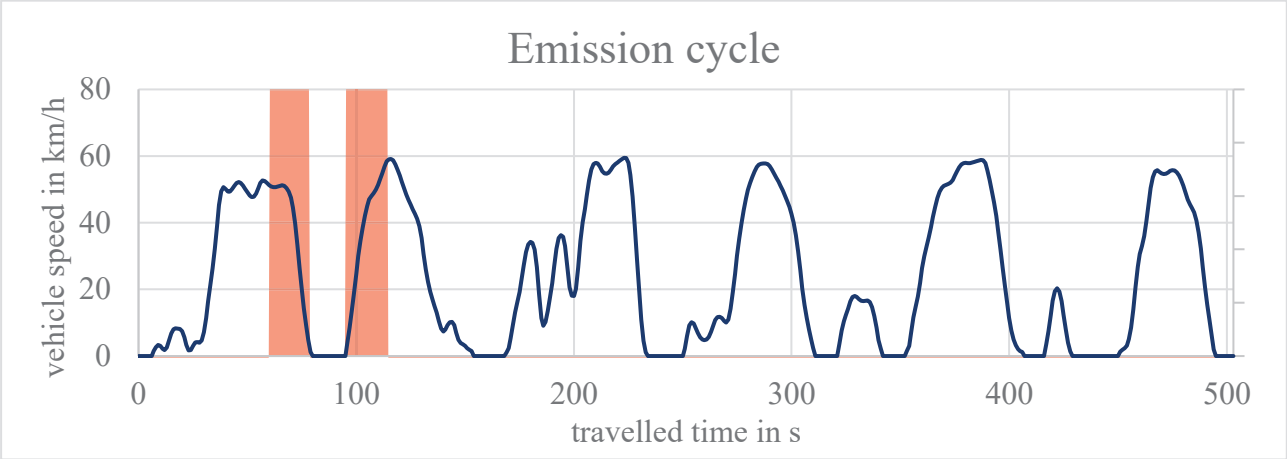


Figure 1: Vehicle speed in the emission cycle, selected section highlighted in red

For a typical intersection on a distributor road the segment as marked red in Figure 1 was chosen. The vehicles approach at a speed of ~50 km/h, come to a full stop and accelerate again to about ~60 km/h.. As just a small section of the cycle was used, the level of service has low influence on the resulting noise emissions.

For more congested traffic, the average speed will be lower, especially if it takes a car more than one cycle to cross the intersection (see Figure 1, second 270-370) . The results of such traffic situations was not simulated, but considered in the conclusions.

In Figure 2, the contribution of rolling noise (green) and propulsion noise (blue) of a typical combustion car is depicted for the selected segment. The corresponding speed (orange) influences both noise emissions. However, the propulsion noise is also dependent on the necessary engine power and the resulting engine speed, giving a lower emission on decelerating (seconds 70-80) than on accelerating (seconds 95-100). The idling of the engine has not (yet) taken into account an automatic start-stop system.

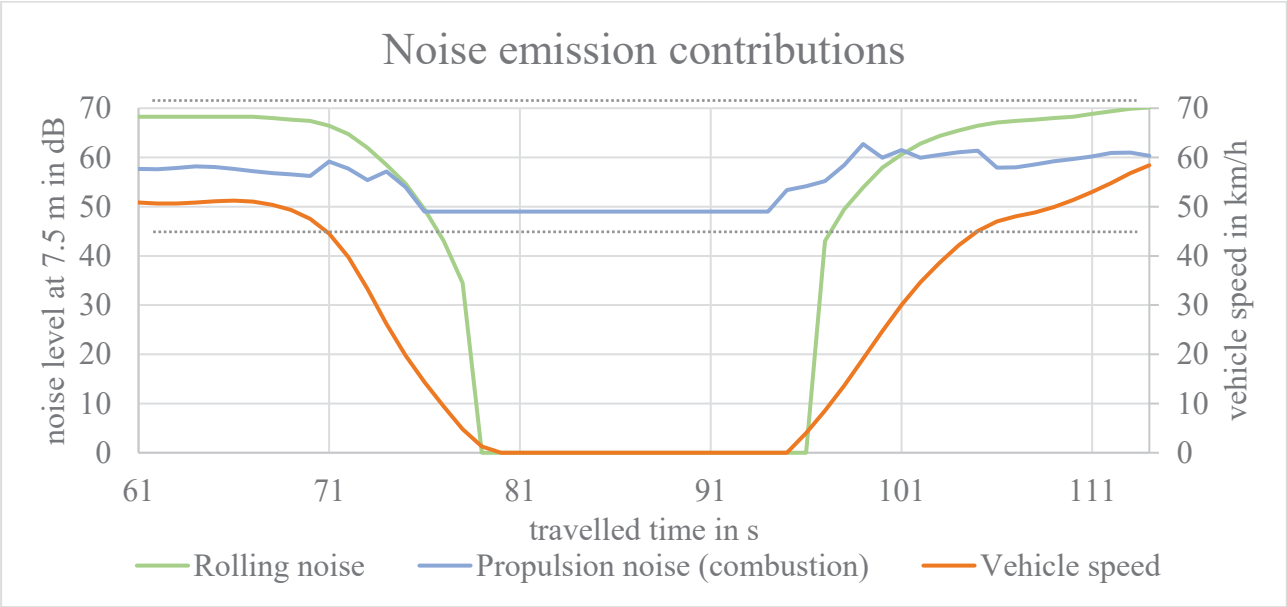


Figure 2: Contribution to the noise emission – rolling noise and propulsion noise

In addition to the emissions of a combustion engine, the limits for AVAS noise emissions are also shown (dotted lines parallel to the x-axis). The values mentioned before were converted to a distance of 7.5 m instead of 2 m, giving a maximum of 66 dB and a minimum of 41 dB. It is clearly visible that typical combustion engine noise is significantly lower than the maximum allowed for AVAS. The minimum noise level allowed for AVAS is in most cases more than 10 dB lower than the combustion engine noise. For low speeds and close to idling, the difference is lower, closer to 8 dB.

2.4 Modelling of AVAS noise emission

Based on the variance of possible noise emissions of vehicles equipped with AVAS (as shown in paragraph 2.2), two variants were modelled besides a vehicle with combustion engine and an electric vehicle without AVAS. The two variants will be referred to as AVAS (60) and AVAS (50) in regard to their AVAS noise level at 7.5 m (60 dB and 50 dB). The electric car without AVAS is modelled with a propulsion noise level of 40 dB at all speeds. For speeds > 20 km/h, this is also used for the vehicles with AVAS. The variant without AVAS was added in order to show the maximum noise reduction potential an electric vehicle fleet would have.

The variant with “maximum AVAS emissions” (66 dB) will not be pursued further, as it can be assumed that this value will only be fully exploited in individual cases. However, individual vehicles that use the AVAS as a "sound design" to stand out from the mass of vehicles will not dominate acoustically on long term average noise levels.

In summary, the modelling takes into account the following scenarios:

- Combustion engine vehicle
- Electric vehicle without AVAS
- Electric vehicle with low AVAS emissions, still about 9 dB above minimum requirements
- Electric vehicle with high AVAS emissions, about 15 dB above minimum, but still 6 dB below maximum

The maximum pass-by levels of TraNECaM were converted to emission (L_{eq}) in accordance to the documentation of TraNECaM [2]. In addition, for the upcoming modelling, the emission was spread on the corresponding segments of a road (1 m) that a noise source is active during a time step. The exposure time for each segment is taken into account. Thereby, at low speeds, the same pass-by level results in higher emission levels due to higher exposure at each segment.

The resulting noise emission levels are depicted in Figure 3. The steps visible are caused by the 1 Hz time resolution of the emissions. The combustion vehicle is shown in blue. The electric vehicles show a slightly lower noise emission at speeds of ~50 km/h due to the slightly lower propulsion noise. Closer to the intersection at decreasing speeds (~ - 40 m), the electric vehicles show a higher decrease in noise emissions.

At a distance of about 10 m to the intersection, the speed is below 20 km/h and the noise emission of AVAS sets in. For the AVAS (50) vehicle, the noise levels of the decelerating car are close to the combustion engine car, whereas the AVAS (60) vehicle shows higher noise emissions.

For accelerating vehicles, the emission of the AVAS (50) vehicle are close to the combustion engine vehicle in the first seconds but quickly decrease. The AVAS (60) vehicle shows a significantly higher noise emission for the first ~10 m and a level comparable to the combustion engine vehicle up to ~25 m from the intersection.

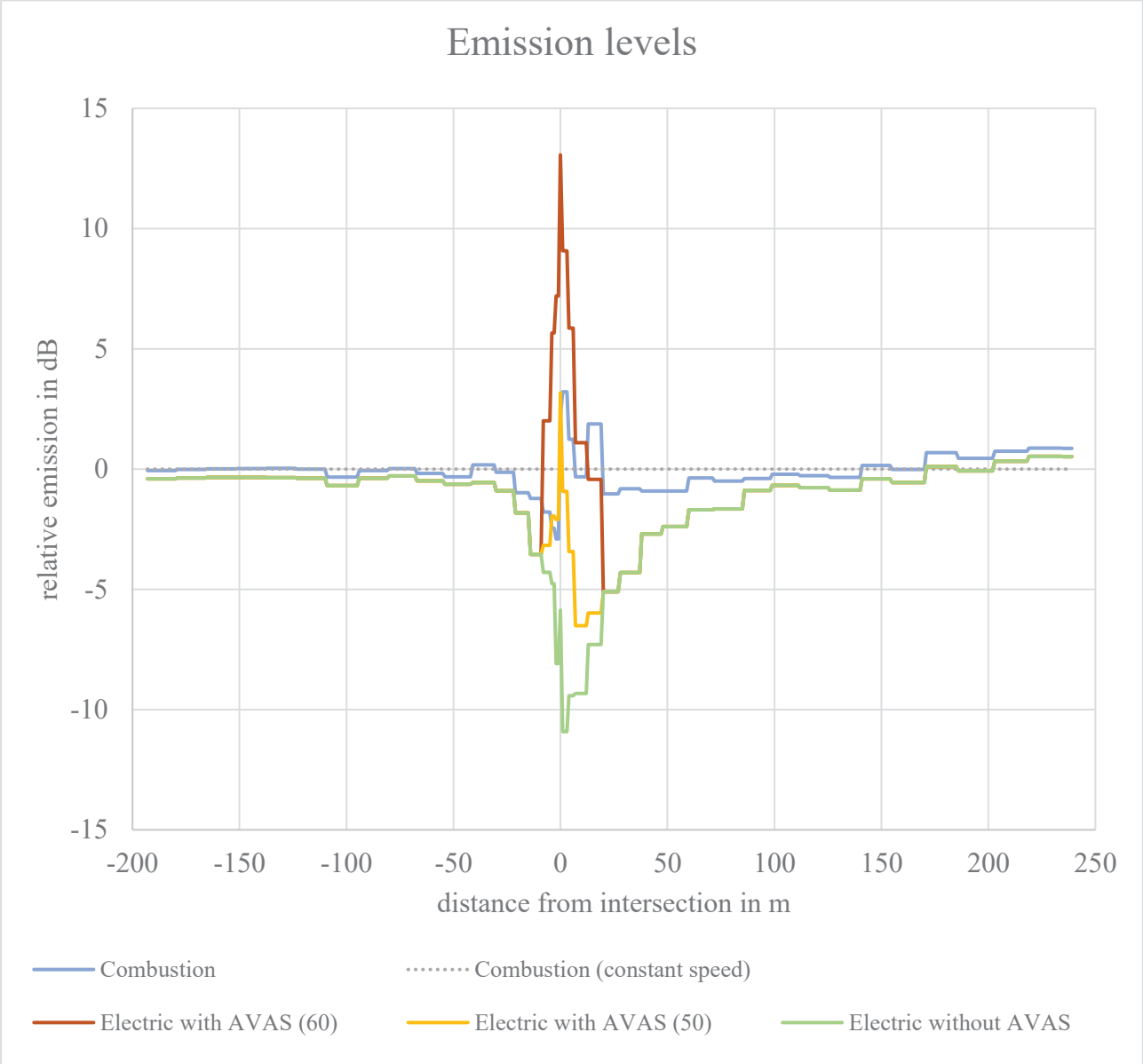


Figure 3: Emission levels (relative to constant speed of combustion engine) for the different vehicle types

In comparison to the combustion vehicle with a stop at an intersection, a combustion vehicle with constant speed is taken into account. The speed equals the speed of the cars before the intersection. At a distance of about 150 m after the intersection, the speed of the other cars exceeds this constant speed and thus the noise emission is slightly higher.

Not taken into account are several other factors that could affect the noise emission close to the intersection:

Start-stop systems for combustion engines compared to AVAS emissions for vehicles at standstill: Most modern cars are equipped with start-stop systems, switching off the engine at standstill and thus reducing the noise emission. For AVAS systems, the behaviour for stopped vehicles is not modelled. Several observations show a possible continuing noise emission even seconds after full halt (e.g. in [8] still active, another example in [9] shows the emission stops when the transmission is in “park mode”). It is presumed in the following calculations that both systems are equal during the standstill phase and have no relevant emissions.

The noise level of AVAS might change with the vehicle's speed. Mandatory is a change in frequency, but an additional change in noise levels could be implemented. A higher emission could either be implemented on higher engine power (e.g. at initial acceleration) but also at increasing speeds. It is presumed that the noise level of both variants is around the given values of 50 dB and 60 dB, but constant.

Regarding AVAS emission above 20 km/h, a sudden "cut-off" is considered in the emissions. Dependent on the implementation, AVAS emissions might still be audible above 20 km/h, but the contribution to the overall noise level will be almost insignificant due to the increasing rolling noise.

Depending on the sound of the AVAS, a tonal component might influence the perception and annoyance reaction.

3 Noise propagation calculation

To assess the influence of the emission changes identified so far, a noise propagation calculation was carried out. The model is a simplified road stretch with segments of 1 m as a noise source. In the figures, the road source is a horizontal line (magenta). A second, vertical line (magenta) indicates the position of the intersection.

3.1 Model and scenarios

Calculations were carried out using SoundPLAN 8.2, with a DIN ISO 9613-2 line source. The height of the noise source was set to 0.5 m, the calculations were carried out for a grid of 1 m in a height of 2 m. The calculation is just approximating the noise levels as different noise calculation models used for road traffic (as the German national RLS-19 and BUB or the European CNOSSOS for environmental noise) may differ in certain specifics. Nevertheless, the results give a good indication on the possible deviations between the considered noise sources, using the difference in noise levels. Noise level changes below 1 dB will be considered almost irrelevant. Absolute noise levels will not be depicted.

The scenarios for the propagation calculations are identical with the variants on noise emissions:

- Combustion vehicle with stop
- Combustion vehicle with constant speed
- Electric vehicle without AVAS
- Electric vehicle with low AVAS emissions (AVAS 50)
- Electric vehicle with high AVAS emissions (AVAS 60)

The differences shown in the following Figure 4 are calculated for three variants of electric cars in comparison to the combustion engine with stop and in one variant for the combustion engine with stop in comparison to combustion engine with constant speed.

3.2 Results

The results are shown in Figure 4 for the different variants.

In the upper left, a clear influence of high AVAS emissions (AVAS 60) is visible. For an area of about 25 m behind the intersection, a noise level increase of more than 1 dB can be seen. In close proximity to the intersection at about 10 m, the increase ranges up to more than 3 dB. In the upper right, the results for lower AVAS emissions (AVAS 50) show a quite different picture. For the same area about 25 m behind the intersection, a decrease in noise levels of 1-3 dB can be seen.

In the lower left, the results for electric vehicles without AVAS show an even higher decrease in noise levels for a larger area.

In comparison of a stopping combustion car to a car with constant speed, shown in the lower right, shows almost no differences in noise levels. This can be derived from the emission characteristics shown in Figure 3: starting about 20 m before the intersection, the noise emission of the stopping car is slightly lower (up to 3 dB), for a stretch of about 20 m behind the intersection the noise emission of the then accelerating car is slightly higher (up to 3 dB).

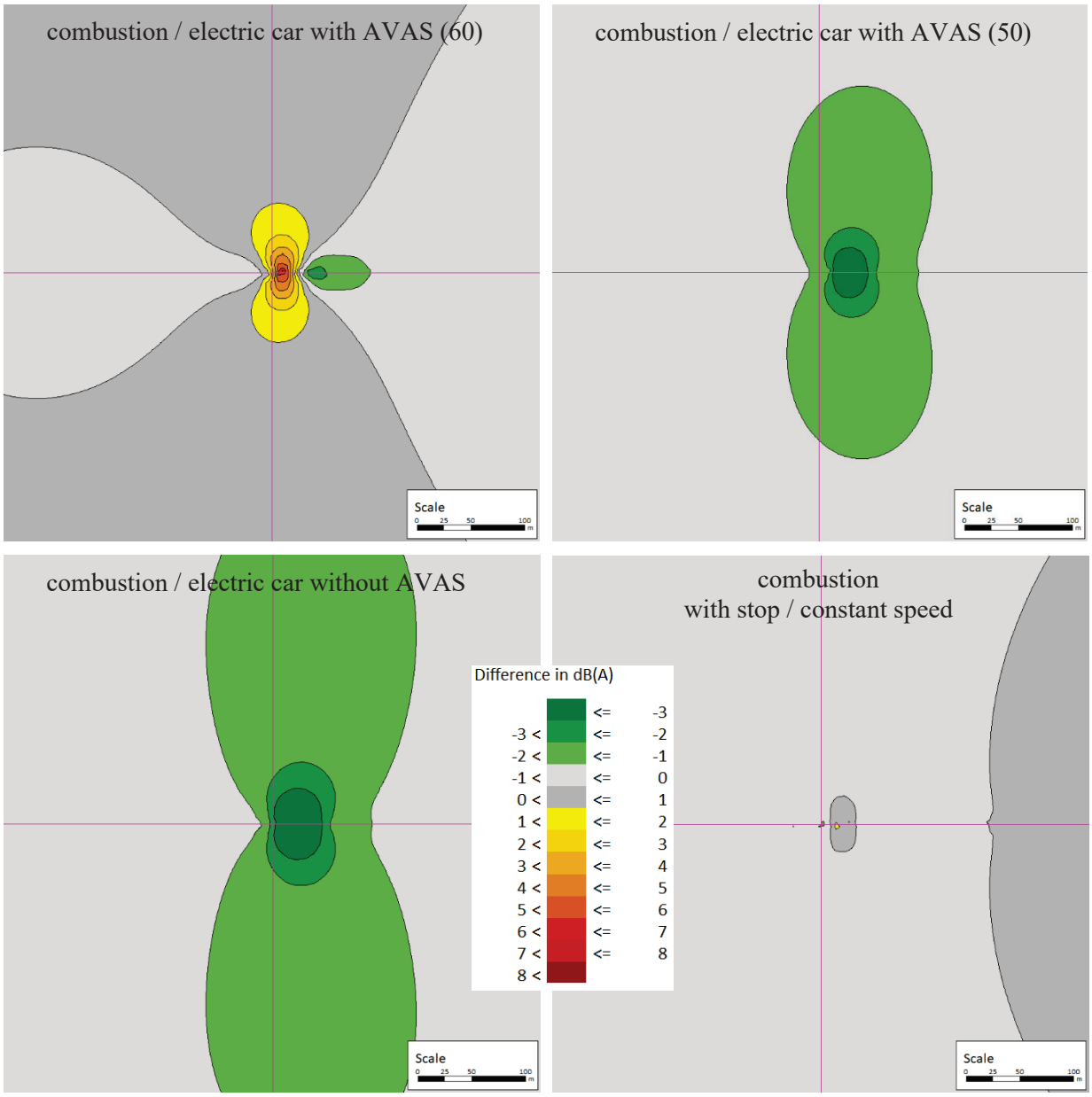


Figure 4: Differences in average noise levels around an intersection – noise source moving from left to right along the horizontal line, intersection at the vertical line. Results for combustion / electric car with AVAS (60) (upper left), combustion / electric car with AVAS (50) (upper right), combustion / electric car without AVAS (lower left), combustion with stop / combustion with constant speed (lower right).

4 Conclusions

A literature review shows that AVAS emissions can have a broad range, caused by the permitted range of about 25 dB given in the regulations. On the basis of the present regulations, it cannot be ensured that the noise emission of an electric vehicle with the use of AVAS is really quieter than a comparable passenger car with an internal combustion engine.

However, the noise emission alone does not give an indication on the noise exposure close to intersections. To assess the resulting effects, noise levels were calculated for different variants of cars. Calculations showed that the influence of a stopping combustion vehicle compared to a vehicle with constant speed is neglectable regarding the average noise levels (annoyance is not taken into account). For electric vehicles, a negative influence can be seen for vehicles with comparable high AVAS noise emissions. With comparable low noise emissions, which still exceed the minimum requirements by almost 10 dB, a slight noise reduction can be achieved.

As final conclusion, an "acoustic gain" cannot be reliably achieved by electric vehicles with the use of AVAS. Rather, the maximum emission level permitted by the AVAS might lead to much higher noise levels in proximity to intersections compared to combustion engine powered passenger car. From a technical point of view, the effect of the AVAS on the emission behaviour of electric vehicles must therefore be assessed as possibly negative in terms of noise. The extent to which noise emissions exceed those of combustion cars in the overall fleet still needs to be examined in further work.

References

[1] UN Regulation No. 138, Uniform provisions concerning the approval of Quiet Road Transport Vehicles with regard to their reduced audibility (QRTV), Revision 1, 2017

[2] TraNECaM, Emissionsmodul im Geräuschbelastungsmodell (Emission module in the noise load model), ARGE TÜV Automotive / LÄRMKONTOR, Vorhaben Nr. 105 02 221, on behalf of the German Federal Environment Agency (Umweltbundesamt), December 2000, see also Traffic Noise Emission Calculation Model - Documentation and Users Manual, 2005

[3] HBEFA, Handbook Emission Factors for Road Transport, see https://hbefa.net/ for further documentation

[4] Laib, F. und Schmid, J.A. 2019. Acoustic Vehicle Alerting Systems (AVAS) of electric cars and its possible influence on urban soundscape. ICA Aachen : s.n., 2019.

[5] Bock, F., et al. 2018. Auswirkungen der AVAS-Gesetzgebung auf elektrifizierte Fahrzeuge. Präsentiert bei DAGA 2018 – 44. Deutsche Jahrestagung für Akustik. München : s.n., 2018.

[6] Berge, Truls Svenn und Haulkand, Frode. 2019. Adaptive acoustic vehicle alerting sound, AVAS, for Electric vehicles Results from field testing. s.l. : SINTEF, 2019.

[7] REGULATION (EU) No 540/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 16 April 2014 on the sound level of motor vehicles and of replacement silencing systems, and amending Directive 2007/46/EC and repealing Directive 70/157/EEC

[8] https://youtu.be/k82rjRxl6Gk?t=55, last access on 22.03.2022

[9] https://youtu.be/MUcbXplAvSc?t=132, last access on 22.03.2022

[10] DIN ISO 9613-2:1999-10, Akustik - Dämpfung des Schalls bei der Ausbreitung im Freien - Teil 2: Allgemeines Berechnungsverfahren (ISO 9613-2:1996) (Acoustics - Attenuation of sound during propagation outdoors - Part 2: General method of calculation (ISO 9613-2:1996))

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Acoustic labels as a simple, easy-to-understand visual representation of the acoustic suitability of unconventional spaces

Kristian Jambrošić*, Marko Horvat

University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia.

*kristian.jambrosic@fer.hr

Abstract

Instead of building many dedicated spaces that are acoustically optimised for specific types of events, many municipalities are forced by financial constraints to use multi-purpose halls and/or unconventional spaces as venues for a variety of different events. Depending on immediate needs, many public spaces are used as conference or lecture rooms, as well as venues for other speech-based events, but also as venues for unamplified or amplified musical performances, exhibition spaces, cinemas, theatres, etc. Many historic buildings are also frequently used in this way due to their attractiveness to visitors and their historical and cultural significance. In all these cases, the acoustic suitability of a space for a particular type of event must be investigated and determined. Many objective room acoustic parameters have been developed and are commonly used to evaluate the acoustic quality of both performance spaces and ordinary rooms. These parameters are not necessarily intended for the evaluation of unconventional spaces or spaces that are used temporarily or occasionally as performance spaces, e.g. an industrial hall used as a theatre, or a stone atrium used as a concert venue. Nevertheless, a rough rule-of-thumb evaluation can be made using these parameters to assess the acoustic suitability of such unconventional spaces for certain types of events. To help non-experts understand the results of such an evaluation, simple methods of visually representing the results are proposed, leading to a kind of acoustic label describing the acoustics of a room, similar to a product label describing a product by listing its main characteristics, or an energy label as the indicator of the energy efficiency of a product. A crude version of this approach has already been used to assess the suitability of several historic and contemporary spaces for different types of events, and is now being refined and extended.

Keywords: room acoustic parameters, acoustic suitability, acoustic quality evaluation.

1 Introduction

Room acoustics is one of the most, if not the most researched field of acoustics since it is experienced and evaluated by anyone having normal hearing. Therefore, people tend to have clear opinions about the quality of acoustics of rooms and spaces, based on life-long experience. The amount of research conducted in room acoustics is best illustrated in review papers such as the one made in 2017 on the influence of room acoustics on human performance [1], with a total of 313 references to corresponding research!

There is certainly no lack of room acoustic parameters defined to help acousticians evaluate the quality of existing spaces, or to design new venues with optimal acoustics for their intended use. For various reasons, many researchers introduce new parameters that better correspond to some specific room acoustic feature. At the same time, the well-known room acoustic parameters are sometimes used in other applications, e.g. in evaluation of the comfort of outdoor spaces [2].

There are numerous recommendations for optimizing room acoustics that define the range of acceptable values for a selected room acoustic parameter. These recommendations always aim towards achieving optimal

acoustics in a room for a specific use. However, not many resources recommend an optimal range for more than one parameter simultaneously, although all room acoustic parameters are connected to a certain degree – if one is changed due to an acoustic intervention, many others will change as well. A good example of a more complex set of recommendations is the Norwegian standard NS 8178 [3] that is focused on performance spaces and gives recommendations on reverberation time for rooms of certain volume, strength and type of music (amplified, weak acoustic and powerful acoustic). Such a complex and multidimensional optimization guideline reveals that there are rooms of certain volume, reverberation time and sound strength that are not particularly suitable as performance spaces at all! Unfortunately, this concept is not easily understood by laymen. Although various charts and diagrams can be helpful when indicating areas of optimum ranges of these parameters, it is hard to believe that such tools will be used by decision makers and investors. However, there are more general recommendations that can be used to indicate important trends in acoustics, such as the one shown in Figure 1.

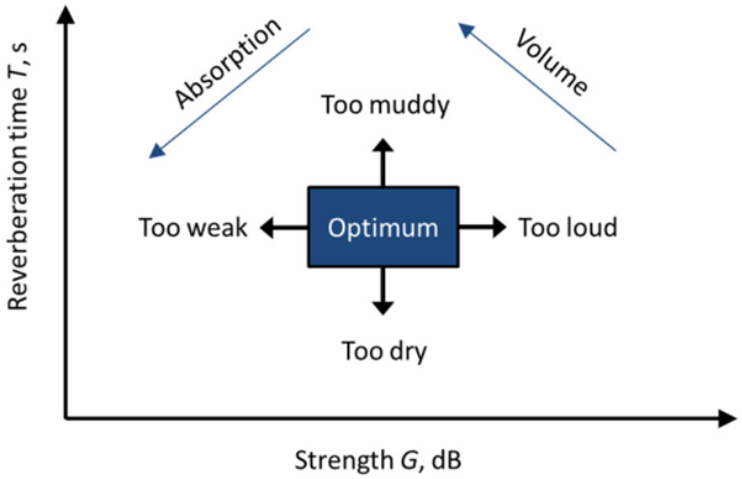


Figure 1: Influence of some acoustic parameters on the reverberation time and strength of rooms, according to Rindel [3].

Since many historical spaces were not built for the purpose they serve nowadays, it cannot be expected that acoustical conditions in these spaces will be appropriate for their current purpose without any adjustments. In some of these spaces it is possible to achieve certain improvement with only minor interventions, whereas others cannot be saved even through thorough reconstruction, as they are fundamentally inappropriate for the events that take place in them. Due to the historical significance of these venues, major interventions that would permanently change their interior are usually not allowed by historians and conservators. The only option that remains is to attempt to improve the acoustic situation through temporary installations that can be easily removed without causing damage.

The usual recommendations about room acoustic parameters are valid for acoustic performances only, without any sound reinforcement system in place. However, the number of performance spaces that require at least occasional, if not regular use of a sound reinforcement system is constantly increasing. Therefore, one must think how the requirements on the acoustics of rooms change to accommodate the presence of a sound reinforcement system, and if certain rooms will be suitable for “electronically assisted” performance without any acoustic redesign. Typically, such rooms serve as venues for contemporary music performances such as pop, rock, jazz or electronic music concerts, etc. A recommendation for the reverberation time for such venues where the sound field generated by the PA system must be controlled by sound absorptive materials, especially in the most critical 125 Hz octave band, can be seen in Figure 2 [4].

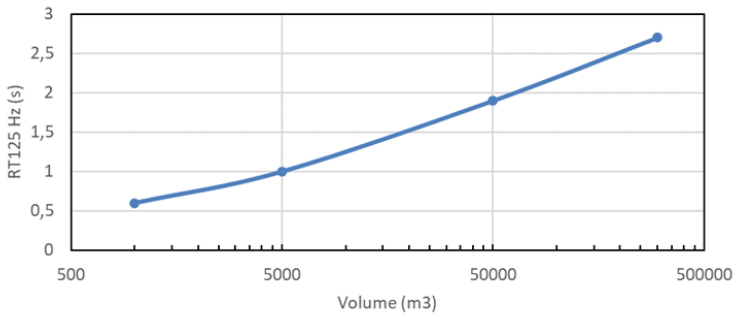


Figure 2: Recommended values of reverberation time in the 125 Hz octave band for performance venues with sound reinforcement, according to Adelman-Larsen [4].

2 Influence of the Covid-19 pandemic on acoustic performances

One of the consequences the worldwide pandemic caused by the SARS-COV-2 virus has on room acoustics of performance spaces is the imposed limitation of the number of visitors that can attend a given event. In some cases, the performances have even been moved from closed to open spaces. Thus, many concert halls, theatres, cinemas and other venues were forced to limit the number of visitors and increase their spacing in the auditorium, and to do the same with performers on stage. Such a change often led to dramatic changes in the acoustics of the room, experienced both by the audience and the performers, particularly if the unused seats were physically removed from the room. A good example can be seen on the left in Figure 3 which shows the concert hall of the Croatian national theatre in the city of Varaždin, built in the usual 19th century shoebox shape, where the main absorption in the hall is coming from the seats and the audience. The right part of Figure 3 shows the same hall during the pandemic, with a significantly reduced number of seats. Consequently, the reverberation time in the hall during musical performances increased to excessively large values, making the hall unsuitable for its usual use, although it was designed for music performances and is well known for its good acoustics.

Figure 3: The concert hall of the Croatian national theatre in Varaždin in its usual setup with performers on the stage and a dense audience seating arrangement (left), and during the Covid-19 pandemic, with a significantly reduced audience (right).

Another example of the influence of Covid-19 pandemics on musical performances can be seen in Figure 4, where musicians had to perform on an improvised stage in front of the Croatian national theatre building in Zagreb. Since this setup had no sound reflectors to direct the sound energy towards the audience, and the venue is surrounded by busy streets, with traffic raising the background noise floor, the only way the concert could

be held was with the help of a sound reinforcement system. Although this was an intermittent solution to the current situation, it was far from ideal in the acoustical sense.

Figure 4: An improvised outdoor concert setup for classical music because of the Covid-19 pandemic.

3 Evaluation of historic spaces in the medieval town of Dubrovnik

Historic spaces with great architectural significance are often used as venues for different kinds of events that tend to enrich the cultural life of a modern city. Built in old city centres, often with considerable restrictions on size and shape, they performed different functions throughout the history of a city, none of which necessarily relates to their current use. A case study was made in four public spaces located in the city of Dubrovnik, a UNESCO World Heritage Site. Its old city is a crown jewel of Croatian coast with shining examples of civil and military architecture, but also a centre of rich cultural life that comes to its peak during high season in the summer.

3.1. Investigated venues

The four selected spaces include the atrium of a historic municipal headquarters, now used for classical concerts of soloists and small ensembles (Rector's Palace), the interior of a stone fortress used as a night club with mostly amplified music reproduction (Revelin), a historic classical theatre for various plays (Marin Držić Theatre), and an improvised concert hall for rehearsals and concerts for a small audience of the Dubrovnik symphonic orchestra located in a stone room of unusual proportions (Slanica). The assessment of their acoustical suitability was carried out through measurements of their impulse responses and calculation of the usual objective parameters: reverberation time RT_{60}, early decay time EDT, clarity C_{50} and C_{80}, strength G, lateral fraction LF_E, binaural quality index BQI, speech transmission index STI, initial time delay gap $ITDG$, and specific volume V/N, where N is the number of seats in a typical seating layout in each venue. Several different criteria were applied to determine the extent to which individual spaces were suitable not only for their primary purpose, but also for other potential uses. Based on the results, recommendations were given for possible improvements of acoustic situation in specific spaces, but such improvements are not likely to be implemented due to the strict rules on preservation of historical buildings. Many more details about the venues and the measurements can be found in [5], while this paper will briefly show only the most interesting outcomes of this case study.

3.2. Results and evaluation of acoustic suitability

The summary of measurement results is shown in Table 1 for all four investigated venues. The usability of these spaces regarding their acoustic performance is evaluated by considering only their primary purpose. The values in Table 1 are coloured according to their suitability for a given purpose. The red-coloured values are considered unsuitable, the yellow-coloured ones are borderline suitable, the green-coloured are suitable, and the black-coloured were not considered for a particular space and purpose. The ranges of suitability for all measured parameters are defined in recommendations found in [6–10].

Table 1: Single-number values of relevant parameters obtained at all four locations

Parameter	Rector's Palace	Revelin Fortress	Slanica Concert Hall	Marin Držić Theatre
RT_{60} (s)	3.80	1.34	1.20	0.78
EDT (s)	3.72	1.30	1.20	0.72
C_{50} (dB)	-8.1	-1.3	-1.3	2.1
C_{80} (dB)	-5.5	1.6	1.6	5.9
G_{mid} (dB)	7.0	0.7	5.7	1.6
LF_E	0.20	0.10	0.23	0.12
BQI	0.74	0.63	0.72	0.71
STI	0.36	0.58	0.56	0.68
$ITDG$ (ms)	24.50	2.56	3.74	4.20
V (m^3)	2900	1500	685	1640
N (s)	300	220	170	283
V/N (m^3)	9.7	6.8	4.0	5.8

This visual aid for evaluation of the acoustic suitability of spaces can be helpful for quickly locating parameter values of spaces that are outside their optimal range. In general, one could conclude that the Revelin Fortress might be acoustically best for its purpose, and the Rector's Palace atrium the worst. In reality, the atrium was measured in the unoccupied condition. During concerts, the acoustical situation changes drastically, since the audience introduces a lot of sound absorption. As a result, the concerts held there are usually described in superlatives. On the other hand, the Revelin fortress might serve fine for its current purpose. However, having a night club in an UNESCO-protected medieval town introduces many other administrative and organizational burdens and problems for the municipality and calls for its relocation to a more suitable location.

Another problem of this colour-coded table is that it gives no additional information about the suitability of these venues for other purposes (which was the goal of this project to start with), and there is no information on the "weight" of the presented numbers. In other words, non-professionals cannot judge if a red-coloured value of the clarity bears the same significance as a red-coloured value of reverberation time.

Another attempt was made to present the suitability of these spaces for other purposes. An example for the Revelin fortress is shown in Table 2 with colour-coded suitability marks (red for unsuitable, yellow for borderline suitable and green for suitable). The biggest difference between this way of presenting the suitability of acoustics spaces and the one shown in Table 1 is that here one can visually compare the difference in suitability between various applications of the space. It is obvious that the Revelin fortress has more green-coloured fields for opera than for chamber music, and even less for symphonic music. But this presentation still gives no information on if some of the table fields are more important than others, or if we just have to "count" the good fields to find the best use for a space.

Although far from perfect, the authors believe that this presentation of measured values and their position within

Table 2: The suitability of Revelin Fortress for various purposes based on the mean values of typical acoustic parameters.

Revelin Fortress	RT_{60} (s)	EDT (s)	BQI	G_{mid} (dB)	C_{80} (dB)	ITDG (ms)	V/N (m^3/pers.)	STI
Measured values	1.34	1.30	0.63	0.7	1.6	2.56	6.8	0.58
Symphonical music								
Chamber music								
Opera								
Speech								
Cinema								

the optimal range of acoustical parameters for various uses of a venue represents a simple, yet a somewhat limited way of showing the suitability of a venue for certain types of events. Such displays are invaluable for non-acousticians who have trouble interpreting the numerical values of the acoustical parameters and grasping their meaning for a given venue. Moreover, such a representation can help the local municipality make decisions on possible use of venues, both historic and contemporary ones. The same visual principle was already applied in other recommendations for room acoustic parameters, or even the physical parameters of a room. For example, Figure 5 from [11] shows a coloured display of favourable and unfavourable ratios of room length to its height, as well as the ratios of room width to its height that lead to minimising the influence of strong room modes in a small rectangular room.

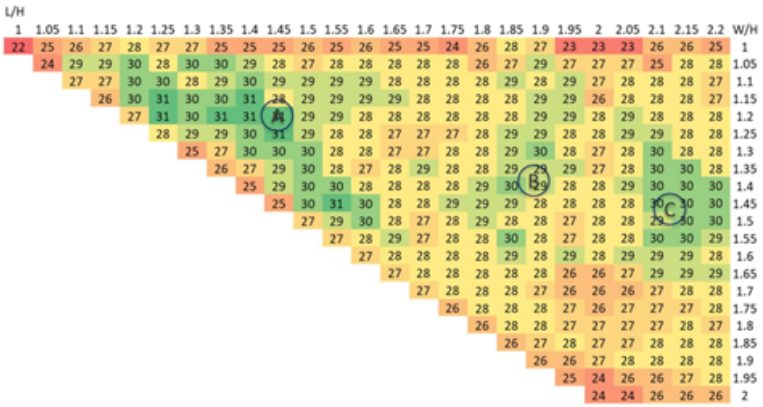


Figure 5: Colour-coded optimization chart for room aspect ratios in order to maximise the number of tones supported by room modes, according to Rindel [11].

4 Acoustic labelling

In this section, the authors tend to establish a simple, easy-to-understand approach to evaluation of suitability of different rooms and halls as venues for various types of events, especially if they were not originally built for a specific purpose. The main goal is to help the non-experts such as managers of such halls understand what their hall is suitable for, and what steps to take, should the evaluation reveal the hall to be unsuitable. To establish the method of evaluation, the authors rely on the German standard DIN 18041:2016 [12]. The earlier version of this standard is in force in the authors' home country of Croatia. The most recent version has not been approved for use at this time due to lack of demand, as stated by the relevant regulatory body. This gives a quite accurate illustration of the present situation in Croatia, with very few experts who use the standard only as a guideline,

as room acoustic design is not covered by any laws or regulations. Nevertheless, the authors find this standard to be an invaluable resource in both the design stage for new halls and the evaluation stage for the existing ones.

The approach to evaluation of a hall illustrated in this section relies on three main questions that need answering:

1. Are the room acoustics conditions in the hall appropriate for a given type of event?
2. Is the seating capacity appropriate for the type of event/activity?
3. Will the sound source be loud enough on its own, or will it need the help of a PA system?

To answer the first question, the approach utilizes the requirements on reverberation time for different types of activities as a function of room volume, as given in [12]. The second question can be answered by finding the specific volume, i.e. the volume per seat, and by comparing it with the typical values given in [12]. The third question can be answered knowing the geometry of the room and the properties of the sound source.

The evaluation procedure shall be shown on an example room originally built and used as a classroom. The room volume is 280 m^3, the mid-frequency reverberation time is 1.15 seconds, and the distance from the source to the farthest listener is 10 metres. The room has 72 seats. For this evaluation, it is assumed that the noise level in the room is 45 dBA, that the people in the audience do not suffer from hearing problems, and that the event will involve communication in native language only.

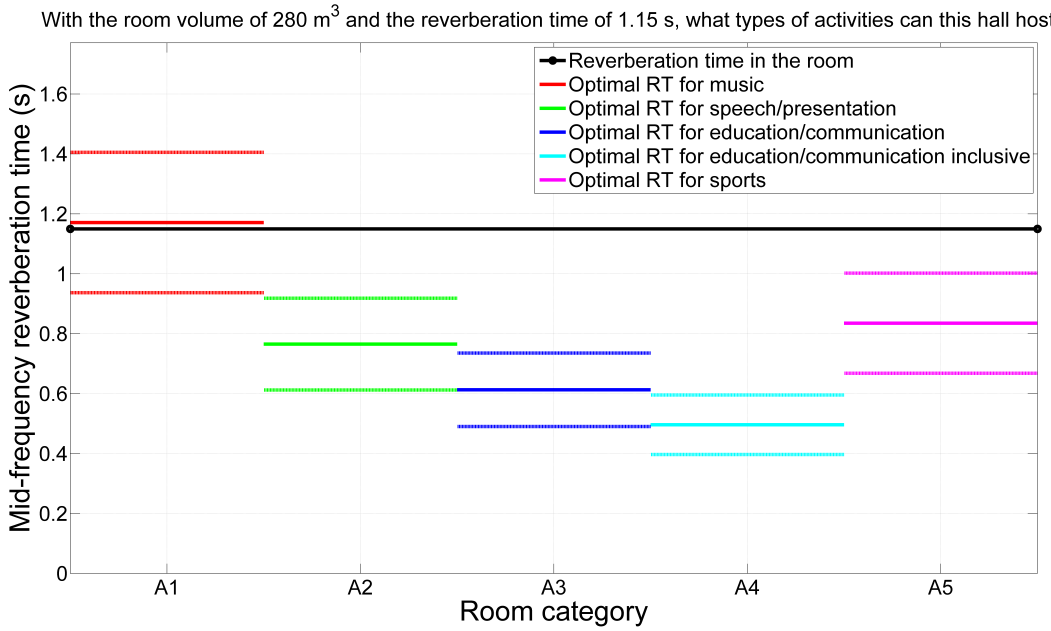


Figure 6: The reverberation time in the example room vs. optimal mid-frequency reverberation times (and tolerance) for different types of activities, calculated for the volume of the example room according to [12]

Figure 6 shows that the reverberation time in the example room makes the room adequate only for music performance, and that the room is too reverberant for any other purpose. If the room is to be used for speech or educational activities, acoustic treatment is needed to shorten the reverberation. To retain the ability of the room as a venue for activities connected with both speech and music, variable acoustics can be implemented up to a point, thus providing the possibility to adjust the reverberation time at least within the tolerance values found for different types of events.

The top chart in Figure 7 shows that the volume of the example room and the current number of seats dictate that the room be used solely for speech-based activities. The bottom chart in Figure 7 offers advice to hall managers and staff by converting the requirements on specific volume shown in the top chart into a range of

seating capacity that is adequate for each type of activity. If the hall is to be used for speech, it will support from 47 all the way to 93 listeners, given that there is room for so many people. If the listeners are to enjoy musical performance, their number should be from 23 to 40. For a mix of speech and music, the number of listeners should range from 35 to 47.

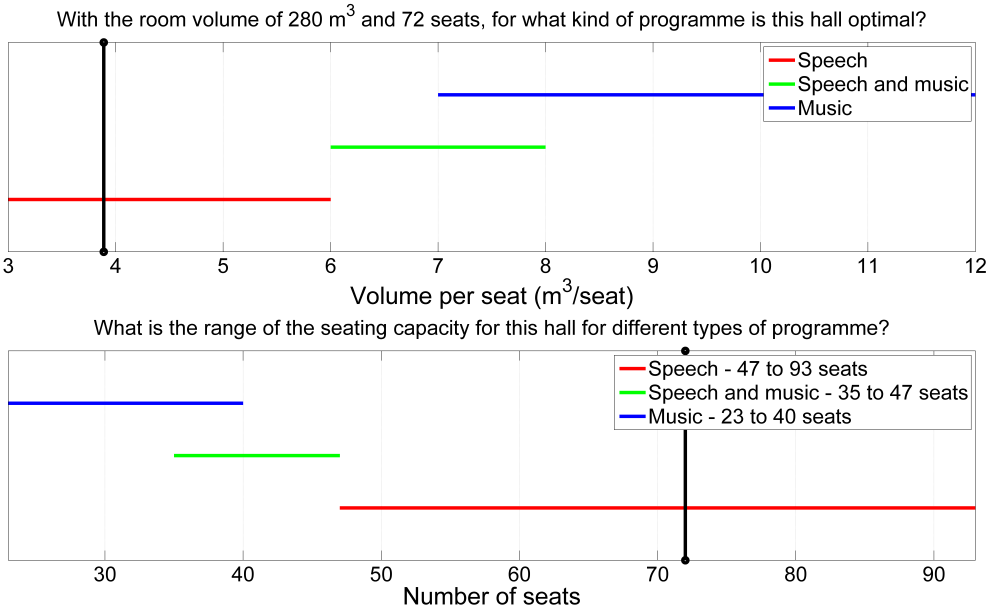


Figure 7: The suitability of the example room with the current number of seats (top), and the ranges of seating capacity for different types of events (bottom), according to [12]

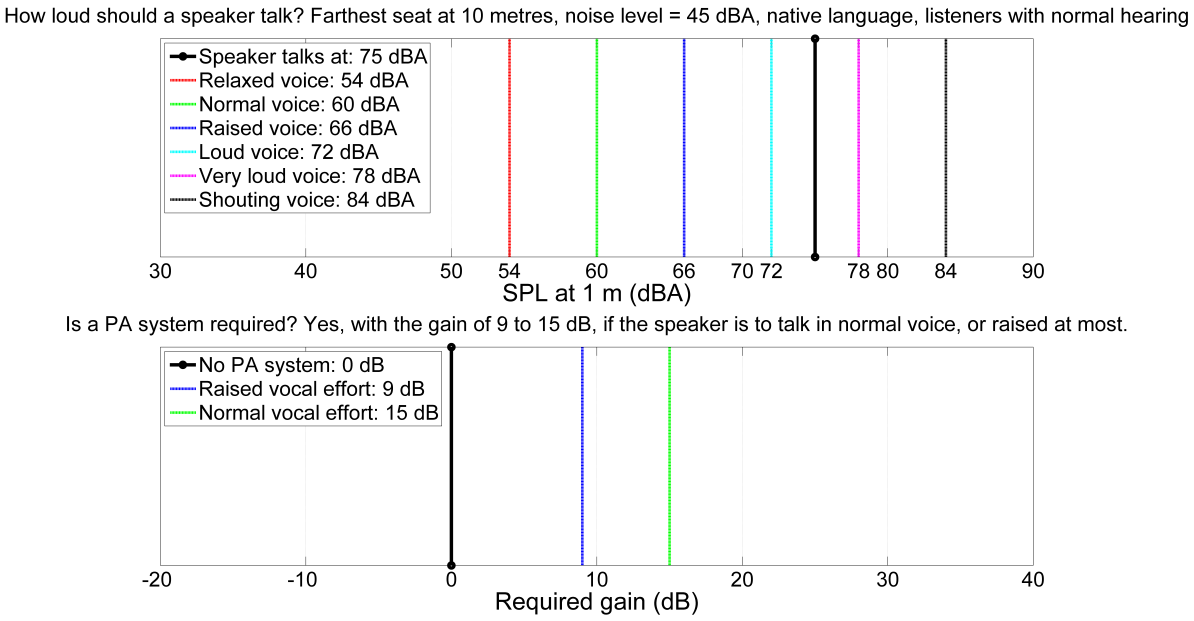


Figure 8: The vocal effort of the speaker without a PA system required to reach even the farthest listener while maintaining the required signal-to-noise ratio (top), and the gain of the PA system (if needed) that allows the speaker to talk in normal/raised voice (bottom).

Finally, Figure 8 shows the vocal effort the speaker will have to put in to reach the listener sitting in the farthest seat of the room. For this evaluation procedure, the authors assume that the level of direct sound of the speaker at the farthest listening position should be 10 dB higher than the level of background noise if native language is used in communication and the listeners have normal hearing (as is the case in this example). For listeners with impaired hearing or for non-native language communication, the difference, i.e. the signal-to-noise ratio should be 20 dB. With the noise level of 45 dBA, the speaker will have to speak at 75 dBA at 1 meter to reach the farthest listener at 10 metres, i.e. they will need to speak loudly or very loudly to accomplish this task, as shown on the top chart in Figure 8. As this effort would either cause fatigue in a rather short time or be impossible for some speakers, a PA system is required, so that the speaker can still accomplish the task while speaking in normal or, at most, raised voice. As the speaker speaking in raised (normal) voice without the PA system would produce 9 (15) dB lower level than needed at the farthest seat, the gain of the PA system needs to be set at these exact values for the PA system to compensate for this shortage. This is shown on the bottom chart in Figure 8.

5 Conclusions

Many researchers in the field of room acoustics already use complex metrics to provide information about the quality of room acoustic design of many types of spaces. They often use a combination of two or more acoustic parameters to form the scientific proof that certain designs will be more efficient than others. This approach, while scientifically sound, is usually very difficult to grasp.

The authors of this paper investigated how the acoustic suitability of spaces used for various applications for which they were not specifically designed and built can be presented to non-acousticians in an easy-to-understand way. The main hypothesis was that the necessity for an acoustic intervention in performance and other spaces can be demonstrated to laymen in a simpler, more intuitive way than by showing numerical values of room acoustical parameters that are beyond the comprehension of non-acousticians.

At the same time, the pandemics of the Covid-19 virus forced many managers of performance spaces to change the usual usage of these spaces by seating capacity and the number of performers, or to use unconventional spaces for acoustic performance events, in compliance with anti-pandemic measures. The experiences gained over the past two years show that these improvisations rarely result in acceptable acoustics. Therefore, an easy way to present these facts to decision makers is required now more than ever.

In this paper, a case study was presented in which the acoustic suitability of several historic spaces in a medieval town was demonstrated to the managing, decision-making bodies. The emphasis was put on the suitability of these spaces for events they do not typically host. The tool used here were easy-to-interpret colour-coded tables. Although this method proved to be quite effective in this particular project, it cannot be universally used since it depends on the known optimal ranges of (too) many specific acoustic parameters, with some of them being appropriate only for specific types of spaces, e.g. for music performance spaces.

To build on this concept, another way of acoustic labelling of the usability of spaces for various applications was presented. The graphical representation was used to check and demonstrate if the reverberation time in a room is appropriate for different types of activities, if the seating capacity is appropriate for the type of event/activity, and, finally, if the sound source in the room will be loud enough on its own or if a PA system should be used. The concept was explained on a prototype room.

The authors believe the presented approach can help acousticians identify the range of events for which a certain room is suitable, according to its design. More importantly, it allows the experts to explain their findings to non-experts such as managers or investors using simple, understandable language, and to explain the basic steps that need to be taken to improve the acoustic conditions in a space if necessary.

References

[1] J. Reinten, E. Braat-Eggen, M. Hornikx, H.S.M. Kort and A. Kohlrausch. The indoor sound environment and human task performance: A literature review on the role of room acoustics. *Building and Environment*, 123: 315-332, 2017.

[2] A. Taghipour, S. Athari, A. Gisladottir, T. Sievers and K. Eggenschwiler. Room Acoustical Parameters as Predictors of Acoustic Comfort in Outdoor Spaces of Housing Complexes. *Front. Psychol.*, 11:344. doi: 10.3389/fpsyg.2020.00344, 2020

[3] J. H. Rindel. New Norwegian standard on the acoustics of rooms for music rehearsal and performance. *Proceedings Forum Acusticum 2014*: 1-6, 2014.

[4] N. W. Adelman-Larsen. Rock and Pop Venues, Acoustic and Architectural Design. *Springer*, New York, 2014.

[5] K. Jambrošić, M. Horvat, H. Domitrović, A. Petošić. Acoustical Suitability of Historic Spaces as Venues for Modern-Day Events. *E-book of reviewed papers - ATF 2014*: 16-22, 2014.

[6] ISO 3382-1:2009 Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces

[7] L. Beranek. Concert halls and opera houses: music, acoustics and architecture. *Springer-Verlag*, New York, 2004.

[8] T. D. Rossing. Springer Handbook of Acoustics, *Springer Science+Business Media*, LLC New York, 2007.

[9] M. Long. Architectural Acoustics. *Elsevier Academic Press*, Burlington, 2006.

[10] M. Barron. Auditorium Acoustics and Architectural Design. *Spon Press*, New York, 2009.

[11] J. H. Rindel. Searching the musical rehearsal room. *Proceedings BNAM 2020*: 1-12, 2020.

[12] DIN 18041: 2016 Acoustic quality in rooms - Specifications and instructions for the room acoustic design

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Alternative Spaces for Learning: Reversible Acoustic Treatments for Transformation of Spaces into Classrooms During COVID Era

Daniela Ilaria Schiavon[1], Louena Shtrepi[1,*], Arianna Astolfi[1]

[1]Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

*louena.shtrepi@polito.it

Abstract
The design of acoustic comfort in classrooms is a challenging issue. It became a crucial aspect during the Covid-19 pandemic regarding class life organization, as for the latter large spaces for teaching were required in order to guarantee the minimum distance between the occupants to limit the spread of Covid-19, and it was important to ensure their acoustics functionality for the performance of lessons. To this aim, this investigation is to provide a guide (for school principals, administrators, safety managers, architects and engineers, acousticians) where easily implementable solutions for different types of school buildings are detailed in various case studies and they are an example in case of environments similar in shape and volume. In particular the focus is put on the transformation of school spaces created for different purposes than normal classrooms (e.g., corridors, atria, sports halls) into classrooms for teaching. Different layouts of the furniture and sound-absorbing materials have been applied and analysed by simulations (Odeon 15) comparing the results with the standardized optimal values and with those of the actual environments obtained through an extensive measurement campaign. Eight schools were taken into account as case studies, representing the Italian school heritage and including elementary schools, middle schools and high schools. A total of 26 different spaces with a volume varying in the range 135-2800 m^3 were analysed. For each of them, both general analyses at the overall environment level and more specific ones for single receivers located over the area occupied by the students were carried out. As a result, it should be highlighted that the proposed solutions could not be fully acoustically optimized for all of the spaces, as the analyses also took into consideration an affordable cost, the speed of realization and the reversibility of the intervention.

Keywords: classroom, Covid, reversible acoustics, simulation

1 Introduction

In Spring 2020, Politecnico di Torino as a leading organization brought together different working groups in order to address the challenges of the new working, learning and living conditions due to the Covid-19 pandemic [1]. The result of these working groups was a series of reports [2] that included guidelines for an optimized use of spaces and devices, as well as good behavioral practices that could limit the spread of the virus and allow the running of the activities, as close as possible, to a normal and inclusive framework. In particular, the work presented here reports the investigation regarding the transformation of school spaces created for different purposes (e.g., corridors, atria, sports halls) into classrooms where speech comprehension is fundamental, by evaluating their acoustic comfort conditions and by proposing possible desks arrangements and acoustic treatments. The goal was to guarantee functional acoustic conditions to properly conduct lessons, when there is the need of having additional and bigger spaces in order to put more distance between desks, as per the new rules originated by the Covid-19 situation.

The investigation of the state of the art on similar actions resulted quite limited within the acoustic field both in terms of case studies and design solutions that could be readily applied in practice. Several works can be found regarding reversible acoustic interventions in historical buildings [3, 4]. In Italy, it is well known the adaptive reuse of existing historical buildings into schools, however the acoustic result is often the outcome of thermal retrofitting actions [5].
The research on reference values of the acoustical parameters that have the greatest influence on student performance is ongoing research [6]. In this investigation, the acoustic results based on "minimal-reversible intervention" criteria were compared to the standardized optimal values [7] and to those of the actual environments obtained through an extensive measurement campaign, in order to obtain a useful guide for school principals, administrators, safety managers, architects and engineers, and acousticians.
Easily implementable solutions for different types of school buildings are detailed in this study and they constitute examples in case of similar rooms in shape and volume to be adapted. The proposed solutions show the resilience of extreme spaces (e.g., gyms) to host the demanding conditions required for teaching-learning performance. However, it should be highlighted that the solutions reported in this study are not fully optimized, as the analyses also took into consideration an affordable cost, the velocity in the setting-up and the reversibility of the treatment.

Table I: List of all the spaces with associated school level, floor area, volume and current function with the number of classes and students hosted in the new configuration. UNI 11532-2 optimal reverberation time value is indicated. The spaces that had already some acoustic treatment are signed with *.

Level of school	**Function**	**Area [m^2]**	**Volume [m^3] (V_{class})**	**Number of classes - Number of students**	**Optimal reverberation Time [s]**
primary	auditorium*	83	324	1 cl. 28 st.	0.6
	gym	84	435	1 cl. 28 st.	0.7
	canteen*	40	197	1 cl. 20 st.	0.6
	canteen*	42	204	1 cl. 20 st.	0.6
	canteen*	39	194	1 cl. 16 st.	0.6
primary	canteen*	44	135	1 cl. 16 st.	0.5
	canteen*	158	499	2 cl. 20+20 st.	0.7
	gym	84	1208	2 cl. 20+20 st.	0.8
primary	computer room*	93	257	2 cl. 16+16 st.	0.6
	reading room*	70	193	1 cl. 20 st.	0.6
	lobby*	98	269 (121)	1 cl. 14 st.	0.6
	lobby	98	293 (130)	1 cl. 14 st.	0.6
	lobby	75	225 (98)	1 cl. 14 st.	0.6
	gym*	298	2018	2 cl. 25+25 st.	0.9
	theatre*	297	990	3 cl. 24+24+24 st.	0.8
	canteen*	287	832	4 cl. 20+20+15+15 st.	0.8
middle	lobby	157	502	1 cl. 18 st.	0.7
	lobby	156	499	1 cl. 16 st.	0.7
middle	canteen	122	371	2 cl. 20+16 st.	0.7
	media room*	98	298	2 cl. 16+16 st.	0.6
	lobby	125	398 (142)	1 cl. 15 st.	0.7
	science room	97	311	1 cl. 18 st.	0.6
	mixed use room	97	310	1 cl. 28 st.	0.6
high	canteen*	194	776	2 cl. 16+24 st.	0.8

2 Methods

Eight schools were analysed in this study. The schools represent the Italian school heritage, which often does not meet up with the current acoustic standards and which is characterised by buildings of the nineteenth and early twentieth century, of the Sixties and Seventies and a small percentage of new buildings recently built up. The actual acoustic conditions of the five case studies were checked through a measurement campaign, later used to calibrate the models in the acoustic simulation phase.
A total of 26 different spaces with a volume varying in the range 135-2800 m^3 were analysed. For each of them, both general analyses at the overall environment level and more specific ones for single receivers located over the area occupied by the students, were carried out. It was not possible to perform acoustic measurements in two spaces, thus the data reported will focus only on 24 spaces.
The investigation was a methodical step by step analysis which started from an inspection of the case studies followed by an extensive in-field campaign in which the reverberation time and background noise levels was measured. A categorisation of each of the 24 spaces was performed by their function and dimensions (Table I). For each space a proper desk arrangement was proposed in line with the Covid-19 distance requirements, and the optimal reverberation time target values were identified for furnished and 80% occupied room. The minimum acoustic treatments were verified by a first analysis with the Sabine formula and later on, simulations were run with Odeon 15 software, after the calibration upon the measurement result was carried out. Noise conditions of each space was set through the NC curves, chosen appropriately with respect to the number of children in the space. After the speaker and the receivers were positioned over the seating area, STI (Speech Transmission Index) punctual specific STI and speech SPL (Sound Pressure Level) values were obtained.

2.1 Case study selection and description

Case studies included three elementary schools, two middle schools, two high schools and one school where all levels were present. They represent typical Italian school buildings, which span from old masonry vaulted buildings to concrete ones, to mixed concrete structure and masonry walls buildings. As mentioned above, since it was not possible to perform acoustic measurements in two spaces in one high school and the one where all levels were present, these were excluded from the data reported here.
Spaces with different volumes and shapes were chosen to include all the possible acoustic scenarios. In some cases, the acoustic treatment was already present, albeit with materials characterised by low-medium sound absorption. It must be highlighted that in case of a lobby, which is not a closed area, the volume and the floor area were an arbitrary chosen in order to have a valid model to run simulation.

2.2 Acoustic parameters

UNI 11532-2 has been used as reference standard for classroom acoustics [7]. It provides updated acoustic parameter and their reference values, which are in line with those applied internationally. In particular, the standard is aimed:

- to guarantee an effective speech comprehension for the students;
- to guarantee the minimum vocal fatigue and vocal effort for the teacher;
- to reduce chatting noise in case of the presence of simultaneous classes.

In the study, which has been carried out in order to be immediately put in practice in the case of emergency due to Covid-12 pandemic, it was decided to provide basic solutions that could be sufficient to obtain between fair and good acoustics in classroom.
Reverberation time (T), Speech Transmission Index (STI) and speech A-weighted Sound Pressure Level, SPL(A), were considered as the main acoustical parameters. Target values for T and STI are reported in the UNI 11532-2, where the school environments spaces are divided in 6 categories. Category A3 was chosen (Eq.1), that is aimed to lesson/communication, with teacher-student interaction.

The optimal value for reverberation time is obtained with the following formula, which can be applied for a volume range 30 $m^3 \leq V < 5000$ m^3:

$$T_{ott,A3} = 0.32 log V - 0.17 \quad (1)$$

Table 1 reports the optimal reverberation time values for each volume. UNI 11532-2 indicates as the optimal values for STI, STI ≥ 0,55 for spaces with a volume V < 250 m^3 and STI ≥ 0,50 for spaces with a volume V ≥ 250 m^3 without sound amplification system. The standard was used a guideline for good practice. These target values were set to be reached as close as possible in order to allow a significant improvement in acoustic comfort compared to the starting condition.

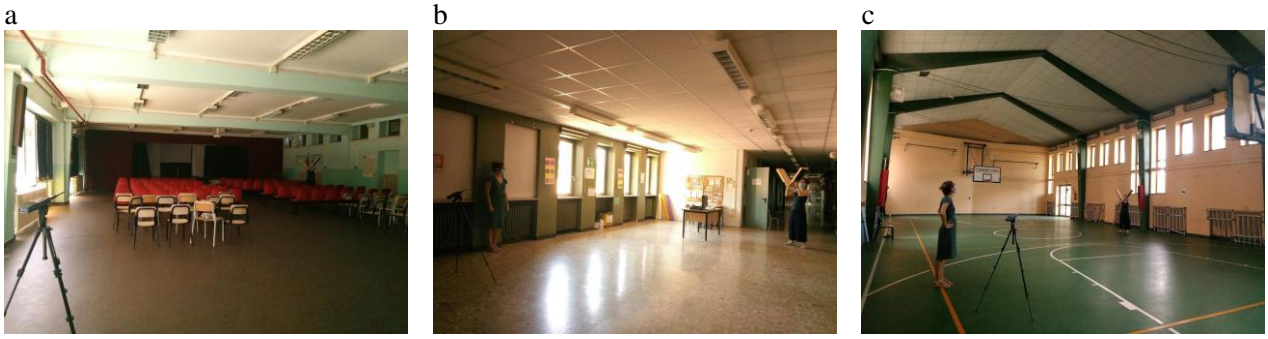


Figure 1: Measurement campaign. Case studies a) auditorium, b) lobby, and c) gym.

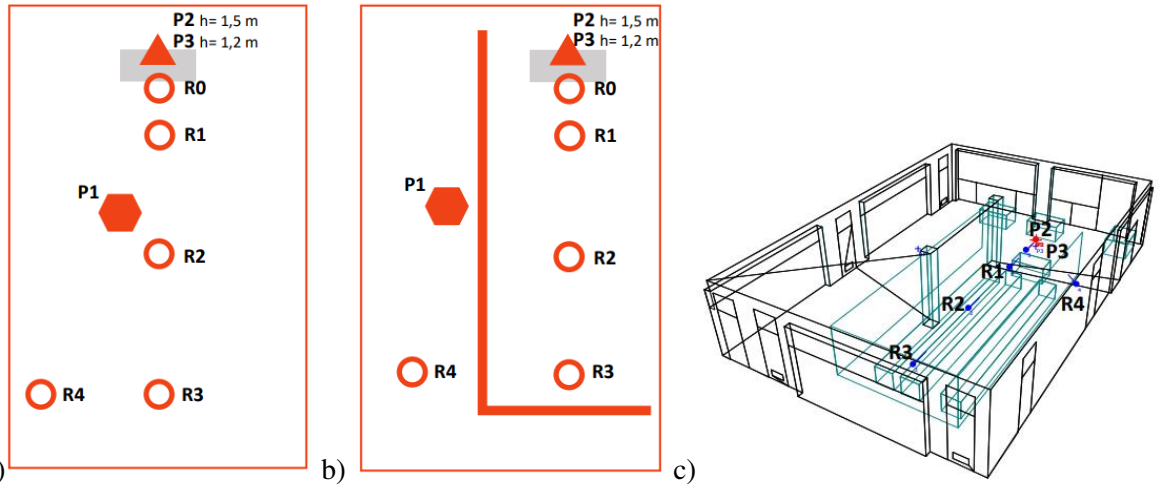


Figure 2: General scheme of the measurement (a) and simulation (c) layout with source P1-P3 (P1-omnidirectional and P2-P3 normal-speaker) and receivers R1-R4 (h=1.1 m in primary school and 1.2m in high school). Receiver R0 is at 1m distance from the source and at the same height h=1.5m. Scheme b) shows the distribution of the receivers when the separation between simultaneous classes occurs.

2.3 Measurement and simulation campaign

A reverberation time measurement campaign was run for 24 out of the 26 analysed spaces (Figure 1). Odeon 15 was used to simulate the different spaces at the different steps: (i) current configuration, which was calibrated based on in-situ measurement, (ii) new configuration without acoustic treatment, and (iii) new configuration with acoustic treatment. A transition order TO=2 and 4000-8000 rays have been used for all the configurations depending on the volume size. Two source and four receiver positions have been considered. One omnidirectional source (P1) and two sources (P2-P3) simulating the human voice directivity at normal speech level (namely TlkNorm in Odeon) at two different heights (sitting and standing up teacher) have been

tested. Only P2 has been used in the present work. STI was calculated assuming a female speech spectrum (Figure 2). Table II shows the NC curves applied for STI calculation both for the presence or the absence of the sound absorbing panels i.e. acoustic treatment and for the presence of one class or two/three/four classes in co-presence. Table I shows the spaces that were divided in more than one classroom.
Figure 3 shows the results of the measurement campaign, i.e., the mean reverberation time across the octave-band center frequencies 250, 500, 1000 and 2000 Hz vs rooms volumes.

Table II: NC curves used to simulate STI.

Number of classes simultaneously in the space	Acoustic treatment	
	yes	no
1	NC 30 – 39.7 dB(A)	NC 30 – 39.7 dB(A)
2	NC 35 – 44.2 dB(A)	NC 45 – 53.4 dB(A)
3	NC 40 – 49.0 dB(A)	NC 50 – 58.1 dB(A)
4	NC 40 – 49.0 dB(A)	NC 50 – 58.1 dB(A)

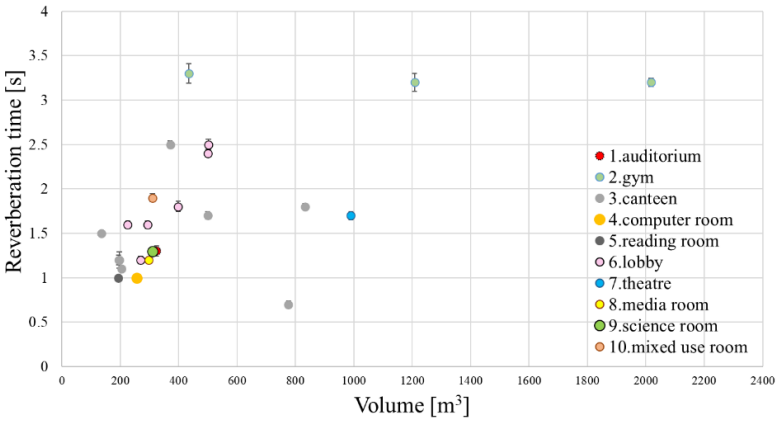


Figure 3: Relationship between the mean reverberation time (Tm) across the octave-band centre frequencies 250, 500, 1000 and 2000 Hz and the volume of each space.

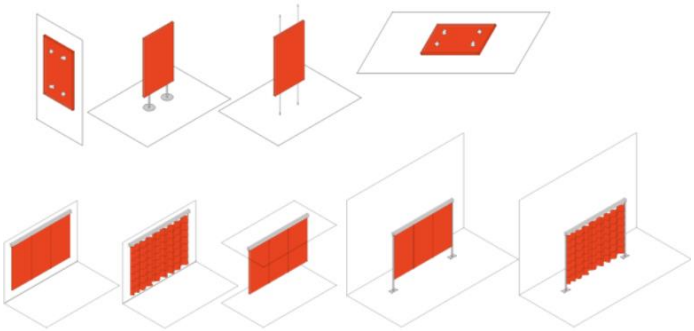


Figure 4: Basic acoustic treatment. These coloured schemes are used to indicate which solution was proposed for each space.

2.4 Selection of minimal acoustic treatments

As a first criteria for the selection of the acoustic treatment, only light and easy-to-handle acoustic materials and systems were chosen to transform different spaces into school environments suitable for teaching. The

main objective of this work was to arrange reversible treatments, fast-settled and well fitted into the existing context.
Only sound absorbing porous material with weighted sound absorption $\alpha_w \geq 0,8$ and sound insulating separation panels with $R_w \geq 18$ dB were considered. An acoustic reflective shell behind the teacher was added in larger spaces with high or vaulted ceilings or in case of absence of the wall behind or in case of lateral walls located at an excessive distance. The proposed solutions (Figure 4 and 5) are stand-alone panels to be hung like paintings on the walls or simply on the ceiling, free standing panels to be placed on the floor, or free standing baffles or acoustic curtains mounted on the ceiling or the walls.
Only in two cases, further acoustic treatment of the ceiling and lateral walls was proposed. These two cases represented an example of further treatment for all similar environments, which with the basic acoustic treatment cannot reach the optimal acoustic values.

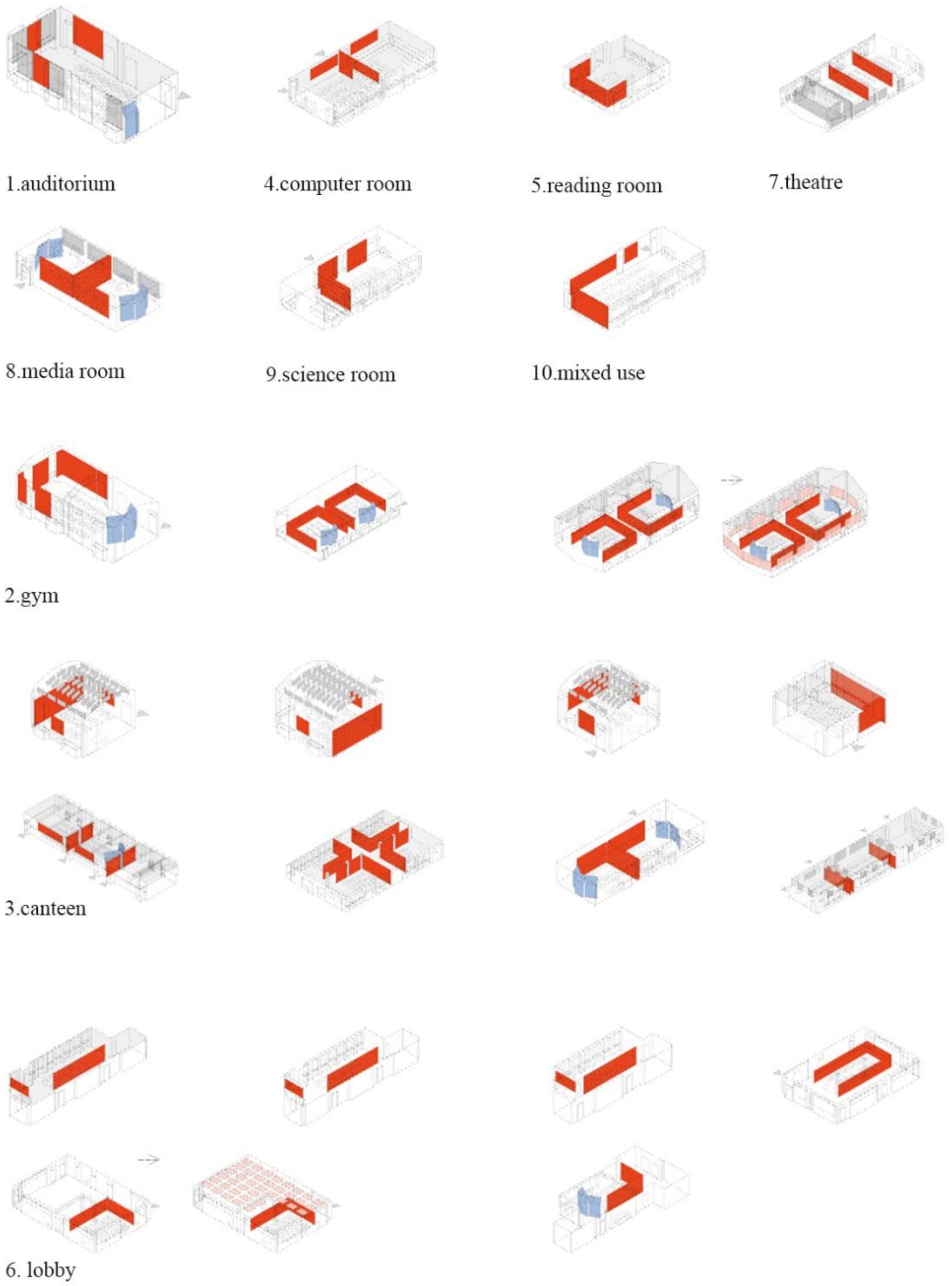


Figure 5: All spaces are divided upon their function with the indication of the extension and position of the acoustic treatment (orange red) and an acoustic reflective shell behind the teacher location in large spaces (light blue).

3 Results

Figure 6 shows the average, across frequencies, reverberation time T for each space with the basic acoustic treatment and the extra optimal treatment, in comparison with the measured reverberation time. Spaces are classified from 1 to 10 based on their function. It can be noticed that the basic treatment applied could reach the optimal values in most of the cases. However, given the unconventional volumes and space divisions it could not be fully achieved in all of them. Two of these cases have been further analysed with additional acoustic treatment (e.g. a 6.lobby and a 2.gym) applied as shown in Figure 5. This second application aimed at showing the significant effect on acoustic parameters (mainly T_m and STI) and the material extension needed to achieve a further improvement. It should be highlighted that this might have a considerable impact on the costs and speed of execution of the project. These two cases of optimal treatment, chosen on a sample basis, represent an example of further treatment in a second step for all similar environments, which with the basic acoustic treatment cannot reach the optimal acoustic values.

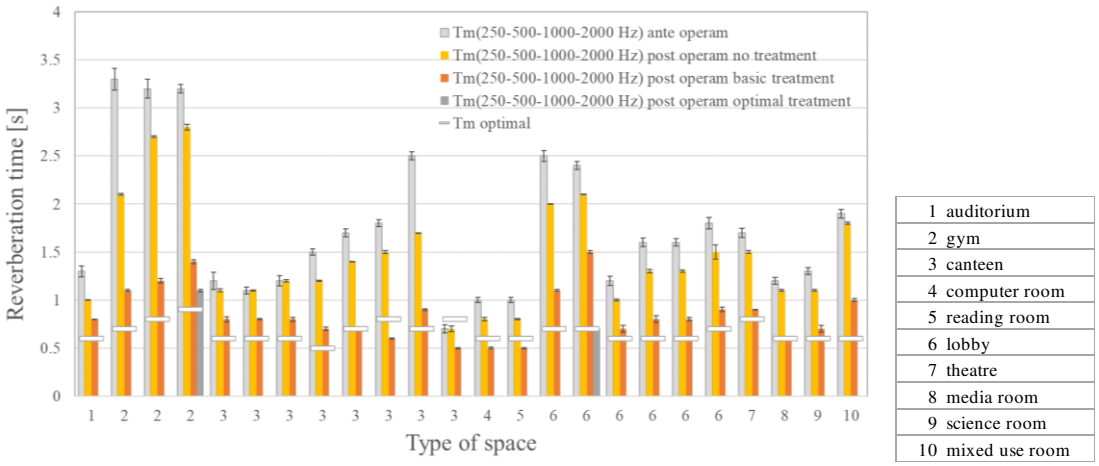


Figure 6: Actual reverberation time and *post operam* reverberation time (results for both the basic and the optimal acoustic treatments are shown for 2.gym and 6.lobby).

Figure 7 shows the average STI values across three or four receivers. The number of receivers was different for each space because in case of lobbies or large spaces with more than one class, a fourth receiver was placed outside the teaching area (Figure 2b), to check for reflections and speech level.

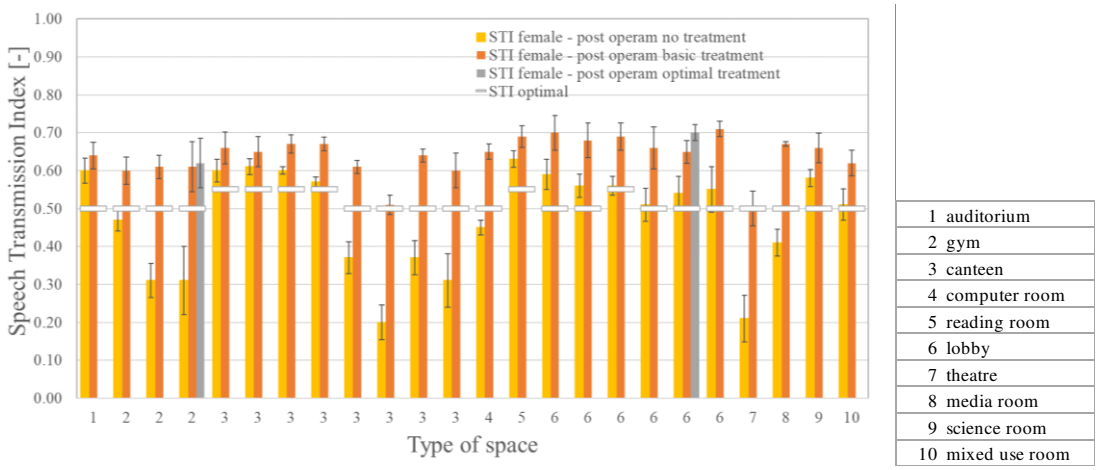


Figure 7: *Post operam* STI (results for both the basic and the optimal acoustic treatments are shown for 2.gym and 6.lobby).

Figure 8 shows the average speech SPL(A) values across three or four receivers simulated with Odeon software using a female speaker with a normal vocal effort.

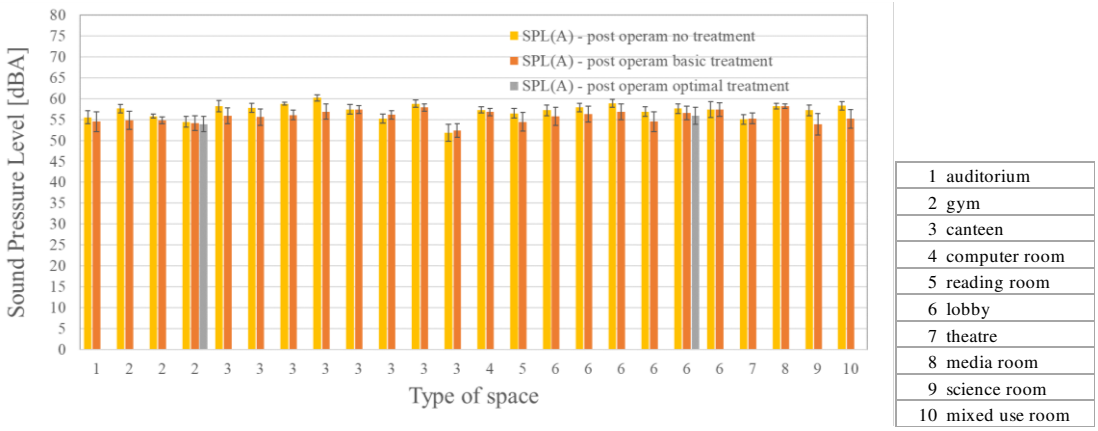


Figure 8: Average value of three of four receivers SPL (results for both the basic and the optimal acoustic treatments are shown for 2.gym and 6.lobby).

4 Discussion and conclusions

This study was carried out from the need of having additional teaching spaces with good acoustics, where it is possible to ensure the required distance between students to limit the contagion from Covid-19, on one side, and an effective teaching condition, on the other side.
The proposed acoustic treatments ensure the reversibility, the fast-settlement, and the good acoustics with respect to the requirements of one of the most updated Standards in the field of classroom acoustics. Some limitations in the achieved target values have been considered acceptable given the unconventional conditions of the treated spaces. However, two cases have been further implemented in order to achieve the standard optimal values. This helped to quantify the design effort and costs for extra acoustic treatment.
The investigation highlighted the absence of acoustic treatment or systems that can be used in emergency conditions as ready-to-use products. However, in conclusion, the proposed solutions can be adopted as guidelines to be used for similar construction types, shapes and volumes. They represent a compromise between good acoustics and reversible and costly affordable treatment. In case of more complex situations, deeper study must be carried out.

Acknowledgements

The authors would like to thank all the Beta Test Schools for their participation to this investigation although it took part during a critical period of the beginning of the pandemic, in Spring 2020. The authors are also grateful to Politecnico di Torino for providing financial support in performing this research.

References

[1] Piano scuola 2020-2021 – Documento per la pianificazione delle attività scolastiche, educative e formative in tutte le Istituzioni del Sistema nazionale di Istruzione, Ministero dell'Istruzione, 26/6/2020.

(School plan 2020-2021 - Document for the planning of school, educational and training activities in all institutions of the national education system, Ministry of Education, 26/6/2020)

[2] Scuole Aperte, Società Protetta (Open Society, Protected Schools) (2021), Available at: https://www.impreseaperte.polito.it/i_rapporti (Accessed: 23 March 2022).

[3] M. Kaššáková, L. Kritly, L. Zelem and M. Rychtáriková. Proposal for reversable intervention to historical monument by means of room acoustic simulations, *Special 30th SKAS Anniversary Issue of Proceedings/2020*, 107-114, 2020.

[4] COST-STSM-TU1301-33175 report "Comparative study of the acoustic comfort in atria, covered by foiland glass-based materials"

[5] S. Secchi, A. Astolfi, G. Calosso, D. Casini, G. Cellai, F. Scamoni, C. Scrosati, L. Shtrepi. Effect of outdoor noise and façade sound insulation on indoor acoustic environment of Italian schools. *Applied Acoustics*, 126, 120-13, 2017.

[6] G. Minelli, G.E. Puglisi and A. Astolfi. Acoustical parameters for learning in classroom: A review. *Building and Environment*, 208, 108582, 2022.

[7] UNI 11532-2, Caratteristiche acustiche interne di ambienti confinati - Metodi di progettazione e tecniche di valutazione - Parte 2: Settore scolastico/Acoustic characteristics of indoor environments – Design methods and evaluation techniques – Part 2: school sector, Ente Italiano di Normazione, 2020.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

The development of modern, interactive acoustic courseware material within the Acoustics Knowledge Alliance project

Marko Horvat[1,*], Karolina Jaruszewska[2], Samuel Raetz[3], Emilie Carayol[4], Yannick Sluyts[5], Andreas Herweg[6], Lukas Aspöck[7], Seweryn Zeman[8]

[1]University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia.
[2]KFB Acoustics sp. z o.o., Wrocław, Poland.
[3]LAUM, Le Mans Université, Le Mans, France.
[4]Kahle Acoustics, Brussels, Belgium.
[5]KU Leuven, Leuven, Belgium.
[6]HEAD Acoustics GmbH, Herzogenrath, Germany.
[7]Institute for Hearing Technology and Acoustics, RWTH Aachen, Aachen, Germany.
[8]Jazzy Innovations sp. z o.o., Gliwice, Poland.
*marko.horvat@fer.hr

Abstract

The impact of sound and acoustic phenomena on daily life and its overall quality is not sufficiently perceived by the general public. In addition, in practically all fields of acoustics, there is a lack of professionals with the necessary knowledge and training to solve the problems that occur in everyday life. In order to raise general awareness of the importance of acoustics and to disseminate knowledge to both professionals and lay people, several universities and companies have joined forces in a consortium with the goal to facilitate the development and dissemination of contemporary teaching materials that go beyond the usual forms and approaches to knowledge transfer. As a result, the freely accessible platform Acoustic Courseware (ACOUCOU) was launched. Over the years, the consortium has developed different courseware on acoustics for specific user groups within the scope of several projects. The latest, ongoing project is the Acoustics Knowledge Alliance (ASKnow) project, funded by the ERASMUS+ programme. The aim is to develop modern, interactive content that will form a well-rounded package of teaching materials. These materials will be used in five specific areas of acoustics, i.e. in five corresponding courses: Acoustic Fundamentals, Psychoacoustics, Acoustic Simulations and Auralisation, Electroacoustics, and Room and Building Acoustics. This paper provides an overview of the ASKnow project, its tasks and objectives. It describes the nature and structure of the developed materials and discusses the development procedures used in their creation. The challenges encountered along the way are briefly outlined. The project has entered the phase of transforming the developed teaching materials into their final form, i.e. the interactive online content. This provides an opportunity to present selected extracts of the raw materials developed during the first half of the project, as well as examples of the finished prototypes of interactive online lessons.

Keywords: online learning platform, interactive teaching materials, freely accessible.

1 Introduction

The quantity and diversity of the overall human knowledge is greater than ever before, and it grows faster and faster. While the fundamental postulates in every discipline and field remain the same, their application

branches out at an impressive rate. As a result, good command of the fundamental knowledge is necessary and is still highly valued, but the ability to improvise, adapt and overcome by quickly adopting and applying new knowledge has become just as important, thus leading to a constant and life-long learning process. In any stage of education, e-learning can be a valuable tool. Although the benefits of e-learning have been studied and known for some time [1, 2, 3], a quantum leap in modernization and diversification of the methods of knowledge transfer has been inadvertently induced by the outbreak of COVID-19, as the teaching staff had to turn to distance learning. As a result, all kinds of online teaching materials including open courseware platforms have become an invaluable resource in the teaching process. In the field of acoustics, there seems to be a serious deficiency of highly skilled and knowledgeable specialists who would have the expertise and the ability to tackle the problems that come along in everyday life. To make a contribution to solving these issues, a consortium has been formed by eight institutions the authors belong to, and various associated partners. This consortium has set the goal to devise and develop a tool that would provide the necessary means for facilitating a different, contemporary approach to knowledge transfer in acoustics. As a result, the Acoustic Courseware (ACOUCOU) platform [4] has been established. At present, several projects have already been successfully implemented under the wing of the ACOUCOU platform, namely, the Architectural Acoustics Multibook (ArAc Multibook), the Acoustic Course for Engineers (ACE) and the Acoustic Course for Industry (ACI). Additional information on this courseware and the platform itself can be found in [4] and [5].
This paper presents the Acoustics Knowledge Alliance (ASKNOW) project [6] as the ongoing undertaking of the consortium, funded through the ERASMUS+ programme. The project strives to cover five topics in acoustics by developing online, interactive teaching materials for the corresponding courses, namely, the Acoustic Fundamentals, Psychoacoustics, Acoustic Simulations and Auralisation, Electroacoustics, and Room and Building Acoustics. The content developed for these courses will be briefly illustrated. A description is given of the methods of developing the teaching materials, and the work strategies implemented by different project partners and working groups are presented. The difficulties and challenges that have arisen mostly due to the outbreak of COVID-19 are discussed. The structure of the lessons as the building blocks of the five courses will be explained in detail. At present, the developed teaching materials are being converted into their final form, so that they be ready for the testing phase. Therefore, selected samples of raw materials will be shown, but also the examples of the finished prototypes of ready-to-be-used online lessons.

2 What is ASKnow?

2.1. The ASKnow consortium

To achieve the goals of the ASKnow project, a consortium has been formed that comprises eight partners. Four of them are academic institutions (universities), namely, the Le Mans Université (France) as project leader, and KU Leuven (Belgium), RWTH Aachen (Germany) and University of Zagreb (Croatia) as academic partners. The remaining four partners are private companies, i.e. small- and medium-sized enterprises (SME), namely, HEAD acoustics (Germany), Kahle Acoustics (Belgium), and KFB Acoustics and Jazzy Innovations (Poland). All academic partners have well-established research and teaching groups that cover different aspects of the five course topics, thus complementing their knowledge and expertise. Three of the four SME partners conduct their core activities in the field of acoustics as well, while Jazzy Innovations is specialized in developing both web and mobile applications.
To complement the knowledge and expertise of the eight partners, a total of 14 associated partners have been asked to take part and support the efforts made in the implementation of this project by taking over the development of very specific parts of course materials, but also by getting involved in different kinds of promotional activities. The associated partners are of very different backgrounds, as follows: the Slovak University of Technology (STUBA), Slovakia, and the Graduate School of Engineering (EPF), France as academic institutions; e.GO Mobile AG (Germany), Harmonia Acústica (Brazil), and R&D Team (Germany) as small and medium enterprises; DOLBY Poland, I.G. Bauerhin GMbH (Germany), Miele & Cie (Germany), PSA Group (France),

STIHL AG (Germany), SOMFY (France), and ZIEHL-ABEGG SE (Germany); the European Acoustic Association (EAA), Spain, and the Audio Engineering Society (AES), USA as professional non-profit associations.

2.2. The goals and outcomes

The main outcome of the ASKnow project is to develop interactive online material for five courses, each of them covering a specific topic of interest: acoustic fundamentals, psychoacoustics, acoustic simulations and auralization, electroacoustics, and room and building acoustics. The material is to be made available on the Acoucou platform. In this manner the ASKnow consortium strives to fulfil the main goal of the project, which is to present the knowledge on the selected fields of acoustics in a modern, interactive way that will facilitate efficient learning and to make the developed materials available to specific user groups. The targeted users are first and foremost the students and teachers in universities and other kinds of higher education institutions, but also professionals (engineers, architects, etc.), as well as corporate researchers. The goal is to help them broaden and extend their set of skills and to raise awareness of the importance acoustics has in many different disciplines and aspects of everyday life. The topics of interest that have been chosen by the consortium complement the knowledge collected within the already developed courses that are available on the Acoucou platform and reflect the fields of interest covered by the members of the consortium, taking into account the needs expressed by potential end users.

2.3. The structure of the project

The project consists of thirteen work packages, defined, ordered and numbered according to their role in the project. The first work package (WP1) deals with the development of guidelines and templates to be used in the development of raw teaching materials. The second one (WP2) is focused on project management. The next five (WP3 to WP7) are dedicated to the development of teaching materials. One work package is defined for each course. The following two (WP8 and WP9) are responsible for compiling all the materials, making them uniform, and transforming them to their final form. The next three work packages (WP10 to WP12) deal with quality control through testing and adjustments of the developed materials, quality assurance and evaluation. The final work package (WP13) tackles the issues of dissemination and exploitation of the developed materials.

3 Implementation

3.1. The structure of the developed materials

Each of the five courses consists of thirty lessons and two practical cases. Lessons are the fundamental particles of each course and each one is dedicated to a specific topic. Depending on the topic at hand, the length and the level of complexity change from one lesson to the next, but all of them consist of three parts.
Part A is the theory section, i.e. text complemented with equations and illustrations. Its purpose is to explain the underlying theoretical concept of the phenomenon presented in the lesson. The form of the material is adapted to online use. Part B is the principle section that strives to illustrate the principles established in the theory part using various types of presentation tools. They include, but are not limited to videos, sound samples, charts, illustrations, dynamic interactive calculations in form of applets, etc. The participation of the user is essential, whether it requires browsing through slides, watching videos, listening to sound samples, changing the parameters of a model of a device or a system and observing the changes in the response, etc. Part C is the task section, and, as such, requires the highest degree of interactivity. The user is encouraged to solve tasks in different forms, such as answering multiple-choice questions, ranking or grouping items, performing calculations, etc. The given solutions are validated, and hints towards the correct solutions can be provided if a wrong solution is initially given.
The final, web-based form of each lesson is designed so that the theory section is always visible on screen and is, therefore, always at the disposal of the users. The principle and the task section are displayed on the remainder

of the screen, but not simultaneously. Instead, they are displayed interchangeably, i.e. a switch button can be employed to change from the principle to the task section and vice versa. The designed format is shown in Figure 1.

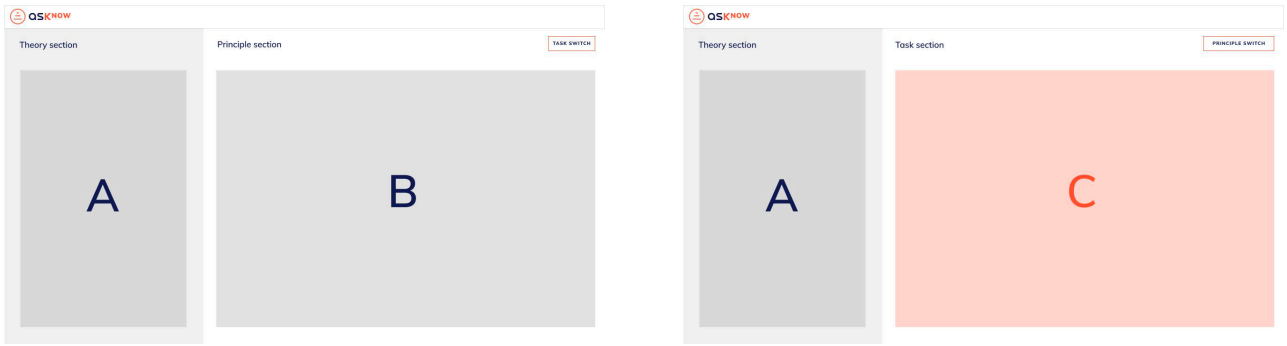


Figure 1: The format of the lessons: the theory part (A) is displayed on the left and is always visible; the principle part (B) and the task part (C) are displayed interchangeably on the right

3.2. Work strategy

Based on experience gained on previous projects, the consortium decided to establish a unique set of tools and templates that would facilitate fast and efficient conversion of developed raw material to the final product in .html format. The theory sections of all lessons have been written and edited in Latex format. To facilitate efficient cooperation in producing these materials, the cloud-based LaTeX editor Overleaf was used. To avoid free-from documents, Jazzy Innovations had designed a .tex template and a conversion routine that automatically transforms the .tex documents to .html format. Given the specific nature and appearance of the principle and task sections, Figma was chosen as a web-based or desktop vector graphics editor and prototyping tool. A set of Figma templates was designed for the same reason as above. When necessary, new templates were created on request to accommodate the diverse range of designs used for both the principle and the task sections. To maintain the visual uniformity of design, the consortium also established a unique graphic design charter upon consultations with graphic designers, as shown in Figure 2.

To extend the means of communication both between and within project groups beyond traditional e-mail messages, a free instant messaging software called Mattermost has been used. Within it, a number of dedicated communication channels has been opened that cover specific topics: organization of meetings, dissemination, connection with the funding agency, the content and form of the course materials, etc. The members of the project staff can join any channels they like. The tool also facilitates private messages sent back and forth between individual members of the project staff. This way, the small issues are resolved directly, and the communication through channels remains as clear as possible. For sensitive and official communication, a mail alias that includes all the project staff has been created by the project coordinator.

To facilitate ”face-to-face” communication while adapting it to pandemic conditions, online meetings were organized both at the level of individual work packages and of the entire project. The meetings were held as required by the schedule of the project and as desired by individual working groups.

The overall progress of the project has been monitored using an online spreadsheet as a progress chart, that allowed the project staff to provide information on the readiness of each part (A, B, C) for each of the thirty lessons in all five courses. Besides this, the spreadsheet has also been used for monitoring the reviewing and the final development process. It also contains the list of meetings, the list of contacts, the data on workload for each partner, etc.

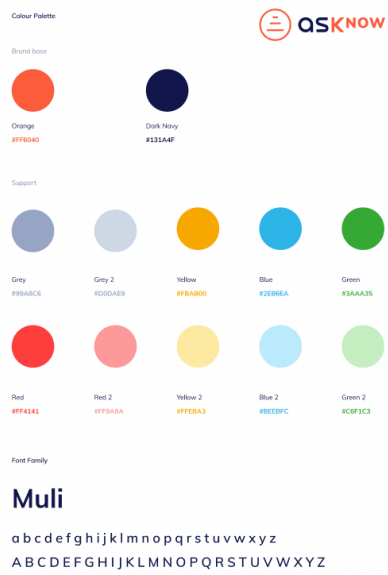


Figure 2: The graphic design charter adopted for the ASKnow project

Each work package dedicated to the development of the teaching materials has chosen and employed the work strategy that promised the best results, depending primarily on the knowledge base of the involved partners. The project partners have chosen the work packages they wish to participate in early on, i.e. during the preparation of the project proposal. The field(s) of expertise of individual staff members was the predominant factor in this work distribution process. The distribution of responsibility was very different from one work package to the next. A centralized responsibility scheme was implemented in certain cases, with one partner being the principal contributor and the developer of an entire course, and other partners being minor contributors, consultants and reviewers. A distributed responsibility scheme had the individual lessons evenly distributed between partners as main contributors. Other partners would review only the content not authored by them and suggest improvements and changes. A highly distributed responsibility scheme had more than one partner working on the development of individual lessons, and strong cooperation between the partners was required and encouraged. All the partners made an effort to unofficially review the developed content.

3.3. Challenges in times of the pandemic

The principal challenge and risk to the implementation of the project according to the planned schedule is the outbreak of the COVID-19 virus and the pandemic situation that has been declared in response. The initial lockdown had stopped all travel, thus preventing any and all in-person meetings. Additionally, the academic partners had no choice but to make a swift switch to online teaching. The effort required to convert all the teaching materials they usually use into an online form that would be useful to students was considerable. Consequently, the majority of project activities had to be postponed, especially the ones tied to the development of teaching materials. Most of the project staff have been stricken by the disease at one time or another, which caused short, but significant work-related interruptions, thus causing additional delays. However, the communication network established by the consortium, as well as the use of online tools and templates allowed the development to continue while maintaining the quality of the resulting material. To cover and make up for the delays caused by *force majeure*, the consortium requested a six-month extension of the project at the end of the first reporting period. The request was granted by the funding agency.

3.4. Quality control

The process of controlling the quality of the developed materials has a crucial role in the ASKnow project, as the ultimate goal is to develop high-quality content. The process is ongoing, with some phases already finished, and others still to come. All these steps are necessary if the developed material is to be easily understood, have the appropriate level of complexity, and be free of errors.
The first phase of the reviewing process was played out within the project teams assigned to each of the five courses. The developed raw material was made available to all team members (but also to all project staff), who reviewed it and suggested improvements and changes, where necessary. Additionally, one or two experts in each team were asked to review the content of the entire course.
The second phase of the reviewing process was carried out by KFB on all the developed materials, with the intent to detect and correct any errors that may have remained, but also to investigate whether the material is suitable for conversion to its final form.
The third and final phase of the reviewing process is ongoing and involves external review by specialists outside the consortium. Reviewers from universities in Spain, Sweden and Poland have already been recruited.
Once all the material has been converted to its final form, a dedicated work package will implement the testing phase on target groups (students and lifelong learners), with the goal of collecting their feedback. The material will then be adjusted according to that feedback.

4 The developed courses

4.1. Acoustic fundamentals

The course is aimed primarily at students and professionals at the bachelor level. It leans more towards theory than other courses, and the main goal is to provide an overview of basic terms and phenomena in acoustics. The basics of sound and waves in one dimension (1D) are described, and the expansion to three dimensions (3D) is made, as well as to bounded spaces such as waveguides and cavities. Basic features of acoustic waves are covered: the reflection/transmission on a discontinuity of the tube section and/or propagating media, impedance, the notion of modes, and the acoustic intensity. Fully interactive tools designed with React enable the visualization of each concept so that the students understand it better. The solutions of the equations in 3D in Cartesian (plane wave) or spherical (spherical wave) coordinates are given. The Snell-Descartes' laws are introduced, and the calculations leading to the mathematical expression of acoustic fields radiated by or diffracted on surfaces are proposed. Acoustics in bounded media such as waveguides and cavities is presented for users with intermediate to advanced knowledge in acoustics. The theoretical parts of these lessons are also supported by interactive illustrations to help the users understand these advanced concepts. The final part of the course focuses on specific concepts of importance for any acoustician and leans towards application. Two case studies have been developed, focusing on characterization of acoustic materials and transparency measurements using acoustic intensimetry.

4.2. Psychoacoustics

The course is aimed at students and professionals who seek better understanding of the human auditory perception. The anatomy of the hearing system and the signal processing in the brain are linked with the way humans perceive sound. A detailed description of psychoacoustic parameters is given. The binaural dimension of our hearing system is extensively illustrated and explained in multiple lessons, revealing the importance of sound localisation and binaural unmasking in our everyday life. The concept of speech intelligibility is introduced. As this phenomenon is influenced by many factors, including room acoustics and noise, the dedicated lessons in this course are interlaced with the selected lessons in the Room and building acoustics course. The final lessons focus on more diverse topics such as soundscapes, noise perception and noise mapping. These lessons examine practical applications of the learned material and demonstrate the use of psychoacoustics for noise control and

sound design. All lessons are conceived in such a way that the learning curve is adequately designed for beginners. The interactive parts of each lesson allow the users to grasp the presented concepts in an intuitive way. This rather unique approach to teaching acoustics will make the course attractive for people with intermediate and advanced knowledge in acoustics as well.

4.3. Acoustic simulations and auralization

The course first explains the basics of signal processing, concerning the impulse response of systems, convolution, Fourier transform and discrete calculations. The second chapter covers different modelling techniques used for room acoustic simulations: The image source model, ray tracing simulations, radiosity methods and wave-based models. With the reverberation time being the most popular parameter of room acoustics, prediction methods of this parameter are discussed as well. Next, the course presents modelling methods for application in noise control and sound design, namely, in areas of environmental sound propagation, airborne and impact sound in buildings, and binaural transfer path analysis. The next chapter of the course deals with auralization as a way of creating sound samples that reflect the behaviour of a system/room and can be listened to. The signal processing methods utilized for (real-time) auralization are presented as well as the acoustical reproduction systems are described, ranging from headphones and binaural loudspeaker setups, through panning techniques such as stereo and Vector-Base-Amplitude Panning (VBAP), as well as panning techniques using object-based audio, to Ambisonics. The final part of the course displays some applications of acoustic simulations and presents the methods of objective and subjective evaluation of the simulations.

4.4. Electroacoustics

As a start, the basic concepts in electricity, mechanics and acoustics are presented. Linear and non-linear characteristics of transducers are described. The most common transduction principles are explained, namely, the electrodynamic and electrostatic coupling. The mechanical-acoustic coupling by membrane is discussed. Models of sensors (geophones, and microphones) and sound sources (shakers and loudspeakers) are developed, with the emphasis on directional microphones and on the Thiele and Small parameters of the electrodynamic loudspeaker. The behaviour of common loudspeaker systems such as closed box, bass reflex and passive radiator systems is investigated. The filtering of multichannel speakers is introduced as well. Advanced knowledge is then presented regarding the characterization of loudspeaker drivers. Finally, the systems that use a large number of transducers are described: microphone arrays or loudspeaker clusters. Applications such as line array systems, 2D and 3D multichannel audio systems, and sound zones are presented.

4.5. Room and building acoustics

This course is aimed mainly at students and professionals working in the building industry (architecture, M&E, structure, etc.). The course is divided into two parts, one dedicated to room acoustics and the other to building acoustics. The difference between both is explained in the first chapter. The basic principles of room acoustics analysis, design and evaluation are then introduced, followed by practical explanations on the behaviour of sound in closed spaces of different shapes, sizes, and applied acoustic treatment. The concept of reverberation and reverberation time is explained followed by the design of absorbers and diffusers. Design criteria and evaluation parameters, as well as guidelines for achieving good acoustics in a space are presented, directly illustrated by acoustic requirements for the design of spaces specifically dedicated to speech and music. The room acoustics part concludes with a lesson on how acoustics can be integrated into architectural design. The next part brings together all topics related to building acoustics. There are lessons on the concepts of airborne and impact (structure-borne) noise, as well as on direct and lateral transmission paths, on definitions of the transmission coefficient and on the noise reduction index. The specifics of various building elements and their sound insulation properties are also given. This leads to the procedures for measuring and assessing airborne and impact sound insulation. A classification scheme that could be used to assess the quality of sound insulation

is also proposed. Special attention is given to typical building materials, but also to special constructions that contribute to improving sound insulation. The ecological aspect is also addressed in terms of building materials useful for acoustics. Finally, the issue of internal background noise is presented as one of the components of acoustic comfort in a space that is closely related to the design of ventilation and other systems that generate noise. Internal noise, sound insulation and acoustic treatment of a space are (or should be) an integral topic of any building project.

5 Examples of developed material

Figures 3 to 7 show examples of the developed raw material. Figure 8 shows how the appearance of raw material is changed as it is converted into a finished prototype.

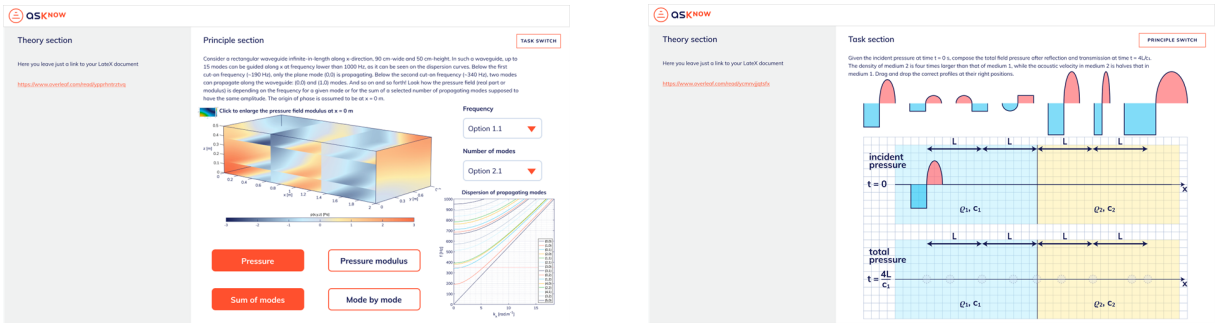


Figure 3: The principle part of the lesson on waveguides (left), and the task part of the lesson on transmission and reflection (right) as examples of material for the Acoustic fundamentals course

Figure 4: Two states in the principle part of the lesson on loudness in the Psychoacoustics course

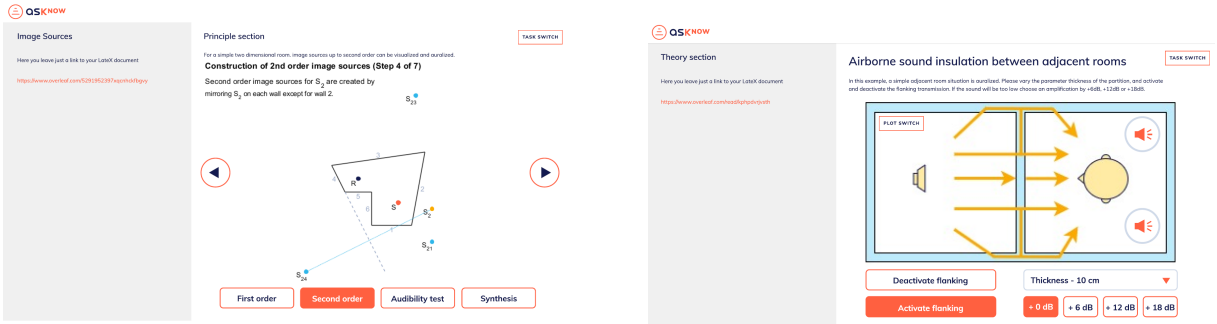


Figure 5: Two examples for principle parts of the course Acoustical simulation and auralization: Image source construction (left) and auralization of sound insulation between rooms (right)

Figure 6: The Electroacoustics course; the principle part of the lesson on electrodynamic loudspeaker (left), the task part of the lesson on transducer limitations (right)

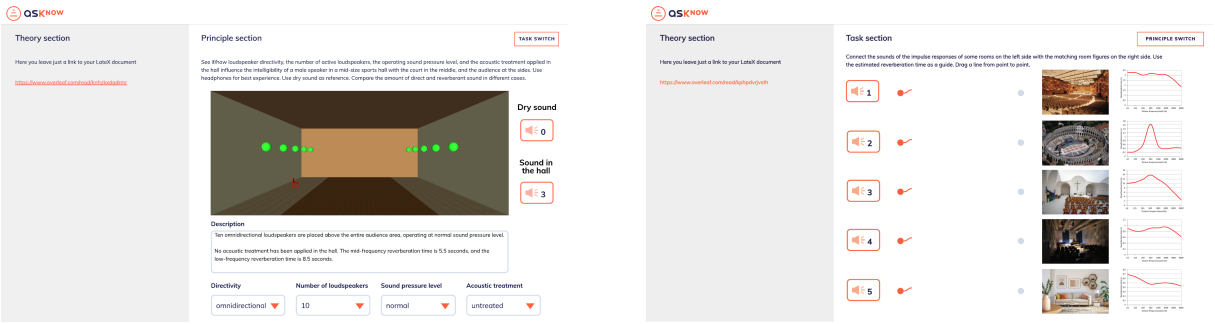


Figure 7: The Room and building acoustics course; the principle part of the lesson on loudspeakers in rooms (left), the task part of the lesson on reverberation time (right)

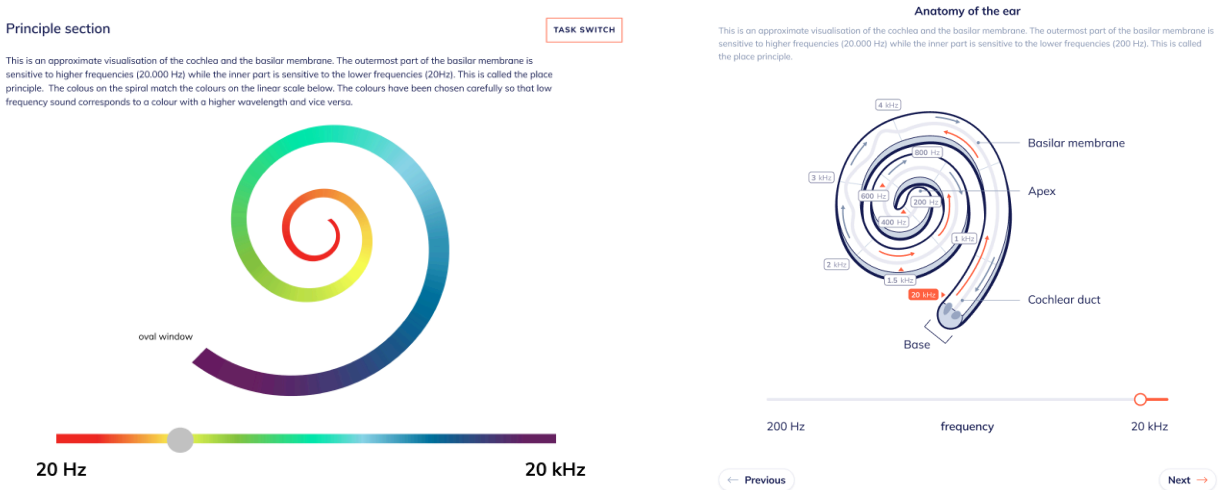


Figure 8: An example of the conversion of the developed material in its raw form (left) to its final form (right)

6 Conclusions

At this time, the ASKnow project has reached and passed its midpoint, as all material is now fully developed and compiled. Work has now been turned over to graphic designers and programmers, whose task is to adjust the visual appearance of the material to the accepted design guidelines and to convert it into its final online form,

respectively. The phase of external review will be completed in parallel. The finished product will be tested on a sample of users, and their feedback will be used to make adjustments, if necessary.

Despite the difficulties caused by the *force majeure* circumstances, the consortium managed to stay on track regarding the activities on the project with only minimal delay, due to good organization and communication.

The finalized material is expected to be ready for use by mid-2023. The consortium will continue to carry out dissemination actions to make the potential users aware of the existence of this material, and to promote its use.

Acknowledgements

All the activities within the Acoustics Knowledge Alliance (ASKNOW) project (project reference: 612425-EPP-1-2019-1-FR-EPPKA2-KA) are funded by the Education, Audiovisual and Culture Executive Agency (EACEA) through the ERASMUS+: Knowledge Alliances programme.

References

[1] P.-S. Chen, A. Lambert and K. Guidry. The impact of Web-based learning technology on college student engagement. *Computers & Education*, 54(4): 1222-1232, 2010.

[2] F. Michau, S. Gentil, and M. Barrault. Expected benefits of web-based learning for engineering education: examples in control engineering. *European journal of engineering education*, 26(2): 151-168, 2001.

[3] F. Manganello, C. Falsetti and T. Leo. Self-Regulated Learning for Web-Enhanced Control Engineering Education. *Educational Technology & Society*, 22: 44-58, 2019.

[4] https://acoucou.org, the homepage of the ACOUCOU project, last accessed on 8 March 2022.

[5] K. Jaruszewska, F. Baranski, M. Piotrowska, M. Melon, O. Dazel, M. Vorländer, L. Aspöck, M. Horvat, K. Jambrošić, M. Rychtáriková, L. Kritly and A. Herweg. ACOUCOU Platform to Acquire Professional Skills and Knowledge in the Field of Acoustics. *Proceedings of the 23rd International Congress on Acoustics*, 4348-4355, DEGA, 2019.

[6] https://asknow.acoucou.org, the homepage of the ASKNOW project, last accessed on 9 March 2022.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

A Two Stage Embedded Genetic Algorithm to Optimise Ceiling Reflections

John O'Keefe

O'Keefe Acoustics, Toronto, Canada.

john@okeefeacoustics.com

Abstract
Recent studies have explored computer aided techniques to optimise individual NURB-based acoustical reflectors. Reconfiguring the shape of an individual reflector with the freedom that only NURB geometry allows. This study, conversely, limits itself to primitive geometries such as triangles, polygons, circles and extruded versions of each, i.e. partial spheres, etc. These primitive shapes are collected in an array as they might be in a theatre or concert hall ceiling. An embedded, two stage Genetic Algorithm (GA) is then employed to optimise their size and orientation to encourage uniform laterally reflected sound. Each of the reflectors in the array is set at a given anchor point, starting out parallel to the receiving surface. To initiate the first GA, the given reflector is oriented to direct sound towards a Target Point on the receiving surface. The first GA then optimises that (individual) reflector to cast as many reflections as possible across the receiving surface, as laterally as possible. This is repeated for each reflector in the ceiling array. Several similar arrays are then created and these form the population for the second GA. An early experiment with this 2-stage GA suggests that it might provide higher and more uniform Lateral Fractions than a human design.

Keywords: Room Acoustics, Genetic Algorithms, Multi-Objective Optimisation.

1 Introduction

The past decade has seen increasing use of Non-uniform Rational B-Spline (NURB) geometry for acoustic reflectors [1,2]. The author has developed methods to optimise NURB-based reflector geometries [3,4] using the Non-dominated Sorting Genetic Algorithm (NSGA-II) [5]. Despite the increased use of NURB-based geometries in architectural and acoustical design, however, many performing arts projects can still not afford the luxury of NURBS. Flat triangular reflectors are easier to build and install compared to the more exotic NURB geometries. This paper will try to address the practicalities of that reality.

2 Reflection Calculations

2.1 Ray Tracing

Reflections are calculated using a modified version of a simple ray tracing routine [6]. The choice of such a simple version of a ray tracer was a deliberate effort to save time on computer optimisation runs that can last several days. The routine assumes, of course, specular reflection and that the effects of diffusion and diffraction are ignored. This was thought a safe approximation as the object of this exercise is only to optimise

the arrival of first order reflections. Diffusion and diffraction are phenomena that occur after the arrival of a first order reflection. A more complete study, using a model calibrated to an existing room and employing both diffusion and diffraction, is currently being undertaken [7].

Reflections in the computer model used in the current study can be cast from either the source location or from the reflector itself. In this experiment, the reflector size was kept constant. Only the angles of inclination for the various reflectors were modified. Thus, it was computationally more efficient to cast reflections from the reflector(s), as opposed to the traditional method of ray bundles emanating from the source location. In the reflector-based method, a number of reflection points are created on the reflector, with a density typically in the range of 10. For each one of these points, an incident line is created between the source and the reflection point. A reflection line is then cast from the reflection point, based on the angle of the incident line.

For each reflection that successfully intersects the receiver surface, its position, time delay, relative amplitude and lateralness is recorded. Lateralness is calculated using a method developed by Protheroe and Day [8] that estimates the early Lateral Fraction (LF) based on a single reflection. Ref. [8] found a reasonable correlation between free-field LFs calculated with this method and measurements performed according to ISO 3382 [9]. We refer to this parameter as the Single Lateral Fraction or sLF.

$$\mathrm{sLF} = \frac{\left(p_{reflection} Sin\phi\right)^2}{p_{direct}^2 + p_{reflection}^2} \tag{1}$$

where: p_{direct} is the direct sound pressure
$p_{reflection}$ is the sound pressure from a single reflection
ϕ is the angle between p_{direct} and $p_{reflection}$

It should be noted that for extreme angles of incidence (ϕ), the sLF will predict Lateral Fractions of 0.5 or higher, which are surely fallacious. In more reasonable ranges, say from 0.0 to 0.4, the sLF has, however, proved to be reliable.

2.2 Initial Genome.

In this study, we shall be using the language of Genetic Algorithms. An array of reflectors such as that shown in Figure 1 will be referred to as the Genome. The individual reflectors that make up the Genome are the Genes. The algorithm is executed in three stages, as described below.

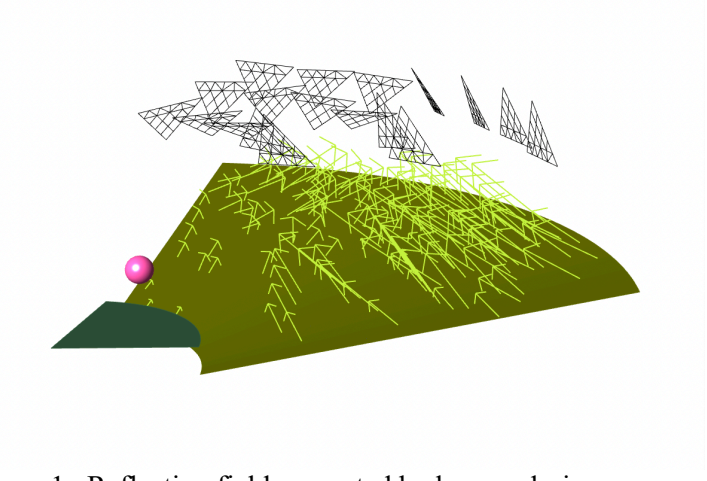


Figure 1. Reflection field generated by human design.

In the first stage of the routine, an array of Anchor Points is established in the vicinity of a receiving surface. In the example we're using here (see Figure 1) the array will form a ceiling above the seating area, but it could be any surface in the room directing reflections to any number of receiving surfaces. A reflector is built at each of the Anchor Points. The options for the six so-called "Primitive Component" shapes that the reflectors might take are shown in Figure 2. The centroid of each reflector (or Gene) is attached to its respective Anchor Point. The size of the 2-Dimensional reflectors is controlled by their edge lengths or, in the case of the circle by its radius. The 3-Dimensional shapes are simply extrusions of the polygons and the circles. Their controlling parameters are the depth of the extrusion or, in the case of the truncated pyramid (middle of bottom row, Figure 2) the depth of the extrusion and the height of the truncation. In all cases, the reflectors can be rotated about their centroids in angles of azimuth and elevation. All of these parameters: edge lengths, extrusion depth, truncation height and angles of inclination can be adjusted, or perturbed, during the Genetic Algorithm optimisation.

Once the array of reflector genes has been created at the Anchor Points, each reflector's centroid point is projected onto the receiver surface. So, in the example we're using – a ceiling of reflectors above a seating area – there will be N_{gene} points projected onto the receiving surface for a ceiling with N_{gene} reflector genes. These are referred to as the Target Points. Each reflector is adjusted to direct a reflection from the (single) source location, to the reflector's centroid and then to one of the (multiple) Target Points. This results in a set of N^2_{gene} Reflector/Target Point combinations and the sLF value is calculated for each one. The set of N^2_{gene} combinations is then sorted according to the sLF values and the combination with the highest sLF becomes the initiating genome.

2.3 Genome Population – 1st Stage

A population of N_{genome} reflector genomes is created by perturbing the individual reflector genes created in the initial genome, as described in Section 2.2. This is the first stage of the embedded Genetic Algorithm. Each of the individual reflector genes is optimised on its own and without regard to what the other reflector genes, that will eventually make up the genome, might do. In this and the following Stage 2, optimisations are performed using the Non-dominated Sorting Genetic Algorithm (NSGA-II) [5] and the SBX cross-over [10]. The number of evolutionary optimisation generations will be referred to as $NumGen_1$ and $NumGen_2$ for Stages 1 and 2, respectively.

Although both reflector gene size and angle of inclination can be perturbed during the optimisations, only the latter was performed in this study. Two methods have been employed to change the inclination of the reflector: Target Point Perturbation and Reflector Normal Perturbation.

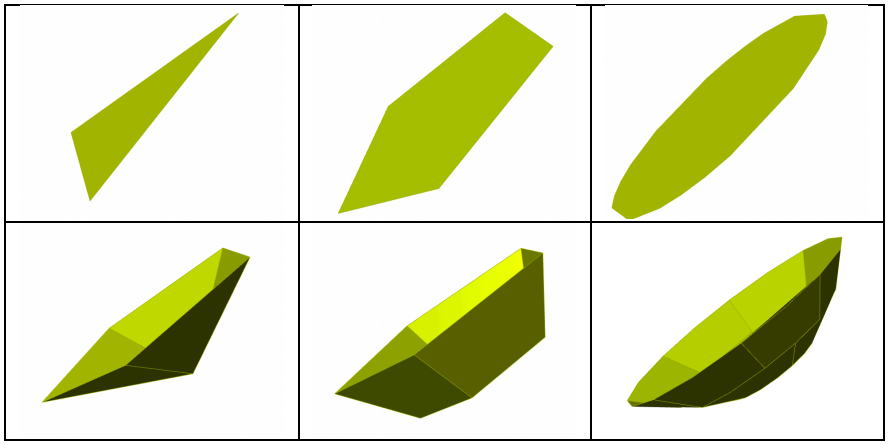


Figure 2. Primitive Component shapes available in the routine. In the top row are the 2-D options of triangles, polygons and circles. In the bottom row are 3-D extrusions of the polygons and the circles.

2.3.1 Target Point Perturbation

Each reflector gene has an original Target Point associated with it. It comes from the initial genome, described in Section 2.2. This point will exist on the receiver surface at the parameterised coordinates (U,V). The U,V coordinates of the Target Point are normalised with respect to the U-V span of the receiving surface. A random number generator then perturbs the normalised U,V parameters, typically over a range of U ± 0.2 and V ± 0.2. The receiver surface is then evaluated at the new U,V parameters to generate the new Target Point coordinates. A new normal for the reflector is then calculated based on the source location, the reflector's centroid and the new perturbed Target Point. The inclination of the reflector is then adjusted according to the new normal.

2.3.2 Reflector Normal Perturbation

Each reflector gene also has an original Reflector Normal derived, again, from the initial genome. The reflector's normal vector is normalized and, consequently, its X,Y and Z coordinates range from 0 to 1.0. A random number, typically ranging between ± 0.2, is then added to each of the vector's XYZ coordinates to create the new perturbed Reflector Normal. The inclination of the reflector is then adjusted according to the new normal.

2.4 Genome Population – 2nd Stage

After an individual reflector gene has been optimised through $NumGen_1$ generations (as described in Section 2.3) it is inserted into the (reflector array) genome. Then the process is repeated until we reach a population of N_{genome} reflector arrays. This population is then optimised using the NSGA-II algorithm for $NumGen_2$ generations.

To summarise, and perhaps clarify, each reflector array (or Genome) will contain N_{gene} reflectors (or Genes). The first stage of the routine optimises each individual reflector for $NumGen_1$ generations then contributes that reflector to the overall reflector array. A population of N_{genome} reflector arrays is created this way and then the second stage of the NSGA algorithm optimises that population of arrays for $NumGen_2$ generations.

3 Fitness Function

Perhaps the most important component of a computer aided optimisation exercise is the Objective Function or, as it called in evolutionary genetic algorithms, the Fitness Function. A Fitness Function for the optimisation of acoustical reflectors has been developed and is described in [3]. A further explanation and extension of this proposed Fitness Function will be presented in [11]. A brief summary is presented here.

The perception of sound in a performing arts venue is a multi-dimensional experience. Any optimisation of the reflected sound in a room must provide the appropriate Reverberance, Clarity, Spatial Impression, Loudness, Warmth and Intimacy, to name a few. Evolutionary algorithms, such as NSGA-II, however, work best with only two or perhaps three optimisation dimensions. There are now some so-called "Many-Objective Evolutionary Algorithms" (MOEA) that have been developed, e.g. NSGA-III [12] but this discussion limits itself to only two objectives. These two objectives, however, address multiple psycho-acoustic dimensions.

3.1 dMean Fitness

The first objective is referred to as the dMean Fitness and its goal is to distribute the reflections as evenly as possible on the receiver surface. To calculate the dMean Fitness, a grid of points is created on top of the receiving surface, the nodes of which are separated by a distance (dMean), derived from the density of received reflections on any given iteration of the optimisation.

$$d_{mean} = \sqrt{\frac{Area\ of\ Receiving\ Surfaces}{\#\ of\ Receiving\ Points}} \quad (2)$$

The dMean Fitness Function then minimises the distance between any given reflection intersection point on the receiver surface and its Nearest Neighbour (NN) grid point.

The goal of this part of the Fitness Function should not be confused with a desire to produce a more "diffuse" sound. Rather, the intention is to prevent the optimisation from concentrating the reflections in one corner or along one edge of the receiving surface.

3.2 sLF Fitness

The second part of the Fitness Function encourages the reflections to arrive from the side before 80 ms. Again, this should not be confused with a desire only to improve the Lateral Fraction and, hence, Spatial Impression. Reflections arriving before 80 ms will also improve Clarity (C80) and early Strength (G80). The sLF Fitness function is designed to minimise the difference between a target Lateral Fraction (sLF_{goal}) and the sLF calculated for each reflection intersection of the receiver surface. sLF values are calculated according to Equation (1). For this study, sLF_{goal} was set at 0.50. This apparently unrealistic number was chosen to avoid sLF Fitness levels too close to zero. In effect, it increases the dynamic range of the sLF Fitness Function as the routine converges towards some of the best solutions.

4 Results – Man vs. Machine

In one of the experiments to test the 2-stage optimisation algorithm described here, a competition was run between a human design of a ceiling reflector array (the author's) and the computer's optimisation efforts. The reflection field resulting from the human design is shown in Figure 1. The yellow arrows indicate the reflections that have intersected the receiver surface. The base of the arrow indicates where the reflection arrived, the direction indicates where it came from and the length of arrow indicates how lateral it is with respect to the source location (i.e. its sLF value).

The comparison between the human and the computer is seen in Figure 3. This is a Pareto Analysis plot of the two part Fitness Function: dMean Fitness vs. sLF Fitness. In the Fitness Function regime that has been adopted here, the goal of the Pareto Analysis is to minimise both the dMean and sLF Fitnesses, which means the better solutions are the ones closer to the origin.

The star at position (0.27, 0.28) indicates the human's performance. The sLF values for the human's design suggest a space averaged Lateral Fraction (LF) of 0.31. The square on top of the star indicates the performance of the computer's initial genome, i.e. before the optimisation has started (see Section 2.2). It appears to have a dMean Fitness and sLF Fitness similar to the human's. But the space averaged sLF value (which is not the same as sLF Fitness) is suggesting a LF of 0.24. The Just Noticeable Difference (JND) for the Lateral Fraction is 0.05 [13]. This suggests that the human's performance is better than the computer's first attempt by slightly more than a single JND. However, after optimisation (N_{genome} = 25, $NumGen_1$ = 25, $NumGen_2$ = 25) the 2-stage genetic algorithm has matched the human's space averaged sLF of 0.31 but, perhaps more importantly, it has provided a better distribution of the sLF values. That is to say, the computer's solutions are closer to the origin on the (horizontal) sLF axis which means that more of the received reflections have values closer to sLF_{goal}.

To explain, the sLF Fitness axis quantifies the uniformity of the sLF values on the receiver surface, with respect to the target value of sLF_{goal}. At its maximum, we might have sLF values very close to sLF_{goal} in only one corner of the receiver surface and very low values everywhere else. At its minimum, i.e. as the values approach the origin, we will have a more uniform distribution of sLF values, i.e. more of them will be closer to sLF_{goal}.

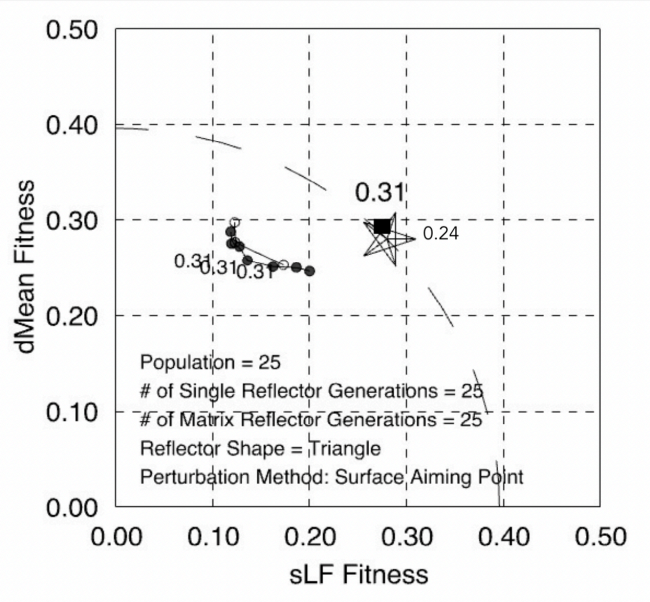


Figure 3. Pareto analysis of the "human" reflector design, compared to the computer's optimisation.

On this run, the computer has only delivered a slightly better spatial distribution of the reflections on the receiving surface. That is, compared to the human's design (the star in Figure 3), some solutions are slightly closer to origin on the (vertical) dMean Fitness axis but others are not. This is an atypical comparison between the human and the computer. In most cases, the human's distribution is worse than what is shown in Figure 3. One might say that the comparison was biased in the human's favour. In either case, the results suggest that the computer can improve on the human's performance, even when the human is at his or her best.

5 Conclusion

A two-stage embedded genetic algorithm to optimise acoustic reflector arrays has been presented. The first stage of the algorithm optimises the individual reflectors before they are allowed into a composite reflector array. The second stage then optimises the composite array itself. The routine limits itself to traditional (i.e. non-NURB) geometries such as triangles, pyramids and partial spheres. In this, the early stages of the algorithm's development, a trial run has shown that it can produce acoustical solutions equal to or slightly better than the human design.

References

[1] Peter Exton, Harold Marshall, The Room Acoustic Design of the Guangzhou Opera House, *Proc. Inst. of Acoustics,* (33) 2:117-124, 2011.

[2] Thomas Scelo, J. Valentine, Harold Marshall, Chris Day, Implementing the Acoustical Concept tor The Philharmonie De Paris, Grande Salle, *Proc. of Inst. of Acoustics*, 37(3), Pt. 3: 118-127, 2015.

[3] John O'Keefe, Geometric Algorithms for Machine Based Optimisation of Acoustic Reflectors, *Proc. of International Conference on Immersive and 3D Audio,* 2021.

[4] John O'Keefe, Applications of Machine Learning Bounding-Breps for Optimised Acoustical Reflectors, *Proc. of Euronoise*, 2021

[5] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multi-objective genetic algorithm: NSGA-II *IEEE Transactions on Evolutionary Computation.* 6 (2), p.182A, 2002.

[6] Krokstad, S Strom, S. Sorsdal, Calculating the acoustical room response by the use of a ray tracing technique, *J. Sound Vib.,* 8 (l), 118-125, 1968.

[7] John O'Keefe, Henrik Moller, Applying Computer Aided Optimization to a Problematic Fan-shaped Auditorium, *Proc. ICA-2022*(to be published)

[8] D. Protheroe, C. Day, Validation of lateral fraction results in room acoustic measurements, *InterNoise Melbourne*, 2014.

[9] ISO 3382-1, 2009. Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces, *International Organization for Standardization*, Geneva.

[10] K. Deb, R.B. Agrawal, Simulated Binary Crossover for Continuous Search Space, *Complex Systems* Vol. 9, pp.115-148. 1995.

[11] John O'Keefe, A Genetic Algorithm Fitness Function for Acoustic Reflectors in Performing Arts Venues, *Proc. ICA-2022*(to be published).

[12] K. Deb, An Evolutionary Many-Objective Optimization Algorithm Using Reference-point Based Non-dominated Sorting Approach, Part I: Solving Problems with Box Constraints, *Proc. of IEEE*, 2013.

[13] T.J. Cox, W.J. Davies, and Y.W. Lam, The sensitivity of listeners to early sound field changes in auditoria, *Acustica* (79), pp. 27-41, 1993.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Orchestra Conductor Acoustics

Magne Skålevik[1,2]

[1]AKUTEK, www.akutek.info , Spikkestad, Norway

[2]Brekke & Strand, Oslo Norway

While the playing and listening conditions for musicians have been studied and taken into consideration in research and design practice since Gade's early work in Podium Acoustics in the 1980s, the impact of room acoustics on the conductors working conditions are only rarely addressed in scientific papers. In preparation of a concert performance, conductors typically need to hear in considerable detail what is going on in the orchestra, perhaps even more during rehearsal than during the actual concert. Importantly, the conductor needs to prepare the orchestra sound and balance to be presented to the audience. In this paper, sound from a symphony orchestra in various rehearsal rooms and a typical concert hall is simulated. Features of orchestra sound at the ears of the conductor's ears and at the ears of the audience is studied. The impact of varying acoustics and the difference between sound at conductor's ears and sound audience ears are commented. In particular, the significance of compliance with ISO-23591:2021 is studied. A questionnaire suggested by this author for a survey among conductors is presented and discussed, the purpose being to provide a basis for finding acoustic measures that can describe critical aspects from the conductors' perspective.

Keywords: orchestra acoustics, orchestra conductor, rehearsal room acoustics, room acoustics

1 Introduction

While the playing and listening conditions for musicians have been studied and taken into consideration in research and design practice since Gade's early work in Podium Acoustics in the 1980s, the impact of room acoustics on the conductors working conditions are only rarely addressed in scientific papers. A noteworthy exception is the work of Jürgen Meyer[1].

In preparation of a concert performance, conductors typically need to hear in considerable detail what is going on in the orchestra, perhaps even more during rehearsal than during the actual concert. Unfortunately, due to the shorter critical distance in most rehearsal rooms, details in the orchestra are more audible for the conductor from the podium of the concert hall than in the rehearsal room.

Importantly, the conductor needs to prepare the orchestra sound and balance to be presented to the audience.

Most conductors would ideally like the sound of the orchestra at the podium to be as similar as possible to that at the audience' ears.

However, this condition is practically impossible to achieve. Then, several questions arise: How can conductors judge and adjust the orchestra balance, articulation, intonation, dynamics, etc., as perceived in the audience?

More practical, but important in the daily work for conductors – how does varying acoustics in rehearsal spaces influence on ease of speech when giving verbal instructions to musicians.

In general: What are the critical subjective aspects of acoustics, from conductors' perspective, and what are the acoustical parameters that would best describe them?

In this paper, some of these questions will be addressed. Acoustical conditions of rehearsal rooms will be categorized with regards to whether they comply with the combinations of volume and reverberation time recommended by the recent rehearsal room acoustics standard ISO-23591:2021 [2].

2 A conductor's perspective

"A conductor's task is to prepare sound and balance for the audience in the best possible way.

Good conductor acoustics would therefore,
- ideally, provide listening conditions for the conductor like those in the audience
- make balance judgements possible

As to Reverberance, preferences vary, from one piece to another, from one conductor to another."

This quote by conductor Eivind Gullberg Jensen[1], from personal communication with the author, represents the perspective of one conductor, and was the assumption of the work presented in this paper. To establish a more general representation of conductors' perspective, this author intends to conduct an online survey among conductors. A questionnaire is proposed and presented for discussion in a paragraph below.

Lehtimäki (2019) investigated the effect of concert hall acoustics on the work of a conductor by means of a questionnaire, answered by 25 conductors among 69 invited. An important result was reported on page 25 in the thesis: "Balance is the aspect that almost all of the participants reported having been affected by the acoustics of a concert hall in their performances"[3]. Moreover, the conductors express the need to judge orchestra balance by listening from the auditorium, and there is no consensus as to an optimum way of combining this with the obvious need for the orchestra to be conducted. Different from the current study, orchestra rehearsal rooms as such was not an objective in the thesis.

3 Symbols and Definitions

The following definitions apply to some frequently appearing symbols in this paper.
- V Volume in cubic meters (m^3)
- T Reverberation Time in seconds (s)
- V-T Diagram in ISO-23591:2021 for indicating recommended combinations of V and T
- G Sound Strength in dB, as defined in ISO-3382
- G_{refl} Sound strength of reflected sound, also referred to as "Room Response"
- D-R Direct-to-reverberant sound level balance, 10*log(d/r) in dB
- STI Speech Transmission Index, dimensionless (1)

Note: In this paper STI is applied for two different purposes. One is for assessing the ease of speech, i.e. the conditions for the conductor to communicate vocal instructions during rehearsals. Less obvious, perhaps, STI from an instrument group to conductor and to audience is used to assess musical articulation. The latter is based on the understanding that musical notes, in addition to tonal characteristics, has transients and transitions that could be represented by amplitude modulation like the algorithm of STI. Values are only used for comparison between rooms and receiver positions, and of course not directly comparable to recommendations for speech. For information about the sources and directivities used to simulate STI measurements, please refer to section 6 Models.

[1] Chief Conductor of the Noord Nederlands Orkest from 2022/23 season and Artistic and General Director of Bergen National Opera. Jensen has previously conducted the Berliner Philharmoniker, Münchner Philharmoniker, Hamburger Symphoniker and WDR Sinfonieorchester in Germany, Amsterdam's Royal Concertgebouw Orchestra, Vancouver, North Carolina and Oregon symphony orchestras in North America, and further in Europe the Royal Stockholm and Netherlands Radio philharmonic orchestras, Orchestre de Paris and Tonhalle-Orchester Zürich

4 Method

To learn more about how room acoustics supports the conductor's work conditions, as described above, and to address some of the questions in the Introduction, the following procedure was chosen.

- 3D-models of altogether 5 rehearsal rooms of various volume and absorption factors and one typical concert hall, all containing one and the same 90-piece orchestra, was built for acoustical simulations in the software Odeon 17
  - Two of them with V and T inside limits given in ISO-23591:2021
  - Three of them with V and/or T outside the limits of the standard (red numbers in Table 1)
- Sound emission from the symphony orchestra in the various room models was simulated by using the multi-source feature in Odeon
- Selected features of orchestra sound at the ears of the conductor's ears and at the ears of the audience is simulated and analysed, including
  - Orchestra balance, i.e. level profile of instrument groups
  - Articulation of music, in terms of the STI profile of instrument groups
  - Reverberance, in terms of the EDT profile of instrument groups
  - Localization, in terms of the Direct-to-Reverberant ratio profile of instrument groups at conductor's ears
- Global parameters of the unoccupied room models, including T and $G_{reflected}$ were simulated
- The difference between sound features at conductor's ears and those at audience' ears, and it's dependency on V, T and $G_{reflected}$ is identified and commented
- Features in orchestra sound at conductor's ears in various rehearsal room models are compared with
  - Listening quality for the conductor in the concert hall
  - Listening quality in audience
- Ease of speech during verbal instructions to the orchestra, in terms of STI, in various rehearsal rooms are compared relatively and assessed with common criteria
- In particular, results are commented relative to whether or not the various rehearsal rooms comply with the requirements in ISO-23591:2021

5 The rehearsal room acoustics standard, ISO-23591:2021

In the recently published rehearsal room standard ISO-23591:2021, requirements are given for minimum volume per musician, in addition to a range of V-T combinations, i.e. combinations of room volume and reverberation times, applying to symphony orchestra rehearsal rooms. In the current study, rooms complying to the standard as well as rooms not complying to the standard are included, to be able to see what significance such compliance has for the conductor acoustics.

6 Models

The geometrical rehearsal room models had the same floor and orchestra while heights varied, to produce the room volumes V=2700m^3, 5400m^3 and 6500m^3. The two bigger rooms have volumes that comply with the ISO-23591:2021 for a 90-piece orchestra (at least 60m^3 per musician). The smaller one had half the volume required by the ISO-standard. A fourth room was a typical concert hall in shoe-box shape, having a volume of V=20.000m^3, in which the rehearsal room floor forms the stage floor. To the orchestra and conductor, the concert hall auditorium would be like an extension to a big rehearsal room where one of the walls were knocked down. The four geometrical string models are shown in Figure 1 below.

Various degrees of absorption was assigned to walls and ceiling of the models in order to generate rooms with various T, and in particular some inside and some outside of the range of V-T combinatins given by the standard. Altogether 7 room models are generated by this process. Their key properties are given in Table 1, and they are plotted by their V-T combinations into the ISO-23591:2021 V-T diagram in Figure 2.

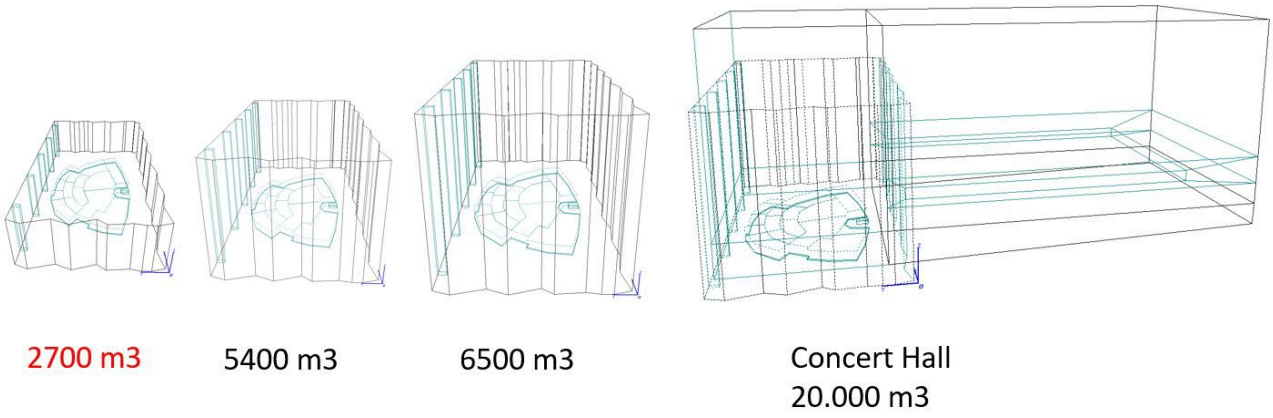


Figure 1 Geometry of the 3D-models in this study

Table 1 Key properties of the models included in the study. Table entries of models not complying with the volume or revrberation time limits in ISO-23591:2021 are in red.

Model #	Volume (m^3)	α walls+ceiling average	T_m empty (s)	G theoretical (dB)
1	2770	0.18	1.93	13
2	2770	0.25	1.43	11
3	5400	0.25	1.87	10
4	6500	0.30	1.63	8
5	5400	0.37	1.23	8
6	6500	0.40	1.20	7
CH	20515	0.24	2.10	5

Sound sources were simulated with the built-in sources in Odeon 17. String groups were simulated with multi-surface sources with spherical radiation characteristics, while other instruments were simulated as point sources. Wind instruments were pointing towards the conductor, with the instrument-specific directivities provided by Odeon, while percussion instruments were simulated by omni-directional point sources.
In this study, the individual sound powers did not have any effect on the results, since all results are either related to levels received by the conductor or in the audience statistics or expressed in terms of strength (G) related to free-field SPL at 10m distance from omni-directional source.
For the assessment of ease of speech for the conductor, the conductors voice was simulated with an omni-directional source, to mimic the always varying head orientation of a conductor, who needs to be understood by the all orchestra members regardless of the orientation of the conductor's head.

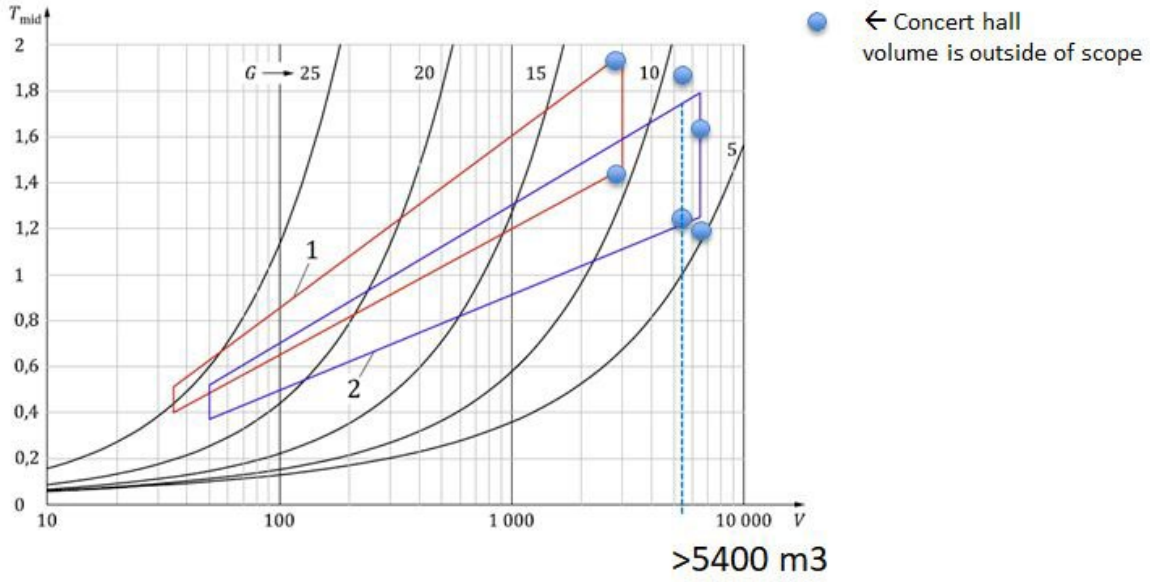


Figure 2 All models in Table 1 plotted in the V-T-diagram of ISO-23591:2021, where the blue box indexed "2" marks the range of V-T combinations applying to rehearsal rooms for Loud Acoustic Music. The vertical dashed line marks the volume limit 5400m^3 implied by the 60m^3 per musician requirement

7 Results and comments

Results from the simulations and analysis are given in diagrams below, together with relevant comments in captions, according to the method presented above. In all the profile diagrams, rooms are ordered by their theoretical Strength, G in dB, decreasing towards the right, and where the rightmost is the concert hall podium. Theoretical values are calculated outside the reverberation radius, thus $G \approx G_{refl}$.

7.1 Orchestra Balance at conductor's ears

From Figure 3, it is concluded that the orchestra balance in the audience cannot be judged and adjusted directly by listening to the orchestra balance at the conductor's ears. For the balance at audience' ears to be perfectly according to the conductor's intention, the string groups 1st & 2nd violins, violas and celli would need to dominate by 3dB over the orchestra average, while brass and percussion would need to sound a bit weaker than the orchestra average. This difference in balance is due to the four mentioned string groups are closer to the conductor than the average instrument, and the brass and percussion are farther away than the average instrument. As a preliminary conclusion it is advised that the conductor needs to find other cues for orchestra balance in the audience than the balance at conductor's own ears.

This observation is very important, and can explain the unison expression from conductors, reported by Lehtimäki [3], that working with orchestra balance is challenging, and that they feel they must listen from the auditorium themselves to confidently judge whether or not the orchestra balance is according to their intention.

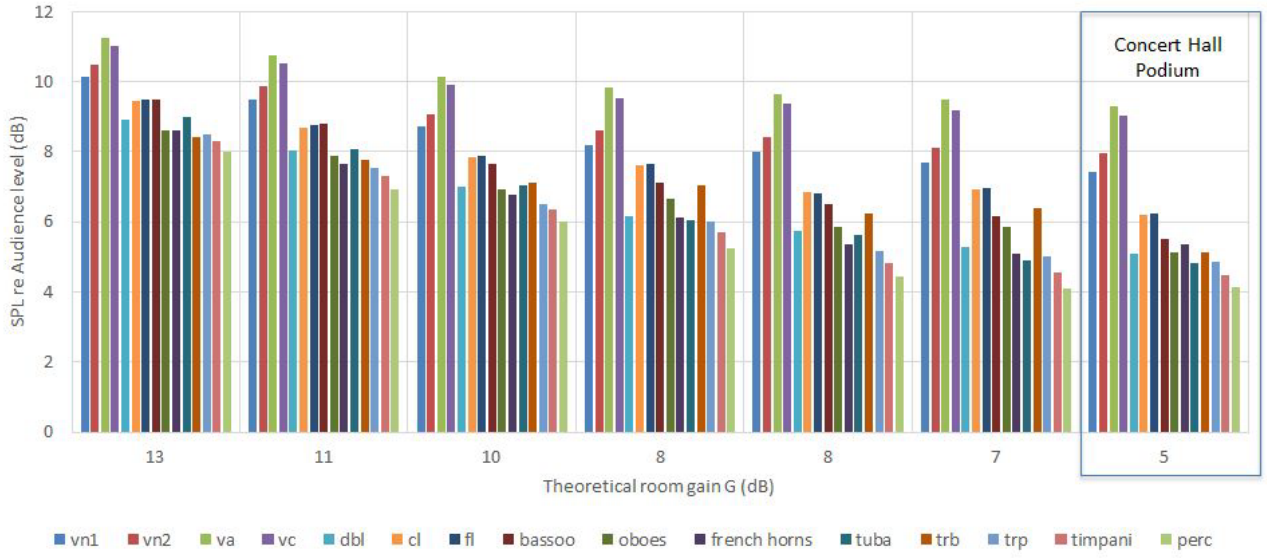


Figure 3 Instrument group level profiles at conductor's ears in the 7 rooms. Vertical axis is level in dB where 0dB is defined as the levels at median (50 percentile) audience ears. Note that on the concert hall podium (room model # 7), strings groups except double basses dominate at conductor's ears by approximately 3dB over ther rest of the orchestra, relative to the balance at audience ears.

In our search for a parameter that for the conductor can serve as a cue of orchestra balance in the audience, Figure 4 reveals that indeed, the room response, i.e. the profile of G_{refl} in dB from the instrument groups to the conductors ears can serve as such a cue. This suggests that in optimisation of rehearsal rooms and podium acoustics for the conductor, one should pay attention to acheiving a flat G_{refl} -profile, with the audience profile defining 0dB like in Figure 4.

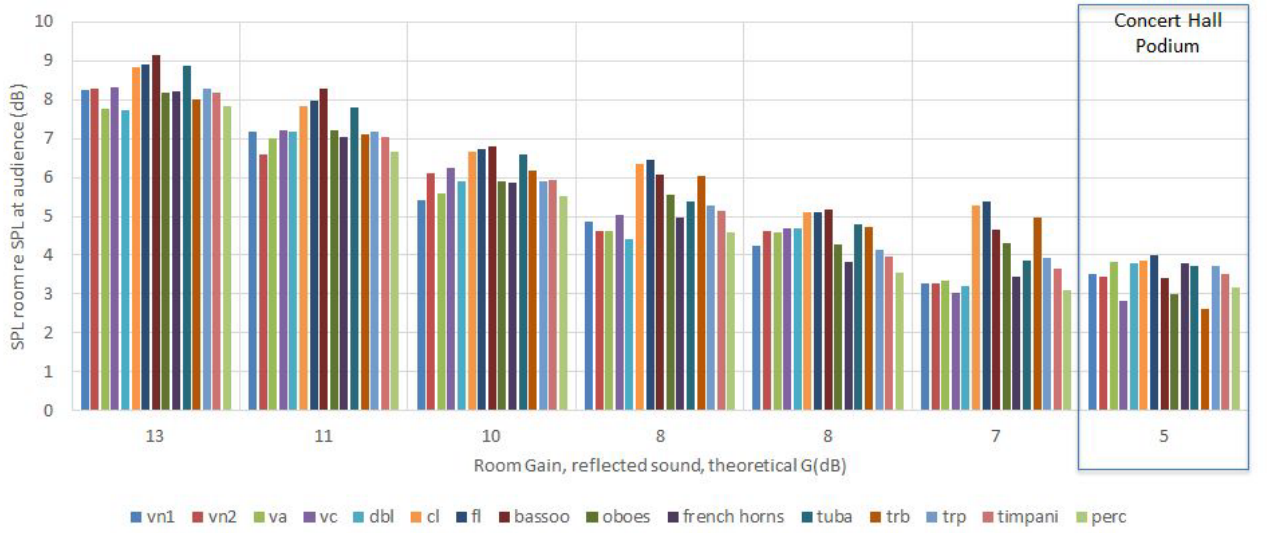


Figure 4 Room response profiles at conductor ears in the 7 rooms. Vertical axis is room response (G_{refl}) in dB where 0dB is defined by the room response heard at median audience ears (50 percentile). Note that on the concert hall podium (room model #7), the profile is effectively flat, i.e. with less than a noticeable difference from the 3dB level above the levels at audience' ears. Moreover it is noteable that the profile in room model #5 is also quite flat, though on average 1.5dB stronger than on the concert hall podium.

7.2 Articulation

Articulation of music, as heard by the conductor, is represented by the STI-profiles in Figure 6. Note that on the concert hall podium, the articulation is from considerably higher to much higher than in the audience.

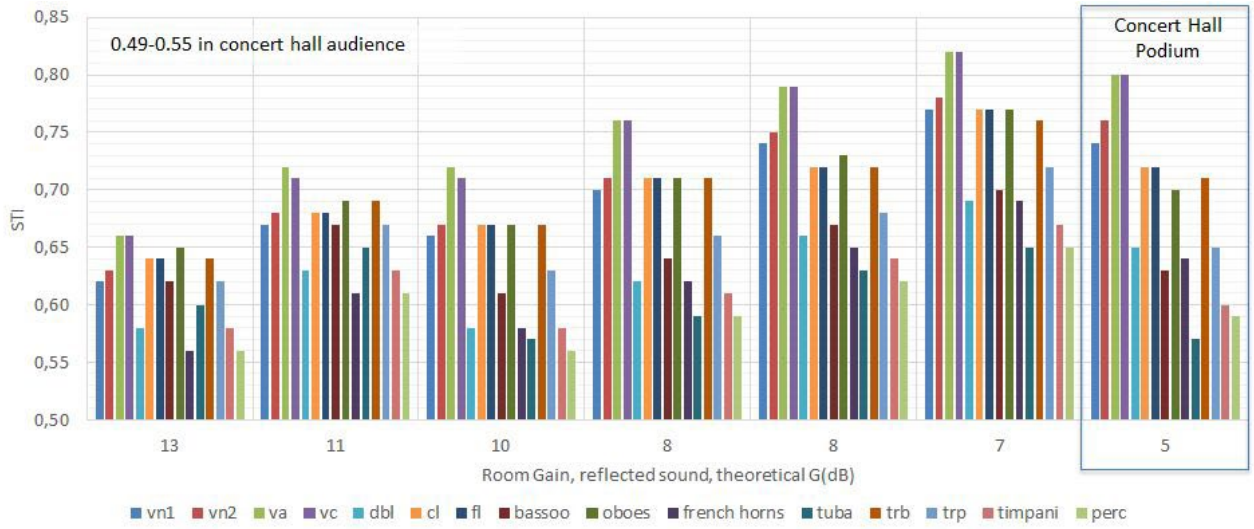


Figure 5 Instrument groups' articulation profiles in terms of STI from each instrument group to the conductors ears in the 7 room models. In the concert hall audience, the articulation is 0.49-0.55, which represent the baseline of the diagram.

The articulation profile received at the concert hall podium is quite uneven, ranging from 0.57 to 0.80, which is more uneven than in most of the rehearsal room models. However, the important fact is that in any of these room models, the articulation of any instrument group is higher at the conductor's ears than in the audience.

This means, 1) that any articulation details audible in the audience, will be detactable by the conductor, but 2) the articultaion will be softer, or more blurred, in the audience than perceived directly by the conductor, and the conductor cannot exactly judge to what degree the musical articulation is conveyed to the audience. Whether this lack of monitoring the musical articulation is an actual problem or not is not clear, but it is hypothised that while conductors sometimes work hard to over-emphasise articulation to fight excessive blending in over-reverberant rooms, they would rather not do it if they could be sure the articulation was sufficient.
So, for now we assume that it is important for a conductor to know whether or not the music is articulated just adequate, but we conclude that the exact judgement of articulation cannot be made from the conductors position in any room, regardless of its acoustics. This means that the conductor in this concern has to move into the concert hall auditorium to make the judgement.

7.3 Reverberance

The reverberance profiles in Figure 6 shows that the reverberance of the instrument groups heard by the conductor on the concert hall podium is quite representative for the reverberance in the audience, with a exceptions. Again, the proximity to the nearest strings, the violas and celli in particular, affects the orchestra EDT-balance, but 1st & 2nd violin groups are not as strongly affected by proximity as was the level balance in Figure 3. Room model #3, which does not comply to the V-T requirements of the standard, is the rehearsal room that has reverberance at conductor's ears closest to that on the concert hall podium. Among the rehearsal room models that comply with the standard, #4 is the one with conductor's reverberance closest to that on the concert hall podium.

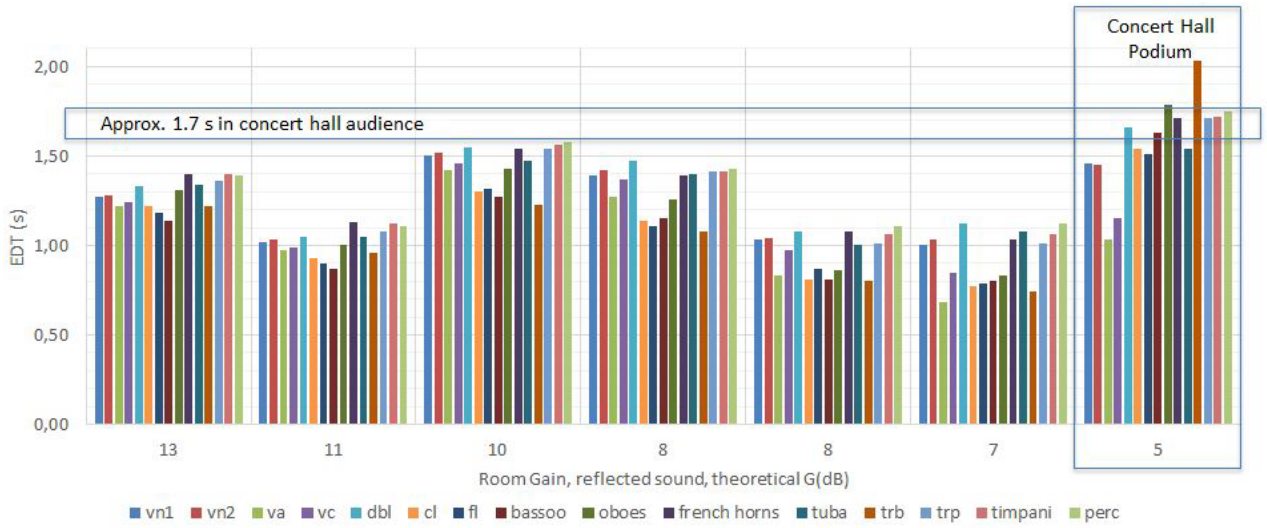


Figure 6 Instrument groups' reverberance profiles, in terms of EDT (s) from each individual instrument group to the conductor's ears, in the 7 room models. In the concert hall audience, EDT is in the 1.6 to 1.8s range. Note that on the concert hall podium, the four string groups, 1st & 2nd violins, violas and celli, are perceived with less reverberance than the audience, while the rest of the instrument groups are perceived with revereberance similar to that in the audience. One exception is the trombones (trb) which sounds more reverberant in the conductor's than in the audience' ears, which in the concert hall model is due to a rather late reflection off the back wall in the auditorium, an effect made stronger by the directivity of the trombone.

7.4 Direct-to-Reverberant level balance

From Figure 8 we see that the closer the theoretical G_{refl} is to 5dB, the closer the D-R is to 0dB, i.e. the direct sound from the orchestra equals the room response. A relatively high D-R makes localization possible, a feature which is very important during rehearsal sessions, in order for the conductor to presicely localize and detect details in individual instruments or groups. D-R should not too high, though, and D-R=0dB is considered an optimum. G_{refl} =5dB is considered to be optimum in concert halls, but since this requires a lot of absorption, such conditions would require volumes of more than 10.000m^3, to provide for other acoustical aspects.

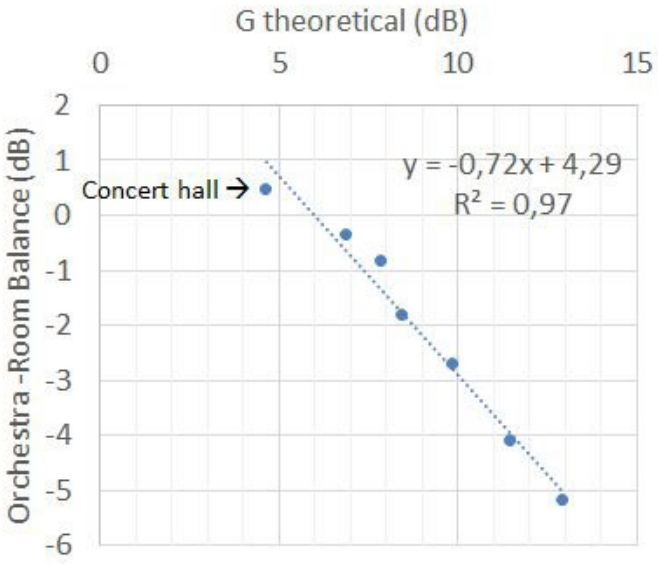


Figure 7 The seven room models plotted with orchestra's D-R level balance on the vertical axis and theoretical G_{refl} on the horizontal axis. The closer the theoretical G_{refl} is to 5dB, the closer the D-R is to 0dB

7.5 Ease of speech during vocal instruction

From Figure 9 we see that conditions for vocal instruction are very good in the rooms complying with the standard. The concert hall deviates negatively from the trend, naturally due to late reflections from the beck of the concert hall auditorium. In the rehearsal rooms, though, the trend is again - the closer to G=5dB, the better.

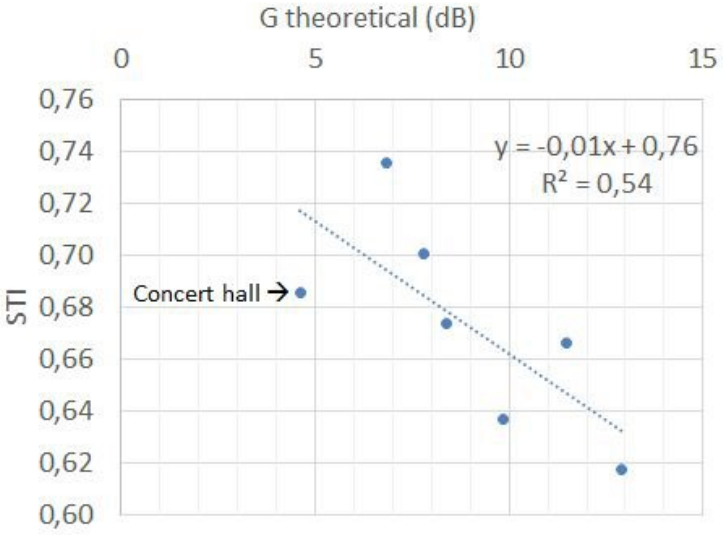


Figure 8 Average Speech Transmission Index (STI) in the communication line between conductor and musicians, in the 7 room models. During rehearsal on the concert hall podium, STI=0.68, whis is considered very good. In the two room models comlying with the standard, #4 and #5, values are 0.68 and 0.70.

8 Conclusion

Orchestra level balance as heard by the concert hall audience cannot be judged directly from the conductor's podium in any of the rooms. However, it can be judged by listening to the instrument balance in the room response, on the concert hall podium and in one of the rooms (#5) complying with ISO-23591:2021. Articulation of music is sufficient in all the rooms studied, but it cannot be optimized in any of the rooms. For this purpose, the conductor needs to visit the concert auditorium. Reverberance in the approved rehearsal halls is too dry to be representative for the concert hall audience' sound, while reverberance on the concert hall podium is quite like that at audience ears, except for the reverberance from the violins nearest to the conductor. When it comes to conductor's localization of individual instruments as well as conditions for vocal instructions, the rule seems to be the closer to G_{refl} =5dB, the better. However, such low room response must be combined with somewhat bigger volume than in the current study, bringing us closer to the point where the rehearsal room volume approaches that of a performance hall.

9 Further work – Questionnaire

To learn more about how conductors at work are influenced by acoustics, a questionnaire is suggested below, as a starting point for discussion. The intention is to simulate acoustics of the performance rooms judged by the conductors, in search for critical features that can be predicted and used to assist the design of new spaces or the problem-solving of existing spaces.

"In this questionnaire, we invite you to judge the conductor acoustics of several well-known performance spaces, specifically how well their acoustic conditions support your work as a conductor.

- In a performance space in the list where you have conducted an orchestra once or more, what is your overall judgement of its Conductor Acoustics, on a scale from 1 to 5?
- On a scale from 1 to 5, how much do the following questions count in your judgement?
    - How easy/difficult is it, at conductor's podium, to prepare the orchestra's sound and balance as you would like it to be heard in the audience?
    - How easy/difficult is it, from the podium, to judge the audience' listening conditions?
    - How easy/difficult is it to hear important details in the orchestra?
    - How easy/difficult is it to give verbal instructions during rehearsal
    - Do you hear an adequate amount reverberance from the room? "

Acknowledgement

The author wishes to acknowledge conductor Eivind Gullberg Jensen for fruitful input and discussions, and the reviewers for their helpful and encouraging comments.

Reference

[1] Meyer, J., Acoustical Demands for the Conductor's Location, Build.Ac., June 2008

[2] ISO, Acoustic quality criteria for music rehearsal rooms and spaces, ISO-23591:2021

[3] Lehtimäki E., The effect of concert hall acoustics on the work of a conductor, MSc Thesis, Aalto University, School of Electrical Engineering, Helsinki, Finland

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Feasibility of inhomogeneous MPPs in multipurpose halls : The case study of Bilkent Concert Hall in Ankara

Ela Fasllija[1*], Semiha Yilmazer[2]

[1]Department of Interior Architecture and Environmental Design, Bilkent University, Ankara, Turkey.

[2]Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN 47907, U.S.A.

*e.fasllija@bilkent.edu.tr

Abstract

For multipurpose halls, there are additional constraints in planning for good visual conditions and sightlines, meeting specific social demands concerning layout, and providing good acoustic conditions. Regarding the latter, they require a compromise between the auditory needs of speech and music. While conventional materials fail to absorb sounds in the low-frequency range, Helmholtz resonators show a very narrow absorption bandwidth. The current study investigates the potential created by inhomogeneous Micro-Perforated Panels (MPPs), which literature demonstrated to have a wider absorption bandwidth in the low-mid frequency spectrum. This work uses Bilkent concert Hall, a 700-seat capacity multipurpose trapezoid-shaped hall located in Ankara, Turkey, as a case study to examine its current acoustic condition and provide material-based solutions for the existing problems. It analyzes the hall by assessing the fundamental acoustic parameters such as reverberation time (T30), sound pressure levels (SPL(A)), and speech transmission index (STI). Comparative simulations between the existing and the proposed material-based solution of the Bilkent Concert hall showed an improvement in the reverberation time and other objective acoustic criteria when inhomogeneous MPPs were used as wall panels instead of conventional materials. This comparison enhances the existing problems in the multipurpose hall and emphasizes the viability of those "next generation" absorbing materials, which are at the same time in demand from architects due to their design feasibility.

Keywords: inhomogeneous MPPs, multipurpose halls, reverberation time, absorption

1 Introduction

The arts of music, drama, and public discourse have both influenced and been influenced by the acoustics and architecture of their presentation environments [1]. Despite the use of halls for more than one purpose, the practical design of those spaces to accommodate more than one type of function is quite recent. A degree of flexibility in use is now becoming the norm rather than the exception. From the Sabine equation, the two variables influencing the reverberation time (RT60) are the internal volume and the amount of acoustic absorbing material. Optimum reverberations values for multipurpose halls should result in an RT60 (1,4 - 2 sec) appropriate for the program that can be speech and/or music. This flexibility is made possible in literature by architectural elements such as mobile bridges, automatic curtains, and retractable walls [2]. In an acoustic analysis of Gottingen Stadthalle [3], where RT was higher than predicted, the authors pointed out that the use of absorbing materials is more effective than geometrical changes because the former is not affected by diffraction.
State of the art in absorbing materials clearly shows that passive absorbers such as fibrous or porous materials are indispensable for high-frequency damping noises. For treating low-frequency problems, resonant structures are usually used. These reactive structures, also known as resonant absorbers, are mass-spring systems with damping to absorb the system's resonant frequency. They are most commonly used in room acoustics in the form of membranes, Helmholtz resonators, perforated, slotted, and micro-perforated panels (MPPs) [4]. This study investigates the potential of inhomogeneous Micro-Perforated Panels, which are known to have a wider absorption bandwidth, especially in the low-mid frequency range, when used as wall panels in a multipurpose hall located in Bilkent University, Ankara, Turkey.

1.1 Micro-Perforated Panels (MPPS)

MPPs were installed for the first time in 1992 to solve acoustic problems during the opening of the German Bundestag in Bonn. They can be made of transparent or colorful plates or membranes, so they are also in demand by architects for sound quality control in auditoriums [5]. As all resonant mechanisms, they come with the handicap of having a very narrow absorption bandwidth. A series of structures based on MPP networks in series and parallel have been explored to introduce multiple resonances [6-7] to broaden the absorption bandwidth. The results yielded excellent agreement between theory and measurement, showing show that the absorption bandwidth is expanded to lower frequencies due to the additional multi resonance peaks. This approach is viable to be integrated into room acoustic applications. The authors of this study have developed and tested an inhomogeneous (4 perforations of 0.7mm, 1mm, and 2mm with different perforation ratios) Micro-Perforated Panels with various cavities (8cm, 6cm, 4cm, 2cm) made of acrylic, having a thickness of 6 mm and density of 1190 kg/m3. The absorption coefficients of the proposed material are tested in an impedance tube by the transfer function method, and the results are as follows:

Table 1. The absorption coefficient of the proposed Inhomogeneous MPP

Inhomogeneous MPP proposed by the authors	***125 Hz***	***250 Hz***	***500 Hz***	***1000 Hz***	***2000 Hz***	***4000 Hz***
	0,09	0,27	0,86	0,76	0,54	0,24

The material will be tested for its performance when used as a wall treatment in room acoustics applications such as multipurpose halls. In the hypothetical case, the proposed material will replace the adjustable wood panels used to enhance different modes of the Bilkent Concert hall located in Ankara.

2 Case study: Bilkent Concert Hall

The Bilkent Concert hall (Figure 1) is a part of the Faculty of Music and Performing Arts building designed during 1993-1994 by architect Ilhan Kural. It is like a trapezoidal room in a plan which is enlarged from the stage through the back rear wall. It consists of the main floor, a second floor with space for the chorus, side aisles, and the main balcony, and the third floor with only a small balcony (Figure 2). It has a volume of nearly 6500$m^{3,}$ and the clear height is 15,2m. For the time being, the hall is used mainly for music purposes by the Bilkent Symphony Orchestra and for speech purposes when important lectures are given. There are 685 seats on the main floor and the second floor of the concert hall. 473 are placed on the main floor, 151 are placed on the second floor for the audience, and the remaining 61 seats are reserved for the chorus.

Figure 1. Interior view of the Bilkent Concert hall

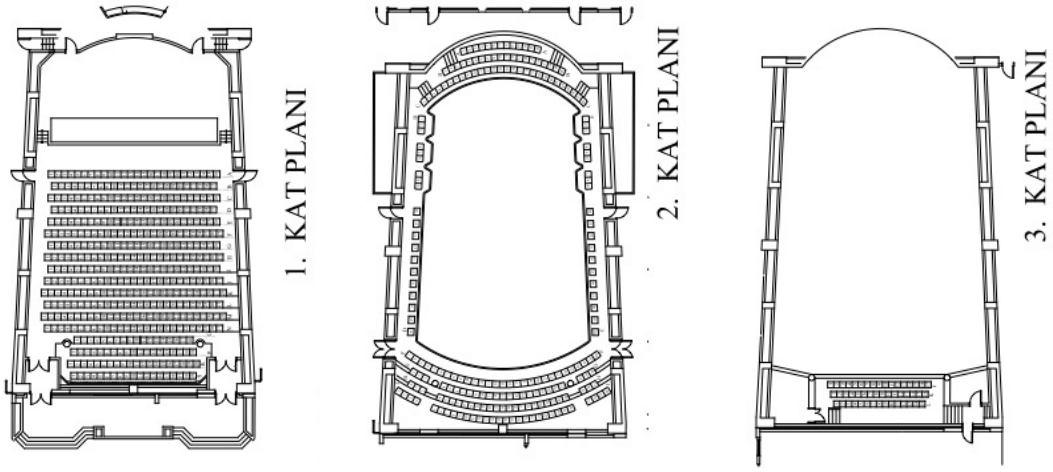


Figure 2. Plans of the Bilkent Concert hall

The materials for the respective surfaces are present in Table 2. The back walls are made of wood/gypsum panels, a layer of fibrous material such as rock wool, and an air cavity. Sidewalls are also separated with an air space and have travertine veneering and a 10 cm rockwool layer. On the sidewalls of the second floor, there are adjustable wooden panels. Simply folding wooden panels on sidewalls has carried out a flexible solution to adjust the reverberation time and other acoustical concepts. For the auditorium mode, the panels are folded on themselves to be able to face the highly absorptive material underneath with the hall medium, whilst concert mode is created by the folding of wooden panels on the highly absorptive material.

The wall behind the chorus area is made from fixed, veneered chipboard and a 10 cm layer of rock wool. The stage has a curved wall made from fixed, veneered chipboard, rock wool, and a concrete backing. There are plastered columns in front of this wall placed at the same angle as the stage wall's curvature. The balcony fronts and the ceiling are all plastered surfaces. The finishing material used for the floor and the stage is wood parquet, and there are upholstered seats for the audience (see Figure 3).

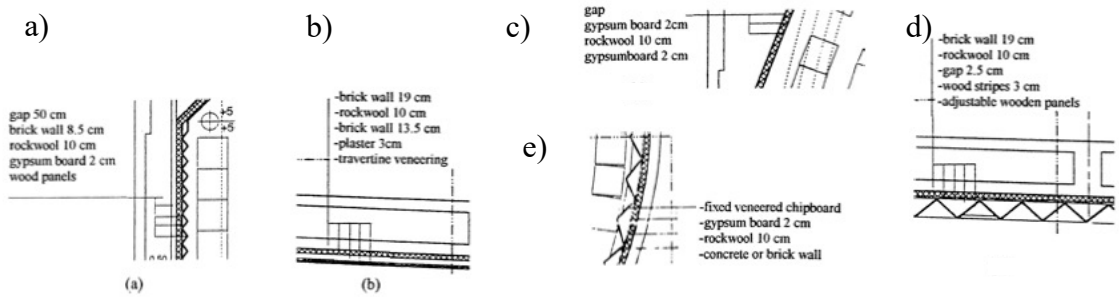


Figure 3. a) back wall (1), b) sidewalls of the main floor (1), c) back wall (2), d) side walls (2), e)chorus wall (2) [8]

3 Comparative simulations

The 3D model of the hall was achieved by using 3D SketchUp and then imported into ODEON Room Acoustics software version 16 Basics version suitable for educational purposes. It was composed of 612 surfaces, resulting in a level of details accurate for acoustics simulations. GA software describes the sound propagation by using ray-tracing methods. The water tightness of the model was verified, as seen in Figure 4.

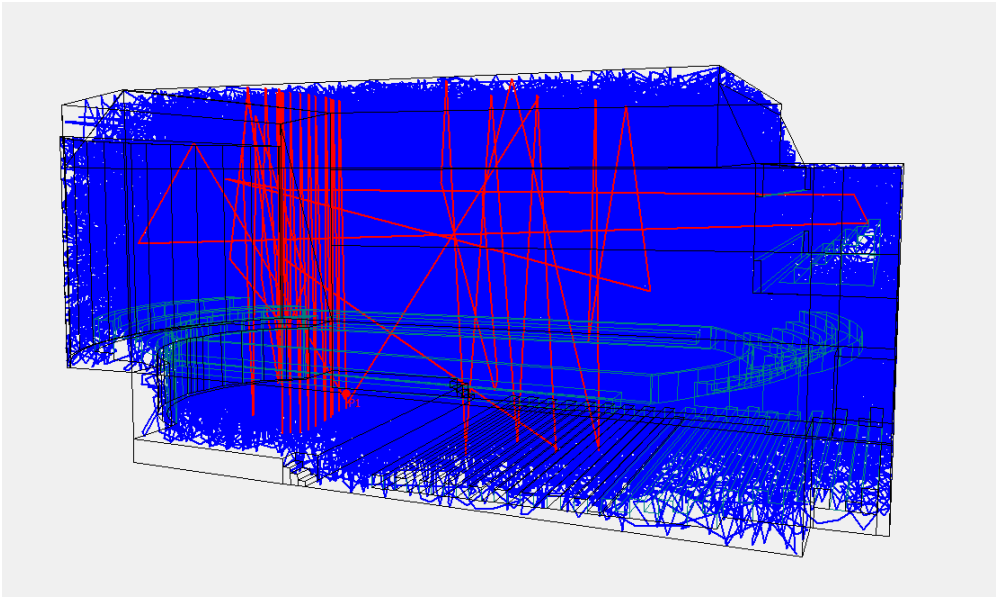


Figure 4. Water tightness of the model

The modeling process was carried out according to the state-of-the-art recommendations, and an omnidirectional point source of 60 dB is located in the middle and 2 m away from the stage front. Six different positions for receivers were also located, representing various audience locations. The source height was decided to be 1,5m, whereas the receivers' height was selected to be only 1,1m, resembling the eyesight of a human head when they sit on the seat (Figure 5).

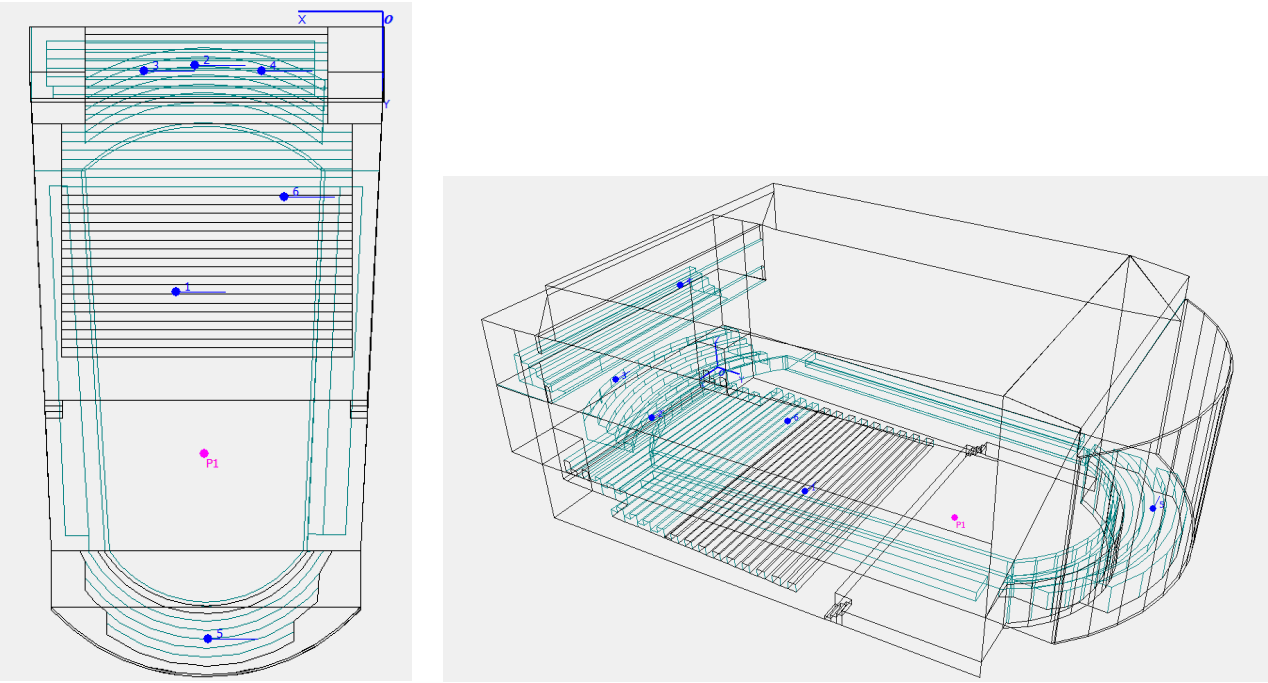


Figure 5. Position of source and receivers

Then existing materials, as shown in Table 2, are assigned respectively to the surfaces. When halls are used for classical music purposes, the Reverberation Time (T30) should be around 1,4-1,6 s as an optimum value. In contrast, when used for theatre and other speech performances, the range should lay 0,8-1,0s. The Speech Intelligibility Index (STI) should be larger than 0,6.

Table 2. Materials of the space

Surfaces	**Materials**	**Absorption coefficient**					
		125	***250***	***500***	***1000***	***2000***	***4000***
Walls	Stage Wall: veneered chipboard	0,10	0,11	0,10	0,08	0,08	0,11
	Chorus Wall: veneered chipboard	0,10	0,11	0,10	0,08	0,08	0,11
	Rear Wall (1): Wood panels + rockwool (10cm) + cavity (50cm)	0,40	0,80	1,00	1,00	0,95	0,80
	Rear Wall (2): Gypsum board + air cavity (2,5cm) + rockwool (10cm)	0,10	0,07	0,05	0,05	0,04	0,04
	Side Walls (1): Travertine veneering	0,36	0,44	0,31	0,29	0,39	0,25
	Side Walls (2): adjustable wooden panels (auditorium)	0,40	0,8	1,00	1,00	0,95	0,80
	Side Walls (2): adjustable wooden panels (concert)	0,10	0,11	0,10	0,08	0,08	0,11
Floor	Wooden flooring	0,20	0,15	0,10	0,10	0,05	0,10
Ceiling	Paster	0,10	0,07	0,05	0,05	0,04	0,04

Other	Balcony fronts: Plaster	0,10	0,07	0,05	0,05	0,04	0,04
	Door: Solid wood	0,20	0,65	1,00	1,00	1,00	0,90
	Audience: Upholstered concert hall chairs	0,56	0,64	0,70	0,72	0,68	0,62

The results show that the hall has a reverberation time of 0,95s in the auditorium mode (Table 3)and 1,45s in the concert mode (Table 4) for mid frequencies. It is seen that the hall is in the optimum levels for both modes due to the adjustable wooden panels on the side walls. Moreover, in the hypothetical simulation, when only an inhomogeneous MPP with no fibrous material and different cavities (max cavity 80mm) is used instead of adjustable wooden walls, the T30 values lie in the optimum range for multipurpose halls (Table 5).

Table 3. (T30) values for Auditorium mode

Existing	125 Hz	250 Hz	500 Hz	1000 Hz	2000Hz	4000 Hz
Reverberation Time (T30) Auditorium mode	1,22	1,04	0,95	0,93	0,9	0,87

Table 4. (T30) values for Concert mode

Existing	125 Hz	250 Hz	500 Hz	1000 Hz	2000Hz	4000 Hz
Reverberation Time (T30) Concert mode	1,45	1,42	1,45	1,42	1,40	1,24

Table 5. (T30) values for hypothetical mode

Hypothetical	125 Hz	250 Hz	500 Hz	1000 Hz	2000Hz	4000 Hz
Reverberation Time (T30) Multipurpose hall	1,37	1,25	1,02	1,03	1,11	0,99

Calculations about (SPL(A)), indicating the sound pressure level perceived by the human ear as a subjective judgment of Sound Pressure Level, clearly show the decrease in the sound energy and intensity with the increasing distance. Table 6 shows the changes of SPL(A) for six receivers in the three modes: existing - Auditorium, existing (Concert), and hypothetical when inhomogeneous MPPs are used. The results show that there exist significant problems with the energy distribution, especially in the receivers positioned under the balconies (R2, R4).

Table 6. SPL(A) levels for the respective receivers in the three modes

SPL(A)	Existing-Auditorium (dB)	Existing-Concert (dB)	Hypothetical (dB)
Receiver 1	43,4	44,2	43,7
Receiver 2	37,4	38,1	37,7
Receiver 3	40,1	42,4	41,2
Receiver 4	38,0	40,9	39.4
Receiver 5	42,0	43,1	42,5
Receiver 6	41,2	42,4	41,8

The following evaluation was carried out about Sound Transmission Index (STI) value among the audience area. The STI of the hall is found to be between 0,48 to 0,60, which satisfies the range for fair speech intelligibility value, as the value lies between 0,45 and 0,60. The slight decrease in the central region can be explained by the sharp ceiling and lateral reflections, which will tend to decrease intelligibility. The slight increase in rear seats occurs due to highly absorbing materials, which tend to absorb sound energy and enhance intelligibility.

Table 7. STI values for the respective receivers in the three modes

STI	Existing-Auditorium	Existing-Concert	Hypothetical
Receiver 1	0,58	0,52	0,55
Receiver 2	0,60	0,54	0,57
Receiver 3	0,58	0,49	0,53
Receiver 4	0,55	0,48	0,51
Receiver 5	0,59	0,51	0,55
Receiver 6	0,57	0,50	0,54

4 Conclusions

This study aimed to investigate and introduce newly designed materials such as inhomogeneous MPPs in room acoustics applications instead of conventional materials. The simulations related to T30, SPL(A), and STI run in the case study showed that inhomogeneous MPPs can substitute large composite walls of considerable thicknesses by accurate design and an overall thickness of less than 10cm. Those materials pose a great potential in interior architecture as fiberless absorbers that can cope with the highest hygiene demands by having a tunable design according to the needs of the space.

5 References

[1] Long, M. *Architectural acoustics*. Boston: Elsevier/Academic Press, 2014.

[2] Cairoli, M. Architectural customized design for variable acoustics in a Multipurpose Auditorium. *Applied Acoustics*, 140, 167–177, 2018, doi:10.1016/j.apacoust.2018.05.026.

[3] Kuttruff, H. A simple iteration scheme for the computation of decay constants in enclosures with diffusely reflecting boundaries. *The Journal of the Acoustical Society of America*, 98(1), 288–293, 1995, doi: 10.1121/1.413727

[4] Cox, T. J., & DAntonio, P. *Acoustic absorbers and diffusers: theory, design, and application*. Boca Raton: CRC Press, 2009.

[5] Fuchs, H. V., & Zha, X. Micro-perforated structures as sound absorbers–a review and outlook. *Acta acustica united with acustica*, 92(1), 139-146, 2006.

[6] Cobo, P., & Simón, F. Multiple-layer microperforated panels as sound absorbers in buildings: A review. *Buildings* (Vol. 9, Issue 2). 2019 MDPI AG. https://doi.org/10.3390/buildings9020053

[7] Li, D., Chang, D., & Liu, B. Diffuse sound absorptive properties of parallel-arranged perforated plates with extended tubes and porous materials. *Materials*, *13*(5). 2020

[8] Sahin E. *Passive sound control in symphony concert hall design*. Master's Thesis, Bilkent University repository, 1995.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Room Acoustic design for electronic enhancement systems

Henrik Möller

Akukon Oy, Helsinki, Finland

henrik.moller@akukon.com

Abstract
Within the last 10-15 years, electroacoustic enhancement systems have become a real tool to improve the acoustic conditions in multi-functional halls, both as full systems, where the total acoustic field is created electronically, and as an extension of existing acoustic conditions in the hall.

However, one common misconception is that it is possible to "repair the acoustics" or fix problems with an electroacoustic enhancement system. The fact is that good acoustic conditions are necessary for the successful implementation of an electroacoustic enhancement system.

This paper will give a brief overview of the current technology and present the typical acoustic requirement and examples of solutions

Keywords: room acoustics, electroacoustic enhancement

1 Introduction

Electroacoustic enhancement systems have been used for more 50 years, mainly as a method of increasing the reverberation time in halls. One of the first examples is the Royal Festival Hall in London, [1] where an Assisted Resonance system was installed soon after the hall was opened to increase the reverberation, in particular at mid and low frequencies. In the last 20 or so years, more elaborate systems have been in use, providing both extra reverberation but also able to generate early reflections, in other words able to create an artificial acoustic environment.

2 Typical topologies

In general, one can distinguish between two principally different approaches, In-line systems and Feedback systems [2]. However most modern systems are a combination of both approaches.

2.1 Inline systems

The Inline system is essentially creating a fully artificial acoustic environment by sampling the sound at the source (on the stage) and producing the acoustic field through loudspeakers. Examples of this are the LARES system and the ACS system. The advantage of this that the system can also produce early reflections and that it can easily be used as a part of a reinforcement system. The disadvantage is that full late reverb is also electronically generated, and at least for some system does not fully create a naturally sounding reverb.

2.2 Feedback System

The Feedback system is in a way a continuation of the early assisted resonance systems. Examples of this approach is the Carmen system. The idea is that the sound in the hall is sampled and reinforced. In other word, the existing reverberation is reinforced. This makes it quite easy to create natural sounding reverberation. The disadvantage is that the system does not provide early reflections, in other words there are no clear control of the spaciousness of the acoustic field. Some newer systems also add actual reverberators, making it possible to create longer reverberation than just by reinforcing the sound field. Examples of such system are for instance the MCR system (Multi Channel Reverberator).

2.3 Current systems

As stated above, most currently commercially available systems, use a combination of inline and feedback approach. Examples of such systems are Constellation, Vivace and Amadeus. Common for them are that they are built around a central processing unit (like the classic InLine systems) but they also use microphones in the auditorium.

3 Acoustic requirements for electroacoustic enhancement systems

As said earlier it is a misconception that an electroacoustic enhancement system can fix all acoustic problems. While it is possible to enhance reverberation in a hall with a feedback system, a system in general requires very “well behaved” acoustic conditions in the space, both concerning reverberation, reflection patterns and background noise levels.

Typically, a hall fully designed for an electroacoustic enhancement system, will be designed with the same guidelines as for a cinema. However, as these halls are sometimes bigger than any cinema (for instance LOGOMO has a maximum capacity of 3000 people), the cinema specs cannot always be used directly or rather scaled up to fit these halls.

3.1 Reverberation

For InLine and combined systems, it is essential that the reverberation time is flat with frequency. The room should not be totally “dead” as some ambience is needed to cover up any reflections from hard surfaces (doors, floor etc) and other audible reflections. But still the ideal reverberation time at least for an InLine system will be close to that of a cinema hall.

If the hall geometry provides sensible (early)reflections, it is possible to just extend the reverberation time of the hall with a feedback system. In this case the reverberation time of the hall is not so important, however again, defects in the reverberation can not necessarily be fixed with the system.

3.2 Reflection patterns

The systems cannot remove unwanted reflections but can add needed reflections. For example, any kind of echoes or flutters, will also be amplified by the system, so in case the room has this kind of problems, it will be necessary to add absorption or scattering to reduce these.

3.3 Background noise

At the systems in question are essentially reinforcement systems, they will also reinforce any background noise. This implies that the maximum permissible background noise levels, will actually be significantly less in a room designed for an electroacoustic enhancement system, than in a normal multipurpose room.

4 Design examples

In the following 3 different halls with rather different approach are presented.

4.1 G-live Lab, Tampere and Helsinki

The G-live labs in Tamper and Helsinki, are restaurants with live music. The enhancement systems used are proprietary systems designed by the Acoustics Group of the Aalto university and build around a Genelec loudspeaker system. The system is essentially an InLine system; however, it is also used for normal amplification. This means that it is possible for the sound engineer to dynamically adjust the system and use it as part of the show.

Both halls were designed with an emphases on “visual impact”, using novel interior ideas, such as a stage floor in Helsinki made from glass boards/strips, glued together to make a 100 mm thick glass stage.
In Helsinki, the main acoustic surface is a thick absorption on the ceiling surface, combined with curtains, both on the stage and along the outside window wall. The height of the space is less than 3 m, meaning that the ceiling absorption proved sufficient to achieve the acoustic requirements.

Figure 1: G-Livelab Helsinki, view towards the stage

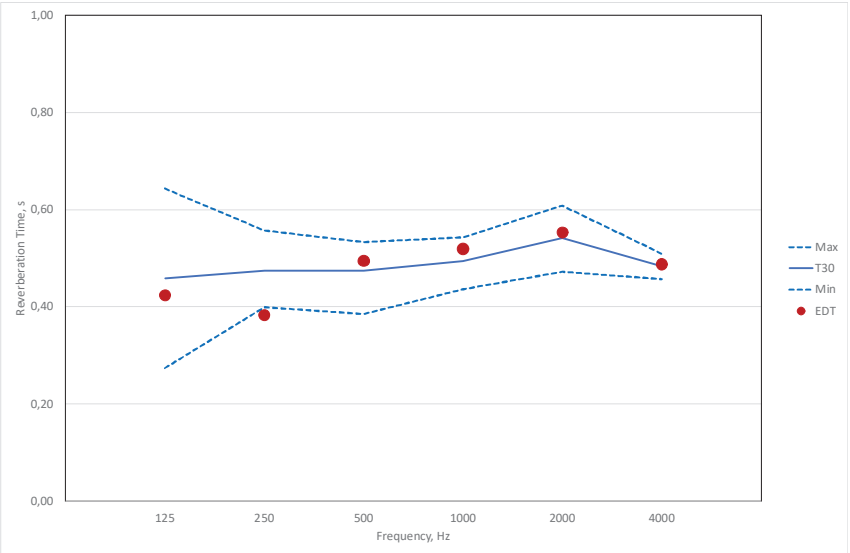


Figure 2: Reverberation time in G-livelab Helsinki

The G-Livelab in Tampere is in two floors with a vaulted ceiling. The whole ceiling is treated with a combination of mineral wool and perforated boards with mineral wool behind. The walls in the space are also treated mainly with perforated boards with mineral wool behind. Furthermore, some of the bended to provide a bit of scattering. The venue was the winner of the mondo*dr 2020 EMEA & APAC as the best concerthall in 2020.

Figure 3: G-Livelab Tampere, view from the balcony towards the stage

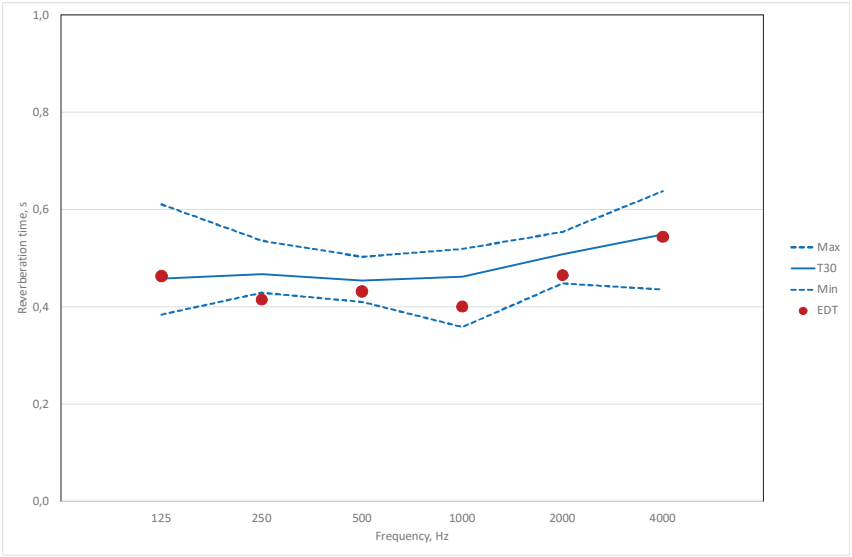


Figure 4: Reverberation time in G-Livelab Tampere

As can be seen from figure 2 and 4, the reverberation time in both cases is quite short. Both measurements are done in the empty hall, so in conditions with audience, the reverberation time at higher frequencies will be somewhat lower, in particular in the Helsinki venue. Also, other acoustic parameters showed very homogenous acoustic conditions throughout he whole space. For example, the STI_{geo} drops less than 0,15 over a distance of 25 m in the Helsinki venue and also both C_{80} and C_{50} are fairly constant within the whole space.

4.2 LOGOMO, Turku

The complex of Logomo is situated in Turku, the oldest city in Finland. Logomo is a new home for various cultural, creative and business events. The main Logomo hall was opened in November 2011. The rest of the

complex, including a further four performance arenas, studio areas, exhibition spaces, office spaces, a restaurant and workrooms for artists, was opened in 2014.
For the main hall to be suitable for differently sized activities physical versatility should be acquired a system with a massive movable seating stand that rides back and forth on compressed air cushions. The stand can be positioned in three different locations to form three configurations for the hall: small (S), medium (M) and large (L) (see figure 5).

The hall has a Meyer Sound Constellation system installed. The system is pre-programmed for the diffe
The objective for the acoustic design was to ensure sufficiently well controlled reverberation time, but also ensuring some “ambience” and, in particular in the large setting, to avoid clear echoes or audible reflections. In order to achieve this, the geometry of the backwall of the stand was optimized to avoid reflections with a combination of absorbing and scattering surfaces.

The side walls are mainly different kinds of perforated metal plates, with some additional scattering surfaces, in particular in the front of the hall.

Figure 5: LOGOMO hall in the Small, Medium and Large settings

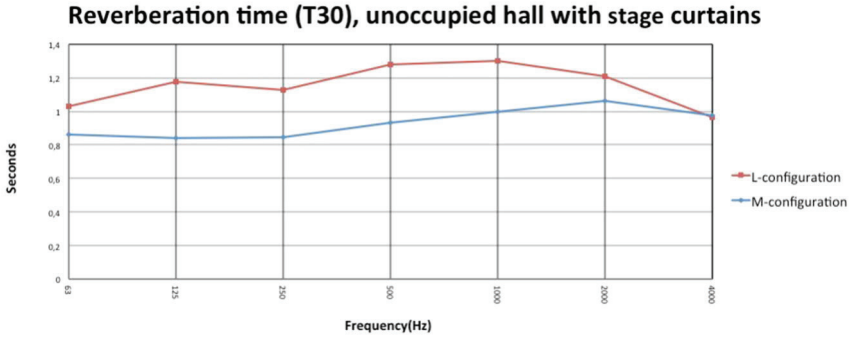


Figure 6: Measured reverberation time for the Medium and Large configuration

As can be seen from figure 6, the reverberation time does not increase dramatically even for the Large configuration. There are no detected discrete reflections or other acoustic defects in the hall.

5 Discussion and Conclusion

In this paper, a short review of the topologies of the electro acoustic enhancement systems as well as review of the room acoustic requirements for these systems are presented and 3 different venues with electroacoustic enhancement systems are presented.

It is clear that electro-acoustic enhancement system has become a tool for creating variable acoustics for multipurpose venues. There are ready made systems that will work for many halls but the use of electro acoustic enhancement should perhaps not be seen as only a question of installing a full system. It is still clear that if one is designing a dedicated concert hall, then an electroacoustic enhancement system is not the first choice.

However, it is also clear that these systems can successfully be used to extend the acoustic/reverberation in halls which are perceived as too dry for acoustic music. The important is that this will only work if the hall does not have actual acoustic defects. In other words, in some case in will be necessary to first add absorption to remove echoes/audible reflections and the use electronics to enhance the reverberation.
The advantage of the current combination systems is that they can be adapted to different conditions, and solve separate issues, such as hearing on stage or overall reverberation. The main difference compared to earlier systems, is that the sound quality is has improved to a point where it can be difficult, even for a trained listener, to detect if the system is on or not.

So electroacoustic enhancement systems is clearly one of the tools which can be used, both to create excellent multipurpose halls and also to some extend to extend the acoustic conditions in existing halls.

References

[1] M. BARRON:Auditorium Acoustics and Architectural Design". E&FN Spon, London 1993

[2] R.BAKKER, S. GILLAN: The histoy of active acoustic enhancement systems, Proceedings of the Institute of Acoustics, Proceedings of the Institute of Acoustics, Vol. 36. Pt. 2. 2014

[3] J. RIIONHEIMO, A. RUUSUVUORI, H. MÖLLER: Logomo – A new multi-purpose hall in Turku, Finland, BNAM 2012, 18-20.6.2012, Odense, Denmark

[4] H. MÖLLER : A review of STI measurements, Forum Acusticum 2020, 7-11.12.2020, Lyon, France, (Online presentation)

[5] S. Tervo, J. Kestilä, T. Ilomäki, J. Pätynen, T. Lokki: Pienen esitystilan huoneakustiset ratkaisut, Akustiikkapäivät 2017, 24-25.8.2017, Espoo, Finland

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Sound insulation, residents' satisfaction, and design of wooden residential buildings

Fredrik Ljunggren[1,*]

[1]Engineering Acoustics, Luleå University of Technology, Sweden.

*fredrik.ljunggren@ltu.se

Abstract
Wood-based multi-family houses continue to gain popularity. Related to acoustics, low-frequency sound insulation as well as appropriate single number quantities for the evaluation of sound insulation have been in focus for a long time. In a series of Swedish research projects running for 12 years, the correlation between rated annoyance from residents and measured airborne and impact sound insulation, with alternative frequency ranges and weightings, have been studied. In total, 38 building cases of various constructions were involved and more than 1200 questionnaire responses were collected.
While the building code's present evaluation parameter for airborne sound insulation, $D'_{nT,w} + C_{50–3150}$, seems to be working well, the situation is different with respect to impact sound insulation. $L'_{nT,w}$ as well as $L'_{nT,w} + C_{I,50–2500}$ show weak correlation with the rated annoyance from the residents. The reason is that frequencies below 50 Hz are overlooked, although they dominate the response from walking in many common, particularly lightweight, floor constructions. The strongest correlation with the rated annoyance from impact sound, including both lightweight and heavyweight constructions, was found when the measured frequency range was extended down to 25 Hz, using $L'_{nT,w} + C_{I,25–2500}$. Because footstep noise rendered the highest degree of annoyance in the survey, a somewhat more restricted requirement than what it used today is suggested to offer a higher degree of protection against unwanted impact sounds.
It is a delicate challenge to design wood-based floor constructions with great sound insulation at low frequencies to meet the higher requirement. A tested innovative floor design based upon two high-density cross-laminated timber plates with an intermediate damping layer may serve as the basis for future constructions.

Keywords: sound insulation, impact sound, residents' annoyance, single number quantity, wooden floor design

1 Introduction

Residential buildings have gained increasingly popular in recent years. It is clear that the acoustic properties of wood-based houses differ from what is known from heavier constructions, like houses made of concrete. As a consequence, sound insulation may need to be evaluated in a different way than what is common today in order to apply to both lightweight (wood) and heavyweight (concrete) buildings. This topic has been in focus in three Swedish research projects – *AkuLite*, *Aku20* and *AkuTimber*, starting in 2009, where the latter project is nearing completion.
The objective of this paper is to summarize the extensive research carried out during more than a decade. The work focus primarily on how to evaluate sound insulation between dwellings to get best possible correlation between the measured airborne and impact sound insulation versus the reported annoyance from the residents. A second is to develop the design of wooden floors for superior low-frequency impact sound insulation.

2 Field measurements

2.1 Method

In total, 38 residential buildings of two to eight floor levels, were included in the study. The buildings represent a variety of mainly modern building techniques. The building case studies are divided into three sub-categories with respect to the construction of the separating floors, mainly related to the mass of the construction. Out of the 38 case studies, 17 are lightweight, 11 are of CLT and 10 are of concrete:

1. *Lightweight* – loadbearing structure of wooden or thin steel beams combined with various types of boards
2. *Cross laminated timber (CLT)* – semi-lightweight structure based upon layers of timber, glued together to form approximately homogenous slabs
3. *Concrete* – homogenous or hollow core slabs

All cases have been examined by field measurements of sound and vibration properties: a) airborne sound insulation, b) impact sound insulation with the standardized tapping machine as well as the ISO rubber ball, c) vibration levels close to junctions (as an indication of flanking transmission) and d) vibration levels of the floor surface at various distances from the impact source. Measurements a–d were performed in the frequency range from 20 Hz. Further details of the measurement procedures are given in [1].
In focus here, is the airborne and impact sound insulation, measured and evaluated according to the appropriate standards: ISO 16283-1 [2], ISO 16283-2 [3], ISO 717-1 [4] and ISO 717-2 [5], or in accordance with the corresponding former standards for the cases investigated at an early stage. The following single number quantities are used: weighted standardized level difference $D'_{nT,w}$ and weighted standardized impact sound level $L'_{nT,w}$, where the reverberation time in both cases is normalised to 0.5 seconds. For simplicity, $D'_{nT,w}$ and $L'_{nT,w}$ are from here on denoted as $D_{nT,w}$ and $L_{nT,w}$ respectively, i.e. without the symbol (′). The definitions of the standardized airborne and impact sound level follow from Equation (1):

$$D_{nT} = \Delta L_p + 10log\left(\frac{T}{T_0}\right) \quad (1a)$$

$$L_{nT} = L_p - 10log\left(\frac{T}{T_0}\right), \quad (1b)$$

where L_p is the sound pressure level, ΔL_p is the sound pressure level difference between two rooms and T_0=0.5 seconds.
A set of spectrum adaption terms are compiled to examine alternative single number quantities using abbreviated denotations. For airborne sound insulation: $D_{nT,w,50}$ corresponds to $D'_{nT,w} + C_{50–3150}$, and for impact sound insulation: $L_{nT,w,50}$ corresponds to $L'_{nT,w} + C_{I,50–2500}$ and $L_{nT,w,25}$ corresponds to $L'_{nT,w} + C_{I,25–2500}$. The latter adaptation term is presently not included in any ISO standard, although it is calculated in the same way as $C_{I,50–2500}$, but in the frequency range starting at 25 Hz. The definition is given by Equation (2)

$$C_{I,X-2500} = 10log\left(\sum_i 10^{\left(\frac{LnTi}{10}-15\right)}\right) - L_{nT,w}, \quad (2)$$

where X is 25 or 50 Hz.
The sound insulation was generally measured in four to six rooms in each building, evenly distributed among living rooms and master bedrooms. The arithmetic mean value of all measurement results within each building case is used for statistical evaluations.

2.2 Result

The single number quantities related to airborne and impact sound insulation among the 38 building cases are presented. The range is 52–69 dB for $D_{nT,w}$ and 47–65 dB for $D_{nT,w,50}$ concerning the airborne sound insulation. Regarding the impact sound insulation, the range is 38–62 dB for $L_{nT,w}$, 46–66 dB for $L_{nT,w,50}$ and 47–68 dB for $L_{nT,w,25}$. The histograms of the sound insulation are presented in Figure 1 and 2. There is a clear trend towards numerically lower airborne sound insulation as the lowest frequency shifts from 100 Hz ($D_{nT,w}$) to 50 Hz ($D_{nT,w,50}$). In a similar way, the impact sound pressure level increases as the lowest frequency is shifted from

100 Hz ($L_{nT,w}$) to 50 Hz ($L_{nT,w,50}$). Applying the extended frequency range from 25 Hz ($L_{nT,w,25}$) can, according to the definition, only increase the weighted impact sound pressure level since more one-third octave band levels are added to the single number quantity, i.e. the total amount of acoustic energy increases. However, the concrete buildings are almost unaffected by the frequency extension from 50 Hz to 25 Hz whereas the CLT and lightweight buildings show a significant increase in single number, which is illustrated by comparing the histograms in Figure 2.

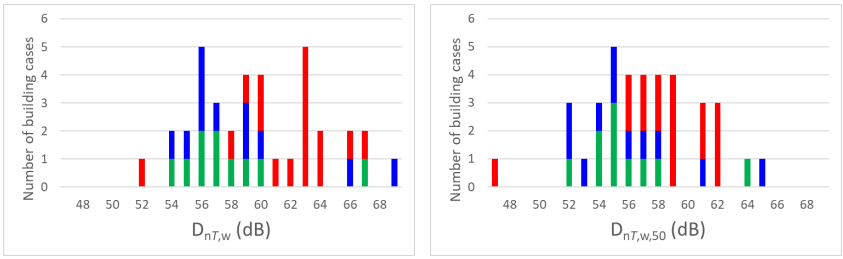


Figure 1. Histograms of the mean airborne sound insulation of the building cases with respect to category: ■concrete, ■CLT and ■lightweight.

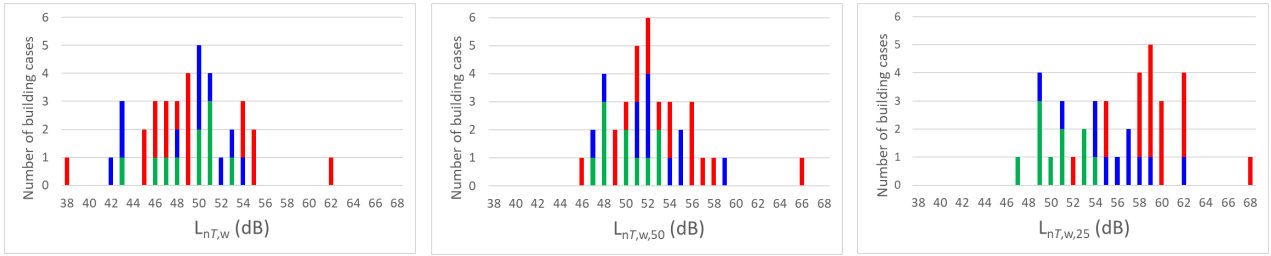


Figure 2. Histograms of the mean impact sound insulation of the building cases with respect to category: ■concrete, ■CLT and ■lightweight.

3 Questionnaire survey

3.1 Method

The questionnaire according to Figure 3 was developed within the European COST action TU0901 [6] based on the technical specification ISO/TS 15666 [7]. The residents have been instructed to rate their perceived annoyance from several potential types of sound and vibration in their home. On a numerical scale ranging from 0–10 where "0" means not at all bothered, disturbed or annoyed and "10" means extremely annoyed. The respondents marked their rating for a total 17 questions.
The questionnaires were distributed by land mail six months or later after occupancy. In total, over 1200 filled questionnaires were returned, which corresponds to an average response rate of about 50%. As a starting point, the mean value from all questionnaires returned from each case was used for the statistical evaluation, but later on, parameters based upon fractions were included.

3.2 Result

Then mean annoyance from the individual sound sources, questions No. 2–14, are presented in Figure 4. The total score, including all buildings as well as the three subcategories, are shown. On average, taking all the sound sources into account, the occupants seem to be fairly satisfied with the acoustical indoor climate. The mean annoyance rating is typically around 1.0–1.5 on the numerical scale ranging from 1 to 10. However, the question regarding footstep noise, No. 5, stands out with an annoyance rating being roughly twice as high for two of the construction categories. The mean annoyance rating of footstep noise from occupants is 3.8 in the lightweight buildings, 3.2, in the CLT buildings and 2.0 in the concrete building. Thus, even though impact

noise is often a severe source of annoyance, this is not the case in the concrete buildings, for which the annoyance is comparable to the other sound sources.

Instructions:
Choose an answer on the 0-to-10 scale for how much noise bothers, disturbs or annoys you when you are in your home.

*if you hear the noise but you are **not at all** disturbed by it, **choose 0***	*if you are extremely bothered, disturbed or annoyed by it, **choose 10***	*if you are somewhere in between, **choose a number from 1 to 9***	*if you do not hear anything at all, the source does not exist or if you cannot answer, **choose "Don't know"***

Thinking about the last 12 months in your home, how much are you bothered, disturbed or annoyed by	Not at all 0	1	2	3	4	5	6	7	8	9	Extremely 10	Don't know
1. Noise from neighbours, technical installations etcetera	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Thinking about the last 12 months in your home, how much are you bothered, disturbed or annoyed by these sources of noise	Not at all 0	1	2	3	4	5	6	7	8	9	Extremely 10	Don't know
2. Neighbours; daily living, e.g. people talking, telephone, radio, TV through the **walls**	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
3. Neighbours; daily living, e.g. people talking, telephone, radio, TV through the **ceilings or floors**	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
4. Neighbours; music with bass and drums	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
5. Neighbours; **footstep** noise, i.e. you hear when they walk on the floor	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
6. Neighbours; **impact or scraping** noise, i.e. from chairs, kitchen sink, lockers, toys, vacuum cleaning etcetera	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
7. Neighbours; rattling or tinkling noise from your **own** furniture when neighbours move on the floor above you	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
8. Stairwells, access balconies; people **talking**, doors closing	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
9. Stairwells, access balconies; **footsteps**, balustrade **impacts**	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
10. Water installations; plumbing, using or flushing WC, shower	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
11. Climate installations; heaters, air condition, air terminals	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
12. Services; elevators, laundry machines, ventilators, heatpumps	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
13. Premises; garages, shops, offices, pubs, restaurants, laundry rooms or other, heard indoors with windows closed	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
14. Traffic (cars, buses, trucks, trains or aircraft); heard indoors with windows closed	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

***Before** moving to your present home, how important to you was **then** the sound insulation, with respect to*	Not at all important 0	1	2	3	4	5	6	7	8	9	Extremely important 10
15. Noise from neighbours, technical installations etcetera	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

*How **sensitive** are you to*	Not at all sensitive 0	1	2	3	4	5	6	7	8	9	Extremely sensitive 10
16. Noise from neighbours, technical installations etcetera	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

*Thinking of **all other aspects** than noise, in your home*	Not at all satisfied 0	1	2	3	4	5	6	7	8	9	Extremely satisfied 10
17. How satisfied are you with your home all in all, overall	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Please comment, describe the types of noise, where the noises come from:

ID:_____

Figure 3: Questionnaire.

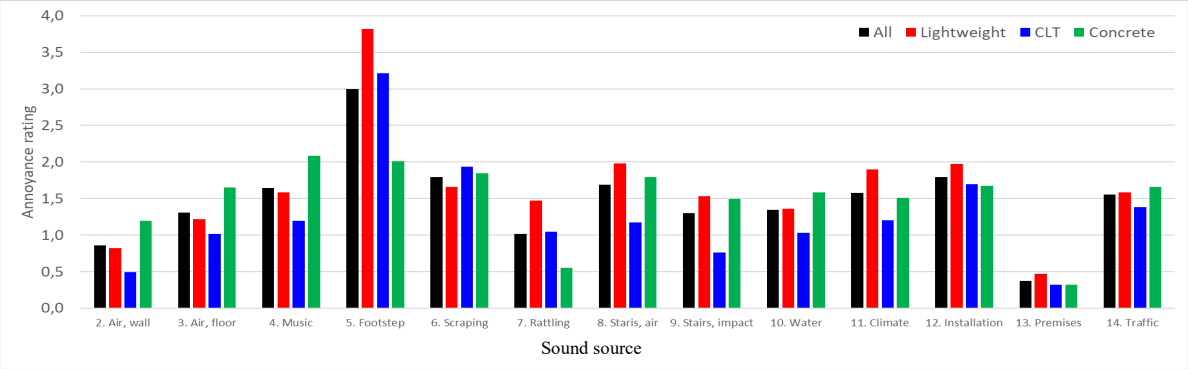


Figure 4. Mean annoyance rating for different building construction types from a variety of sound sources.

The responses from the survey questions related to airborne and impact sound insulations are presented below in the alternative way of using an ordinal scale with five alternatives: not at all, slightly, moderately, very and extremely annoyed. The numerical scale was translated into the ordinal scale according previous research [8–9]. The rating 0 on the numerical scale corresponds to not at all, ratings 1–3 correspond to slightly, 4–6 to moderately, 7–9 to very and 10 to extremely.

Three questions can be directly linked to airborne sound insulation. These include No 2: daily-living sound from the neighbours through the walls, No 3: daily-living sound from the neighbours through the floors or ceilings and No 4: music with bass and drums. As can be seen in Figure 5, music is somewhat more annoying than the other daily-living sounds from neighbours. Taking all cases into account, 17% of the occupants are moderately, very or extremely annoyed by music, while 13% are the corresponding annoyance ratings for other daily-living sounds through floors. The occupants of the concrete buildings are generally slightly more annoyed by airborne sounds than occupants living in lightweight or CLT buildings. The annoyance related to daily-living sounds through walls (question No 2) was found to be lower than through floors.

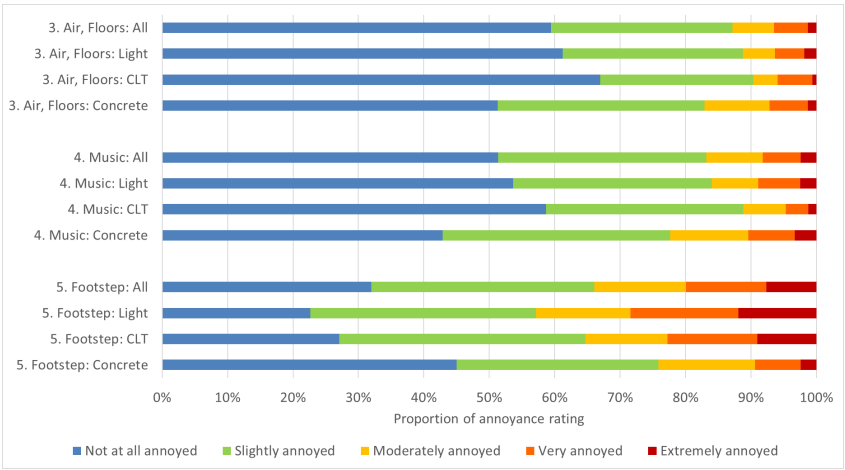


Figure 5. Proportion of annoyance rating from airborne sound sources; 3) daily-living sounds from neighbours through floors, 4) music and from 5) footstep.

Impact sound insulation is directly associated with question No 5: footstep noise from the neighbours. The annoyance difference between the lightweight and concrete categories is considerable. The results show that 2% of the occupants in the concrete buildings, but 12% in the lightweight buildings, are extremely annoyed. In the latter category, 29% are very or extremely annoyed while as many as 77% are annoyed by footstep to some extent, i.e. being at least slightly annoyed.

Two more questions in the questionnaire can be related to impact sound insulation. These concern rattling or tinkling noise from the furniture in the respondent's own apartment, and impact or scraping noise from e.g. chairs, kitchen work and toys. The reported annoyance related to these sources were lower than those from

walking neighbours and are not further discussed here, but analysis at an earlier stage was presented in [10].

4 Single number quantities for impact sound insulation

4.1 Method

Since footstep noise generated the highest rated annoyance, it should be examined whether this could be due to inadequate impact sound insulation descriptors, i.e. single number quantities. The statistical relationship between the rated annoyance from impact sound (question No 5) and various single number quantities derived from the standardized impact sound levels measured in one-third octave bands was investigated using linear regression analyses. The model is expressed by the relation $Y=a+bX$ where Y is the annoyance rating and X is the single number quantity in dB. The rated annoyance is represented by the mean value for each building case. From the statistical model, the coefficient of determination, R^2, (equivalent to the square of the correlation coefficient r) is determined. This way of handling the statistical analyses follows the methods used in the previous analysis of subsets of the complete data now available [1, 10]. The results presented within in this section presume that the sound pressure levels at frequencies below 50 Hz are evaluated without normalization of reverberation time, i.e. they represent L_p rather than L_{nT}. This approach may be justified from a practical point of view, since the procedure to measure the reverberation time at such low frequencies may be cumbersome and induce a great uncertainty in the related single number quantity [11].
There is no common standardized method for statistical evaluation of a subjective dose-response relationship to noise annoyance. One kind of analyses may be based on the mean annoyance according to the European Commission Services' [12]. Another alternative is to define a cut-off point as a basis for the analysis [9, 12], an alternative that divides the respondents into two groups with a specified rated annoyance level – at a given cut-off point. In this case, the fraction ≥ 7 has been used, showing the percentage of respondents who rated their annoyance ≥ 7, corresponding to a 64% cut-off point. Although not included in the paper, complementary analysis using other cut-off points as well as using the mean value instead of fraction were performed, and the results were found to be consistent when assessing the single number quantities.

4.2 Result

The resulting linear regression analyses are presented in Figure 6 including the coefficient of determination, R^2. Note that all symbols representing the building cases, share the same numerical value on the y-axes (the rated annoyance) in all the three figures, while the measured single number quantity on the x-axes varies with respect to frequency range. Evaluating the impact sound insulation by $L_{nT,w}$, results in the coefficient of determination of just 12%. A dramatic improvement is seen if $L_{nT,w,50}$ is used as R^2 increases to 40%. The correlation increases further as lower frequencies are included and evaluation from 25 Hz, $L_{nT,w,25}$, resulted in the strongest correlation, R^2=62%. It should be mentioned that in the foregoing studies [1, 10], the strongest correlation was instead found with evaluation from 20 Hz, which is not the case here after that the database has been extended. Evaluated from 20 Hz, $L_{nT,w,20}$, R^2 reached 49% (not covered in the figure).

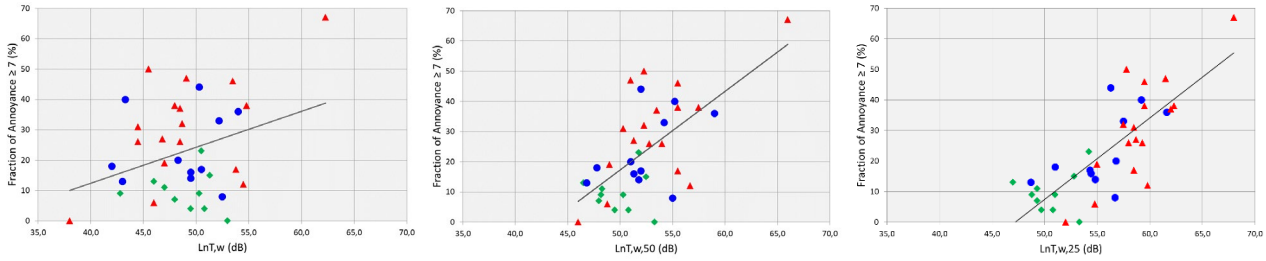


Figure 6. Linear regression of annoyance from footstep vs. the single number quantities $L_{nT,w}$, $L_{nT,w,50}$ and $L_{nT,w,25}$ with corresponding R^2 of 11, 40 and 62% respectively. Building types: ▲ lightweigth, ● CLT and ◆ concrete. No normalization of impact levels at 20-40 Hz to the reverberation time 0.5 s was applied.

5 Single number quantity limits to be considered

5.1 Method

The questionnaire survey is assumed to reflect the overall perspective of the occupants' satisfaction/annoyance regarding the acoustical comfort in modern apartments. But since the sound insulation varies considerably between building cases, variation in the acoustical satisfaction as well as in the rated annoyance is expected as a consequence. The relation between sound insulation and rated annoyance according to the ordinal scale is studied by logistic regression analysis. This is the probability that a given single number quantity of sound insulation will lead to annoyance within one or more of the ordinal categories. Mathematically, this is described by Equation 3:

$$Prob(Annoyance) = \frac{e^{k \cdot SNQ + m}}{1 + e^{k \cdot SNQ + m}}, \tag{3}$$

where k and m have their counterparts in the coefficients of the linear regression $Y(x) = kx + m$ and SNQ refers to the dB-value of the single number quantity. Logistic regression analysis is applied to airborne and impact sound insulation separately. Multiple single number quantities and their respective probability to result in annoyance are studied for each specific survey question.

5.2 Result

The rated annoyance following the survey questions of airborne sound insulation is relatively low when $D_{nT,w,50}$ fulfils the Swedish minimum requirement of 52 dB. For daily-living sound sources and for music, 10% and 12% of the occupants respectively are expected to rate those sources as very or extremely annoying, and 21% and 26% respectively as at least moderately annoying, see Figure 7.

Figure 5 shows that 13% are moderately, very or extremely annoyed by daily-living sounds through floors and ceilings. The corresponding proportion for daily-living sounds through walls is 7% (not covered in the figure) and for music 17%. Overall, it is suggested that the current Swedish legislation, $D_{nT,w,50} \geq 52$ dB, works satisfactorily and offers adequate airborne sound protection to the occupants.

As shown, regarding the correlation between measured impact sound insulation and rated annoyance, $L_{nT,w,25}$ gives a stronger correlation than $L_{nT,w,50}$. According to the previous sections, despite fulfilling the present Swedish impact sound insulation class A ($L_{nT,w,50} = 48$ dB), the annoyance may still be greater than for airborne sound insulation that only meets the minimum requirement ($D_{nT,w,50} = 52$ dB). This indicates the need for a more restrictive requirement that offers a higher degree of protection against unwanted impact sounds, and for this purpose $L_{nT,w,25}$ is preferred.

The logistic regression analyses involving $L_{nT,w,50}$ and $L_{nT,w,25}$ in Figure 8 are compared. For $L_{nT,w,50} = 56$ dB, the Swedish minimum requirement, 32% of the occupants are expected to rate the annoyance as very or extremely annoying and 44% as at least moderately annoying. This level of annoyance corresponds to $L_{nT,w,25} = 59$ dB.

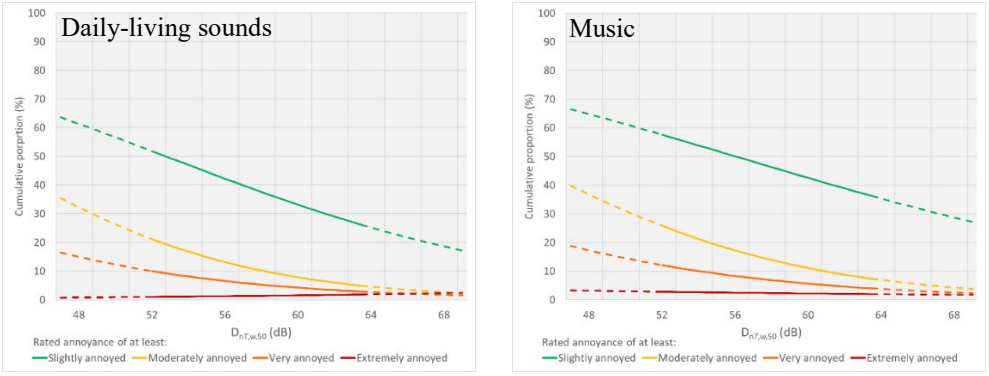


Figure 7. Cumulative proportion of rated annoyance due to daily-living sounds through floors (left) and music (right) as a function of $D_{nT,w,50}$. Solid lines represent 95% of the available data.

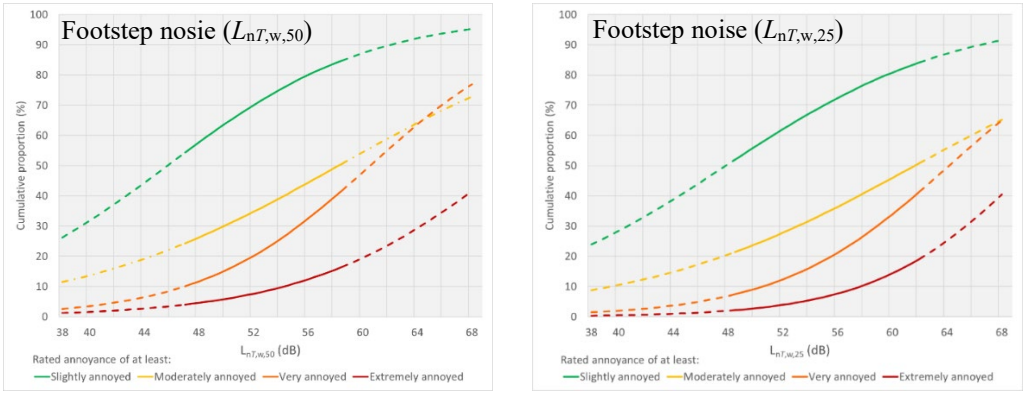


Figure 8. Cumulative proportion of rated annoyance due to footstep noise as a function of $L_{nT,w,50}$ (right) and $L_{nT,w,25}$ (left). Solid lines represent 95% of the available data.

If instead $L_{nT,w,25}$ is set to 56 dB, i.e. the same numerical value as today's Swedish requirement for $L_{nT,w,50}$, this expects to generate the proportion of 21% being very or extremely annoyed, and 36% being a least moderately annoyed.
Another further step further is to find a level of expected annoyance from impact noise that equals the annoyance from airborne daily-living sounds. Then the annoyance should match 21% and 10% being at least moderately, and very or extremely annoyed, occupants respectively. This level of annoyance corresponds to $L_{nT,w,25}$ being as low as about 50 dB. Even though arguments exist, such a dramatic shift compared to today's standard would probably be too difficult for the building industry to handle.

6 Floor design

It has been shown that wood-based floors show high impact sound levels at low frequencies, and thereby high annoyance rating among the residents. The issue is then how to design the floors to decrease the impact sound level. As the mass is a key factor, a solution could be to use a screeded floor, e.g. a layer of concrete on top of the timber construction. Although such a hybrid solution may be effective from a sound insulation perspective, the building industry search for solutions free from cement-based products.
Two ideas were tested in the *AkuTimber* project: to increase the density of a CLT floor and to increase the damping.

6.1 Method

Spruce, with a density of 400-440 kg/m^3 is the dominating wood species used for CLT panels in the Scandinavian market. As an alternative, panels made of birch having a density of 600-650 kg/ m^3 is tested. Another tested configuration is to use densified spruce. The planks are then compressed to ~55% of their original height before the CLT panels are manufactured. Such a panel have a density of about 1.8 times the density of a plate made from untreated spruce.
To increase the damping, a viscoelastic sheet [13] of various thickness, 2-12 mm, was put in between two 60 mm CLT panels. The viscoelastic material is designed to work by shear and to be efficient at low frequencies. Panels of the size 0.6x4.0 m and/or 0.6x2.4 m, were tested for the mobility averaged over the surface and modal analysis was performed to check for natural frequencies and damping ratios. The tested plate was hooked up by resilient ropes and was excited by an electromagnetic shaker. The response was measured by accelerometers in 27 positions (3x9) over the surface. All individual panels were 60 mm thick but were also tested in pairs, 2x60 mm, either glued together or using viscoelastic foil in between.

6.2 Result

The fundamental frequency of the 4.0 m panels was ~20 Hz for the 60 mm panels made of spruce or birch and ~32 Hz when two panels were glued together, 2x60 mm (no panels of densified spruce was available in 4.0 m length). Fir the shorter 2.4 m panels, 60 mm, the fundamental frequency was found to be 52 Hz for spruce, 50 Hz for birch and 44 Hz for densified spruce. With 2x60 mm, the fundamental frequency was measured to 87 Hz and 80 Hz for the original and densified spruce respectively. See Figure 9 for natural frequencies and mode shapes found in a 60 mm spruce CLT panel of 4.0 m length.

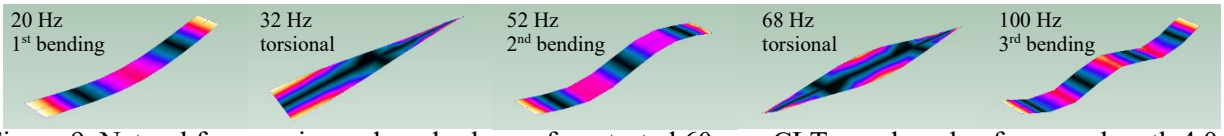


Figure 9. Natural frequencies and mode shapes for a tested 60 mm CLT panel made of spruce, length 4.0 m.

In Figure 10a), the surface average of the mobility for three different 2x60 mm panel configurations, glued together and 2.4 m in length, are presented. As expected from the mass relation, the panel made of densified spruce shows the lowest mobility level followed by the birch panel.
In Figure 10b), 2x60 mm panels are used to study the effect of the viscoelastic sheet of different thicknesses compared to the referenced case were two panels are glued together. The thinnest viscoelastic layer, 2 mm, has hardly any effect while thicker layer, 4-12 mm, gradually lower the mobility as the thickness increases.

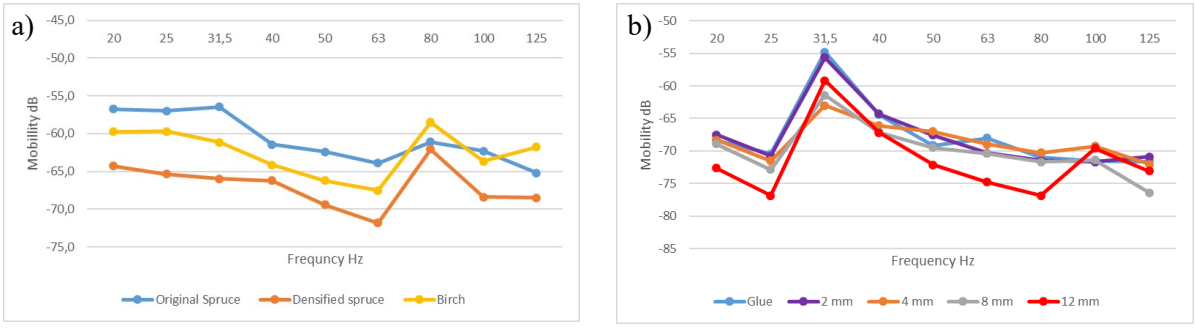


Figure 10. Mobility, averaged over the surface, with respect to a) different kinds of wood and b) viscoelastic layers of different thickness.

7 Conclusions

The importance of impact sound insulation, below 50 Hz has been actualised. The evaluation methods used today rely upon a frequency range starting from 100 or 50 Hz. At least the range starting from 50 Hz may be appropriate dealing with heavy constructions but for lightweight constructions, the correlation between measurement and residents' perception in terms of rated annoyance is weak. Among the 38 building cases of mixed constructions, the coefficient of determination between rated annoyance and measurement was just 11% when the impact sound insulation was evaluated from 100 Hz ($L_{nT,w}$), 40% evaluating from 50 Hz ($L_{nT,w,50}$), but 62% evaluating from 25 Hz ($L_{nT,w,25}$).
Based upon the results, evaluating impact sound from 25 Hz, by using $L_{nT,w,25}$, is better than using $L_{nT,w}$ or $L_{nT,w,50}$ in markets where both heavyweight and lightweight building constructions occur. However, using $L_{nT,w,25}$ means that sound energy over more third-octave bands are summed compared to $L_{nT,w,50}$, and any numerical single number quantity limit should therefore not be used without adjustment when changing from $L_{nT,w,50}$ to $L_{nT,w,25}$. Besides, when comparing the estimated annoyance at the levels of the present Swedish minimum requirements, $D_{nT,w,50} \geq 52$ dB and $L_{nT,w,50} \leq 56$ dB, it was found that footstep noise is expected to generate higher level of annoyance than airborne daily-living sounds and music. From this perspective, a somewhat stricter impact sound insulation requirement makes sense.

Lightweight constructions often show poor low-frequency impact sound insulation compared to heavy constructions. This is to high extent related to the mass difference, and it is a great challenge to design a wood-based floor with great sound insulation at low frequencies. A couple of potential solutions were tested experimentally and compared to a standard CLT panel made of spruce. First, the panel´s mass was increased, either by using birch, an alternative type of wood with about 50% greater density, or by using densified – compressed – spruce with 80% greater density than the original. Secondly, a layer of viscoelastic damping material was put in between two CLT panels to form a new unit. Both methods affected the measured mobility which was found to be up to 8 dB lower in single one-third octave bands below 100 Hz. The techniques do have potential, and further investigations – including a larger specimen (12 m^2) to be tested in a test house – are planned.

Acknowledgements

The author gratefully acknowledge the financial support by the Swedish research council *Formas* and the *Swedish Energy Agency*.

References

[1] Fredrik Ljunggren, Christian Simmons, Klas Hagberg. Correlation between sound insulation and occupants' perception – Proposal of alternative single number rating of impact sound. *Appl. Acoust.* 85: 57–68, 2014.

[2] ISO 16283-1:2014. Acoustics – Field measurement of sound insulation in buildings and of building elements – Part 1: Airborne sound insulation.

[3] ISO 16283-2:2020. Acoustics – Field measurement of sound insulation in buildings and of building elements – Part 2: Impact sound insulation.

[4] ISO 717-1:2020. Acoustics – Rating of sound insulation in buildings and of building elements – Part 1: Airborne sound insulation.

[5] ISO 717-2:2020. Acoustics – Rating of sound insulation in buildings and of building elements – Part 2: Impact sound insulation.

[6] COST TU0901. Integrating and harmonizing sound insulation aspects in sustainable urban housing constructions. http://www.costtu0901.eu, 2022-03-18.

[7] ISO/TS 15666:2003. Acoustics – Assessment of noise annoyance by means of social and socio-acoustic surveys.

[8] Bernd Rohrmann. The use of verbal labels in noise annoyance scales. Theoretical deliberations and empirical findings. *Proceedings of Noise as a Public Health Problem*, Sydney, 22–26 November 1998

[9] Mark Brink, Dirk Schreckenger, Danielle Vienneau, Christian Cajochen, Jean-Marc Wunderli, Nicole Probst-Hensch, Martin Röösli. Effects of scale, question location, order of response alternatives, and season on self-reported noise annoyance using ICBEN scales – A field experiment. *Int. J. Environ. Res. Public Health* 13: 1163, 2016

[10] Fredrik Ljunggren, Christian Simmons, Rikard Öqvist. Correlation between sound insulation and occupants' perception – Proposal of alternative single number rating of impact sound, part II. *Appl. Acoust.* 123: 143–151, 2017.

[11] Fredrik Ljunggren, Rikard Öqvist, Christian Simmons. Uncertainty of in situ low frequency reverberation time measurements from 20 Hz – An empirical study. Noise Control Eng. J. 64(6): 707–715, 2017.

[12] European Communities. Position paper on dose response relationships between transportation noise and annoyance. ISBN 92-894-3894-0, 2002.

[13] Swedac AB. http://www.swedac-acoustic.com/products/stomljudsdampande/viskoelastisk-folie-vs/, 2022-03-18

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

A light-weight wireless omni-directional sound source for room acoustic measurements

Claus Lynge Christensen[1,*], **George Koutsouris**[1], **Antoine Richard** [1], **Jens Holger Rindel** [1], **Andreas Kornelius Nørgaard**[1]

[1]Odeon A/S, DTU Science Park, Lyngby, Denmark.

[*]clc@odeon.dk

Abstract

One of the biggest challenges in room acoustic measurements, according to the ISO 3382 standard (parts 1-3), is to use sources that maintain omnidirectional directivity (with small deviations) across the frequency range, from the 125 Hz to 4 kHz octave bands. Conventional omnidirectional sources can achieve adequately omni-directional directivity with the use of more than one loudspeaker unit (e.g. twelve units in a dodecahedron). However, these sources tend to have poor performance at low frequencies, because of the restricted volume shared by the units, as well as non-uniform directivity at high frequencies. In addition, their size makes them far from ideal point sources and discourages acousticians from performing measurements more often, due to their reduced portability. In this paper, a novel sound source is presented, which solves the aforementioned problems. The source consists of a single loudspeaker unit, which keeps the size small and close to an ideal point source. The loudspeaker unit is backed by a cavity with sufficient volume for efficient low-frequency performance. In front of the loudspeaker unit is an acoustic lens with a doughnut-like shape, designed in such a way that the high-frequency radiation is close to omni-directional. This acoustic lens contains all electronics required to make the sound source fully portable and wireless. Free-field measurements of the sound source show that the frequency response is fairly flat over a wide frequency range and that the directivity pattern is well within the limits of ISO 3382, also for frequencies outside the default range (63 Hz and 8 kHz octave bands). An application example is presented, showing that the source is well-suited for room acoustic measurements in small to medium-sized rooms. In addition, the source can be used for measuring the room frequency response, e.g. from 20 Hz to 200 Hz, when positioned in a corner.

Keywords: room acoustics, measurements, loudspeakers, omnidirectional, calibration.

1 Introduction

In general, room acoustic measurements refer to the process of measuring a series of objective parameters that characterize the acoustics of a variety of spaces, such as offices, classrooms, auditoria, concert halls etc. ISO 3382 provides the guidelines for room acoustic measurements in performance spaces, ordinary rooms and in open-plan offices [1]. Such guidelines include the definitions of room acoustic parameters (for example the variations of reverberation times T_{15}, T_{20}, T_{30}; clarity, C_{80}, gravity time, T_s) the positioning of sound sources and microphones in space, as well as the requirements for the source. Specifically for the source, the following requirements must be fulfilled:

1. The source must be able to produce at least 45 dB above the background noise level at the microphone position, to ensure a sufficient impulse response decay range. If T_{20} is to be measured, it is sufficient to use 35 dB above the background noise level.

2. The source must be as omnidirectional as possible. To evaluate omnidirectionality, the source is driven with pink noise in free-field conditions and the level is measured at a fixed radius every 5° for a full 360° circle. This is done on the plane with the greatest asymmetry. From these measurements a *reference level* is calculated as an energetic average. Then, a 30° arc moving average is calculated and subtracted from the reference level. Finally, the difference is compared with the maximum acceptable deviation from Table 1. ISO 3382 covers the range between the 125 Hz and 4 kHz bands. However, in several situations an extended range from 63 Hz to 8 kHz is more relevant, especially in musical applications or auralisations of measured impulse responses. For these extra bands, the values from 125 Hz and 4 kHz are used.

Table 1: Maximum deviation of source directivity in dB per octave band when pink noise is used for measurement in free field (anechoic chamber). The bands from 125 Hz to 4 kHz are provided by ISO 3382 [1], while the 63 Hz and 8 kHz bands are extrapolated.

Frequency (Hz)	63	**125**	**250**	**500**	**1000**	**2000**	**4000**	8000
Maximum deviation (dB)	± 1	± 1	± 1	± 1	± 3	± 5	± 6	± 6

ISO 3382 does not specify the size or the actual shape of the omnidirectional source. However, it is desired that the size is as small as possible and close to that of an ideal point source, which makes it consistent with the assumptions used in room acoustic theoretical models [2]. Therefore, using an omnidirectional source that is close to a point makes it easier to compare measurements with simulations. On top of that, a small, lightweight source encourages acousticians to make more measurements due to the convenience provided.

Traditional omnidirectional sources involve many loudspeaker units that point towards different angles in order to ensure sufficiently uniform sound radiation. A typical example is a *dodecahedron* source, which consists of 12 faces with one loudspeaker unit mounted on each of them, as the models manufactured by NorSonic [3] and Brüel Kjær [4]. Although these type of sources fulfil the directivity requirements of Table 1, they tend to be large in size (usually requiring external amplifiers) which makes them far from ideal point sources and harder to transport. There are other types of sources, that use fewer units and have a much smaller volume. The model *Qoms2* by Qsources [5] uses two loudspeaker units in a vertical configuration, facing each other. However, the need of external amplification and cabling make these sources less portable. Recently, a low-cost alternative method has been suggested, by averaging the response from three common directional loudspeakers [6]. This can be useful when it is possible to make use of existing loudspeakers in a room, instead of bringing a separate measuring source.

In this paper, a light-weight wireless omnidirectional source is presented. The source is called *OdeonOmni* and has been designed and developed by Odeon A/S, specifically for room acoustic measurements with the sine-sweep method [7]. It consists of a single loudspeaker unit (driver), which keeps the size small and close to an ideal point source. The device was developed primarily for portability and ease of handling, to encourage acousticians to perform more room acoustic measurements. Therefore the device is wireless and includes a built-in amplifier. The project was initiated in 2015 and was also part of Danish Tech Challenge 2016.[1]

2 Description of *OdeonOmni* source

Figure 1 shows a drawing of the *OdeonOmni* source. It has an overall height of 35 cm, diameter of 14 cm and a total weight of 1.5 kg. The device consists of two main elements:

[1] https://dtusciencepark.dk/futurebox/danish-tech-challenge/

1. A loudspeaker cabinet with sufficiently large and damped volume for low-frequency performance. On top of the cabinet, a single full-range loudspeaker driver is mounted.
2. An acoustic doughnut-like shaped lens, mounted at approximately 1.5cm above the loudspeaker driver (measured from the lowest part of the lens to the centre of the driver). Its shape has been optimized for omnidirectional high-frequency radiation, according to Table 1. This acoustic lens contains all electronic components required to make the sound source fully portable and wireless. Mainly, a *digital amplifier*, a *Bluetooth module* and *rechargeable batteries*.

The source is mounted on a tripod at vertical position. Henceforth, we will consider the top side as 0° and the bottom side as 180°. Audio signal is transferred to the source by wireless Bluetooth connection or by a 3.5 mm jack audio cable. The source can be used with any measuring software, such as the module included in ODEON [8].

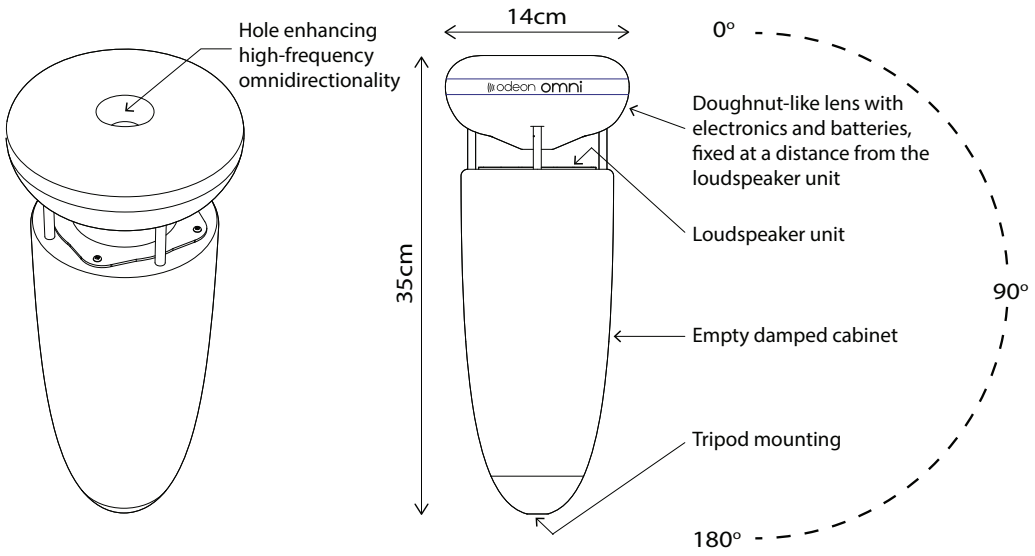


Figure 1: Sketch of the *OdeonOmni* source. The main elements include the cabinet, as well as an acoustic lens that encompasses all electronics and batteries for the operation of the device, and helps radiating sound evenly.

3 Specifications and design of *OdeonOmni* source

The design of *OdeonOmni* source has been determined following three main goals: 1) Sufficient omnidirectionality, according to the directivity limits provided in Table 1. 2) As flat as possible frequency response. 3) High portability. The final design has been a result of many iterations both in terms of calculations and in terms of building and measuring a series of prototypes. The first prototypes were 3D-printed in plastic material and used external amplification via cable connection. The two separate parts of the device, the *cabinet* and the *acoustic lens*, made it possible to accelerate the prototype evaluation process, by testing combinations of them (eg. cabinet v.1 combined with lens v.2 and so on). In the initial stage, simple analytic formulas were used to determine basic loudspeaker parameters, such as the volume of the cabinet and the resonance frequency of the system [9]. The resonance frequency was required to be low, while the overall frequency response was required to be as flat as possible. Therefore, it was decided that *OdeonOmni* should be a closed system, as opposing to vented (open) systems that can have lower resonance frequency but very steep roll-off below that. Vented systems also require larger cabinet volume, which was not desired in terms of omnidirectionality and portability.

To keep the source portable and lightweight enough, the cabinet volume had to remain below 2 L. At the same time a low system resonance frequency was desired, so that the device could be efficient at the lower octave

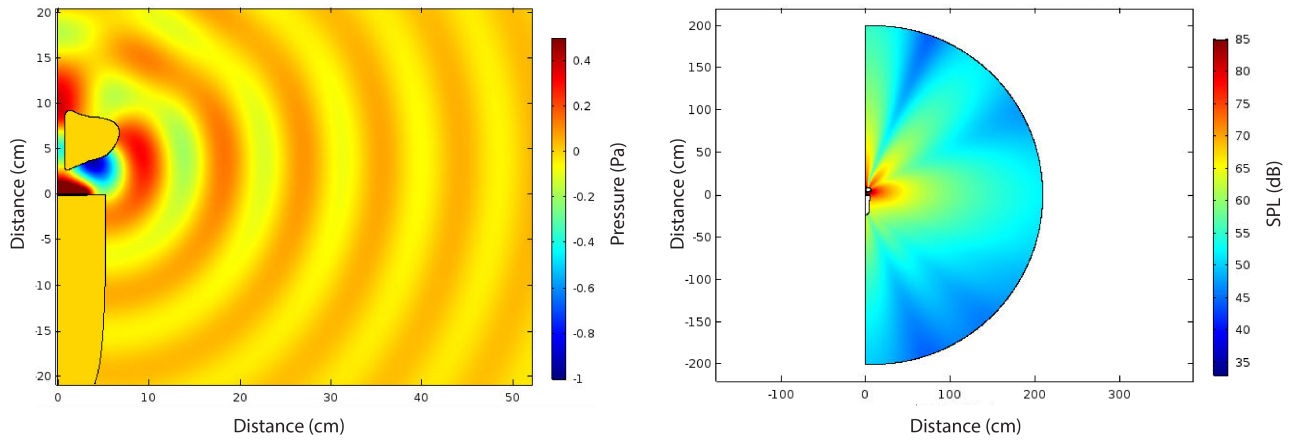
bands. For these reasons, a loudspeaker driver of compliance equivalent volume $V_{as} = 1.4$ L and resonance frequency $f_s = 74$ Hz was chosen. The driver has a high excursion capability X_{max} of 4.6 mm and a total quality factor Q_{TS} of 0.41. Choosing a system total quality factor Q_{TC} of 0.547, resulted to an active cabinet volume of approximately 1.8 L and a system resonance frequency $f_c \approx 99$ Hz. Using damping material inside the cabinet helps increasing the active volume by approximately 15%. Therefore only a physical volume of $1.8/(1 + 15\%) = 1.56$ L was necessary in the end.

3.1. Omnidirectionality

After fixing the cabinet volume, different geometries were tested to comply with the directivity limits of Table 1. Any plain loudspeaker system - with a cabinet and a driver on one of the sides - is highly directive as frequency increases. Therefore, to achieve sufficient omnidirectionality for *OdeonOmni* the main concept was to use an acoustic lens in front of the driver, that could reduce some radiation at high frequencies and reflect some sound towards the back of the cabinet. In this way a more uniform radiation was obtained.

Initially, a commercial Finite Element Method (FEM) software was used by B.Eng candidate A. K. Nørgaard [10] to simulate the directivity patterns obtained using different designs of the cabinet and lens, as well as the distance between them. The overall shape of the device was decided to be cylindrical at a very early stage. This made the directivity problem axially symmetric, therefore only the vertical plane had to be studied.

From numerous different designs, a long conical shape was chosen for the cabinet and a doughnut-like shape for the lens. These shapes provided the most uniform radiation on the vertical plane, while specifically **a central hole** in the lens offered better radiation at the high frequencies (4 and 8 kHz). Without the hole, the directivity shifted to the other side of the loudspeaker driver, as most of the sound was reflected towards the bottom. The hole helped mitigating that problem by letting some sound being radiating naturally in front of the driver. Figure 2 shows a screenshot of 2D FEM simulations. Radiation is uniform enough at 4 kHz, while the sound wavefront approaches that of an ideal point source at 0.5 m distance.



(a) Acoustic pressure up to 50 cm from source.

(b) Sound pressure level up to 2m from source.

Figure 2: A 2D FEM simulation with the cabinet and lens at 4 kHz, on the vertical plane, where directivity problems are more critical.

The FEM simulations provided only a basic indication of the actual cabinet and lens shape. After that, a *trial and error* iteration process took place using real prototypes and directivity measurements, according to Sec. 1, at the Technical University of Denmark. In these prototypes, several manufacturing issues (e.g. thickness of material, placement of electronics) were taken into account as well. At least 10 prototypes were 3D-printed, measured and improved, each of them with several distance variations between cabinet and lens. In Figure 3 the directivity plots of the final shape (Figure 1) are shown. Due to axial symmetry of the source (when placed in normal vertical position) only the part from 0° to 180° is displayed.

The majority of the directivity plots show that the directivity curve is well within the ISO limits. The most critical band in the directivity plots is at 500 Hz, where the directivity curve approaches - but does not exceed - the ISO limits at 90°. This is partly due to the fact that the limits have been set tight for 500 Hz (± 1 dB), while they become much broader at 1 kHz (± 3 dB). It is interesting to observe that with the current ISO 3382 limits, the *OdeonOmni* source performs very well at the highest bands (4 and 8 kHz). This is generally very difficult to achieve with a traditional dodecahedron and most of this behaviour comes from the actual design of the acoustic lens that evens out the highly directive components.

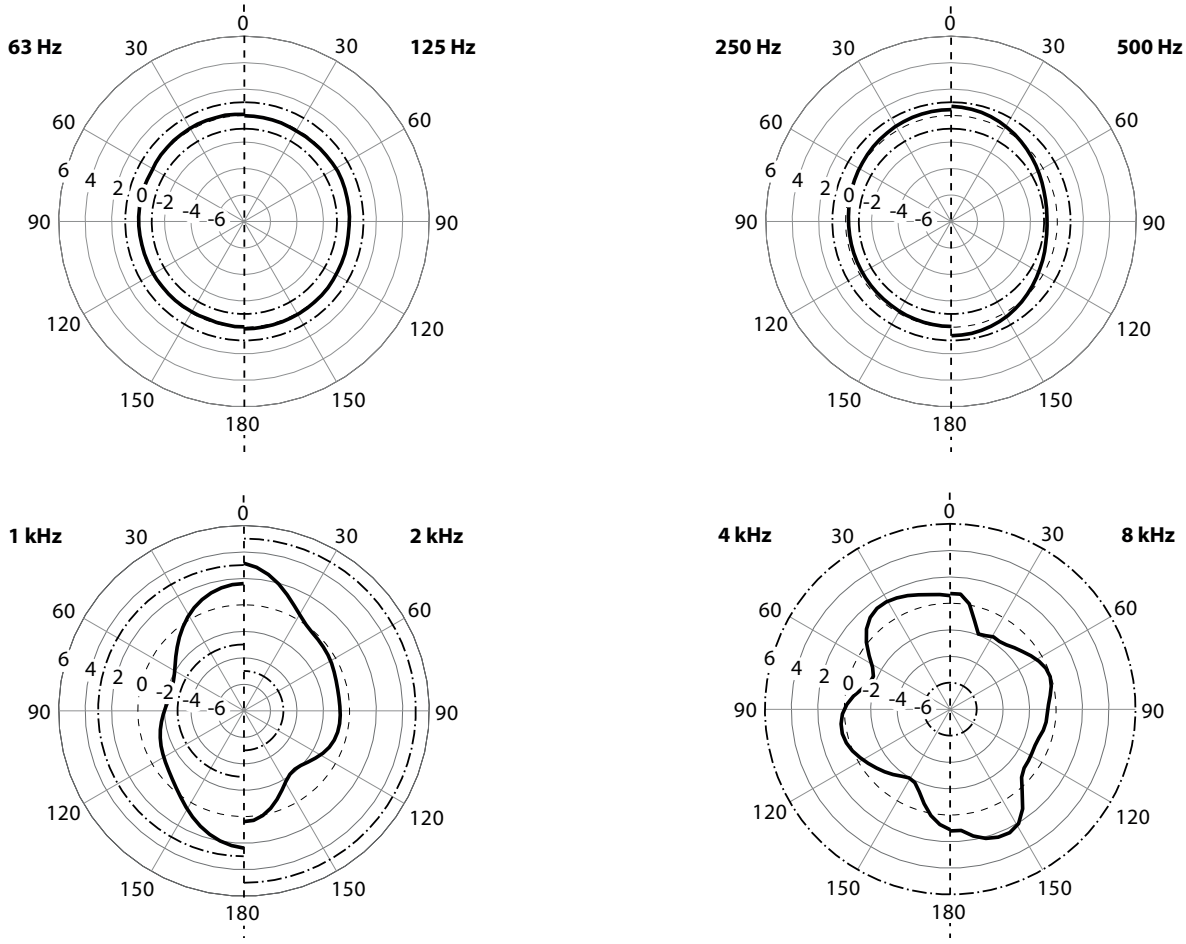


Figure 3: Directivity plots of *OdeonOmni* across the vertical plane for each octave band between 63 Hz and 8 kHz. Since the geometry is axisymmetric, only the part from 0° to 180° is displayed. ISO limits (-·-·). Directivity curve (—). 0 dB reference (- - -).

Using the minimum and maximum values from Figure 3, the largest deviations, together with the ISO 3382 limits can be plotted as a function of frequency (octave band), as shown in the simplified directivity plot of Figure 4. An interesting observation is that the current ISO limits between 2 and 4 kHz are very wide in comparison to the capability of *OdeonOmni*. This shows that these limits could be reduced in the future to encourage the manufacturing of better sound sources in terms of omnidirectionality.

3.2. Flat frequency response and sound power output

Although there is no specific requirement by the ISO 3382 standard, *OdeonOmni* was designed to provide as flat as possible frequency response between 125 Hz and 4 kHz. The main reason is that a flat frequency response provides an even weighting of all modal reverberation times within an octave band. Besides that, having enough

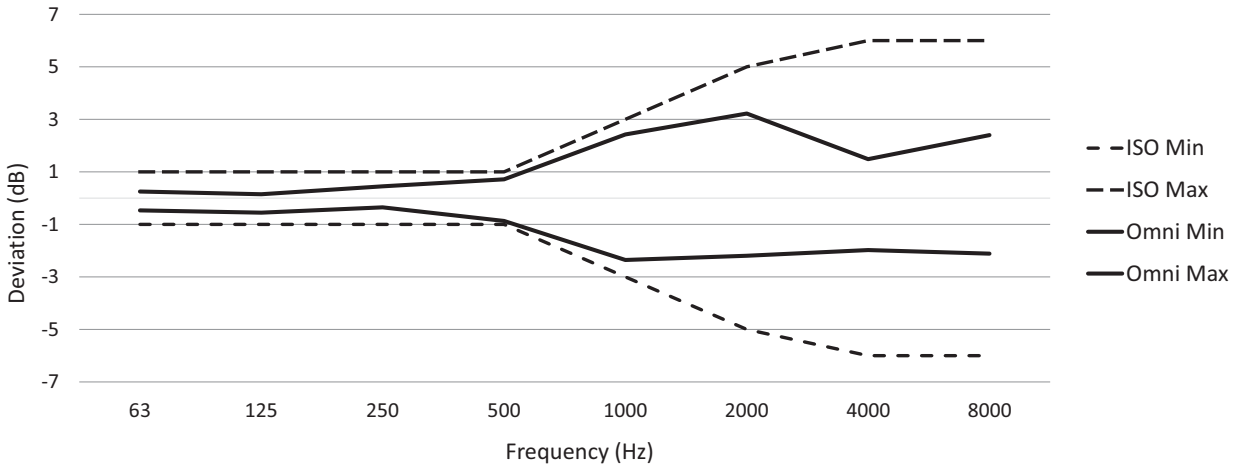


Figure 4: Simplified directivity plot showing the largest deviation from the reference level (0 dB) in each octave band. The plot uses the min and max values from each polar plot in Figure3.

power at 63 Hz and 8 kHz was considered an advantage for auralisations using measured impulse responses. For these reasons it was preferred to use only one high-quality full range loudspeaker driver that could work with the entire volume of the cabinet and provide good performance at low frequencies for its size. In contrast, dodecahedron sources use 12 drivers which share a common cabinet volume, resulting to a steeper decline of the frequency response at low frequencies. They also introduce narrow lobes at higher frequencies, which makes the response less uniform around the source. This results to certain directions having poor performance at the highest octave bands.

The drawback of using a single driver is that the overall sound power output is lower than when using multiple drivers. The maximum power level of the *OdeonOmni* source was measured inside a reverberation chamber at DTU, according to ISO 3741 [11]. The source was connected wirelessly to a laptop, via the available Aptx Bluetooth connection, meaning that the internal sound card and amplifier of the source were used. In such a case the output level of the source can be controlled digitally by the laptop's volume slider, which was set to max. Finally, the source was driven with pink noise of max amplitude 0.9 to ensure that distortion remained at low levels. The results are shown in Figure 5 and offer a good picture of the frequency response as well. The power levels vary approximately between 79 dB and 87 dB, from 80 Hz to 4 kHz, with a small roll-off at the lowest bands. This corresponds to a 8 dB dynamic range and it is a way to quantify the flatness of the response. Typical ranges for dodecahedron sources in the market are about 15 dB, which is much larger than the range of *OdeonOmni*. The last column shows that the total power level from all octave bands is 95 dB, which can be considered a limitation of the source for measuring large rooms, compared to certain dodecahedron sources that can output a total power level of 120 dB. However, due to its relative flat frequency response, the overall power of *OdeonOmni* is evenly distributed at all frequencies, especially at lower ones, where it is typically more difficult to achieve sufficient signal-to-noise ratio. Therefore, a higher overall power output might not be as crucial for *OdeonOmni*, as it is for other sources with less flat frequency response. On top of that, the source is designed for measurements with sweep signals. This offers the advantage of increasing the signal-to-noise ratio by increasing the length of the sweep.

3.3. Portability

The final design consideration of *OdeonOmni* was the high level of portability, with low weight and small size. Several lightweight materials were taken into account at the initial stages of the design process. These included: 1) Aluminum for the entire cabinet and lens, which would require an extensive turning machining process and significant waste of material. 2) Injection molded plastic for the cabinet and lens, which was also rejected

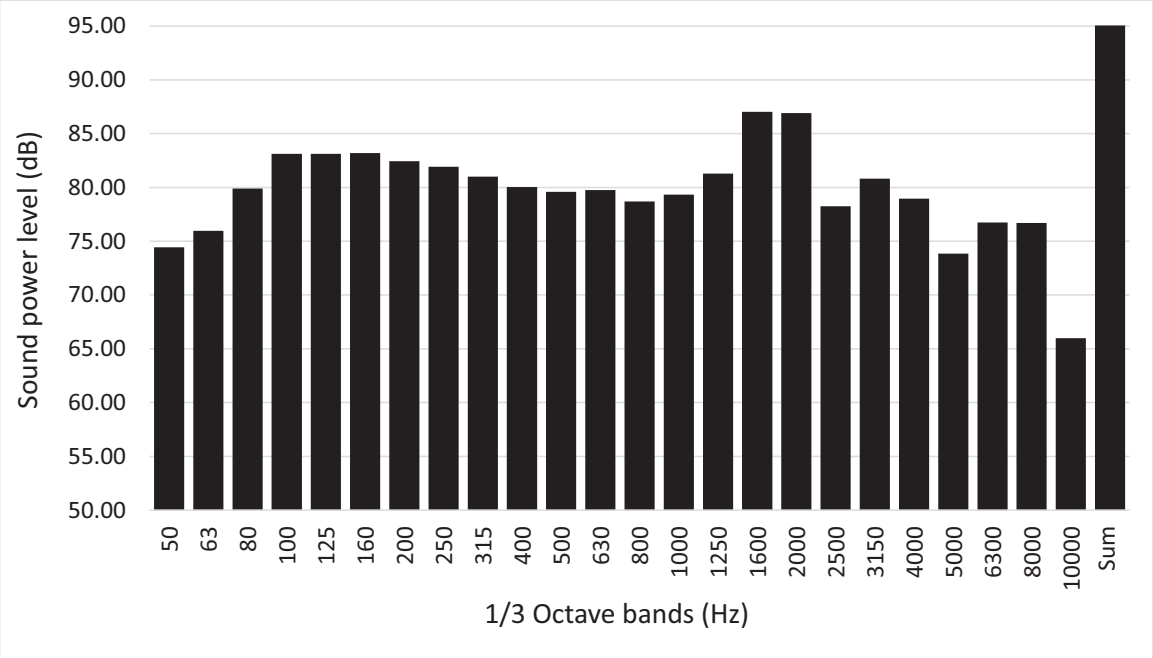


Figure 5: Max sound power of *OdeonOmni* per 1/3 octave. Last column: broadband sum.

because of high cost for the limited number of pieces that were planned in the beginning. 3) Carbon-fiber or glass-fiber composite, which allows a structurally strong, lightweight construction that is cost-effective for a limited number of pieces. The final choice was a coated version of glass-fiber composite.

All functional components were finally placed inside the so-far unused volume in the acoustic lens, in order to leave the loudspeaker cabinet as empty as possible for optimum performance. To enhance portability, the source is connected wirelessly using Bluetooth 5th generation, with *AptX* technology, that ensures low distortion and high resolution compared to standard Bluetooth. An antenna inside the device can maintain a robust connection up to **30m** without obstacles, and up to **10m** with obstacles.

4 Real-case measurements in a cafeteria

The *OdeonOmni* source was tested inside the cafeteria at DTU Science Park (Lyngby, Denmark) together with a traditional Dodecahedron source, built at the Acoustic Technology Group at DTU. This source operated with an external amplifier. The cafeteria has several coupled spaces, a dining area and a large hallway. The overall acoustics is quite challenging because of non-uniform absorption and unstable background noise from ventilation, refrigerators and other machines. For the test measurements, a Zoom H3-VR ambisonic microphone has been used, together with a laptop using ODEON's measuring system [8].

Figure 6 shows a comparison of the cafeteria frequency responses, measured with *OdeonOmni* and the dodecahedron source at the same microphone-source position. The distance is 3 m between them (position S1 in Figure 7), which is considered close enough to reduce the influence from the room. Although the measured responses include both the response from each source plus the response from the room, they can still be used for comparison of the two sources since all other factors remain fixed (room, microphone-source positions). The responses in Figure 6 have been normalized to the peak value of each source. It can be seen that *OdeonOmni* has better performance both at low and high frequencies, as well as flatter overall response.

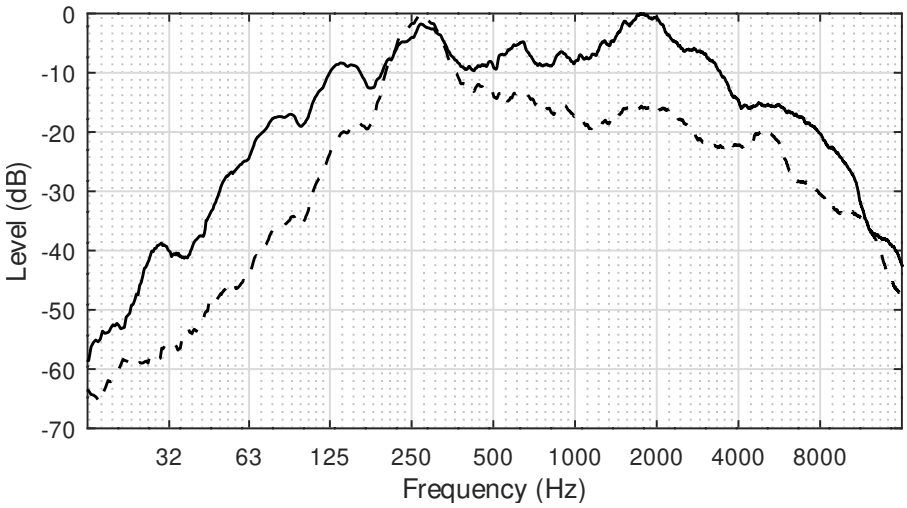


Figure 6: Measured frequency response of the same room with *OdeonOmni* (—) and the Dodecahedron source (- - -), normalized to peak value. The source is at 3m from the receiver.

4.1. Comparison of Room Acoustic Parameters

A line of impulse response measurements was used inside the lunch area of the cafeteria, as shown in Figure 7. It was decided to fix the microphone position at one end and move the source away from the microphone at 5 discrete positions with a total distance of 19 m. From the measurements the ISO 3382 room acoustic parameters [1] were derived in 8 octave bands, from 63 Hz to 8 kHz.

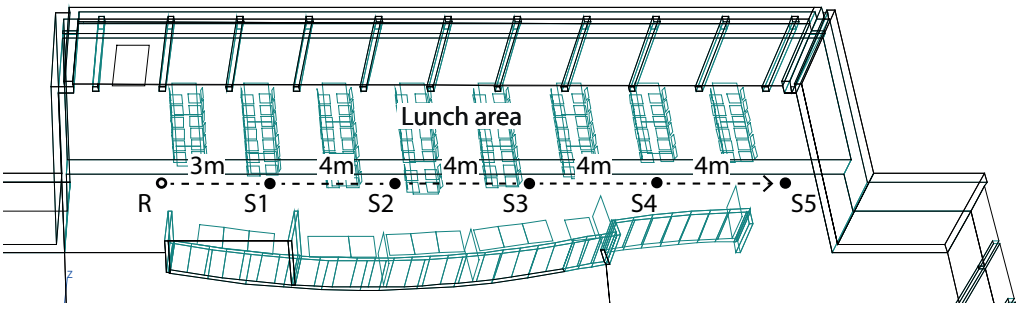


Figure 7: A sketch of the cafeteria at DTU science park used for comparing the performance of *OdeonOmni* with a Dodecahedron source. The receiver (R) position was fixed, while the sources were moved at five locations, with a total distance of 19 m.

To ensure that the output levels from both sources were comparable, a 1 kHz tone was recorded for 15 s, with the sources at the closest position, S1. The level of *OdeonOmni* was set to -4.5 dB digitally from within ODEON. The external amplifier level of the dodecahedron source was then adjusted to match that of *OdeonOmni*. Since the initial level was set below maximum, it allowed for some increase during the measurements, as the sources were placed further from the receiver. This ensured better Signal to Noise Ratio (SNR) and sufficient decay range for the derivation of room acoustic parameters. The advantage with adjusting the level digitally in ODEON is that any change is compensated automatically and does not modify the calibration of the initial level.

Figure 8 shows the measurements of T_{20} and C_{80}, at S3 and S4. The sources derive similar results for both parameters at position 3, while there are some significant differences for C_{80} at S4. Such differences could be attributed both to varying background noise conditions, as well as the properties of the sources. *OdeonOmni* has a flatter frequency response (Figure 6) which means that modes are excited more uniformly within each octave band. Therefore it is expected that measurements are of better quality. The C_{80} curve from the Dodecahedron

seems more unstable with noticeable fluctuations, while the one from *OdeonOmni* looks smoother, following a stable and more reliable trend which is comparable to the C_{80} curve at S3.

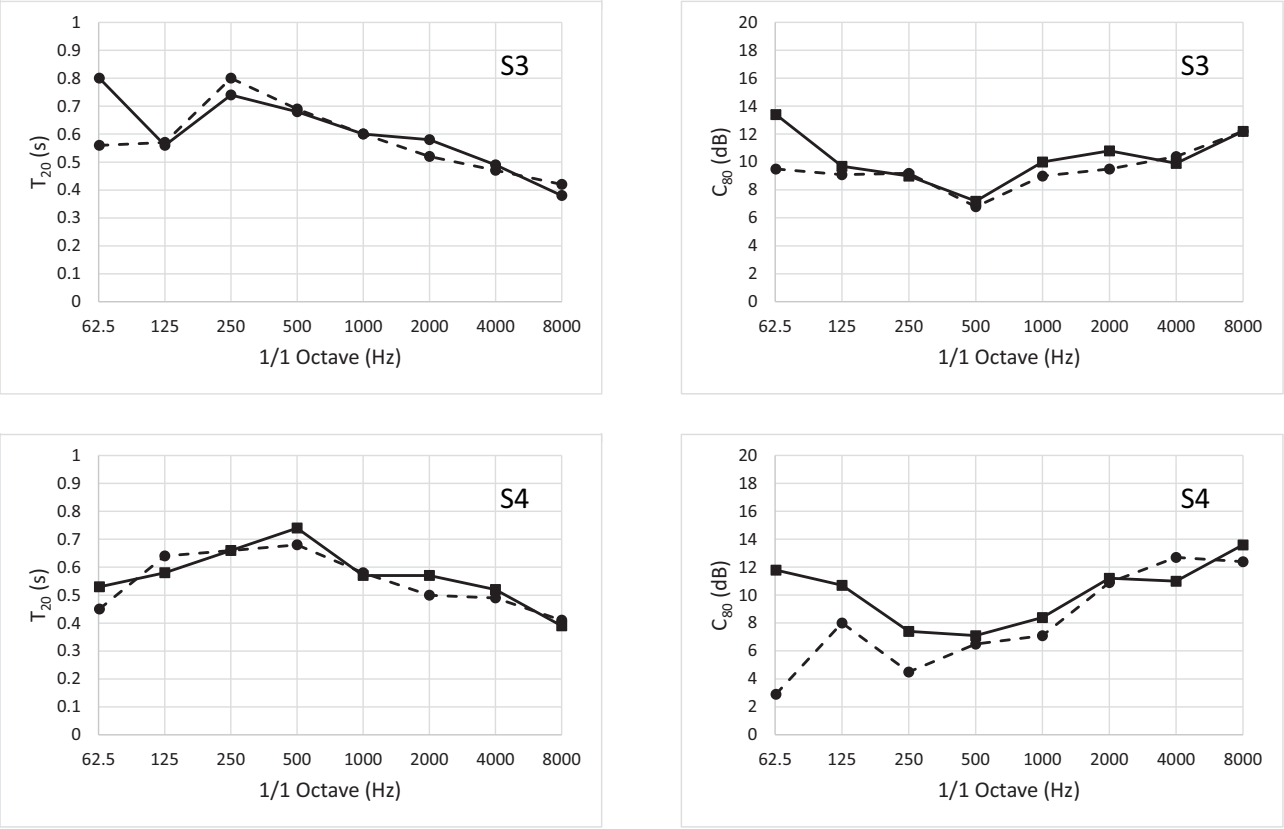


Figure 8: Room acoustic parameters T_{20} and C_{80} as a function of frequency (octave bands), measured at source positions 3 and 4 with *OdeonOmni* (—) and the Dodecahedron source (- - -).

OdeonOmni was also tested at the maximum distance of 30 m in another part of the cafeteria. Between 125 Hz and 8 kHz it was possible to derive all room acoustic parameters, while at 63 Hz, some of them were missing, mainly because of high background noise levels at lower frequencies. It is worth mentioning that a similar experiment with a conventional Dodecahedron would require very long cables, which makes it less practical.

5 Discussion

The design of *OdeonOmni* has been based on single loudspeaker driver which keeps the source small and lightweight. The biggest drawback of this design is a maximum power that is limited to 95 dB. However, with a relatively flat frequency response and use of sweep signals, the source seems to be capable for room acoustic measurements in rooms up to 30 m with moderate background noise levels. Two more advantages result from the specific design. First, the use of one loudspeaker driver makes it easier to know the exact shift in level between several *OdeonOmni* items, in contrast to a dodecahedron source were more drivers are involved and differences in levels cannot be tracked. This facilitates sound strength measurements using the *two-step calibration method* [12]. Using this method, only a prototype *OdeonOmni* needs to be calibrated in the laboratory. Any other items can go through an *in-situ* correction process, which compensates for any shifts in level, due to the different driver. In addition, because of its small overall size and good performance at low frequencies (below 63 Hz), *OdeonOmni* can be placed close to corners in a room for accurate measurement of modes [13].

6 Conclusions

In this paper, a new design of an omnidirectional source for room acoustic measurements has been presented. The source has been developed and manufactured by Odeon A/S as a light-weight, wireless and overall portable solution to encourage acousticians to perform more measurements in the future. The directivity pattern of the *OdeonOmni* has been measured in an anechoic chamber and it has been shown that it fulfills the requirements given in ISO 3382. Evaluation measurements in a large cafeteria space have shown that the impulse responses as well as common room acoustic parameters are similar to those obtained with a conventional dodecahedron source. In such a space, the *OdeonOmni* source has been able to make measurements up to 30 m with Bluetooth connection, with a power level sufficient to provide room acoustic parameters from at least 125 Hz and higher. Finally a direct comparison between the frequency response of *OdeonOmni* and the frequency response of a dodecahedron source, shows that *OdeonOmni* has much better performance at low frequencies and much flatter response throughout the range between 63 Hz and 8 kHz, providing more uniform excitation of the space under measurement and better auralisation when the measured impulse response is convolved with an anechoic signal.

Acknowledgements

The authors wish to thank the two Odeon student assistants, Caroline Gaudeoso and Mario Alfredo Montoya Sandoval for their help during measurements and post-processing, as well as the committee at Danish Tech Challenge 2016 that offered valuable feedback during the first stages of the project.

References

[1] ISO 3382. *Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces, Part 2: Reverberation time in ordinary rooms, Part 3: Open-plan offices*. European Standard, 2009.

[2] M. Vorländer. *Auralization, Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. Springer, Berlin, Germany, 1st edition, 2008.

[3] Norsonic official page. URL https://norsonic.be/.

[4] Brüel & kjær official page. URL https://www.bksv.com/.

[5] Qsources official page. URL https://www.qsources.be/.

[6] N. M Papadakis and Stavroulakis E. Low cost omnidirectional sound source utilizing a common directional loudspeaker for impulse response measurements. *Applied Sciences*, 8, 2018.

[7] A. Richard, C. L. Christensen, and G. Koutsouris. Sine sweep optimization for room impulse response measurements. *Proceedings of Forum Acusticum*, 2020.

[8] *ODEON User's Manual*. Odeon A/S, 2022. URL https://odeon.dk/download/Version17/OdeonManual.pdf.

[9] W. M. Leach. *Introduction to Electroacoustics Audio Amplifier Design*. Kendall/Hunt Publishing, 4050 Westmark Drive, Dubuque, IA 52002, 3rd edition, 2003.

[10] A. K. Nørgaard. *Omnidirectional Loudspeaker for Room Acoustic Measurements*. B.Eng Thesis, Electrical Engineering, Technical University of Denmark, 2017.

[11] ISO 3741. *Acoustics - Determination of sound power levels and sound energy levels of noise sources using sound pressure - Precision methods for reverberation test rooms.* European Standard, 2010.

[12] C. L. Christensen, G. Koutsouris, J. Gil, and J. H. Rindel. Applying in situ recalibration for sound strength measurements in auditoria. *Proceedings of the Institute of Acoustics, year = 2015,*.

[13] H. K. R. Berg. *In-situ analysis of room modes and absorption coefficients at low frequencies*. MSc Thesis, Electronic Systems Design, Norwegian University of Science and Technology, 2021.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Effective absorption of architectural ETFE membranes in the lab

Yannick Sluyts[1], Christ Glorieux[2], Monika Rychtarikova[1,3]

[1] KU Leuven, Department of Architecture, Campus Brussel & Gent, Paleizenstraat 65, 1030 Brussel / Hoogstraat 51, 9000 Gent, Belgium

[2] Laboratory of Acoustics, Soft Matter and Biophysics, Department of Physics and Astronomy, KU Leuven, Celestijnenlaan 200D, B3001 Heverlee, Belgium.
[3] STU Bratislava, Faculty of Civil Engineering, Radlinskeho 11, 81108, Bratislava, Slovakia.

yannick.sluyts@kuleuven.be

Abstract

In this paper the determination of the sound absorption characteristics of architectural ETFE (Ethylene tetrafluoro ethylene) membranes by an adapted ISO354 [1] measurement method is reported. Despite the large-scale deployment of these systems in the built environment, little work has been published on this matter. Architectural ETFE membranes and cushions, which are often based on a multilayer foil structure, are gaining popularity amongst architects and designers. As the sustainability topic evolved from an elusive goal to a stringent requirement, over the past decades, designers have been seeking for new material solutions. ETFE membranes fits this paradigm. These lightweight membranes allow for a more efficient structural design that leads to a reduction of material use for structural purposes. ETFE foils themselves require less energy for production and transportation than glass panes [2]. The lightweight nature of the membranes results in acoustic transparency at low frequencies. In a typical setting where the membranes are used in a roof or wall structure, this behaviour leads to low frequency acoustic waves not being reflected. From a room acoustic perspective, the high one-way transmission to the outdoor world can be seen an “effective absorption” by the membrane. By the reducing effect on the sound pressure level and reverberation time, this is highly beneficial for the indoor acoustic comfort.
In terms of determination of this “effective absorption” in a reverberant chamber, this acoustic transparency poses a challenge: acoustic waves that pass the membrane, reflect on the hard wall or floor behind it, and re-enter the room after having been partially absorbed. This situation is quite different from the one where ETFE membranes are encompassing the building interior, in which the part of the acoustic wave energy that passes the membrane does not re-enter the room. This complication was solved by mounting the ETFE membrane on a frame and place it on a box filled with mineral wool, so that the box space under the membrane acted as an acoustic sink, mimicking an outdoor (100% absorbing) environment. Despite the small size of the sample, reliable values were found for the frequency dependence of the “effective absorption”, and implemented in the framework of simulations, auralizations and measurement interpretation of real spaces based on ETFE membranes and cushions.
Keywords: ETFE, room acoustics, cushions, absorption, ISO354

1 Introduction

ETFE membranes are thin extruded films that are often used in architectural projects. In medium to large sized projects, the membranes are used as an alternative to glass panes as a separation between outside and inside. While the membranes are often used to form air-inflated cushions in buildings, in this paper single layer membranes are examined exclusively. The surface area of these constructions often exceeds 1000 m^2. ETFE membranes/cushions are usually applied as roof structures in shopping malls, public transit terminals and hotel lobbies. Due to the design flexibility of these structures, the shapes and sizes of individual cushions and entire constructions vary widely.

2 Conceptual approach: room acoustics

The room acoustic behaviour of ETFE membranes is unconventional for a building material [3]. The acoustic transparency of the ETFE membranes is high in the low frequencies and low in high frequencies. The transparency contributes to a high "effective absorption" (in room acoustic modelling) at low frequencies.

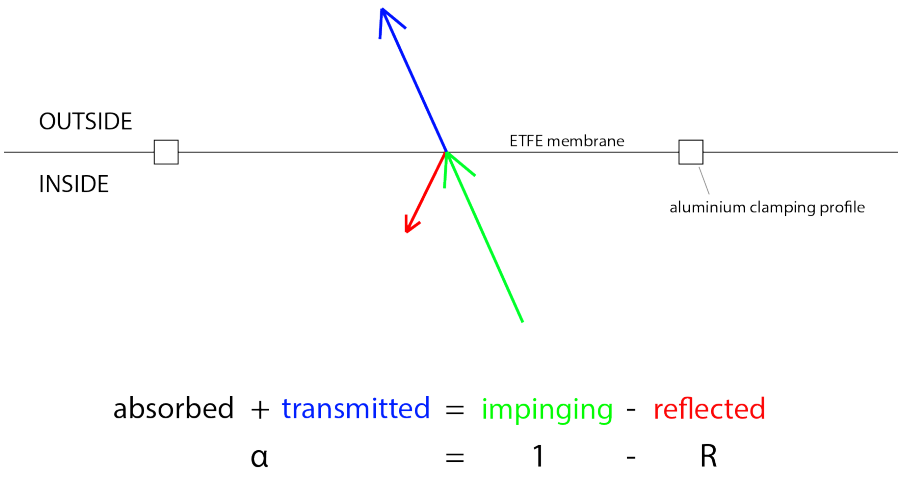


Figure 1. Transmission and reflection of an ETFE membrane. At low frequencies, most of the energy is transmitted. In the setting where the membrane is encompassing a building interior, from a room acoustic point of view, this transmitted energy can be considered as "absorbed". At high frequencies, membranes are quite reflective.

3 Experimental approach

The effective absorption was measured using the conventional ISO354 method in the reverberation chamber. However, the transparency of ETFE membranes made it challenging to accurately measure their effective absorption. For this measurement campaign, two 2.08 m² aluminium frames with tightened ETFE membranes were available. The surface tension was comparable with the one in architectural projects.

Since it was expected to measure a relatively large absorption in the low frequencies, a 60 cm thick layer of glass wool was placed in a 61 cm high MDF box under the transparent ETFE membrane (pictured in Figure 2). The 60 cm of glass wool effectively mimics an outdoor environment that doesn't reflect any sound energy (transmitted through the membrane). A measurement was performed to check this assumption and the results will be discussed later on in this paper (setup "C").

Figure 2. Basic experimental setup in the reverberation chamber (A): a 61cm high wooden box filled with 60cm of mineral wool is used to absorb the acoustic energy that is transmitted through the membrane.

The wooden box was made such that the frame fitted flush on the box. The walls of the wooden box were sufficiently thick (according to the mounting type A description in ISO354).

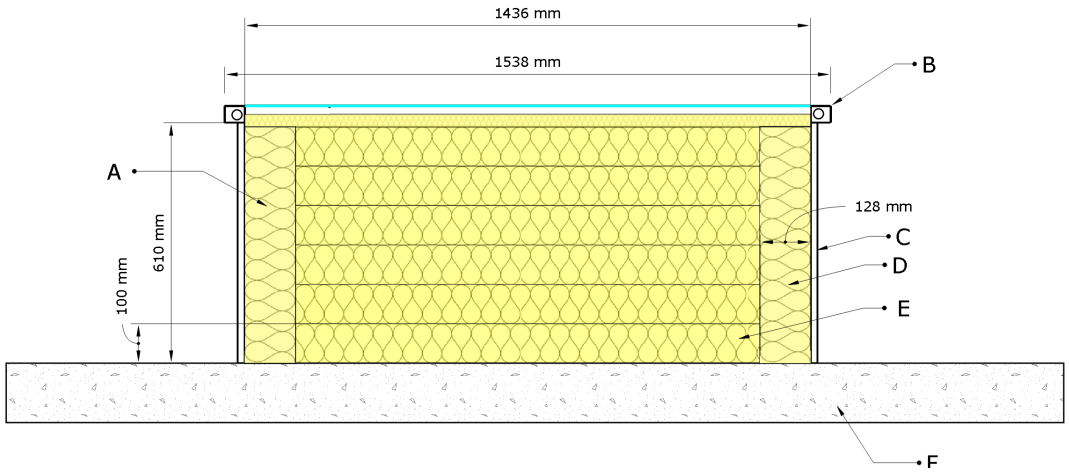


Figure 3. cross-section of the box. A: Compacted foam mats, B: Aluminium frame, C: MDF panels 18 mm thick, D: rolled up foam mats, E: glass wool panels 100 mm thick, F: laboratory floor (concrete)

While the reverberant room is suitable for ISO354 measurements, the sample surface size (2.08 m²) is below the required size (10-12m^2). Therefore, to validate the first measurement, the second ETFE frame was used to double the surface area of the test sample. A second box was constructed, the result is shown in Figure 4.

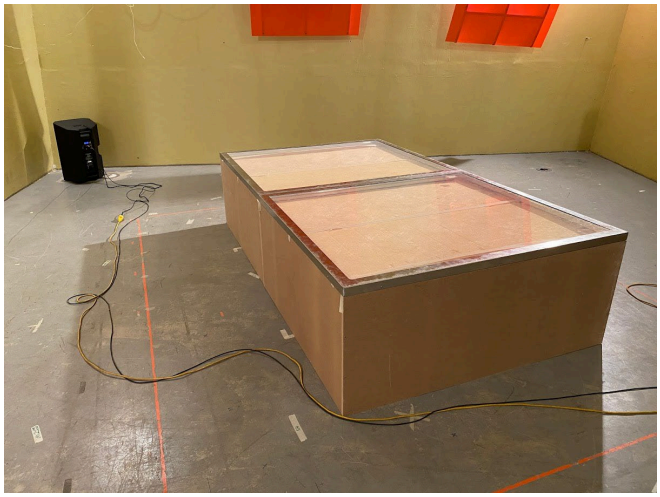


Figure 4. Supplementary measurement configuration with two boxes, each filled with glass wool absorption, both covered with an ETFE frame. This configuration was constructed to increase the absorption behind the ETFE structure of interest (to simulate open air), and thus to enhance the accuracy of the measurement (B).

4 Calculation

Reverberation time was measured by using the interrupted noise method. The samples were positioned in the middle of the room. The sample absorption information was extracted from two sets of measurements: a reference measurement 1 with the room and empty box (in the analysis, the small absorption of the inner box surfaces was negligible), and a measurement 2 with the sample covering the mineral wool filled box. For each set, 8 microphone positions and 3/4 speaker positions were used. Three speaker positions were used with sample and four without sample. The sound production and signal acquisition were controlled remotely, and the temperature and humidity were closely monitored and averaged over time.

The effective absorbing surface of the mineral wool backed ETFE was determined in the standard way:

$$A_T = 55.3V\left(\frac{1}{c_2 T_{30,2}} - \frac{1}{c_1 T_{30,1}}\right) - 4V(m_2 - m_1) \quad (1)$$

where c_1 and c_2 (m/s) are the values of the speed of sound in air during the measurements 1 and 2 at temperatures T_1 and T_2 respectively. $T_{30,1}$ and $T_{30,2}$ (s) are the reverberation times per third octave band in the empty reverberation chamber with empty box and with sample respectively. m_1 and m_2 [1/m] are the frequency and humidity dependent power attenuation coefficients according to ISO 9613-1 [3]. V=197m^3 is the room volume.

The effective absorption coefficient was determined as:

$$\alpha_s = \frac{A_T}{S} \quad (2)$$

where: S (m^2) is the surface area of the test sample (2.08m^2).

5 Results

Figure 5 shows that in the model prediction of the reference situation with the empty box in the room, an absorption peak occurs in the frequency range in the 125 and 160 Hz third octave bands. This peak corresponds with a resonance due to standing waves in the box that are most likely not present when the box is filled with glass wool. This peak, shown in Figure 5, corresponds with the half-wavelength of the standing waves in the box (1.53m => 223 Hz/2 = 111 Hz, 1.35m => 253 Hz/2 = 126 Hz). For further processing, it was manually removed by interpolating between the absorption values of the adjacent 1/3 octave bands. A simple linear interpolation was performed between the values obtained for 100 Hz and 200 Hz to neutralize this effect.

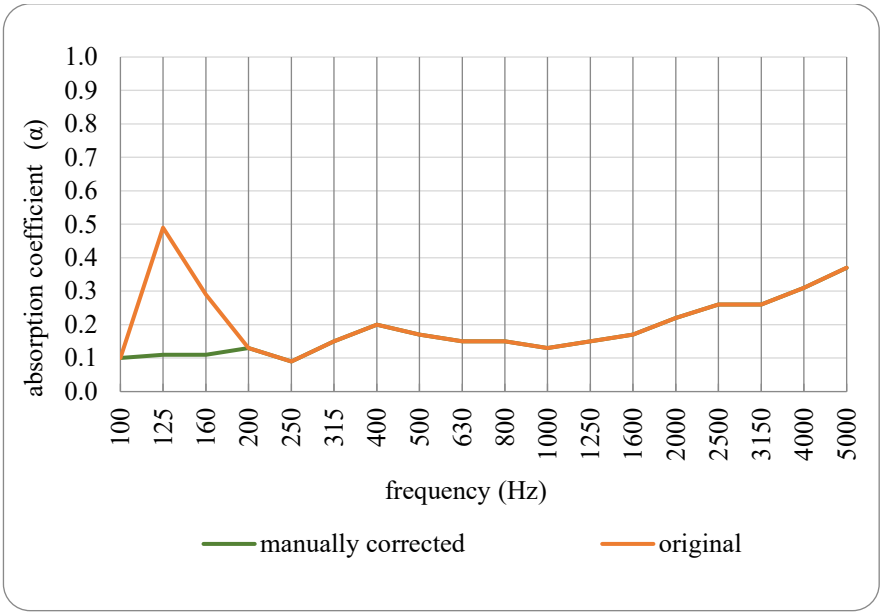


Figure 5. Absorption coefficient of the reference scenario, the empty wooden frame (box)

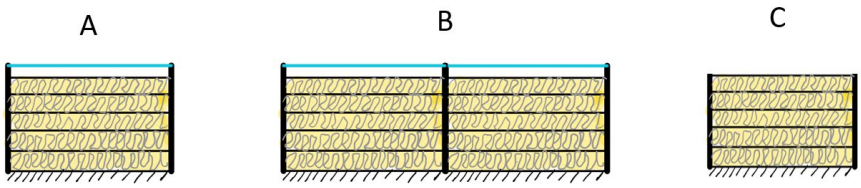


Figure 6. schematic configuration of the 2 measurement configurations A and B, with respectively one (A) and two (B) glass wool filled boxes covered by an ETFE membrane. (C) corresponds to a glass wool filled box without an ETFE membrane or aluminium frame on top. The yellow areas indicate the presence of a 60cm thick layer of glass wool. The blue lines indicate a 250 µm layer of ETFE. Between the membrane and the upper surface of the mineral wool, there was 3-4 cm of air.

In the two measured configurations “A” and “B” (Figure 6), there were respectively one and two boxes, with 2.08 and 4.16 m² of ETFE in the room; the area of the aluminium test frames was neglected in the analysis. The two results for the absorption spectrum in Figure 6 are very similar, with a somewhat higher deviation below 160 Hz, which can be attributed to effects of the above-mentioned resonance.

To check the assumption of broad band total absorption in the observed frequency range by the 60 cm of glass wool, setup ”C” was also measured individually (with the empty box as a reference, box with absorption but no ETFE). In Figure 6 the high absorption of this thick layer of glass wool is displayed, it hovers around 1 in

all measured frequencies. This confirms the hypothesis of very high broad band absorption of the thick layer of glass wool.

To further validate the experimental results, an analytical model for the absorption was used to determine the random incidence absorption on the 2.08 m^2 ETFE membrane. The analytical model was implemented in Odeon(R) and was based on the transfer matrix method [4]. Figure 6 shows that the model result corresponds quite well with the experimental absorption spectra, with a limited excess absorption of about 0.05-0.1 across the whole spectrum.

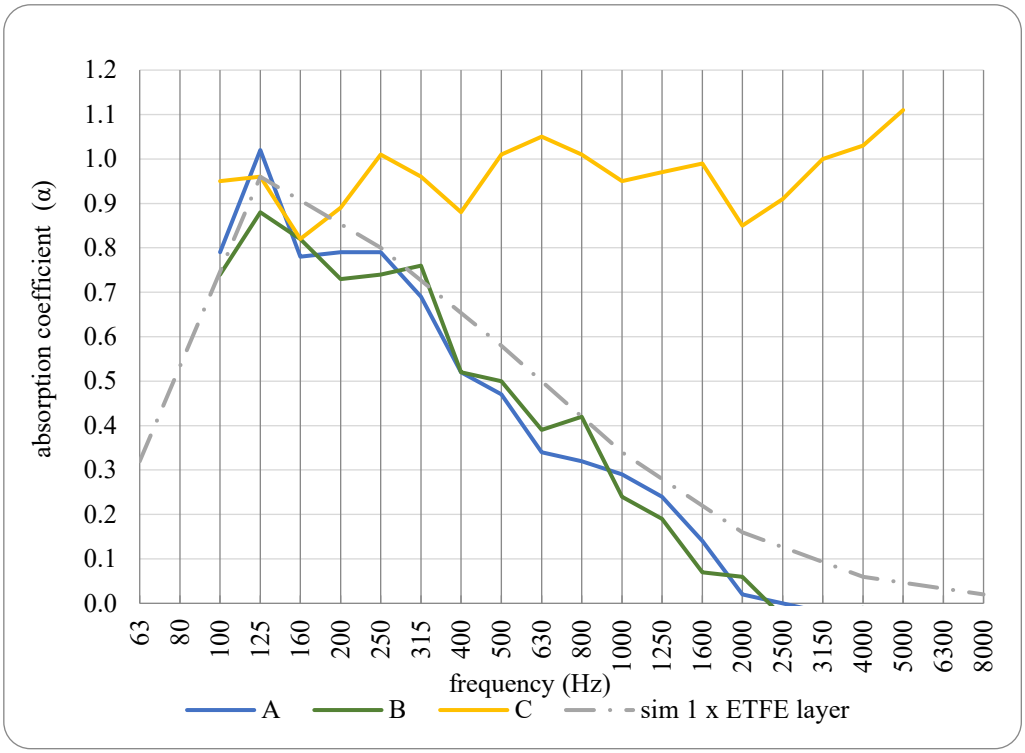


Figure 7: spectra of the absorption coefficient for the measured configurations, in the three configurations (full lines), together with a simulated spectrum (dash-dot line) of the absorption of one ETFE covered, glass wool filled box.

Table 1: Relative standard deviation of the reverberation times used for absorption coefficient determination (in %).

measurement	100	125	160	200	250	315	400	500	630	800	1000	1250	1600	2000	2500	3150	4000	5000
A	0.16	0.19	0.09	0.10	0.06	0.05	0.04	0.04	0.04	0.03	0.04	0.03	0.02	0.03	0.03	0.02	0.03	0.03
B	0.16	0.16	0.08	0.07	0.07	0.08	0.07	0.05	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.02	0.02	0.03
C	0.15	0.17	0.12	0.09	0.08	0.07	0.07	0.04	0.04	0.04	0.04	0.02	0.02	0.02	0.02	0.03	0.03	0.01
empty box	0.16	0.16	0.08	0.07	0.07	0.08	0.07	0.05	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.02	0.02	0.03

Finally, all relative standard deviations of the reverberation time measurements are shown in Table 1 (calculated according to ISO354). Outliers are removed manually but a minimum of 18 individual measurement points were used. The variability of the reverberation time is relatively high in the low frequencies but similar for all measurements in each third-octave band.

6 Conclusion

In spite of its small surface area, the spectra of the absorption coefficient of a flat ETFE membrane of about 2 m^2, determined experimentally using a modified ISO354 method, by using one and two samples, correspond quite well with each other and with the spectrum calculated by the transfer matrix method. The result, which was obtained by placing the ETFE membrane on top of a strongly absorbing mineral wool filled box, confirms that ETFE membranes are transparent at low frequencies and reflective at high frequencies. Further research is ongoing to determine the absorption coefficient of the more complex multi-layered ETFE cushions.

7 References

[1] International Standards Organisation: ISO 354:2003

[2] Lamnatou, C., Moreno, A., Chemisana, D., Reitsma, F., & Clariá, F. (2018). Ethylene tetrafluoroethylene (ETFE) material: Critical issues and applications with emphasis on buildings. In *Renewable and Sustainable Energy Reviews*. https://doi.org/10.1016/j.rser.2017.08.072, pp 2193

[3] International Standards Organisation: ISO 9613-1:1993

[4] Allard, J., & Atalla, N. (2009). *Propagation of Sound in Porous Media: Modelling Sound Absorbing Materials 2e*. John Wiley & Sons.

[5] Chiu, S., Noble, D., & Valmont, E. (2015). *Acoustics in architectural fabric structures: The case of ETFE pillows* (pp. 241–256). Elsevier Inc.

[6] Rychtarikova, M., Urban, D., Maywald, C., Zelem, L., & Kassakova, M. (2017). Advantages of ETFE in terms of acoustic comfort in atria and large halls. *Proceedings of Advanced Building Skins 2017*, 646–654.

[7] Rychtarikova, M., Urban, D., Kassakova, M., Maywald, C., & Glorieux, C. (2017). *Perception of acoustic comfort in large halls covered by transparent structural skins*. https://doi.org/10.1121/2.0000540

[8] Jagt, S. B. D., Laudij, C., Gerretsen, E., & Raijmakers, T. (2015). *Description of the acoustic characteristics of ETFE roof structures* (Vol. 50).

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Simple dynamic measurement system for testing IMU sensor precision in spatial audio

Petar Franček[1,*], Kristian Jambrošić[1], Marko Horvat[1], Vedran Planinec[1]

[1]Department of Electroacoustics, Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia.

*petar.francek@fer.hr

Abstract
Spatial sound is used in a variety of systems these days. The application of high-quality audio systems with real-time processing ranges from hi-fi systems to systems for live sound and virtual reality. The data needed to (re)create realistic surround sound audio is often collected using Inertial Measurement Unit (IMU) sensors. With technological progress, the number of such sensors in many devices is constantly increasing. By combining the data from all sensor devices, a dynamic virtual audio field can be created that provides an immersive experience for the user. When processing the collected raw data, attention must be paid to the precision and tolerance of the sensors. The IMU sensors used in this type of measurement usually consist of an accelerometer, a gyroscope and sometimes a magnetometer, all of which are triaxial. When processing and editing the measured data, the effects of low precision and accuracy, parameter drift and latency can greatly affect the measurement uncertainty and the overall user experience. Sensor devices used for spatial audio fall into three groups: VR headsets, smartphones and DIY sensor systems. However, the cost of a device does not necessarily guarantee the accuracy of the measured data.

This paper presents the results of dynamic measurements on sensors used in embedded systems. Several different types of sensors with different shapes and sizes are used. Some are built on PCB boards with processing units, while others are implemented as modules on dedicated PCBs. A microcontroller is used for data processing. The most widely used open-source platform Arduino is used for the measurements. To avoid the magnetic influence of the moving parts in the setup, a simple aluminium/plastic pendulum is used. To minimise the measurement error, all measurements are taken simultaneously. The measurement results are compared, and the quality of the measured sensors is evaluated.

Keywords: IMU sensor accuracy, binaural synthesis, spatial audio

1 Introduction

Nowadays, devices capable of processing and determining its own orientation are used by many people on a daily basis. These types of devices, e.g. smartphones, cameras, wristbands, trinkets, etc., often consist of a microcontroller for processing purposes and an Inertial Measurement Unit (IMU) sensor for determining their orientation. The usage of this type of devices is growing rapidly, resulting in an increasing number of features built into the devices for optimal user experience and, at the same time, maximizing its utilization. This paper presents the results of measurements conducted on such devices as part of the research focused on determining whether such hardware can be used for binaural synthesis. The data from IMU sensors consist of accelerometer and gyroscope sensor output generated during the measurements. All measurements were taken simultaneously to minimise measurement error, and to facilitate direct comparison of measurement results.
Two previous publications authored by this research team report the results of static and dynamic measurements performed on various smartphones and Arduino microcontrollers. In [1], the focus was put on testing the static performance of the devices, i.e., their stability and resilience to drift of the measured

parameters. In [2], the same devices were investigated in dynamic conditions, i.e. they were placed on turntables with constant angular speed and the same set of measurements was performed. The goal of the research presented in this paper is to investigate the behaviour of these devices in a different, but still controlled dynamic conditions to determine whether their quality is sufficient for use in auralization.

2 Methodology

During the design of the measurement setup, extra attention was paid to minimizing any kind of disturbing magnetic field close to the device under test (DUT). This setup is used for measuring the accelerometer, gyroscope and magnetometer quality of DIY sensors, smartphones and VR sets as a part of the AUTAURA research project [3]. The design of this setup proved to be a comprehensive task, as it had to facilitate dynamic measurements with variable speeds of moving parts, while avoiding the use of any kind of devices that generate magnetic fields. Forces needed for the dynamic measurements can be generated using electrical motors (with interference on magnetic field) or by using other types of actuators (compressed air, etc.). Both of solutions have complexity and cost of a design exceed limitations of a project resources. The optimal solution that meets these specific demands was a setup that utilizes a simple mathematical pendulum (gravity pendulum). Similar solutions have been implemented in [4, 5]. An example of a gravity pendulum is shown in Figure 1.

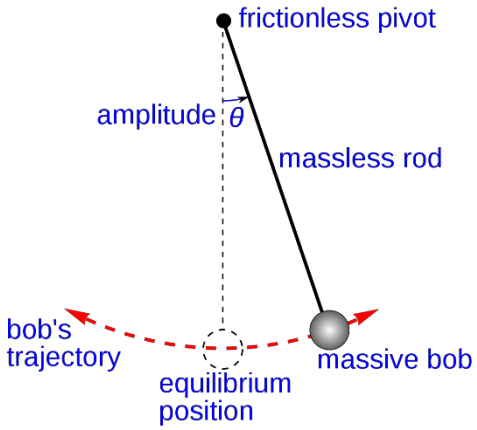


Figure 1: A simple gravity pendulum [6].

Besides minimizing the disturbance caused by magnetic fields, there are several other benefits that come from using a pendulum. It is quite easy to determine the starting parameters (e.g. the height of the bob, the maximum deflection angle θ) and to calculate the force that causes the pendulum to swing back and forth, passing through its equilibrium position. The force F that acts on the bob perpendicular to the rod is expressed as:

$$F = -mg \sin\theta \tag{1}$$

where m represents the centre of mass, g is the gravity constant, and θ with its maximum deflection angle of the pendulum represents starting position. This force is linear to the sensor readings of one accelerometer axis. The period of mathematical oscillation T is constant and can be determined as (2):

$$T = 2\pi\sqrt{\frac{l}{g}} \tag{2}$$

where l represents the length of the pendulum. It can be observed that the period T depends only on the length of the rod (or thread) and the gravity constant. On the other hand, mass m and angle θ have no influence on the period of the pendulum.
Extra consideration was given to the comparability of measured results. It is important to note that the measurements were conducted using three different types of microcontrollers that have been obtained as a project resource for these measurements. That means that the internal clocks in the controllers could exhibit a mismatch during measurements. The process of calibration has been designed with the assumption that the processing of the same set of instructions consumes the same amount of time for each microcontroller [7]. All the devices (microcontrollers with sensors) were left in operating condition for a long period of time, i.e. much longer than the duration of the actual measurements (48+ hours compared to 30 seconds). The time base data was extrapolated by calculating the input parameters. In addition, the use of three different microcontrollers with different hardware and software setups led to different initialization times on each microcontroller-sensor combination. To overcome the problem of inconsistent starting time of sensors, a pushbutton switch was installed. Each microcontroller/sensor has a different booting and warmup time. After powering up and finishing the boot/warmup sequence, time synchronization was done by activating the pushbutton. The code is executed outside of the continuous loop in the microcontroller, thus removing any effect on speed or load on microprocessor [8]. This approach solves the problem of inconsistent internal clocks of each microcontroller and their internal time.
To determine the latency of the sensors, a specific approach is used. A microcontroller with an LED display is used only to display time in the millisecond range, while other microcontrollers are used to perform sensor measurements at the same time. A high-speed video camera [9] is positioned in the equilibrium position, i.e. the lowest point of the pendulum, where the velocity of the pendulum reaches its highest values. Using the recorded high-speed video, it is easy to determine the position of the pendulum where the forces and the velocity are at their maximum. The timestamp can be recorded by reading the display on the recorded video. Using the calibration described above, the latency between the time where the values of measured sensor data reach their maximum, and the timestamp logged with camera can be calculated.

3 Measurement setup

The pendulum was assembled using aluminium, plastic, and wooden parts to minimize its influence on the magnetic field and to lower the mass of the pendulum. A 3-meter-long aluminium rod with a thin cross-section was suspended from the ceiling by means of a small steel axle, which forces the pendulum to swing in a single plane. The rotational friction that appears during the movement of the pendulum (and its axle) was minimized using high-class ball bearings. At the other end of the aluminium rod, a wooden plate is fixed to the rod as the bob, i.e. the moving mass. To increase this mass, a steel counterweight is fixed to the plate as well. The surface of the wooden plate is used to mount the sensor devices with Arduino microcontrollers. Figure 2 shows the measurement setup used in this research, with the pendulum at its maximum displacement from equilibrium. Pictures of setup are taken with loudspeakers positioned behind pendulum, but measurements were conducted without presence of loudspeakers.
The pendulum was made as long as physically possible and its length was limited by the ceiling height of 3.2 m, in order to maximize the period of the motion and the generated forces during measurements. The pendulum moves along the axis (swings on the axle) stretching in the east-west direction, thus restricting its movement to the north-south direction. A set of six microcontrollers and IMU sensors is located on the wooden plate. There are three types of microcontrollers and four different sensor modules. The wooden plate with mounted microcontrollers and sensors is displayed in Figure 3.
One microcontroller is dedicated to displaying the elapsed time with millisecond precision, two microcontrollers have built-in IMU sensors, and three microcontrollers have external connected IMU PCBs. Microcontrollers and sensors are mounted from north to south and labelled as follows: display, LSM6DS3 (external sensor module), NanoInt (internal), UnoInt (internal), ICMAK (external), and DOFv2 (external). The labels correspond to their IMU sensor chip, respectively: LSM6DS3 [10], LSM9DS1 [11], LSM6DS3 [10], ICM-20600 [12], and MPU-9250 [13]. In Figure 3 it is easy to identify the order and the position of each

device. Extra effort was made to position the sensor chips on the marked axis, thus assuring direct comparability of all measured results.
The orientation of the sensor chips is determined by their position on the PCB. The mounting points of the PCBs are used to fix the sensors in the position where minimal wiring to electrical modules is needed. The minimal wiring setup of all sensors mounted on the wooden plate means that power supply is directly connected to sensors, and communication cables are directly connected from each sensor to specific microcontroller. This setup also be seen in Figure 3. In these positions, the sensor modules output the data according to the orientation of axes shown in Figure 4.

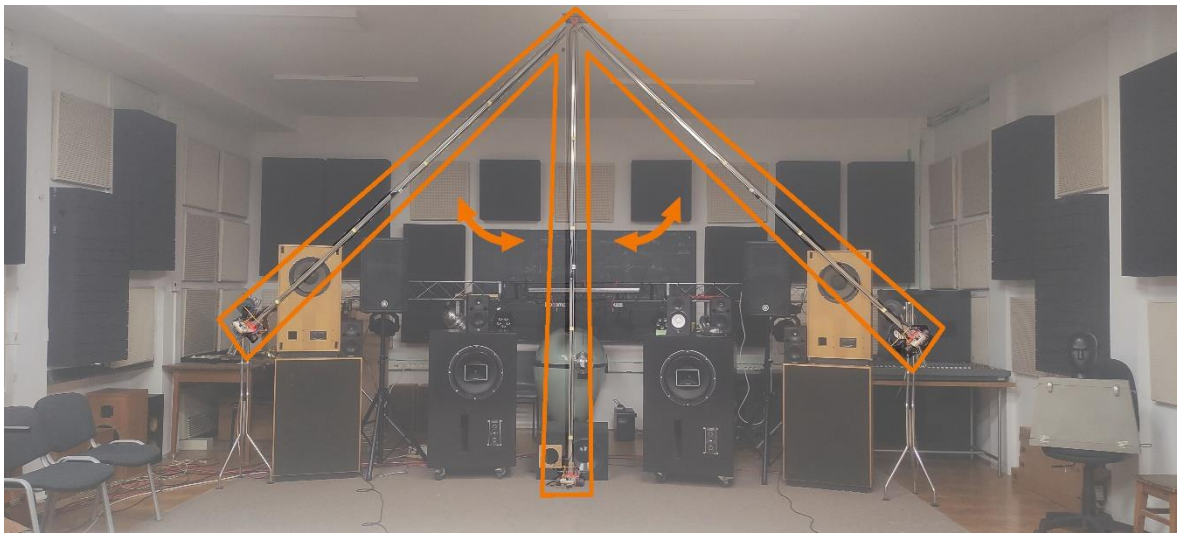


Figure 2: Measurement setup based on a pendulum, with maximum displacement positions on both sides and the equilibrium position.

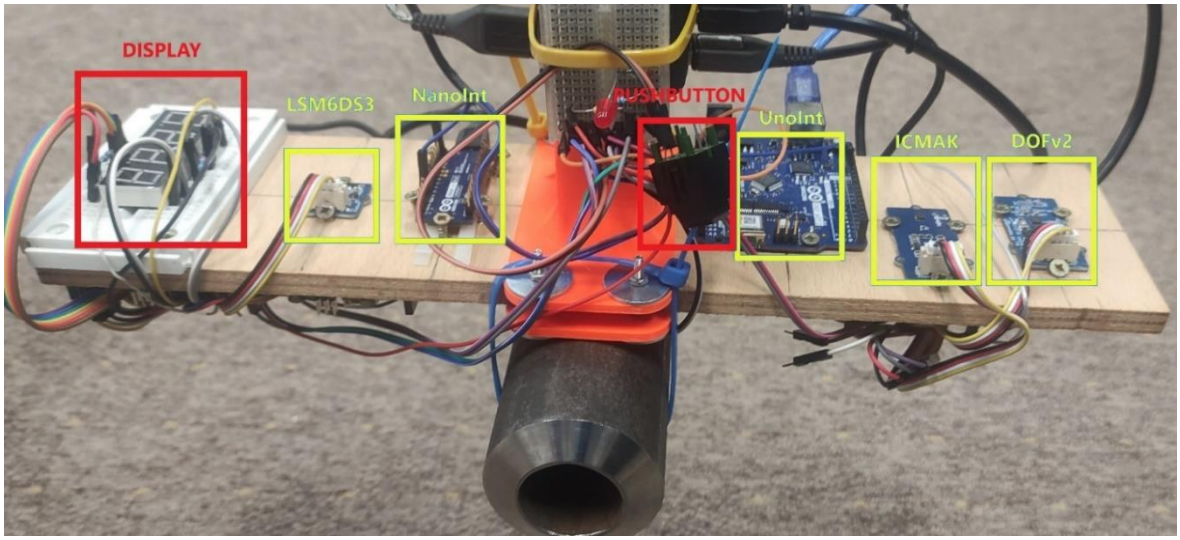


Figure 3: Wooden plate with mounted microcontrollers and sensor modules.

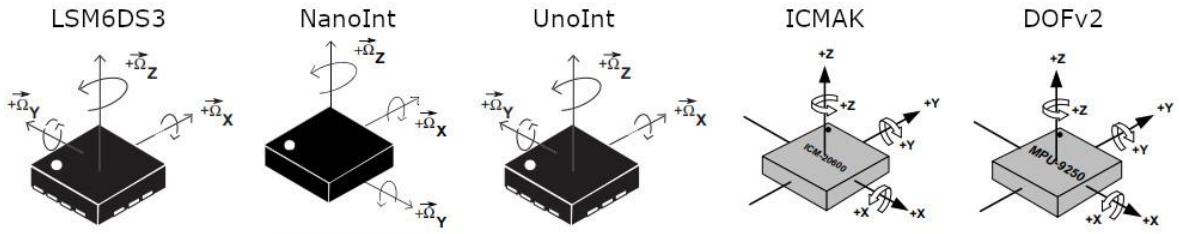


Figure 4: The orientation of the axes for the investigated sensors [10-13].

The orientation of the DOFv2 is used as a reference. All other sensor data was processed with reference to the orientation of this sensor. The mismatch of the axes of different sensors can be seen in Figure 4.
Used setup has hardware limitations. The built-in IMU sensors are connected to the processing unit using two different protocols. In particular, the NanoInt uses the I2C protocol, while the UnoInt uses the SPI protocol. These connections are determined and cannot be changed. As no significant difference in the sensor communication time was observed, the authors decided to use the I2C protocol for all other sensors.
The measured data include the internal time of the microprocessor, the accelerometer data, and the gyroscope data. The magnetometer data was recorded as well during the measurements, and it will be analysed and published as the comparison of IMU magnetometer data obtained from DIY sensors, smartphones, VRs and other sensing devices. The microprocessor internal time is used to avoid latency in communication between the microcontrollers and the data storage system. The lack of the microprocessor RTC module is also justified by minimizing microprocessor load, so that the highest possible sample rate of a IMU sensor can be achieved. Each microprocessor is optimized for a specific task, i.e. to read and send sensor data without any additional computation.
In this setup stock libraries from sensor manufacturers were and parameters (if there is any) for highest refresh rate and precision were selected.

4 Results

This section displays the results of the conducted measurements. All the measurements were performed simultaneously using six microcontrollers. Several sets of measurements were made for different starting height (deflection angle) of the pendulum. T1 and T2 parameters were determined using readings from a record of a high-speed camera video in lowest position of a pendulum where acceleration of Z axis and angular speed of X axis has its maximum. Using this parameters latency of each sensor is calculated. It is possible to observe T1 and T2 as a red vertical line in Figure 6 and Figure 8 around T=4 sec and T=38 sec. For a very precise and accurate comparison, pre-processing of measured data is done. The calibration described above is followed by normalization of the measured data.

4.1 Calibration and normalization of data

The calibration process ensures that the starting and the ending moment of the measurement are aligned in time for all microcontrollers and sensor combinations. The effect of calibration can be observed using graphical representation of the measured data. Without calibration, the data coming from different DUTs neither start nor end at the same moment in time because of different internal clocks of processing units. The time differences with and without calibration are shown on the left chart in Figure 5.

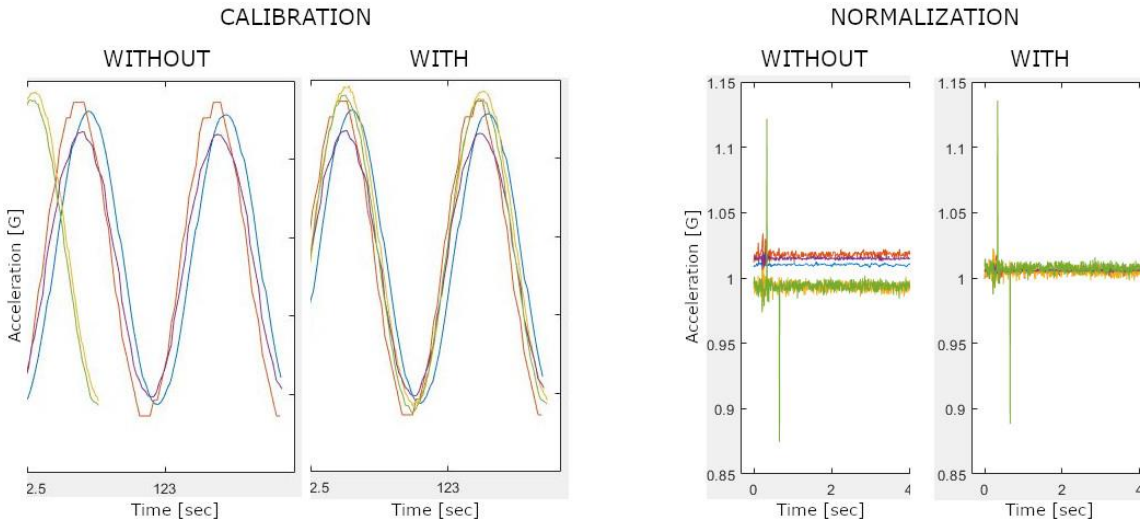


Figure 5: Comparison of measured data with and without pre-processing: left – calibration, right – normalization.

The goal of the normalization procedure is to align the amplitude of raw sensor data obtained from different DUTs. For example, stable static sensors readings of accelerometers without drift are in range from 0.99 to 1.02G. To overcome this difference during the comparison of results, process of normalization is used. The result of normalization is displayed on the right chart in Figure 5.
The described pre-processing applied on measured data facilitates the direct comparison of different DUTs and makes all subsequent calculations much simpler. Nevertheless, all raw data has been saved in a unique database for further calculations and manipulation.

4.2 The results of acceleration measurements

The first measured parameter to be presented is acceleration. It was measured with X axis pointing to the west, Y axis to the south, and Z axis to the floor. The results of acceleration measurements are visible in Figure 6.

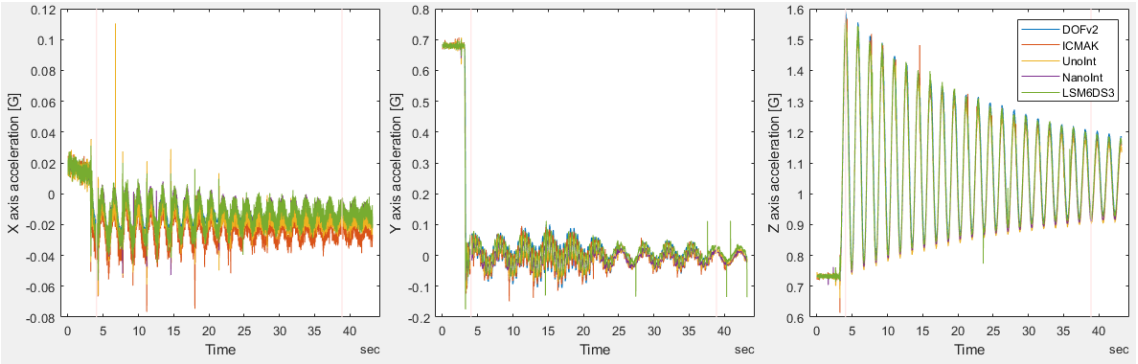


Figure 6: Acceleration results for three axes.

As expected, the biggest amplitudes are found for the Z axis, while on other axes the acceleration values are much lower, but still present. By zooming in on one part of the chart that shows the acceleration on the X axis, it is noticeable that the measured signal of NanoInt and LSM6DS3 is, in fact, out of phase compared to the signals obtained from DOFv2, ICMAK and UnoInt. Figure 7 shows this behaviour.

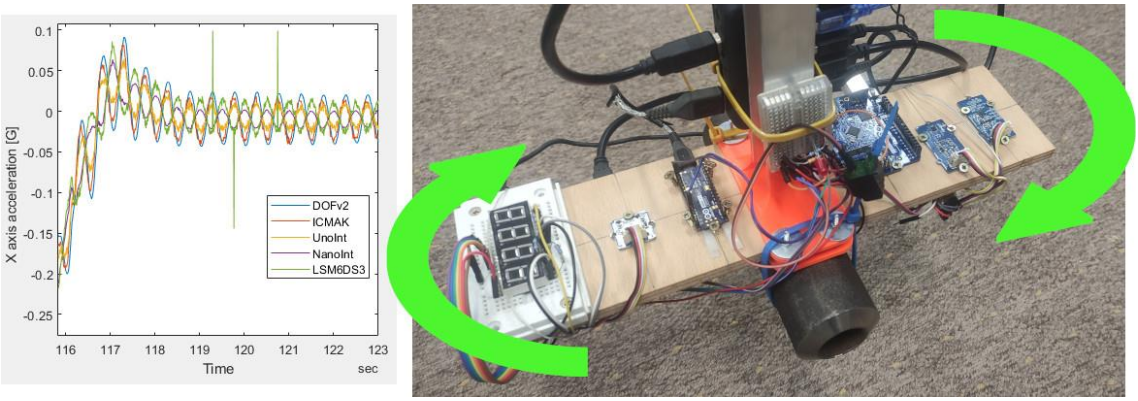


Figure 7: Opposite acceleration forces on X axis in period of harmonic axial rotation.

On the left side of Figure 7 (until timestamp 117 sec) it is possible to observe that all five sensor measurements have a negative acceleration and same trend. From 117 sec onwards, the direction of oscillations of LSM6DS3 and UnoInt sensor are opposite from the direction of the remaining sensors. On the right side of Figure 7, green arrows indicate the rotation direction of the measurement setup (from west to north), thus revealing the reason for such behaviour. The same effect was observed in the data for all three acceleration axes. Table 1 shows the numerical results of the measured parameters with maximum values for Z axis of setup and latency calculated using values T1 and T2 from video record.

Table 1: Numerical representation of recorded acceleration data for Z axis.

Parameter	LSM6DS3	NanoInt	UnoInt	ICMAK	DOFv2
Amplitude (T1) [mG]	1565.80	1557.40	1549.55	1574.05	1591.01
Amplitude (T2) [mG]	914.81	910.39	902.05	905.05	928.11
Latency (T1) [ms]	32.5	38.0	5.3	0.5	22.0
Latency (T2) [ms]	43.0	49.0	28.0	1.0	32.5

It can be noted that in this measurement, ICMAK is a sensor with lowest latency.

4.3 The results of gyroscope measurements

The gyroscope data was recorded during the same measurements as the data obtained for the acceleration. As expected, the maximum amplitude of the angular velocity was found for the rotation along the X axis. As a consequence of torsional forces, the frequency for the other two axes is much higher due to a significantly shorter radius of oscillation. Figure 8 shows the raw data obtained from gyroscope measurements for all three axes.

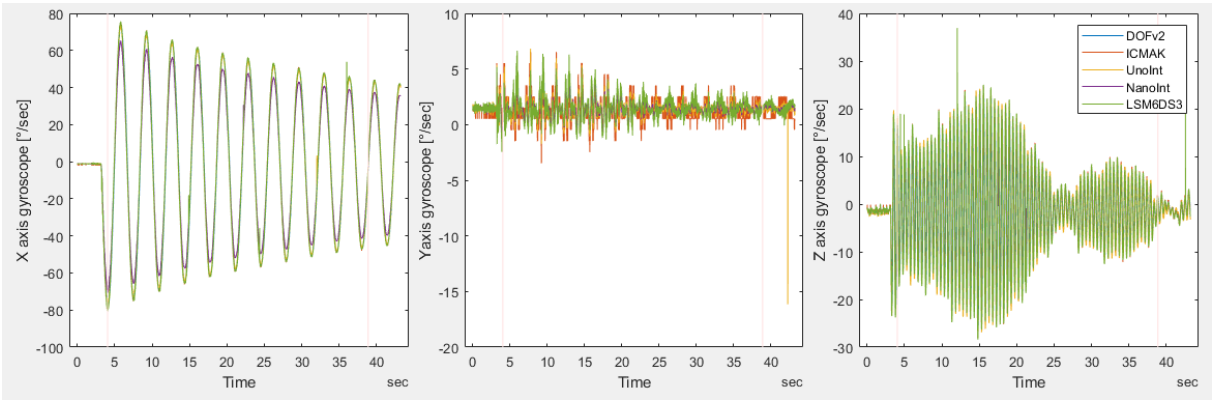


Figure 8: Gyroscope results for three axes.

In this case, the torsion of the wooden plate does not affect sensor readings as it does for the accelerometer data, i.e. in this case no out-of-phase readings have been observed. All rotations are in the same direction, as expected. Values for Y and Z axis are as expected different from 0 because of a torsion movement explained in right part of a Figure 7. Table 2 shows the numerical results of the measured gyroscope parameters for the X axis.

Table 2: Numerical representation of recorded gyroscope data for X axis.

Parameter	LSM6DS3	NanoInt	UnoInt	ICMAK	DOFv2
Amplitude (T1) [°/sec]	80.79	71.06	78.39	80.02	78.91
Amplitude (T2) [°/sec]	44.16	37.58	42.98	43.98	43.28
Latency (T1) [ms]	5.5	6.0	5.3	0.5	33.9
Latency (T2) [ms]	9.8	7.5	16.0	1.0	41.5

5 Conclusions

The described process of investigating the quality of the IMU sensors has raised research questions that must be highlighted and addressed. First of all, this setup has mentioned hardware limitations. The built-in IMU sensors are connected to the processing unit using two different protocols and the authors decided to use the I2C protocol for all sensors where it was possible.

It is important to mention that in some cases the NanoInt sensor had the data refresh status register at false state, although the values were changing, and new ones were updated. This problem was solved by not checking the state of the status register and simply reading the values from it.

The lack of high precision in some sensors can be the consequence of hardware limitations and/or suboptimal software implementations. In our examples where stock libraries were used, and problems were detected when the processing speed was compared with alternative libraries written by the community. Some community-written libraries enabled a quadruple increase in speed, and improved precision. The problem with such types of stock libraries is that the code is often undocumented and without any references, so the accuracy of the received data might be questionable.

One of the sensors does not have direction marks on its PCB nor any marks on the chip, so its true orientation was ambiguous. Things get more complex when community forums are contacted, because the marks on the PCB are not consistent for the accelerometer and the gyroscope. After rethinking and checking the datasheet, an agreement was reached regarding the orientation of the sensor based on common sense. The data obtained from the conducted measurements confirmed that the selected directions/orientations were properly chosen.

The process of calibrating the time base of each microcontroller can be omitted by recording the data directly in real time. This approach requires more hardware resources, and it can be used in real time processing. It was not implemented as a solution to be examined in this paper because the authors wanted to record data in the microcontroller, thus avoiding the unnecessary additional lag and uncertainty that stem from communication between the microprocessor and the data storage unit.

The results of measurements show that the ICMAK module exhibits the lowest latency. On the other hand, that same module also has the lowest resolution. The best resolution and the highest latency were found for the DOFv2 module. The features of the optimal device are the result of a compromise between responsiveness (as low latency as possible) and resolution (as high as possible).

Further measurements of sensors built into smartphones and VRs or implemented in some other ways are planned in the future using the same pendulum setup. Upon finishing all measurements, a comparison with representative samples of each group will be done to assess the quality of different devices and their usability in binaural head-tracking systems.

Acknowledgements

The authors acknowledge financial support by the Croatian Science Foundation, (HRZZ IP-2018-01-6308, "Audio Technologies in Virtual Reality Systems for Auralization Applications (AUTAURA)".

References

[1] Kristian Jambrošić, Miljenko Krhen, Marko Horvat and Tomislav Jagušt. and Peter Aspinall. Measurement of IMU sensor quality used for head tracking in auralization systems. *Proceedings of e-Forum Acusticum 2020*, December 2020.

[2] Kristian Jambrošić, Vedran Planinec, Marko Horvat and Petar Franček. Precision of inertial measurement unit sensors in head-tracking systems used for binaural synthesis. *INTER-NOISE and NOISE-CON Congress and Conference Proceedings, pp. 2634-2645(12),* August 2021. doi: 10.3397/IN-2021-2190.

[3] [Online, Accessed 03/03/2022], https://autaura.fer.hr/autaura

[4] S. E. R. Charel, E. H. Binugroho, M. A. Rosyidi, R. S. Dewanto and D. Pramadihanto. Kalman filter for angle estimation using dual inertial measurement units on unicycle robot. *2016 International Electronics Symposium (IES),* 2016

[5] Kee-Young Choi, Se-Ah Jang, Yong-Ho Kim. Calibration of Inertial Measurement Units Using Pendulum Motion. *International Journal of Aeronautical and Space Sciences, 11(3), pp. 234-239*, 2010. doi: 10.5139/IJASS.2010.11.3.234

[6] [Online, Accessed 03/03/2022], https://en.wikipedia.org/wiki/Pendulum

[7] Bruce Jacob, Spencer W. Ng and David T. Wang. Memory Systems, Cache, DRAM, Disk. *Elsevier Inc., 2008*, doi: 10.1016/B978-0-12-379751-3.X5001-2

[8] [Online, Accessed 04/03/2022], https://www.arduino.cc/reference/en/language/structure/sketch/loop/

[9] [Online, Accessed 04/03/2022], https://www.samsung.com/global/galaxy/what-is/frames-per-second/

[10] [Online, Accessed 08/03/2022], https://botland.store/index.php?controller=attachment&id_attachment=1307

[11] [Online, Accessed 08/03/2022], https://www.st.com/resource/en/datasheet/lsm9ds1.pdf

[12] [Online, Accessed 09/03/2022], https://3cfeqx1hf82y3xcoull08ihx-wpengine.netdna-ssl.com/wp-content/uploads/2021/05/DS-000184-ICM-20600-v1.1.pdf

[13] [Online, Accessed 10/03/2022], https://invensense.tdk.com/wp-content/uploads/2015/02/MPU-9250-Datasheet.pdf?ref_disty=digikey

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

A parameter-conditional neural network framework for modelling parameterized auditory models

Peter Asbjørn Leer Bysted [1,3,*] **Jesper Jensen**[2,3], **Zheng-Hua Tan**[3], **Jan Østergaard**[3], **Lars Bramsløw**[1]

[1]Eriksholm Research Centre, 3070 Snekkersten, Denmark.
[2]Oticon A/S, 2765 Smørum, Denmark.
[3]Department of Electronic Systems, Aalborg University, 9000 Aalborg, Denmark
[*]pelb@eriksholm.com

Abstract

The use of auditory models is important for designing speech and audio processing algorithms for hearing assistive devices. These auditory models are often parameterized by a set of parameters relating to auditory function, e.g., hair cell loss or synaptopathy. In practice, the computational load of these auditory models can be very high thus limiting the feasibility of using the models as bio-inspired loss functions for deep learning based hearing loss compensation strategies or denoising strategies. Previous efforts have addressed this problem by training a neural network for each parameter configuration of the auditory model which greatly reduces the computation time of the auditory model but requires a new network to be trained whenever the parameterization changes. In this paper we propose an approach where a single neural network is trained, once and for all, to accurately simulate auditory models across their parameter spaces by conditioning the weights of the network on the parameters of the respective auditory models. This approach enables greater flexibility than training a single model for each parameter configuration, as any parameterization can be acquired on the fly. The accuracy of the neural network is shown to be robust across both unseen inputs and standard audiograms.

Keywords: Auditory models, deep learning, neural networks, computational modelling

1 Introduction

The mammalian auditory system is a complex, dynamic and non-linear system which processes and extracts information from acoustic signals. In order to gain understanding and predictive capability of the auditory system, there has for several decades been undertaken a large effort to model the experimental data. Traditionally the auditory system is broken down into smaller sub-systems corresponding to a functional or anatomical sub-structure, e.g. the cochlea or the auditory nerve. These models are themselves often complex and might be computationally expensive. The models are therefore limited in their use-cases for real-time signal processing or deep learning applications. Previous approaches have solved this problem by training a neural network, here denoted as an Auditory Model Simulator (AMS), to simulate the input-output relations of the sub-systems of the auditory systems [1][2], which can simulate the auditory models in real-time. The drawback of this approach is that an individual AMS has to be trained for each parameter configuration of a given auditory model. To circumvent this problem, we present a framework that can simulate the inner representation of the auditory model, by extending the previous described approach to include a Weight Generating Network (WGN), a neural network that generates the weights for the fixed structure of the AMS conditioned on the parameter space of the auditory model.

2 Methods

In this section the overall framework and the different components of the proposed framework will be introduced, including the chosen example auditory model, the architecture of the AMS and WGN, the training procedure and the parameters of the networks.

2.1. Modelling framework

We denote the auditory model, as $f_\theta : X \rightarrow I$, with X being the signal space, I the inner representation space of the model and θ the parameters of the model. We approximate this model by the structure shown in Figure 1. We generate the parameters for the AMS from the WGN: $\Theta \rightarrow W$, where Θ is the weight space of the model, and W is the parameter space of the AMS that approximates f_θ, which we denote as $\hat{f}_\theta$. There are thus two disjoint paths in the framework, each of which is represented by a neural network: The signal path (represented by the AMS) and the parameter path (represented by the WGN). By decoupling the audio and the audiogram inputs through two networks, the complexity at inference is smaller than training one large network, as the functional relationship between the auditory models and audiogram is disjoint from the audio processing path.

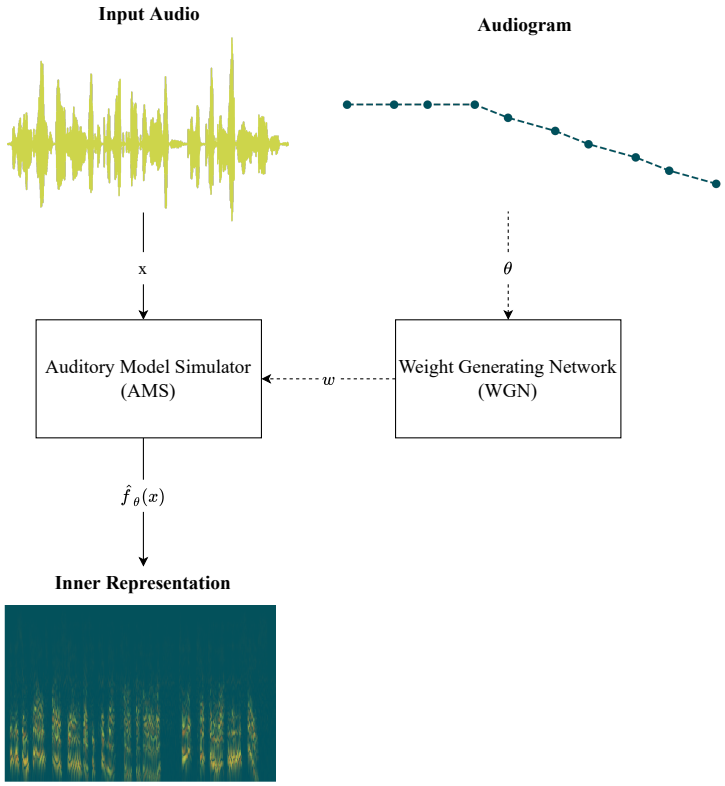


Figure 1: An acoustic input signal, x, is passed through the Auditory Model Simulator (AMS). The weights of the neural network are generated by the Weight Generating Network (WGN), that produces the neural network weights based on model parameters, θ. The bold lines denote the signal path, and the dashed lines denote the parameter path. The result is a model inner representation.

2.2. Auditory model

In this work we consider the Inner Hair Cell (IHC) stage of the UR EAR 2020b model[3][4] as an exemplary model. The model is illustrated in Figure 2. The model consists of K parallel non-linear auditory filters modelling the IHC transduction induced by the movement of the basilar membrane, with filters placed successively along the basilar membrane. The center frequencies of the filters are logartihmically distributed from 125 to 8000 Hz. For each filter, there are 2 parameters related to hearing loss, c_{OHC} and c_{IHC}, relating to outer hair cell function and inner hair cell function, respectively. The model is supplied with a function that for any given audiogram computes the full set of appropriate model parameters.

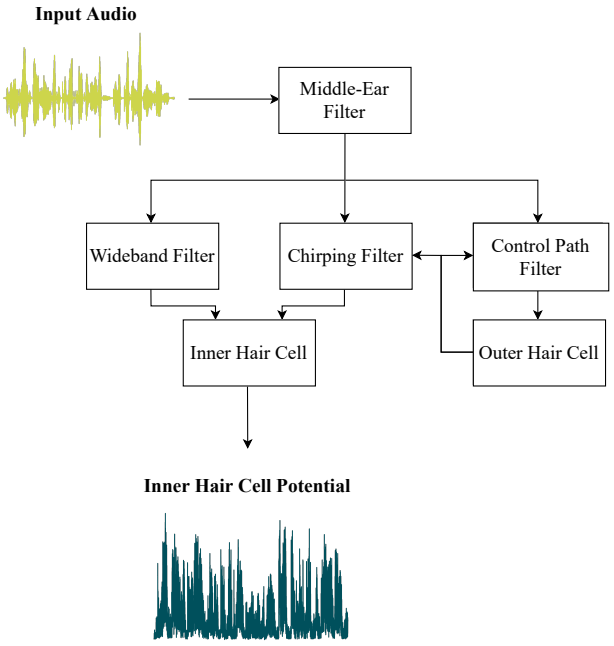


Figure 2: Example auditory model [3][4].

2.3. Auditory Model Simulator

For the AMS we use the Wave-U-Net structure [5]. The structure is a U-shaped convolutional auto-encoder, that for each layer performs a convolution, followed by downsampling or upsampling and a non-linear activation function. The network is subdivided into blocks according to their function in the network: The input block, downsampling block, embedding block, upsampling block and output block. All blocks except for the input block include a 1D-convolution operation parameterized by the WGN, resulting in a total of 2N+2 1D-convolution operations, where N is the number of downsampling blocks. The architecture is illustrated in Figure 3. The Wave-U-Net is very similar to previously used models for simulating cochlea respones [1]. In the previous work the structure was found appropriate for simulating different hearing losses, and different models, such as models of the cochlea, the inner hair cells and the auditory nerve, motivating this choice of network.

2.3.1 Input Block

This block resamples the input to 20 kHz and crops the input such that the input length is a multiple of 2^N. This ensures that the length of the output of the network and skip connections are consistent with the input.

2.3.2 Downsampling Block

The downsampling block is a 1-dimensional convolution block, i.e. convolution over time and an activation function, followed by decimating the signal by a factor 2. Note, there is no anti-aliasing filters in this setup. Skip connections are connected to the upsampling blocks, which might help restore the phase information that might be lost during down-sampling and improve gradient flow during training.

2.3.3 Embedding block

The embedding block consists of a convolutional layer and an activation function. The parameters of the convolution is the same as of the downsampling block, but there is no skip connection and downsampling.

2.3.4 Upsampling Block

The upsampling block is structured inversely to the downsampling block. The input is upsampled by linear interpolation, followed by a concatenation with the skip connection from the downsampling block and a convolution block.

2.3.5 Output Block

The output block concatenates the input and performs a point wise convolution with the last upsampling layer and the input. The inner representation is the output of this layer.

2.4. Weight Generating Network

The WGN generates the neural weights for the AMS by a linear combination of parameter tensors. The WGN is illustrated in Figure 4. Denote the tensors containing the weights for the nth layer in the AMS by $\mathbf{w}_n$, then the WGN generates $\mathbf{w}_n$ as a function of audiogram parameters θ, by a linear combination of K tensors:

$$\mathbf{w}_n = \sum_{k=1}^{K} \alpha_{n,k} \mathbf{w}_{n,k} \tag{1}$$

The $\alpha_{n,k}$ are generated by the WGN, by feeding the model parameters found from the audiogram through a 3-layer Multi Layer Perceptron (MLP), e.g. a fully connected feedforward network with 3 layers. The output of the MLP has dimensions $(K(2N+2))$, and is reshaped into a matrix of dimension $(2N+2, K)$. This matrix is split into 2N+2 K-dimensional vectors, and the softmax function is computed across the K elements in each vector, outputting $0 \leq \alpha_{n,k} \leq 1$, which are used as in Equation (1). Note that the method is similar to the Dynamical Convolution approach in [6] and Conditional Convolution approach in [7], the difference being that these approaches generate the weights by conditioning on the input and therefore solves a different problem.

2.5. Training

In order to train the network, a dataset is created by generating 7500 input-output pairs, $\{T \coloneqq (x, f_\theta(x)) | x \in X, \theta \in \Theta\}$, where X is 7500 random sentences from the LibriTTS [8] database, and Θ is generated by sampling from 10 different standard audiograms [9] and multiplying the audiograms by a random scalar between 0.5 and 1. The audiograms are defined for 10 frequency bands and are interpolated linearly on a log/dB scale for center frequencies between these frequency bands. For the results presented here we use J=90 center frequencies, or frequency channels, of the auditory model. For all audiograms we contribute 2/3 of the threshold shift to the outer hair cells (OHC) and 1/3 to the IHC. The inputs are normalized to 80 dB SPL, and the input-output pairs

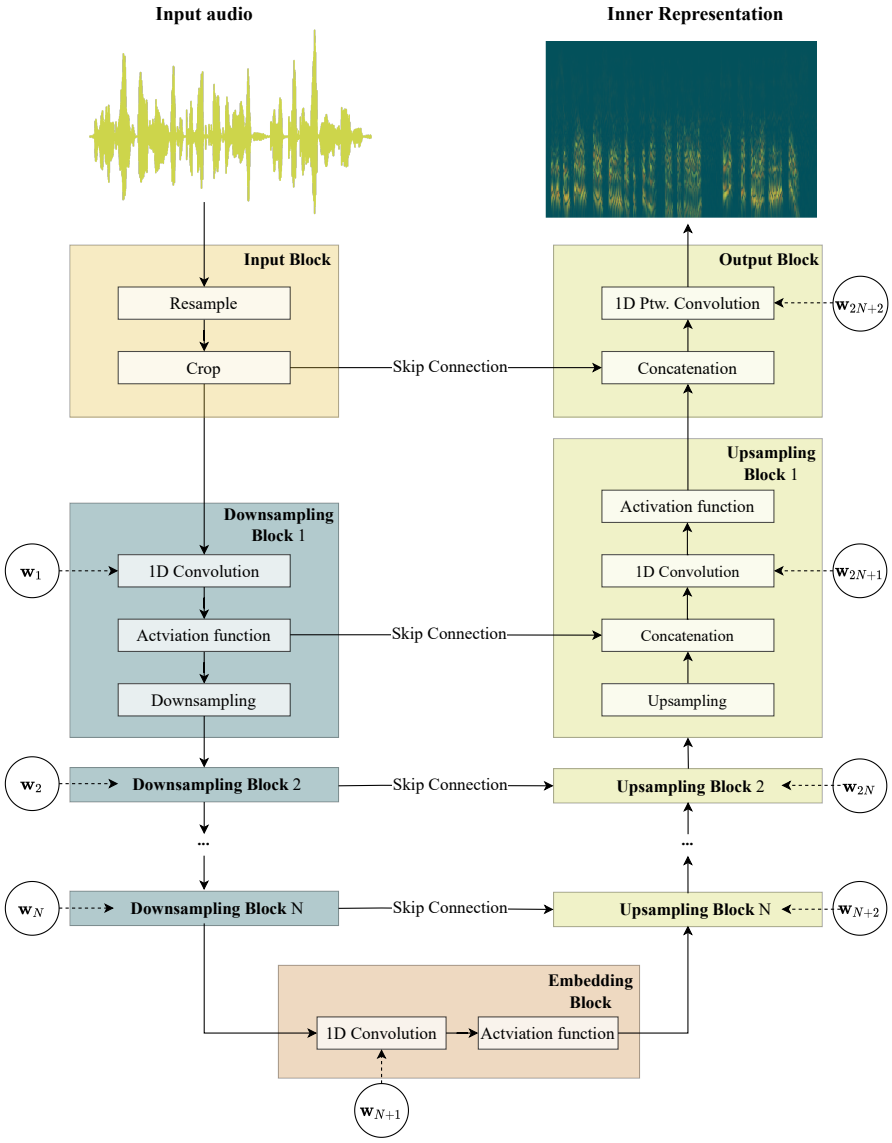


Figure 3: Overview of the auditory network simulator. The bold lines denote the signal path and the dashed lines denote the parameter path. The circles denote the weights generated by the Weight Generating Network (WGN).

are sampled at 20 kHz and segmented into windows of 2048 samples with 256 samples of temporal context on each side of the window. The AMS and WGN are trained jointly with respect to a scaled L1-loss function:

$$L(f_\theta(x), \hat{f}_\theta(x)) = \sum_{j=1}^{J} ||\beta_j f_{\theta,j}(x) - \hat{f}_{\theta,j}(x)||_1 \tag{2}$$

where j is the j-th channel of the inner representation and

$$\beta_j = \frac{1}{|T|} \sum_{(x, f_\theta(x)) \in T} \frac{1}{||f_{\theta,j}(x)||_1} \tag{3}$$

The weighting term β_j allows better optimization of the higher frequency channels, where the absolute energy is orders of magnitude lower than the lower frequency channels which affects the integration of the gradient

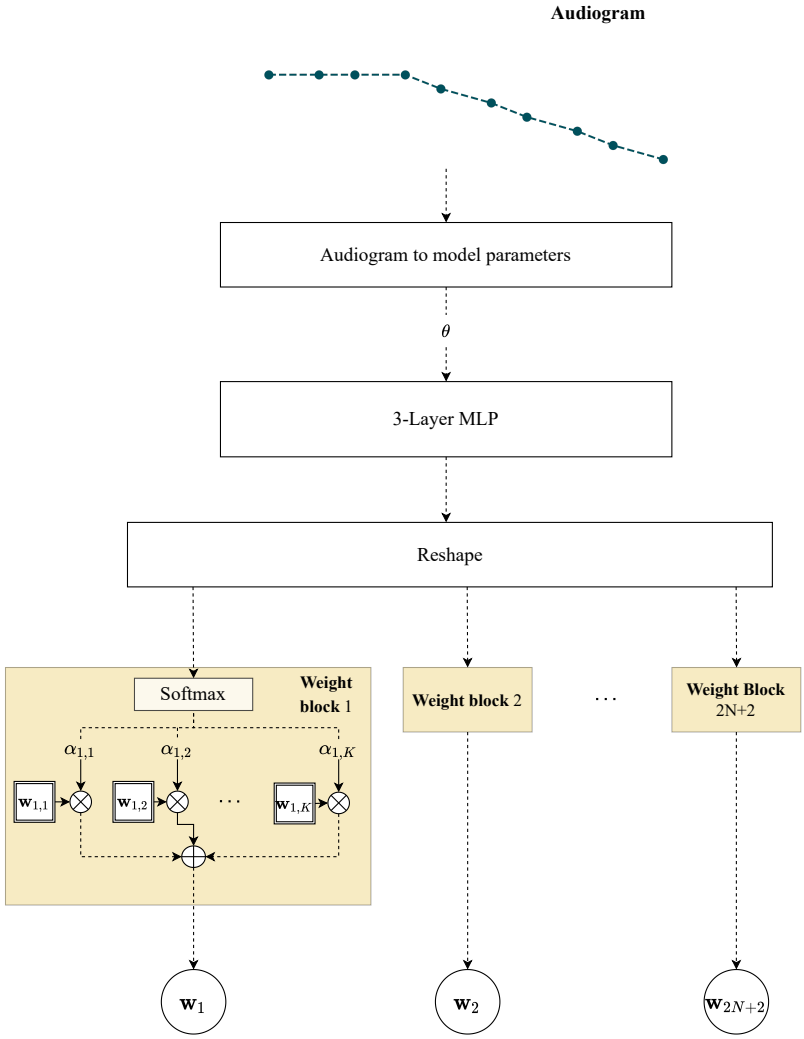


Figure 4: Overview of the Weight Generating Network. The outputs of this network is used as parameter inputs for the Auditory Model Simulator

with respect to the different channels during training. This weighting is especially important for simulating sloping hearing losses, since the shape of a conventional hearing loss exacerbates the effect. During inference, $\hat{f}_{\theta,j}$ is multiplied by the reciprocal of β_j. The ADAM [10] optimizer is used together with back propagation and gradient descent using a batch size of 256 and a learning rate of 0.0001. The network was trained for 75 epochs.

2.6. Model parameters

For the results in this paper we use the AMS and WGN structures given in Tables 1 and 2.

Table 1: Structure of the Auditory Model Simulator

N	8
Kernel size	21
Depth	160
Encoder activation	Tanh
Decoder activation	PReLU

Table 2: Structure of the Weight Generating Network

Layer 1 of MLP	(180×100)
Layer 2 of MLP	(100×100)
Layer 3 of MLP	(100×48)
Activation of MLP	Tanh
K	3

3 Results

In this section we will present the results in three different ways: (i) Through visualization of the inner representations, (ii) a comparison of the error between the training set and test set, and (iii) an error measure of the model across different audiograms and levels.

3.1. Visualizing the results

In Figure 5 we illustrate the inner representation of the ground truth (top row), our simulation (middle row) and the difference between the two (bottom row), for three different audiograms, for the same speech signal, x. From the figure it can be seen that the model fairly well approximates the general characteristics of the reference model, although discrepancies still exists.

3.2. Generalization to unseen speech

In order to test if the model can generalize to unseen speech, we compare the Mean Absolute Error (MAE) on a set of 100 audiograms and sentences of speech that was in the original training dataset with a testset consisting of the same audiograms but different sentences, with the speech scaled to the same level of the training set. The results are shown in Table 3. The test error is 7% larger than the training error, suggesting that the model is quite robust to unseen input signals.

Table 3: Training and test Mean Absolute Error (MAE) for the model, computed across different audiograms

	Train	Test
MAE	0.00081	0.00087

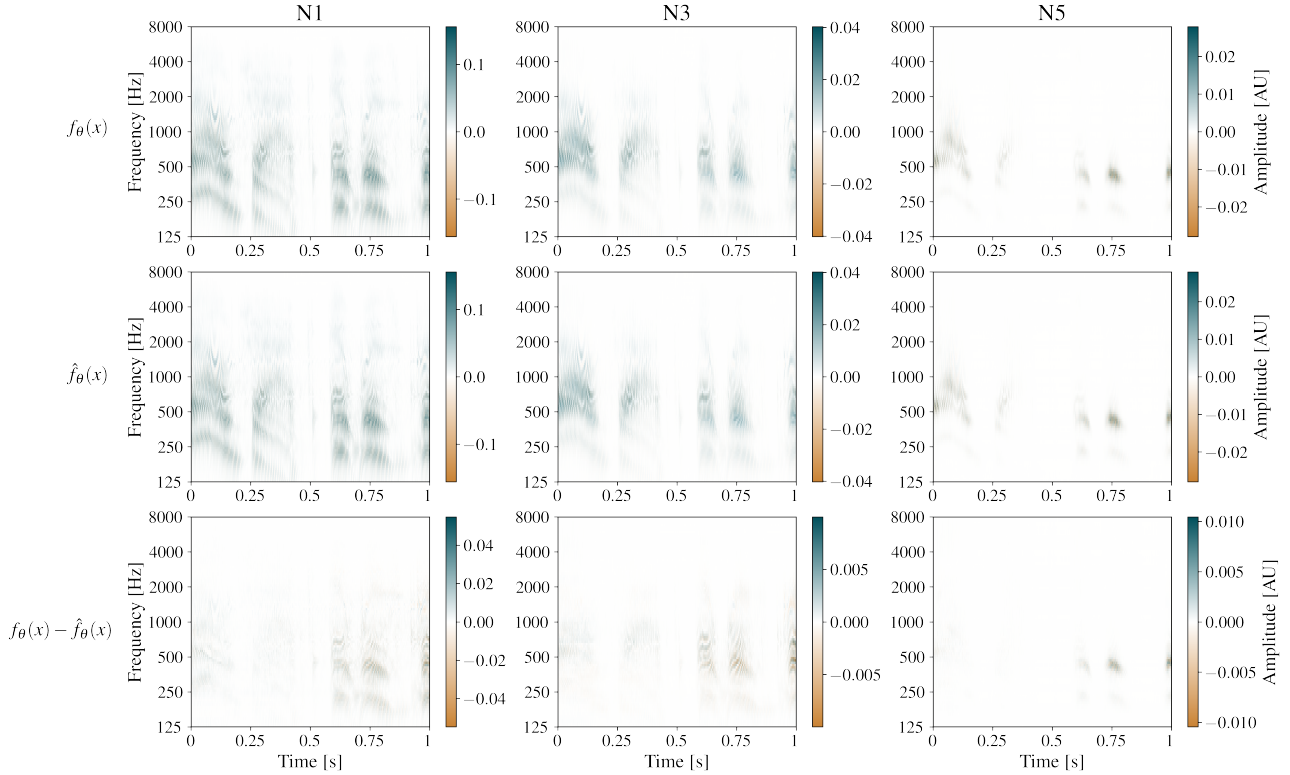


Figure 5: An example of the outputs of the Auditory Model Simulator conditioned on three different standard audiograms (N1,N3,N5) [9]. The top row is the ground truth, the middle row is the output of the AMS and the bottom row is the difference between the ground truth and our approach. The input is a speech signal that was not part of the training set, scaled to have an RMS value of 80 dB SPL. Notice different scales across audiograms and the bottom row to enhance visibility.

3.3. Testing error across audiograms and levels

In order to quantify the performance across audiograms and speech levels, we measure the error on the 10 standard audiograms [9] across 10 sentences that were not in the training set. We scale the absolute error, so that the errors are comparable across different levels and audiograms by introducing a Relative Mean Absolute Error (RMAE):

$$RMAE = \frac{|f_\theta(x) - \hat{f}_\theta(x)|_1}{|f_\theta(x)|_1} \tag{4}$$

From Table 4 it is clear that the error becomes large at low SPL levels relative to the level which was used to train the networks for audiograms with severe/profound broadband hearing losses, i.e N5-N7, meanwhile the error stays relative constant across different audiograms for the SPL levels on which the networks were trained. We note that N6 and N7 have identical errors, which is due to the parameters of the auditory model being identical for these two audiograms. It can be seen that the error is relatively large for the 60 dB SPL for the N5-N7 audiograms. Upon inspecting the output, we observe that the waveform looks to be the correct but has a large DC offset. This DC offset disappears for the 70 and 80 dB SPL condition. The DC offset could be due to low amplitudes and energy relative to the other audiograms and input levels resulting in a small effect on the gradient of the weights, and in particular the bias units, during training of the network and a higher sensitivity to small perturbations of the weights during inference. The significance of this error will largely depend on the application. We hypothesize that by using an architecture that discards the bias units, extending the training set to include additional input levels and modifying the loss function to accommodate the additional input levels,

that one can alleviate the problem of the DC offset.

Table 4: Relative Mean Absolute Error plus/minus the standard deviation for the 10 standard audiograms [9], measured across the same 10 sentences for each audiogram

Audiogram/Level dB[SPL]	60 dB	70 dB	80 dB
N1	0.23 ± 0.04	0.20 ± 0.03	0.17 ± 0.02
N2	0.21 ± 0.04	0.18 ± 0.03	0.16 ± 0.03
N3	0.21 ± 0.03	0.19 ± 0.03	0.19 ± 0.02
N4	0.36 ± 0.04	0.32 ± 0.03	0.31 ± 0.02
N5	1.7 ± 0.14	0.39 ± 0.04	0.21 ± 0.04
N6	1.84 ± 0.16	0.39 ± 0.04	0.21 ± 0.04
N7	1.84 ± 0.16	0.39 ± 0.04	0.21 ± 0.04
S1	0.23 ± 0.04	0.20 ± 0.03	0.17 ± 0.02
S2	0.20 ± 0.04	0.18 ± 0.03	0.16 ± 0.03
S3	0.27 ± 0.02	0.24 ± 0.02	0.23 ± 0.02

4 Conclusion

In this work we propose a neural network modelling framework that simulates an auditory model across its parameter space by processing the input and the parameters disjointly through two neural networks. The networks are trained on a set of speech and audiograms, and the model is shown to generalize well to unseen speech inputs. The error of the model is measured across a set of standard audiograms and input level, and is shown to have robust performance, except for a few cases where the model had a significant DC offset. Further work could investigate architectural improvements, parameter tuning and optimization techniques for this kind of model.

5 Acknowledgements

This work is partly supported by Innovation Fund Denmark Case no. 0153-00091B

References

[1] Deepak Baby, Arthur Van Den Broucke, and Sarah Verhulst. A convolutional neural-network model of human cochlear mechanics and filter tuning for real-time applications. *Nature machine intelligence*, 3(2):134, 2 2021. doi: 10.1038/S42256-020-00286-8. URL /pmc/articles/PMC7116797//pmc/articles/PMC7116797/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC7116797/.

[2] Anil Nagathil, Florian Göbel, Alexandru Nelus, and Ian C. Bruce. Computationally efficient DNN-based approximation of an auditory model for applications in speech processing. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2021-June:301–305, 2021. ISSN 15206149. doi: 10.1109/ICASSP39728.2021.9413993.

[3] Muhammad S. A. Zilany, Ian C. Bruce, and Laurel H. Carney. Updated parameters and expanded simulation options for a model of the auditory periphery. *The Journal of the Acoustical Society of America*, 2014. ISSN 0001-4966. doi: 10.1121/1.4837815.

[4] Auditory Models - Publications - Carney Lab - University of Rochester Medical Center. URL https://www.urmc.rochester.edu/labs/carney/publications-code/auditory-models.aspx.

[5] Daniel Stoller, Sebastian Ewert, and Simon Dixon. Wave-U-Net: A multi-scale neural network for end-to-end audio source separation. *Proceedings of the 19th International Society for Music Information Retrieval Conference, ISMIR 2018*, pages 334–340, 2018. doi: 10.5281/zenodo.1492417.

[6] Yinpeng Chen, Xiyang Dai, Mengchen Liu, Dongdong Chen, Lu Yuan, and Zicheng Liu. Dynamic Convolution: Attention over Convolution Kernels. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 11027–11036, 12 2019. ISSN 10636919. doi: 10.1109/CVPR42600.2020.01104. URL https://arxiv.org/abs/1912.03458v2.

[7] Brandon Yang, Google Brain, Gabriel Bender, Quoc V Le Google Brain, and Jiquan Ngiam. CondConv: Conditionally Parameterized Convolutions for Efficient Inference. URL https://github.com/tensorflow/tpu/tree/master/.

[8] Heiga Zen, Viet Dang, Rob Clark, Yu Zhang, Ron J. Weiss, Ye Jia, Zhifeng Chen, and Yonghui Wu. LibriTTS: A Corpus Derived from LibriSpeech for Text-to-Speech. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2019-September:1526–1530, 4 2019. ISSN 19909772. doi: 10.21437/Interspeech.2019-2441. URL https://arxiv.org/abs/1904.02882v1.

[9] Nikolai Bisgaard, Marcel S M G Vlaming, and Martin Dahlquist. Standard Audiograms for the IEC 60118-15 Measurement Procedure. *Trends in Amplification*, 14(2):113–120. doi: 10.1177/1084713810379609. URL http://tia.sagepub.com.

[10] Diederik P. Kingma and Jimmy Lei Ba. Adam: A Method for Stochastic Optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 12 2014. URL https://arxiv.org/abs/1412.6980v9.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Free-field correction values for RadioEar DD65 v2 Circumaural Audiometric Earphones

Christian Sejer Pedersen[1,*], **Rodrigo Ordonez**[1], **Dorte Hammershøi**[1]

[1]Department of Electronic Systems, Aalborg University, Aalborg, Denmark.

[*]cp@es.aau.dk

Abstract

Speech audiometry with earphones requires frequency dependent calibration based on free-field correction values for the particular earphone used. The free-field correction values are derived from a loudness matching experiment between free-field frontal incidence third-octave-band noise signals from a loudspeaker and the required earphone levels as measured on an IEC 60318-1:2009 acoustic coupler. This paper describes the determination and results of free-field correction values for RadioEar DD65 v2 circumaural audiometric earphones for all third-octave band frequencies from 125 Hz to 8 kHz, following the procedure described in the IEC 60268-7:2010 standard using 23 test subjects who also participated in the determination on Equivalent Threshold Sound Pressure Levels for the same earphones. The standard contains details on setting up the loudness comparison to have a controlled sound field and great care was taken to conform with this in the experimental setup. The standard also contains practical details on approximate stimulus and pause durations, and on presentation order; first loudspeaker then earphones. However, when it comes to the choice of psychometric method for obtaining loudness match i.e. how the level of the earphones is varied between presentations and how the final loudness match level is obtained it is just generally stating that the level of the earphones is varied until equal loudness is obtained. As this choice is up to the experimenter, care was taken to eliminate potential biases and a two-alternative forced choice paradigm was chosen. The results are compared to data from another laboratory and a similar type of earphones and possible explanations for minor differences are discussed.

Keywords: earphones, audiometry, equal loudness, calibration.

1 Introduction

Speech audiometric assessment with earphones requires frequency dependent calibration based on free-field correction values for the particular earphone used. The free-field correction values are derived from a loudness matching experiment between free-field frontal incidence third-octave-band noise signals from a loudspeaker and the required earphone levels as measured on an IEC 60318-1:2009 [1] acoustic coupler. When free-field correction values are available for a specific earphone it is possible to determine an overall calibration value for speech audiometers using spectral levels of a simulated speech signal with an overall level of 0 dB. The procedure is described in [2] and involves adding the free-field correction values to the spectral values of the simulated speech signal and calculating the total level of the combined third-octave bands.

Free-field correction values are reported in the literature for different audiometric earphones, for example [2, 9, 11] and following the tradition in the field, this paper describes the determination and results of free-field correction values for the RadioEar DD65 v2 circum-aural audiometric earphone. The free-field correction values are determined for all third-octave band frequencies from 125 Hz to 8 kHz, following the procedure described in the IEC 60268-7:2010 [3] standard, section 'Free-field comparison frequency response', using 23

test subjects who also participated in the Equivalent Threshold Sound Pressure Levels (ETSPL) study by [4]. The IEC 60268-7:2010 [3] standard contains details on setting up the loudness comparison to have a controlled sound field (e.g. minimum distance from loudspeaker is 2 m) and maximum permissible level variations and great care was taken to conform with this in the experimental setup. The standard also contain practical details on approx. stimulus and pause durations, and on presentation order; first loudspeaker then earphone. However, when it comes to choice of psychometric method for obtaining loudness match i.e. how the level of the earphone is varied between presentations and how the final loudness match level is obtained it is just generally stating that the level of the earphone is varied until equal loudness is obtained, and *"The adjustment may be performed by the test person or the test supervisor, or automatically under computer control"* (IEC 60268-7:2010 [3] page 25). As this choice is up to the experimenter, care was taken to eliminate potential biases and a two-alternative forced choice paradigm was chosen. The results are compared to data from another laboratory and a similar type of earphone, possible explanations for differences are discussed.

2 Method and material

2.1. Psychometric method for loudness matching

The point of subjective equality (PSE) corresponds to the earphone sound pressure level (SPL) (as measured in an IEC 60318-1:2009 [1] acoustic coupler) needed in order to be perceived equally loud as the free-field frontal incidence 70 dB SPL signal from the loudspeaker (as measured in the listening position with the subject absent). When performing loudness matching there is a level range around the PSE where the test subject may not be completely certain which sound is the louder, as described by the psychometric function. In this range around the PSE a test subject may, for the same presentation levels, at one time judge the earphone signal louder, but at another time the loudspeaker signal louder. The PSE is defined as the SPL of the earphone at the 50 % point on the psychometric function i.e. the SPL, where either earphone or loudspeaker signal is judged louder 50 % of the time.

Numerous methods for finding the PSE exist (see e.g. [5]), ranging from classical methods like constant stimuli method and method of adjustment, to adaptive methods like staircase method or sequential maximum likelihood estimation. For this study the PSE is found using an adaptive up/down staircase method ([6], [7]), which according to [5] is one of the most used methods of loudness matching. The method uses a two-alternative forced choice paradigm (2AFC), where the test subject must choose whether the earphone signal is softer or louder than the loudspeaker signal. This deals with potential bias from the test subject's internal criterion on how certain they are of "equal loudness", which would be a concern if the test procedure was simply to adjust the level until the subject reports that the earphone is equally loud as the loudspeaker signal.

As specified in IEC 60268-7:2010 [3] standard the level of the loudspeaker signal is fixed at 70 dB SPL while the earphone level is varied in steps, and as such the reference signal is always that from the loudspeaker. The loudness of the earphone signal is judged compared the that of the loudspeaker signal. The staircase method specifies that if the earphone signal is judged louder then it will be decreased in level on the next presentation, and if it is judged softer, then it is increased on the next presentation. When the direction in level is changed i.e. a level increase after a level decrease or visa versa, this is called a reversal. Consequently, the method will converge on the 50 % point on the psychometric function i.e. the PSE.

In order to make the method efficient it was chosen to starts with a larger step size of 4 dB in order to quickly reach the level range around the PSE and serve as a familiarization with the new pair of stimuli. After two reversals the stepsize is reduced to 2 dB in order to converge more precisely at the PSE. A stepsize of 1 dB was first considered, but from initial testing it was determined that it makes the task more difficult for the test subject, as it would increase the number of presentations with similar loudness. It would also increase the duration of the experiment without providing more accurate answers.

As specified in the IEC 60268-7:2010 [3] standard a pause of 2.5 s was inserted before the next free-field

presentation. This pause functioned as a natural reminder for the test subject that a new loudness comparison had initiated and that they should forget the previous presentation. The start level for the earphone signal was 70 dB ±5 dB by random, i.e. either 65 dB or 75 dB. This was done in order to prevent start levels too far from the expected PSE for efficiency and to minimize potential systematic bias in start point, i.e. always approaching the PSE from below could potentially cause a small bias. The stop criterion is set at 6 reversals. The mean value of the SPL of the last 4 reversal points is used as an estimate of the PSE. Additionally, for additional verification purpose a psychometric function is fitted to the data using maximum likelihood estimation.

2.2. Maximum likelihood estimation

In order to evaluate the robustness of the PSE obtained by using the average of the last four reversal points and the staircase method itself a maximum likelihood estimation (MLE) procedure is used to fit a theoretical curve to the data. As the cumulative normal distribution is often used to describe the shape of the psychometric function (see e.g. [8]) it is used in this study. The MLE procedure estimates parameters of the psychometric function by fitting a cumulative normal function to all the answers given by the subject during a loudness match (excluding the initial answers before two reversals which are considered familiarisation and therefore not included). The MLE procedure fits the cumulative normal distribution through several iterations of varying the two parameters of the cumulative normal distribution, mean μ and standard deviation σ until optimum values for the two parameters are found. The MLE of the mean is the 50 % point on the psychometric function i.e. the PSE.

2.3. Test setup

The system used for loudness matching consisted of: a Lenovo notebook running a custom made Matlab test program (Matlab 2018a with Audio System Toolbox using ASIO drivers), an external USB 24 bit sound card (RME Fireface UFC), a power amplifier (ROTEL RB-976 MKII, modified to give a constant 0 dB gain and lower noise), an active loudspeaker (Genelec 1031A), and the RadioEar DD65 v2 earphones.

2.4. Transducers

One pair of RadioEar DD65 v2 earphones were included in the tests. According to the IEC 60268-7:2010 [3] standard the left and right transducer should have a similar frequency response within 2 dB. This was fulfilled for the earphones and additionally each transducer was calibrated to give the same level in each ear (as measured on the IEC 60318-1:2009 [1] acoustic coupler - see section 2.8). The free-field sound is reproduced using an active loudspeaker, Genelec 1031A, that has a quite flat frequency response in the frequency range used. This is a requirement in the IEC 60268-7:2010 [3] standard, as it states in Annex C: *"The frequency response of the loudspeaker, measured with sinusoidal signals, should be free from sharp peaks and dips, so as to avoid errors due to colouration"*. Also the harmonic distortion of the loudspeaker at the playback levels of 70 dB SPL used in this study is measured to be well below the limits of 2 % stated in the standard.

2.5. Signals

Third-octave-band filtered pink noise were created in Matlab at the center frequencies from 125 Hz to 8 kHz (48 kHz sampling frequency) and stored as 32 bit wave-files. The duration of each signal was 2.7 s and start and stop shaping was made with half a Blackman-Harris window of 100 ms in duration which was sufficient to prevent clicks at the onset and offset of the signal. Thus, each tone had onset and offset ramps of 50 ms, giving a 2.6 second duration with constant amplitude. The signal level was controlled by digitally attenuating the signals before sending them to the output of the sound card. This was deemed sufficient and free of influence of quantization errors as only a small range of levels were required for the loudness matching and the lowest level needed was in the order of 50 dB from full scale leaving more than 15 bit for the lowest earphone levels on the 24-bit soundcard. The order of frequency was the same for all subjects (see Table 1 for the order in

each session 1-4). The order followed the guidelines from the IEC 60268-7:2010 [3] standard, by starting at 1 kHz and progressing up in frequency to the highest frequency (8 kHz), going down in frequency to the lowest frequency (125 Hz) and then increasing the frequency up to the 1 kHz band. The familiarization was done with 1 kHz and not included in the data.

Session 1	1000 Hz	1600 Hz	2000 Hz	5000 Hz	8000 Hz
Session 2	6300 Hz	4000 Hz	3150 Hz	2500 Hz	1250 Hz
Session 3	800 Hz	500 Hz	315 Hz	200 Hz	125 Hz
Session 4	160 Hz	250 Hz	400 Hz	630 Hz	1000 Hz

Table 1: The four sessions and the order of the frequency bands. The familiarisation round before session 1 used 1 kHz, and is not included in the data.

2.6. Response System

In typical loudness matching experiments the test subject would push a button corresponding to which sound was the loudest. However, in this experiment both hands of the subject are occupied with taking earphones on and off repeatedly. Therefore, it would be impractical to use a button test interface. Instead the test subject responded to the presentation by saying out loud whether the earphone sound was ”softer” or ”louder” than the loudspeaker sound. The experimenter would then click the corresponding button on the graphical user interface of the custom made Matlab test program that was also used to supervise the experiment. The Matlab test program loads the stimuli wave-files, adjust levels, presents the stimuli, collects user responses, plots the data in a figure and saves data to files. The signals are send through the sound card to the power amplifier (for earphone signals) or to the active loudspeaker.

2.7. Sound field calibration and verification

All measurements were carried out in the anechoic room of the Acoustics Laboratory of Aalborg University, which is anechoic down to approx. 63 Hz and has a very low background noise level (typically below 0 dB A-weighted SPL).

The reference sound field was calibrated and verified using a free-field microphone (B&K type 4165) connected to a microphone preamplifier (B&K type 2669) connected to a conditioning amplifier (B&K Nexus type 2690) which was connected to the input of the sound card. A custom made Matlab program was used to reproduce the calibration stimuli and record the signals from the microphone. Before the measurements the entire electroacoustic chain (from microphone to digital value) was calibrated using a 1000 Hz calibrator (B&K 4230). The recorded signals were compensated for the microphone sensitivity and conditioning amplifier gain, obtaining amplitude values in Pascals. The reference point is on axis of the loudspeaker at a distance of 2.48 m. The reference level of 70 dB is measured with the test subject absent. The homogeneity of the sound field was verified and within ± 1 dB at all positions 15 cm front/back/left/right/up/down from the reference listening position (see Fig 1). The reference level of 70 dB in the reference listening position was regularly verified by measurement during the days of the experiment.

2.8. Earphone calibration and verification

The SPL from the DD65 v2 (both left and right ear) were determined using an IEC 60318-1:2009 [1] acoustic coupler (B&K type 4153), mounted on the side of a cube so that the distance between each earphone cushion was 135 mm and the measured side was placed centred on the circum-aural earphone adapter plate of the coupler. The headband length was set to give 130 mm from the centre-line between cushion midpoints and the inside top of the headband, giving a coupling force of 10.3 ± 0.7 N. A pressure field microphone (B&K type 4134),

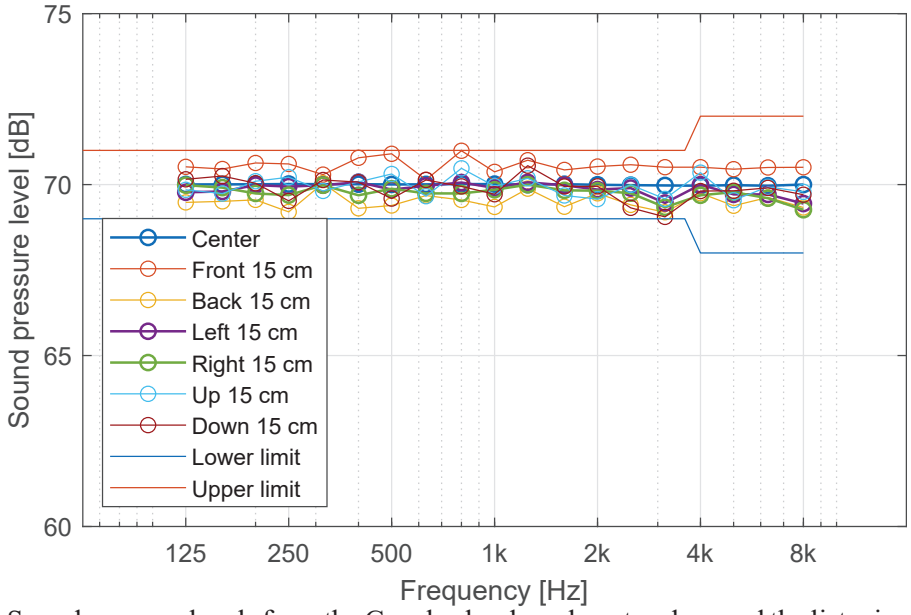


Figure 1: Sound pressure levels from the Genelec loudspeaker at and around the listening positions

microphone preamplifier (B&K type 2669) and conditioning amplifier (B&K Nexus type 2690), were connected to one of the analog inputs of the sound card. A Matlab program was used to reproduce the calibration stimuli and record the sound pressure level produced by the earphone in the acoustic coupler. Before the measurements the entire electroacoustic chain (from microphone to digital value) was calibrated using a 1000 Hz calibrator, B&K 4230. Measurement were made by reproducing different levels from the Matlab program and verifying the levels. The recorded signals were compensated for the microphone sensitivity and measurement amplifier gain, obtaining amplitude valued in Pascals. The earphone calibration for both left and right was regularly verified by measurement during the days of the experiment.

2.9. Fitting of the earphones

Spectacles and earrings were removed and hair-style was rearranged if it might influence the position of the earphones as described in the IEC 60268-7:2010 [3] standard. The earphones were fitted by the test subjects themselves under the supervision of the experimenter. They were instructed to make certain that the earphones rested properly and comfortably around the ears. The experimenter was in control of starting the sound reproduction in the earphones, and would not initiate this before the test subject's hands had left the earphones, which signalled that the earphones were sitting comfortably on the ears. While listening to the loudspeaker in the free field the test subjects held the earphones in their lap.

2.10. Subjects

The 25 normal-hearing and otologically normal test subjects from the ETSPL study [4] were invited to participate. Two test subjects withdrew giving a total of 23 subjects (12 male and 11 female). They were between 18 and 25 years of age. All subjects were students at Aalborg University and received remuneration for their participation.

2.11. Procedure

All subjects were given written instructions outside the anechoic room supported by oral instructions inside the anechoic room. The subject was positioned on the chair and the height was adjusted to get the entrance of the right ear canal in the reference position (marked by a red laser cross) and facing the loudspeaker (marked with red vertical laser on nose). The vertical loudspeaker laser was turned off during the experiment as to not disturb

the eyes of the test subject, but the laser cross at the ear was on during the test and the test subject position was continuously monitored and if necessary corrected during the experiment.

Test subjects were instructed to look at the loudspeaker and keep the head position during the experiment. When not listening to the earphones they would hold them in their lap. Following the instructions, a familiarization trial with one frequency band (1000 Hz) was made, before data collection began with session 1.

The experimenter was sitting inside the room at a considerable distance (3.7 m from reference position off-axis) as to not disturb the test subject or the sound field. From this position the experimenter had a good view of the earphones and the laser cross placed at the reference position which should hit the subject right ear canal entrance during free-field exposure from the loudspeaker.

During the experiment, the investigator could monitor the test progress from the Matlab controlled system. Four sessions of five center frequencies were used (see Table 1). Between sessions the subject had a break before proceeding with the next session. The whole test including familiarization and breaks lasted approx. 1 hour per subject.

3 Results

3.1. Free-field correction values

Table 2 shows the earphone level that gives equal loudness to a free-field level of 70 dB (frontal incidence), and Table 3 shows the corresponding free-field correction values. These are shown in Figure 2 and Figure 3 together with maximum likelihood estimates based on the individual answers after the two initial reversals.

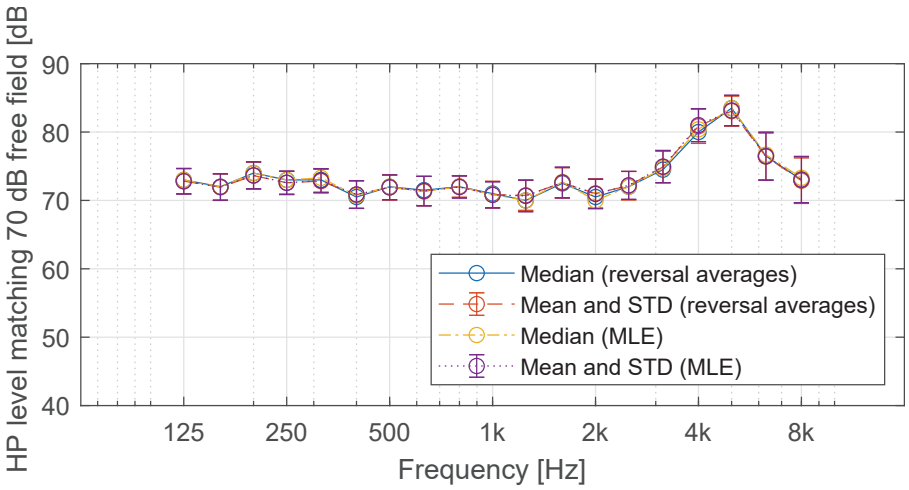


Figure 2: RadioEar DD65 v2 SPL needed to match the loudness of a 70 dB free field third-octave-band noise. Error bars show the standard deviation.

3.2. Determination of SPL of speech presented via earphones

Overall calibration of speech audiometers can be determined for the specific earphone using the free-field correction values and spectral levels of a simulated speech signal with an overall level of 0 dB. The procedure is described in [2] and involves adding the free-field correction values to the spectral values of the simulated speech signal and calculating the total level of the combined third-octave bands. The calculations are based on Impulse time weighting and a fast time weighting reduces the level by 0.5 dB as described by [9]. For the free-field correction values (median) in Table 3 this gives a level of 1.8 dB for impulsive and 1.3 dB for fast time weighting.

Freq. [Hz]	Median [dB]	Mean [dB]	STD [dB]	95% CI [dB]
125	73.0	72.8	1.9	0.8
160	72.0	71.9	1.9	0.8
200	74.0	73.6	2.0	0.8
250	73.0	72.6	1.7	0.7
315	73.0	72.8	1.7	0.8
400	70.5	70.8	2.0	0.9
500	72.0	71.9	1.9	0.8
630	71.5	71.4	2.2	0.9
800	72.0	72.0	1.6	0.7
1000	71.0	70.8	1.9	0.8
1250	70.0	70.7	2.2	0.9
1600	72.5	72.6	2.2	1.0
2000	70.5	71.0	2.1	0.9
2500	72.0	72.2	2.1	0.9
3150	74.5	74.9	2.3	1.0
4000	80.0	80.8	2.5	1.1
5000	83.5	83.1	2.2	1.0
6300	76.5	76.4	3.4	1.5
8000	73.0	72.9	3.3	1.4

Table 2: Earphone levels in coupler in order to loudness match a free-field frontal incidence level of 70 dB SPL reported as median, mean, standard deviation (STD) and 95% confidence interval (CI) for all frequencies

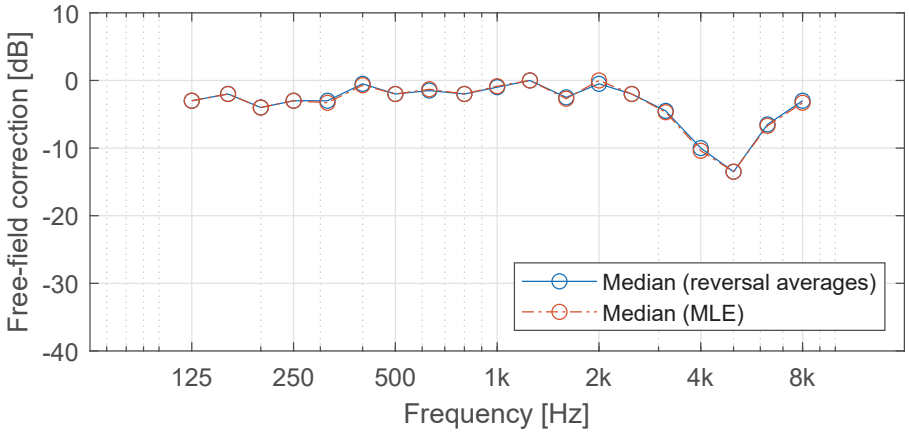


Figure 3: Free-field correction values for RadioEar DD65 v2.

4 Discussion

The median PSE based on the average of the last four reversals and the median PSE based on maximum likelihood estimates from individual answers after the two initial reversals agree with each other. This serves as a confirmation that the reversal averages from the stair-case method give a good estimate of the PSE and the

Freq. [Hz]	Median [dB]	Mean [dB]	STD [dB]	95% CI [dB]
125	-3.0	-2.8	1.9	0.8
160	-2.0	-1.9	1.9	0.8
200	-4.0	-3.6	2.0	0.8
250	-3.0	-2.6	1.7	0.7
315	-3.0	-2.8	1.7	0.8
400	-0.5	-0.8	2.0	0.9
500	-2.0	-1.9	1.9	0.8
630	-1.5	-1.4	2.2	0.9
800	-2.0	-2.0	1.6	0.7
1000	-1.0	-0.8	1.9	0.8
1250	0.0	-0.7	2.2	0.9
1600	-2.5	-2.6	2.2	1.0
2000	-0.5	-1.0	2.1	0.9
2500	-2.0	-2.2	2.1	0.9
3150	-4.5	-4.9	2.3	1.0
4000	-10.0	-10.8	2.5	1.1
5000	-13.5	-13.1	2.2	1.0
6300	-6.5	-6.4	3.4	1.5
8000	-3.0	-2.9	3.3	1.4

Table 3: Free field correction values reported as median, mean, standard deviation (STD) and 95% confidence interval (CI) for all frequencies

free-field correction values.

The free-field correction shows a frequency dependent correction, which is relatively close to 0 dB at the low-to medium frequencies. The difference in this frequency region can possibly be explained by the physical difference between real subject's head and ear and the artificial ear with plate of the IEC 60318-1:2009 [1] acoustic coupler. For example is it expected that there are differences between fit and internal volume, where leakages is mostly influencing the lowest frequencies. In addition, there is a perceptual difference between listening to an external frontal incidence plane wave from a visible loudspeaker and a diotic earphone presentation that will give an auditory event located in the center of the head. It can be hypothesized that this could potentially affect the perceived loudness. At the higher frequencies the pinna will amplify the frontal incidence plane wave, thereby requiring substantial higher levels from the earphone to obtain equal loudness. This is clearly seen in Figure 2 and the free-field correction values where the largest correction is at 5000 Hz, where a median of 13.5 dB increase in level of the earphone signal is required. The standard deviation is generally largest at the highest frequencies 6300 Hz and 8000 Hz, where individual differences in head, and pinna shape, and earphone fit is expected to have a significant impact on the levels at the eardrums for the free-field and earphone playback respectively. For the lower frequencies the standard deviation is in the order of 2 dB, which gives good confidence in the estimates and resulting in 95 % confidence intervals for the mean values below ± 1 dB in this frequency range.

The free-field correction for the RadioEar DD65 v2 has also been determined by another group of researchers [10][11] in another laboratory. A comparison between the results from this study and their results can be seen in Figure 4. Several differences in methodology exist, mainly: The loudspeaker level was not kept

at 70 dB SPL but at levels of equal loudness was chosen meaning that the level was at a level in the range 61-72 dB depending on frequency (here it should be mentioned that IEC 60268-7:2010 [3] standard suggests keeping it fairly constant ±5 dB and approx. 70 dB). A stimulus duration of 2 s (IEC 60268-7:2010 [3] suggests approx. 2.5 s) and an identical start level for earphone and loudspeaker was used. The psychometric method for finding the equal loudness is different as they used two series of 2 dB steps (up-down, down-up) after an initial 6 dB and 3 dB step size staircase procedure. The subjects indicated their responses using two-foot switches and presentations were repeated if no response was giving. Finally, they used only one mandatory break (as compared to three breaks in this study), which means that the test subjects had to perform during longer time periods. These methodological differences and the use of different equipment, test subjects, laboratory, sound field, etc. can easily explain the relatively small differences (up to 4.9 dB at 6.3 kHz) seen. Similar for all corrections are a dip at 5 kHz which corresponds to the amplification of the loudspeaker signal due to the pinna.

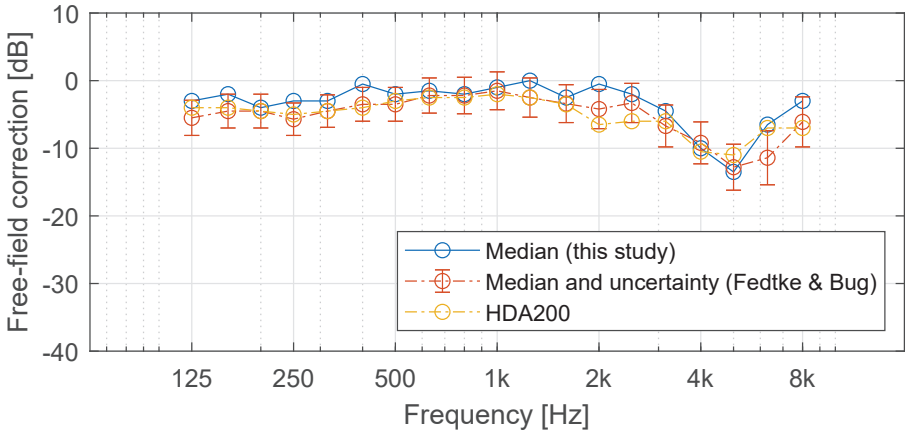


Figure 4: Free-field correction values for RadioEar DD65 v2 compared to values from another study [10][11]. For reference also HDA200 from [2] free-field correction values are plotted.

In the Equivalent Threshold Sound Pressure Levels (ETSPL) study by [4] the RadioEar DD65 v2 performs quite similar to the Sennheiser HDA200, which is also a circumaural audiometric earphone. A similarity trend can also be seen in the free-field correction values for the two earphones as seen in Figure 4.

Using a fixed presentation order with always presenting the loudspeaker signal before the earphone signal may potentially cause a systematic bias to the loudness matches, as there are known time-order effects in loudness judgements (see e.g. [5]). A common way to minimize the influence of these effects is to randomize the presentation order. However, this is not practical for the present study as randomizing whether or not the test subject should wear the earphones first would add a considerable complexity to the task of the test subject and experimenter. This is probably the reason why the presentation order is also fixed in the procedure described in the IEC 60268-7:2010 [3] standard.

Another potential source of bias is using a fixed presentation order of frequencies. It is possible that effects like carry-over effects and fatigue effects could have influenced the results in a systematic way. I.e. in the beginning the test subject is fresh and have full attention to the task, but in the end of the experiment attention may be lower, which could introduce more variation in responses and affect the accuracy of the results in the last session. Breaks between each session was used to minimize this potential effect. Going from a frequency with higher loudness to another frequency where 70 dB SPL has a lower loudness can potentially bias the initial loudness perception. However, this bias is expected be equal to both loudspeaker and earphone presentation, and as such not cause bias on the last number of reversals of a test run. In order to minimize the effect of any carry-over effects and fatigue effects etc. it would have been beneficial to randomize the frequency presentation order. However, this would not follow the guidelines in the IEC 60268-7:2010 [3] standard (see section 2.5).

5 Conclusion

Free-field correction values for the RadioEar DD65 v2 circum-aural audiometric earphones have been determined and the sound pressure level of speech presented via earphones have been determined. For overall calibration of speech audiometers this gives a level of 1.8 dB for impulse and 1.3 dB for fast time weighting. The results are similar to results from another laboratory and the small differences can be explained by methodological differences as well as different physical setups and subjects.

6 Acknowledgement

Funding and collaboration with Interacoustics A/S who also supplied the RadioEar DD65 v2 circum-aural audiometric earphones is gratefully appreciated. The authors would also like to thank all the test subjects for their participation in the present and accompanying study [4]. The authors also thank engineering assistant Claus Vestergaard Skipper for technical support and physical setup in the laboratory, and M. Sc. E. E. Oliver Erik Palmhøj Jokumsen for his assistance with running the experiments.

References

[1] IEC 60318-1. Electroacoustics – Simulators of human head and ear – Part 1: Ear simulator for the measurement of supra-aural and circumaural earphones. Technical report, International Electrotechnical Commission, Geneva, 2009.

[2] Klaus Brinkmann and Utz Richter. Ensuring reliability and comparability of speech audiometry in germany. In M. Martin, editor, *Speech audiometry*, chapter 6, pages 106–130. Whurr Publishers, 1997.

[3] IEC 60268-7. Sound system equipment - Part 7: Headphones and earphones. Technical report, International Electrotechnical Commission, Geneva, 2010.

[4] Rodrigo Ordoñez, Christian Sejer Pedersen, and Dorte Hammershøi. Equivalent threshold sound pressure levels (etspl) and insertion loss for the radioear dd65 v2 circum-aural audiometric earphones. In *EUROREGIO BNAM2022*, pages 1–10, Aalborg, Denmark, May 2022.

[5] Lawrence E. Marks and Mary Florentine. Measurement of loudness, part i: Methods, problems, and pitfalls. In Mary Florentine, Arthur N. Popper, and Richard R. Fay, editors, *Loudness*, chapter 2, pages 17–56. Springer-Verlag New York, 2011.

[6] W. J. Dixon and A. M. Mood. A method for obtaining and analyzing sensitivity data. *Journal of the American Statistical Association*, 43(241):109–126, 1948.

[7] Tom N. Cornsweet. The staircase-method in psychophysics. *The American Journal of Psychology*, 75(3): 485–491, 1962.

[8] Bernhard Treutwein. Adaptive psychophysical procedures. *Vision Research*, 35(17):2503–2522, 1995.

[9] Torben Poulsen. Free-field correction values for interacoustics dd 45 supra-aural audiometric earphones. *International Journal of Audiology*, 6(50):361–366, 2011.

[10] T. Fedtke and M. Bug. Equivalent threshold sound pressure levels for pure tones and difference between free-field sensitivity level and ear simulator sensitivity level (free-field equalization) for an audiometric headphone. Technical report, PTB, Braunschweig, Germany, 2018.

[11] Marion U. Bug and T. Fedtke. Free-field correction values for radioear dd65v2 circumaural audiometric headphones. *International Journal of Audiology*, pages 1–6, 2021.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Equivalent Threshold Sound Pressure Levels (ETSPL) and Insertion Loss for the RadioEar DD65 v2 Circum-aural Audiometric Earphones

Rodrigo Ordoñez[1,*], **Christian Sejer Pedersen**[1] **and Dorte Hammershøi**[1]

[1]Department of Electronic Systems, Aalborg University, Aalborg Denmark.

[*]rop@es.aau.dk

Abstract

Assessment of hearing thresholds requires proper calibration of the audiometric equipment that considers the sensitivity of the earphones used. For this purpose, the ISO 389 series defines the Reference Equivalent Threshold Sound Pressure Level (RETSPL), as the sound pressure level produced by the earphone actuated by the voltage corresponding to the median auditory threshold of a young, healthy population, measured in a standardized acoustic coupler. As such RETSPL depend on the type and sensitivity of a given transducer. This paper describes the assessment of the Equivalent Threshold Sound Pressure Levels (ETSPL) and the measurement of insertion loss of the RadioEar DD65 v2 audiometric earphones. Hearing thresholds were determined for twenty-nine test subjects at eleven audiometric frequencies from 125 to 8000 Hz according to the specifications given in ISO 389-9 and ISO 8253-1. Insertion loss of two sets of earphones was measured in third-octave bands between 50 and 16000 Hz using an acoustic test-fixture in a reverberant sound field according to ISO 4869-3. The results are reported as ETSPL measured in an acoustic coupler conforming with the specifications given in IEC 60318-1 and compared to reference data used as a basis for the RETSPL for circum-aural audiometric earphones, ISO 389-8. Results show that the auditory thresholds obtained with the RadioEar DD65 v2 circum-aural audiometric earphones are in close agreement with reported values for the Sennheiser HDA200 circum-aural audiometric earphones and the reference values given in ISO 389-8. The insertion loss shows minimum values at the low frequency range (below 160 Hz) of 5 dB, then it gradually increases up to 45 dB at 2000 Hz and above.

Keywords: Audiometry - Earphones - Insertion Loss - Calibration

1 Introduction

Reliable audiometric assessment requires earphones to be calibrated according to a standardized reference, such as the reference equivalent threshold sound pressure levels (RETSPLs) of the ISO 389 series. The RETSPL for a given audiometric earphone type are derived from a number of independent studies of equivalent threshold sound pressure levels (ETSPLs) from normal hearing subjects at different laboratories.

ETSPLs express threshold values as sound pressure level (SPL), decibels (dB) with reference to 20μPa, or dB SPL. By relating the voltage applied to earphones at threshold to the SPL produced by the earphone measured in a standardized calibration fixture or artificial ear, as described in IEC 60318-1 [4], it is possible to compare hearing thresholds assessed with different equipment as calibrated hearing levels (HL). 0 dB HL represents the median hearing threshold of a normal hearing population.

ETSPLs are reported in literature for different earphones [for example, 6, 7, 8, 9, 10, 11, 12, 13], signal types [for example, 14, 15, 16, 17] and extended high-frequency ranges [for example, 18, 19]. Following the tradition in the field, this article describes the determination of pure-tone ETSPLs, and insertion loss for the RadioEar DD65 v2

circum-aural audiometric earphones in accordance with the guidelines of the ISO 389-9 [1] and ISO 4869-3 [3] standards. Results are presented in view of comparable data.

2 Materials and Methods

2.1. Electro-acoustic system

The system used for threshold measurements and for calibration consisted of: a computer running a MATLAB program, two programmable attenuators (PA5, Tucker Davis Technology), an external USB full duplex sound card (RME Fireface UFX 2), a power amplifier (ROTEL RB-976 MKII, modified to give a constant 0 dB gain and low noise), a custom-made earphone-measurement interface with constant attenuation (41.6 dB with output impedance of 1 Ohm), a response button, and the RadioEar DD65 v2 earphone (Part. 8510117 – WTG00009). Only one set of earphones (two transducers, left and right) were used in the threshold measurements, that is, all left ears were measured with the same left transducer (blue colour) and all right ears were measured with the same right transducer (red colour).

During the threshold experiments, the signal level was controlled by the programmable attenuators, where each channel of the earphone was controlled by a different attenuator. Each attenuator covers a range from 0.0 to 120.0 dB and can be altered in 0.1 dB increments. The MATLAB program generates and presents the stimuli, collects user responses, controls the programmable attenuators and saves data to files. The stimuli, in the program, is generated as a full scale digital signal (48 kHz sampling frequency and 24 bits) and is sent through the sound card with 0 dB gain. The output of the sound card is fed through the programmable attenuators which in turn connect to the amplifier and attenuation box that powers the earphones.

1.1 Coupler measurements

The ETSPLs for the DD65 v2 were determined using an IEC 60318-1 [4] acoustic coupler (B&K type 4153), mounted on the side of a cube so that the distance between each earphone cushion was 135 mm with the measured side placed centred on the circum-aural earphone adapter plate of the coupler. The headband length was set to give 130 mm from the centre-line between cushion midpoints and the inside top of the headband, giving a coupling force of 10.3 $\pm$0.7 N. A pressure field microphone (B&K type 4134), microphone preamplifier (B&K type 2639) and measurement amplifier (B&K type 2636), were connected to one of the analog inputs of the sound card. The MATLAB program was used to reproduce the calibration stimuli and record the sound pressure level produced by the earphone in the acoustic coupler. Before the measurements, the entire electroacoustic chain (from microphone to digital value) was calibrated using a 1000 Hz calibrator, B&K type 4230. Measurements were made by setting the programmable attenuators to 0 dB and reproducing full scale tones from the MATLAB program. The recorded signals were compensated for the microphone sensitivity and measurement amplifier gain, obtaining amplitude values in Pascals.

In order to avoid high attenuation levels from the programmable attenuators, the maximum output level of the system was set below 80 dB SPL. The minimum RETSPL for this type of earphone is 2.5 dB at 3 kHz according to ISO 389-8 [5], thus the attenuation needed by the system would be in the order of 80 dB.

1.2 Earphone Insertion Loss

Measurements of insertion loss for the set of earphones used in the experiment were carried out according to the ISO 4869-3 [3] standard. The measurements were carried out in the reverberant room of the Acoustic Laboratory at Aalborg University, that complies with the diffuse field requirements stated in ISO 3741 [20]. The system used for the measurements consisted of: a computer running a MATLAB program, an external USB full duplex sound card (Roland Quad-Capture), a power amplifier (Pioner 616), an acoustic test fixture (GRAS type

Table 1: Test subjects age and gender distribution

Age	20	21	22	23	24	25	**Total**
Male	1	2	0	5	3	4	15
Female	3	3	3	3	2	0	14
Total	4	5	3	8	5	4	29

45CA) configured with two GRAS 40AD 1/2" pressure microphones mounted without an ear simulator, two loudspeakers (Electro-Voice S200) and a calibrated hand held sound level meter (Brüel & Kjær type 2270).

The sound field was produced by the two loudspeakers reproducing uncorrelated pink noise placed at different locations in the room. The sound pressure level at the location of the test fixture in the room with the sources active was measured with the sound level meter at 1 kHz to a level of 87.8 dB SPL.

Sound pressure level in dB at 1/3-octave bands from 50 Hz to 16 kHz were measured with the test fixture without the earphones and the sound sources on (Test Fixture), the sound sources off (Noise Floor), and with the earphones mounted on the test fixture and the sound sources on (Earphones). Measurements were repeated 6 times repositioning the earphones on the text fixture each time.

2.2. Threshold determination

2.1 Subjects

29 test subjects (15 male and 14 female) participated. Test subjects were between 20 and 25 years of age, see Table 1. All subjects were students at Aalborg University and received remuneration for their participation. Subject selection was made following the guidelines and questionnaire given in ISO 389-9 [1]. Subjects willing to participate in the experiment responded the questionnaire and were selected based on their previous hearing history. All subjects that indicated undue sound exposure (noisy working environment, loud music listening or playing, etc.) or previous hearing problems (drain operation, chronic middle ear infection, etc.) were not included in the test. Selected subjects were then screened with otoscopy and tympanometry. Only subjects that had an unobstructed ear canal (no excessive ear wax), a visible tympanic membrane, and middle ear pressure between $\pm$ 50 daPa, were included in the measurements. According to subject selection criteria from ISO 389-9 [1], no exclusion was made in terms of hearing levels and all results are reported here. From the 29 subjects that participated in the experiments results from one left ear were not registered, due to technical issues. Giving a total of 29 right ears and 28 left ears in the data set.

2.2 Procedure

For each subject ETSPLs were calculated by relating the programmable attenuator setting at threshold to the sound pressure levels measured in a earphone coupler at 0 dB attenuation.

The thresholds were determined using the ascending method complying with ISO 8253-1 [2] using a step size of 5 dB. The method was fully automated and controlled by a computer, supervised by the operator that monitored the progress of the test from the adjacent control room. The first stimuli was presented at 30 dB above the RETSPL for the Interacoustics DD45 earphones reported by [10]. The only exception was for the frequency of 125 Hz, where the maximum output level was below the desired start level. For that frequency the maximum output level of 68 dB was used as the starting level. The general principle of the ascending method is as follows: If a valid response is obtained, the level is reduced by 10 dB and a new presentation is made at that level. If no valid response is obtained, the level is increased by 5 dB and a new presentation is made at that level. This ascending series continues until a valid response is obtained. The threshold is defined as the level where valid responses were obtained three times out of a maximum of five ascents. If a threshold was not obtained within

these five ascents a "second" try was performed. If a valid threshold was not obtained in the second try, threshold determination for that frequency was repeated at the end of the test.

Whether the left or the right ear was measured first was balanced across subjects. The sequence of frequencies followed the order stated in ISO 8253-1 [2] and the 1000 Hz tone was repeated at the end. The other ear was measured after a 5 minutes break following the first ear. The complete threshold determination (for all frequencies and for both ears) lasted between 20 and 30 minutes.

2.3 Signals

Pure tones needed for the hearing threshold measurements were created in MATLAB at the frequencies of interest complying with the requirements stated in ISO 8253-1 [2]. Each tone was 1.1 s in duration and the start and stop shaping was made with half a Blackman-Harris window of 100 ms in duration. Thus, each tone had onset and offset ramps of 50 ms, giving a 1 second section with constant amplitude. The inter-stimulus intervals were random in lengths varying from 1 to 3.3 s.

2.4 Test Site

All measurements were carried out in the audiometric room of the Acoustics Laboratory at Aalborg University. This room complies with the requirement of minimal background noise as stated in ISO 8253-1 [2]. An intercom system and a camera were used to communicate and monitor the subjects from the control room (adjacent to the audiometric room). Additionally the MATLAB program running the experiment displayed the presented levels and the subjects' responses, allowing the experimenter to follow the progress of the experiment. The earphones were fitted by the test subjects themselves under the supervision of the facilitator, all subjects were asked to remove glasses and earrings (if any) prior to fitting the earphones.

2.5 Response System

The audiometric room was equipped with a push button connected to the computer. In order to indicate that a tone was heard, the test subject pressed and released the push button immediately. Responses were only considered valid if the subject's response was registered while the test tone was playing or up to 0.5 s after the end of the tone. Otherwise the response was marked as not valid.

2.6 Instructions and familiarisation

All subjects were given verbal instructions on how to operate the response button and how to correctly fit the earphones. Following the instructions, a familiarisation trial with three frequencies (1000, 4000, and 500 Hz, presented in this order) was made, before data collection began.

3 Results

3.1. Equivalent Threshold Sound Pressure Level

Tables 2 show an overview of the audiometric data in dB ETSPL measured in the acoustic coupler that complies with IEC 60318-1 [4]. The table shows results for all ears, for left and right ears, for female and male subjects as well as for subjects between 20 and 22 yeas of age and subjects between 23 and 25 years of age.

The final ETSPL values are given rounded to the nearest half decibel and are shown in Table 3.

According to the general guidelines for expression of measurement uncertainty, and using the recommended model outlined in Annex A from ISO 8253-1 [2], the expanded measurement uncertainty is calculated to be

Table 2: Equivalent threshold sound pressure level (ETSPL) for RadioEar DD65 v2 earphone and an acoustic coupler that complies with IEC 60318-1 [4]. Data is shown as the median, average and standard deviation for all ears, each side, gender and age groups.

Frequency [Hz]	125	250	500	750	1000	1500	2000	3000	4000	6000	8000	N
All median [dB]	28.5	15.5	6.8	4.9	4.3	2.1	1.9	2.5	9.6	19.1	21.4	57
All mean [dB]	30.8	17.4	8.8	4.5	5.0	3.6	2.4	2.7	9.0	19.0	22.4	57
All Std. Dev. [dB]	5.6	5.1	6.9	6.4	5.5	6.6	5.9	6.4	6.0	6.8	8.2	57
Left median [dB]	28.5	15.5	6.8	5.0	4.3	1.5	1.9	1.9	9.6	19.1	21.4	28
Left mean [dB]	31.5	18.0	9.5	4.5	4.0	2.3	2.3	2.6	8.7	18.4	21.6	28
Left Std. Dev. [dB]	5.3	5.9	7.5	7.1	5.4	6.1	6.5	7.4	5.9	6.6	9.2	28
Right median [dB]	27.7	14.7	6.3	4.9	5.1	2.1	3.1	2.5	9.6	20.2	26.1	29
Right mean [dB]	30.1	16.9	8.2	4.6	6.0	4.7	2.5	2.7	9.4	19.7	23.1	29
Right Std. Dev. [dB]	5.8	4.3	6.3	5.7	5.5	6.9	5.4	5.3	6.0	7.1	7.3	29
Female median [dB]	28.5	15.5	6.3	4.9	4.3	2.1	1.9	1.9	4.6	19.1	21.1	27
Female mean [dB]	31.4	16.9	7.1	2.6	3.6	2.7	1.8	1.3	7.5	17.4	21.6	27
Female Std. Dev. [dB]	5.2	4.6	4.9	3.8	4.2	5.8	5.1	4.9	5.4	5.4	6.7	27
Male median [dB]	28.5	15.5	9.1	5.0	5.1	2.1	1.9	2.5	9.6	19.1	21.4	30
Male mean [dB]	30.2	17.9	10.4	6.3	6.3	4.3	3.0	3.9	10.4	20.5	23.1	30
Male Std. Dev. [dB]	5.9	5.6	8.1	7.6	6.2	7.2	6.6	7.3	6.2	7.7	9.4	30
20-22 years median [dB]	32.7	15.5	6.8	4.9	4.3	1.8	3.1	2.5	9.6	19.6	23.6	26
20-22 years mean [dB]	31.7	16.6	7.3	4.2	4.5	3.5	2.7	3.4	9.6	19.1	22.6	26
20-22 years Std. Dev. [dB]	5.6	4.9	4.7	3.7	3.3	7.4	5.6	6.4	5.7	5.1	9.3	26
23-25 years median [dB]	28.5	15.5	11.3	0.0	5.1	2.1	1.9	2.5	9.6	19.1	21.4	31
23-25 years mean [dB]	30.0	18.1	10.1	4.8	5.4	3.6	2.2	2.1	8.6	19.0	22.2	31
23-25 years Std. Dev. [dB]	5.4	5.3	8.2	8.0	6.8	5.9	6.3	6.3	6.2	8.1	7.4	31

Table 3: Equivalent threshold sound pressure level (ETSPL) for RealEar DD65v2 earphone and an acoustic coupler that complies with IEC 60318-1 [4], rounded to the nearest 0.5 dB.

Frequency [Hz]	125	250	500	750	1000	1500	2000	3000	4000	6000	8000
All Median [dB]	28.5	15.5	7.0	5.0	4.5	2.0	2.0	2.5	9.5	19.0	21.5

7.2 dB for frequencies below and up to 4 kHz and 10.4 dB for frequencies above 4 kHz. This measurement uncertainty takes into account standardised uncertainty values given in ISO 8253-1 [2] for: 1) variations observed in repeated threshold determinations due to the subjects' criteria; 2) variations to calculated threshold level due to equipment calibration and methodological parameters; and 3) variations in presented sound pressure level due to differences in fitting of the earphones.

3.2. Insertion loss

Figure 1 shows the average SPL measured with the acoustic test fixture in the reverberant sound field with and without the earphones mounted on the test fixture and the sound sources active, and the background noise in the room. Table 4 shows the calculated insertion loss for each cup of the earphone used in the experiment. The measurement uncertainty was calculated using the standard deviation for the repeated measurement of the open and occluded test fixture, and tabulated values (Appendix B in ISO 4869-3 [3]) for: uncertainties associated to the sound field, equipment and acoustic test fixture. The expanded uncertainty of the insertion loss measurements reported here is calculated to be 1.9 dB.

Table 4: Average insertion loss [dB] in 1/3-octave bands calculated as the level difference between the SPL at the test fixture microphones without and with the earphones mounted on the test fixture, and the sound sources active (Columns 2 and 3). Standard deviation over 6 repetitions of the measurements (Columns 4 and 5).

	Insertion Loss		Standard Deviation	
Freq. [kHz]	L [dB]	R [dB]	L [dB]	R [dB]
0.05	6.8	5.8	1.8	1.7
0.063	5.4	4.4	1.5	1.3
0.08	5.5	4.7	2.5	2.3
0.1	8.5	6.7	0.9	1.2
0.125	9.6	8.2	0.9	1.1
0.16	9.0	9.5	0.7	0.6
0.2	11.3	12.0	0.5	0.4
0.25	15.4	15.7	0.4	0.2
0.315	19.1	19.3	0.6	0.6
0.4	23.2	23.4	0.3	0.4
0.5	26.2	25.8	0.3	0.3
0.63	28.4	27.6	0.2	0.3
0.8	29.2	26.2	0.4	0.2
1	32.1	31.8	0.3	0.2
1.25	29.7	30.6	0.1	0.2
1.6	34.5	35.1	0.5	0.2
2	42.8	46.8	0.5	0.6
2.5	44.4	48.9	0.9	0.6
3.15	41.0	42.7	0.3	0.4
4	43.5	44.3	0.6	0.4
5	46.1	46.1	1.1	0.9
6.3	45.0	46.6	1.1	0.7
8	45.4	45.8	1.2	0.7
10	42.9	43.2	1.0	0.4
12.5	44.0	44.5	0.5	0.4
16	43.9	46.0	1.5	1.1

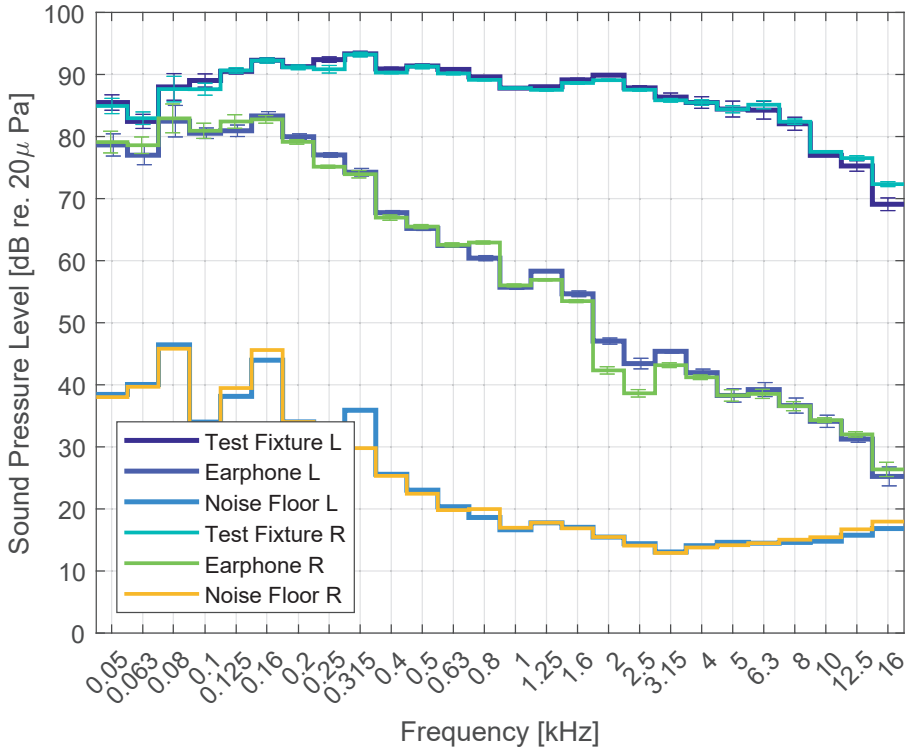


Figure 1: Average sound pressure level (SPL) [dB re. 20μPa] in 1/3-octave bands measured in the reverberant room with the text fixture microphones, with the sound sources active and no earphones (Test Fixture); with the sound sources active and the earphones mounted on the test fixture (Earphone); and with the sound sources disabled and no earphones (Noise Floor). Error-bars show $\pm$ standard deviation between repeated measurements.

4 Discussion

4.1. Threshold determination method

Thresholds were determined using the ascending method, defining the threshold as the level heard by the subject in three out of a maximum of five ascending tracks according to ISO 8253-1 [2]. This method has the inherent bias of approaching the threshold from below, leading to an overestimation of the threshold compared to the classical definition: the 50% detection of the psychometric function. This is because the threshold is estimated using only the last level the subject hears. This method has the advantage that it does not require complex calculations, and it is easily implemented with manual audiometers readily available in clinics. The problem is that it makes comparison with thresholds obtained with other methods that aim at a threshold around the 50% detection problematic [21, p. 24–25]. In order to illustrate this, maximum likelihood estimates (MLE) of the thresholds for each ear were calculated using all levels and answers given by the subjects after the initial familiarisation descend of each trial. This includes all levels presented to the subjects after the first "not heard" answer. The thresholds estimated with the MLE method represent the 50% detection probability of the estimated psychometric function modelled as a cumulative normal distribution. The MLE threshold estimates are significantly lower than the ones obtained using three ascents at the same level, for all frequencies tested ($t_{df=56} > 8, p < 0.8 \times 10^{-11}$). The average difference between the methods is of 3.7 dB with a minimum of 1.8 dB and a maximum of 5.6 dB. A comparison of the average estimate across subjects for both methods is shown in Figure 2.

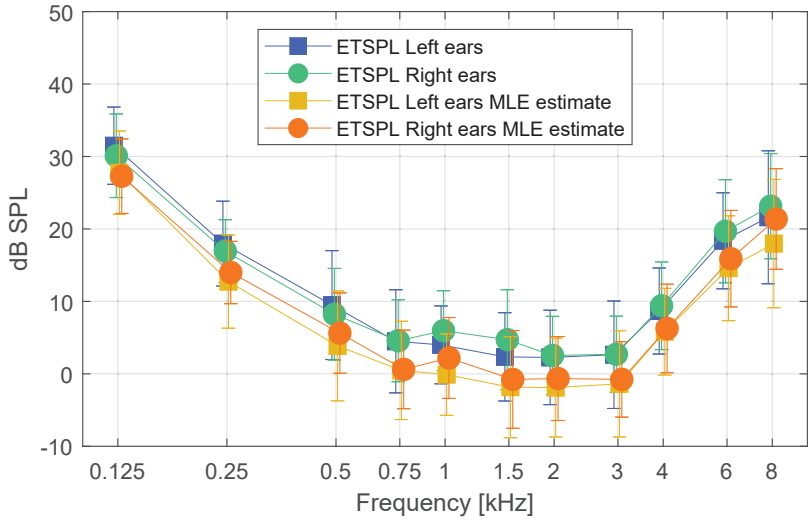


Figure 2: Average Equivalent Threshold Sound Pressure Levels (ETSPL) for DD65v2 earphones for left and right ears. Determined using the ascending method's stop criteria (light grey symbols); and using a Maximum Likelihood Estimate (MLE) procedure (dark grey and black symbols). Error-bars represent ± standard deviation. Symbols have been shifted in frequency for clarity.

4.2. Reference data

Figure 3 shows the present data with the data of five other studies reporting ETSPLs for circum-aural earphones. The first four studies serve as part of the reference material for ISO 389-8 [5] reporting ETSPL values for the Sennheiser HDA 200 earphones [6, 22, 23, 8]. The last study reports ETSPL values for RadioEar DD65 v2 earphones [24], the same type used here. These results confirm that the populations of subjects used in this

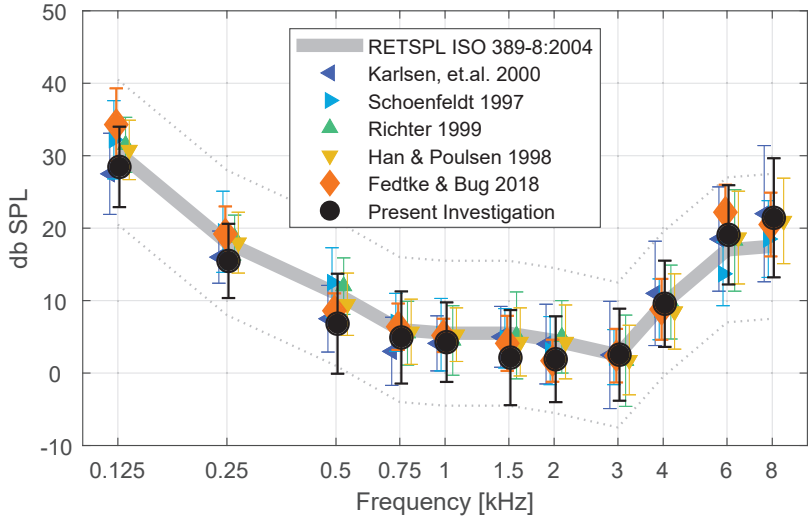


Figure 3: Median Equivalent Threshold Sound Pressure Levels (ETSPL) reported as reference data for ISO 389-8 [5], using Sennheiser HDA 200 earphones (grey symbols); and using RadioEar DD65 v2 earphones (dark grey and black symbols). Error-bars represent ± standard deviation. Symbols have been shifted in frequency for clarity. Also shown in the figure is the RETSPL from ISO 389-8 [5] (thich line) and ±10 dB (dotted lines).

experiment can be considered of normal hearing, and the equivalent threshold sound pressure levels presented here are in good agreement with reference data given in ISO 389-8 [5].

Equivalent threshold sound pressure levels, ETSPLs, for the RadioEar DD65 v2 circum-aural audiometric earphones are similar to the ETSPLs for Sennheiser HDA200 circum-aural audiometric earphones, and should be considered as having the same RETSPL for the calibration of audiometric equipment.

Acknowledgements

The authors would like to thank Claus Vestergaard Skipper for his technical assistance with equipment and laboratory facilities, as well as Oliver Erik Palmhøj Jokumsen for his assistance with data collection. Finally, thanks to all the subjects that participated in this experiment.

References

[1] ISO 389-9. Acoustics – Reference zero for the calibration of audiometric equipment – Part 9: Preferred test conditions for the determination of reference hearing threshold levels. Technical report, International Organization for Standardization, Geneva, 2009.

[2] ISO 8253-1. Acoustics – audiometric test methods – part 1: Pure-tone air and bone conduction audiometry. Technical report, International Organization for Standardization, Geneva, 2010.

[3] ISO 4869-3. Acoustics – Hearing Protectors – Measurement of insertion loss of ear-muff type protectors using an acoustic test fixture. Technical report, International Organization for Standardiazation, Geneva, 2006.

[4] IEC 60318-1. Electroacoustics – Simulators of human head and ear – Part 1: Ear simulator for the measurement of supra-aural and circumaural earphones. Technical report, International Electrotechnical Commission, Geneva, 2009.

[5] ISO 389-8. Acoustics - Reference zero for the calibration of audiometric equipment - Part 8: Reference equivalent threshold sound pressure levels for pure tones and circum-aural earphones. Technical report, International Organization for Standardization, Geneva, 2004.

[6] L.A. Han and T. Poulsen. Equivalent threhsold sound pressure levels for the sennheiser HDA 200 headphone and the etymotic research ER-2 insertearphone in the frequency range 125 Hz to 16 kHz. *Scan. Audiol.*, 2:105–112, 1998. doi: 10.1080/010503998420342.

[7] T. Poulsen. Reference thresholds for Eartone-3A insert earphones. *Scan Audiol*, 20(3):205–207, 1991. doi: 10.3109/01050399109074955.

[8] B. L. Karlsen, M. Lydolf, and A. O. Santillán. Audiometric tests with various earphones, signals of short duration and signals of high frequencies. Technical report, ISO/TC 43/WG 1, Doc. Nr. 309, 2000.

[9] T. Poulsen and S. Oakley. Equivalent threshold sound pressure levels (ETSPL) for Sennheiser HDA 280 supra-aural audiometric earphones in the frequency range 125 Hz to 8000 Hz. *Int J Audiol*, 48(5):271–276, 2009. doi: 10.1080/14992020902788982.

[10] T. Poulsen. Equivalent Threshold Sound Pressure Levels (ETSPL) for Interacoustics DD 45 supra-aural audiometric earphones. *Int. J. Audiol.*, 11:850–855, 2010. doi: 10.3109/14992027.2010.500625.

[11] M. Zebian, J. Hensel, T. Fedtke, and S Vollbort. Equivalent hearing threshold levels for the etymotic research er-10c otoacoustic emission probe. *Int JAudiol*, 51(7):564–568, 2012. doi: 10.3109/14992027.2012.669050.

[12] Torben Poulsen. Equivalent threshold sound pressure levels (ETSPL) for Sennheiser HDA 280 supra-aural audiometric earphones in the frequency range 125 Hz to 8000 Hz (vol 48, pg 271, 2009). *Int J Audiol*, 53 (4):287–288, 2014. doi: 10.3109/14992027.2013.860559.

[13] Cheng-Yu Ho, Pei-Chun Li, and Shuenn-Tsong Young. Reference equivalent threshold sound pressure levels for Apple EarPods. *J Acous Soc Am*, 141(2):115–119, 2017. doi: 10.1121/1.4976110.

[14] U. Richter and T. Fedtke. IReference zero for the calibration of audiometric equipment using 'clicks' as test signals. *Int J Audiol*, 44(8):478–487, 2005. doi: 10.1080/14992020500060230.

[15] T. Fedtke and U. Richter. Reference zero for the calibration of air-conduction audiometric equipment using 'tone bursts' as test signals. *Int J Audiol*, 46(1):1–10, 2007. doi: 10.1080/14992020601050361.

[16] Torben Poulsen and Soren Vase Legarth. Reference hearing threshold levels for short duration signals. *Int J Audiol*, 47(10):665–674, 2008. doi: 10.1080/14992020802203330.

[17] Kristian Gotsche-Rasmussen, Torben Poulsen, and Claus Elberling. Reference hearing threshold levels for chirp signals delivered by an ER-3A insert earphone. *Int J Audiol*, 51(11):794–799, 2012. doi: 10.3109/14992027.2012.705901.

[18] T. Frank. High-frequency (8 to 16 kHz) reference thresholds and intrasubject threshold variability relative to ototoxicity criteria using a Sennheiser HDA 200 earphone. *Ear Hear*, 22(2):161–168, 2001. doi: 10.1097/00003446-200104000-00009.

[19] A. Rodriguez Valiente, J. R. Garcia Berrocal, A. Roldan Fidalgo, A. Trinidad, and R. Ramirez Camacho. Earphones in extended high-frequency audiometry and ISO 389-5. *Int J Audiol*, 53(9):595–603, 2014. doi: 10.3109/14992027.2014.903339.

[20] ISO 3741. Acoustics – Determination of sound power levels and sound energy levels of noise sources using sound pressure – Precision methods for reverberation test rooms. Technical report, International Organization for Standardization, Geneva, 2010.

[21] Morten Lydolf. *The threshold of hearing & contours of equal loudness – a study of measuring methods and normal hearing*. Department of Electronic Systems, Aalborg University, Aalborg Universitetsforlag, 1999.

[22] U. Richter. Determination of the equivalent threshold sound pressure levels for a circumaural earphone in the conventional frequency range. Technical report, ISO/TC 43/WG 1, Doc. Nr. 284, 1999.

[23] U. Schönfeld, W. Reuter, R. Fischer, and M. Gross. Hearing thresholds of otologically normal subjects in the extended high frequency range using the earphone hda200. Technical report, ISO/TC 43/WG 1, Doc. Nr. 263, 1993.

[24] M. Bug and T. Fedtke. Equivalent threshold sound pressure levels (ETSPLs) for RadioEar DD65v2 circumaural audiometric headphones. *Int J Audiol*, 59(8):624–630, 2020. doi: 10.1080/14992027.2020.1727034.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Data-driven Optimization of Parametric Filters for Simulating Head-Related Transfer Functions in real-time Rendering Systems

Fenja Schwark[1,*], **Marc R. Schädler**[1], **Giso Grimm**[1,2]

[1]Medizinische Physik, Universität Oldenburg, and Cluster of Excellence "Hearing4all", Oldenburg, Germany.

[2]Hörzentrum gGmbH, Oldenburg, Germany.

[*]fenja.schwark@uni-oldenburg.de

Abstract

In binaural virtual acoustics, the head-related directional properties of the receiver, i.e., the listener, are often modeled by convolution of the signal with measured head-related impulse responses (HRIRs). However, the computational cost of HRIR-convolution is rather high, especially when targeting interactive low-delay reproduction. Furthermore, depending on the spatial resolution of the HRIR data set, interpolation needs to be applied to simulate all source directions. In order to reduce the computational cost in low-delay real-time virtual acoustic rendering, this study uses a parametric digital filter model with delay lines to approximate the direction-dependent features of the head. A data-driven optimization method for the direction-dependent design function parameters is introduced that aims at matching the direction-dependent features of modeled and measured HRIRs using a spectral distance metric. Using an objective binaural speech intelligibility model, it was shown that the estimate of the speech intelligibility for the optimized model approaches the estimate for the measured HRIRs. This suggests that the parametric HRIR model may be sufficient to enable plausible spatial perception in virtual acoustic environments. With the parametric HRIR model a reduction of computational cost in the order of two magnitudes is possible for virtual acoustic environments with a small number of objects. Further work will include subjective testing of the model, compared against measured HRIRs.

Keywords: virtual acoustic environments, binaural rendering, parameterization.

1 Introduction

Virtual acoustics is used in a wide range of applications, including hearing aid research and audiology. Compared to conventional setups with only a few sound sources, the use of virtual acoustic environments (VAEs) allows more complex acoustic configurations to be tested while remaining in a predefined and reproducible environment. Thus, virtual acoustic environments form a kind of intermediate level between simple laboratory conditions and real-life environment [1].

Binaural reproduction of VAEs is most commonly realized by convolution with measured head-related impulse responses (HRIRs) [2, 3, 4, 5]. Virtual acoustic scene rendering with measured HRIRs results in an accurate reconstruction of the direction- and frequency-dependent filter effect of our head on impinging sound. At the same time, the limited number of discrete incidence directions for which measurements are available leads to spatial limitations or – when interpolation is applied – to undesired spectral artifacts. Furthermore, convolution-based systems are quite inefficient when it comes to low-delay real-time applications that require short block sizes.

Parametric models can resolve these limitations by resembling the main features of the head-related transfer functions (HRTFs) with low-order digital filters with a continuous filter design. A simple and structured modeling approach aimed at modeling physiologically motivated features such as shadow effects of head, pinna

and torso or interference effects arising from reflections within the pinna structure is structural decomposition [6, 7, 8]. A structured implementation of a small number of acoustic filters related to the geometric characteristics of the body and the physical properties of sound propagation makes such models quite understandable. In addition, model-based reproduction of VAEs is more efficient and less time-consuming than reproduction using convolution-based rendering methods for real-time applications.

An analysis of the directional magnitude characteristics revealed that such parametric models insufficiently resemble the characteristics of measured HRTFs. Moreover, the spatial fine structure of binaural intelligibility level differences cannot be reproduced correctly.

To overcome the limitations of computational performance and spatial fine structure in speech recognition threshold (SRT) predictions, we propose a framework for data-driven optimization of parametric models. The design parameters of the filters are optimized to achieve better congruence with the directional characteristics of measured HRTFs. The performance of this data-driven optimization framework is analyzed using an objective binaural speech intelligibility model [9]. Furthermore, computational performance is analyzed for a set of typical simulation parameters.

2 Methods

2.1. Parametric HRTF Models

In this study, a parametric HRTF model based on the spherical head model (SHM) introduced by Brown and Duda in 1998 [6] with additional filters modeling shadow effects and reflections [10] was used. This model consists of three stages – a *spherical head model*, a stage to reduce *front-back confusions*, and filters to reduce *up-down confusions*. The filter design functions for each stage are rotation-symmetric around a filter axis, and depend on an opening angle θ between this axis and the incident direction, see Figure 1 for an overview.

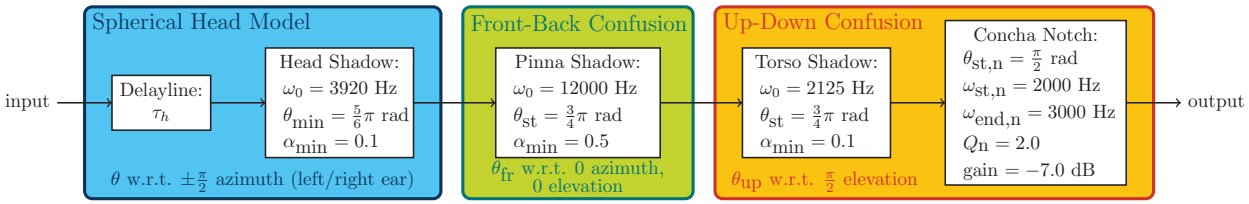


Figure 1: Structure of the spherical head model (blue box) and the extensions to reduce front-back confusions (green box) and up-down confusions (orange box), with an overview of the chosen design parameters.

Spherical Head Model

The SHM provides a simple model to describe the diffraction behavior of sound waves around the head by modeling the propagation of sound around a rigid sphere. The interaural time difference τ_h is calculated according to the frequency-independent formula of Woodworth and Schlosberg [11], which gives the time delay of a wave traveling to the observation point compared to a wave traveling to the center of the sphere in the free field:

$$\tau_h(\theta) = \begin{cases} -\frac{a}{c}\cos(\theta) & \text{if } 0 \le |\theta| < \frac{\pi}{2} \\ \frac{a}{c}(|\theta| - \frac{\pi}{2}) & \text{if } \frac{\pi}{2} \le |\theta| < \pi \end{cases} \tag{1}$$

Here θ is the angle of incidence relative to the axis of the ears at azimuth angles of $-\frac{\pi}{2}$ (left ear) or $\frac{\pi}{2}$ (right ear).

The shadow effect of the rigid sphere is approximated by the following stable single-pole, single-zero filter:

$$H_{\text{hs}}(z) = \frac{(\omega_0 + \alpha f_s) + (\omega_0 - \alpha f_s)z^{-1}}{(\omega_0 + f_s) + (\omega_0 - f_s)z^{-1}} \tag{2}$$

The frequency ω_0 is related to the head radius a and the speed of sound c by $\omega_0 = c\,a^{-1}$. The filter coefficient α is calculated as a function of the incidence angle θ according to the design function $\alpha(\theta)$:

$$\alpha(\theta) = \left(1 + \frac{\alpha_{\text{min}}}{2}\right) + \left(1 - \frac{\alpha_{\text{min}}}{2}\right) \cos\left(\frac{\theta}{\theta_{\text{min}}} \cdot \pi\right) \tag{3}$$

The original model parameters [6] are $\alpha_{\text{min}} = 0.1$ and $\theta_{\text{min}} = \frac{5}{6}\pi$, with a head radius of $a = 8.75$ cm. The directional magnitude characteristics of the SHM in the transverse and median plane is shown in Fig. 2a and 2e.

Extended Spherical Head Model

Ewert et al. [10] introduced extensions of the original SHM to account for pinna effects and torso shadow. First, a pre-warping method was added to better control the attenuation behavior for small azimuth angles (equal to equation (7) of [10] with $w_{exp} = 0.5$ and $w_{\text{a}} = 0.5$):

$$\theta_{\text{wrp}} = \frac{\theta}{3}\left(2 - \cos\left(\sqrt{\pi\theta}\right)\right) \tag{4}$$

θ was replaced by θ_{wrp} in Equation (1) and Equation (3).

To reduce *front-back confusions* a filter that attenuates high-frequency components of sources behind the head was added. This filter has the same structure as the filter for the head shadow effect, Equation (2). However, by adopting the design function, the filter operates only for a restricted range of incidence directions $\theta > \theta_{\text{st}}$:

$$\alpha(\theta) = \left(\frac{1 + \alpha_{\text{min}}}{2}\right) + \left(\frac{1 - \alpha_{\text{min}}}{2}\right) \cos\left(\frac{\theta - \theta_{\text{st}}}{\pi - \theta_{\text{st}}} \cdot \pi\right) \tag{5}$$

This is equal to equation (8) in [10] with $d = 1$. For this filter, the angle θ_{fr} is defined with respect to the frontal direction [azimuth 0, elevation 0] and the design function $\alpha(\theta_{\text{fr}})$ is evaluated. The original parameter choice was $\alpha_{\text{min}} = 0.5$, $\theta_{\text{st}} = \frac{3}{4}\pi$ and $\omega_0 = 12$ kHz.

Analogously to the shadow effect of the pinna, the torso attenuates high-frequency components from signals impinging from the lower hemisphere, resulting in less *up-down confusion*. The torso shadow filter has the same structure and design function as the pinna effect filter, with a filter axis pointing upwards, i.e., the incident angle θ_{up} is defined relative to the upward direction and the design function $\alpha(\theta_{\text{up}})$ is evaluated. The filter is applied for incidence angles $\theta_{\text{up}} > \theta_{\text{st}}$ with an original value of $\theta_{\text{st}} = \frac{3}{4}\pi$. The other initial parameters of this filter were $\omega_0 = 2125$ Hz and $\alpha_{\text{min}} = 0.1$.

Another important cues for elevation perception are created by reflections in the pinna structure. The effect of such reflections can be modeled by a parametric filter that produces a notch whose gain and center frequency change linearly as a function of elevation angle. This notch was realized by a parametric equalizer. For the original filter, the applied gain in the transverse plane ($\theta_{\text{st,n}} = \frac{\pi}{2}$) was zero and a maximum attenuation of 7 dB was applied for the sounds from above. The notch frequency changed from 2 kHz in the transversal plane to 3 kHz at $\theta_{\text{up}} = 0$. The Q-factor of the notch was set to 2.

2.2. Optimization of the Parametric Model by Measured HRTFs

For the data-driven optimization of the parametric model, the OlHeaD-HRTF Database [12] was used. From this database, the ear drum microphones of the KEMAR head-and-torso simulator were selected as a reference. For a uniform weighting of all input directions, HRIRs measured for a directionally balanced subset of 47 incidence directions were extracted from the provided measurements. The measured HRIRs were Fourier-transformed and smoothed with a function that mimics auditory filters. The magnitude of the smoothed HRTFs was extracted at center frequencies spaced by 0.5 ERB in the range between 250 Hz and 16 kHz.

The parametric filters of the extended SHM do not account for the direction-independent components of the HRTFs, e.g. the ear canal. In order to model these direction-independent filter characteristics of the ear canal,

an FIR filter for diffuse field equalization was applied. For this purpose the measured diffuse field response of the KEMAR dummy head – which is included in the OlHeaD-HRTF Database [12] – was used.

The cost function to be minimized in the optimization process was the mean square difference between the logarithmic magnitude of the sampled and simulated HRTFs, for all frequencies and directions.

For the optimization, the Nelder-Mead simplex method [13] implemented in the predefined GNU Octave [14] function 'fminsearch' was used. Simplex methods provide a numerical solution to find the minimum of multidimensional functions. While the convergence speed of these methods is rather slow, the implementation is relatively easy and the system is robust. For this study, the function 'fminsearch' was modified to only allow values in parameter-dependent predefined intervals that make sense in the context of our model, e.g., an angle should fall in the interval of 0 and π. After each operation of the Nelder-Mead Method, it was verified whether the parameters exceed the set boundaries. If any vertex element exceeded the defined boundaries, a mechanism that mirrors all values beyond the boundaries periodically back into the defined interval was executed.

The simplex was initialized with the original design parameters (see Figure 1 for an overview). As the features are generated by the entirety of all filters, it is sensible to optimize all filter parameters in parallel. However, to achieve faster convergence, the parameters of each filter were pre-optimized separately before parallel optimization of all parameters begins, taking into account the interactions of the filters. First, the parameters of the head shadow model were optimized, then the Pinna model was optimized with these parameters, and finally the notch filter was optimized, keeping all other parameters fixed.

Since the OlHeaD-Database does not provide enough measurement data in the lower hemisphere, the filter parameters of the high-shelf filter that describes the shadow effect of the torso were not included in the optimization.

2.3. Speech Intelligibility Prediction

Speech intelligibility measures are important for the description of acoustic scenes as well as signal processing methods related to acoustic communication. A parametric HRTF model which replaces measured HRTFs for rendering VAEs, should result in comparable speech intelligibility like convolution-based methods. The simulation framework for auditory discrimination experiments (FADE) [9] is a framework to perform different kinds of acoustical discrimination experiments. FADE has been designed to work as an automatic speech recognizer specialized to the simple grammar and limited speech material of the Matrix-Tests [15]. The automatic speech recognizer is trained and tested with noisy signals at different signal-to-noise ratios (SNRs) and that SNR at which a threshold of speech intelligibility, i.e. SRT, of 50% is reached is determined.

FADE was used to predict speech intelligibility of speech signals from the German Matrix Test presented in the *icra5-250* masking noise [16]. As the long term spectra of target and noise signal equal each other this noise masks the speech signal well. However, the pauses in the noise signal enable for "listening in the dips". Simulations were performed with FADE (version 2.1.1) using contralateral inhibition (FADE-kain).

FADE performs speech intelligibility predictions on the sound signals at the ear drum. In the convolution-based rendering technique, the HRIRs measured at the ear drum can be used to directly obtain the desired signals. For the parametric model, the direction-independent diffuse field response provided by the OlHeaD-HRTF Database [12] was used to obtain the simulated ear-drum audio signals.

It was possible to reproduce empirical data such as the results of a study by Beutelmann and Brand [17], in which the SRT was measured for a speech signal impinging from the front and noise presented at different azimuth angles – the so-called S_0N_φ condition – by conducting speech intelligibility prediction using FADE. This particular condition was therefore used to compare the speech intelligibility obtained with the parametric models and the measured HRTFs.

2.4. Computational performance

For a comparison of computational performance, the time required to simulate a virtual acoustic environment of 60 seconds duration was measured. The simulation was performed using the Toolbox for Acoustic Scene Creation and Rendering (TASCAR) [18, 19], version 0.223.1. This toolbox, including an implementation of the parametric HRTF model with optimized parameters, is written in C++ and optimized for real-time interactive processing with low delay.

Six different rendering methods were compared: A direction-independent method called 'omni' served as a baseline indicating the overhead of the simulation toolbox, e.g. due to the position calculation. The second method tested was the parametric HRTF simulation with the proposed optimization ('hrtf'). The non-parametric spatial rendering was implemented as 2-dimensional or 3-dimensional VBAP [20]. The 2-dimensional VBAP was rendered on eight virtual loudspeakers arranged on a circle. The 3-dimensional VBAP was rendered on 47 speakers arranged on a sphere, as used for optimization. Both methods were tested without and with convolution using the corresponding impulse responses from the HRIR database [12]. The labels for the nonparametric methods were 'vbap2d', 'vbap2dconv', 'vbap3d', and 'vbap3dconv'.

To estimate the factors influencing the required computing time, the block size P (64, 128, 256, 512, 1024) and the number of virtual sound sources N (1, 10, 100, 1000) were varied. The sampling rate was set to 48 kHz. The air absorption model as well as the time-varying delay line were disabled to isolate the effect of the rendering method as much as possible.

3 Results

Optimized Model

Starting from the original parameters of the extended SHM model as given in Table 1, the data-driven optimization was performed and the final parameters were computed.

Table 1: Original and final filter parameters of the data-driven optimization.

(a) Head Shadow

Parameter	original	final
ω_0 / Hz	3920	3100
θ_{min} / rad	$\frac{5}{6}\pi$	$\frac{8}{9}\pi$
α_{min}	0.1	0.14

(b) Pinna Shadow

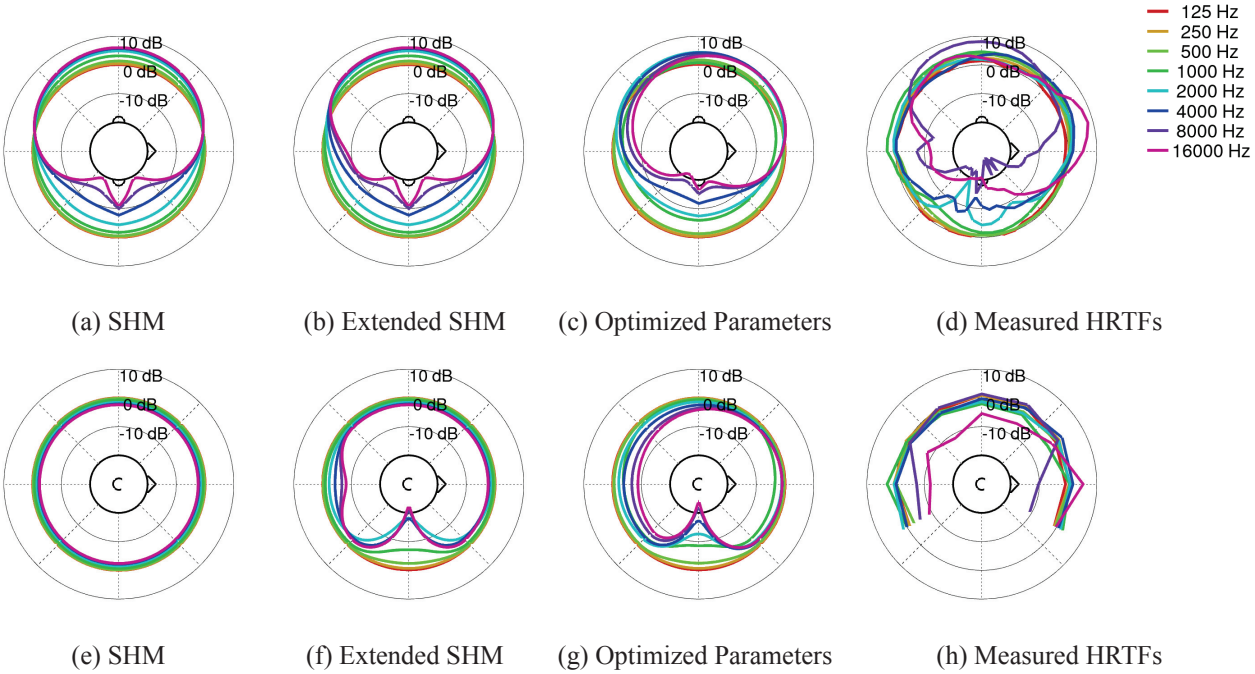
Parameter	original	final
ω_0 / kHz	12.0	11.2
θ_{st} / rad	$\frac{3}{4}\pi$	0
α_{min}	0.50	0.39

(c) Concha Notch

Parameter	original	final
$\theta_{\text{st,n}}$ / rad	$\frac{\pi}{2}$	$0.5\overline{6}\pi$
$\omega_{\text{st,n}}$ / Hz	2000	1300
$\omega_{\text{end,n}}$ / Hz	3000	650
Q_{n}	2.0	2.3
gain / dB	-7.0	-5.4

The directional magnitude characteristics of the model with the final parameters found by the data-driven optimization are shown in Fig. 2c and 2g.

The directional magnitude characteristics in the transverse plane obtained with the optimized parameters (Fig. 2c) compared to those obtained with the original parameters (Fig. 2b) show an attenuation behavior on the contralateral side that is in better agreemnet with the feature observed in the measured data (Fig. 2d). This difference is mainly related to the change in the parameters of the head shadow filter. Increasing the angle θ_{min} results in a less pronounced back lobe at high frequencies on the contralateral side. In addition, shifting the frequency ω_0 to a lower frequency leads to a shift of the transition region of the high-shelf filter to lower frequencies, i.e. attenuation of lower frequency components.



(a) SHM (b) Extended SHM (c) Optimized Parameters (d) Measured HRTFs

(e) SHM (f) Extended SHM (g) Optimized Parameters (h) Measured HRTFs

Figure 2: Directional magnitude characteristics of the left ear in the transversal plane (upper panel) and median plane (lower panel) for the SHM, the original extended SHM, the extended SHM with optimized parameters and measured HRTFs from the OlHeaD-HRTF Database [12] (diffuse field equalized measurements at the ear drum of the KEMAR head-and-torso simulator).

After optimization, the pinna shadow filter to limit *front-back confusion* operates in a completely different range than originally intended. This can be seen in the extreme change of the stating angle θ_{st}. The obtained filter leads to a general attenuation of high-frequency components in the posterior hemisphere, which is particularly visible in the characteristics in the median plane (see Fig. 2g). In addition, a general directionality of the high frequencies is observed in the transversal plane, which is slightly rotated compared to the low frequencies (Fig. 2c). These effects are in very good agreement with the characteristics of the measured HRIRs.

The features modeled by the parametric equalizer are barely visible in the directional magnitude characteristics. However, it can be seen that after the optimization, the notch is modeled in a lower frequency region than before. A detailed analysis of the transfer functions reveals that this notch obtained with the optimized parameters models a notch observed in the measured HRTFs.

An implementation of the proposed extension of the model including the optimized parameter set for the OlHeaD-Database is part of the Toolbox for Acoustic Scene Creation and Rendering (TASCAR) [18] and available as open source [19].

Speech intelligibility prediction

In addition to the directional magnitude characteristics as a physical measure, speech intelligibility is analyzed as a perceptual measure. In Fig. 3, the predicted SRT is plotted as a function of azimuth angle φ, equivalent to the direction of incidence of the noise signal. Since contralateral inhibition was used for the SRT prediction, the SRT function is (almost) axisymmetric with respect to $\varphi = 0$. It can be seen that the predicted SRTs of the measured HRTFs were better reproduced by the parametric model with the optimized parameters than by the model with the original parameters.

Differences in performance between the original and optimized parametric model are most visible in the region near full lateralization, i.e. $\pm\frac{\pi}{2}$. For these incidence directions, the SRTs predicted from the measured HRTFs

show two minima and a maximum. This behavior is not reproduced by the SRTs predicted from the model with the original parameter set, but the minimum is quite blurred. The deviation from the SRT predicted for the measured HRTFs is quite high, with up to 6 dB. In contrast, for the SRTs predicted for the model with optimized parameters, two minima and one maximum are clearly visible. The depth of the minima is less pronounced than for the SRTs obtained by the measured HRTFs. The maximum deviation from the predicted SRTs for the measured HRTFs is about 3.5 dB.

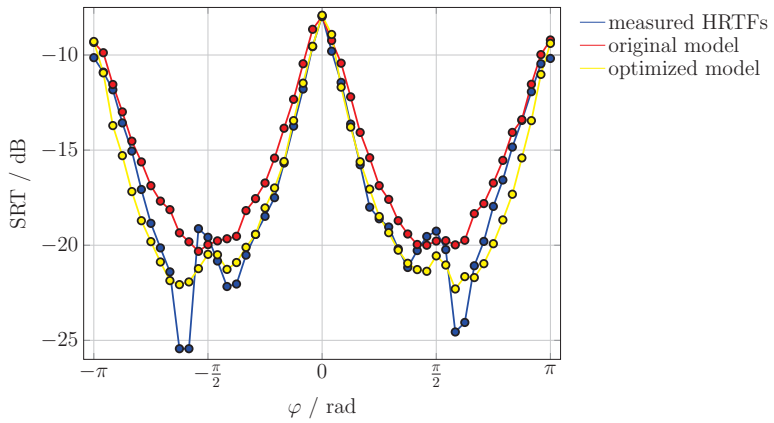
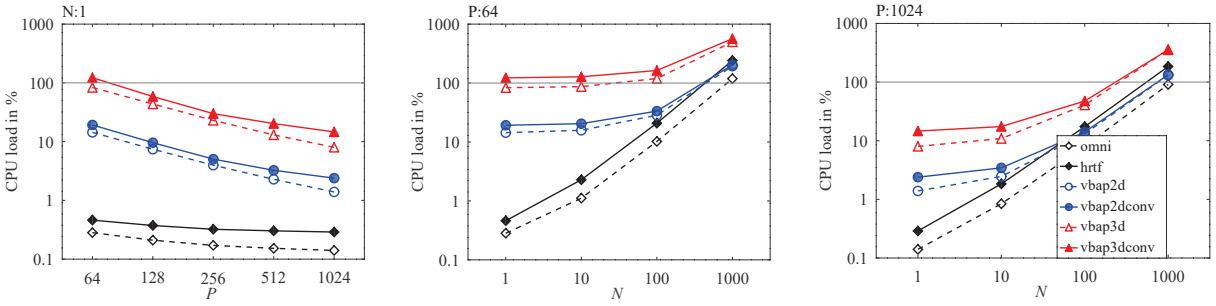


Figure 3: Predicted SRTs for the S_0N_φ condition for measured HRTFs as well as the parametric model with original and optimized parameters as a function of φ at $\frac{\pi}{24}$ resolution.

Computational Performance



(a) Influence of P (b) Influence of N, small P (c) Influence of N, large P

Figure 4: CPU load needed while rendering a virtual acoustic scene, for different rendering methods and parameters.

Figure 4a shows the CPU load, i.e., the time needed to compute an audio segment divided by the duration of that segment, as a function of block size P, for the simulation from one sound source. The CPU load of the parametric model is largely independent of the block size, while for the other rendering methods the CPU load is roughly inversely proportional to the block size. It can be seen that the parametric model has a performance advantage over 3-dimensional simulation via convolution of about a factor of 100 for small block sizes. This advantage decreases as the block size increases.

Figures 4b and 4c show the CPU load as a function of the number of sources N, each for one block size (4b for 64 samples, 4c for 1024). Here it can be seen that for the parametric model, the computational load increases linearly with the number of sources, largely independent of the block size. For the speaker-based methods, there is a lower bound independent of the number of sources.

Under all conditions tested, the parametric model was more efficient than 3D VBAP with convolution, which

corresponds to the training data set.

4 Discussion

Model Optimization

Through the proposed optimization, the directional characteristics of the parametric model approximate those of the measured HRIRs. However, the influence on the different filter stages is different: In the median plane, optimization leads to a better reproduction of the back lobe at high frequencies on the contralateral side. In addition, the directional characteristics of the high-frequency components are slightly rotated relative to those of the low-frequency components, which is consistent with the measured data. In the median plane, the general attenuation of the high-frequency components also leads to a higher agreement with the measured data.

Due to the limited number of available HRTF data points in the OlHeaD-HRTF Database [12] in the lower hemisphere, it is not possible to optimize the torso shadow modeled by the third high-shelf filter to measurement data. However, the relatively strong attenuation of high frequencies by this high-shelf filter seems appropriate because – compared to the pinna – the torso is a large object around which the sound is diffracted. This leads to an increased attenuation of high frequencies as well as to an attenuation of a wider frequency range compared to the high-shelf filter modeling the effect of the pinna. In addition, noise sources from below are rare under real conditions and therefore not very relevant.

Speech intelligibility prediction

Speech intelligibility prediction analysis shows that the proposed optimization reduces the difference between predicted SRT of the parametric model and predicted SRT with measured HRTFs. The most pronounced difference between the predicted SRTs is found for angles close to full lateralization. While the change of SRT in this region is negligible using the original filter parameters, the optimized model reproduces the two minima and the maximum in SRT as found with the measured HRTFs. However, these are more pronounced with the measured HRTFs. The prediction of SRTs shows that the agreement between parametric model and measured data is improved by the optimization, both in terms of directional characteristics and speech intelligibility.

Computational Performance

Computational performance analysis reveals that the parametric HRTF model outperforms the convolution with HRIRs for all 47 directions used in the optimization under all conditions tested. It is found that the CPU load of the parametric model is mostly independent of the block size P and increases linearly with the number of sources N. This can be explained by the implementation as a time-domain filter: the filter design function contains only few operations and is called only at the block boundaries. The computationally more intensive filter stages operate on every audio sample, thus the CPU load scales linearly with the number of audio samples, i.e., with the number of sources. In the case of the convolution, the situation is more complex: here, the audio samples are first panned according to the incidence direction. The result is a multi-channel audio, with one channel per direction. In a next step, each channel is convolved with the HRIR. If the number of sources is lower than one third of the number of channels, this implementation is less efficient than an explicit convolution for each source. However, the complexity depends on the interpolation method – here VBAP is chosen, which requires three channels for each source in the 3-dimensional case, but for Higher Order Ambisonics all channels are required even for a single source.

5 Conclusion

The physiologically motivated parametric HRTF model [10] reproduces the fundamental characteristics of binaural hearing. However, without adjustment of its parameterization, the directional characteristics of individual HRTF databases, e.g., the OlHeaD-HRTF Database [12], can be modeled only poorly. This can be demonstrated with directional SRT predictions: Prediction differences between convolution-based rendering methods

and parametric HRTF modeling indicate an effect at perceptual and even behavioral levels.

By applying the proposed data driven optimization method, the mean square deviation between the parametric HRTF model and the measured HRTF data is reduced by 1 dB from 4.1 to 3.1 dB, averaged across all directions. Likewise, the deviation of the predicted SRT based on the parametric model to the convolution-based rendering was reduced from a maximum of 6 dB to 3.5 dB without changing the structure of the model, thus maintaining the computational efficiency.

The computational performance of the parametric HRTF model is about two magnitudes higher than that of a convolution-based approach, as used for training, when the number of sources is small and the block sizes are short. As the number of sources and block size increases, this advantage decreases.

Acknowledgements

Stephan Ewert provided a Matlab implementation of an earlier version of the model described in [10]. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 352015383 – SFB 1330.

References

[1] Volker Hohmann, Richard Paluch, Melanie Krueger, Markus Meis, and Giso Grimm. The Virtual Reality Lab: Realization and Application of Virtual Sound Environments. *Ear & Hearing*, 41(Supplement 1):31S–38S, nov 2020. ISSN 0196-0202. doi: 10.1097/AUD.0000000000000945. URL https://journals.lww.com/10.1097/AUD.0000000000000945.

[2] Florian Denk, Birger Kollmeier, and Stephan D. Ewert. Removing reflections in semianechoic impulse responses by frequency-dependent truncation. *AES: Journal of the Audio Engineering Society*, 66(3): 146–153, 2018. ISSN 15494950. doi: 10.17743/jaes.2018.0002.

[3] Joachim Thiemann and Steven van de Par. A multiple model high-resolution head-related impulse response database for aided and unaided ears. *EURASIP Journal on Advances in Signal Processing*, 2019(1):9, dec 2019. ISSN 1687-6180. doi: 10.1186/s13634-019-0604-x. URL https://asp-eurasipjournals.springeropen.com/articles/10.1186/s13634-019-0604-x.

[4] Benjamin Bernschütz. A spherical far field hrir/hrtf compilation of the neumann ku 100. In *Proceedings of the 40th Italian (AIA) annual conference on acoustics and the 39th German annual conference on acoustics (DAGA) conference on acoustics*, page 29. AIA/DAGA Merano, 2013.

[5] Fabian Brinkmann, Alexander Lindau, and Stefan Weinzierl. A high resolution head-related transfer function database including different orientations of head above the torso. *Proceedings of the AIA- ...*, pages 596–599, 2013.

[6] C. Phillip Brown and Richard O. Duda. A structural model for binaural sound synthesis. *IEEE Transactions on Speech and Audio Processing*, 6(5):476–488, 1998. ISSN 10636676. doi: 10.1109/89.709673. URL http://ieeexplore.ieee.org/document/709673/.

[7] V Ralph Algazi, Richard O Duda, Reed P Morrison, and Dennis M Thompson. Structural composition and decomposition of hrtfs. In *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575)*, pages 103–106. IEEE, 2001.

[8] V Ralph Algazi, Richard O Duda, and Patrick Satarzadeh. Physical and filter pinna models based on anthropometry. In *Audio Engineering Society Convention 122*. Audio Engineering Society, 2007.

[9] Marc René Schädler, Anna Warzybok, Stephan D Ewert, and Birger Kollmeier. A simulation framework for auditory discrimination experiments: Revealing the importance of across-frequency processing in speech perception. *The journal of the acoustical society of America*, 139(5):2708–2722, 2016.

[10] Stephan D. Ewert, Oliver Buttler, and Hongmei Hu. Computationally Efficient Parametric Filter Approximations for Sound-Source Directivity and Head-Related Impulse Responses. In *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, pages 1–6. IEEE, sep 2021. ISBN 978-1-6654-0998-8. doi: 10.1109/I3DA48870.2021.9610923. URL https://ieeexplore.ieee.org/document/9610923/.

[11] Jens Blauert. *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.

[12] Florian Denk, Stephan MA Ernst, Stephan D Ewert, and Birger Kollmeier. Adapting hearing devices to the individual ear acoustics: Database and target response correction functions for various device styles. *Trends in hearing*, 22:2331216518779313, 2018.

[13] John A Nelder and Roger Mead. A simplex method for function minimization. *The computer journal*, 7 (4):308–313, 1965.

[14] John W. Eaton et al. Gnu octave. https://www.gnu.org/software/octave/, 1998-2022.

[15] K Wagener, T Brand, and B Kollmeier. Development and evaluation of a german sentence test part iii: Evaluation of the oldenburg sentence test. *Zeitschrift Fur Audiologie*, 38:86–95, 1999.

[16] Kirsten Carola Wagener, Thomas Brand, and Birger Kollmeier. The role of silent intervals for sentence intelligibility in fluctuating noise in hearing-impaired listeners. *International Journal of Audiology*, 45(1): 26–33, 2006.

[17] Rainer Beutelmann and Thomas Brand. Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 120(1):331–342, 2006.

[18] Giso Grimm, Joanna Luberadzka, and Volker Hohmann. A toolbox for rendering virtual acoustic environments in the context of audiology. *Acta acustica united with acustica*, 105(3):566–578, 2019.

[19] Giso Grimm, Tobias Herzke, Fenja Schwark, and Merle Gerken. Toolbox for Acoustic Scene Creation and Rendering (TASCAR). http://tascar.org/, https://github.com/gisogrimm/tascar, 2022.

[20] Ville Pulkki. Virtual Sound Source Positioning Using Vector Base Amplitude Panning. *J. Audio Eng. Soc*, 45(6):456–466, 1997.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Linguo-Acoustic Aspects of Speech

Marianna Hudcovičová[1], Božena Petrášová[1], Monika Rychtáriková[2]

[1]Department of British and American Studies, University of Ss. Cyril and Methodius, Trnava, Slovakia.

[2] Department of Architecture, KU Leuven, Leuven, Belgium.

marianna.hudcovicova@ucm.sk, bozena.petrasova@ucm.sk, monika.rychtarikova@kuleuven.be

Abstract
Speech is the ability making human beings unique. Linguistic diversity reflecting wide variety of speech segments does not only bring about different speech sound production mechanisms, but because of lack of appropriate acoustic research, reminds us of the need to document acoustic properties of individual sounds, as well as their possible combinations in speech in various languages. The paper presents the issues connected with the process of preparing a digital database of Slovak sentences that will be used for testing speech intelligibility in various acoustic conditions. The sentences produced as spoken utterances will be selected according to the most significant linguistic criteria to meet all the requirements for creating the relevant database. The standardised sentences will be recorded and subsequently tested in simulated laboratory conditions, while paying attention to variables in different acoustic environments. The phonetic features of the material analysed will be evaluated by acoustic analyses using the computer program PRAAT (Praat software Version 6.1.47; Boersma & Weenink) which is also used to create auditory spectrograms. The program helps to test speech intelligibility, as it measures pitch, formants, intensity and duration of sounds. These aspects are fundamental for perceiving speech sounds and decoding them properly. Based on the rules of phonotactics, combinations of vowel and consonant phonemes (in open/closed syllables, mono-/polysyllabic words) reflecting higher or restricted level of speech intelligibility in different acoustic environments will be studied. The linguistic data will be collected via the Slovak National Corpus, version prim 9.0. In order to construct the sentences from the most frequent words, the electronic statistical tools will be applied: ARF (average reduced frequency) and potential occurrence in one million words. For detecting the co-occurrence of the words, the statistical tools of the electronic corpora will be applied: MI Score and T Score.

Keywords: speech, intelligibility, digital database, acoustic conditions, linguistic criteria

1 Introduction

Communication is a fundamental and most natural way of interaction between living creatures. Not only people communicate together and exchange messages, animals or even plants can interact like this as well. The spoken form of language is dominant and prevails over the written form. Its historical, structural, functional and biological priority is inevitable and as such also proved by many scholars. The process of exchanging information could be presented by words (verbal communication) and/or by gestures or other body movements

representing non-verbal communication. Speech makes use of the smallest units of language which are manifested as concrete, audible realisation of phonemes.

2 The sound system

The sound system of languages is studied by phonetics and phonology. These linguistic disciplines analyse speech properties from various points of view. Phonetics is more practical, it deals with a concrete realisation of sounds in speech, phonology is more theoretical and studies an abstract side of a language. Three main branches of phonetics - articulatory, auditory and acoustic – study phases of communication, as it is depicted in a speech chain:

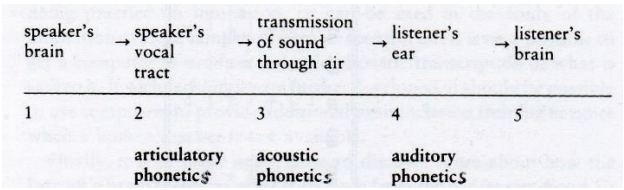


Figure 1: Speech chain [10]

In phonemic inventories of English and Slovak, vowel and consonant phonemes are manifested. The symbols used in transcription are based on the International Phonetic Alphabet (IPA).

The English language - phonemes:

English short vowels: */ʌ e ɪ ɒ ʊ ə æ/*
English long vowels: */a: i: ɔ: u: ɜ:/*
English diphthongs: */eɪ aɪ ɔɪ əʊ aʊ ɪə eə ʊə/*
English triphthongs: */eɪə aɪə ɔɪə əʊə aʊə/*
English consonants: */p t k b d g f v θ ð s z ʃ ʒ h ʧ ʤ l m n ŋ r j w/*

The Slovak phonemic inventory comprises 11 vowels (6 short and 5 long vowels), 4 diphthongs and 27 consonants. Their pronunciation is notated in slant brackets:

Slovak short vowels: phonemes */a e i o u/*, two more graphemes are used: *y* (pronounced */i/)* and *ä* (pronounced /æ/ - this phoneme is very rarely used in common speech, it is present in a higher style only)

Slovak long vowels: */a: e: i: o: u:/*, one more grapheme is used. It is written *ý* but pronounced the same way as */i:/*

Slovak diphthongs: */ĭa ĭe ĭu ŭo/*

Slovak triphthongs: none

Slovak consonants: */p t ť k b d ď g f v s z ʃ ʒ c dz ʧ ʤ h x l ľ m n ň r j l: r: /* (ť, ď, ľ, ň are palatalised t, d, l, n)

In English, each short vowel and its long counterpart differ in both quality and quantity. In Slovak, the only difference between short and long vowels lies in their quantity.

The English and the Slovak language differ from various perspectives. Regarding their origin and classification, English comes from the family of the Germanic languages and Slovak is a Slavic one. Another difference lies in their structural point of view. English is analytical, while Slovak is a synthetic language. In other words, grammar structures, owing to either morphological or syntactic issues, are quite different. Describing the sound level, English, unlike Slovak pronunciation, is based on a phonological principle between

the written and spoken form of words. It means that in Slovak, in each word a specific grapheme is always pronounced the same, i.e. represented by the same phoneme, whereas in English, no exact rules for pronunciation of single graphemes or their clusters are stated. Thus the already mentioned correspondence between the written and spoken form is a matter of convention. A typical example which can be found in many English textbooks is the pronunciation of English *oo*. Following the occurrence of this vowel cluster in English words, its different pronunciation can be really confusing for learners. Although the written form is identical, spoken equivalents of the written form vary, e.g. *blood /blʌd/, book /bʊk/, door /dɔ:/, cool /ku:l/*, etc. Despite the graphical parallelism, i.e. occurrence of the identical combinations of the graphemes, the spoken equivalents are different. In Slovak, no such discrepancies between the written and the spoken form of words exist.

2.1 Speech intelligibility

Speech intelligibility in architectural and/or civil engineering practice is generally represented by STI (speech transmission index) parameter which can be measured or simulated in different acoustic conditions [1]. The intelligibility of natural speech can be explored in various ways. For example, noise can be added to a speech signal at various pitch levels with variable bandwidth and varying intensity. Some parts of the speech signal can be filtered out, some sounds can be deleted, speech intensity can be changed. The basic results of natural speech intelligibility research prove the following facts: human natural speech is intelligible in the sound level range of about 80 dB, it is fully intelligible even at very low sound levels down to 40 dB, depending on Signal to noise ratio. According to some already done experiments, if natural speech is mixed with noise of the same intensity over the entire frequency range of hearing, noise has no effect on speech intelligibility when the ratio of speech intensity to noise is 100: 1. Speech can remain intelligible even when the noise is more intense than speech or if the speech signal comes from a different direction than the noise [2]. It is a consequence of directional hearing. Other results were obtained by speech filtering. Speech is intelligible when it passes through a low-pass filter (it means that it passes only in low frequencies), but its upper limit must be so high that the most important frequency components of speech pass through the filter. It has also been found that speech is 67% intelligible either when only low frequencies up to 1800 Hz are transmitted, or it is 67% intelligible when only frequencies above 1800 Hz are transmitted. Other experiments have shown that a 1000 Hz wide band in the region of about 1500 Hz is sufficient for 90 % speech intelligibility. The rest of low and high frequency speech sounds may be filtered out. Remarkable speech intelligibility is achieved if only the 100 - 3000 Hz band is passed. Speech remains intelligible even in adverse conditions and when an appropriate signal is not generated [3]. Recent studies have also shown the effect of sound filtering on speech intelligibility, when wearing the face masks during pandemics of COVID-19 of speech as well [4, 5].

Owing to linguistic aspects, the intelligibility of sentences is greater than the intelligibility of words, and the intelligibility of words is consequently greater than the intelligibility of syllables. The context and situation in which speech takes place is of great importance. The specific communicative situation that corresponds with the given information is also crucial, because in certain situations we expect a specific response as well as an adequate reaction. Therefore, the expected word is much easier to recognize than the unexpected word. The key point of the information is carried by the phonetic context, knowledge of grammar and vocabulary of the language, as well as possible combinations of its segments (e.g. phonotactics which means possible speech sound combinations in a certain language). Speech recognition is therefore based on a combination of acoustic, grammatical, semantic and situational information sources. Speech intelligibility is based on many sources of information, not just acoustic information itself. Everything that happens in spoken communication is derived from acoustic information, which is the basic stimulus for recognizing the speech mechanism. In natural speech, there are also information signal components that can complement and replace each other [3].

2.2 Digital database

One of the ways how to test speech intelligibility is to provide acoustic and linguistic tests focused on listening comprehension of a set of carefully constructed sentences which are recorded first in silence in laboratory

conditions and then in changed acoustic environments with increased noise components of various intensity. The selected standardised sentences form a so-called digital database which is prepared for a specific language by experts. To prepare the database as well as possible, the experts are supposed to specialise both in acoustics as well as in linguistics. The databases already exist and are used for testing speech intelligibility in many languages, e.g. the Dutch language, the Danish language, the Finnish language or the German language [6]. Our prior objective of the research is to prepare a digital database for testing speech intelligibility of such selected sentences in the Slovak language [7, 8].

3 Methodology

The research material has already been partially evaluated through description, observation, statistic methods, and synthesis. Linguistic data collection was done via the Slovak National Corpus, version prim 9.0. In order to create sentences, the electronic statistical tools were applied: ARF (average reduced frequency) and potential occurrence in 1 million words [9]. In order to detect co-occurrence of the words, the statistical tools of the electronic corpora were used: MI Score and T Score. The sentences comprised in a Slovak digital database were selected and recorded. The research sample used for this paper includes 760 Slovak sentences which differ from the structural and functional point of view. The sentences were recorded by a native Slovak speaker and then tested for their intelligibility. The whole procedure of this test is in [8]. Recordings were performed in silent environment of semi-anechoic room at the Faculty of Civil Engineering STU Bratislava, Slovakia. The intention of the performed research is to verify the STI parameter for the Slovak language using developed standardized sentences. In the following research phase, the phonetic features of analysed material will be evaluated by acoustic analyses using the computer program PRAAT (Praat software Version 6.1.47; Boersma & Weenink) which can be also used to create auditory spectrograms (or cochleagrams). Praat is a free software created by Paul Boersma and David Weenink from the University of Amsterdam. The software allows the user to do an accurate analysis of spectrograms, cochleagrams, pitch, formants, intensity, along a multitude of other functions. The software is used to recognize F1 and F2 values of vowel sounds, which furthermore allows the identification of the tongue position during the pronunciation of vowels. The program helps to test speech intelligibility by measuring pitch, formants, intensity and duration of sounds. These aspects are fundamental for perceiving speech sounds and decoding them properly.

Regarding issues of acoustic phonetics, the basic acoustic parameters of individual vowel sounds are frequency, intensity, duration and quality, i.e. wave structure determined by formants in vowels. Spectrograms allow the visualisation of three main properties of sounds: frequency, length, amplitude. The F1 (i.e. *formant 1*) and F2 values are directly connected to the height and location of vowels: for high (close) vowels, low F1 values are recognised, low (open) vowels are defined by high F1 values. On the other hand, for front vowels, it is typical to have high F2 values and back vowels have low F2 values.

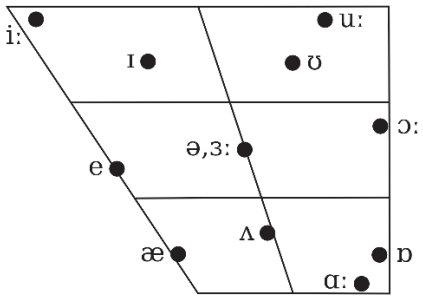


Figure 2: English distinctive vowels [10]

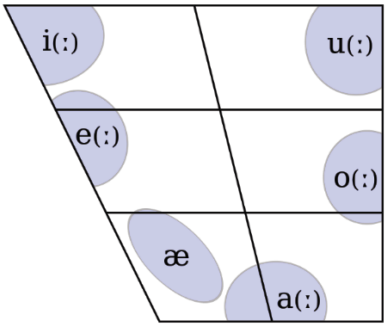


Figure 3: Slovak distinctive vowels [3]

F3 parameters reveal the level of nasality in vowel pronunciation. For consonants, the manner and place of articulation are the most relevant factors for their classification. According to the F1 value, the size of the constriction gives a clue to the manner of articulation, F2 and F3 values are related to the place of articulation. Formant structure is not recognized for voiceless consonants. Frequency corresponds to resonance in the vocal tract. Its length determines the pitch, i.e. the value of the frequency. The average formant frequencies vary when it comes to differences in a person´s gender or age. The pitch of a sound is of the highest value in children's voice (300 Hz) and lowest in adult males (110 Hz). Female voice (120-180 Hz) occurs in the middle of the frequency span. The reason is length of the vocal tract, women generally have a shorter vocal tract than men. A child's vocal tract is undoubtedly shortest causing the highest values of formant frequencies.

Table 1: Average formant frequencies of English vowels pronounced by an adult male [12]

Vowel	F1 (Hz)	F2 (Hz)	F3 (Hz)
i:	280	2620	3380
ɪ	360	2220	2960
e	600	2060	2840
æ	800	1760	2500
ʌ	760	1320	2500
ɑ:	740	1180	2640
ɒ	560	920	2560
ɔ:	480	760	2620
ʊ	380	940	2300
u:	320	920	2200
ɜ:	560	1480	2520

Table 2: Average formant frequencies of Slovak vowels [3]

Vowel	F1 (Hz)	F2 (Hz)	F3 (Hz)
i/ i:	285	1916	2656
e/ e:	452	1718	2365
a/ a:	682	1315	2293
o/ o:	481	1084	2194
u/ u:	326	967	2059
æ	700	1510	2300

Acoustic properties of speech are traditionally depicted by spectrograms which show three dimensions of the sound: time, amplitude and frequency. "A spectrogram is a graphical analysis of a sound and offers an analysis in three dimensions. The first (horizontal) dimension represents time, the second (vertical) dimension

represents frequency, and the third dimension (distinguished according to the colour intensity of the vertical lines) represents amplitude" [5]. "The intensity of the speech sounds is shown by the relative darkness of the marks. The vowels and vowel-like sounds are darkest, and the different vowel qualities can be clearly seen in the changing pattern of black bands (formants), which represent varying concentrations of acoustic energy in the vocal tract" [6]. Frequency which is measured in Hertz depicts the number of cycles per one second. According to the information mentioned by R.Pavlík, "the human ear is capable of hearing the sounds ranging from 20 to 20 000 Hz, however, most human sounds are recognizable within the frequency scope as narrow as 80 – 350 Hz" [10].

The other way how to depict speech parameters is using oscillograms. In the specialised literature, the term *waveform* is sometimes used instead. There is the horizontal axis, representing time, that directs from left to right. The curve presents the pressure increasing and decreasing in the speech signal.

Regarding linguistic criteria, acoustic parameters of speech sounds, especially vowels, were crucial for us. We stated the hypotheses:

H1: High vowels in Slovak sentences (i/i:, u/ u:) lower speech intelligibility.
H2: Polysyllabic words lower speech intelligibility of the sentences.

The other presumptions took into consideration different types of vowels/ consonants (segmental problems), tempo, rhythm, sentence stress, intonation (prosodic features), whether the words in the sentences were mono- or polysyllabic and if the words included open/ closed syllables. The other relevant criteria were their position in sentences, classification of words according to their word classes, as well as their frequency in the Slovak lexicon and sentence form from the functional and structural point of view – statements, questions, commands, exclamations, simple or multiple sentences. According to the rules of phonotactics (a branch of phonology), combinations of vowel and consonant phonemes (in open or closed syllables, mono- or polysyllabic words) reflecting higher or restricted level of speech intelligibility in different acoustic environments will be described. From the technical point of view, attention was paid on variables in different acoustic environments simulated in laboratory conditions. It was proved that selection of the sentences into the Slovak digital database required meeting both linguistic and technical criteria and parameters.

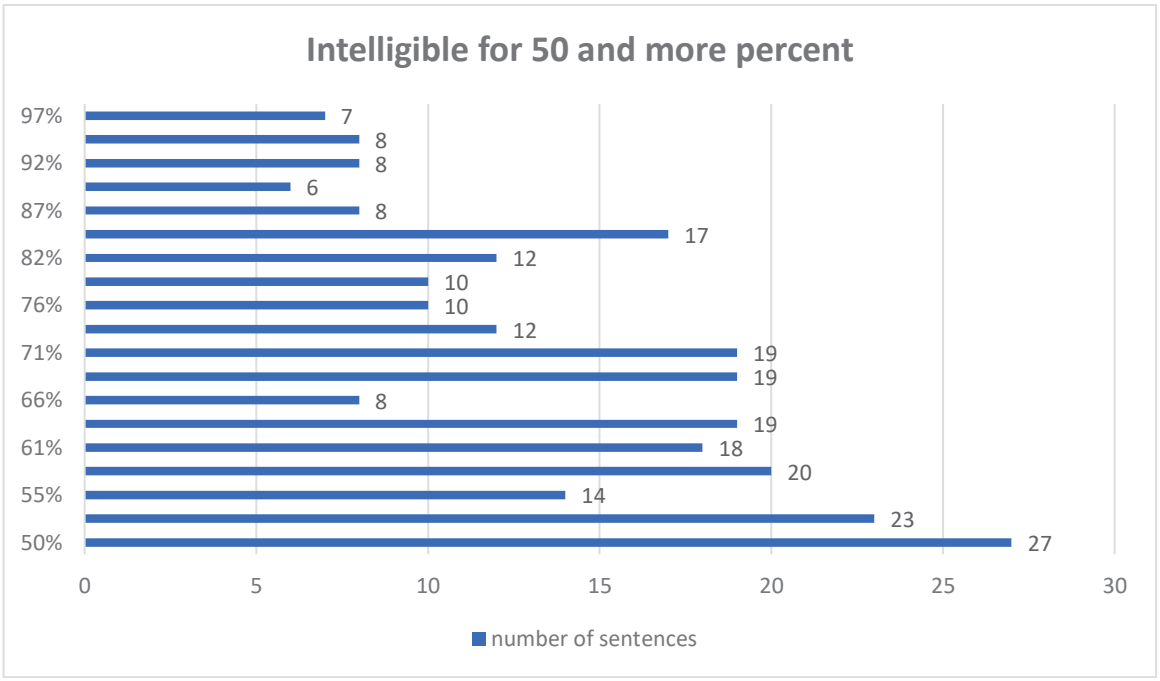
It does not matter whether the acoustic or linguistic criteria for describing the aspects of the analysed sentences in the research are used, it is crucial to use correct and adequate terminology in all disciplines of science. I.Cíbiková speaks about *terminology literacy* in this context and defines it as "the ability of a language users and specialists to use consistent terminology in functional specialised communication" [13, p.6]. According to this specialist, mastering terminology thoroughly must follow eight interrelated levels: 1) basic, 2) functional, 3) conceptual oriented, 4) defining, 5) interdisciplinary, 6) comparative, 7) documentary and 8) perspective. She emphasises the fact that scientific and terminology literacy must be consistent and reflecting a systematic documentation of human knowledge [13].

4 Research results

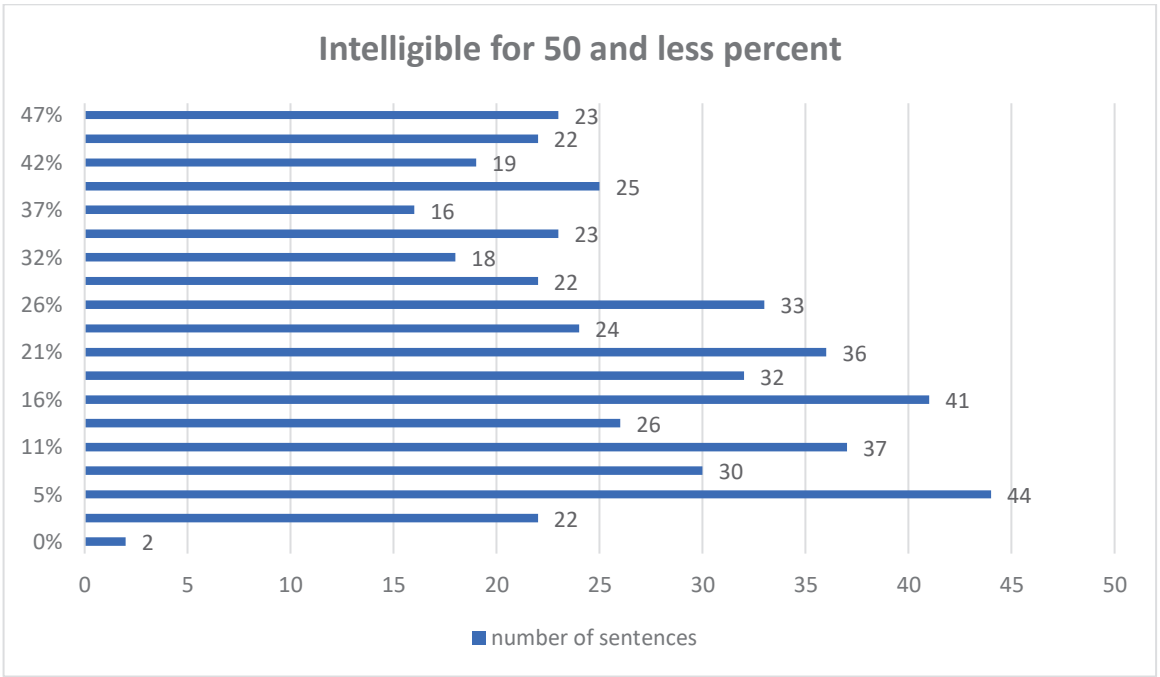
The research sample included 760 Slovak sentences which differ from the structural and functional point of view. They were recorded by a Slovak native speaker and tested in laboratory conditions with the added noise component during comprehending their intelligibility in simulated changed acoustic conditions. Their intelligibility has been evaluated in percentage:

7 sentences were intelligible for 97%, 8 sentences for 95%, 8 ones for 92%, 6 sentences for 89%, 8 for 87%, 17 for 84%, 12 for 82%, 10 for 79%, 10 for 76%, 12 for 74%, 19 for 71% as well as 19 for 68%, 8 for 66%,

19 for 63%, 18 for 61%, 20 for 58%, 14 for 55%, 23 for 53%, 27 for 50%, in TOTAL: **265 sentences** out of 760 were **intelligible for 50 and more percent**:



The rest (495 sentences in TOTAL) were intelligible for less than 50%: 23 sentences for 47%, 22 for 45%, 19 for 42%, 25 for 39%, 16 for 37%, 23 for 34%, 18 for 32%, 22 for 29%, 33 for 26%, 24 for 24%, 36 for 21%, 32 for 18%, 41 for 16%, 26 for 13%, 37 for 11%, 30 for 8%, 44 for only 5%, 22 for 3% and – what was surprising - 2 sentences were intelligible for 0%. It means that **495 sentences** were **intelligible less than for 50%:**



5 Conclusions

Partial analysis has already been done focusing on intelligibility of the Slovak sentences in a digital database. The total number of 760 sentences were evaluated according to their acoustic and linguistic properties, so the research results can be considered relevant owing to a high amount of the sentences analysed. All of the sentences were tested for intelligibility in changed, simulated acoustic laboratory conditions with the added background noise component, so the conditions were not ideal for comprehending every word comprised. The two hypotheses were stated, both were proved. Slovak high vowels (i/i:, u/ u:) were mostly present in sentences with the level of intelligibility less than 50%, so they really lower speech intelligibility. The sentences that were better understood and the meaning was clearly deciphered, included vowels a/ a:, o/ o: mainly. These vowels are considered low in the Slovak language.

Regarding Hypothesis 2: Polysyllabic words lower speech intelligibility of the sentences was proved as well. The sentences with the level of intelligibility less than 50% included mainly polysyllabic words. The sentences with a higher degree of comprehension comprised predominantly shorter words than the sentences with a lower degree of intelligibility. It is a great challenge to explore also the other aspects supporting or breaking intelligibility of Slovak sentences, specifically the ones involved in the already prepared database. The research is still ongoing.

References

[1] HOUTGAST, Tammo – STEENEKEN, Herman. Past, present and future of the Speech Transmission Index. Publ. TNO Human Factors, Soestenberg, 2002. ISBN 90-76702-02-0.

[2] RYCHTÁRIKOVÁ, Monika, et al. (2011): *Perceptual Validation of Virtual Room Acoustics: Localization and Speech Understanding.* Applied Acoustics 72 (4), p. 196-204, ISSN 0003-682X.

[3] KRÁĽ, Ábel – SABOL, Ján. Fonetika a fonológia. Bratislava: SPN, 1989. ISBN 80-08-00036-8.

[4] BOTTALICO, Pasquale, et al. Effect of masks on speech intelligibility in auralized class-rooms. J. Acoust. Soc. Am. 2020, 148, 2878–2884.

[5] CHMELÍK, Vojtech, et al. Effect of Mouth Mask and Face Shield on Speech Spectrum in Slovak Language. *Appl. Sci.* 2021, *11*, 4829. https://doi.org/10.3390/app11114829

[6] VAN WIERINGEN, Astrid – WOUTERS, Jan. Sentences and numbers for quantifying speech Understanding in severely impaired listeners for Flanders and the Netherlands. Publ. International journal of Audiology. London, 2008.

[7] PANOCOVÁ, Renáta – GREGOVÁ, Renáta. Designing the Slovak matrix sentence test, International journal of applied languages and cultural studie. Volume 2, Issue 2, 2019.

[8] HÚDOKOVÁ, Dominika et al. Conception of the listening test procedure for quantifying speech intelligibility in Slovak language – a preliminary study, *EUROREGIO/BNAM2022.* Aalborg, Denmark, 2022.

[9] Slovak National Corpus, version prim 9.0

[10] PAVLÍK, Radoslav. Phonetics and Phonology of English. Bratislava: PdF UK, 2000. ISBN80-88868-67-X.

[11] CRYSTAL, David. The Cambridge Encyclopedia of the English Language. New York: Cambridge University Press, 2011. ISBN 978-0-521-53033-0.

[12] WELLS, John Christopher. Sounds Interesting. CUP. 2014. ISBN 978-1107074705.

[13] CÍBIKOVÁ, Ingrid – SIANTOVÁ, Gabriela – MITAĽOVÁ, Katarína. Specialised communication. Scientific and terminology literacy. In: International Masaryk Conference for Ph.D. Students and Young Researchers. Volume XII. Hradec Králové: Magnanimitas akademické sdružení. 2021. ISBN 978-80-87952-35-1.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Transformation of office space to laboratory listening room

Lukáš Zelem[1*], Vojtech Chmelík[1], Daniel Urbán[1] and Monika Rychtáriková[2,1]

[1] STU Bratislava, Faculty of Civil Engineering, Radlinského 11, 810 05 Bratislava, Slovakia

[2] KU Leuven, Department of Architecture - Campus Brussel and Gent, Hoogstraat 51, 9000 Gent/Paleizenstraat 65, B1030 Brussel, Belgium

*lukas.zelem@stuba.sk

Abstract
The issue of sustainability in society leads to efforts associated with changing the purpose of conventional spaces while maintaining the essence of the original building. Therefore, the change of use of the spaces is a common process during the buildings' life cycles. However, if the purpose of the room with lower requirements is transformed to a room with higher requirements, in terms of building physics, it is necessary to look at the given issue from multiple points of view. One of the basic points of view is to critically evaluate, whether there is a space in the building that has the potential to fulfil the requirements for a new purpose. This article is aimed at the transformation of the so-called office spaces to an acoustic laboratory designed for subjective laboratory listening tests. Within a given transformation, it was necessary to select a room, which eliminated possible sources of interfering sounds by its location. Subsequently, construction adjustments were necessary. These treatments were focused on the improvement of the building and room acoustics. The aim was to create a room with as low a background noise level as possible and high sound absorption at the same time. During the conversion process, we had to face several issues resulting from the original design of the space. Airborne sound insulation is improved using a gypsum board lining system. The sound attenuation of the newly designed listening room is ensured by materials with high sound absorption. Thanks to these adjustments, it was possible to significantly increase airborne sound insulation and reduce the average reverberation time. The laboratory is recently used for research as well as the teaching process.

Keywords: listening room, office, reverberation time, equivalent sound pressure level, sound insulation

1 Introduction

Nowadays trend of using unconventional spaces for a conventional purpose (or vice versa) has been enhanced by the pandemic of COVID-19 since the beginning of 2020 [1]. There has been a need for adaptation of rooms with large volumes to fulfil the requirements of local hygiene measures and sometimes even transformation of large atriums of schools into lecture halls or gymnasia into classrooms, etc [2]. The other consequence of pandemic can be seen in many administrative buildings that remained empty for a couple of months, after companies chosen to go for home office. Recently, retrofitting administrative buildings into dwellings became a very discussed topic [3-5]. Different (higher) requirements need to be fulfilled in dwellings in comparison with offices. Not all changes in space function are caused by a new release of hygienic requirements due to the COVID-19 pandemic. The need for room purpose changing could be also due to a need to create specialized space which is easily reachable from the office. In the following paper, the change of two offices into an acoustic laboratory dedicated for listening tests was done. During the process of conversion, we had to face several challenges, since the requirements on sound insulation as well as background noise levels for listening tests rooms are stricter in comparison with common office [6-8]. Standard for designing of listenig room for laboratory listening tests with loudspeakers exists [9].

The purpose of this case study was to create a laboratory space, well insulated from surrounding rooms, in order to obtain low background noise levels, and to treat the interior surfaces with sound absorbing materials, both necessary for the performance of high-quality laboratory listening tests. The laboratory consists of two parts, the mentioned listening room and adjacent control room where the noisy technical equipment, as well as operator of the listening test, are typically situated.

2 Analysis of potential sources of noise

The building-up of listening room was planned in the building of the Faculty of Civil Engineering at the Slovak University of Technology in Bratislava (STU SvF). The idea behind was to build a well-insulated acoustic laboratory which would be at the same time easily accessible for students of civil engineering and architecture (e.g. for presentation and demonstration purposes). The transformation of the original office space to a new purpose, required extended analysis of the potential indoor and outdoor noise sources (e.g. traffic, position of the elevator, activities performed inside the building, etc.), and detailed information about the building (e.g. type of structure, construction style and inspection of surrounding walls, floors and ceilings, etc.). Based on the mentioned information, the most suitable place in this multi-storey building was chosen. A scheme of the building and its orientation towards surrounding environment is shown in the Figure 1.

2.1 Exterior sound sources

The building is situated in the city center of Bratislava and thus close to a busy street (Radlinského) with a two-way tram line. The office space, chosen to be converted to the laboratory space, has been chosen at the 22nd floor (almost the highest floor in this building) in approximately 73 m above ground to reduce the levels of direct sound coming from the traffic. On the 23rd floor there are only storage places and silent infrastructure, which act as an extra buffer space between the roof and ceiling of the lab. On one hand, the traffic noise in front of the building façade is reduced by altitude but on the other hand the noise caused by unexpected wind impact on the building facades sometimes occurs.

2.2 Interior sound sources

The Figure 1 (in a green rectangle) shows a particular typological scheme of the 22nd floor of the Faculty of Civil Engineering. The black colour represents the positions of elevators, turquoise colour are classrooms and rooms with 3D printers. Red colour shows the position of designed listening room of the acoustic laboratory and blue colour is the control room of the acoustic laboratory. Grey colour shows corridors and public spaces. It is clear, that the highest acoustic load from interior sound sources will be caused by the movement of elevators, communication of students and moving of people down the hallway adjacent to the laboratory.

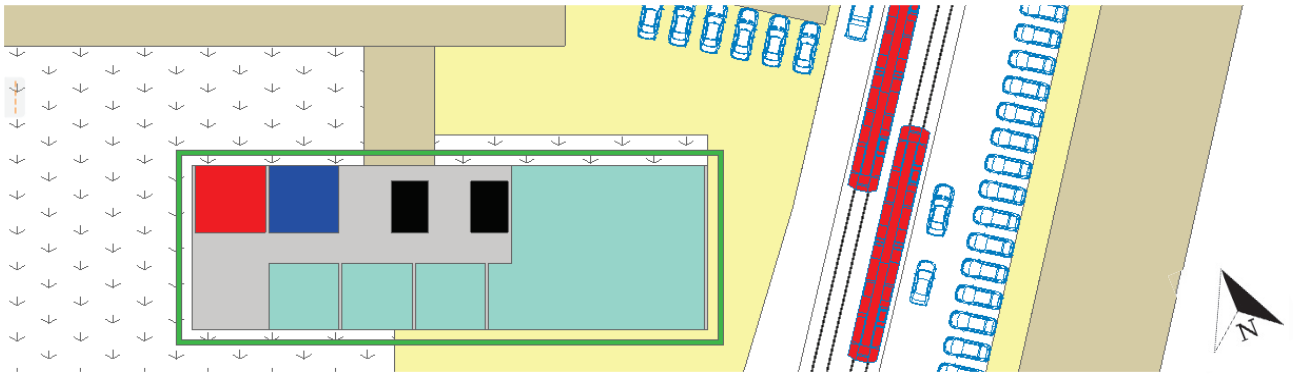


Figure 1 Schematic illustration of the STU SvF complex in Bratislava and the typology of 22nd floor in the high-rise building (the green rectangle) and its location related to the traffic.

3 Adaptation of the building constructions around the listening room

The basic structure of the building is based on reinforced concrete. The exterior walls are made out of autoclave concrete with thickness of approximately 300 mm. These walls ensure a high level of airborne insulation at low frequencies of the noise generated outdoors. Problem is a structure of a lightweight transparent façade from the north side of the building, which has a low mass and includes acoustic bridges. In addition, it is a subtle structure that is sensitive to wind impacts.

For this reason, an additional glazed wall (Glass Solutions), with increased acoustic properties, was built to improve the sound insulation of façade and to guarantee a day light in the listening room. The extra glazed wall creates a corridor, which thanks to its width increases airborne sound insulation at low frequencies too. The floor plan of the acoustic laboratory and its control room is shown in the Figure 2 left.

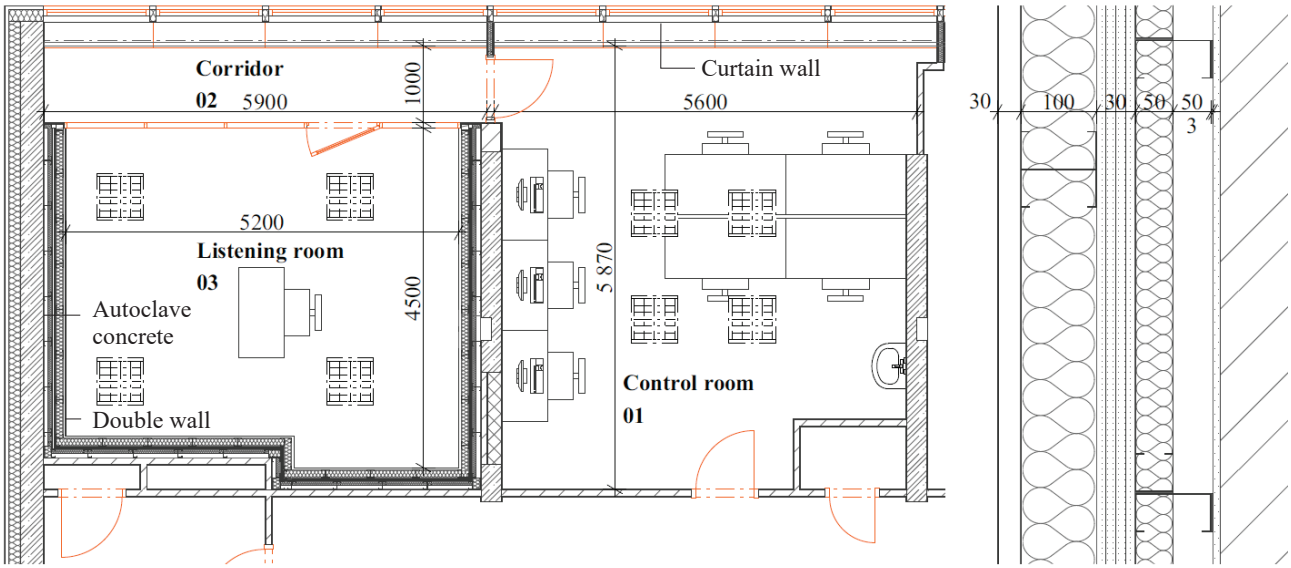


Figure 2 (Left) The floor plan of the acoustic laboratory - control room (01), corridor (02) and listening room (03). (Right) The composition of the additional lining of the wall in the listening room.

To increase the airborne sound insulation against interior sound sources, the double wall was designed. It consists of 4 layers of acoustic gypsum boards (Rigips ACTIV'AIR® MA AA - thickness of one board is 12.5 mm) built with distance of 100 mm from the original wall. The resulting air gap is filled with mineral wool (with thickness 50 mm). The frame for double wall is created from acoustic CW and UW steel profiles. Detailed composition of the listening room wall is shown in the Figure 2 (right). The whole frame is flexibly connected to the original structure to prevent transmission of vibration to plasterboards. Subsequently, the material with high sound absorption coefficient (mineral wool with thickness of 100 mm - Tab. 1) is placed on the whole wall surfaces, from interior side (except the glazed wall) using additional frame from steel profiles. Finishing of walls, i.e. covering of mineral wool is done by textile with high airflow coefficient to ensure the high absorption of the system. The ceiling is covered by acoustic Ecophon panels (Tab. 1) and floor is covered by carpet with thickness of 5 mm. It helps achieving the optimal room acoustic parameters in the listening room.

The above-described construction modifications, aimed at increasing of the airborne sound insulation in the considered listening room, don't solve problems with impact noise. Given relatively low height of ceiling, the impact noise insulation was solved especially typologically. Laboratory is located on the highest floor. Therefore, only the technical floor with minimum occurrence of persons is located above the listening room. Thanks to this, the impact noise from the floor above our protected room is not an issue. The second problem is the movement of the elevator in the elevator shaft and noises from the engine room of the elevator.

This noise source has been partially reduced by the positioning of the listening room at further distance from elevator. The same approach was applied for solving of the impact noise within one floor. Teaching rooms are in relation to the laboratory's positions as far as possible. If necessary, the classroom can be equipped with carpet as well as in the laboratory or control room.

Table 1 Sound absorption of materials used inside the listening room

Material	Product	Thickness (mm)	Frequency (Hz)						α_w (-)	Sound absorption class
			125	250	500	1 k	2 k	4 k		
Mineral wool panels	Isover Akuplat	100	0.55	1.00	1.00	1.00	1.00	1.00	1.00	A
Acoustic panels	Ecophon Master SQ	40	0.25	0.80	0.95	0.95	1.00	1.00	1.00	A

4 Objective assessment of acoustic adjustment

To determine the impact of acoustic treatments, the acoustically treated listening room was compared with the untreated control room. Both spaces were analysed in terms of sound insulation of walls and reverberation time. The background noise determined from equivalent sound pressure level measurements in the listening room compared with the control room too.

4.1 Room acoustic parameters

Room acoustics parameters were derived from impulse response measured by software Matlab - using ITA toolbox. The exponential sweep signal with a length of 5.46 s (5 repetitions) was used as excitation signal. Measurements were performed by means of omnidirectional loudspeaker and microphones Behringer ECM8000 with a flat frequency response from 20 Hz to 20 kHz. The impulse response was measured at 3 positions of sound sources and 6 positions of microphones in two heights - 1.2 and 1.8 above the floor and sound sources was placed at the height of 1.5 m above the floor (Fig. 3) (altogether 36 microphone positions). The measurements were performed according to the standard EN ISO 3382 - 2 [10] as "precision measurement" category (considering the number of measured sound sources and microphones positions).

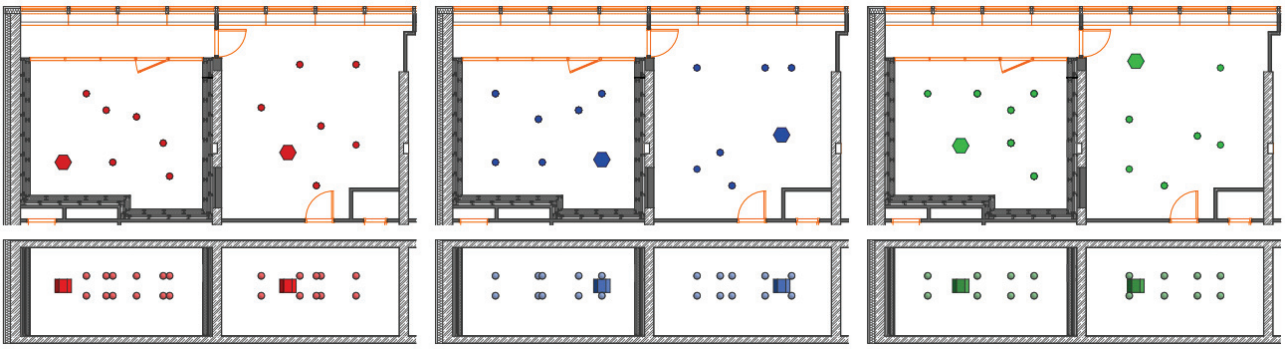


Figure 3 Floor plans (pictures above) and sections (pictures below) of sound sources positions (hexagons) and associated microphones (circles).

The values of early decay time *EDT* and reverberation time T_{10}, T_{20} and T_{30} for both rooms (listening room and control room) were derived from the impulse responses (Fig. 6). In the listening room, the course of the reverberation times and the *EDT* is almost identical (Fig. 4 left), in the control room the course of the *EDT* is different in the frequency spectrum from 1.6 kHz to 10 kHz (Fig. 4 centre). For comparison of the reverberation

time of the two rooms, the T_{20} was chosen (Figure 4 - right). The reverberation time with indication of standard deviations in the control room is indicated by black triangles and the reverberation time in the listening room is shown in red circles. The effect of sound absorbing material in the listening room is very clear.

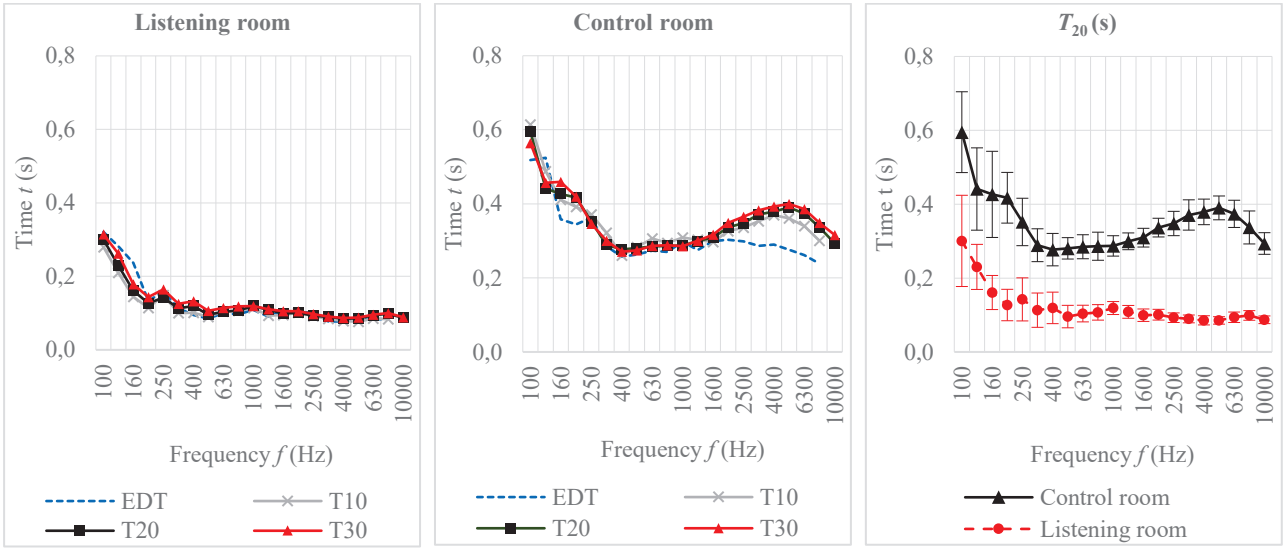


Figure 4 Comparison of *EDT* and reverberation time T_{10}, T_{20}, T_{30} in the listening room (left) and in the control room (middle) and the course of the reverberation time T_{20} with the standard deviation in the listening and control room (right).

4.2 Building acoustic parameters

Measurement was performed according to standard ISO 16283 - 1 [11]. The positions of microphones and sound sources are shown in Figure 7. The airborne sound insulation D_{nT} (standardized level difference) and D (level difference) values were derived from the measurements according to the procedure in the standard ISO 717 - 1 [12]. The pink noise signal in the frequency range from 20 to 20 kHz was generated by the omnidirectional sound source (Fig. 5 - right).

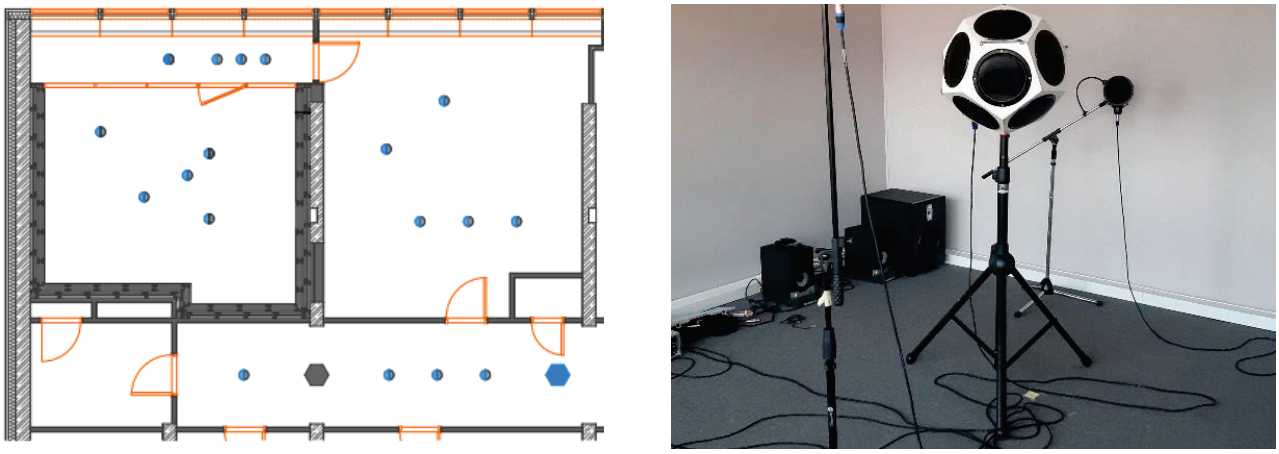


Figure 5 Positions of sound sources (hexagons) and associated microphones (circles) when measuring airborne sound insulation (left) and omnidirectional sound source (right).

Monitoring of sound pressure level was done by the Norsonic Nor140. The measurement of airborne sound insulation was made in the completely closed listening room. Further the door in the glazed wall was opened to investigate the influence of the glazed wall. The airborne sound insulation measurement between the control

room and the corridor was realized only with completely closed door (Fig. 6). In Figure 6 on the left, we can see values of D_{nT}, both in the listening room and in the control room. The total $D_{nT,w}$ (weighted standardized level difference) value of the partition wall between corridor and listening room reaches 67 dB compared to the wall between corridor and control room, which reaches 33 dB. Thanks to the use of building modifications, an overall increase in airborne sound insulation $\Delta D_{nT,w}$ of 34 dB was achieved in case of listening room compared to the control room where original partition wall is built. In Figure 6 on the right, we can see the sound insulation between the corridor and the two rooms (listening room and control room) expressed in D values.

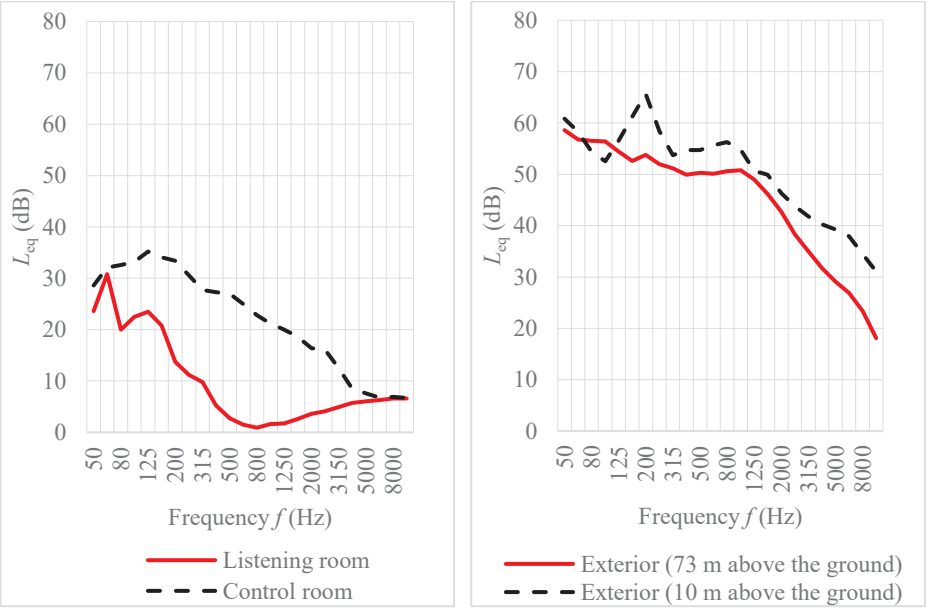


Figure 6 Equivalent sound pressure level L_{eq} in the listening room and control room in frequency domain. (left). Equivalent sound pressure level L_{eq} in the building extarior (2 m in front of fasade) at 2 different heights (right).

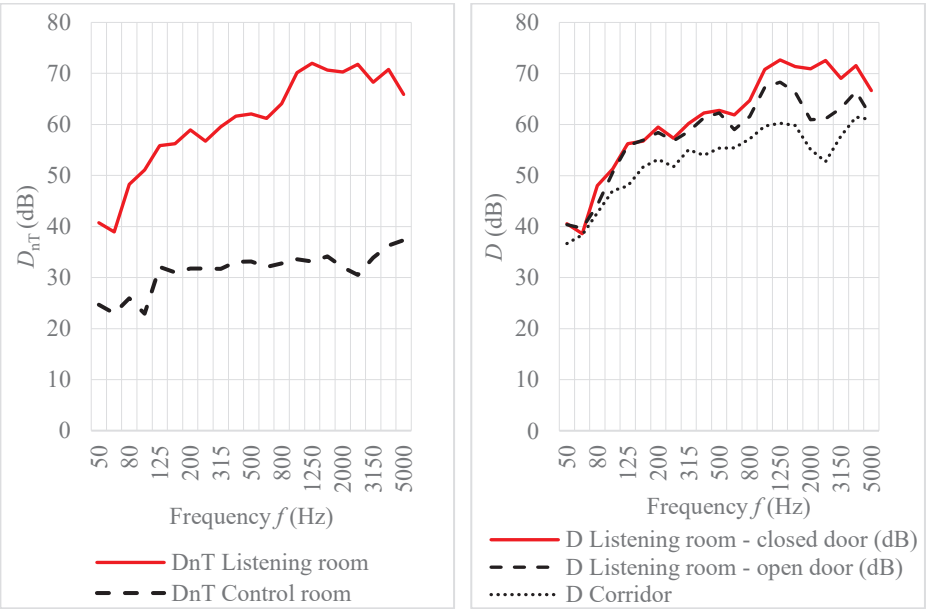


Figure 7 Frequency-dependent values of airborne sound insulation D_{nT} of the walls between corridor and the listening room and control room respectively (left) and the influence of the glass wall to reduce the noise level D (right).

4.3 Equivalent sound pressure level

The background noise was measured during a most busy hours, i.e. from 6:00 to 11:00. The measurement was performed using a Norsonic Nor140 analyser. The values of the equivalent sound pressure level $L_{A,eq}$ and percentile values of noise levels $L_{A,95}$ and $L_{A,10}$ were calculated.

In Figure 7-left we can see the equivalent sound pressure level measured in the listening room and control room. The effect of the increase in airborne sound insulation of the listening room compared to the control room is reflected in all third octave bands from 50 Hz to 20 kHz. However, the highest influence is from the frequency band 50 Hz to 5000 Hz. This interval corresponds to the maximum spectral width in the evaluation of structures in terms of building acoustics [12]. The total A-weighted equivalent sound pressure level reaches 17 dB in the listening room and 31 dB in the control room.

The level of almost constant noise $L_{A,95}$ is 16 dB in the listening room and 31 dB in the control room. The noise levels $L_{A,10}$ was measured as high as 20 dB in the listening room and 35 dB in the control room. We can see that the noise dispersion in the listening and control room is the same, i.e. 4 dB. However, thanks to adaptation of surrounding constructions, the noise level in the listening room was reduced by 14 dB in the case of $L_{A,eq}$ and by 15 dB in the case of the levels $L_{A,95}$ and $L_{A,10}$.

The equivalent sound pressure level in the building exterior was also measured at a distance of 2 m in front of the façade at a height of 10 m above the ground (3rd floor) and at a height of 73 m above the ground (22nd floor). Figure 7-right shows the frequency-dependent course of equivalent sound pressure level. The A-weighted equivalent sound pressure level in height of 10 m was in average 63 dB; and it reached 58 dB in hight of 73 m. Thanks to the maximum possible increase in the distance of the listening room from traffic noise, the equivalent sound pressure level was reduced by 5 dB.

5 Conclusion

This article focuses on a conversion of an office space to an acoustic laboratory dedicated to performance of listening tests with minimal costs. Challenges on adaptation of a space with mild acoustic requirements to a space with increased acoustic requirements are shown.
First, the analysis of potential noise sources has been performed and most convenient position of the listening room in the building has been chosen. The measured difference in equivalent sound pressure level at a distance of 2 m in front of the facade was at the 22nd floor (where the lab is finally situated) 5 dB lower in comparison with pedestrian ground floor level.
Later, several steps in terms of improvement of airborne sound insulation of partition walls and building façade were made. The sound insulation after the improvements have reached a D_{nT} value of 67 dB, while the D_{nT} of the original partition wall was only 33 dB.
The background noise expressed in equivalent sound pressure levels $L_{A,eq}$ and statistical values of $L_{A,95}$ and $L_{A,10}$ were also significantly reduced. The $L_{A,eq}$ in the listening room does not exceed 17 dB, while the background noise in the control room (representing the situation without any interventions) is around 31 dB. The $L_{A,95}$ in the listening room is around 16 dB, while in the control room it reaches 31 dB. The $L_{A,10}$ = 20 dB in the listening room and 35 dB in the control room.
Finally, adjustments were made also in terms of room acoustics, by placing the highly sound absorbing material on wall and ceiling surfaces. Floor is covered by carpet.
It can be concluded that the transformed room is now quiet enough and suitable for the research where laboratory listening test are used.

6 References

[1] "Director-General's Opening Remarks at the Media Briefing on COVID-19-11 March 2020," WHO, [Online]. Available: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-oncovid-.

[2] Y. Sluyts, L. TingChun, D. Urbán and M. Rychtarikova, "The effect of mask wearing on speech intelligibility in various architectural environments in schools," in *In proceedings of Euronoise*, Madeira, 2021.

[3] M. Buschka, J. Bischof, C. Meier-Dotzler and W. Lang, "Developing non-residential building stock archetypes for LCI - a German case study of office and administration buildings," *The International Journal of Life Cycle Assessment,* vol. 26, no. 9, pp. 1735-1752, 2021.

[4] H. Ogawa, K. Kobayashi, N. Sunaga, T. Mitamura, A. Kinoshita, S. Sawada and S. Matsumoto, "A study on the architectural conversion from office to residential facilities - through three case studies in Tokyo," pp. 171-178.

[5] B. Clifford, J. Ferm, N. Livingstone and P. Canelas, "Overview of Office-to-Residential Conversion in England and Our Case Studies," *Understanding the Impacts of Deregulation in Planning,* pp. 47-60, 2019.

[6] C. Monteiro, M. Machimbarrena, D. de la Prida and M. Rychtarikova, "Subjective and objective acoustic performance ranking of heavy and light weight walls," *Applied Acoustics,* vol. 110, pp. 268-279, 2016.

[7] P. Virjonen, V. Hongisto and J. Radun, "Annoyance penalty of periodically amplitude-modulated wide-band sound," *The Journal of the Acoustical Society of America,* vol. 146, no. 6, pp. 4159-4170, 2019.

[8] C. Calleri, L. Shtrepi, A. Armando and A. Astolfi, "Evaluation of the influence of building façade design on the acoustic characteristics and auditory perception of urban spaces," *Building Acoustics,* vol. 1, no. 25, pp. 77-95, 2018.

[9] IEC TR 60268-13 Sound system equipment - Part 13: Listening tests on loudspeakers, 1998.

[10] ISO 3382-2 Acoustics - Measurement of room acoustic parameters - Part 2: Reverberation time in ordinary rooms, 2008.

[11] ISO 16283-1 Acoustics - Field measurement of sound insulation in buildings and of building elements - Part 1: Airborne sound insulation, 2014.

[12] ISO 717-1 Acoustics — Rating of sound insulation in buildings and of building elements - Part 1: Airborne sound insulation, 2013.

[13] V. Chmelík, M. Rychtáriková, H. Müllner, K. Jambrošić, L. Zelem, J. Benklewski and C. Glorieux, "Methodology for development of airborne sound insulation descriptor valid for light-weight and masonry walls," *Applied Acoustics,* vol. 160, 2020.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Sound producing mechanism in temperature inversion layer and its sensitivity to geomagnetic activity

Unto K. Laine[1,*]

[1]Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland.

* unto.k.laine@aalto.fi

Abstract
Sounds associated with aurora borealis, so called Auroral Sounds (AS), are explained by electrical discharges occurring in Temperature Inversion Layer (TIL) approximately 75 meters above the ground [1]. It is assumed that under favorable weather conditions the number of AS events depends on the activity of the geomagnetic (GM) storm. The historical testimonies of AS indicate that the GM storm should be strong enough to rise the bright and lively auroras to the zenith before these sounds can be observed. The goal of this study is to test the relevancy of this claim by studying the sensitivity of the AS producing mechanism under a moderate GM activity. A four-hour long period of the sound measurements done around the local magnetic midnight Jan 25–26, 2022 at Fiskars village, Finland, is analyzed. Sixty AS event candidates were manually selected and their temporal distribution and period histogram constructed for statistical comparisons with the GM data measured simultaneously by the Finnish Meteorological Institute (FMI) at Nurmijärvi Geophysical Observatory (NGO). The data was also used for regression models to study possible causal relationship between the GM activity and the AS events. A strong causality between the GM activity and the AS events was found. The sounds were predicted by the GM activity with 90% accuracy after a delay of 21 min. The results show that the AS events are much more common than previously thought.
Keywords: aurora borealis, solar wind, magnetosphere, auroral sounds, geomagnetism, temperature inversion layer, regression models, period histogram.

1 Introduction

Based on the present knowledge, Auroral Sounds (AS) have physically existed around auroral zones as long as the Earth has had an atmosphere and magnetosphere similar to the present ones. Thus, the problem of AS is not just *a millennial problem*, as often referred to, but has probably surprised many species before the humans. The question of the first descriptions of aurora borealis and also of the sounds associated, have been discussed a long time. Based on earlier studies [2,3], it is probable that one of the first documented observations was made by prophet Ezekiel around 593 BCE. For some reason the earlier studies of this event are limited entirely to the *visual* part of Ezekiel's observations falling silent in front of his *sound* perceptions. Ezekiel not only described the visual art of auroras, which he called angels, but continued with the sounds. He tells, that he heard "*the roar of rushing waters*".

When the Auroral Acoustics project [1] started and the author made one of his first AS observations, a TV-reporter asked for a description of the sounds. The answer was: "*they sounded like Niagara Falls from a distance of one kilometre*". Afterwards, when the author became familiar with the text of Ezekiel, the similarity of the descriptions surprised. The recent AS observations include descriptions, where a strange sound from the sky woke up the attention and after looking to the direction of the sound source an aurora was observed. The sound focused the mind to search its reason from the direction it was perceived, on the open sky. In Same culture we find opinions that an aurora without a sound is not a real aurora. A 'real

aurora' should be bright, colourful, lively with changing structures. Exactly these features are often mentioned as conditions for auroral audibility.

Until around 1930's some authors still believed that under special conditions auroras may come down below the altitude of the clouds and therefore they could be audible [4]. Even after the hundreds of altitude measurements made by Norwegian *Carl Størmer* [5], who first time showed that only in extremely rare cases the lowest edge of the auroral curtain may be a little below 80 kilometres, some authors kept asking, could they still sometimes come much closer to the ground and be therefore audible. Because the *acoustics* is not in the focus of geophysics, the *location of the sound source* was never studied or revealed.

At the end of the 18th century, the question of whether electricity could explain the auroras garnered increasing attention. The earliest texts proposing that the AS events are caused by electric discharges are probably from the years 1831 and 1834 [6,7]. Sometimes AS were even used as a proof for the electric nature of the aurora. In the 1920's *C. A. Chant*, editor of *The Journal* (JRASC) interested about this problem and asked the readers to send their sound observations to the editorial [8,9]. Professor *William C. Baker*, from Queen's University, Kingston, sent a letter to Chant where he described an AS event which occurred in mid-winter in 1884 or 1885. Baker's story was exact and rich in ideas. He described the sounds as *"the crumpling of stiff paper"*. He assumed that the sounds are produced by *the same electro-magnetic* (EM) source that also controls the visible aurora and describes the sound; "*It reminds of the noise heard when discharging the glass jar*", referring to a capacitor known as *Layden jar*. Baker summarizes that because the delay between the auroral movements and the sounds was small (1–2 s), the sound source must be closer to the observer than the light source. It is obvious that Chant learned much of these accurate observations and embraced Baker's brilliant ideas. However, the problem of *where such capacitor is and how it is charged*, remained unresolved almost one hundred years during which, with a few exceptions [10], more speculation was produced than genuine research to solve this problem.

However, the electric fields caused by the GM activity are never large enough to produce *directly* discharges and sounds in the lower atmosphere. Tree tops are quiet. The present Inversion Layer Hypothesis (ILH) provides a physical explanation for the sound production [1]. It assumes that after a calm and sunny day the warm air at the ground rises carrying negative ions up forming a Temperature Inversion Layer (TIL) about 70-100 metres from the ground. Above this layer positive ions are 'raining' down from the upper atmosphere accumulating in the layer above the warm air. Because vertical movement of the air is minimal at the inversion, these two layers continue to accumulate charges forming finally a structure like a large plate capacitor. The activation of a GM storm generates the needed additional potential between the layers and triggers the discharges. It should be noted that *the energy for the sounds* is loaded to the capacitor during the evening and night, meanwhile the GM activity works just as an *activator* for the discharges.

The main goal of this paper is to study the sensitivity of the sound production mechanism of the inversion layer under *moderate* GM activity. The statistical connections between the magnetic field measurements made by the FMI and the recorded AS events are studied. Four different methods to reveal these connections are introduced. They all provide similar results showing *a causal relationship* between the magnetic field fluctuations and the sounds. Surprisingly many periodic events known in the field of geomagnetism have left their fingerprints to the sound sequence. The new findings support the presented ILH.

Next chapter starts with a description of the measuring system, the weather, and the GM conditions during the night of measurements. Then the selection of sixty AS events is described with examples of their temporal and spectral properties. Chapter 3 describes the four approaches to show the causal and statistical connections between the magnetic field measurements made by FMI and the recorded AS events. The paper ends with a conclusion that summarises not only the new results, but the whole Auroral Acoustics project.

2 Sounds and magnetic field measurements Jan 25–26, 2022

The same measuring system as described in [1] was used in this recording, too. Its 'soul' is the high–quality measuring microphone of Brüel & Kjær (4179 with preamp 2660) that was mounted at the focus of a parabolic dish antenna directed vertically to the sky. The internal noise of this microphone is around -3 dBA and it is carefully shielded against all EM-interferences. Additionally, two other microphones and a VLF loop antenna were connected to a four–channel digital recorder thus allowing the sound source localization.

2.1 The weather and geomagnetic conditions

A moderate GM activation occurred on Jan 25–26, 2022. The Kyoto Dst index [11], shown in Figure 1, stayed at -35 during the hours 19–20 LT. The intensity of the GM activation degreased during the period of data collection. The local magnetic field activity, measured at NGO (60° 30.5' N, 24° 39.3' E), was moderate, too. The FMI has estimated that the GM activity meter should rise over 0.3 nT/s before the probability of auroras in the southern Finland is high. The highest measured activity was only 0.14 (see Figure 5 chart in blue) indicating that during the night in question aurora was probably not visible in the southmost Finland.

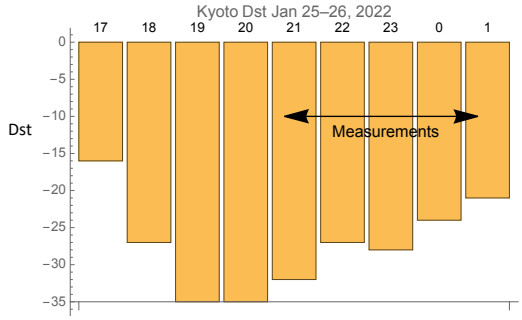


Figure 1. Kyoto Dst in the night Jan 25–26 2022. The arrow shows the time window of The data collection for this work (21:20–01:20 LT).

The quasi-K-index calculated for the three last hours before the local midnight gives K=4 and for the next three hours after the midnight K=3 (in decimals: 4.6 and 3.7). The probabilities of these occurrences at Nurmijärvi are correspondingly: 8.1% and 19.3% [12]. However, still the developed automatic analysis of the recorded audio material was able to find around two hundred AS candidates and a careful manual selection (carried out before the automatic analysis) sixty cases in the 4 h time window. This means that *the sound production in the TIL is much more common than ever imagined*. Under favourable weather conditions, TIL is sensitive to even relatively small geomagnetic disturbances as will be seen. The probability for similar geomagnetic activities is over 25% in south Finland, however, the build–up of a strong TIL is not equally common and both of them are needed for the sound production.

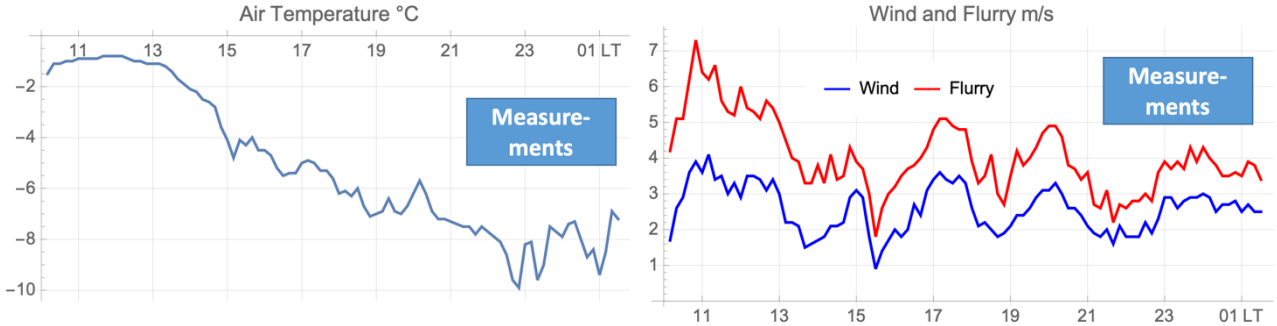


Figure 2. Wind and temperature conditions at Jokioinen Jan 25–26, 2022. Blue boxes indicate the time window of the measurements used in this study. Data from the FMI.

The closest fully armed FMI weather station to the Fiskars village locates at Jokioinen, about 90 km to the North. According to the data of that station, the air pressure degreased from 1020 to 1017.6 hPa during the measurements, and the sky was clear. The clear night produced about 9 °C temperature drop at the ground from the highest midday reedings indicating that an inversion layer was formed. The wind speed was around 2 m/s at the beginning of the data collection and increased later close to 3 m/s (see Figure 2).

2.2 Sixty AS event

Sixty AS event were manually collected from the 4 h long recording of the B&K microphone signal based on their auditive, temporal and spectral properties. Twenty examples of them are shown in the Figure 3. The AS events have large temporal variability. Some of them are like impulses, very compact in time, meanwhile some others are like clap or pop sounds with noisy echoes. In the left frame a sequency of cracks is shown (the fourth signal from the top). Due to the amplitude (peak) normalization, it looks like the background noise is varying. In reality it varied quite little. The night was calm and quiet.

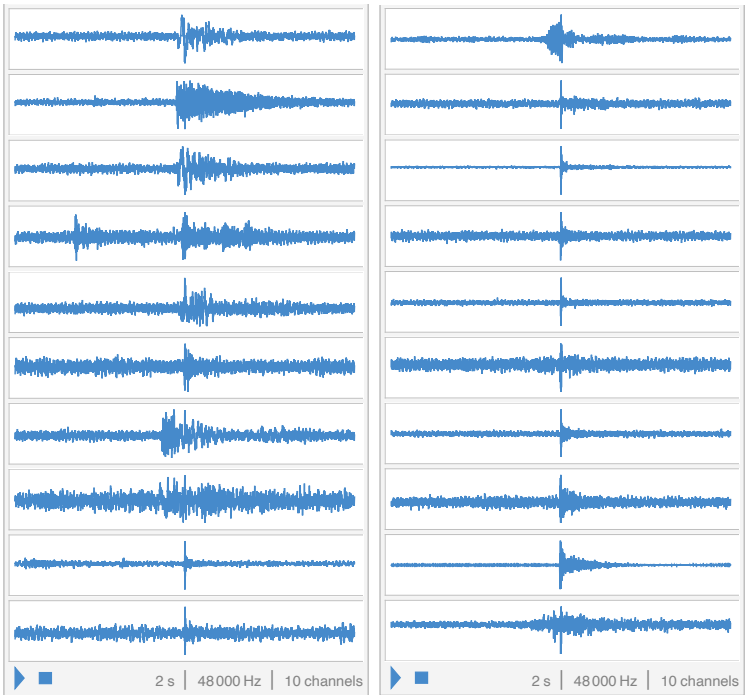


Figure 3. Twenty randomly selected AS events from the set of sixty cases.

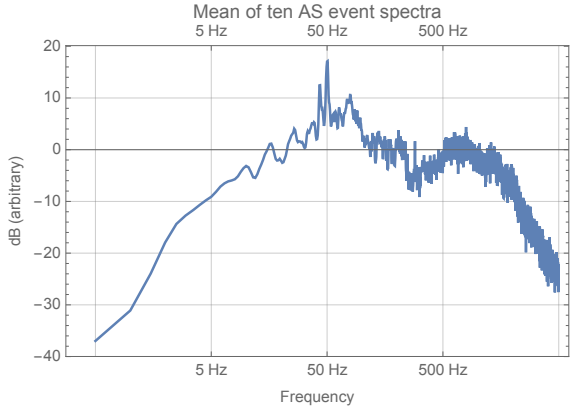


Figure 4. Mean of ten randomly selected AS event spectrum.

Figure 4 shows the mean of spectra of ten AS events that occurred in a row. The signal of the B&K measuring microphone is high-pass filtered before the recording. The filter with -6 dB/oct roll–off and 500 Hz cut–off is used to hinder strong infra sounds to clip the signal by driving it out of its dynamic range.

The spectral peaks below the 300 Hz occur often at the Schumann resonances. They are sometimes clearly intensified around the AS events. These components from a constant background hum on which the AS events are superimposed. A strong 50 Hz component is seen. It cannot be caused by any direct electric or magnetic field leakage from the mains to the microphone which is carefully shielded and has a low output impedance. The present supposition is that also this acoustic hum is created at the inversion layer. The observation that this peak often varies between 47–52 Hz supports this idea. The electric network in Finland is based on 50 Hz AC and this frequency can't vary so much out of its standard. However, these details in the low-frequency range need still more work. The main spectral peak of the AS events is located around 700–2500 Hz. Sometimes spectral peaks may appear even around 7–9 kHz.

In the following the time instances of the AS events are studied in comparison to the GM activity and in more detail to the magnetic field measurements made by FMI at NGO.

3 Causality between GM activity and AS events

During the first phase of the Auroral Acoustics project one attempt was made to build a bridge between AS events and geomagnetic measurements made by the FMI. This was published in the Master's Thesis of *Janne Hautsalo* [13]. Without extracting the individual AS events from the audio signals the whole material was analysed by a 1/3–octave filter bank and the RMS values of the outputs correlated with those of the GM field at different delays. Statistically significant correlations were found at many frequency bands and at many delays. Since that analysis it was clear that the variations in the geomagnetic fields are the cause for the sounds, even though that time the localizations of the sound sources were not solved, neither the physical mechanism of the sound production. The experience of the past years provides methods to extract the AS events from the noisy background both manually and algorithmically. An automatic method is under development, but it is not yet working sufficiently well to serve the research like this.

After the extraction of the individual AS events their *temporal histogram* in 10 min windows is constructed (the right histogram in red in Figure 5). The data starts at 21:20 and continues to 01:10 local time. The data for the histogram in blue is taken from the web–page of the FMI [14]. Both histograms have a clear peak of maximal activity. The GM activity seems to precede and predict the AS events.

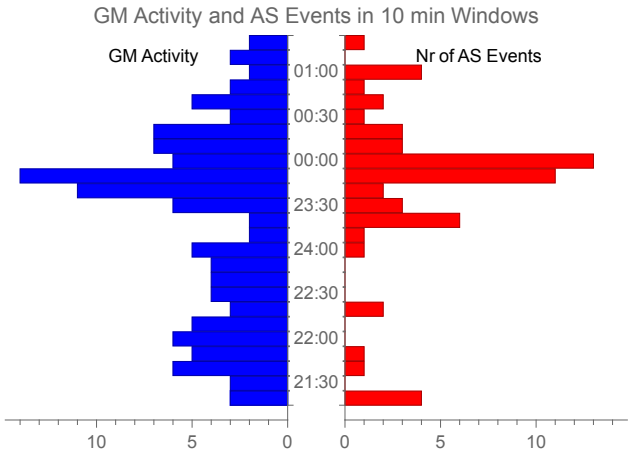


Figure 5. Comparison of the number of AS events and the GM activity at NGO in 10 min windows. The unit of the GM data is 10 pT/s. The time scale in the middle is the local time.

The direct correlation between the GM activity and the number of AS events gives $r = 0.44$ with a large $p = 0.38$. When the GM activity is delayed by 10 and 20 minutes, these numbers are: $r = 0.69$, $p = 0.4$ and $r = 0.47$, $p = 0.09$. This indicates that a statistically significant fit should occur at a delay around 20 minutes. A closer picture of the connections between the magnetic field fluctuations and the AS events can be created by applying the regression analysis.

3.1 Regression Model

A standard method to reveal possible *causality* between a set of independent variables and a dependent variable, or a response of a system to variations in the independent variables, is the *regression analysis*. Now, the GM data provided by the FMI containing three magnetic field components (*B*x, *B*y, *B*z) and sampled with ten second periods is used for a regression model to predict the number of the AS events counted in overlapping 12–34 minute windows moved in one–minute steps. The mean of the magnetic field components are removed and the sequences partitioned to ten–minute segments consisting of 60 samples each. The RMS values of each segment of each component is computed. The obtained RMS sequences form the independent variables for the regression model.

First, eleven time-sifted variants of the AS event sequence are produced moving the window's centre starting from 1 to 12 min after the first AS event in steps of one minute. Additionally, there was a constant delay of 10 min between the magnetic field data and the first AS event. Now, each time–sifted variant represents the AS event sequence with a delay varying from 11 to 22 minutes providing the solving of the optimal lag for the regression model. The optimal regression model was found with a lag of 21 minutes. It predicts the AS events with 90% accuracy (R2 = 0.895). The correlation between the best regression model and the optimal AS event shift–variant is $r = 0.95$ with $p < 0.0001$ when the 32 min window size is used.

The regression coefficients for the {*B*x, *B*y, *B*z} sequences are {0.91, 0.22, -0.12}. When the mean RMS values of these sequences are multiplied with their coefficients, we get: {6.58, 3.15, -2.49}. This indicates that the *B*x has more active role in the sound production than the *B*y and the effect of *B*z is negative. This leads to a conclusion that *the variation in the horizontal magnetic field component H={Bx, By} has an important role in the sound generation in the temperature inversion layer*. This outcome provides a deeper insight to the Inversion Layer Hypothesis: Variations in the *H*–component induces vertical potentials (electromotive forges) which may increase the ionization in and between the charged layers thus increasing the probability of a discharge between the layers in the form of an avalanche of electrons [15].

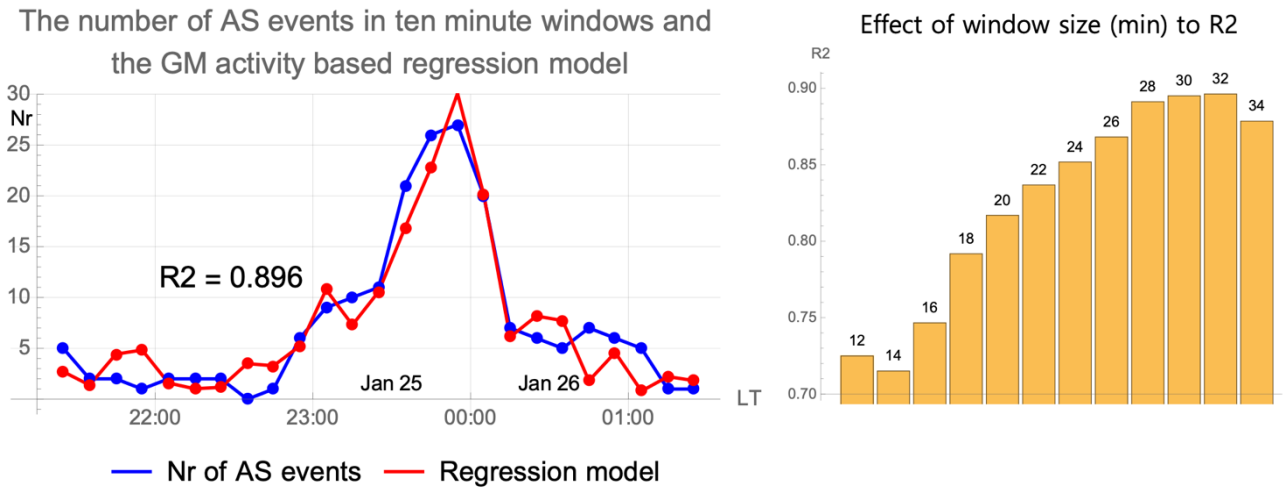


Figure 7. Smoothed temporal distribution of the sixty AS events (in blue, delayed 21 min) predicted by a magnetic–field–based regression model (in red). R2 values as a function of the window size (on the right).

3.2 Period histogram of the AS event sequence

A novel, not earlier discussed aspect related to AS is their *period histogram*. Knowing the time instances of the sixty AS events, a histogram of all time periods between the events is easily constructed by subtracting every time instant from all the later ones. The number of periods in a gliding four–minute window is calculated at one–minute intervals to construct the histogram shown in Figure 8 in blue. Also, period histograms of one hundred randomized cases with the same number of events were constructed. Their mean and deviations from the mean by one standard deviation are correspondingly illustrated in gray, red, and green. The AS events are clearly not located randomly in the four–hour window, but certain periods are occurring more often than the others. Especially periods of 10, 17, 26, 37, and 53 minutes arise strongly above the random statistics. These peaks deviate from the random mean 15.2, 9.4, 10.8, and 12.0 times the random STD. Their largest p–value is smaller than 0.007 indicating the statistical significancy of these peaks. The histogram shows also how the number of the shortest periods is high. These cases include clap–pairs, known as "klip–klap"–events. These cases together with the crackling sounds have been analysed earlier and their statistical connections to Schumann resonances shown [16].

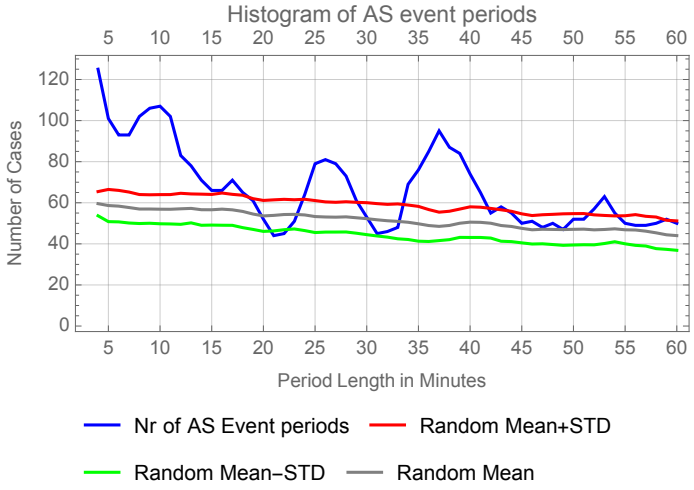


Figure 8. Period histogram of the AS events shows peaks at durations of 10, 17, 26, 37, and 53 minutes. Statistics of one hundred random cases are presented as a reference (see text).

The found peaks at durations of 10, 17, 26, 37, and 53 min happen to be almost identical to those found in a study of the magnetosphere and the solar wind by Huang &al. [17]. They summarize: 'It is known that the solar wind has significant effect on the magnetospheric convection and that magnetospheric electric fields can penetrate into the midlatitude and low-latitude ionosphere. We suggest that the midlatitude ionospheric electron density perturbations were caused by the penetration of magnetospheric electric fields which were controlled or modulated by the oscillations in the IMF (Interplanetary Magnetic Field) or solar wind pressure.' In that study the loss time of the oxygen ions due to recombination at the *F* region altitudes were estimated to be 12, 18, 26, 37, and 52 min at the altitudes of 220, 230, 240, 250, and 260 km. Now a question arises as to whether these processes, occurring at 220–260 km above the ground, could cause the same pattern in geomagnetic field on the ground and further in AS event period histogram or, is this just a coincidence? This cannot be answered in this short study. However, next the period histograms of the magnetic field measurements at NGO are analyzed and compared to the AS event period histogram.

3.3 Period histograms of the magnetic field sequences *B*x, *B*y, and *B*z

Period histograms were constructed for the same magnetic field data that was used for the regression model in section 3.1 by using practically the same algorithm as in the construction of the period histogram for the AS events. In this case the amplitude peaks of the magnetic field components must first be solved. The cross-correlation between the obtained period histograms of the components (*B*x, *B*y, *B*z) and the period histogram

of the AS events are: r={0.58, 0.60, 0.58} with the largest p< 0.0001. The period histogram of By and that of AS events are shown in Figure 9. The period histogram of the By component has the three first peaks approximately at the same period length as the AS event histogram, whereas the fourth peak of the AS event histogram doesn't have any clear counterpart in the By histogram. The same is true with the Bx and Bz period histograms.

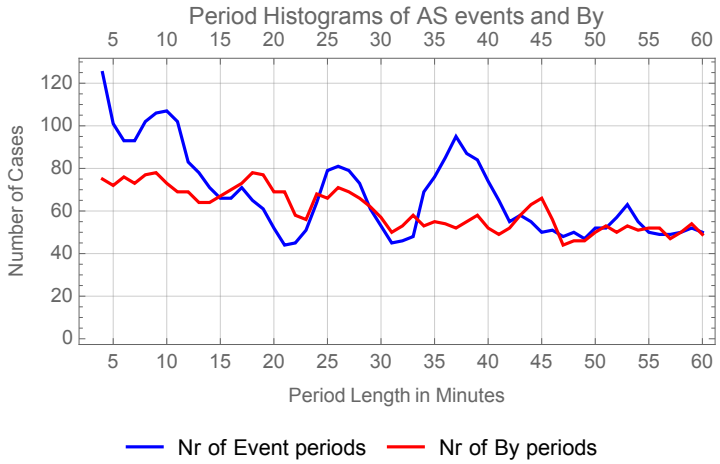


Figure 9. Period histograms of AS events and simultaneous magnetic field By-component at NGO.

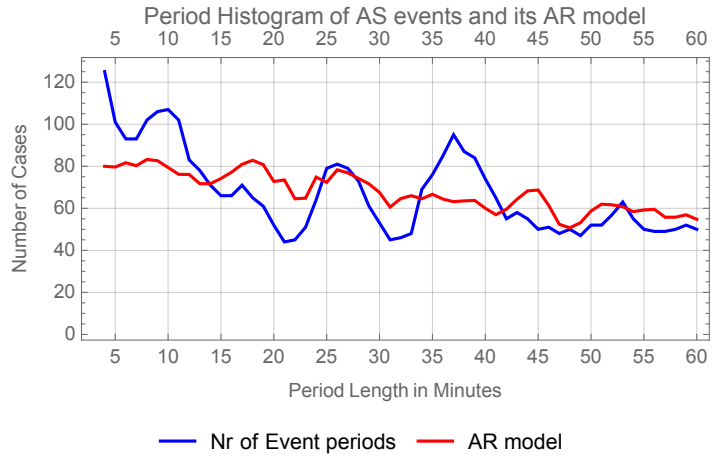


Figure 10. Period histogram of AS events and AR model (see text).

The statistical connections between the period histograms can be illuminated further by constructing a regression model from the magnetic field component histograms to model the AS event histogram. The obtained regression model is depicted in red in Figure 10.

The cross-correlation between the AR model and the period histogram of the AS events is r=0.64 with p<<0.0001. Thus, the magnetic field components explain 41% of the AS events period histogram. However, when the cross-correlation is calculated in a smaller window having period lengths between 5–30 min, the correlation improves to r=0.72 (p<<0.0001) and R2=0.52. The three first peaks of the AS events period histogram are explained by the regression model that is based on the period histograms of the magnetic field components by 52%.

3.4 *H*-component of the magnetic field

The role of the horizontal (*H*) component activity in the AS event production was discussed above. The magnetic field components Bx and By were interpolated for solving the ΔH vectors at 1300–50 seconds

before the sound events in steps of 50 s. The obtained results are clustered to five classes by using Gaussian mixture method (see Figure 11). The ΔH values of the first and largest cluster with 22 cases are amplified by ten for a better visualization. This cluster has a sharp variation in the Bx-component before the sound events. The analysis shows that about 2/3 of the cases (#48) have a slow and large variation of 20–40 nT and 1/3 (cluster 1) a smaller but fast variation before the AS event.

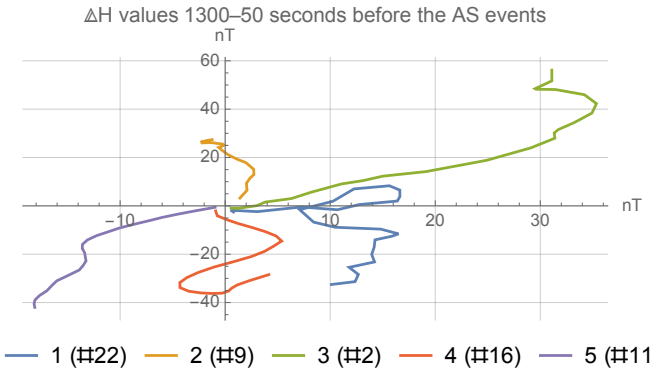


Figure 11. Five cluster means of the ΔH = {ΔBx, ΔBy} before the AS events (at origin).

4 Conclusions

The Auroral Acoustics project started over twenty years ago [18]. That time there was not much hope that this 'millennial problem' could ever be solved. Many geophysicists around the globe had fixed their opinions: "Auroral sounds are illusions" and "They are physically impossible". However, a new, previously unknown sound producing mechanism in the temperature inversion layer was discovered [1]. The present study connects the sound production to the independent magnetic field measurements. The results show that the sound producing mechanism is far more sensitive to the magnetic field variations and the auroral sounds are much more common than has ever been understood. The key issue is the formation of a strong temperature inversion layer. This occurs only during an excellent "*fox weather*" conditions meaning; calm, serene sky, temperature drop around 8°C, and at least a moderate geomagnetic activity.

AS events have a complex relationship to the different types of geomagnetic activities. It was discovered earlier that the patterns of the crackling sounds are connected to the Schumann resonances [16]. The present work shows new connections through period histograms between the magnetism and the sounds. The discovered important periods are on the band 10–40 min belonging to the Ps 6 type magnetic variations that are especially connected to *B*y–component during the recovery phase of magnetic substorms [19]. The discovered 21 min delay between the magnetic fluctuations and the sounds has not yet explanation.

Auroral sounds are produced with a separate process from the auroral light production. However, both have the same geomagnetic background causing synchronism between these phenomena. The creation of auroral sounds may occur and be audible even without any visual aurora, which is one of the revolutionising conclusions of this long project. Still, due to the historical and perceptual (cognitive) reasons, we may still use the term *auroral sounds*, even though the sounds are not 'auroral', but *electric discharge sounds* created at 70–100 m above the ground.

Acknowledgements

The author is grateful to the FMI for providing the geomagnetic data of NGO and *Okko Räsänen* as well as an anonymous reviewer for the constructive comments on the manuscript.

References

[1] Unto K. Laine, Auroral Acoustics project – a progress report with a new hypothesis, Baltic-Nordic Acoustic Meeting, BNAM 2016, Stockholm, Sweden, June 2016.

[2] George L. Siscoe, S. M. Silverman, and K. D. Siebert, Ezekiel and the Northern Lights: Biblical Aurora Seems Plausible, Eos, Transaction, American Geophysical Union, Vol. 83, No. 16, 16 April 2002, pp 173–179.

[3] Oleg M. Raspopov, Ezekiel's Vision: Visual Evidence of Sterno–Etrussia Geomagnetic Excursion? Eos, Transaction, American Geophysical Union, Vol. 84, No. 9, 4 March 2003, pp. 77–88.

[4] Sydney Chapman, The Audibility and Lowermost Altitude of the Aurora Borealis, *Nature* Vol. 127 No. 3201, March 7 1931, pp. 341-342.

[5] Carl Størmer, Altitudes of Aurorae, *Nature*, Vol. 97, No. 2418, March 2 1916, p. 5.

[6] W. Dunbar, Notice of the Aurora Borealis of Last Winter. Edinburgh Journal of Natural and Geographical Sciences (New Series), Vol III, ART. V., April 1831, pp. 225-226.

[7] Anonymous, Aurora Borealis, The Knickerbocker 1834 – New-York Monthly Magazine Vol. IV. August, 1834, pp. 98-107.

[8] William C. Baker, a letter to the editor of Journal of the Royal Astronomical Society of Canada, JRASC, Vol XVII, No. 7, September 1923, p. 279.

[9] Fiona Amery, The disputed sound of the aurora borealis: Sensing liminal noise during the first and second international polar years, 1882–3 and 1932–3. https://royalsocietypublishing.org/doi/10.1098/rsnr.2021.0031

[10] S. M. Silverman, T. F. Tuan, Auroral audibility, Adv. Geophysics, Vol 16, pp. 155–266, 1973.

[11] http://wdc.kugi.kyoto-u.ac.jp/dst_realtime/index.html

[12] https://space.fmi.fi/image/realtime/K/quasi_K_NUR.html

[13] Janne Hautsalo, Master's Thesis, Study of Aurora Related Sound and Electric Field Effects, Helsinki University of Technology, 2005. http://lib.tkk.fi/Dipl/2005/urn007898.pdf

[14] https://www.ilmatieteenlaitos.fi/revontulet-ja-avaruussaa

[15] Rudolf A. Treumann, Zbigniew Klos, Michel Parrot, Physics of electric discharges in atmospheric gases: an informal introduction, arXiv:0711.1672v2, 10 Apr 2008.

[16] Unto K. Laine, Auroral crackling sounds and Schumann resonances, 26th International Congress on Sound and Vibration, Montreal, Canada, 7-11 July 2019.

[17] Chao-Song Huang, J.C. Foster, P.J. Erickson, Effects of solar wind variations on the midlatitude ionosphere, Journal of the Geophysical Research Space Physics, 16 August 2002. https://doi.org/10.1029/2001JA009025.

[18] Unto K. Laine, Twenty Years Hunting for Auroral Sounds, Project document, 2021. https://www.researchgate.net/project/Auroral-Acoustics

[19] Robert L. McPherron, Magnetic pulsations: Their sources and relation to solar wind and geo-magnetic activity, Surveys in Geophysics (2005) 26:545–592, 2005.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Development of a Technical Memorandum Describing Optimal Room Acoustic Parameter Ranges for Musical Performance and Rehearsal Spaces

Evan Green[1,*,°], Michael Barron[°], Eckhard Kahle[1,°], Tor Halmrast[°], Winfried Lachenmayr[°], Henrik Möller[°], Uwe Stephenson[°], Louena Strepi[°], Bård Støfringsdal[°]

[1]Kahle Acoustics, Brussels, Belgium.

[°]Contributing primary author to the EAA Technical Memorandum.
A full list of Working Group members is provided in the Acknowledgements.

*egreen@kahle.be

Abstract
A working group has been convened by the European Acoustics Association to develop optimum Room Acoustic Parameter (RAP) ranges for musical performance and rehearsal spaces. A Technical Memorandum, outlining the proposals of the working group is in production: the main contents are summarized in this article. The main focus has been to develop optimum parameter ranges for concert halls, as well as for other spaces for music rehearsal and performance, such as chamber music halls and orchestra rehearsal halls.

Room acoustic quality is a multi-dimensional phenomenon: multiple RAPs are therefore required to quantify acoustic quality. The RAPs as currently implemented do not however provide a robust differentiation of acoustic quality – both good and bad sounding halls can fall within the RAP ranges currently proposed in the literature. The Technical Memorandum describes a selection process for RAP ranges, with the aim of setting projects on a good path towards acoustical excellence, and to improve the quantification of the outcome.

While the possibility to develop novel parameters was at first discussed, is has been agreed that the RAPs already defined in ISO 3382 Part 1 will be used. The RAPs are physically linked; therefore, the RAP optimum ranges must form an internally consistent set, taking into consideration their variation throughout the space. Other important aspects that must be considered in the definition of optimal RAPs that are often overlooked are the orchestra/ensemble size, music type or genre, room volume available, ideal loudness during loud (forte) passages, variation of parameters with acoustic volume and with distance from the source.

It has also been established that measurement protocols should be refined to improve the differentiation between spaces of differing acoustical quality.

Keywords: room acoustics, optimal parameters

1 Introduction

It has long been recognised that acoustical quality is a multi-dimensional phenomenon. Sabine established this 120 years ago in his paper "*Reverberation*" [1] describing three necessary acoustical properties "in order that hearing may be good in any auditorium":

- sufficient loudness;
- components of a complex sound should maintain their relative intensities;
- successive sounds in music or speech should be clear and free from extraneous noises.

Various studies over the last century have indicated numerous independent subjective acoustical aspects significantly exceeding Sabine's three separately perceivable subjective acoustical qualities. In parallel, a search for measurable objective acoustical parameters, each maximally correlated with a subjective acoustical quality, has led to the commonly used set of Room Acoustics Parameters (RAP) defined in the International Standard ISO 3382 "*Measurement of room acoustic parameters*", Part 1: Performance Spaces [2]:

Strength, G
Reverberation Time, T
Early Decay Time, EDT
Clarity Parameters, C50, C80
Definition Parameter, D50
Centre Time, Ts
Early Lateral Energy Fraction, JLF
Late Lateral Sound Level, LJ
Interaural Cross-Correlation Coefficient, IACC
Stage Support, STearly and STlate

Establishing a set of optimal RAPs for a particular space is not always straightforward. While there exists a wealth of information in the acoustics literature, it has been up to individuals to combine resources from a number of places, often stretching across decades, to understand which combination of RAPs would be optimal for a particular space based on the type of space, the ensemble(s) that will use it, the type of music (and other uses) envisaged, the need for variable acoustics, along with the space and volume available (or that can be afforded). In addition, while there might be substantial information on a particular parameter for halls of a particular size, there may be a lack of information for other situations. The Early Lateral Energy Fraction and Late Lateral Sound Level are good examples of this: while there is substantial literature on large concert halls, much less information is available on the ideal ranges for smaller rooms.

Nor is it simple to discriminate between spaces and their acoustical quality on the basis of a small number of RAPs, in particular if only room-averaged values are provided. While room-averaged data may point to very fundamental issues with a space – such as an excessively long or short reverberation time (for the use intended) – rarely does room-averaged data provide enough information to indicate location-sensitive issues. Room averages can hide the fact that parts of a room may exhibit very high values of a particular parameter while others are lacking. Taking Strength as an example: the fact that the room-average Strength is in an acceptable range does not help the listeners for whom the sound is excessively loud in some seats, yet too quiet in other seats. The same logic applies to any and all parameters that are sensitive to the local acoustical conditions at particular seats or groups of seats in a hall.

For these reasons, a Working Group has been convened by the European Acoustics Association to develop optimum Room Acoustic Parameter (RAP) ranges for musical performance and rehearsal spaces. A Technical Memorandum, outlining the proposals of the working group is in production: the main contents are summarized in this article. The main contributors to the text of the Technical Memorandum are listed as authors

here. A full list of the members of the Working Group is included at the end of this paper. The group is open to additional members – please contact the corresponding author of this paper for more information.

2 Current Sources for Room Acoustic Parameter Ranges

The mathematical definition of the above RAPs, along with requirements for their measurement, are contained in the Standard ISO 3382 Part 1. The standard also includes "typical ranges" for each parameter. These typical values are not optimal values, they are a record of the full range for each parameter, as measured in performance spaces. However, experience from the Working Group indicates that in some cases these parameter ranges have been come to be understood as "ideal" or "optimum" values. The “typical ranges” have appeared as requirements in briefs for new halls, unaltered from the ranges listed in the standard.

The “typical ranges” are however not optimal ranges: a performance space with room-average measured parameters within the "typical ranges" outlined in the standard could on the one hand have exceptional acoustical quality, on the other hand the space could equally well exhibit major acoustical flaws. In the current discussions to update ISO 3382, it should be considered whether the table containing “typical ranges” should be removed.

Furthermore, while it is clear that the RAPs for a particular space are physically related, it is common to see requirements for RAPs that are not physically possible to achieve in the same space. The RAP optimum ranges set out for a project must form an internally consistent set, taking into consideration the variation of the RAPs throughout the space.

Currently, accepted optimum ranges for RAPs tend to come from research papers and books. Well-known sources are, amongst others, the books by Beranek, Barron [3,4], who have independently surveyed many halls and correlated the measured parameters with the results of questionnaires into the subjective acoustical qualities. Some standards such as the German DIN 18041 [5] include recommendations for reverberation time dependent on volume, but do not provide recommendations for other parameters or indeed mention the important link between reverberation time, room volume and loudness (acoustic strength). The Norwegian Standard NS8178 [6], now adapted as ISO 23591 *Acoustic quality criteria for music rehearsal rooms and spaces* [7], has established a method for determining reverberation time, volume and strength, primarily to avoid that smaller rooms for music rehearsal are not too loud.

Previous research has tended to focus on individual RAPs, with "big picture" studies covering multiple aspects of music acoustics and the interconnections and relationships between the RAPs being much less common. Even though the relationships between RAPs - at least for a diffuse field case - are relatively simple to derive theoretically and are included in many textbooks, recommendations for RAPs tend to be "unlinked" with the mutual effects of RAPs not being taken into account (leading to the possibility, if not careful, of selecting physically impossible combinations).

Traditionally RT has been the main RAP to be set, and indeed for a long time it was the only RAP. Typically, RT is first set without reference to other acoustical aspects and as the main or only target for a project this can lead to problems, in particular if the room volume and the resulting acoustic strength and loudness are not considered in parallel. This reliance on the RT as the primary RAP is therefore to be considered.

3 Aim and Approach

The aim of the Working Group is to develop optimum ranges for RAPs for spaces for music, while simultaneously adding context to the RAPs and providing explanation regarding the implications of certain choices.

Early on in the process, it was discussed whether to expand the number of RAPs beyond those listed in ISO 3382. While there are valid criticisms regarding the common RAPs, it was decided to restrict the contents of the primary chapters of the Technical Memorandum to the parameters defined in ISO 3382. In addition, the standard does not include some commonly used RAPs such as bass and treble ratios, a chapter has been included in the Technical Memorandum to discuss alternative parameters and aspects of acoustics not covered by the common parameters.

The broader aim is to help acousticians – along with the clients, architects and artists that they work with – to put projects on a good track from the very conception of the project. In particular, the choices that are made at the start of a project are especially critical, since these fundamentally establish the acoustics and the potential acoustical quality of the space. As mentioned above, a major discussion point has been the relationship between room volume, reverberation time and strength. Room volume (as well as height, width, and other spatial attributes) is not an acoustic parameter in the typical sense, but it is critical to the acoustical quality. The balance between the size, loudness and reverberance of a space is an extremely important aspect to determine at the start of any project. Based on the approach developed in the Norwegian Standard NS8178 [6] and ISO 23591 [7], a method for establishing optimal relationships between these factors also for larger halls has been developed.

A further aim is to describe (and promote) methods and guidelines for conducting measurements, as well as for the presentation of measurement results, so that RAPs can be more effectively used to understand the detailed acoustics of a space and to discriminate between spaces of differing acoustical quality. A discrimination of many aspects of acoustical quality on the basis of the common RAPs, their implementation and presentation (typically as room averages) is extremely challenging, if not almost impossible. Experience from renovation projects, where measurements have been carried out before and after acoustical changes to a hall have been implemented provide an important insight [8]. In some cases, very significant, positive, audible changes to a hall have been made, yet the room-average RAPs indicate either a negligible change or a change in the opposite direction to that expected.

The conclusion of the Working Group is that measurements must in general be analysed and presented in a more detailed fashion, avoiding room averages. ISO 3382 Part 1 already includes a statement to this effect:

> "*The measurement results for the measures described in this annex [ISO 3382 Part 1 Annex A] should normally not be averaged over all microphone positions in a hall because the measures are assumed to describe local acoustical conditions. In the case of a large hall, it can be useful to average the results in some sections of the hall (stalls, first balcony etc.). Some measures such as sound strength, G, tend to vary with the distance, and a graphical plot of G as a function of source-receiver distance can be useful.*" [2]

In addition to plotting G as a function of distance (with comparisons to Barron's Revised Theory of Sound Level [9] providing a highly useful theoretical comparison), applying this approach to other parameters can elucidate more about the acoustics of a space. Mapping measurement results across audience and stage areas also provides a further method by which to gain understanding about the acoustics of a hall.

The intention is also to increase the reliability of comparisons made between measurements by different groups. A chapter in the Technical Memorandum is dedicated to refinements in measurement methods and the presentation of results.

A final aspect of the approach taken with the Technical Memorandum is to provide a logical framework for establishing optimal RAPs. Rather than presenting tables of parameter values, the intention is to describe methods and strategies by which optimum parameter value ranges can be determined based on the subjective aims and constraints of a particular project or space. The discussion in the Technical Memorandum is not intended to limit or restrict: it is still up to concert hall designers, acousticians, architects and clients to make their own decisions. The hope however is that with the information in the Technical Memorandum, more informed decisions can be made, based on a fuller understanding of the implications of the chosen RAP values, thereby putting the design on a good track from the start.

4 Scope and Structure

The scope of the Technical Memorandum focusses on spaces for unamplified music. At the lower limit are spaces for chamber music where a small audience can be accommodated – a lower limit of an audience of 50 has been discussed, leading to a space with a floor area of around 50m^2 and an acoustical volume of 200m^3. At the upper end of the scale, the Technical Memorandum intends to cover the largest concert halls for symphonic music with volumes up to 20,000 to 40,000m^3. Within this total range of 200m^3 to 40,000m^3 fall a whole host of room types and sizes, including recital halls, smaller concert halls, orchestra rehearsal halls and large recording spaces.

There is an overlap in the scale of rooms covered by the Technical Memorandum and the rehearsal spaces covered by the standard ISO 23591. This is intentional and members of the group who have developed NS8178 and ISO 23591 are part of the Working Group. It is intended that the documents align for these smaller spaces.

The Technical Memorandum is structured as follows, with a primary author determined for each chapter:

Introduction (Evan Green): Description of the scope of the document, aim, intent, approach and background information and definitions.

Selection of V, G, and RT (Uwe Stephenson): Discussion of the relationships between volume, strength and reverberation time leading to proposed strategies for establishing optimal values.

Reverberance (Winfried Lachenmayr): Selection of EDT and RT parameters. A point of discussion has been whether the EDT target should take precedence, with RT following the EDT.

Balance of Early and Late Sound (Michael Barron): Aspects relating to the selection of early and late energy balances and the parameters C80, Gearly, Glate.

Spatial Parameters (Louena Shtrepi): Discussion of the optimal values of Early Lateral Fraction, IACC and Late Lateral Level.

Stage Parameters (Eckhard Kahle): Analysis of stage support measurements, primarily STearly (ST1) leading to guidelines for STearly values dependent on hall volume and orchestra size and the distribution of STearly across the stage.

Variable Acoustics (Henrik Möller): Discussion regarding the range of variability for various parameters in multipurpose venues along with strategies for adjusting the acoustics.

Acoustics Aspects Not Covered by the Parameters (Tor Halmrast): Discussion of acoustics aspects not accounted for by the common parameters.

Measurement and Presentation of Parameters (Bård Støfringsdal)**:** Guidelines for measurements and presentation of results, including presentation as plots with source-receiver distance and mapping.

The Technical Memorandum is in production, and additional members to the Working Group are welcome to join for further discussion and review.

5 Conclusion

We hope that the Technical Memorandum will become a useful reference for the acoustics community, leading to more assurance regarding the acoustical quality of future music spaces.

The Working Group welcomes any comments and input to the development of the Technical Memorandum.

Acknowledgements

The production of the Technical Memorandum is the result of many interesting (and long!) discussions and important input from the entire Working Group. Many thanks to all who have contributed!

The following are members of the Working Group (at the time of writing and in alphabetical order):

Arianna Astolfi
Mike Barron
Vojto Chmelik
Jens Jørgen Dammerud
Daniel De la Prida Caballero
Dusan Dlhy
Jan Dolejší
Andor Tamas Fürjes
Anders Christian Gade
Evan Green
Ignacio Guillén
Tor Halmrast
Luc Jaounen
Eckhard Kahle
Winfried Lachenmayr
Ana Llopis
Tapio Lokki
Eckard Mommertz
Henrik Möller
Herbert Muellner
Christian Nocke
Jens Holger Rindel
Ivana Rossell
Lounea Shtrepi
Uwe Stephenson
Bård Støfringsdal
Daniel Urban
Michael Vorländer
Stefan Weinzierl
Stefan Weigand
Remy Wenmaekers

References

[1] W.C. Sabine. Reverberation. *The American Architect*, 1900.

[2] ISO 3382-1:2009 Acoustics – Measurement of Room Acoustics Parameters. Part 1: Performance Spaces

[3] L Beranek. Concert Halls and Opera Houses: Music, Acoustics, and Architecture. 2nd Edition. *Springer*, 2003.

[4] M. Barron. Architectural Acoustics and Auditorium Design. *Spon Press*, 2009.

[5] DIN 18041 Acoustic quality in rooms - Specifications and instructions for the room acoustic design. *Deutsches Institut Für Normung*, 2016

[6] NS 8178 Acoustic criteria for rooms and rooms for music performance. 2014 (Discontinued)

[7] ISO 23591:2021 Acoustic quality criteria for music rehearsal rooms and spaces.

[8] T. Halmrast. Acoustical aspects not covered by the common, standardised room acoustic parameters. *Proceedings BNAM*, 2022.

[9] Barron, M. and Lee, L.-J. Energy relations in concert auditoriums, I. *Journal of the Acoustical Society of America, 84, 618–628*, 1988.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Comparison of Speech Reception Thresholds for diotic, dichotic and antiphasic headphone presentations of digits-in-noise triplets using Dantale I material

Palle Rye[1,*], Rasmus Overgaard Andersen[1], Gert Ravn[2], Dorte Hammershøi[1]

[1]Department of Electronic Systems, AI and Sound, Aalborg, Denmark.

[2] Teknisk Audiologisk Laboratorium (TAL), FORCE Technology, Odense, Denmark.

*par@es.aau.dk

Abstract
A complete and accurate diagnosis is the prerequisite of efficient treatment of any disorder. For hearing disorders, the primary diagnostic tools for characterizing the type and severity typically include a pure-tone audiogram describing the sensitivity loss **across** the frequency range typical for speech and one or more speech recognition tests. While there is typically a high level of correlation between the audiogram and the speech recognition tests, the speech test may reveal different characteristics of the hearing disorders. The research interest in out-of-clinic versions of speech in noise has increased in recent years. Different implementations exist in various languages, e.g. the digits-in-noise triplet test made available by the World Health Organization. The present study examines a similar triplet test featuring digits in Danish originating from the DANTALE I material. The original masking noise of the DANTALE I material did not contain sufficient high-frequency content to mask the digits effectively and equally well, so a new masking noise was derived from the digit material. Three different spatial configurations of the presented speech in noise were tested: diotic speech in dichotic noise, and antiphasic speech in both diotic and dichotic noise. Initial test on 18 subjects (5 subjects with better hearing ear pure-tone average (PTA_{BE}) > 30dB) in the age range 53-81 years (mean 64.4 years) shown reasonable correlation ($r > 0.86$ for all three configurations) between PTA_{BE} and the estimated speech reception thresholds.
Keywords: speech audiometry, digits-in-noise, antiphasic speech

1 Introduction

For decades air conducted pure tone audiometry has remained the gold standard in screening and the mainstay in diagnosing hearing disorders. Methods, equipment and environment requirements are standardised[1], and are widely accepted to provide useful, accurate and repeatable results. Typically performed in a clinical environment, it relies on a well-controlled acoustic environment with low noise levels, regularly calibrated transducers, and hearing health-care professionals for administering the test with suitable patient guidance. Although many implementations of automatic procedures for estimating the pure tone threshold exist, some even self-administered, it has yet to gain widespread in-clinic use. The goal of hearing rehabilitation is predominately improving speech understanding. A high level of correlation between the pure tone thresholds and the performance in speech recognition tests allows either to be used as an estimator for the other. Many factors influence speech perception[2], from the type of speech material, the presenting speaker, acoustic environment, hearing characteristics and cognitive abilities of the subject, response collection and scoring method. As with pure tone threshold audiometry, the basic methods for speech recognition tests are standardised[3], with an important difference of being inherently language-dependent.

In current Danish clinical practice, two speech audiometric procedures are routinely used[4]: the supra-threshold Word Recognition Score (WRS), counting the percentage of words discerned correctly and the Speech Reception Threshold (SRT), estimating the speech level where the test subject understands at least half of the presented material. The speech material for these test use lists of unrelated monosyllabic words from DANTALE I [5] and are typically performed in quiet without competing noise. The change in performance as a function of intensity change is described by the intelligibility slope, regarded as a smooth function from zero to full intelligibility over a range of intensities. The intelligibility slope steepness has been found to increase when a competing noise is presented simultaneously with speech sentences[6], [7]. Annex C of ISO8253-3[3] lists typical reference SRT using example data from [7], where the steepness of the intelligibility slope at the 50% point was 11%/dB in quiet and 19%/dB in competing noise. Sentence lists are more complex than monosyllabic word lists, and it is challenging to compile sentence lists that provide equivalent results. Notable examples of this type of hearing-in-noise-tests (HINT) for different languages include English[8], Swedish[9], Danish[10], and German[11]. Currently, HINT relies on operator interaction for response collection and scoring, inhibiting its use in a self-administered scenario.

The digits-in-noise (DIN) test, or digit-triplet test, was developed as a self-administered speech-in-noise test [12], which sought to screen the Dutch population over landline phones[13]. Essentially a measure of the SRT in competing noise, the suprathreshold test does not rely on carefully calibrated transducers or low ambient noise levels yet exhibits high test-retest reliability and high correlation with pure tone threshold average (PTA) calculated from four frequencies 0.5, 1k, 2k and 4kHz. Following the initial Dutch advances, the DIN test was adapted to different languages, targeted populations, and ubiquitous test platforms, such as smartphones apps and internet browsers. A recent scoping review[14] provides a comprehensive overview and notes the intelligibility slope steepness being in the range of 15-20%/dB across most of the reviewed studies and language adaptations. The earliest DIN studies tend to use the monaural presentation of both speech and noise sequentially testing each ear. Later studies used diotic tests with the same stimuli simultaneously in both ears (the nomenclature N_0S_0 signifies that there are no interaural differences for neither the noise, N, nor the stimuli, S, in this case, speech). While the diotic test halves the test time compared to each ear in succession, it relies strongly on the better hearing ear, and asymmetric hearing losses may be missed in the screening. A dichotic noise, N_u, timewise completely uncorrelated between the two ears yet has the same long-term spectrum as diotic speech S_0, appears to have different spatial characteristics than a diotic signal fully correlated between the ears. For the same volume setting, binaural summation of loudness means N_u is perceived to be louder than N_0. Nevertheless, due to binaural unmasking, the reception threshold of a N_uS_0 spatial configuration can be expected to be a few dB lower than N_0S_0. Such threshold improvement due to binaural unmasking is called the binaural masking level difference (BMLD)[15]. For diotic noise combined with antiphasic stimuli, N_0S_π, meaning the stimuli is in opposite phase between ears, binaural unmasking also enhances the stimuli detection relative to N_0 S_0 resulting in BMLD. Recent studies have found a higher correlation between worse hearing ear PTA (PTA_{WE}) and antiphasic DIN, which improves the screening sensitivity to asymmetric hearing impairment[16]. A mobile application featuring antiphasic DIN was adopted by the World Health Organization[17].

In recent years, the interest in eHealth research and out-of-clinic hearing healthcare has been surging[18]. Examination of the performance of supervised clinical methods versus self-administered methods in ecologically valid out-of-clinic settings can help reveal potential critical issues and benefits of out-of-clinic approaches. Previous research established several benefits of favouring antiphasic speech over the diotic speech with diotic noise. The primary goal of the present study is to examine differences between three other spatial configurations to determine any potential advantages of using dichotic uncorrelated noise. Secondly, it is of interest whether some spatial configurations of DIN correlate with HINT measurements encouraging further research into applying unsupervised DIN speech audiometry as a diagnostic measure or even as hearing aid fitting validation tool, duly noting that relying on presenting antiphasic speech could prove challenging in the latter case.

2 Methods and Materials

2.1 Participants

Eighteen listeners (8 female) aged between 53 and 81 years (mean 64.4 years) participated in the study. Four participants were experienced hearing aid users, two with moderate impairment (PTA 35-50 dB)[19], one with moderate-to-severe (PTA 50-65dB) and one with strongly asymmetric hearing impairment. Eleven participants had normal hearing (PTA<20), one with moderate hearing impairment and the remaining with mild hearing impairment (PTA 20-35dB).

2.2 Test platform and setup

The DIN test was programmed in C#, compiled on the Unity platform and installed on an Asus Zenpad 3S10 Android tablet. The participants entered age and gender data and answered a 12-item computer self-efficacy questionnaire before starting with the DIN tests, followed by various other self-administered audiological tests and questionnaires. A familiarisation phase with five triplet presentations preceded data collection of the first spatial configuration. The total test time of the complete test battery including exit interviews was designed to be conducted in less than one hour per participant. A set of active noise-cancelling headphones BOSE QC35II connected to the tablet via Bluetooth was used for presenting the stimuli. All tests were conducted in an out-of-clinic environment typical to the intended use case, and ambient noise levels were monitored. A rest period was provided between the self-administered test battery and the more traditional audiometric tests. For eight participants the tests were conducted on the same day after the rest period, while the other ten had a minimum of two weeks between the self-administered and traditional tests.

Pure tone thresholds were measured manually according to ISO 8253-1[1] using the audiometer functionality of the InterAcoustics Affinity Test Suite connected to a RadioEar DD45 headphone. The HINT was evaluated using a set of MATLAB scripts from the original author of the Danish procedure[10]. The scripts were executed on a laptop controlling a soundcard RME Fireface UCX connected to Sennheiser HDA200 headphones. Headphones for all experiments were calibrated using an artificial ear coupler B&K Type 4153.

2.3 Speech material and adaptive procedure

Seven monosyllabic numbers (0, 1, 2, 3, 5, 6 and 7) were extracted from DANTALE I list and recombined into random digit triplet sequences. Each triplet sequence contains three different digits, and the time between each digit is 0.5 seconds. The playback noise level was fixed to 60dB SPL throughout the test unless the adaptive algorithm should converge towards providing a positive SNR above +20dB. In such a case the noise level would be reduced instead to avoid unnaturally loud speech levels. The initial SNR was set to 0dB. An adaptive 1up-1down procedure was used for each correct triplet. A digit-scoring approach was used where a triplet was considered correct when two out of the three digits were correctly identified. Antiphasic speech was simple phase-reversal of one channel. Due to BMLD, considerable variation in SRT across the three spatial conditions was likely to occur. Indeed pilot tests revealed SRT differences in excess of 10dB between the three spatial conditions. A procedure with an adaptive step size was chosen to start from the same initial SNR across the spatial conditions, where the initial step size of the SNR was set to 5dB to allow fast descent toward the threshold. A total of 8 reversals was chosen as a compromise between a manageable number of trials for each spatial condition and an adequate number of trials around the threshold. Following each of the first three reversals the step size was reduced by 1dB, with three reversals of 2dB step size before finishing 1dB for the final two reversals. The threshold was calculated as the average of the last five reversals.

2.4 Masking noise

Following analysis of psychometric curves of individual digits in a pilot test, changing the speech-shaped masking noise from DANTALE I was deemed necessary, since digits 3, 6 and 7 were recognisable at

significantly lower SNRs than the other digits. A similar issue was reported earlier for the DANTALE I sentence material[20], recommending that certain word lists be avoided when using DANTALE I masking noise. For the present study a new speech-shaped noise was generated by filtering Gaussian random noise to have the same long-term spectrum as a sequence of all digits, thus containing more high-frequency energy. Uncorrelated channels were achieved by using random playback points in the noise wave file separated by at least 5 seconds. Third-octave analysis of the speech and noise signals are shown in Figure 1 for DANTALE I and Figure 2 for the noise used in this study.

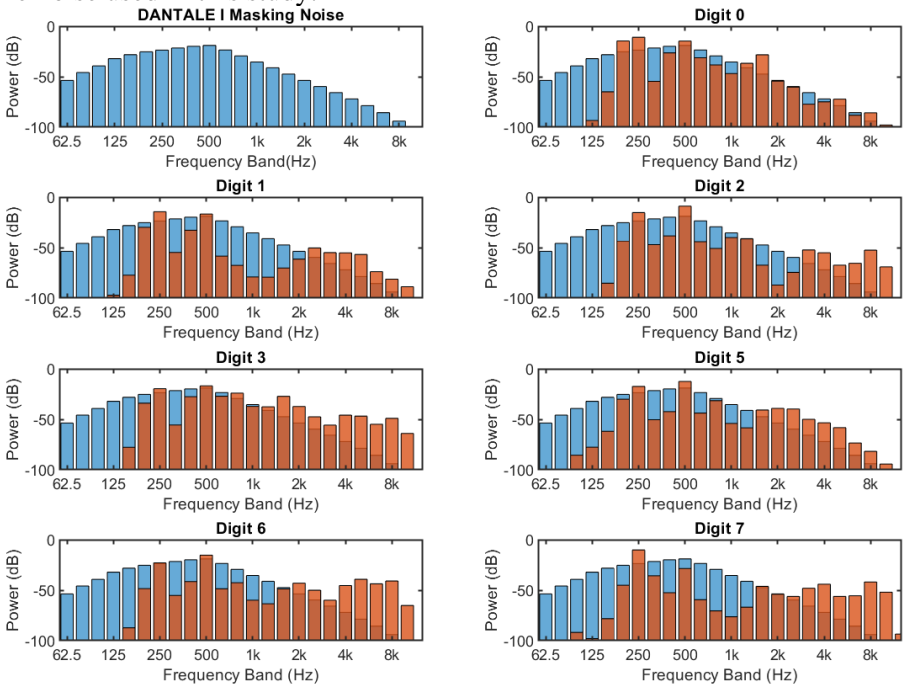


Figure 1: Third-octave spectra of the individual digits with the masking noise from DANTALE I

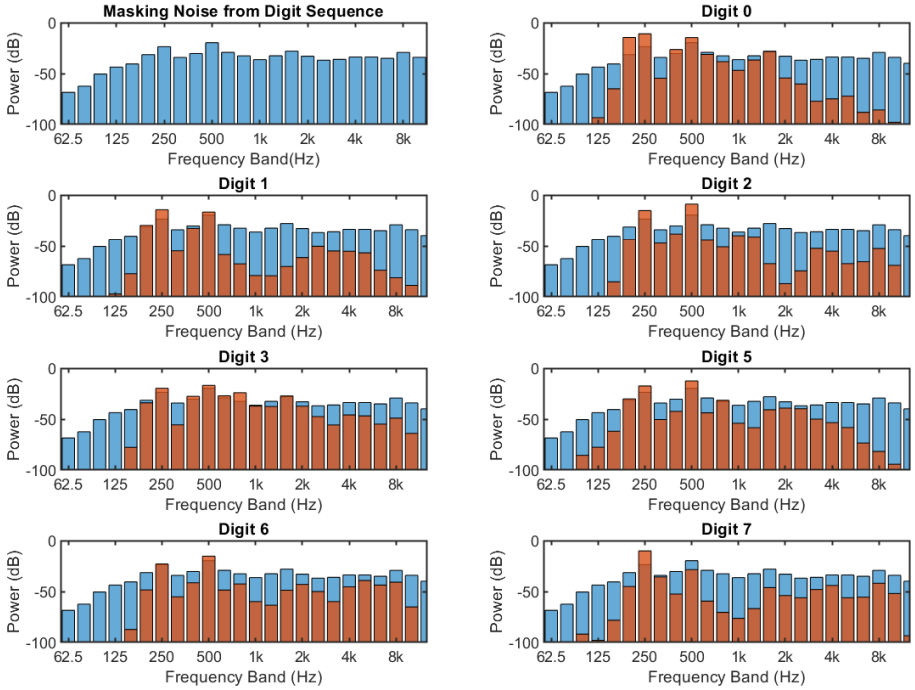


Figure 2: Third-octave spectra of the individual digits with the speech shaped noise used in this study

3 Results

The average number of digit triplet trials for determining the SRT was 15.3 (SD=2.9). The total trial time for the three spatial conditions was, on average 6.6 (SD=1.4) minutes. Two typical examples of the adaptive staircase for the three spatial conditions are shown in Figure 3.

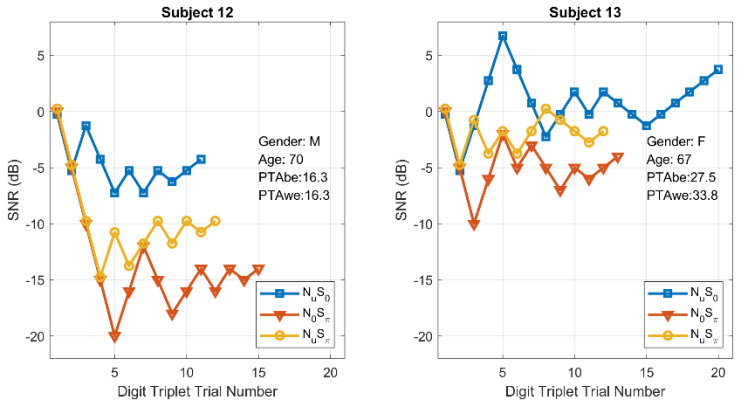


Figure 3: Typical examples of the adaptive staircases for the three spatial configurations.

Linear regression fit including 95% confidence intervals (CI) are shown in Figure 4, where each column represents a spatial condition so that BMLD may be revealed as a systematic offset across columns. Intercepts of N_uS_0 and N_uS_π differ by 2.3 dB for PTA_{BE} and 2.8 dB for PTA_{WE}. The middle column N_0S_π is clearly offset downwards compared to the other conditions and features a steeper regression slope allowing for greater separation of PTA in a screening scenario. Top row (better hearing ear) regression features steeper slopes than bottom row (worse hearing ear) and higher correlation in two of three conditions. Regressed DIN SRT from PTA (black lines of Figure 4) illustrates potential separation of normal hearing (<20dB), from mild (20-35 dB), moderate (35-50dB) and moderate-severe (50-65dB) hearing configurations.

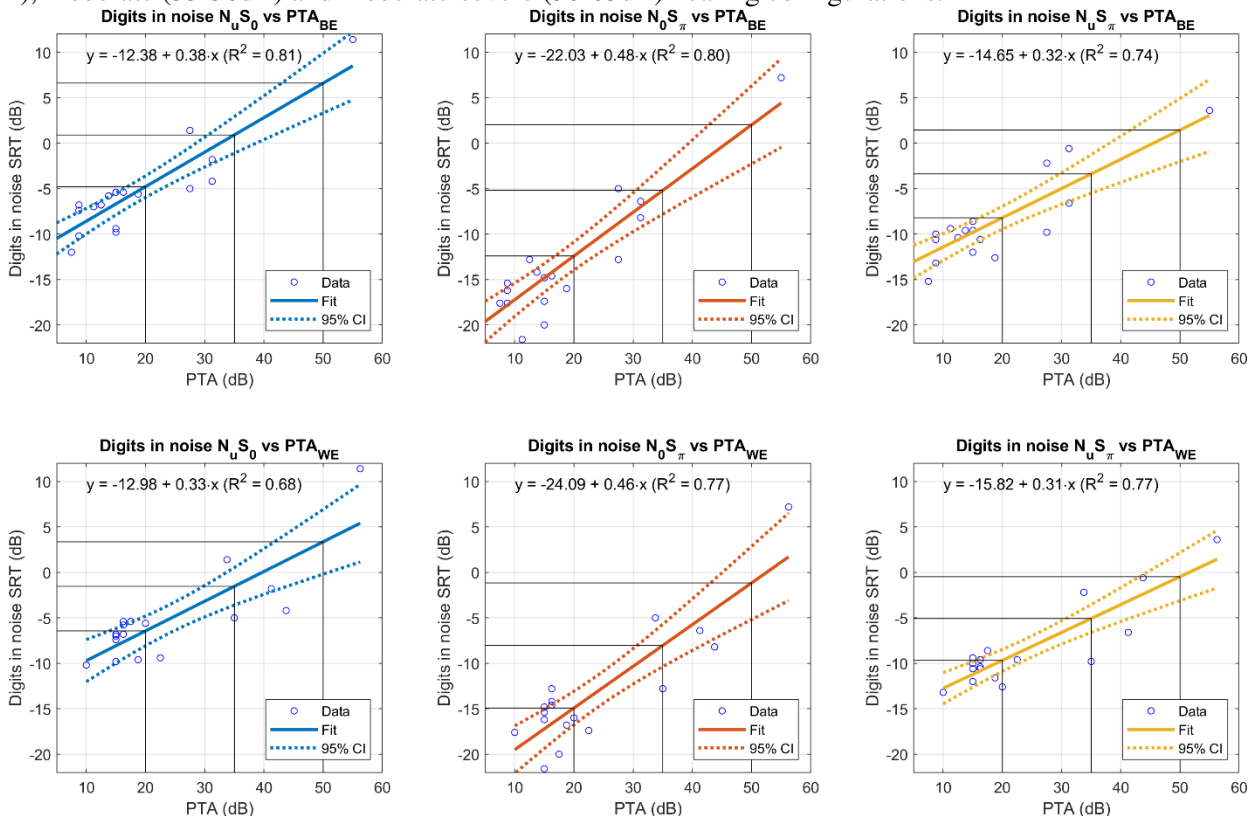


Figure 4: Linear regression plots with PTA as predictor and DIN SRT as the response variable.

Linear regression relating DIN SRT for all three spatial conditions across columns to HINT thresholds for diotic $HINT_{Diotic}$ (N_0S_0), better hearing ear ($HINT_{BE}$) and worse hearing ear ($HINT_{WE}$) across rows is shown in Figure 5. Shallow slopes are generally observed along with relatively poor correlation over narrow ranges (5.8, 5.9 and 7.4 dB) of HINT SRT with sizable variance (1.9, 3.4 and 5.0 dB) determined for this population. The steepest slopes and highest correlations are found in the bottom row featuring the $HINT_{WE}$. Here the groups with and without hearing impairment also appear more clearly separated.

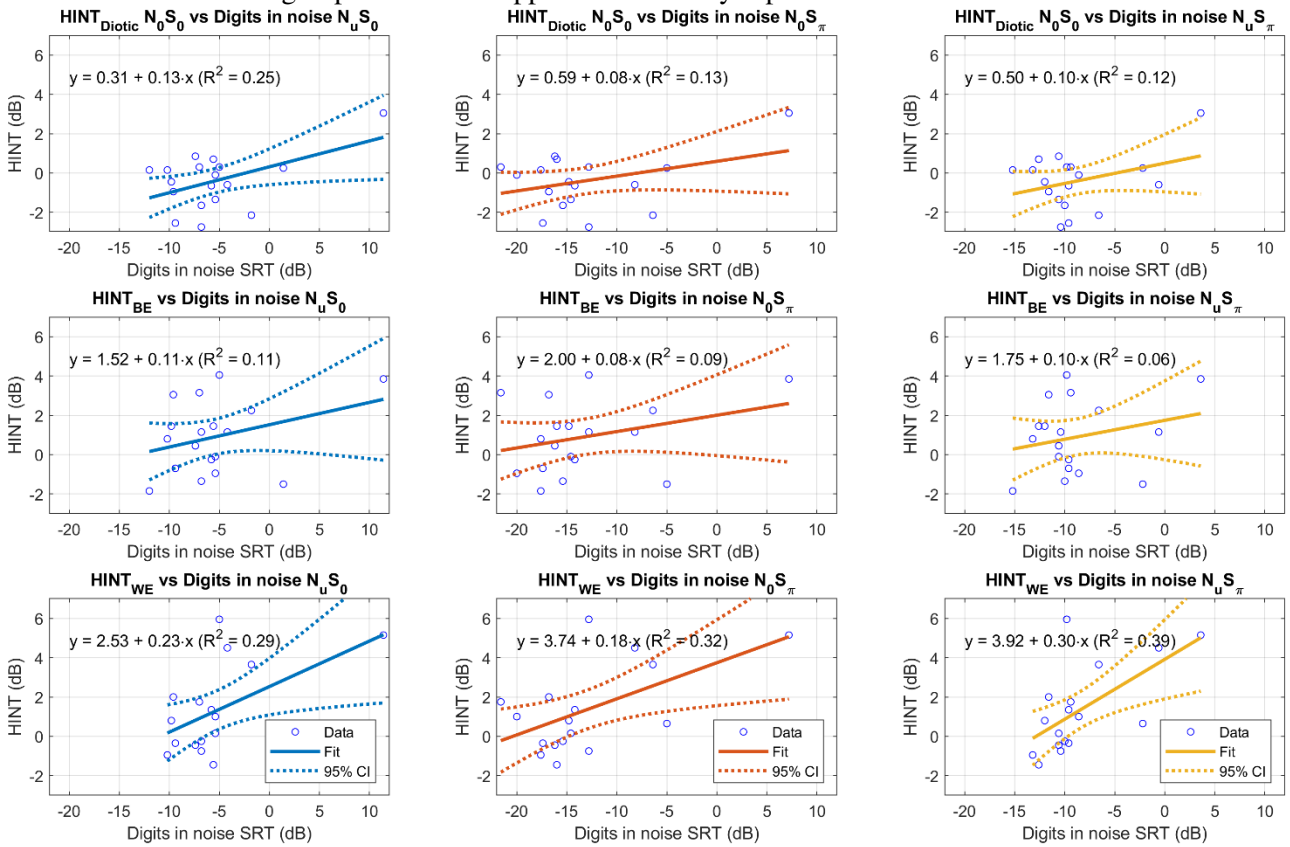


Figure 5: Linear regression plots with DIN SRT as predictor and HINT as the response variable

The correlation matrix is shown in Table 1. The studied population has significant PTA symmetry between ears (r = 0.93) and between each of the three spatial conditions (r < 0.85). Significant correlations are found between all DIN spatial configurations and PTA_{BE} (r > 0.86) and PTA_{WE} (r > 0.83). The $HINT_{WE}$ correlates significantly with PTA_{WE} (r = 0.79), while the $HINT_{BE}$ does not correlate significantly with PTA_{BE} (r = 0.35).

Table 1: Correlation between PTA at the better and worse hearing ear, the measured DIN SRT levels for the three spatial configurations, and HINT thresholds for better hearing ear, worse hearing ear and diotic condition; Correlations in bold are significant at the 0.001 level.

	PTA_{BE}	PTA_{WE}	N_uS_0	N_0S_π	N_uS_π	$HINT_{Diotic}$	$HINT_{BE}$	$HINT_{WE}$
PTA_{BE}	1							
PTA_{WE}	**0.93**	1						
N_uS_0	**0.90**	**0.83**	1					
N_0S_π	**0.89**	**0.88**	**0.90**	1				
N_uS_π	**0.86**	**0.88**	**0.87**	**0.85**	1			
$HINT_{Diotic}$	0.47	0.35	0.56	0.39	0.41	1		
$HINT_{BE}$	0.35	0.41	0.24	0.27	0.14	0.38	1	
$HINT_{WE}$	0.68	**0.79**	0.54	0.56	0.62	0.32	0.64	1

4 Discussion

For the sake of reducing test time in a large test battery, short trial lengths (average of 15.3 triplet trials per spatial configuration) were achieved by using an adaptive procedure that was based on 8 reversals and reducing step size. The correlation between SRT and PTA was reported to change only slightly after 15 trials of 2dB in an earlier study[21]. A recent study[22] reported SRT test-retest differences for adults of 0.98 dB (SD = 2.91) with only 6 reversals and an average of 23.2 (SD = 4.1) trials with an adaptive step size changing from 6 dB initially to 3 dB after first reversal. The favourable number of trials per reversal was achieved by changing from a 1-down-1-up to a 2-down-1-up paradigm after the first two reversals.

A similar approach could be adopted for future implementations of the DIN in the present study. This could prove beneficial, especially if the slopes of the psychometric function for the DIN are found to be less steep for individuals of a larger population than examined here. Various impairments and characteristics of hearing may be imagined where the SRT estimation of a test subject would benefit from a more trials around threshold. This could also be achieved by selecting more than 8 reversals, fixing a minimum number of trials, or using a stopping criteria based on observed individual threshold variance[23]. The high correlation between antiphasic speech in diotic noise and PTA_{WE} (r = 0.88) is similar to the findings of antiphasic DIN studies in other languages, e.g. South-African English (r = 0.82)[16] and French (r = 0.82)[24].

The observed offset between intercepts of N_uS_0 and N_uS_π spatial configurations of 2.3 dB and 2.8 dB for better and worse hearing ear respectively are in agreement with a BMLD of about 2 dB described in [15] for tonal stimuli in uncorrelated noise. However, uncorrelated noise, N_u, yielded flatter regression slopes with PTA in both conditions compared to N_0 suggesting no benefit for screening purposes. Due to the steeper slopes found in the correlation with PTA, the present study suggest antiphasic speech in diotic noise (N_0S_π) is more appropriate for screening than the other spatial configurations.

For diagnostic purposes it may be beneficial to find a measure that is not directly correlated with PTA yet may reveal other deficits than loudness recruitment. A sentence-based HINT, arguably a more ecologically valid candidate, might suit such purpose for in-clinic use. With DIN as a potential candidate for self-administered out-of-clinic diagnostic test any significant correlation with HINT would encourage further research into the suitability. However, low correlation between DIN and HINT measures was observed for the tested population, with possible exception for DIN and HINT worse hearing ear performance. The HINT SRT for the tested population exhibit sizable variance up to over a relatively narrow range (up to 5.0 dB over a 7.4 dB range). This could be indicative of very steep psychometric slopes for sentence-based paradigm, where a complete trial fails even when getting single word wrong (except for tense, order and so on). In HINT some test subjects might not feel entirely comfortable with guessing and opt to pass if they do not understand the complete sentence. Even when one or more digits are not understood, the DIN method forces the subject to guess the unheard digits in order to proceed. Inspection of responses well below threshold revealed how some subjects entered triplets, which could be interpreted as opting to pass a specific trial: e.g. 1-2-3 or 0-0-0 in case of completely missing a triplet.

5 Conclusions

The goal of the present study was to examine differences between three spatial configurations of digits in noise and determine potential advantages of using dichotic uncorrelated noise with diotic or antiphasic speech. For screening purposes no advantages of dichotic uncorrelated noise were found. In line with previous research, antiphasic speech in diotic noise exhibited significant correlation worse ear PTA and remains the strongest candidate of the three examined spatial configurations for screening purposes. The study did not find convincing correlation between monosyllabic DIN and sentence-based HINT measurements, however the research interest in using unsupervised DIN speech audiometry as a diagnostic measure or hearing aid fitting validation tool remains.

Acknowledgements

Collaboration and support by Innovation Fund Denmark (Grand Solutions 5164-00011B); Oticon, G.N. Hearing, Widex-Sivantos Audiology, and other partners (Aalborg University Hospital, Odense University Hospital, Aalborg University, Technical University of Denmark, FORCE Technology and Copenhagen University Hospital) is sincerely acknowledged

References

[1] ISO 8253-1, *Acoustics – Audiometric test methods – Part 1 : Pure-tone air and bone conduction audiometry*. International Organization for Standardization, 2010.
[2] Working Group on Speech Understanding, "Speech understanding and aging," *J. Acoust. Soc. Am.*, vol. 83, no. 3, pp. 859–895, Mar. 1988, doi: 10.1121/1.395965.
[3] ISO 8253-3, *Acoustics – Audiometric test methods – Part 3: Speech audiometry*. International Organization for Standardization, 2022.
[4] A. Wolff, *Health-Related Quality of Life Following Hearing Aid Treatment a large Cohort study*. Aalborg Universitetsforlag, 2019.
[5] C. Elberling, C. Ludvigsen, and P. E. Lyregaard, "Dantale: A New Danish Speech Material," *Scand. Audiol.*, vol. 18, no. 3, pp. 169–175, Jan. 1989, doi: 10.3109/01050398909070742.
[6] R. Plomp and A. M. Mimpen, "Improving the Reliability of Testing the Speech Reception Threshold for Sentences," *Int. J. Audiol.*, vol. 18, no. 1, pp. 43–52, Jan. 1979, doi: 10.3109/00206097909072618.
[7] B. Kollmeier and M. Wesselkamp, "Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment," *J. Acoust. Soc. Am.*, vol. 102, no. 4, pp. 2412–2421, Oct. 1997, doi: 10.1121/1.419624.
[8] M. Nilsson, S. D. Soli, and J. A. Sullivan, "Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise," *J. Acoust. Soc. Am.*, vol. 95, no. 2, pp. 1085–1099, Feb. 1994, doi: 10.1121/1.408469.
[9] M. Hällgren, B. Larsby, and S. Arlinger, "A Swedish version of the Hearing In Noise Test (HINT) for measurement of speech recognition," *Int. J. Audiol.*, vol. 45, no. 4, pp. 227–237, Jan. 2006, doi: 10.1080/14992020500429583.
[10] J. B. Nielsen and T. Dau, "The Danish hearing in noise test," *Int. J. Audiol.*, vol. 50, no. 3, pp. 202–208, Mar. 2011, doi: 10.3109/14992027.2010.524254.
[11] J. Joiko, A. Bohnert, S. Strieth, S. D. Soli, and T. Rader, "The German hearing in noise test," *Int. J. Audiol.*, vol. 60, no. 11, pp. 927–933, Nov. 2021, doi: 10.1080/14992027.2020.1837969.
[12] C. Smits, T. S. Kapteyn, and T. Houtgast, "Development and validation of an automatic speech-in-noise screening test by telephone," *Int. J. Audiol.*, vol. 43, no. 1, pp. 15–28, Jan. 2004, doi: 10.1080/14992020400050004.
[13] C. Smits, P. Merkus, and T. Houtgast, "How we do it : The Dutch functional hearing – screening tests by telephone and internet," *Clin. Otolaryngol.*, vol. 31, no. 5, pp. 436–440, 2006, doi: https://doi.org/10.1111/j.1749-4486.2006.01195.x.
[14] E. Van den Borre, S. Denys, A. van Wieringen, and J. Wouters, "The digit triplet test: a scoping review," *Int. J. Audiol.*, vol. 60, no. 12, pp. 946–963, Dec. 2021, doi: 10.1080/14992027.2021.1902579.
[15] J. F. Culling and M. Lavandier, "Binaural Unmasking and Spatial Release from Masking," in *Binaural Hearing*, vol. 73, R. Y. Litovsky, M. J. Goupell, R. R. Fay, and A. N. Popper, Eds. Cham: Springer International Publishing, 2021, pp. 209–241. doi: 10.1007/978-3-030-57100-9_8.
[16] K. C. De Sousa, D. W. Swanepoel, D. R. Moore, H. C. Myburgh, and C. Smits, "Improving Sensitivity of the Digits-In-Noise Test Using Antiphasic Stimuli," *Ear Hear.*, vol. 41, no. 2, pp. 442–450, Mar. 2020, doi: 10.1097/AUD.0000000000000775.
[17] D. W. Swanepoel, K. C. De Sousa, C. Smits, and D. R. Moore, "Mobile applications to detect hearing impairment: opportunities and challenges," *Bull. World Health Organ.*, vol. 97, no. 10, pp. 717–718, Oct. 2019, doi: 10.2471/BLT.18.227728.

[18] A. Paglialonga, A. Cleveland Nielsen, E. Ingo, C. Barr, and A. Laplante-Lévesque, "eHealth and the hearing aid adult patient journey: A state-of-the-art review," *Biomed. Eng. Online*, vol. 17, no. 1, pp. 1–26, 2018, doi: 10.1186/s12938-018-0531-3.
[19] G. Stevens, S. Flaxman, E. Brunskill, M. Mascarenhas, C. D. Mathers, and M. Finucane, "Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries," *Eur. J. Public Health*, vol. 23, no. 1, pp. 146–152, Feb. 2013, doi: 10.1093/eurpub/ckr176.
[20] G. Keidser, "Normative Data in Quiet and in Noise for 'Dantale'—A Danish Speech Material," *Scand. Audiol.*, vol. 22, no. 4, pp. 231–236, Jan. 1993, doi: 10.3109/01050399309047474.
[21] C. S. Watson, G. R. Kidd, J. D. Miller, C. Smits, and L. E. Humes, "Telephone Screening Tests for Functionally Impaired Hearing: Current Use in Seven Countries and Development of a US Version," *J. Am. Acad. Audiol.*, vol. 23, no. 10, pp. 757–767, Nov. 2012, doi: 10.3766/jaaa.23.10.2.
[22] D. R. Moore *et al.*, "FreeHear: A New Sound-Field Speech-in-Babble Hearing Assessment Tool," *Trends Hear.*, vol. 23, p. 233121651987237, Jan. 2019, doi: 10.1177/2331216519872378.
[23] H. Dillon, E. F. Beach, J. Seymour, L. Carter, and M. Golding, "Development of Telscreen: a telephone-based speech-in-noise hearing screening test with a novel masking noise and scoring procedure," *Int. J. Audiol.*, vol. 55, no. 8, pp. 463–471, 2016, doi: 10.3109/14992027.2016.1172268.
[24] J.-C. Ceccato *et al.*, "French Version of the Antiphasic Digits-in-Noise Test for Smartphone Hearing Screening," *Front. Public Health*, vol. 9, p. 725080, Oct. 2021, doi: 10.3389/fpubh.2021.725080.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Whole glass facade in office building – Measured noise level and requirement for facade

Bernt Mikal Larsen[1*,]

[1]Department of Acoustics, Multiconsult Norway, Kristiansand, Norway.

*Bernt.Mikal.Larsen@multiconsult.no

Abstract

The presentation will summarize calculated and measured noise level from road traffic in office building with whole glass facade. Measurements from project show that the difference between laboratory value and field value is as much as 8-10 dB for a whole glass facade. The difference is explained by the fact that a whole glass facade needs a total correction for both the effect of weakening due to profile system and area correction due to the size/dimension of the glass. Due to both of this effects, the required sound isolation from laboratory should normally be at least 8-10 dB higher than the value achieved for the facade in field. The experience is based on a new office building called Baneheia Park in Kristiansand in Norway. With a whole glassfacade with R_w+C_{tr} 46 dB in a noisy situation, both calculated and measured noise level from road traftic was L_d 39-40 dB. In Norway the required noise level in offices is L_d 35 dB from road traffic. If effect of reduced sound isolation due to profile system with polyethylene inside of the rebate (4-6 dB) and reduced effect of sound isolation due to area/dimension correction (3-4 dB) were included, the requirement for the facade in the given situation should have been minimum R_w+C_{tr} 51 dB. With such facade, the indoor noise level of L_d 35 dB would have been achieved.

1 Introduction

A new office building in Kristiansand (Baneheia Park) was designed with a whole glass facade. In the phase of designing there were different views between two acoustic companies about necessary requirement for the facade. This document gives a summary of the noise levels, the chosen facade and conclusions to be useful for similar future projects.

2 Design parameters – outdoor and indoor level

2.1 Outdoor noise level – free field value

Baneheia Park (Fjellgata 6) is located by a tunnel along a road with heavy traffic. The outdoor noise level at the office building was calculated by two companies to be L_d 74-75 dBA free field. For design of the facade L_d 75 dB was used. The traffic situation on the main road (E18) is approximately the same today and with future road system. In 2019 outdoor registration of noise for a couple of days on the building Banehaven 4A showed L_d of 69-70 dB (figure 1a and 1b) in periods with normal traffic, and very good match between this measured level and calculated L_d of 68 dBA in a calculation model on the facade/position where measurements were done.

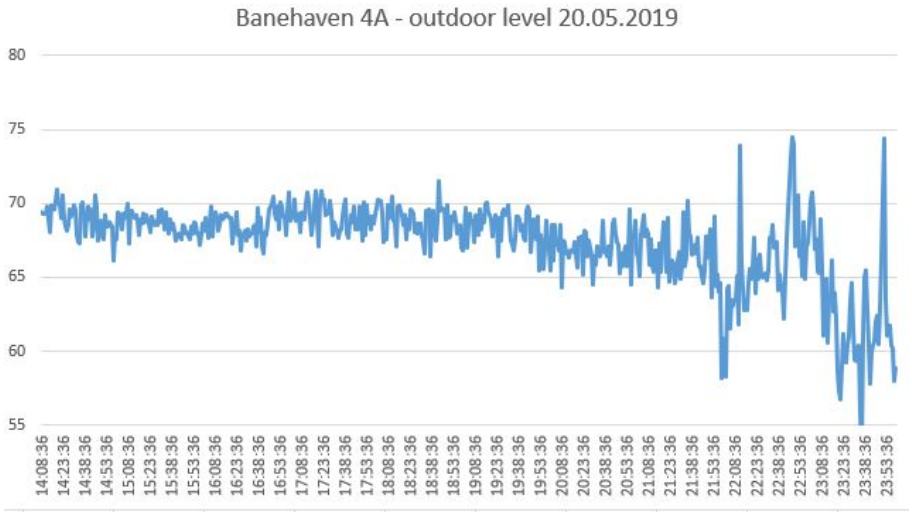


Figure 1a Outdoor measured level for Banehaven 4A at 20th of May 2019 – L_d 69 dBA from 16-19

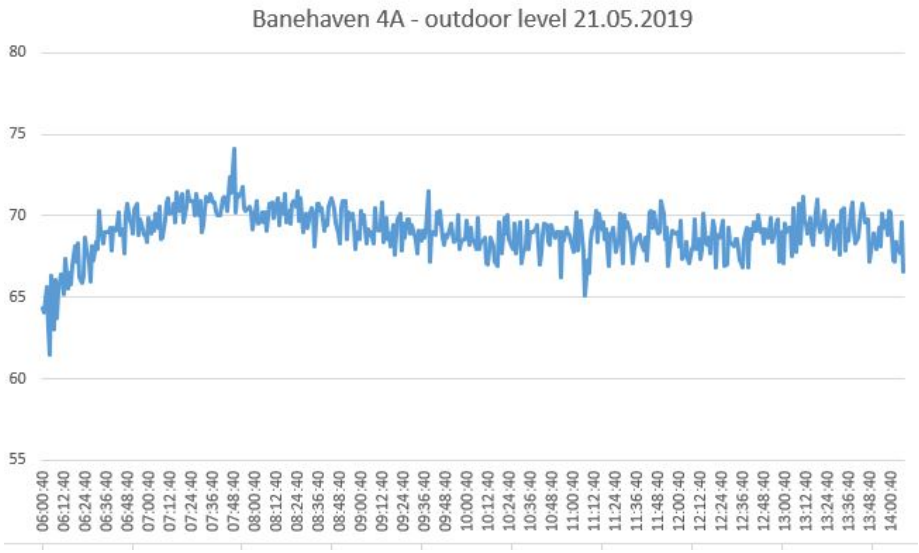


Figure 1b Outdoor measured level for Banehaven 4A at 21th of May 2019 – L_d 69-70 dBA from 08-10

The calculation model (figure 1c) also shows that representative L_d of 68 dBA at Banehaven 4A corresponds very well with representative L_d of 74 dB outside Fjellgata 6 (Baneheia Park).

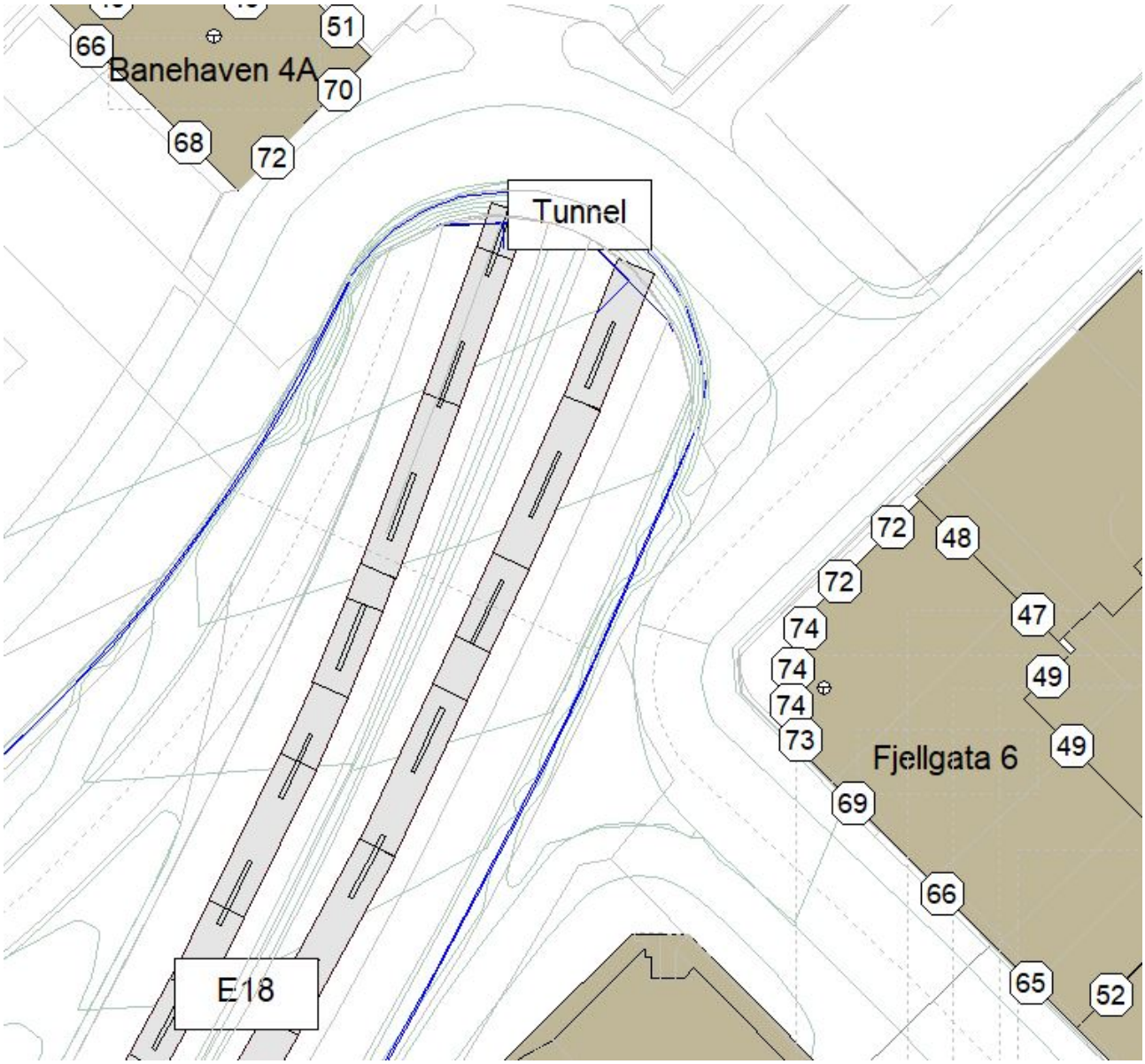


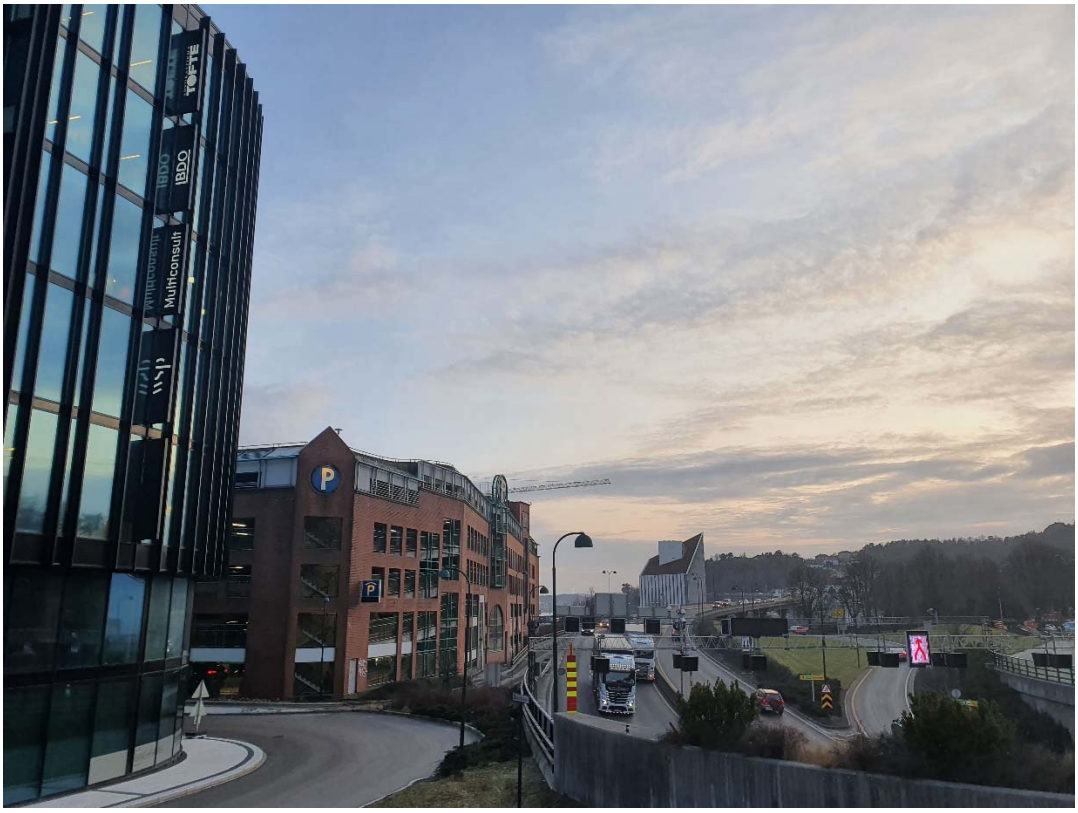
Figure 1c Outdoor calculations with Cadna for Banehaven 4A and Fjellgata 6 (Baneheia Park)

Measurements at Banehaven 4A probably show higher level than calculated due to complicated situation for modelling with the tunnel.

The free field outdoor noise level for Baneheia Park with todays traffic is (because todays and future situation are similar) L_d 74-75 dB as assumed above. The office building and road are shown in picture 1.

Picture 1: Baneheia Park, office building with whole glass facade

According to guideline 421.425 [3] from SINTEF in Norway, the sound field is assumed to be diffuse from movable noise sources like road traffic. In this case there is heavy traffic, and the noise on the facade at Baneheia Park comes from both tunnel and the bridge. The guideline from SINTEF states that the angle between the direction incidence of the noise source and the flat normal vector of the facade must be at least 63° or more to affect the sound insulation of the facade in a negative way. From the whole situation and with movable source this is not assumed to be relevant for Baneheia Park. A calculation based on distance to the road and height of 3rd floor indicates as "worst case" that the angle of incidence between traffic by the tunnel and the flat normal vector of the building is approximately 50° (picture 2).



Picture 2: Baneheia Park, traffic situation and diffuse noise field

2.2 Requirement to facade and indoor level

Other consulting company did calculations in early phase, where Multiconsult was responsible company for acoustics in the building process. There were two different views on the facade:

1. The other consulting company set requirement $R_w + C_{tr}$ 46-47 dB from laboratory with 4 dB reduction due to profile system and no other correction
2. Multiconsult set requirement $R_w + C_{tr}$ minimum 51 dB from laboratory due to 4 dB reduction for profile system and 4 dB further reduction due to area/dimensions of the facade, i.e. a total correction between field and laboratory value of minimum 8 dB.

The other consulting company distinguished between areas and set $R_w + C_{tr}$ of 47 dB as requirement for facade to cell office and $R_w + C_{tr}$ of 46 dB as requirement for facade to office landscape.

The indoor noise limit for offices in Norway is L_d 35 dBA. Measurements (continuous registering of noise level over approximately one hour) show indoor noise level of 39 dBA. This was similar to what Multiconsult predicted beforehand. The noise level was measured on days with dry asphalt February 2022:

- 9th of February 16^{44} - 18^{59} L_d 39,4 dBA
- 10th of February 06^{19} - 07^{19} L_d 39,3 dBA

The graph for indoor measurement in the afternoon (with no employees in the office landscape) on the 9th of February is given in figure 2.

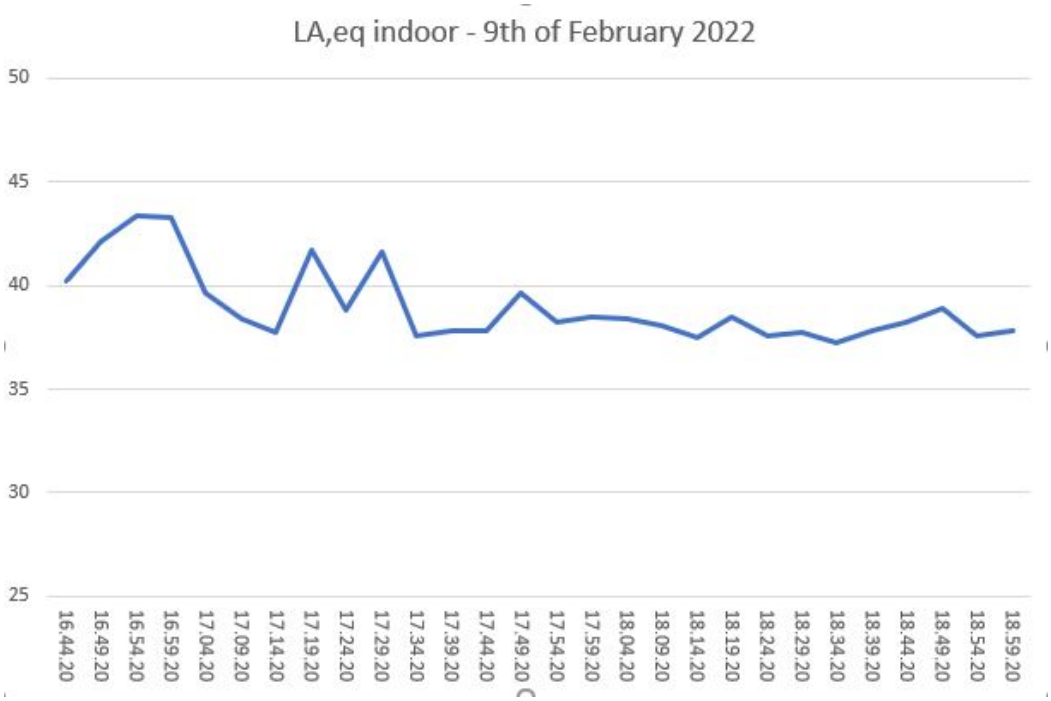
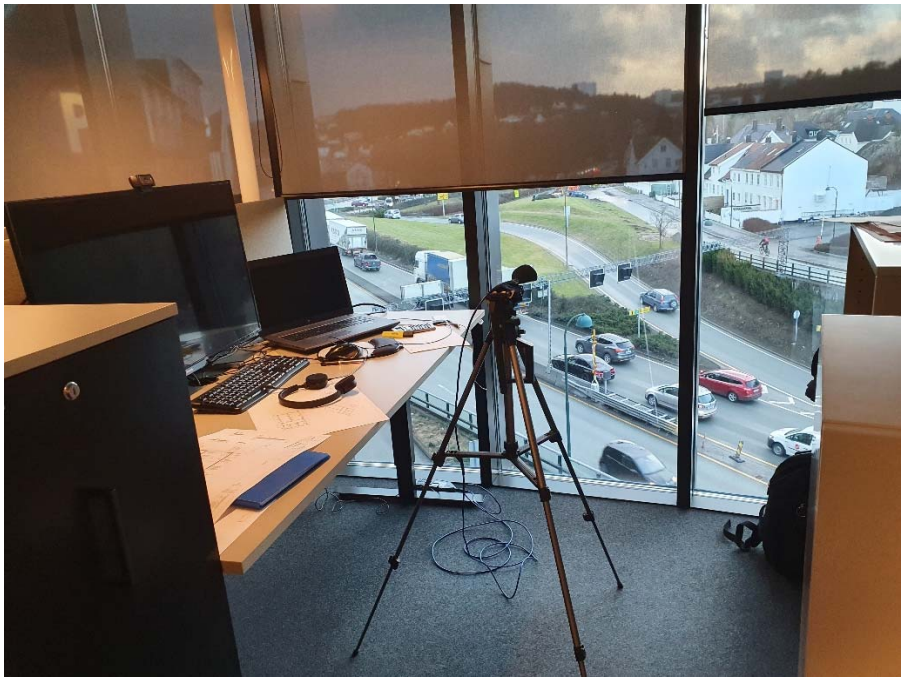


Figure 2 Measured indoor noise level in office landscape 9th of February

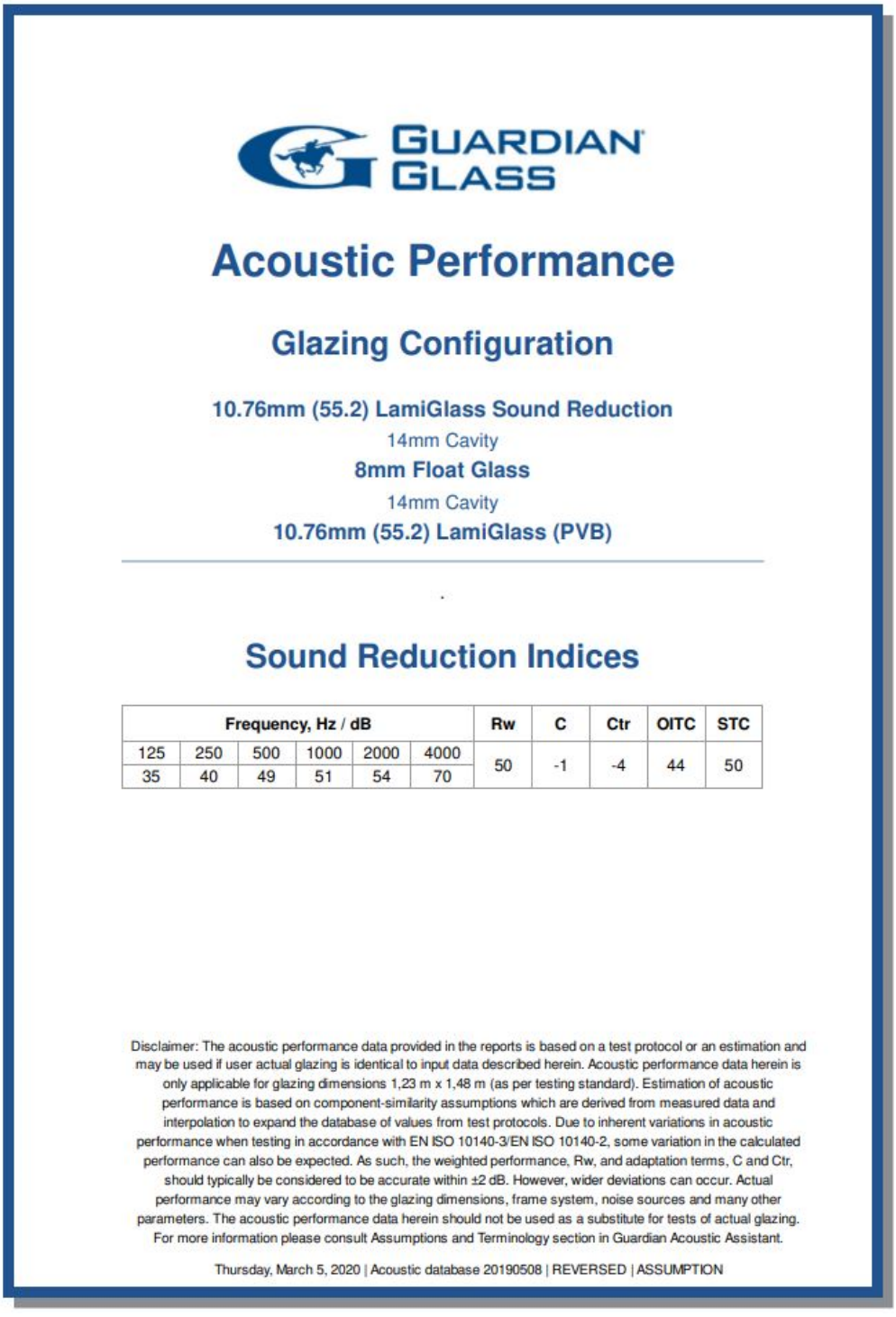
Earlier measurements outside (in 2019 as mentioned above) have shown constant noise level outside the facade from 06 to 19, except in rush hours. The indoor measurements morning and afternoon (as shown in picture 3) therefore represents the true value of L_d.



Picture 3: Baneheia Park, noise measurements inside in office

3 Laboratory tests of facade

Datasheet for chosen facade is shown below in figure 3.



GUARDIAN GLASS

Acoustic Performance

Glazing Configuration

10.76mm (55.2) LamiGlass Sound Reduction
14mm Cavity
8mm Float Glass
14mm Cavity
10.76mm (55.2) LamiGlass (PVB)

Sound Reduction Indices

Frequency, Hz / dB						Rw	C	Ctr	OITC	STC
125	250	500	1000	2000	4000	50	-1	-4	44	50
35	40	49	51	54	70					

Disclaimer: The acoustic performance data provided in the reports is based on a test protocol or an estimation and may be used if user actual glazing is identical to input data described herein. Acoustic performance data herein is only applicable for glazing dimensions 1,23 m x 1,48 m (as per testing standard). Estimation of acoustic performance is based on component-similarity assumptions which are derived from measured data and interpolation to expand the database of values from test protocols. Due to inherent variations in acoustic performance when testing in accordance with EN ISO 10140-3/EN ISO 10140-2, some variation in the calculated performance can also be expected. As such, the weighted performance, Rw, and adaptation terms, C and Ctr, should typically be considered to be accurate within ±2 dB. However, wider deviations can occur. Actual performance may vary according to the glazing dimensions, frame system, noise sources and many other parameters. The acoustic performance data herein should not be used as a substitute for tests of actual glazing. For more information please consult Assumptions and Terminology section in Guardian Acoustic Assistant.

Thursday, March 5, 2020 | Acoustic database 20190508 | REVERSED | ASSUMPTION

Figure 3 Datasheet for the glass facade at Baneheia Park

Laboratory tests are all done according to the standard ISO 10140-2 [2]. In Norway the test area for windows in laboratory is a height of 1,5m and a width of 1,25 m. Guideline sheet 533.109 [4] from SINTEF points at a negative area correction for windows with bigger area, and this guideline also mentions that more special dimensions (than standard) should be measured to ensure the real sound reduction of the facade.

The chosen facade is from Poland with a Ponzio system, and the laboratory test was done in Poland. The datasheet (figure 3) shows the following construction:

- 10.76 mm Lami Glass Sound Reduction
- 14 mm Cavity
- 8 mm Float Glass
- 14 mm Cavity
- 10.76 mm LamiGlass

The test from Poland was done according to ISO 10140-2 [2]. The datasheet (figure 3) states that the noise reduction of the facade is $R_w + C_{tr}$ 46 dB. It is also said that the acoustic performance is only applicable for glazing dimensions 1,23 m x 1,48 m. There is a general comment that actual performance may vary according to the glazing dimensions, frame system, noise sources and many other parameters.

From the information above, different countries seem to use the same test size of windows in laboratory. There is a general comment on different performance due to other dimensions, but the real effect of this is not clearly stated. There is today no standard method of applying test results to constructions with changed parameters such as area or size.

Investigations have shown that laboratory tests are done with a profile system on all edges in laboratory. This is done in similar ways in both Norway and Poland. The Norwegian laboratory has confirmed that tests in Norway are done in the way described in Poland, i.e.:

- Glass mounted in test opening and held on both sides by glazing beads (25 mm x 25 mm)
- Glass edge sealed on both sides with plastic sealant

Due to the given information about laboratory test (same dimension and same mounting), there is strong support for the “weakening” due to area/size of the glass in the case of Baneheia Park. Because the laboratory tests are confirmed to be similar in Norway and Poland, the experience of a facade with $R_w + C_{tr}$ of 46 dB in laboratory being 8 dB weaker in field has to be explained by the combined effect of profile system and big area/dimension of the glass. Laboratory tests cannot include flanking transmission as this is a characteristic of a given building.

4 Used profile system

The company H-glass has told that the profile system is designed by Ponzio, and is a high isolated profile system as shown in figure 4.

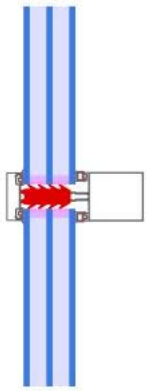


Figure 4 Facade profile used at Baneheia Park

The covering with pressure strip is from aluminium, and there is insulation of polyethylene inside of the rebate. According to this information, there is no strong weakening of the glass facade due to the chosen profile system, but comparable with isolated profiles. The weakening of the high isolated Ponzio profiles seems to correspond with the experience from Euronoise 2018 [1], where isolated profiles with mineral wool gave 4 dB reduction of noise for the combined system of glass with profiles, when the glass itself has laboratory value of $R_w + C_{tr}$ 50 dB.

5 Dimension of glass facade and guidelines

Both the Norwegian guidelines and the datasheet from tests in Poland states that other dimensions than used in laboratory may give other acoustic performance of the window. The phenomena is explained by other lateral resonant frequencies for a window of bigger size, which is mentioned and explained in theoretical books of acoustics. In a bigger window there will be other resonant frequencies than in a smaller window, and the modal density will also be highest in a bigger window.

The actual facade of Baneheia Park is divided in elements of following size:

- Width 1,25 m and height 3,15 m
- Width 0,80 m and height 3,15 m
- Width of 1,55 m and height 3,15 m

In this case, the height is much longer than “standard dimension” and seems to be the reason for other and more resonant frequencies than for standard window size.

The article from Euronoise in 2018 [1] only focuses on the effect of the profile system without considering area/size of the glass facade. For practical use, both the effect of profile system and the effect of area/dimension need to be considered when deciding what construction to use in a given situation.

6 Conclusions

A whole glass facade needs a correction for both weakening due to profile system and weakening due to area/dimension of the glass. All laboratory tests are done for "standard area". The suggested total correction for a facade with high isolated profile system is 8-10 dB due to both effects mentioned. Further on more research should be done to clarify the "lower noise reduction" related to area/dimensions of glass facade.

Acknowledgements

May I offer my deepest thanks to the company H-glass who gave me information about the chosen solution, and Christianholm Eiendom who allowed me to present this experience for their office building.

References

[1] Hans-Walter Bielefeld and Tejav DeGanyar. Acoustical Performance of Aluminium Framed Facade Systems. *Euronoise 2018*

[2] ISO 10140-2:2021 Acoustics – Laboratory measurement of sound insulation of building elements – Part 2: Measurements of airborne sound insulation

[3] SINTEF Byggdetaljblad 421.425 Isolering mot utendørs støy. Beregningsmetode.

[4] SINTEF Byggdetaljblad 533.109 Lydisolerende egenskaper for vinduer.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

The sound insulation of façades at infrasound frequencies

Valtteri Hongisto[1*], Jukka Keränen, Jarkko Hakala

Turku University of Applied Sciences, Acoustics Laboratory, Turku, Finland.

*valtteri.hongisto@turkuamk.fi

Abstract
Our purpose was to provide experimental information on the indoor-outdoor level difference, DL, of typical façade constructions in Finland and to present feasible estimation of DL that can be used to assess the indoor SPL of environmental noise in any building. The DL of 26 façade constructions were measured for 13 residential houses within 4–200 Hz. Statistical methods were used to derive a feasible estimate of DL that is exceeded in 84% of cases. The estimate can be used to assess the sound pressure level of environmental noise indoors in cases when the level difference of the façade cannot be measured but the outdoor sound pressure level of environmental noise is available.

Keywords: low-frequency noise, sound insulation, façades

1 Introduction

In Finland, the sound pressure level (SPL) of low-frequency environmental noise indoors shall not exceed the values of Table 1 [1]. The SPL caused by environmental noise inside a dwelling at low frequencies is calculated by subtracting the outdoor-indoor level difference (*DL*) of the façade from the outdoor SPL. During area planning stage, it is usually not possible to measure the façade *DL* of every dwelling of the inspected area since the number of inspected dwellings can be very large. Therefore, politically accepted estimations of *DL* need to be used. They are most probably exceeded in most façade constructions. There is very little data available of the *DL* of façade constructions during the last decades as reviewed by Keränen et al. [2]. Our purpose was to provide experimental information on the *DL* of typical façade constructions in Finland and to present feasible estimation of *DL* that can be used to assess the indoor SPL of environmental noise in any building. This paper is a summary of Ref. [2].

Table 1: The action values of equivalent unweighted SPL, $L_{pZ,eq,1h}$, in rooms used for sleeping for night-time hours (22–07) and daytime hours (07–22) given by the Finnish Ministry of Social Affairs and Health [1].

f [Hz]	20	25	31.5	40	50	63	80	100	125	160	200
Night [dB]	74	64	56	49	44	42	40	38	36	34	32
Day [dB]	79	69	61	54	49	47	45	43	41	39	37

2 Methods

We chose 13 single-family houses built during different decades (1921–2016) and built using different materials (timber, brick, hybrid) under inspection. If the house had different constructions on different façade sides, two or three façades of the same house were measured. Altogether 26 façades were measured.

Test sound was produced using a custom-made infrasound subwoofer (5–20 Hz), a standard subwoofer (Genelec 7050B, 20–80 Hz), and an omnidirectional sound source (NOR276). The three sound sources were located 5 meters from the measured façade. Since the loudspeakers were located on the ground, we assume

that the vertical façade was the dominant sound transmission path. Measurements were conducted under low wind speed conditions to avoid wind-generated pseudo-noise due to airflow turbulence in the microphone. Despite of low wind speed, we used windscreen around the microphone.

ISO 16283-3 describes a method to determine the sound reduction index of façade within 50–5000 Hz. We focused on low frequencies and used a non-standard method. The level difference to a corner position i of the room, DL_{Ci} [dB], was determined by

$$DL_{Ci} = L_{p1} - L_{p2,Ci} - 6 \quad (1)$$

where $L_{p,1}$ [dB] is the energy average of outdoor SPL in the vicinity of the façade surface, and $L_{p2,Ci}$ [dB] is the SPL in the corner position i of the receiving room. We used four positions in the room corners, C1−C4, 0.3–0.4 m from the corner. The level difference to a middle position j of the room, DL_{Mj}, was determined by

$$DL_{Mj} = L_{p1} - L_{p2,Mj} - 6 \quad (2)$$

where $L_{p2,Mj}$ [dB] is the SPL in the middle position j of the receiving room. We used five positions in the middle of the room: M1−M5. The average level difference of four corner positions, DL_C, was determined by

$$DL_C = -10 \cdot \log_{10}\left[\frac{1}{4}\sum_{i=1}^{4} 10^{-DL_{Ci}/10}\right] \quad (3)$$

Correspondingly, the average level difference of five middle positions, DL_M, was determined by

$$DL_M = -10 \cdot \log_{10}\left[\frac{1}{4}\sum_{j=1}^{5} 10^{-DL_{Mj}/10}\right] \quad (4)$$

We determined the 84th percentile level from the 26 measured level differences $DL_{C\sigma}$ and $DL_{M\sigma}$ to represent sophisticated lower limit of level difference among the sample as suggested by Hoffmeyer and Søndergaard [3]. Because the level difference values are planned to be used in assessing the SPL of environmental noise indoors, and the occupant can spend time in any position in the room, we determined first the data-based value, called $DL_{\sigma,data}$, by giving a weight of 1/3 to $DL_{C\sigma}$ and a weight of 2/3 to $DL_{M\sigma}$, as Hoffmeyer and Søndergaard [3] did. Both outcomes had a specific frequency dependence. Because of political reasons, we derived the final DL_σ by applying a polynomic fit to $DL_{\sigma,data}$. The outcome, DL_σ, was then monotonically growing as a function of frequency. It was also rounded to the nearest integer.

3 Results

The DL values of the 26 façade constructions and the 84th percentile values are shown in Fig. 1. The main outcome of this work, DL_σ, as a function of frequency, is shown in Table 2.

4 Practical utilization

Because the values of DL_C and DL_M varied significantly between the studied 26 façade constructions, it is important to know the actual DL if the indoor SPL of environmental noise needs to be known precisely. However, the measurement of DL for every façade is not realistic during area planning. In 84% of cases, DL is larger than the DL_σ of Table 2. The indoor SPL can be assessed by

$$L_{p,in} = L_{p,out} - DL_\sigma \quad (5)$$

where $L_{p,out}$ [dB] is the SPL of environmental noise outdoors and DL_σ [dB] is obtained from Table 2. It is assumed that indoor SPL exceeds $L_{p,in}$ only in 16% of cases. $L_{p,out}$ is obtained either by measurement or simulation. It shall not contain any reflections from the studied building. In Finland, $L_{p,in}$ is compared to the action values of Table 1. If they are exceeded in a specific dwelling, it is justified to determine the façade level difference in that dwelling by measurements.

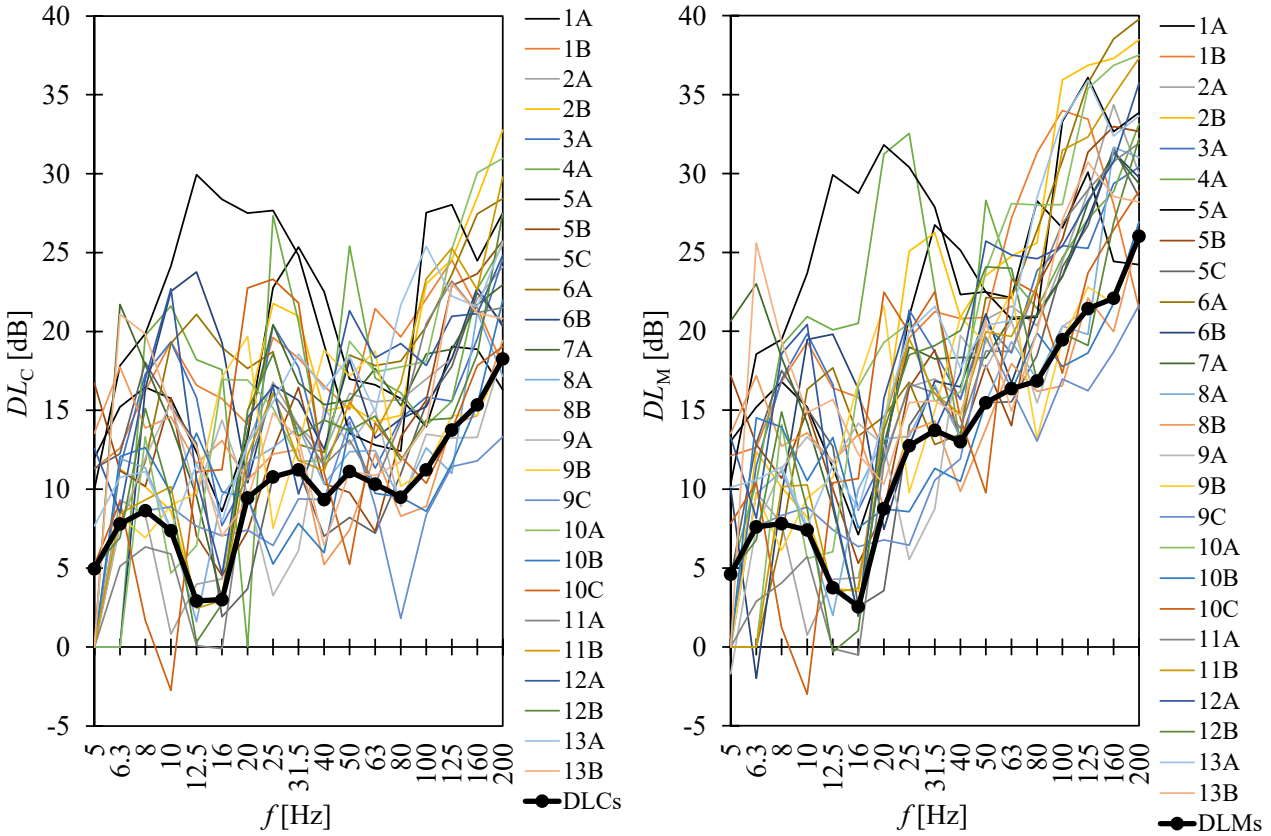


Figure 1: Left) Outdoor-indoor level difference measured in the corner positions, DL_C, as a function of frequency, f, for the 26 measured façade constructions. DLCs shows the 84th percentile of the 26 façades. Right) Outdoor-indoor level difference measured in the middle area positions, DL_M, as a function of frequency, f, for the 26 measured façade constructions. DLMs shows the 84th percentile of the 26 façades.

Table 2: Outdoor-indoor level difference, DL_σ, as a function of frequency, f.

f [Hz]	5	6.3	8	10	12.5	16	20	25	31.5	40	50	63	80	100	125	160	200
DL_σ [dB]	6	6	6	6	7	7	8	8	9	10	12	13	15	17	19	21	23

Acknowledgements

This study was part of Anojanssi project 2016–2019 funded by Business Finland (Tekes grant 828/31/2015), Turku University of Applied Sciences, and ten external parties (ministries, companies, and associations).

References

[1] Ministry of Social Affairs and Health (2015). Decree 545-2015 on Health-Related Conditions of Housing, Helsinki, Finland. http://www.finlex.fi/en/laki/kaannokset/2015/en20150545.pdf.

[2] Keränen, J., Hongisto, V., Hakala, J. (2019). The sound insulation of façades at frequencies 5–5000 Hz. Building and Environment 156 12–20.

[3] Hoffmeyer, D., Søndergaard, B. (2008). EFP-06 Project - Low Frequency Noise from Large Wind Turbines, Measurements of Sound Insulations of Façades, DELTA AV 1097/08, Apr 30, 2008, Denmark.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Numerical investigation on sound transmission behaviour of multi-layered panels with periodic arrays of spring-mass resonators

Milica Jovanoska[1,*], Todorka Samardzioska[1]

[1]Faculty of Civil Engineering, Ss. Cyril and Methodius University, Skopje, North Macedonia

*m.jovanoska@gf.ukim.edu.mk

Abstract
Recent developments in engineering and sustainability initiatives have resulted in building trend of lightweight partitions, utilizing new materials and technologies. At the same time, noise pollution is becoming a growing problem across the globe, as a result of present-day life dynamics. The impact of the increased noise levels on human health and wellbeing is perceived. All this implies the need for sound insulation enhancement of the lightweight partitions. Theoretical and numerical investigation of multi-layered panels under acoustic excitation are conducted using different methods. The sound transmission loss (STL) of unbounded multi-layered panels is calculated using the theory of three- dimensional elasticity and well-known transfer matrix method (TMM). Additionally, the finite element method (FEM) is also used for sound transmission loss calculation and for obtaining the dispersion diagrams. Based on the dispersion curves, detection of the band gaps is made possible. In order to improve the sound transmission loss in a specific frequency region, periodic resonant units (spring-mass resonators) are tuned and introduced to the considered panels. It can be shown that the resonance mode of the spring-mass resonators couples with the plate vibration in the way of breaking the mass law and overcoming some phenomena like coincidence effect and mass-air-mass resonance. These units can be attached or embedded in the host panel. Such periodic structures recently have been recognized as acoustic metamaterials. Acoustic resonant metamaterials are artificial periodic structures with unique acoustic wave manipulation properties, owing to the dynamic influence of their local resonant units. The drawback of this concept is that this ”unordinary” power for wave manipulation works only in specific narrow frequency band associated with the resonant frequency of the resonant units. Nevertheless, this phenomenon has potential to be useful ”tool” for correcting/complementing lightweight partition systems.

Keywords: acoustic metamaterials, sound transmission loss, mass-spring resonators

1 Introduction

The scientific path of acoustic metamaterials was driven by the findings in the quantum mechanical band theory of solids, followed by the concept of photonic crystals and later the concept of phononic/sonic crystals, due to the analogy between electromagnetic and elastic waves [1], [2]. Phononic/sonic crystals are periodic scatterers incorporated in matrix/medium that interfere with acoustic travelling waves, resulting in destructive interference when Bragg’s condition is met: $n\lambda=2asin\theta$, where λ is the wavelength, a is the periodicity constant and θ is the incident angle of incoming wave, [3], [4]. Unfortunately, the concept of phononic/sonic crystals becomes impractical to implement in the low frequency region, where a large distance between the scatterers is required. This limitation was overcome by introduction of the dynamic effects of periodic resonant units to a host, by Z. Liu et al., [5], and the acoustic resonant metamaterials started to establish as artificial periodic structures with unique acoustic wave manipulation properties, owing to the dynamic influence of their local resonant units. The drawback of this concept is that this interesting

power for wave manipulation works only in specific narrow frequency band associated with the resonant frequency of the resonant units, [6]. Nevertheless, this phenomenon can be useful "tool" for correcting/complementing the classical systems in the region of coincidence [7], [8], [9], mass-air-mass resonance [10], [11], ring frequency [12], low-mid frequencies [13], [14], [15], [16] etc.
The possibilities and limitations of traditional materials for noise control treatment are well explored. Recent advances in technology make room for new materials with interesting properties, and thus the opportunity for more complex but smarter solutions. The concept of metamaterials is now in the scientific focus. From the extensive research in the last decade, their use for sound insulation treatment is feasible, but still, there is a vast area for exploring ahead of the researchers.
This paper is concerned with theoretical investigation of the potential of using periodically attached resonant units, with subwavelength size of the unit cell, for sound insulation enhancement of lightweight partitions i.e., gypsum walls.

2 Modelling approach

2.1 Effective mass density method (EM)

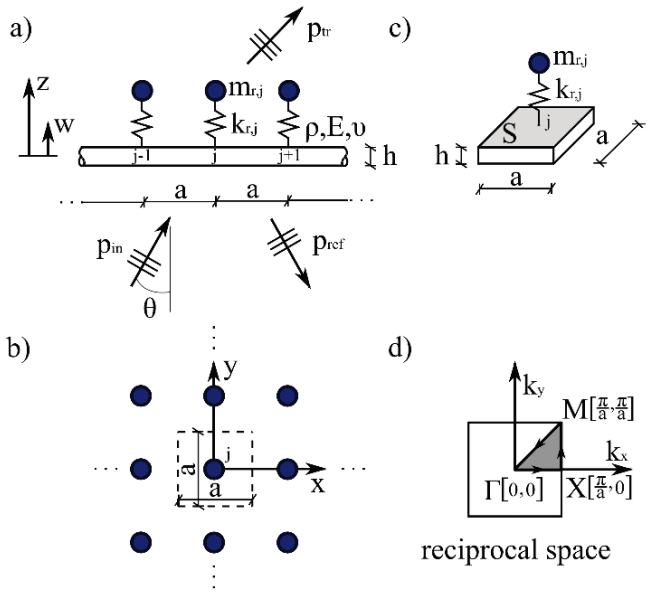


Figure 1: Metamaterial model: a) infinite thin plate with periodic resonators, b) periodic arrangement of the resonators, c) unit cell, d) reciprocal space, irreducible Brillouin zone Γ-X-M-Γ.

For elastic host plate with periodically attached single degree of freedom mass-spring units with periodicity constant much smaller than the wavelength of the motion in the plate, effective mass density representation is possible, [7]. This concept allows the use of equations derived for the host plate simply by replacing the mass density term with a frequency dependent effective mass density that incorporate the dynamic influence of the resonant units. The dynamic equivalent mass of the j-th spring-mass resonator, is given by the relation, [17]:

$$m_{eq,j} = \frac{m_{r,j}}{1-\frac{\omega^2}{\omega_{r,j}^2(1+i\eta_{r,j})}} \tag{1}$$

where $\omega_{r,j}$ is the resonance frequency of the j-th resonator, damping is introduced by complex spring constant $k_{r,j}(1+i\eta_{r,j})$, $\eta_{r,j}$ is the loss factor. By averaging the mass over the unit cell, the effective mass density is equal to:

$$\rho_{eff} = \rho + \frac{\rho_{r,j}}{1-\frac{\omega^2}{\omega_{r,j}^2(1+i\eta_{r,j})}} \tag{2}$$

2.2 Transfer matrix method (TMM)

For modelling the unbounded multi-layered systems, transfer matrix method is used. Based on the theory of elasticity and Biot's theory, 4x4 transfer matrix for elastic-solid layer and 6x6 transfer matrix for elastic-porous layers are given in [18], and for orthotropic solid layer is 6x6, [19],[20]. The transfer matrix for stationary fluid layer is 4x4. The derived transfer matrices relate acoustic pressure and particle velocity on both sides of a layer considering plane wave propagation. The fluids on both sides of the partition are semi-infinite. The continuity conditions between two adjacent layers are used to build the global transfer matrix. From this, the reflection and transmission coefficient can be easily calculated hence the sound transmission loss. For diffuse sound field, where the waves are incident from all direction with equal probability, the sound transmission coefficient can be obtained through integration:

$$STL_d = -10log_{10}\frac{\int_0^{2\pi}[\int_0^{\pi/2}\tau(\theta,\phi)sin\theta cos\theta d\theta]d\phi}{\int_0^{2\pi}[\int_0^{\pi/2}sin\theta cos\theta d\theta]d\phi} \tag{3}$$

2.3 Finite element method (FEM)

For comparison and expanded analysis, finite element method will be used for calculation of the sound transmission loss (STL) and dispersion diagrams. By using simple 1D model of infinite panel and 3D model of finite panel, sound transmission loss is obtained. Perfectly matched layers are employed for terminating the infinite domains, Figure 2. Diffuse field is modelled as pressure load (incident + reflected) calculated as a sum of N random plane waves (random angles of incidence and phases) multiply by 1/sqrt(N). Band diagrams i.e., dispersion diagrams are calculated using 3D FEM model of the unit cell by setting periodic boundary conditions based on Floquet-Bloch theory. In order to obtain the band structures, it is sufficiently to sweep the k-vector through the contour of the irreducible Brillouin zone Γ-X-M-Γ, Figure 1. The main characteristic of the acoustic resonant metamaterials is their ability of creating stop bands - that is frequency range where no free propagation of waves exists. These stop bands can be easily detected via dispersion diagrams. Periodic boundary conditions are employed for the 1D FEM model as well.

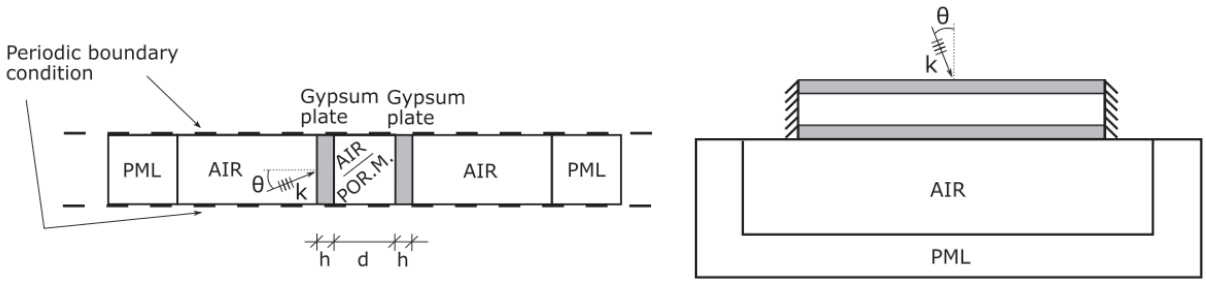


Figure 2: 1D FEM model (left) and plane section of 3D FEM model (right) for calculation of sound transmission loss.

2.4 Problem description

The potential of metamaterial concept with attached resonators for sound insulation will be discussed through gypsum walls. Single gypsum panel GP and double-leaf gypsum wall, without cavity absorption, GW1, and with cavity absorption, GW2, will be considered, Figure 3. Material properties of the walls are given in

Table 1. The panels are analyzed without and with periodically attached single degree of freedom mass-spring resonators, incorporated in the models via effective mass density EM. When considering the sound transmission loss of single and double panels, the main "weak spots" are the coincidence region and mass-air-mass resonance. The target of this research is to analysis the possibility of "correcting" these weaknesses.

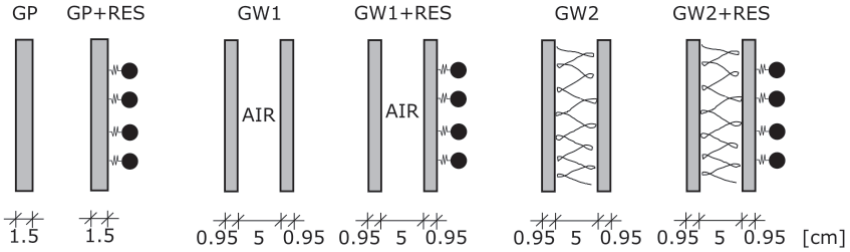


Figure 3: Configuration of considered panels.

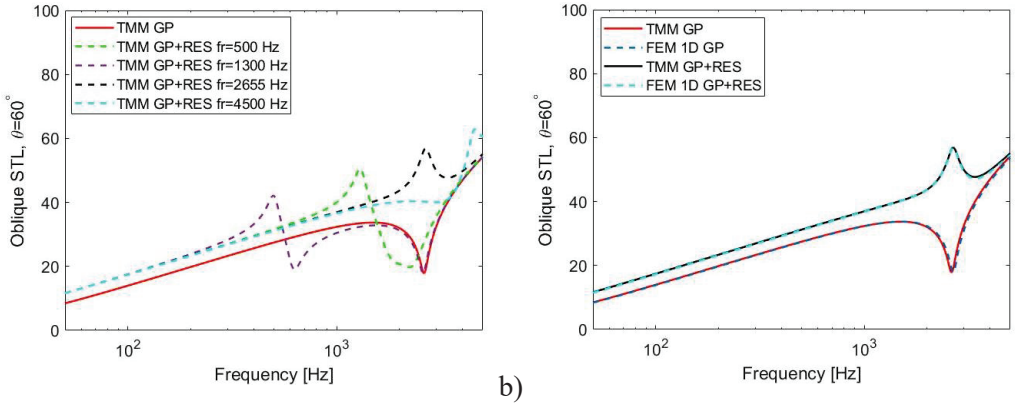
Table 1: Material properties

	GP	GW1		GW2		
		Gypsum panels	Air cavity	Gypsum panels	Porous material	
					Properties of the solid elastic frame	Air saturated pore - fluid properties
Density [kg/m3]	850	700	1.2	700	30	1.2
Modulus of elasticity [GPa]	4.1	2.25		2.25		
Poisson's ratio	0.3	0.3		0.3	0.3	
Flow resistivity [Ns/m^4]					6000	
Porosity					0.98	
Tortuosity					1.06	
Viscous dimension [μm]					150	
Thermal dimension [μm]					300	

3 Results and discussion

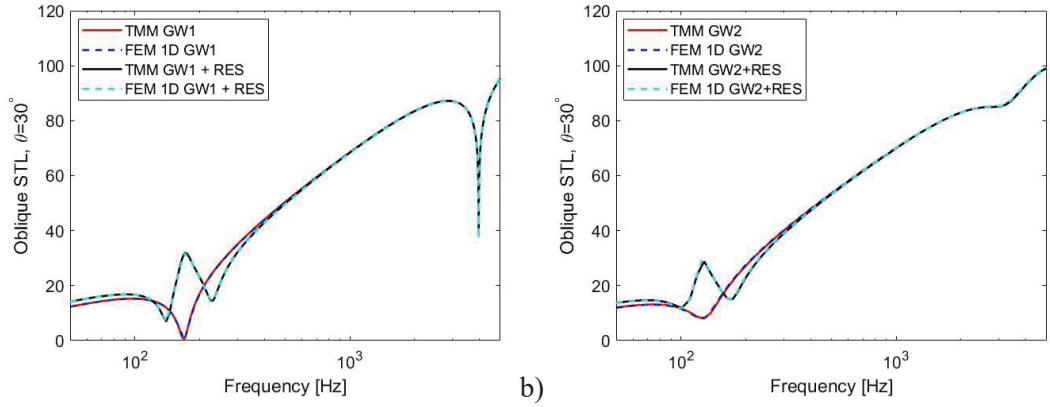
For GP, GW1 and GW2, single degree spring-mass resonators are selected to be attached periodically. The mass of the resonators m_r provides mass addition of $m_r/(\rho ha^2)=0.5$, i.e., *50%,* while the stiffness of the spring is tuned in order to obtain the desired resonant frequency. The loss factor for the resonators is set to be 10%. The panels are exposed to oblique and diffuse incidence.

The sound transmission loss for oblique incidence, $\theta=60°$, for bare GP and GP with attached resonators with different resonant frequencies is shown in Figure 4a. From the STL curve of the bare GP, the classical behaviour of a single panel can be observed. The mass law is "interrupted" by the coincidence dip, determined by the critical frequency, fc. When the resonators are introduced, picks/dips occur in the sound transmission loss. If the resonant frequency of the resonators is way below the coincidence dip, one sharp pick exists, then one sharp dip, followed by the coincidence dip. But if the resonant frequency is set to be in the coincidence dip, dominant role has the pick and depending on the position of the pick, the coincidence dip is more or less controlled. In Figure 4b, the comparison between TMM and 1D FEM infinite model is shown. The curves match perfectly. Here the resonators are tuned to the critical frequency of the host panel. For the frequencies before the resonant frequency of the resonators, there is improvement in the mass law region due to the additional weight from the resonators, but for the frequencies after the resonator effect, the curve approaches the curve of the bare panel.



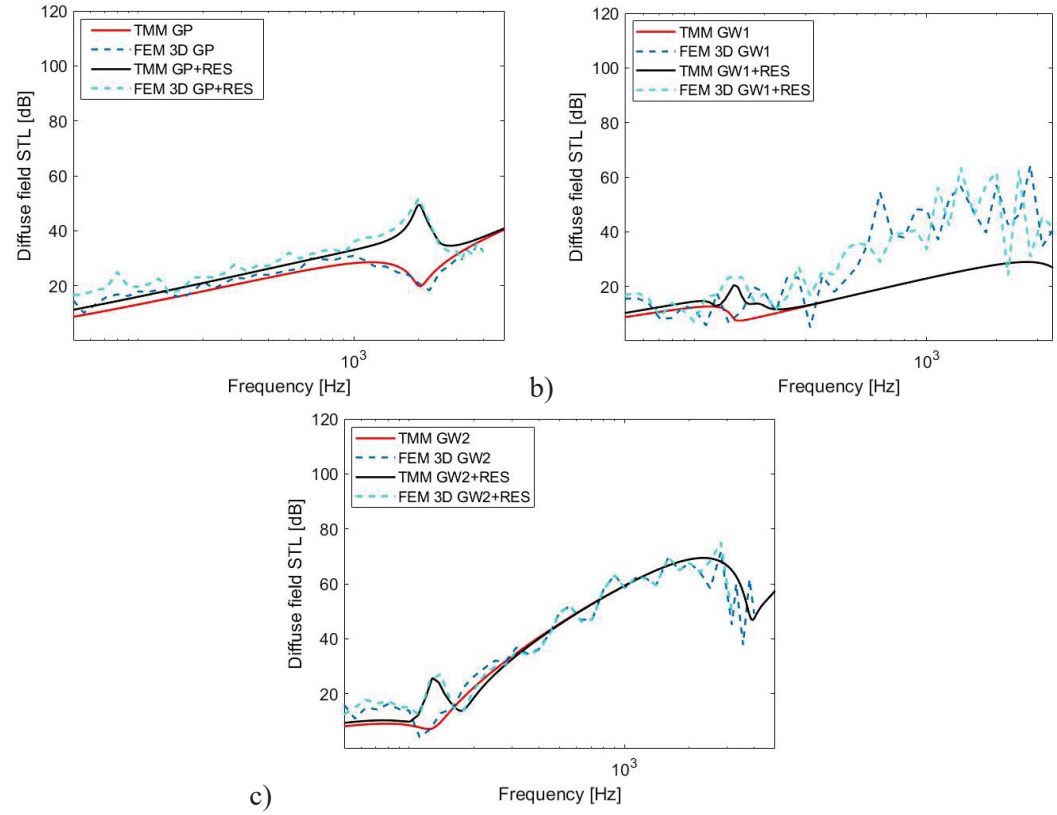
a) b)

Figure 4: STL for oblique incidence of GP without and with periodically attached resonators.



a) b)

Figure 5: STL for oblique incidence of GW1 and GW2 without and with periodically attached resonators.



a) b)

c)

Figure 6: STL for diffuse incidence of GP, GW1 & GW2 without and with periodically attached resonators.

In Figure 5, the STL curves for GW1 and GW2 for oblique incidence, $\theta=30°$, without and with periodically attached resonators are presented. Here the characteristic behaviour of double walls under acoustic excitation is evident, where besides the mass law region and coincidence dip, mass-air-mass resonance occurs. The resonators are tuned to the mass-air-mass resonance and the improvement is evident. For the solid-porous-solid GW2 wall, in the TMM model, very thin fluid layer is introduced between layers. The 1D model and TMM results are in agreement.
For diffuse field, the STL calculation based on the finite 3D FEM model shows good agreement with the infinite TMM model, for GP and GW2, Figure 6. There are some discrepancies in the low frequency region because of the finite dimensions of the panel in the 3D FEM model. For GW1 there is significant difference regarding the impact of the infiniteness of the air layer to the coupling between the plates in the TMM model.
Materialization of one resonator targeting the critical frequency of GP is proposed. This solution also applies to double wall with GP panels. The resonator is composed of steel mass and PLA 3D printed base. The resonators are attached to the plate in periodic schemes with periodicity constant $a=6$ *cm*. Mass addition is set to be 50% (of the host plate) and the stiffness is controlled through the PLA base. For the bare panel it is obvious that no band gap exists, but for the metamaterial design, band gap with width of nearly 250 Hz appears, between 1900-2150 Hz. These results shows that by tunning the dynamic behaviour, a favourable interaction between resonators and bending wave in host plate is possible, resulting in band gap creation, Figure 7.

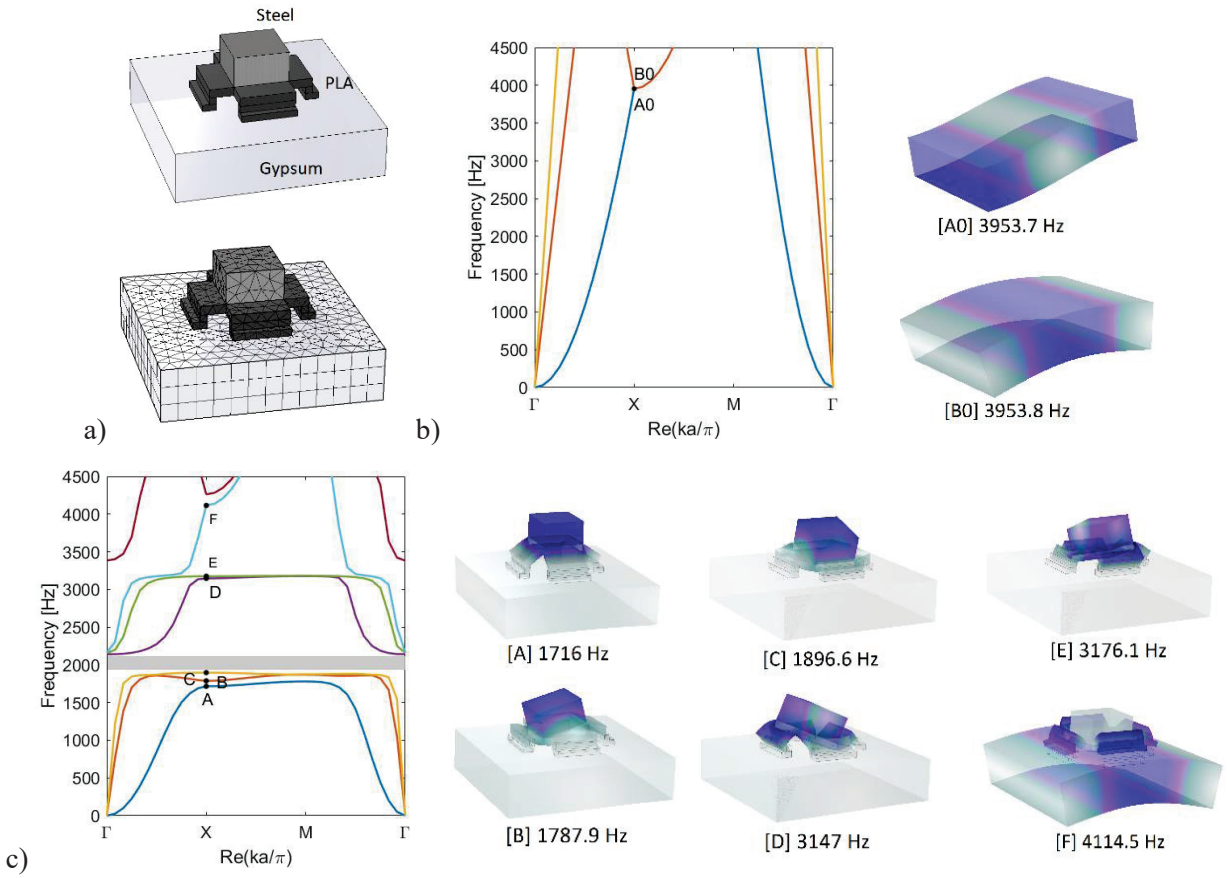


Figure 7: a) Materialized resonator targeting the critical frequency of GP, dispersion diagrams for b) bare plate and c) plate with resonators.

The proposed gypsum plate with resonators from Figure 7, is used to build a double wall, with cavity absorption, 15mm(50mm)15mm. The resonators are attached on both panels. Using the information provided in the dispersion diagrams as input for the TMM model with effective mass density, the STL behaviour is predicted, Figure 8.

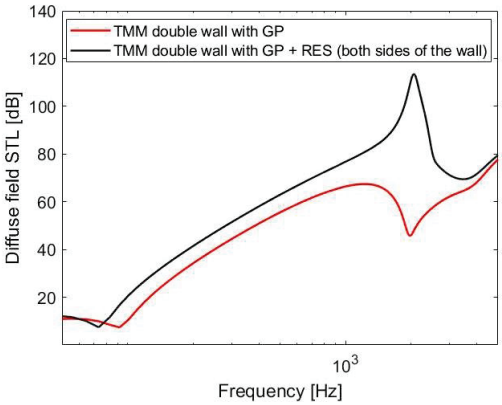


Figure 8: Double wall 15(50)15 with proposed resonators applied on both panels.

The main drawback of this concept is that the positive effect is in the narrow frequency region. One possible improvement is to introduce two or more resonators in one unit cell that have slightly different resonant frequencies, that can create two or more consecutive picks in the STL curve, Figure 9.

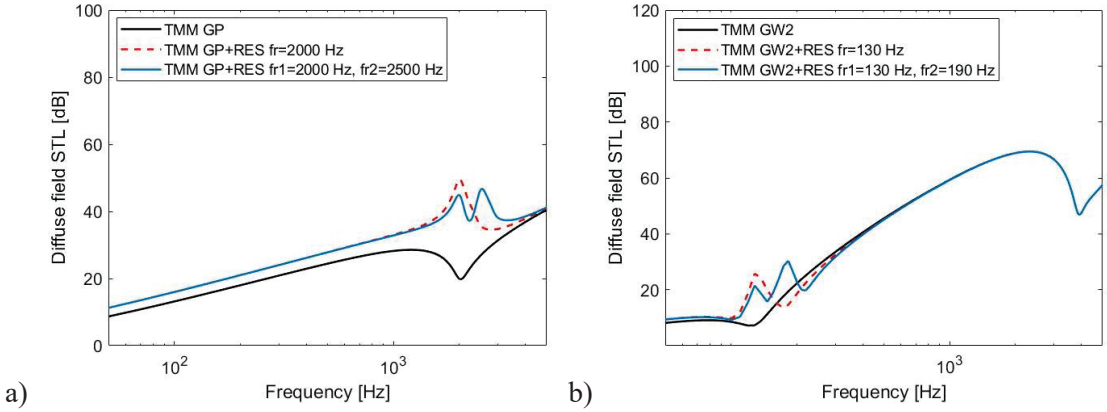


Figure 9: Two mass-spring resonators in one unite cell for GP and GW2.

4 Conclusions

With introduction of periodic mass-spring resonators to a lightweight partition, with proper design, improvements in the specific regions like coincidence region or mass-spring mass resonance can be obtained. Steel-PLA resonators can be one possible solution for materialization. By tunning the dynamic behaviour of the resonant unit, a favourable interaction between resonators and bending wave in host plate is possible resulting in band gap creation. Of course, there are many drawbacks of this concept, like, the effectiveness is only in narrow frequency range, also if the design of the resonant units is not precise, significant dips can arise, then, complicated design of the panels etc. But new technologies, like 3D printing, open the door for the metamaterials, also the intense research in this area directed towards optimization of the width of the band gaps, for example by introducing tailored different resonant units in one unit cell. In that way, metamaterial design can offer possible solution for noise treatment.

References

[1] G. Ma, P. Sheng: Acoustic metamaterials: From local resonances to broad horizons, Science Advances, 2(2), e1501595, 2016, doi: 10.1126/sciadv.1501595

[2] J. Liu, H. Guo, T. Wang: A Review of Acoustic Metamaterials and Phononic Crystals, Crystals, 10, 305, 2020, doi: 10.3390/cryst10040305

[3] C. Kittel: Introduction to Solid State Physics, John Wiley & Sons, Inc, 2005

[4] S. Kumar, H. P. Lee: The Present and Future Role of Acoustic Metamaterials for Architectural and Urban Noise Mitigations, Acoustics, 1(3), 590-607, 2019, doi.org/10.3390/acoustics1030035

[5] Z. Liu, X. Zhang, Y. Mao, Y. Y. Zhu, Z. Yang, C. T. Chan, P. Sheng: Locally Resonant Sonic Materials, Science, 289, 17341736, 2000, doi:10.1126/science.289.5485.1734

[6] B. Assouar, M. Oudich, X. Zhou: Acoustic metamaterials for sound mitigation, Comptes Rendus Physique, 17(5), 524-532, 2016, doi.org/10.1016/j.crhy.2016.02.002

[7] Y. Xiao, J. Wen, X. Wen: Sound transmission loss of metamaterial-based thin plates with multiple subwavelength arrays of attached resonators, Journal of Sound and Vibration, 331(25), 5408-5423, 2012, doi.org/10.1016/j.jsv.2012.07.016

[8] M. Oudich, X. Zhou, M. B. Assouar: General analytical approach for sound transmission loss analysis through a thick metamaterial plate, Journal of applied physics, 116, 193509, 2014

[9] M. Jovanoska, L. Godinho, P. Amado-Mendes, P. H. Mareze, M. Pereira, E. Ramis Claver: Overcoming the Coincidence Effect of a Single Panel by Introducing and Tuning Locally Resonant Structures, Internoise 2019, Madrid

[10] N. G. R. De Melo Filho, L. Van Belle, C. Claeys, E. Deckers, W. Desmet: Dynamic mass based sound transmission loss prediction of vibroacoustic metamaterial double panels applied to the mass-air-mass resonance, Journal of Sound and Vibration, 442, 2844, 2019, doi: 10.1016/j.jsv.2018.10.047

[11] N. G. R. De Melo Filho, C. Claeys, E. Deckers, W. Desmet: Metamaterial foam core sandwich panel designed to attenuate the mass-spring-mass resonance sound transmission loss dip, Mechanical Systems and Signal Processing, 139, 106624, 2020, doi: 10.1016/j.ymssp.2020.106624

[12] Z. Liu, R. Rumpler, L. Feng: Investigation of the sound transmission through a locally resonant metamaterial cylindrical shell in the ring frequency region, Journal of Applied Physics, 125, 2019

[13] Z. Yang, H. M. Dai, N. H. Chan, G. C. Ma, P. Sheng: Acoustic metamaterial panels for sound attenuation in the 50–1000 Hz regime, Applied Physics Letters, 96(4), 041906, 2010

[14] Y. Ye, X. Wang, T. Chen, Y. Chen: Step-by-step structural design methods for adjustable low-frequency sound insulation based on infinite plate-type acoustic metamaterial panel, Modern Physics Letters B, 2050220, 2020, doi: 10.1142/S0217984920502206

[15] H. Zhang, S. Chen, Z. Liu, Y. Song, Y. Xiao: Light-weight large-scale tunable metamaterial panel for low-frequency sound insulation, Applied Physics Express, 13, 067003, 2020

[16] A. Hall, G. Dodd, E. Calius: Diffuse field measurements of Locally resonant partitions, Acoustics: Sound, Science and Society, 2017, Perth

[17] J. P. Den Hartog: Mechanical Vibration, Dover publication, Inc, New York, 1985, ISBN 0-486-64785-4

[18] J. F. Allard, N. Atalla: Propagation of Sound in Porous Media, 2009, Wiley, ISBN: 978-0-470-746615-0

[19] Y. M. Kuo, H. J. Lin, C. N. Wang: Sound transmission across orthotropic laminates with a 3D model, Applied Acoustics, 69 (2008) 951–959, 2007, doi:10.1016/j.apacoust.2007.08.002

[20] C. Huang, S. Nutt: Sound transmission prediction by 3-D elasticity theory, Applied Acoustics, 70, 730-736, 2009, doi: 10.1016/j.apacoust.2008.09.003

[21] A. Dijckmans, G. Vermeir: Hybrid wave based - transfer matrix modeling of sound insulation problems, PROCEEDINGS OF ISMA 2014

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Laboratory test method to determine noise from wastewater installations

Simone Conta[1,*], Anders Homb[1]

[1]SINTEF Community, Trondheim, Norway.

*simone.conta@sintef.no

Abstract
We are working on a project which aims at improving sewage pipe material and piping system to reduce the radiated noise. One challenge we faced initially was to identify a suitable experimental method that fulfilled following requirements: allows for testing of the complete system and single components, fits our standard laboratory, suitable for measuring both airborne and structure-borne noise, reliable and cost effective. EN 14366 describes a laboratory measurement procedure to determine noise from wastewater installations. This method requires a setup comprising of two rooms: a source room to determine the airborne sound and a receiving room to indirectly determine the structure-borne sound by measuring the sound pressure level. As the test method was designed to compare pipe systems, the system layout is strictly defined, and the setup is not well suited to investigate system components or selected system sections for R&D purposes. The requirements on the laboratory facility are extensive and test of single components is not foreseen in the standard. In this paper we briefly present findings from a literature survey on the measurements of noise from wastewater installations. Moreover, we present an alternative setup and test method that we designed for our purposes. This setup allows the investigation of vertical and horizonal pipes or even just single components (e.g., bends). We will present sample results and are looking forward to an active discussion of pros and cons.

Keywords: wastewater pipes, sound radiation, measurement setup.

1 Introduction

The noise emissions from waste-water pipe installations can be a source of annoyance in buildings. The densification of the urban space, the limited space available for technical installations together with the use of lightweight constructions increase the risk of unwanted noise from the waste-water installations. Building users are increasingly aware towards noise levels [1]. Building constructors look constantly for options to reduce space requirements for technical installation and reduce installation costs without increasing the risk for complaints. This makes low noise piping systems attractive and the major wastewater system manufacturers responds accordingly developing low noise pipes and piping systems [e.g., Geberit, Wavin, Pipelife]. Increased development effort demands more effective and flexible test method than the established standards as, e.g., the EN 14366:2004 [2].

The vibrations induced by the water flow on the pipe walls produce directly airborne sound and structure-borne sound when they are transmitted to other building components by the pipe clamps. Field measurement results shows that structure-borne sound often dominates the sound pressure level in practical application. However, reducing the direct airborne sound might offer advantages in terms of lower requirements on the pipe enclosure or even allow for visible pipes in certain applications. Having to control both aspects makes measurement setups rather demanding. Moreover, different pipe system sections are exposed to different flow

and are excited in different ways. These generates the different characteristics sounds as e.g., the splash sound at the bend or at the branch and the flow noise in a vertical or horizontal section [3]. Also, the flow noise appears to be different in the vertical section where the water flow tends to lean on the walls in a helicoidal shape and in the horizontal section where the flow fills the lower half of the pipe.

We are currently involved in a research project leaded by a Norwegian pipe manufacturer aiming at developing even quieter pipes by gaining a better understanding of the noise generating mechanisms and sound radiation from the pipe and pipe components. A significant part of the project comprises the acoustic testing of material and components. For this purpose, a flexible yet reliable and accurate test method was needed.

In this paper, after a brief literature review, we present the setup we designed with the aim of overcoming the limitations of the current measurement standard and aim towards the project goals.

2 Literature

EN 14366:2004 [2] describes the laboratory measurement procedure to determine noise from waste-water installations and it is the reference standard for the topic. The standard test setup requires a rig comprising two rooms as shown in Figure 1. In this way it is possible to determine the directly radiated airborne sound and the structure-borne sound radiated in an adjacent room. The standard focuses on the comparison of pipe system with a very specific setup: it is limited to vertical pipes only, and the bend shall be mounted specifically below and outside the test room. The setup allows for different height differences between the water inlet and the bottom bend. The flow shall be continuous, ranging from 0,5 to 8 l/sec. The standard was often criticized because of the comprehensive requirements and because the data collected is not well suited for the available prediction models (see Amendment A1:2019). A review of the standard is ongoing to overcome the latter limitations.

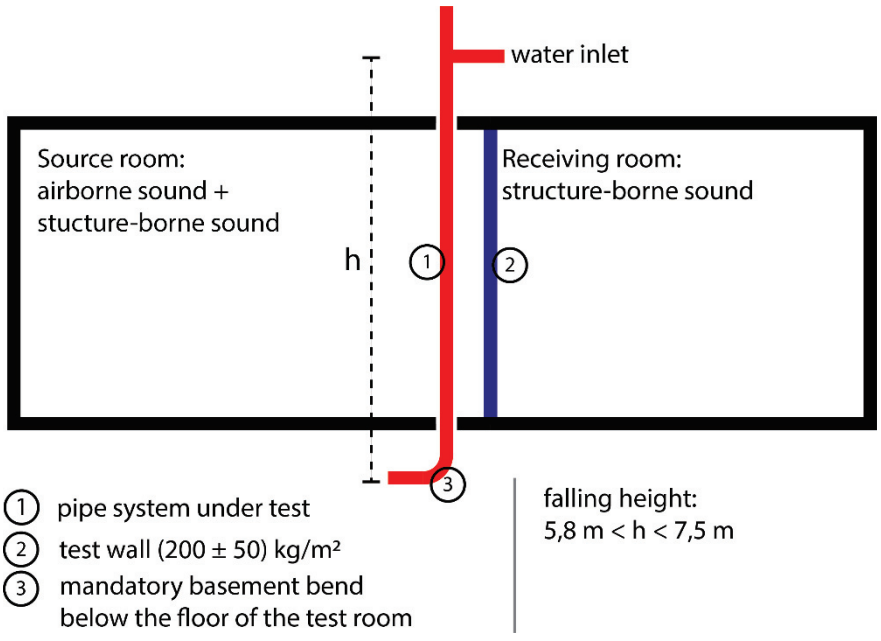


Figure 1 Measurement setup concept from EN14366.

The standard EN 15657:2017 [4] can be successfully used to determine the structure borne sound generated by a pipe system as described in [5] but the method does not cover the direct sound radiation.

A very interesting alternative measurement method is presented in [6]. The method combines the application of sound intensity measurements according to ISO 9614 [7] and structure-borne sound power according to EN 15657 to determine the airborne and the structure-borne sound directly on a power basis without the need of any specific room. The method appears to be very flexible and effective. The major drawbacks relate to the need of sound intensity measurement equipment and the corresponding requirements.

The paper [8] was a precious resource in this work: they installed a pipe system in a reverberation chamber and performed measurements on vertical and horizontal pipes, including also vertical offsets (i.e. a shift of the pipe axis achieved through two consecutive bends) as one of the tests. They tested products based on different materials and the effects of various types of pipe shielding (gypsum enclosure, external acoustic insulation, ceiling tiles). They measured with constant flow and with toilet discharge and conclude that a constant flow rate of 3 l/s gives sound pressure levels comparable with the level of a flushing toilet.

The work from Van Der Jagt [9] must be included in this brief overview since it offers many important insights on possible measurement setups. However, the analysis presented therein is about the applicability of the SEA method and focuses on mode count rather than on the sound radiation properties of the pipes and components.

Earlier works present measurement setups and theory that might be relevant for studying the material properties, e.g., [10] and [11]. Finally, it is worth mentioning the textbook [12] which covers the underlying theory of sound radiation from pipes.

3 Alternative measurement setup

We developed a measurement setup aimed at comparing different pipe materials and components for R&D purposes. Besides the obvious criteria that the setup needed to fit our available facilities, key requirements were a) allows the investigation of vertical and horizonal pipes or even just single components (e.g., bends), b) cost effective and c) easy to adapt to different research aspects.

Figure 2 shows a schematic of the test rig that we used. The measurement room has a volume of 200 m^3 and fulfil the requirements of ISO 10140 [13]. We have two tanks; an upper one that serve as source and a lower one that serve as sink. The flow regulation is made by a passive siphon to avoid pumping noise. In the current setup we used 1 l/s, 2 l/s, 3 l/s and 4 l/s. Other flow rates are possible but the measurement time available at higher rates than 4 l/s becomes relatively short due to the volume of the tanks. Up to now we used a constant flow rate. It would be possible to adapt the setup to measure with transient excitation as well.
Two alternative inlets are available. Both inlets are designed according to the specifications from EN 14366. Using the higher inlet, it is possible to obtain a falling height of 5,4 m. The minimum falling height that we can use is 0,5 s determined by the ceiling construction of the measurement room. A vibration decoupling element is used to prevent vibration transmission along the pipe outside of the measurement room. This allows for quick shifting of the pipe system and reduces the effect of the inlet and outlet fixtures. The pipe system can be installed hanging free in the room or installed by means of clamps on the two available reception plates (vertical and horizontal). In the latter case measurements can be carried out according to EN 15657.

We measured the sound pressure level in the room by means of two microphones mounted on rotating booms. In addition, we measured the reverberation time. In this way, it is possible to correct the sound pressure level to a standardized equivalent sound absorption area as described in EN 14366. Alternatively, it is possible to calculate the sound power which could be used as input data to prediction models.

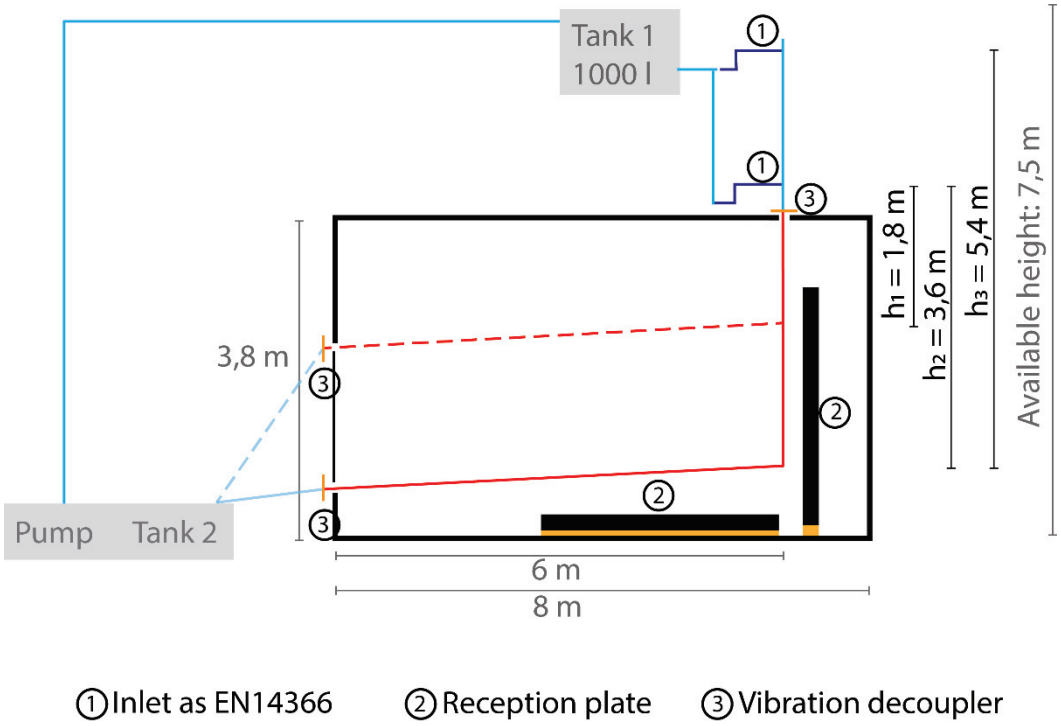


Figure 2 Proposed measurement setup

4 Some results

4.1 Reproducibility and mounting details

Figure 3 shows a picture of one of the configurations we measured. The pipe system was installed without brackets, resting on resilient material pads at two points on the horizontal section. We installed the very same system from scratch twice to get a rough estimate of the reproducibility of the setup. In addition, we performed a third measurement where at the bends we pulled apart the pipe by 10 mm. The aim was to verify how installation details might affect the overall sound radiation.
The results are shown in Figure 4. The values of the airborne sound pressure level $L_{a,A}$ are corrected with the equivalent absorption area according to EN14366.

We see that the variation between repetitions is below 1 dB at the higher flow rates while it is slightly above at lower flow rates. We observe a similar behaviour when we look at the test with bend and pipe pulled apart: very little difference at the higher flow rates (0,2 dB to 0,3 dB at 3 l/s and 4 l/s) and larger difference at low flow rates (1,4 to 1,7 dB at 1 l/s and 2 l/s). We have no good explanations for this effect, but we observe in other papers that measurements at low flow rates deviate slightly from general trends, as e.g. in [14]. We assume that this might be due to the water flow which is less stable at low flow rates.

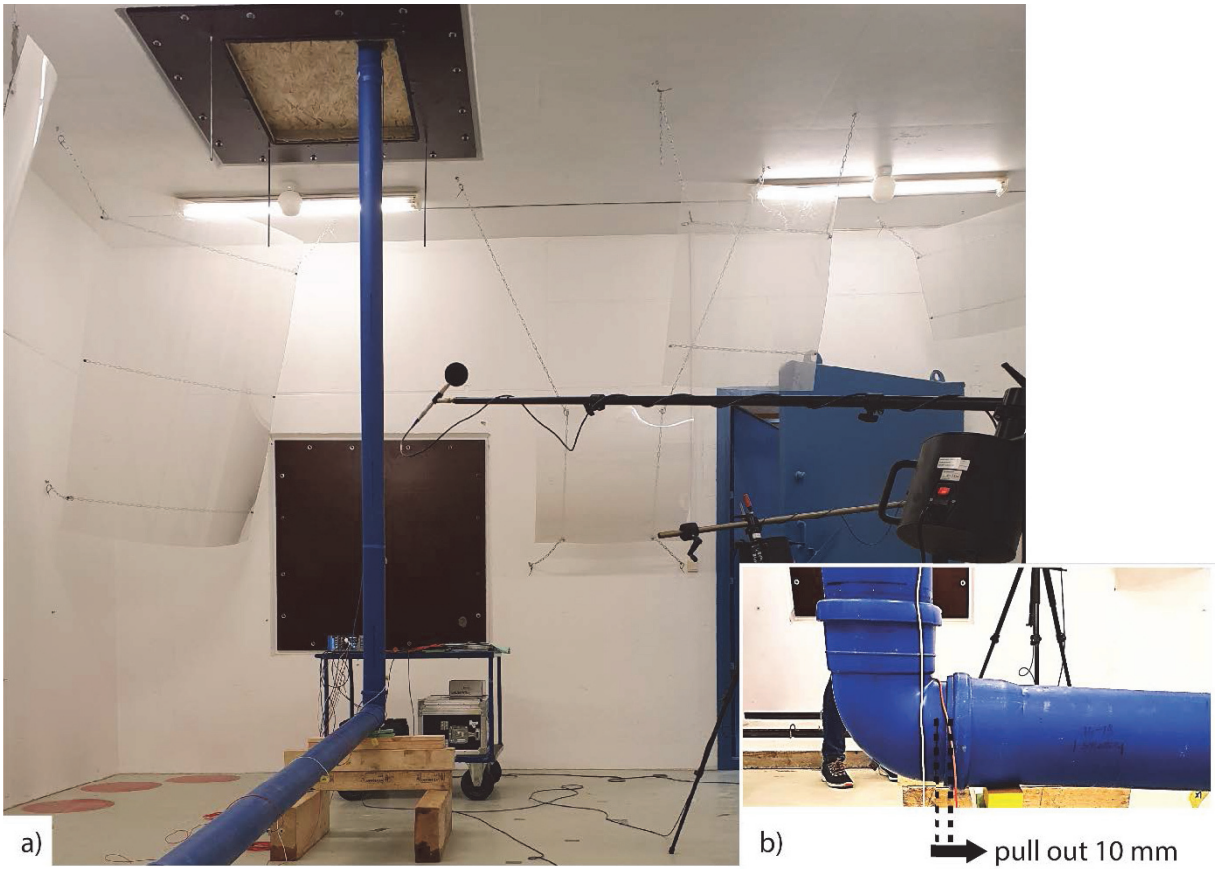
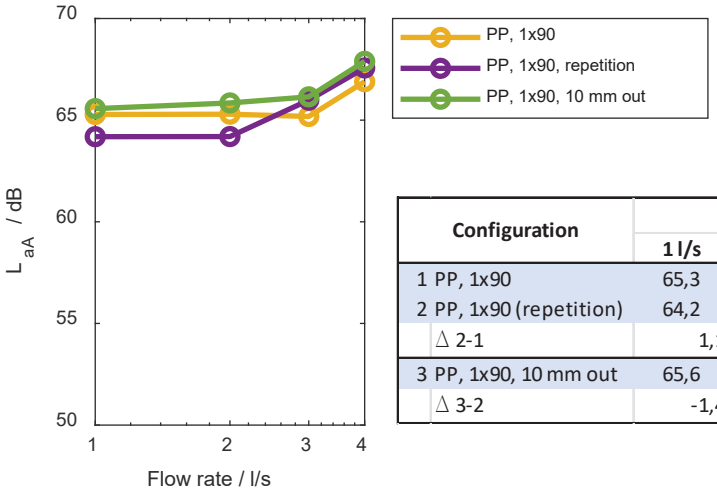


Figure 3 a) Configuration with pipe system installed without brackets. b) Test with pipe and bend pulled apart by 10 mm.



Configuration	$L_{a,A}$ / dB			
	1 l/s	2 l/s	3 l/s	4 l/s
1 PP, 1x90	65,3	65,3	65,2	66,9
2 PP, 1x90 (repetition)	64,2	64,2	66,0	67,6
Δ 2-1	1,1	1,1	-0,8	-0,7
3 PP, 1x90, 10 mm out	65,6	65,8	66,1	67,9
Δ 3-2	-1,4	-1,7	-0,2	-0,3

Figure 4 Measured airborne sound pressure levels at four different flow rates. Two repetitions and one measurement with bend and pipe pulled 10 mm apart.

4.2 Sound radiation from different components

Each section of the pipe system contributes to the radiated sound to different extents. One might experience the larger contribution to the sound pressure level from the bend. The radiated sound power from the vertical section is not the same as the one from the horizontal section. To investigate in detail these aspects, we shielded different sections of the pipe and measured the radiated sound pressure level. Figure 5 shows the measurement setup. The picture in a) was taken while installing the shield around the bend. The box made of OSB boards is still open: you recognize the mineral wool used as absorption (50 mm – 100 mm placed in contact with the wall box, without pressing on the pipe). The picture b) shows a configuration with the shields in place on the bend and 1 m before and after it. The openings around the pipe were sealed with mineral wool so to prevent direct mechanical contact between the pipe and the box and airborne transmission. In Figure 6, we show the measurement result obtained for 4 different configurations:

1) no shielding: we measure the sound pressure level from the whole system,
2) shielding of a section of the horizontal pipe: the sound pressure is now dominated by the sound radiation from the bend and the vertical pipe,
3) same as 2) with an additional shield on the bend: the sound pressure is determined by the sound radiation from the vertical pipe,
4) the shields prevent sound radiation from the bend itself and 1m before and after it: the measured sound pressure in the room is now dominated by the flow noise from the pipe sections away from the bend.

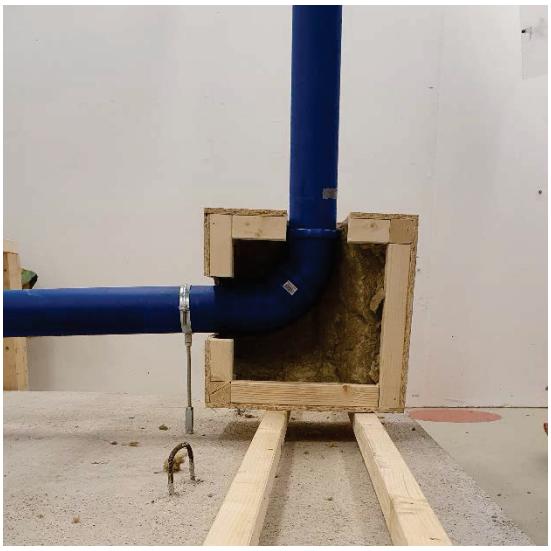
a)  b) 

Figure 5 a) Mounting the shielding around the bend. b) Installed shielding on the bend and 1m before and after it.

The results confirm that the bend and the pipe section close to it dominates the sound radiation as described, e.g., in [8]. The shielding appears to have little effect or even an amplification effect below 400 Hz. This might be due to following effects or a combination of them: a) the shield is not good enough at frequencies below 400 Hz, b) the vibration field in the pipe at frequencies below 400 Hz decreases at a lower rate with distance compared to higher frequencies and therefore sound radiation from the straight sections is more relevant at further distance from the bend and/or c) modal resonances of the shielding box. An improved version of the shields or longer shielding would clarify these aspects.

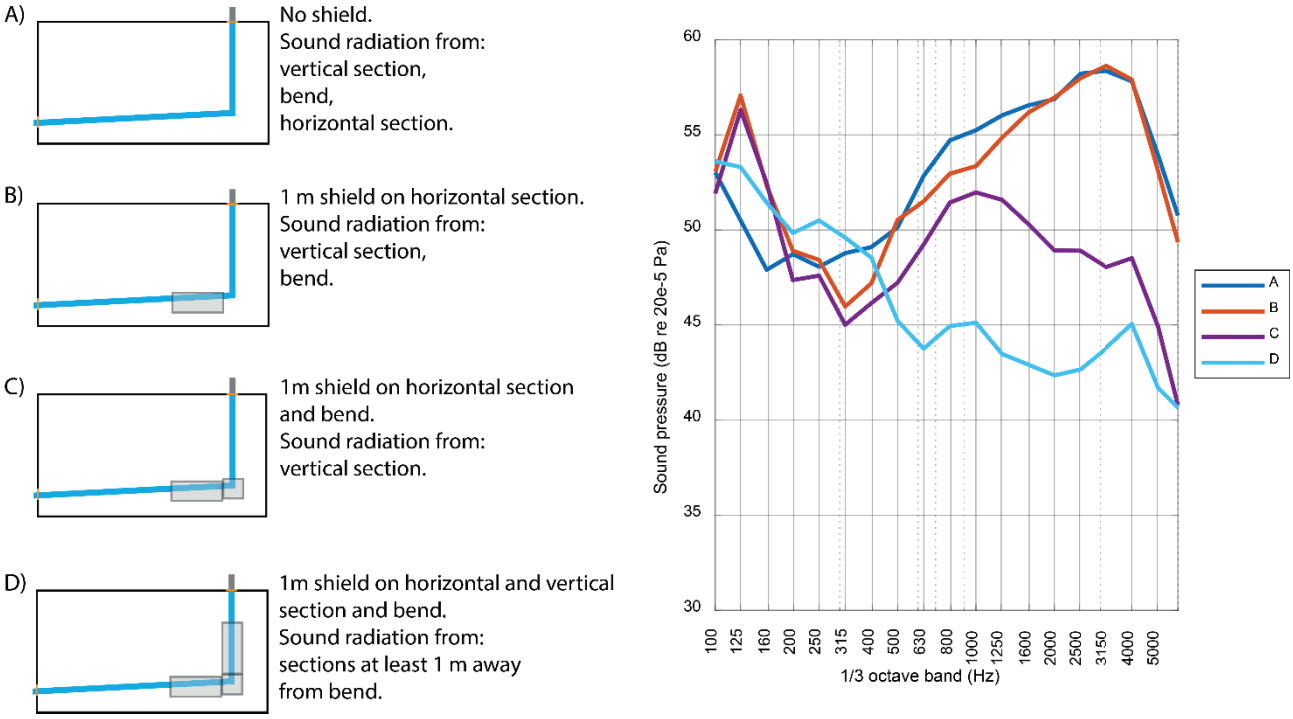


Figure 6 Measurement results obtained shielding different parts of the pipe system.

5 Conclusions and discussion

Increasing effort to reduce the sound radiated by drainage pipes requires a flexible and accurate measurement setup that allow for the investigation of both components and the full system. In this paper we presented the measurement setup we are using in a current research project. The setup was realised in our acoustic laboratory and overcomes some of the limitations of standard measurement setup realised according to EN14366.
We presented and discussed some selected results, focusing on the reproducibility of the measurements results (< 1 dB at higher flow rates) and showing how it possible to investigate sound radiation from different sections of the pipe system within one single installation.

Acknowledgements

This study has been carried out as a part of the IPN project Tyst, funded by The Research Council of Norway. The project owner is Pipelife AS. Research partners in the project are SINTEF Industry and SINTEF Community. Bjørn Ludvigsen at the SINTEF Acoustic laboratory in OSLO supported us developing and building the measurement setup.

References

1. Homb, A., *Noise from Waste Water Pipes in Apartment Houses*, in *Proceedings, The Sixteenth International Congress on Sound and Vibration*. 2009: Krakow, Poland.
2. NS-EN 14366:2004+A1:2019, *Laboratoriemåling av støy fra avløpsinstallasjoner*. 2020.
3. SINTEF Byggforsk, *553.182 Støy fra avløpsinstallasjoner*, in *Byggforskserien*. 2010, SINTEF: Oslo.
4. NS-EN 15657:2017, *Acoustic properties of building elements and of buildings Laboratory measurement of structure-borne sound from building service equipment for all installation conditions*. 2017.
5. Bailhache, S., et al. *Characterization of a water drainage pipe using the EN 14366 and EN 15657 methods*. in *Proc. Euronoise*. 2018.
6. Weinzierl, J., et al. *Approach for a new laboratory method to determine noise from waste water installations*. in *e-Forum Acusticum 2020*. 2020.
7. series, N.-E.I., *Acoustics - Determination of sound power levels of noise sources using sound intensity*. 2009.
8. van der Schee, W., *Measurement of the noise production in drainage pipes*, in *CIB W062 Symposium 2012*. 2012.
9. Bron-Van Der Jagt, G.S., *Sound Transmission through Pipe Systems and into Building Structures*. 2007, PhD thesis, Eindhoven, The Netherlands.
10. Manning, J.E. and G. Maidanik, *Radiation properties of cylindrical shells.* The Journal of the Acoustical Society of America, 1964. **36**(9): p. 1691-1698.
11. Holmer, C. and F. Heymann, *Transmission of sound through pipe walls in the presence of flow.* Journal of Sound and Vibration, 1980. **70**(2): p. 275-301.
12. Fahy, F. and P. Gardonio, *Sound and structural vibration*. 2007, Elsevier.
13. NS-EN ISO 10140 - serie, *Acoustics — Laboratory measurement of sound insulation of building elements*. 2021.
14. Guigou-Carter, C., et al. *Acoustic measurements on waste water installations and solutions to limit residents annoyance*. in *Euronoise 2018*. 2018. Crete: EAA – HELINA.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Modeling of multiple reflections between noise barriers and trains using the boundary element method

Christian H. Kasess[1,*], **Thomas Maly**[2], **Wolfgang Kreuzer**[1]

[1]Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria.

[2]Institute for Transportation, TU Wien, Vienna, Austria.

[*]christian.kasess@oeaw.ac.at

Abstract

While noise barriers are built to act as an acoustical barrier, they also cause an optical barrier effect, which could be reduced using transparent elements. Typically, such elements are acoustically highly reflecting. Unfortunately, in particular for railway traffic highly reflecting surfaces potentially reduce the insertion loss by multiple reflections between train and barrier. Current noise mapping methods such as the EU directive 2002/49/EC (Annex II) provide means to approximately consider such reflections. There is, however, the need for detailed investigations how the placement and dimensions of sound hard elements, the distance of the train to the barrier, the train cross-section, the distance of the observer, and other variables influence the insertion loss of the barrier. Measurements of passbys for 4 different barrier variants combining highly absorbing and reflecting materials were performed. These measurements provide the basis for the validation of the 2.5D boundary element calculations. In a next step, based on calculations of many different configurations a simplified, practical calculation model will be derived which may be used in noise mapping applications.

Keywords: noise barriers, railway traffic, reflecting elements, boundary element method, insertion loss.

1 Introduction

Noise barriers are an important tool for environmental noise control, in particular for traffic noise. However, in addition to the desired acoustical effect, noise barriers also are an optical barrier obstructing the view for passengers as well as for residents. Transparent elements could reduce this visual barrier effect. Typically, transparent elements for noise barriers are sound hard. The highly reflecting surfaces of common transparent barrier elements, however, pose a large problem in particular for railway traffic as they may produce considerable multiple reflections between train and barrier, potentially reducing the insertion loss of the barrier [1].

Current noise mapping methods such as the EU directive 2002/49/EC (Annex II) [2] provide means to approximately consider such multiple reflections including retrodiffraction assuming parallel planes and constant absorption over the entire height of the barrier. The main question being, whether the framework in the Annex II [2] is sufficient to deal with a variety of situations that occur in railway traffic where e.g. transparent elements are often placed only in the upper sections of the barrier. For example, in Morgan et al. [3] the effect of absorption and the train cross-section was investigated using the boundary element method (BEM) in 2D. They showed a large effect of absorption on the noise barrier which was larger for a box-shaped train than for a rounded vehicle body. This study, however, only considered barriers which were either absorbing or reflecting as a whole, although different barrier shapes were investigated. Recently, Bustos et al. [4] investigated the effect of mixing reflecting an absorbing parts on a noise barrier using a combined boundary element-finite element approach. There it was found, that the vertical placement, the barrier height as well as the distance of the

source to the barrier influence the insertion loss. However, for computational reasons both studies used pure 2D approaches which implies a fully coherent, infinitely long line source, which is computationally efficient but unrealistic. Furthermore, only overall changes are reported although it seems clear that frequency dependency effects play an important role. Finally, neither study was validated using measurements as for instance the study of Kirisits et al. [1] who investigated the influence of retrodiffraction on the calculation results, however, for highly absorbing noise barriers only. Thus, there is still the need for more detailed investigations how the placement and dimensions of sound hard elements, the distance of the train to the barrier, the train cross-section, the distance of the observer, and other variables influence the insertion loss of the barrier.

The work presented here aims at investigating the different factors systematically using the 2.5D BEM [5, 6]. The advantage of the 2.5D approach is, that under the assumption of a constant cross-section point sources as well as incoherent line sources can be modeled. To validate the computational approach extensive measurements campaign were carried out consisting of acoustic passby measurements of four different barrier variants combining highly absorbing and reflecting materials. Using the so determined source model calibration, simulations of many different configurations will be used to derive a simplified, practical calculation model which may be used in noise mapping applications.

2 Methods

2.1. Measurement setup

For the measurements a site was chosen, where two adjacent measurement cross-sections with and without a barrier were available and where the surroundings were essentially flat. The site was located close to Vienna along the eastern line. The noise barrier was a concrete noise barrier covered with a highly absorbing layer of wood-fibre concrete which was present on both sides of the barrier with a rippled diffuser profile oriented towards the train. The height of the barrier was 2 m above the top of the rail. The noise barrier's cross-section was placed 250 m away from the end of the barrier. A reference cross-section with unobstructed sound propagation was located 200 m away from the end of the barrier in the other direction.

The noise barrier cross-section was equipped with 9 microphones: 2 between train and noise-barrier (Pre-Sonus PRM-1, 1.2 m and 2.8 m above the top of the rail), 5 microphones at 7.5 m (G.R.A.S. 46AE, equally spaced heights from 1.2 m to 3.6 m above the top of the rail), and 2 microphones at 25 m (B&K UA-1404, heights 2.0 m and 4.5 m above the top of the rail). Two inductive wheel sensors (one per track) were used to determine the time of axle passbys as well as train speeds.

The reference cross-section was equipped with 2 microphones (G.R.A.S 46AE) located at the standard emission point (7.5 m, 1.2 m above the top of the rail) and in the upper 25 m position as in the noise barrier cross-section. In this cross-section train speeds and axle spacings were determined using two light barriers. The sampling rate of the acoustic measurements in the barrier cross-section was 48 kHz and in the reference section (due to different recording hardware) 51.2 kHz which were resampled to 48 kHz for further analysis. For manual analysis of train types and vehicle composition of freight trains all train passbys were recorded with a video camera. To ensure the comparability of the noise emissions between the cross-sections rail roughness and track decay rates of the rails of both tracks and cross-sections were measured.

2.2. Noise barrier modifications

In order to investigate the effect of highly reflective material, 40 plywood panels (27mm thick, 3-layers, spruce, covered with a thin layer of melamine resin) were used. The panels were placed in front of the rippled absorber using a custom made steel construction. There was a distance of around 6 cm between absorber and wood panel. Each panel was 2.5 m long and 1 m high. A total of 100 m of the noise barrier were covered, centered around the measurement cross-section. Three different panel placements were investigated: 1) panels were placed at the top of the barrier such that the upper meter was covered (Fig. 1); 2) panels were placed 0.5 m below the

barrier edge; 3) starting from state 2, every second panel was removed. For the top positioning (variant 1), the gap on top of the barrier due to the distance between barrier and panels was covered using small wooden panels to avoid absorption at the barrier edge which can be seen in Fig. 1.

Figure 1: Modified noise barrier. Shown is the upper panel configuration.

2.3. Boundary element method

Boundary element calculations were performed using the 2.5D BEM [5, 6]. The assumption is, that the scatterer is infinitely long and has a constant cross-section. However, in contrast to pure 2D BEM the source can be a point or an incoherent line source. The ballasted track was modeled using a 3-layer impedance model [7]. The highly absorbing noise barrier was defined from absorption measurement using real (zero phase) impedance values. The absorption values at high frequencies were around 0.9. Similarly, the wood panels were modeled using absorption coefficients for OSB boards [8]. A number of different source positions and train cross-sections were modeled (Fig. 2 shows the passenger train). Source positions were located at the top of the rail and at roughly the center of the wheel (points in Fig. 2). Omnidirectional as well as dipole sources were investigated. Point sources were placed at a spacing of 1.25 m along the track. The energy was integrated using a Simpson rule to obtain the result of an incoherent line source of about 1.5 times the length of the train. Due to the restrictions on the scatterer for the 2.5D BEM a number of important simplifications had to be applied. First, the bogies and wheels cannot be modeled properly, thus similar to previous work these lower parts of the wagons were not present in the model [9, 10]. Second, for the setting with alternating panels computationally involved methods utilizing the periodic structure based on the Floquet transform e.g. [11] were not considered. Instead, a somewhat diffuse field between barrier and train was assumed and the frequency dependent mean of the absorption of the wooden panels and the noise barrier was used to model the panel absorption for the alternating case. Third, gaps between wagons cannot be modeled and were thus ignored. This is mostly relevant for certain freight trains. Fourth, the steel structure on which the panels were mounted was not considered.

2.4. Data analysis

During regular operation, various train types were recorded including passenger trains, commuter trains, as well as a variety of different freight trains. For the latter the type of vehicle body contained cylindrical tanks, container or other wagons with vertical reflecting structures, car transporters, and other specialized types of wagons. Flat wagons (e.g. for containers or wood) not carrying any load were also observed.

The main focus here is on passenger and commuter trains due to their homogenous exterior shape, however two types of freight trains are also reported. For the passenger trains the procedure was to define a segment of the recording on the basis of the axles such that acoustic effects of the ends of the train are reduced. The passenger train (4 to 6 passbys per barrier configuration, speeds from 119 to 140 km/h) was about 200 m long

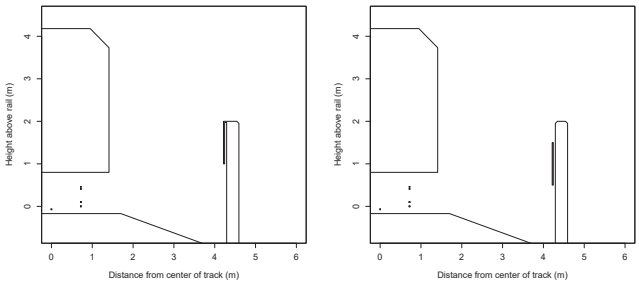


Figure 2: Cross-sections used for the simulation of the passenger train. The left panel shows the upper panel configuration, the right panel the middle configuration.

with a locomotive located at one end. At both ends 12 axles were ignored considering only the middle section. For the two types of observed commuter trains (6 to 21 passbys per configuration, length of 140 m in total, speeds from 130 to 140 km/h) 8 axles were ignored at each end. For freight trains, the video recordings were used to identify sections of freight trains with homogenous wagon types.

Using the Large Time-Frequency Analysis Toolbox (LTFAT, [12]) spectrograms were generated (function *dgtreal*, 4096 frequency channels, 1000 samples hop size). One-third octave band spectra for each point in time were generated by energetically summing up the respective frequency bins. Using the axle timing data the first and last axle of the middle segment were determined and an energetic average over the so defined section was calculated. For freight trains, due to their heterogeneity and varying analysis section lengths an axle-centered energetic averaging over 21 adjacent spectral estimates was performed for each axle. The axles considered had to be at least 30 m away from an axle of a different wagon type and 60 m away from the end of the train. At least 26 and up to 158 axles were determined this way per type and barrier configuration (speeds from 72 to 100 km/h). For the reference cross-section the time points for the spectra were spaced similarly, taking into account differences in speed.

The spectra of the barrier cross-section were divided by the respective spectra of the reference cross-section. For the nearer microphones up to 7.5 m the emission microphone at the reference cross-section was used. For 25 m the immission position was used. The arithmetic mean of the so determined attenuation terms was then calculated over passbys or axles in the case of freight trains.

3 Results

3.1. Measurements

Fig. 3 shows the differences between the absorbing and the respective reflecting condition for all nine microphone positions as a function of frequency in terms of the one-third octave band center frequencies. Positive values imply a decrease in barrier attenuation for mounted panels. It is obvious, that there is a clear effect of the reflecting panels which is highly dependent on the measurement position. There are also differences between the panel configurations of up to 2 dB for the same number of panels. Reducing the reflecting surface reduces the adverse effect on the attenuation considerably.

3.2. Model calculations

Fig. 4 shows the comparison between BE calculations and measurements for six of the nine microphone positions. Shown is the top reflecting barrier compared to the absorbing barrier. Different source positions are color coded. Clearly, all source positions yield results comparable to the measurements across all microphone positions with some degree of over- or underestimation. Fig. 5 shows the same for the middle reflecing panels.

Figure 3: Measured reduction in attenuation for passenger train. Different settings are color coded. Dashed lines show the standard error of the mean. Positive values indicate a reduction in attenuation when using reflecting panels compared to a fully absorbing barrier.

Here, the omnidirectional wheel source (red lines) leads to an overestimate of the effect of the reflecting panels whereas the rail source (green lines) still yields a good agreement with the measurements. Using the mean absorption of the panels and the barrier also leads to a similar agreement with the situation where only every other panel was mounted. Using a horizontal dipole source at the wheel position (blue lines) in general leads to a further overestimation. This is most likely due to the horizontal directivity which leads to a higher weighting of positions close to the measurement cross-section, where the effect of multiple reflections is stronger.

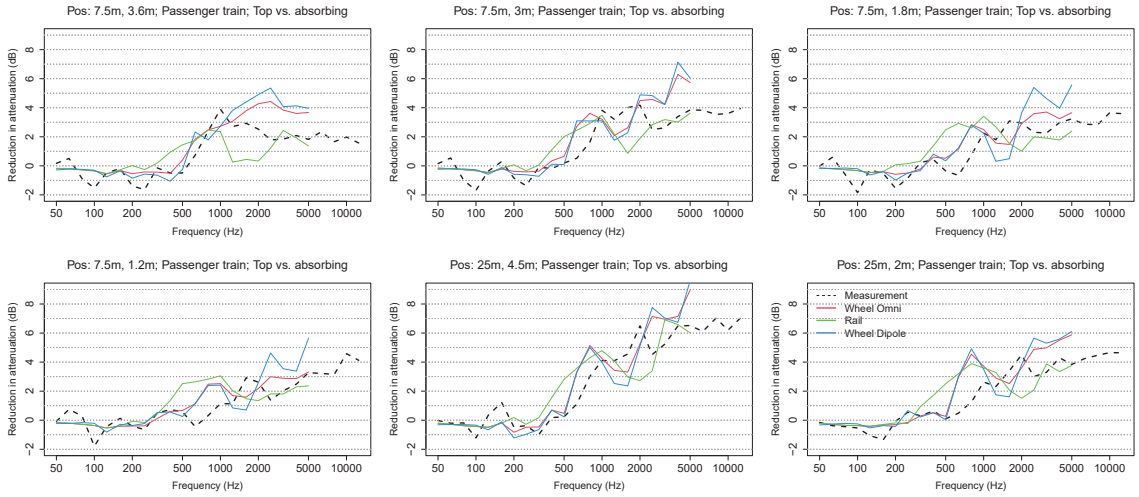


Figure 4: Measured and simulated reduction in attenuation for the top reflecting barrier compared to the fully absorbing barrier for the passenger train. The panels illustrate different microphone positions. Thin colored lines show calculation results for different source positions.

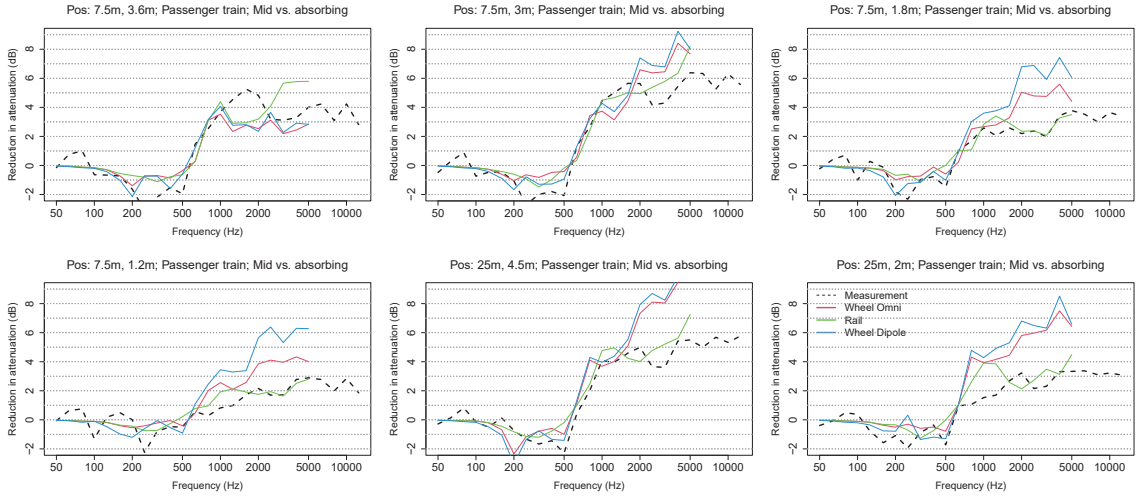


Figure 5: Measured and simulated reduction in attenuation for the middle reflecting barrier compared to the fully absorbing barrier for the passenger train. The panels illustrate different microphone positions. Thin colored lines show calculation results for different source positions.

As mentioned, there were two main types of commuter trains operating on the track. Both trains had a similar total length and height. The main difference between the two trains was that one type had a slightly curved vehicle body cross-section (referred to as commuter train 2). Looking at commuter train 1 (flat exterior), the measurement and calculation results are partially in good agreement (middle panel configuration in Fig. 6). The cross-section used for the calculations was very similar to the one for the passenger train. Gray dashed lines show the measurement results of the passenger train for comparison. Again, positioning the source at the center of the wheel leads to a considerable overestimation of the effect of the reflecting panels. Note, that the wheel source was positioned 50 mm lower due to the smaller wheels of the commuter trains.

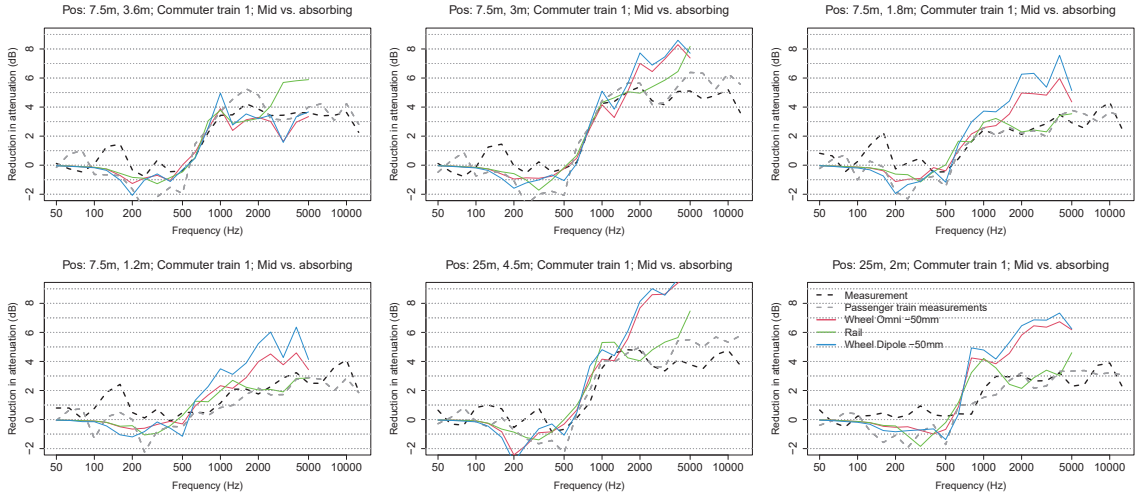


Figure 6: Measured and simulated reduction in attenuation for the middle reflecting barrier compared to the fully absorbing barrier for commuter train 1. The panels illustrate different microphone positions. Thin colored lines show calculation results for different source positions.

In contrast, for the commuter train 2 the panels caused a considerably smaller effect for all placements across most microphone positions (black vs. gray dashed lines in Figs. 7 and 8). For the top panel placements (Fig. 7),

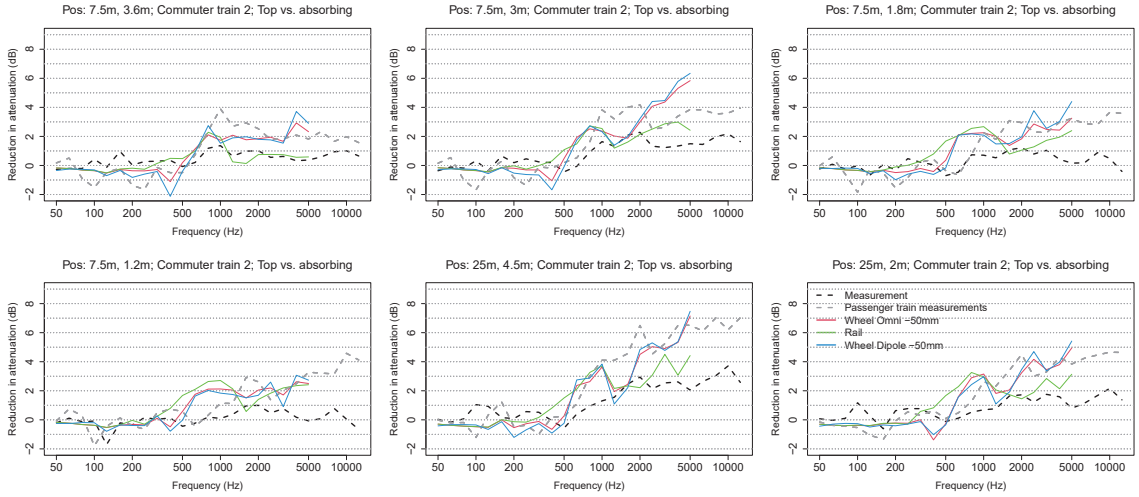


Figure 7: Measured and simulated reduction in attenuation for the top reflecting barrier compared to the fully absorbing barrier for commuter train 2. The panels illustrate different microphone positions. Thin colored lines show calculation results for different source positions.

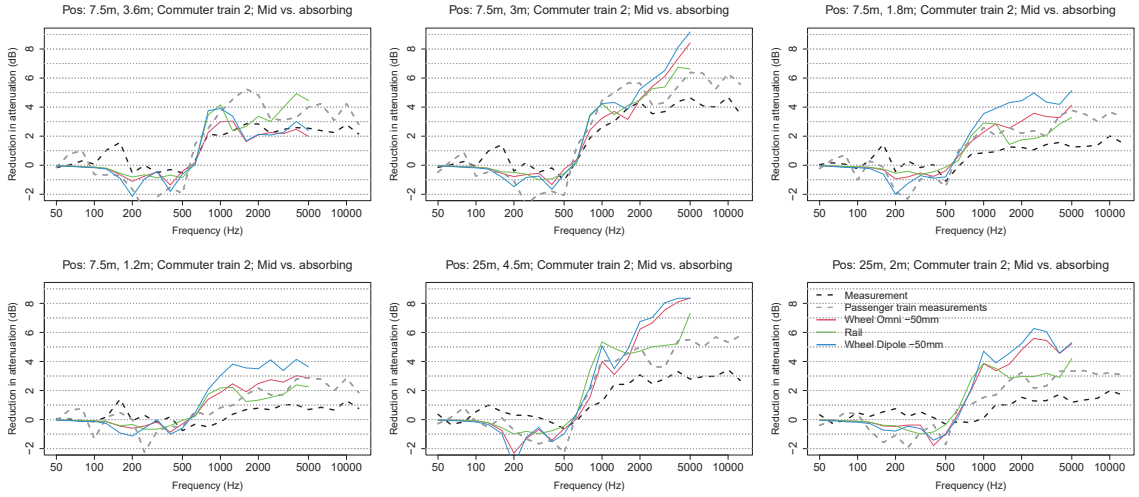


Figure 8: Measured and simulated reduction in attenuation for the middle reflecting barrier compared to the fully absorbing barrier for commuter train 2. The panels illustrate different microphone positions. Thin colored lines show calculation results for different source positions.

the effect is reduced by up to 4 dB compared to the passenger train. Even though the curved vehicle body was considered in the simulations for this train type, all source positions lead to an overestimation of the panel effect. Again, the source positioned on top of the rail leads to a better agreement. For the middle panels (Fig. 8) the reduction of the effect of the reflecting panels is up to 3 dB compared to the passenger train. Overall, the results for the calculations with the curved vehicle body were more similar to the passenger train.

Fig. 9 summarizes the findings for the passenger train and the commuter trains. Clearly, positioning the source at the rail (gray boxes) leads to an overall better agreement of simulations and measurements with a smaller median deviation as well as a smaller range of deviation indicated by reduced inter-quartile ranges. For commuter train 2 the overestimation at high frequencies is also clearly visible. The data presented also include the alternating panel configuration. For the alternating variant the approach using an average absorption coefficient also seems

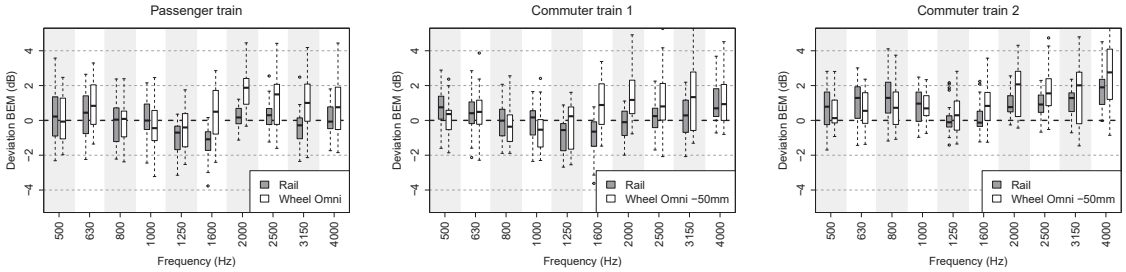


Figure 9: Deviation of simulations from measurements. Each panel shows the deviation of the simulations across microphones and variants (top, middle, and alternating reflecting panels) as a function of source position and frequency for the passenger train and the two commuter trains. The shading of the boxes indicates the source position. Only the omnidirectional sources are shown. Positive values imply an overestimation of the panel effect in the calculation.

to lead to similar agreements as for the top and middle variant.

Compared to passenger trains (including commuter trains) which have a mostly constant and known cross-section, freight trains vary much more in shape and can in practice only be modeled using rough estimations of the vehicle body. Here, results for extreme deviations from the box-shaped passenger train are shown: a flat wagon without any load, i.e., a wagon with essentialy no acoustically relevant vehicle body and a tank wagon with a circular cross-section (Fig. 10).

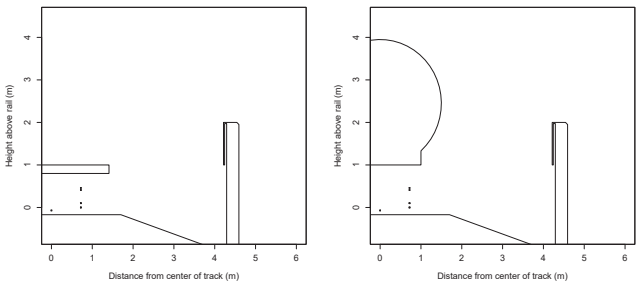


Figure 10: Cross-sections used for the simulation of freight trains. The left panel shows the flat freight wagon without load, the right panel a tank wagon.

Fig. 11 illustrates the results for a train comprising flat wagons without load (i.e. presumably without any acoustically relevant vehicle body for top reflecting panels, left panel in Fig. 10). As was to be expected, neither measurements nor simulations exhibit any significant effect of the reflecting panels.

As an extreme deviation from the box-shaped exterior of many passenger trains, tank wagons were also considered. As the exact exterior could not be acquired, a rough estimate of the diameter was done from the videos recordings (right panel in Fig. 10). The reflecting panels have less effect than for a box-shaped exterior, in particular for the middle panel position and elevated microphone positions (Fig. 12). Although, the calculations agree for most of the microphone positions, for the high position in 25 m, the results for the different panel positions are slightly contradictory, as the rail position leads to a good agreement for the top (not shown) and the wheel position leads to a good agreement for the middle position (Fig. 12). Positioning the source in between (100 mm above the top of the rail or 50 mm lower than the large wheel) leads to intermediate result.

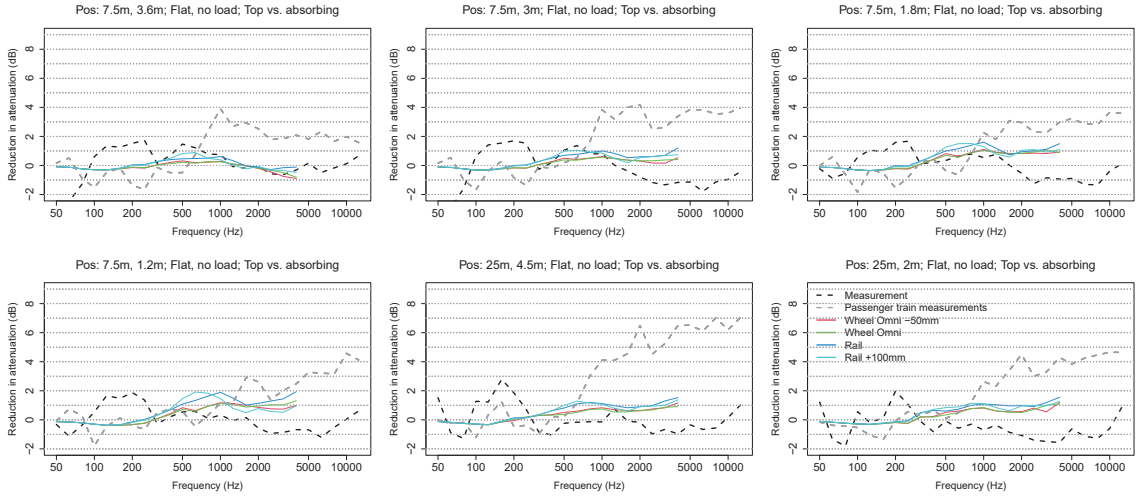


Figure 11: Measured and simulated reduction in attenuation for the top reflecting barrier compared to the fully absorbing barrier for flat freight wagons without load. The panels illustrate different microphone positions. Thin colored lines show calculation results for different source positions.

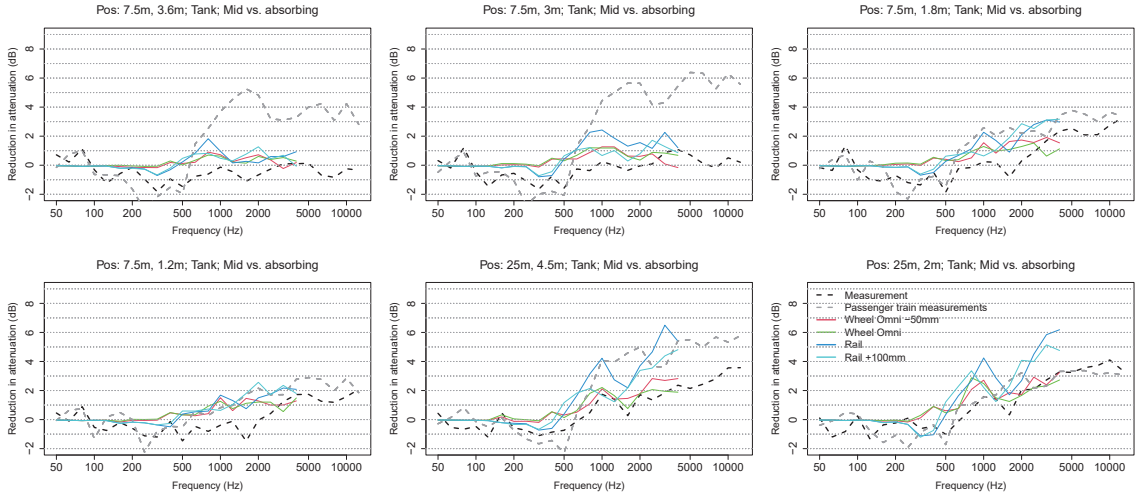


Figure 12: Measured and simulated reduction in attenuation for the middle reflecting barrier compared to the fully absorbing barrier for tank wagons. The panels illustrate different microphone positions. Thin colored lines show calculation results for different source positions.

4 Conclusions

Summarizing, the data illustrate that the panel positioning as well as the the train type have a considerable influence on the effect of reflective elements on a noise barrier. The exact reasons for the influences are not entirely clear. For commuter train 2, which has a slightly rounded shape which was taken into account in the simulations, the comparison between measurements and simulations showed larger deviations than for the passenger train and commuter train 1. Thus, the shape seems to be only a part of the reason for the altered effect of reflecting panels. Preliminary simulations show that the deviations between measurements and simulations when only rolling noise is used can be explained to a large degree by an additional source on top of the train (e.g. air conditioning units).

Overall, the 2.5D boundary element method provides a good tool to investigate such specific research questions. Considering the uncertainties present in in-situ passby measurements the agreement with the measured data is in parts very good with median deviations mostly below 1 dB. For the alternating panel configuration the simplifying assumption of an average absorption coefficient turned out to be sufficient for the situations considered here.

Using this calculation model, the current work is to simulate a large number of different settings (barrier heights, panel placement, train types) and compare the results to noise mapping calculations in order to identify any potential shortcomings of the currently used prediction method.

Acknowledgements

This work was in part supported by the Austrian Research Promotion Agency (FFG, project 873177), the Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology, and the Austrian Federal Railways (ÖBB-Infrastruktur AG).

References

[1] C. Kirisits, H. Meidl, G. Dinhobl, H. Gutschelhofer, J. Punk, and H. Kirisits. Comparison of measurements and calculations to investigate the effect of multiple-reflections between absorptive noise barriers and trains. In *Proceedings of Internoise*, 2013.

[2] European Union. Directive 2002/49/EC of the European Parliament and of the Council, 2021.

[3] P. Morgan, D.C. Hothersall, and S.N. Chandler-Wilde. Influence of Shape and Absorbing Surface—a Numerical Study of Railway Noise Barriers. *Journal of Sound and Vibration*, 217(3):405–417, 1998.

[4] C. Bustos, V. Jurdic, C. Sharp, and D. Hiller. Optimisation of Railway Noise Barrier Design Using Finite Elements and Boundary Element Modelling Methods. In *Euronoise 2021*, pages 390–399, 2021.

[5] C.H. Kasess, W. Kreuzer, and H. Waubke. An efficient quadrature for 2.5D boundary element calculations. *Journal of Sound and Vibration*, 382:213–226, 2016.

[6] D. Duhamel. Efficient calculation of the three-dimensional sound pressure field around a noise barrier. *Journal of sound and vibration*, 197(5):547–571, 1996.

[7] R.A. Broadbent, D.J. Thompson, and C.J.C. Jones. The acoustic properties of railway ballast. In *Euronoise*, pages 3307–3316, 2009.

[8] J. Smardzewski, T. Kamisiński, D. Dziurka, R. Mirski, A. Majewski, A. Flach, and A. Pilch. Sound absorption of wood-based materials. *Holzforschung*, 69(4):431–439, 2015.

[9] C.H. Kasess, H. Waubke, M. Conter, C. Kirisits, R. Wehr, and H. Ziegelwanger. The effect of railway platforms and platform canopies on sound propagation. *Applied Acoustics*, 151:137–152, 2019.

[10] 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*, 141(5):3151–3162, 2017.

[11] S.M.B. Fard, H. Peters, N. Kessissoglou, and S. Marburg. Modelling the acoustic performance of a noise barrier with a quasi-periodic boundary condition. In *Proceedings of the Forum Acusticum*, 2014.

[12] Z. Průša, P.L. Søndergaard, N. Holighaus, C. Wiesmeyr, and P. Balazs. The Large Time-Frequency Analysis Toolbox 2.0. In Mitsuko Aramaki, Olivier Derrien, Richard Kronland-Martinet, and Sølvi Ystad, editors, *Sound, Music, and Motion*, Lecture Notes in Computer Science, pages 419–442. Springer International Publishing, 2014.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Measured and Computed Reflection from a Finite Plate

Sondre Utmo Vikøren[1,*], **U. Peter Svensson**[1]

[1]Acoustics group, Dept. of Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway.

[*]sondre.u.vikoren@ntnu.no

Abstract

Inaccuracies in computation of sound propagation through complex geometries are related in part to inadequate modelling of edge diffraction. Prediction methods need to be evaluated against reference results which could be benchmark measurements, of adequate accuracy. This paper presents impulse response measurements for a rigid 1x1m plate in an anechoic room. The plate was made of 20 mm plywood with a smooth and hard surface. The measurements used a 25 mm diameter tweeter as a source and a 1/2-inch microphone. Using fixed source and receiver positions, the plate was suspended from the ceiling, and rotated in steps of 15 degrees around its vertical symmetry axis. This gave cases with and without specular reflections, with and without direct sound. A signal-to-noise ratio larger than 30 dB was found for the frequency range of 125 Hz to 20 kHz. Calculations were made with the Matlab Edge diffraction toolbox, modeling the loudspeaker as a point source, and the plate as thin or thick. Measured and predicted 1/6-th octave band levels were compared and for the 1/6-th octave band from 125 Hz to 16 kHz, and across 12 plate rotation angles, the 2.5% and 97.5% percentiles and median values were computed. For one of the two setups, using a thick plate model and diffraction order 3, the percentile values were -1.1 dB and +1.8 dB, respectively, with a median value of 0.36 dB. For diffraction order 2, and a thick plate model, these results were only marginally worse, with the percentile values -1.4 dB and +1.8 dB.

Keywords: Edge-diffraction, measurements, simulations

1 Introduction

Geometrical acoustics (GA) is the dominating computational method in room acoustics, and outdoors in city environments [1]. A major challenge is that GA can not handle finite reflectors, but edge diffraction modeling can improve that [2]. Other, more accurate, methods such as BEM, FEM, FDTD, are too computationally demanding for many cases, with computational load scaling poorly with problem size and frequency.

The addition of edge diffraction modeling adds significantly to the complexity of GA software. How important is it? Some benchmarks were published recently [3] that included some cases that were particularly testing different software algorithms' abilities to handle edge diffraction. Measurements and simulations with different GA softwares were presented. The differences between measurements and simulations were quite substantial already for such a simple case as reflection from a finite rigid plate. A finite rigid plate is a classical test case which has been studied earlier [4], and simplified modeling approaches have been suggested as well [5].

The goal of this paper is to compare carefully conducted measurements of scattering from a rigid plate, with calculations made with the freely available edge diffraction Matlab toolbox [6].

2 Theory

As an extension of GA the Edge Source Integral Equation (ESIE) method decomposes a soundfield generated by a convex scattering object into three components, namely the direct (p_{dir}), the specular reflected (p_{spec}) and

the diffracted sound pressure components of first and higher orders (p_{FOD}, p_{HOD})[7].

$$p = p_{\mathrm{dir}} + p_{\mathrm{spec}} + p_{\mathrm{FOD}} + p_{\mathrm{HOD}} \tag{1}$$

The direct sound and specular reflections are the same components as calculated by GA methods. The first-order diffraction term, p_{FOD}, is computed for all edges that are visible from both the source and the receiver, using time-domain [8] or frequency-domain expressions [9]. The final term, p_{HOD}, can be computed in two ways. A time-domain order-by-order method can be used for low orders of diffraction [2], while an integral equation formulation can be used to compute arbitrarily high diffraction orders for convex objects using a frequency-domain formulation [7]. The order-by-order method has a computational cost which increases exponentially with diffraction order whereas the integral equation formulation has a computational cost which grows linearly with diffraction order, with a higher ”startup cost”.

These calculations are implemented in to various degrees in the edge diffraction Matlab toolbox version B1 [10] as well as in the version 0.216 [6]. In this study, only low orders of diffraction will be used, which means that time-domain calculations, using the order-by-order method, are more efficient than the frequency-domain integral equation method. The time-domain first-order diffraction is computed with a numerical integration scheme which is very accurate for low frequencies and with a gradually reduced accuracy for high frequencies. Therefore, a high sampling frequency is typically used. For the second- and higher-order time-domain diffraction calculations, each edge is pre-divided into elements that are made larger for each higher diffraction order. For second- and higher-order diffraction, each element’s contribution is computed for the center points of the elements. Calculation settings are specified below.

3 Scattering measurements

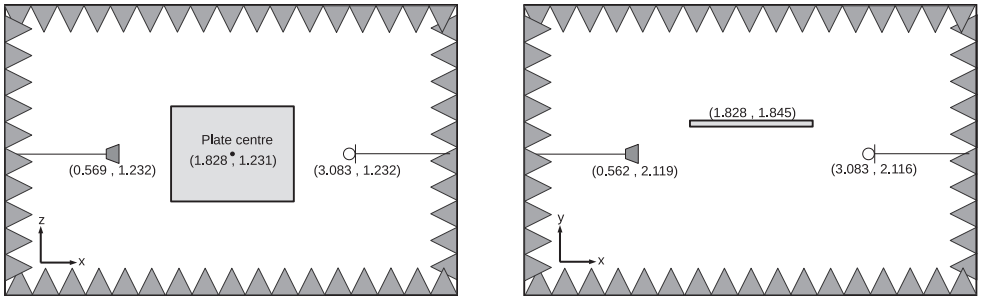
The measurements were performed in the anechoic chamber at NTNU Gløshaugen. The setup consisted of a 1-inch dome tweeter (SEAS®H614) as sound source, a 1/2-inch microphone (Brüel & Kjær®4190) and a 1x1 m plywood plate with 20 mm thickness. The plywood plate had phenolic resin smooth surface layers, and strips of plexiglass were glued to the cut edges. Figs. 1 a) and b) show the geometry of the setup. The loudspeaker element was screwed to a horizontal metal rod and the microphone was placed at the end of a horizontal metal tube. Between the source and receiver the plate was suspended from the ceiling by two metal wires.

The software AFMG Easera®, version 1.2.16.4, and a Lynx Aurora®audio interface, were used to measure impulse responses, using exponential sine sweeps, ”chirps”, as stimuli. Chirps covered the full range from 0 Hz to half the sampling frequency, which was 44.1 kHz. The recommended frequency range for the used tweeter is 3 kHz to 25 kHz, but here the tweeter was used down to around 118 Hz. A low amplitude was used so that repeated measurements did not show any trend in frequency response amplitude, for the frequency range of interest. A trend of increasing or decreasing magnitude, over time, could have indicated heating effects in the voice coil or suspension.

The measurements were performed for 12 different plate positions. The different positions consisted of rotating the plate around its central vertical axis in 15 degree increments. For each plate position impulse responses were measured with two different combinations of source-receiver positions. Source-receiver position one had both the source and the receiver placed at the same height as the centre of the plate. Source-receiver position two had the source being 125 mm higher than the centre of the plate while the receiver was placed 250 mm lower. Figure 2 illustrates this. Additionally, for both source-receiver positions, the impulse response was measured without the plate. The used loudspeaker is small but with a directivity which increases with frequency. For the setup shown in Figs. 1 a) and b), the largest radiation angle which reaches the plate is around 30 degrees relative to on-axis radiation. The directivity for the same type of loudspeaker was measured in a previous study [11], averaged across octave-band-wide frequency ranges, at a few distinct angles and frequencies, as reproduced in Table 1.

Table 1: Octave-band-wide directivity values for the used loudspeaker element, at the non-standard center frequencies as indicated.

Radiation angle [deg.]	**625 Hz**	**1.25 kHz**	**2.5 kHz**	**5 kHz**	**10 kHz**	**20 kHz**
0°	0 dB	0 dB	0 dB	0 dB	0 dB	0 dB
10°	-0.1 dB	0.0 dB	-0.2 dB	0.0 dB	-0.5 dB	-1.4 dB
20°	-0.4 dB	+0.2 dB	-0.6 dB	-0.6 dB	-1.7 dB	-6.0 dB
30°	-0.6 dB	+0.3 dB	-1.1 dB	-1.6 dB	-3.4 dB	-15.9 dB



(a) (b)

Figure 1: Coordinates of physical measurement setup. (a) xz-plane. (b) xy-plane

4 Numerical simulations

The numerical simulations were performed with the Edge Diffraction Matlab toolbox version B1, and executed with Matlab 2021a. The most essential settings used for the calculations are summarized in table 2. Here the ”element size” gives the number of elements per wavelength for the shortest wavelength (i.e., at half the sampling frequency). As mentioned in the Theory chapter, first-order diffraction uses another integration scheme than higher-order diffraction.

The toolbox uses omnidirectional sound sources with an ideal Dirac impulse response in the computations. As can be seen in Table 1, the (octave-band wide) directivity of the measurement loudspeaker is within [-0.6,0] dB for a large part of the spectrum and radiation angles. The increasing directivity for high frequencies and large radiation angles will certainly limit the accuracy of the numerical simulations, but to a small degree. To remove the influence of the measurement loudspeaker’s non-flat frequency response, all measured frequency responses (transformed from the measured impulse responses using the discrete Fourier transform) were divided by the free-field loudspeaker response, which largely eliminated the effects of the loudspeaker’s response. The same was done for the simulations, which made the measured and simulated relative responses comparable. However, a time windowing was used before transforming the measured impulse responses to frequency responses (see subsection 1), and this windowing will inevitably affect the response, and primarily the low-frequency end of the response. To get similar effects for the simulated responses, the simulated impulse responses were convolved with *a model* of the measurement loudspeaker. A model was used, rather than the measured response, in order to remove the effects of noise and small scattering details that might change with radiation direction.

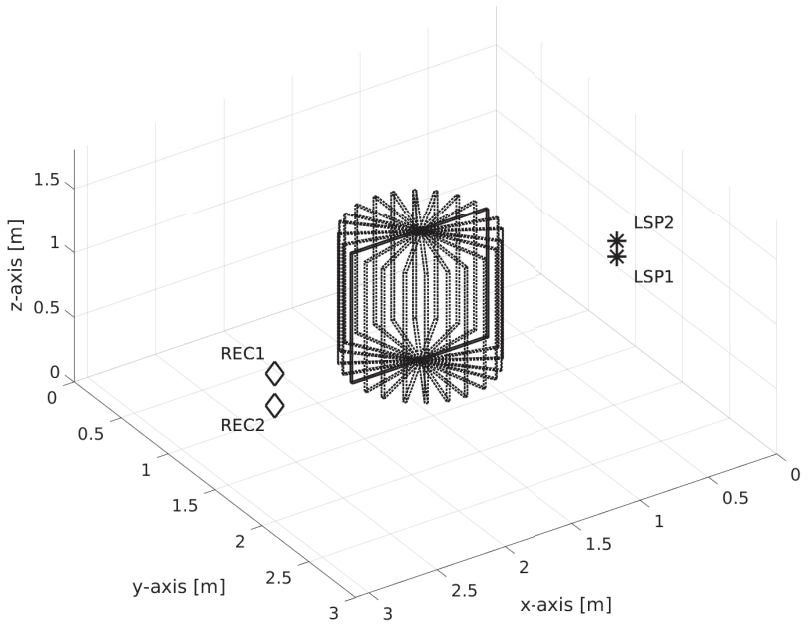


Figure 2: Measurement setup. All plate angles indicated. LSP1 and LSP2 refer to the source position for setup 1 and 2 respectively. Similarly REC1 and REC2 refer to the receiver position for setup 1 and 2. For setup 1, both source and receiver had the same z-coordinate as the plate centre, i.e. no vertical difference between the two positions. For setup 2, the source was shifted 125 mm in positive z-direction while the receiver was shifted 250 mm in negative z-direction direction, making the vertical difference between the two 375 mm.

4.1. Developing a model of the measurement loudspeaker

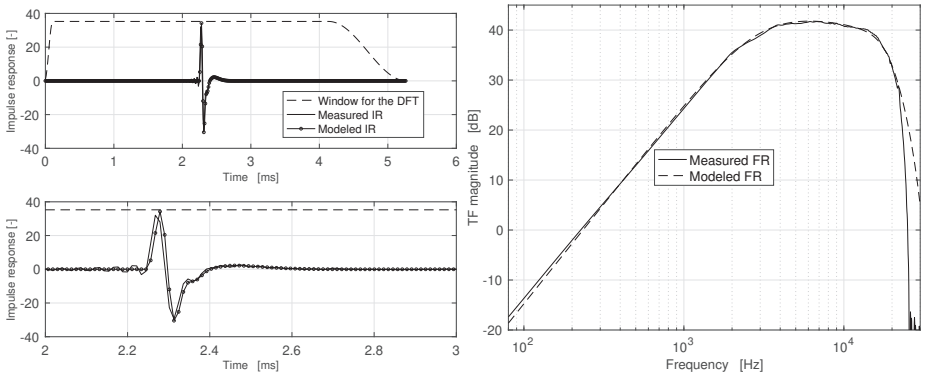
The free-field impulse response was measured in the main direction. A short window was applied before transforming the impulse response to a frequency response.

An IIR filter was fit to the magnitude of the frequency response, as follows. First, a second-order high-pass filter was designed by tweaking manually the cutoff frequency and the Q-value. Then a third-order IIR filter was fit to the difference between the measured response and the modeled high-pass filter response, using Matlab's `yulewalk` function. Finally, an extra second-order low-pass filter at 22 kHz was introduced. These filters could be combined to an 7-th order IIR filter. The results are shown in Fig. 3. Since the measured response was up-sampled from a sampling frequency of 44.1 kHz to 88.2 kHz (which was the sampling frequency of the simulations), a very steep fall-off can be observed above 22 kHz. In Fig. 3 (a) a small fraction of a sample time difference can be observed between the measured IR and model IR, since no attempt was made to adjust the delay of the model impulse response to less than integer samples.

The window used for this free-field response is also indicated in Fig. 4 (a): a 10-sample long half-Hanning window started at 200 samples (2.27 ms) before the max peak of the measured impulse response. At the end, a 100 samples long (1.1 ms) half-Hanning window ended 3 ms after the max peak.

Table 2: Key settings used in the toolbox

Variable	Value
Sound speed [m/s]	342.74
Sampling rate [Hz]	88200
Diffraction order	1-3
Element size (1./2./3. diffraction order)	-/1/0.5 elements per shortest wavelength



(a) (b)

Figure 3: (a) The measured impulse response of the loudspeaker used in the scattering measurements, and the model impulse response, with the used window indicated, and (b) the corresponding frequency responses.

5 Processing results

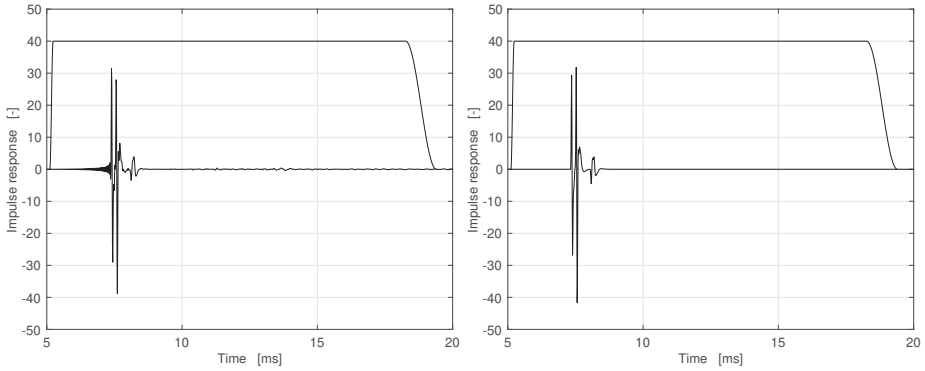
5.1. Windowing the measured response

To get as comparable results as possible, both the measured and the simulated IRs were windowed before transforming them to the frequency domain via a discrete Fourier transform (DFT). Prior to the windowing, the simulated IRs were convolved with a loudspeaker model IR described in section 4. The peak of the impulse response was detected and 2.27 ms before, and 12 ms after, the peak were kept. In addition, a half Hanning window of length 0.11 ms (10 samples) was applied at the start, and another half Hanning window of length 1.1 ms (100 samples) was applied to the end of the window, as illustrated in Fig. 4.

Examples of simulated and measured frequency responses are shown in Fig. 5. They were computed using a DFT size which gives a frequency resolution of 1 Hz. The measured noise was determined by repeating an impulse response measurement without sending an excitation signal to the loudspeaker. From this single noise sample, it seems like a signal-to-noise ratio of at least 40 dB results from around 200 Hz, and at least 30 dB from around 100 Hz.

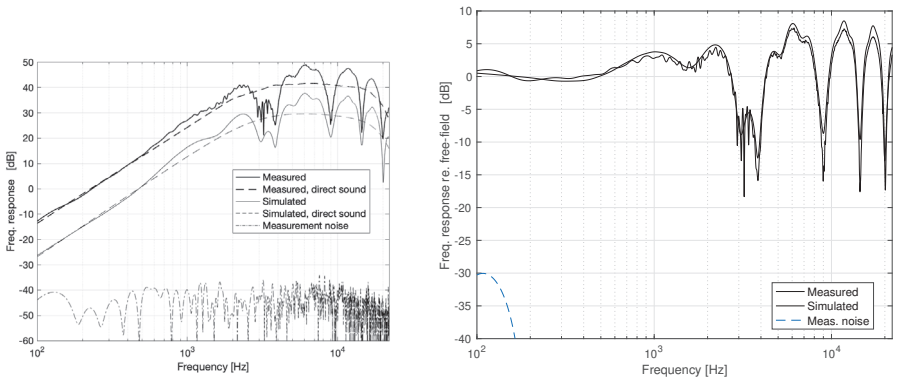
5.2. Fractional-octave band smoothed responses

Rather than analyzing the difference between the simulated and measured narrowband frequency responses in Fig. 5(b), the fractional octave-band smoothed versions are analyzed instead. $H(\omega_k)$ denotes the (complex-valued) frequency response for the set of discrete angular frequencies ω_k, and $H^2(\omega_b)_{1/n}$ denotes the $1/n$-



(a) (b)

Figure 4: An example of (a) a measured impulse response and (b) the corresponding simulated impulse response, both with the used window indicated. The simulated impulse response has been convolved with a model impulse response that mimics the loudspeaker response.



(a) (b)

Figure 5: An example of (a) measured and simulated frequency responses, and (b) frequency responses relative to the direct sound (= free-field response), for the same case as in Fig. 4. The simulated response includes the high-pass filter response that roughly mimics the loudspeaker response. The measured noise is computed from an IR measurement and processing, without sending an excitation signal to the loudspeaker.

octave-band smoothed version, calculated for the $1/n$-octave-band center frequencies ω_b,

$$H^2(\omega_b)_{1/n} = \frac{1}{N}\sum_{k=1}^{N} |H(\omega_k)|^2$$

where the summation is understood to include the frequency bins that fall within each fractional octave band. Thus, the relative smoothed responses are given by

$$H^2(\omega_b)_{rel.,1/n} = \frac{H^2(\omega_b)_{1/n}}{H^2(\omega_b)_{direct,1/n}}.$$

It is very common to use 1/3rd-octave band levels in acoustical measurements. Here 1/6th-octave band smoothing has been used to get more detailed comparisons.

6 Results

To display the difference between the two different measurement setups, Fig. 6 shows the median and the 97,5 and 2,5 percentiles of the differences between measured and simulated fractional-octave band levels for the two different measurement setups. Each dataset, for which the statistical parameters are computed, involves 12 plate angles and 43 1/6-th octave bands (125 Hz - 16 kHz), that is, 516 individual level differences.

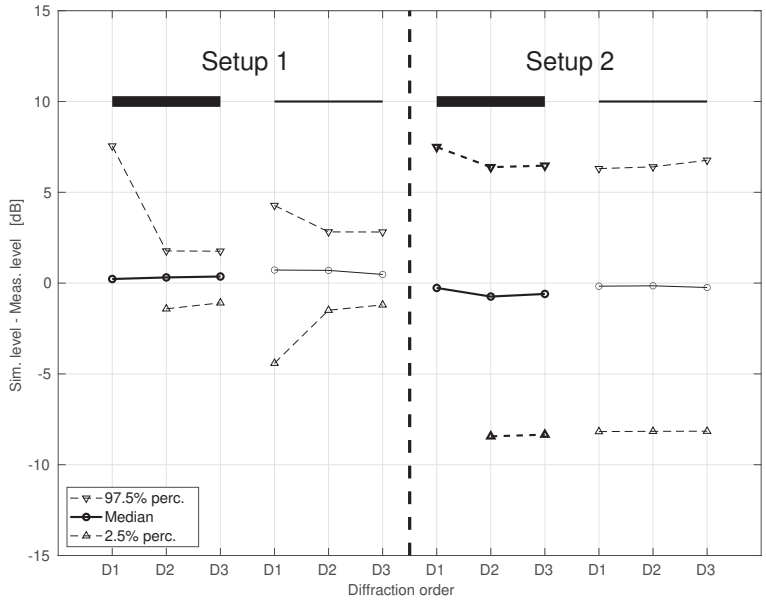


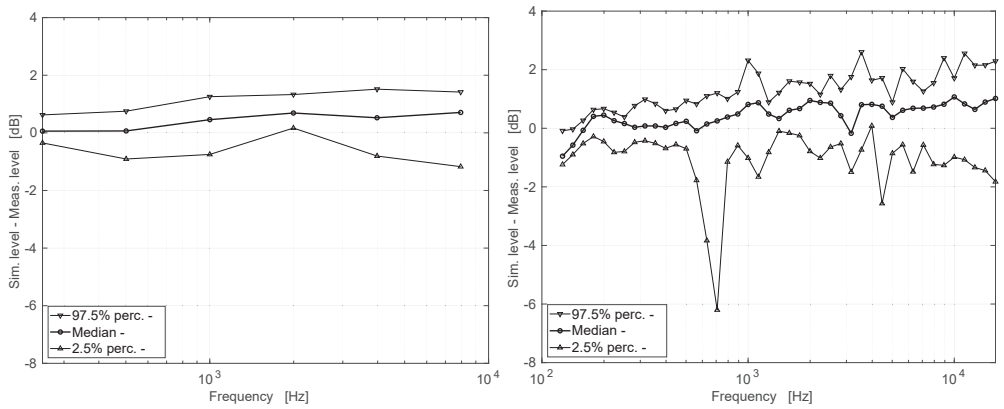
Figure 6: Median and percentiles of the difference between measured and simulated results. The values are calculated with 1/6th-octave band smoothing, and 1/6th-octave band sampling, for each simulation case: thin and thick plate, and diffraction order 1-3. The setup and plate thickness is indicated above the curves in the figure while the diffraction order is indicated on the horizontal axis.

Fig. 6 suggests that measurement setup 1 has generally been simulated with greater accuracy than setup 2. The median values are very similar, within $\pm$ 1 dB for all diffraction orders, for both setups 1 and 2. However, the spread in agreement, illustrated by percentile values, is substantially greater for setup 2 than for setup 1. We draw the intermediate conclusion that the positioning uncertainty was higher for setup 2 than for setup 1. In this short paper, we focus on setup 1. For setup 1 one can also make a few observations. First, diffraction order 1 gives less accuracy for both the thin and thick plate modeling. Second, using the thick plate model gives greater accuracy.

6.1. Frequency dependence of results

In Fig. 6, results for all frequency bands were pooled together. However, it can be assumed that the difference between simulations and measurements is larger for middle and higher frequencies than for lower, for several reasons. One reason is that a certain uncertainty of the plate dimension, and of the transducer positioning, leads to larger and larger uncertainties for interference effects as the wavelength gets shorter. Another reason is that the used loudspeaker and, to a smaller degree, microphone, show directivity effects above some frequency. The interference effects would largely be smoothed out and be smaller for octave-band smoothed responses than for fractional-octave band smoothed results. Fig. 7 show the octave-band and 1/6th octave-band smoothed results,

and the variation range, expressed as the 95%-range, increases gradually from around 4 kHz. Also, the median values in Fig. 7(b) indicate a systematic overestimation that gradually increases from around 600-1000 Hz, up to 1 dB around 10-20 kHz. This gradual increase might be caused by the slowly increasing directivity of the used loudspeaker element. For each frequency value in Fig. 7, the results for 12 different angles contribute to the median and percentile values.

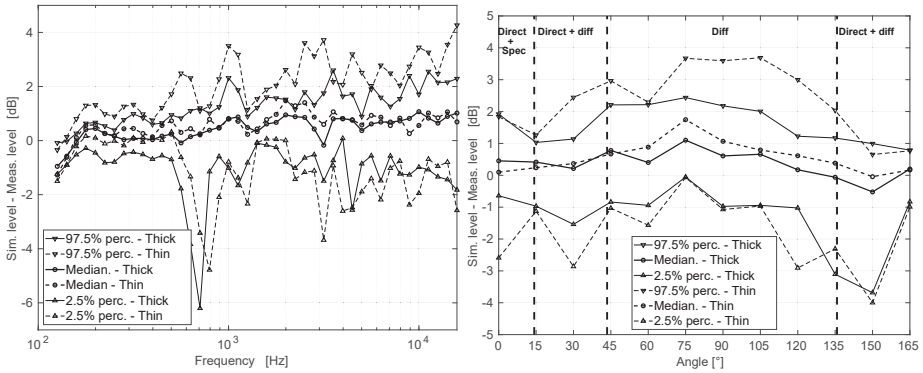


(a) (b)

Figure 7: Simulated level - measured level as function of frequency, for (a) octave-band smoothed responses, and (b) 1/6-th octave band smoothed responses. Results are shown only for setup 1, and for the simulations using a thick plate model, and third-order diffraction.

6.2. Thin vs. thick plate

The plate used in the measurements had a size of 1 m by 1 m, and a thickness of 20 mm. It seems reasonable that the plate could be modeled thin or thick for low frequencies, but that a thin plate modeling might be inadequate at high frequencies. To investigate that, Fig. 8 shows the median and percentiles of the difference between simulations and measurements, using both thick and thin plate, as a function both of frequency and angle.



(a) (b)

Figure 8: Median and percentiles of the difference between simulated and measured levels for the thin and the thick plate model, for source-receiver position 1 and diffraction order 3, 1/6th-octave band frequency smoothing, (a) as function of frequency, and (b) as function of angle. Dashed lines indicate which angles represent different combinations of soundpaths.

In Fig. 8 a), the median values vary between 0 dB and 1 dB (up to 1.5 dB for the thin plate), and no clear difference can be observed between the thin and the thick plate results. For the most part both percentile figures

are closer to zero for the thick plate across the spectrum. There is no clear tendency that the thick plate results are more accurate than the thin plate results for the higher frequencies. When the comparison is plotted as function of plate angle, in Fig. 8 b), it can be observed that the differences, for both the thin and the thick plate, tend to be more positive for the plate angles where only diffraction occurs. A positive difference implies that the simulated levels were higher than the measured ones, which is consistent with a directional loudspeaker - the sound radiated from the measurement loudspeaker towards the edges was somewhat weaker than for a truly omnidirectional loudspeaker. For the plate angle 90 degrees, one of the four edges of the plate will be hit by a wave radiated at 30 degrees radiation angle from the loudspeaker and therefore the influence of the loudspeaker directivity would be strongest around the 90 degree plate angle.

7 Uncertainties

During the measurements, great care was taken to determine the positions of the source, receiver and plate. It was estimated that the coordinates of the source, receiver and plate centre could be determined with an accuracy of $\pm$ 2 mm. The coordinates of the corners of the plate have a higher degree of uncertainty for two reasons. First, the plate was rotated by hand, using a large protractor placed on the floor beneath the plate and vertical laser to aim the plate at the desired angle. This operation was both tedious and difficult to get right. In a separate study, where simulations of the plate were repeated with the plate rotated in steps of 0.1° around the *nominal* rotation angle in the measurements indicated that the error associated with the rotation angles might have been up to $\pm$ 1.2°, which for a plate of a 1 m size translates to $\pm$ 10 mm. Second, the plate itself was measurably skewed, by a few mm, which is difficult to simulate in the current version of the ED toolbox. It is difficult to estimate the total error of these sources of uncertainty, but it is reasonable to estimate the total positioning uncertainties to be on the order of $\pm$ 5-20 mm. The reason for the manual adjustment of the rotation of the plate was a desire for the suspension of the plate to be as thin as possible, minimizing spurious reflections. However, the difficulties associated with the adjustments suggest that the advantages with the increased precision of an electronic turntable might outweigh the disadvantages of spurious reflections.

The second greatest error source is the sound speed, as it affects the arrival of the different components in the impulse response. The greatest influence on the sound speed in this case is the temperature, which was recorded before and after each measurement. The range of the different temperatures recorded was 2 degrees, from 16,9 to 18,9 degrees Celsius. This gives rise to the sound speed varying between 342,05 m/s to 343,30 m/s. The simulations though were performed with the temperature of 18 °C for all measurements and thus the simulated sound speed was set to 342.74 m/s.

8 Conclusions

Measurements and simulations of reflections from a 1x1 m plate rotated around its vertical axis in increments of 15 degrees have been performed. Numerical edge-diffraction based simulations were performed with both a thin and a thick plate model.

The most important, and unsurprising, finding is that the inclusion of second and higher order diffraction improved the accuracy of the simulations substantially, especially for the case of the thick plate. When looking at the difference between measured and simulated levels, see Fig. 8, the thick plate for the most part performs better than the thin plate across the spectrum. Naturally the thick plate is performing worse than the thin plate for first order diffraction, as the cases where the direct path between source and receiver is obstructed produce no sound in the simulations. The most accurate simulation results were found for the thick plate model, diffraction order 3, and source-receiver setup 1. Across 516 comparisons of 1/6th-octave band levels, from the 125 Hz band to the 16 kHz band, and 12 plate angles, the 2.5% and 97.5% percentiles were -1.1 dB and +1.8 dB, respectively, with a median value of 0.36 dB. For diffraction order 2, these results were only marginally worse,

with the percentile values -1.4 dB and +1.8 dB. Uncertainties associated with the exact positioning of the plate have affected the results to a non-negligible degree. Also, the loudspeaker element's directivity have probably also affected the results to some degree. Further measurements of this type will be performed with the inclusion of a electronic turntable to reduce the amount of plate position error, at the cost of more spurious reflections.

Acknowledgements

The authors would like to thank Tim Cato Netland for assistance in the preparations for, and conduction of, the measurements.

References

[1] Lauri Savioja and U. Peter Svensson. Overview of geometrical room acoustic modeling techniques. *The Journal of the Acoustical Society of America*, 138(2):708–730, 2015. doi: 10.1121/1.4926438. URL https://doi.org/10.1121/1.4926438.

[2] U. P. Svensson, R. I. Fred, and J. Vanderkooy. An analytic secondary source model of edge diffraction impulse responses. *The Journal of the Acoustical Society of America*, 106(5):2331–2344, 1999. doi: http://dx.doi.org/10.1121/1.428071. URL http://scitation.aip.org/content/asa/journal/jasa/106/5/10.1121/1.428071.

[3] Fabian Brinkmann, Lukas Aspöck, David Ackermann, Steffen Lepa, Michael Vorländer, and Stefan Weinzierl. A round robin on room acoustical simulation and auralization. *The Journal of the Acoustical Society of America*, 145(4):2746–2760, 2019. doi: 10.1121/1.5096178. URL https://doi.org/10.1121/1.5096178.

[4] Trevor J Cox and YW Lam. Evaluation of methods for predicting the scattering from simple rigid panels. *Applied Acoustics*, 40(2):123–140, 1993.

[5] J. J. Rindel. Attenuation of sound reflections due to diffraction. In *Proceedings of the Nordic Acoustical Meeting, Aalborg, Denmark*, 1986.

[6] U. Peter Svensson. Edge diffraction toolbox for matlab, v. 0.216 [computer program]. . URL https://github.com/upsvensson/Edge-diffraction-Matlab-toolbox.

[7] A. Asheim and U. P. Svensson. An integral equation formulation for the diffraction from convex plates and polyhedra. *The Journal of the Acoustical Society of America*, 133(6):3681–3691, 2013. doi: http://dx.doi.org/10.1121/1.4802654.

[8] Svensson U. P. Torres, R. R. and M. Kleiner. Computation of edge diffraction for more accurate room acoustics auralization. *The Journal of the Acoustical Society of America*, 109(1):600–610, 2001. doi: http://dx.doi.org/10.1121/1.1340647.

[9] U. Peter Svensson, Paul T. Calamia, and Shinsuke Nakanishi. Frequency-domain edge diffraction for finite and infinite edges. *Acta Acustica united with Acustica*, 95:568–572, 2009.

[10] U. Peter Svensson. Edge diffraction toolbox for matlab, version edb1 [computer program]. . URL https://folk.ntnu.no/ulfps/software/index.html.

[11] A. Lovstad and U. P. Svensson. Diffracted sound field from an orchestra pit. *Acoust. Sci. Tech.*, 26:237–239, 2005.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Mesh2HRTF / NumCalc: An Open-Source Project to Calculate HRTFs and wave scattering in 3D

W. Kreuzer[1,*], **K. Pollack**[1], **P. Majdak**[1], **F. Brinkmann**[2]

[1]Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria,

[2]Audio Communication Group, Technical University of Berlin, Berlin, Germany.

[*]wolfgang.kreuzer@oeaw.ac.at

Abstract

Mesh2HRTF is an open source project originally developed to provide users with an easy-to-use platform for the calculation of the head-related transfer functions based on the boundary element method. However, its boundary element module NumCalc with its easy-to-use input file is flexible, and thus, it can be used as a standalone application for solving the Helmholtz equation in various other contexts such as calculating the sound field inside a car or a duct, or the calculation of the scattering of sound waves in an open field around objects in 3D. NumCalc is a combination of the collocation approach based on the Burton-Miller formulation of the boundary integral equation and the fast multipole method. This paper discusses novel features of NumCalc such as the definition of frequency-dependent admittance boundary conditions, and point to its current limitations with respect to the numerical accuracy and memory consumption.

Keywords: boundary element method, software, acoustic scattering

1 Introduction

In acoustics, the simulation of wave propagation around objects is an important task. Examples include the calculation of head-related transfer functions (HRTFs), which model the filtering effect of the human anthropometry on incoming sound, the simulation of noise mitigation measures, or the modelling of musical instruments [1, 2, 3]. Especially for wave propagation in open domains, the boundary element method (BEM) plays a special role, because compared to other numerical methods like the finite element method or the finite difference method, the BEM only needs a discretization of the surface of the scattering objects, but not the space inside or around the objects.

In this paper, we want to introduce the open source project Mesh2HRTF and its subproject NumCalc. Of course, one may ask "Why another BEM software project?". First of all, Mesh2HRTF is open source, and second, it is relatively easy to use without the need of deep knowledge of BEM topics like boundary integral equations, singular integral operators or methods to deal with critical frequencies.

Mesh2HRTF was originally planned and designed as an easy-to-use tool for people working in 3D audio to calculate HRTFs. The idea behind the Mesh2HRTF-chain is, that users only need to load the geometry of the individual's head and pinnae into the open source project Blender [4], mark a few elements there, and Mesh2HRTF takes care of the rest. To that end, Mesh2HRTF comes with an export script for Blender that generates a file called `NC.inp` that serves as an input for the BEM-code NumCalc. NumCalc then numerically solves the Helmholtz equation, that models the propagation of sound waves, and calculates the sound pressure on the head and at an evaluation grid around the head. Finally, Mesh2HRTF contains Matlab/Octave and Python scripts that calculate HRTFs using the BEM results and that store the filter functions in the SOFA format, a

standardized format for HRTFs [5]. Interested readers are referred to [6] for some tutorials on how to use the pipeline provided with Mesh2HRTF. However, the structure of the input file for NumCalc is very simple, and therefore, the BEM code can be applied to many problems in acoustics. Also, as convenient as Blender and the export script may be, since `NC.inp` is a text file, there are multiple ways to create this file. E.g., for the sphere in Section 4, a simple Octave script was used, see [7] for the octave script used to create Fig. 3.

To be clear, among all the other free and open source BEM projects, Mesh2HRTF and NumCalc are probably not the most elaborate or the most accurate implementations of the Helmholtz BEM, but they provide an easy way to achieve sufficiently accurate simulation results for users not experienced in numerical mathematics or partial differential equations. Nevertheless, we also feel that users should know a bit about the software they are using, about possible caveats and the accuracy they can expect from the calculation.

The aim of this paper is twofold. First, we will use the opportunity to introduce a new feature of the soon-to-be-released version 1.0 of Mesh2HRTF: Frequency-dependent admittance boundary conditions to model sound-absorbing materials. Second, we will talk a bit about the input for NumCalc and its implementation, and provide results for a benchmark example to give users an impression of the accuracy they can expect when using Mesh2HRTF and/or NumCalc. We are fully aware that this paper can only give a little glimpse into some aspects of Mesh2HRTF and its application. A more detailed paper is currently in preparation, we also refer interested readers to the Mesh2HRTF Wiki [8] and to [9].

2 General

The BEM implementation in NumCalc is based on collocation with constant elements, i.e., on each element of the mesh (= discretization of the surface of the scatterer using quadrilaterals or triangles), the acoustic pressure and the particle velocity are assumed to be constant. As a general rule of thumb that works very well for most practical problems, the mesh should contain about 6 to 8 elements per wavelength [10]. See Fig. 1 for an example of a mesh of a human head and pinna that is fine enough to be used for frequencies up to 20000 Hz.

On one hand, it is known that the accuracy of collocation with constant elements is not high compared to other methods like, e.g, Galerkin with higher order elements. However, on the other hand, the implementation of the collocation is easier, the generation of the system matrix uses (in general) less numerical operations and the fact that we are using discontinuous elements allows us to subdivide elements without having to consider neighbouring elements. Also, for a lot of practical applications, a relative error in the range of 10^{-3} to 10^{-2} , which can be achieved by NumCalc, is more than enough.

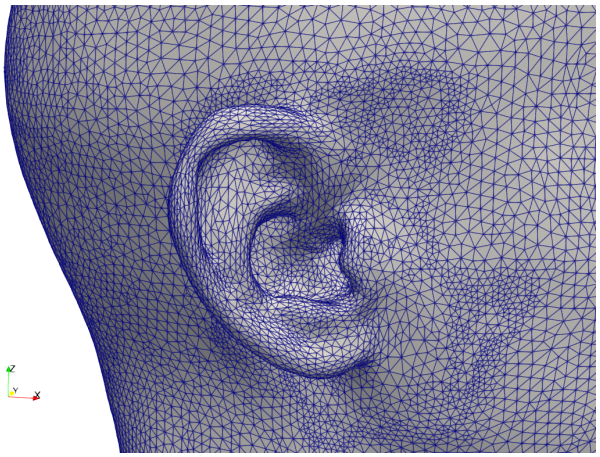


Figure 1: Mesh of the human pinna and parts of the head. The mesh (including parts of the torso) contains about 74000 triangular elements with an average edge length of 2.4 mm.

One drawback of the BEM is the fact the system matrix to be solved in connection with the BEM is densely populated and of size $N \times N$ where N denotes the number of elements. For the mesh in Fig. 1, for example, the system matrix would be a densely populated matrix with about 5.4 billion complex valued entries, which puts a strain on memory and calculation time. To deal with this fact, NumCalc uses the (multilevel) fast multipole method (FMM, cf. [11]) to speed up matrix-vector multiplication and to reduce the amount of memory needed to store the system matrix. In a nutshell, the FMM can be described using a man-in-the-middle principle, where parts of the mesh are grouped into different clusters, and point-to-point interactions are replaced by local expansions between point and center of the cluster, and expansion between cluster centers. This helps in reducing the number of operations from $O(N^2)$ to $O(N \log_2 N)$, also, the system matrix is never constructed in full, which reduces the time to set up the linear system of equations and the memory consumption. The linear system itself is solved using an iterative conjugated gradient solver, cf. [12, Chapter 7.4.1]. However, users should be aware that the FMM is just an approximation of the original system, and that its accuracy and efficiency is dependent on the clustering and the frequencies used.

On a side note: It is also known that the BEM for the Helmholtz equation has numerical difficulties at certain critical frequencies [13, 14]. To deal with these instabilities at critical frequencies, a formulation proposed by Burton and Miller [13] is used, however, this is handled automatically by NumCalc and users do not need to bother about the different integral equations.

3 Important input parameters for NumCalc

In order for NumCalc (and any other Helmholtz BEM code) to work, one needs to provide at least the following information:

- The frequencies for which the scattering problem should be solved,
- the geometry of the scatterer using small triangular or quadrilateral elements,
- boundary conditions for the elements,
- optional: Sound sources and evaluation points, that do not lie on the surface of the objects and where the sound field should be calculated too.

There are also general parameters containing, e.g., information about the speed of sound and the density of the medium, a flag if an interior or an exterior problem needs to be solved, and some advanced parameters that allow experienced users to change details with respect to the clustering used for the FMM. A detailed discussion of these parameters would be out of the scope of this manuscript, the default values set for these parameters cover a wide range of use cases, users do not need to worry about these settings in general. The interested reader is referred to [8] for an in-depth description of these parameters.

3.1. Frequency-defining curves

NumCalc numerically solves the Helmholtz equation in the frequency domain for a given set of frequencies provided by the user. Frequency-defining curves are piecewise linear functions that map a frequency step to a frequency or a frequency-dependent value of a boundary condition. The user has to provide the nodes of these curves. An example: If one would want to calculate the sound field for 10 frequencies between 1000 Hz and 10000 Hz in steps of 1000 Hz, a simple way to describe the curve is given in Lst. 1. A line in the input file that starts with the hash tag '#' is ignored by NumCalc and can be used for comments. The first non-comment line consists of a dummy parameter (`dummy = 1`) followed by the number of frequency steps to be calculated (`Number_of_freqsteps = 10`) followed by the distance h_i between two frequency steps (`stepsize = 1.0`). The last entry in this line is given by the start index i_0 of the curve (`Index0 = 0.0`). The next line contains the

unique number of the curve (in case of a frequency definition curve the curve number is always `Curvenumber = 0`) and the number of nodes defining the piecewise linear function mapping the frequency step n to the respective frequency f_n. In Lst. 1 this map is defined as the line between the two nodes $(S_1, F_1) = (0.0, 0\,\text{Hz})$ and $(S_2, F_2) = (10.0, 10000\,\text{Hz})$. The frequency for the n-th frequency step is then given as

$$f_n = F_1 + \frac{F_2 - F_1}{S_2 - S_1}((i_0 + h_i \cdot n - S_1) = 0\,\text{Hz} + \frac{10000\,\text{Hz}}{10}(0.0 + 1.0 \cdot n) = 1000\,\text{Hz} \cdot n, n = 1, \ldots, 10. \quad (1)$$

```
# '#' denotes a comment line
# dummy_parameter_for_future_use Number_of_freqsteps stepsize Index0
1 10 1.0 0.0
# Curvenumber  Nr_of_nodes_defining_the_curve
0 2
# nodes of the curve (Si, Fi): freq-step vs frequency
0.0 0.0
10.0 10000.0
```

Listing 1: Definition of 10 uniform frequencies between 1000 Hz and 10000 Hz

A more elaborate example is given in Lst. 2. This code snippet contains the definition of 8 third-octave frequency bands (see Tab. 1) where three frequencies are chosen in each band.

```
#dummy n-steps stepsize
1      24   1.0 0.0
# Curvenr. Nr._of_Nodes
0       9
#define the curve
0.0   35.481
3.0   44.668
6.0   56.234
9.0   70.795
12.0  89.125
15.0  112.202
18.0  141.254
21.0  177.828
24.0  223.872
```

Listing 2: Definition of a frequency-defining curve for the third octave band example

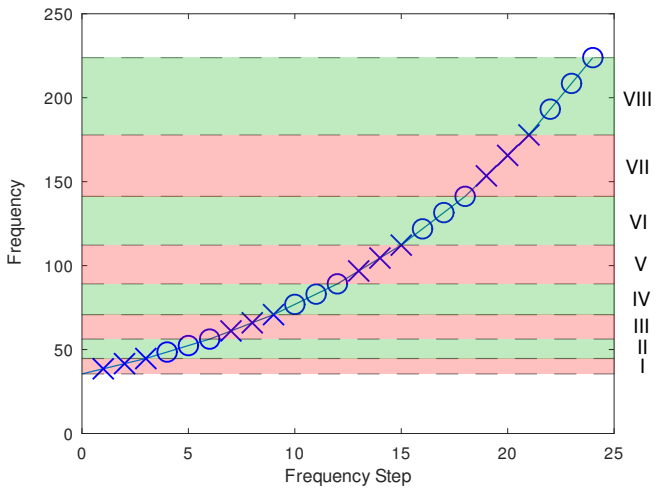


Figure 2: Frequency-defining curve for the third octave band described in Tab. 1.

Band:	I	II	III	IV	V	VI	VII	VIII
Low:	35.5	44.7	56.2	70.8	89.1	112	141	178
Center:	40	50	63	80	100	125	160	200
High:	44.7	56.2	70.8	89.1	112	141	178	224

Table 1: Lowest, center, and highest frequencies in Hz of the third-octave band used in the example.

3.2. Defining Geometry in NumCalc

To describe the geometry of the scattering objects, their surfaces need to be discretized using small triangles or plane quadrilaterals. To that end, users need to provide a list of the vertices of these elements, and a list which vertices make up which element. In this list, it is essential that the vertices of each element are ordered *counter-clockwise*. If users create the mesh using Blender, the order is correct if they ensure that the normal vector on each face points to the outside. If users are also interested in the sound field at (evaluation) points that do not lie on the surface of the objects, these points as well as an element list containing them need to be provided too. It may seem odd at first that evaluation points also need to have an element associated with them, but this has two reasons: First, the definition of the elements also contains the information about the type of the element, i.e., object element or evaluation element. Second, most of the time, users may want to display the results of their computations, and in this case, elements provide an easy way to interpolate and display results between nodal points.

For NumCalc this information needs to be provided using one or more text files containing the number and the coordinates of all the vertices/evaluation points, and one or more text files to describe the different elements, i.e. a list containing the number of its vertices and its type. Using multiple files to describe the geometry has the advantage that, e.g., BEM mesh and evaluation grid can be defined individually, and different combinations of object surface and evaluation grids can be used. If the export script provided by Mesh2HRTF is used in combination with one of the available evaluation grids, these text files are generated automatically.

The content of the files describing the nodes is as follows. The first line always needs to contain the number of nodes in the file, the following lines contain the node number and its x, y, and z coordinates. The files describing the elements have a similar format. The first line always needs to contain the number of elements in the file, the following lines are given by `element-number, node-number of each vertex, type, 0, groupnr`. It is possible to mix triangular and quadrilateral elements, NumCalc checks the length of a line to determine the form of the element. The `element-number` is later needed to assign a boundary condition to each element, the `type` of the element determines if the element is a BEM element on the surface of the scatterer (`type = 0`) or if it contains evaluation nodes outside or inside the scatterer (`type = 2`). Additionally, elements on object surfaces can be collected into different groups. However, as his grouping also affects the clustering of the FMM, users are advised to use only two groups. Group 0 for BEM elements, and Group 1 for evaluation elements. Again, if Mesh2HRTF is used to generate the input file, this is done automatically. In Lsts. 3 and 4, a few lines of an example node and element definition are given. The mesh contains 6602 vertices and evaluation points and 10464 elements. Element 1 is a triangle on the surface of the scatterer belonging to Group 0 , element 10464 is a quadrilateral evaluation element belonging to Group 1.

```
# number of nodes
6602
# node_number  x y z coordinate
1 -5.773503e-01 -5.773503e-01 -5.773503e-01
...
6602 3.000000e+00 0.000000e+00 1.200000e+00
```

Listing 3: First and last lines of an example node file.

```
# number of elements
10464
# elem_nr  vert1 vert2 vert3 (vert4) elemtyp 0 grpnumber
1 703 704 808 0 0 0
...
10464 6583 6584 6602 6601 2 0 1
```

Listing 4: First and last lines of an example element file.

3.3. Boundary conditions

Boundary conditions describe the acoustic behavior of the scattering objects. In NumCalc, three types of boundary conditions (`PRES`,`VELO`,`ADMI`) can be defined for each element. The general syntax for the boundary conditions is

```
ELEM e1 TO e2 bctype real(v0) curve1 imag(v0) curve2
```

This line means that all elements with element numbers in the interval [`e1`, `e2`] have the boundary condition `bctype` $\in$ {`PRES`,`VELO`, `ADMI`} with value `v0`. Using `PRES` the sound pressure can be fixed for the element, if the pressure is set to zero, the element is sound soft. Using `VELO` the particle velocity can be fixed for the given element, if the velocity is set to zero, the element is sound hard. With `ADMI` an admittance can be assigned to the element to, e.g, model sound absorbing materials. In general, elements can only have one type of boundary condition, however, it is possible, that elements can have a `VELO` as well as an `ADMI` boundary condition. If no condition is assigned to an element, it is automatically assumed to be sound hard.

As sound-absorbing behavior can be frequency dependent, `ADMI` conditions can be combined with a piecewise linear curve (see Sec. 3.1) that maps each frequency step to a scaling factor for the admittance.. In case of the `ADMI` condition `curve1` and `curve2` can contain the number of the curves used to describe the frequency-dependence of the real and imaginary part of the admittance. The frequency steps themselves have already been defined for the frequencies (see Sec. 3.1). If no frequency-dependence is needed, the number for the curve is set to -1. The boundary condition section is closed using the keyword `RETURN`.

3.3.1 Example

In the following, we assume an example with different boundary conditions to illustrate the different options in `NC.inp`. Elements 0 to 10 are sound soft, element 11 has a velocity of 1.0 m/s, elements 50 to 52 have a constant admittance of $0.1 + 10^{-4}\mathrm{i}$ $\mathrm{m}^3/(\mathrm{P}\cdot\mathrm{s})$ and element 65 has a frequency-dependent admittance that is defined by the frequency-defining curves 1 and 2. Per default, all other elements are sound hard.

```
BOUNDARY
ELEM 0 TO 10 PRES 0.0 -1 0.0 -1
ELEM 11 TO 11 VELO 1.0 -1 0.0 -1
#frequency independent admittance
ELEM 50 TO 52 ADMI 0.1 -1 1.0e-4 -1
#frequency-dependent admittance using curves 1 and 2
#the real part is scaled with 1.0, the imaginary part with 0.5
ELEM 65 TO 65 ADMI 100.0 1 10.0 2
# Definition of the boundary conditions is finished,
# denote this by the keyword 'RETURN'
RETURN
#
CURVES
# number_of_curves max_number_of_points_per_curve
2 4
# curve_number  number_of_nodes_defining_curve
1 4
0.0 0.0
1.0 1.0
3.0 4.0
11.0 8.0
# curve_number number_of_nodes_defining_curve
```

```
2 3
0.0 0.0
2.0 3.0
10.0 6.0
```

Listing 5: Definition of boundary conditions plus frequency-defining curves for the admittance.

For the admittance two curves are defined for real and imaginary part, respectively. Curve 1 consists of 3 linear segments with nodes (S_j, V_j) at (0.0, 0.0), (1.0,1.0), (3.0, 4.0) and (11.0,8.0). Curve 2 for the imaginary part consists of 2 linear segments with nodes (S_j, V_j) at (0.0,0.0), (2.0,3.0), and (10.0,6.0). Thus, different curves do not necessarily need to have the same number of nodes. The parameters for `Number_of_freqsteps`, `stepsize` and `Index0` are have already been defined in the frequency section Lst. 1. To provide an example we choose two arbitrary frequency steps ($n = 1$ and $n = 6$). Similar to Eq. (1) the real part of the admittance α_n is given by

$$\mathrm{Re}(\alpha_n) = \mathrm{Re}(v_0)\left(V_j + \frac{V_{j+1} - V_j}{S_{j+1} - S_j}(i_0 + n \cdot h_i - S_j)\right), \quad nh_i \in [S_j, S_{j+1}], \tag{2}$$

the imaginary part is defined in the same manner. The definition of the curves in Lst. 5 implies that at $f_1 = 1000\,\mathrm{Hz}$ (frequency-step $n = 1$) element 64 has an admittance of $\alpha_1 = 100.0 \cdot 1.0 + 10 \cdot \frac{3}{2}\mathrm{i} = 100 + 15\mathrm{i}$, at 6000 Hz ($n = 6$), for example, the element has an admittance of $\alpha_6 = 100.0\left(4 + \frac{8-4}{11-3}(6-3)\right) + 10.0\left(3 + \frac{6-3}{10-2}(6-2)\right)\mathrm{i} = 550 + 45\mathrm{i}$.

3.4. Sound Sources

External sound sources can be provided by a combination of point sources and plane waves. A point source $p_{\text{point}}(\mathbf{x}) := P_0 \frac{e^{\mathrm{i}k||\mathbf{x}-\mathbf{x}_0||}}{4\pi||\mathbf{x}-\mathbf{x}_0||}$ at a source point $\mathbf{x}_0$ is defined by its number, the coordinates of its origin `(X0,Y0,Z0)` and its source strength `P0`:

```
POINT
Nr X0 Y0 Z0 Real(P0) -1 Imag(P0) -1
```

A plane wave $p_{\text{planewave}}(\mathbf{x}) := P_0 e^{\mathrm{i}k\mathbf{d}\cdot\mathbf{x}}$, $||\mathbf{d}|| = 1$ with direction `(DX,DY,DZ)` and strength `P0` is given by

```
PLANE
Nr DX DY DZ Real(P0) -1 Imag(P0) -1
```

Additionally, it is also possible to define arbitrary elements of the scatterer as a "vibrating" source by setting the appropriate `VELO` boundary condition; in this case the source strength is defined by `VELO = v0`.

4 Benchmark example

In our opinion, it is important that users of NumCalc are aware of the accuracy they can expect from the numerical solutions with NumCalc. In the following we look at the benchmark example of a scattered plane wave from a sound-hard sphere with radius $r = 1\,\mathrm{m}$. For this example, the acoustic field on the sphere and outside the sphere is known cf. [15, Chapter 6.10.3] and can be compared with the numerical BEM solution. In Fig. 3, the sound field for two frequencies ($f_1 = 200\,\mathrm{Hz}$ and $f_2 = 1000\,\mathrm{Hz}$) with speed of sound $c = 340\,\mathrm{m/s}$ is shown. The sphere was discretized using 20480 triangular elements, the average edge length of the elements was about 0.037 m. The triangularization itself is based on an icosphere, thus the triangles are almost equilateral.

On a computer with an Intel Core i5-4430 processor and 8 GB RAM, the calculation for $f_1 = 200\,\mathrm{Hz}$ took about 40 s, where 15 s were used for calculating the field on 24874 points outside the sphere, the iterative solver was

finished after 26 iterations with an relative error of $2.3 \cdot 10^{-10}$. At $f_2 = 1000$ Hz the iterative solver was stopped after 17 iterations (rel error $5.8 \cdot 10^{-10}$), the total amount of calculation time was 68 s, where 41 s were used for calculating the sound field outside the sphere. In Fig. 4 the difference of the dB value of the sound pressure at

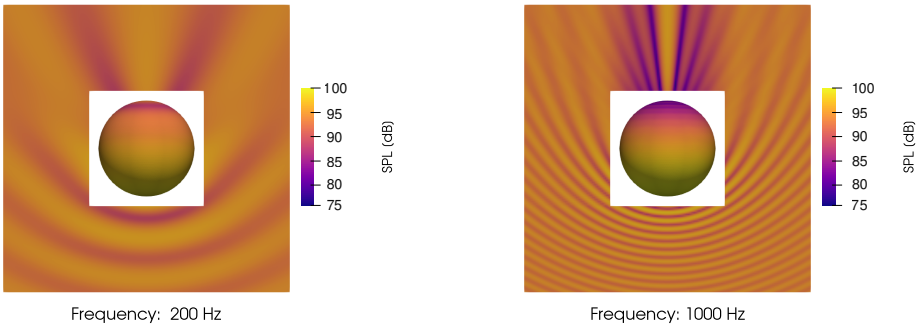


Figure 3: Acoustic field around a sound hard sphere at 200 Hz and 1000 Hz.

the sphere is depicted for two different types of discretizations at a frequency of $f_2 = 1000$ Hz. On the left and the middle panel the sphere was discretized using 5292 triangles that were constructed using a projection of a cube to the sphere, on the right panel the sphere was discretized with 5120 elements, where the triangularization was based on an icosphere. It is clearly visible that the regularity of the triangles has a lot of impact on the difference between BEM and analytic solution. In the case of the very regular icosphere triangularization a maximum difference of about 0.5dB could be observed on a few elements, the mean error over all elements of this discretization was about 0.114 dB. For the other triangularization the maximum difference was even in the range of 1.5 dB, the mean error was about 0.183 dB, and in contrast to the icosphere triangularization there was a clear distinction between "sunny" and "shadow" side, i.e. the side oriented towards the plane wave and the side oriented away from the plane wave.

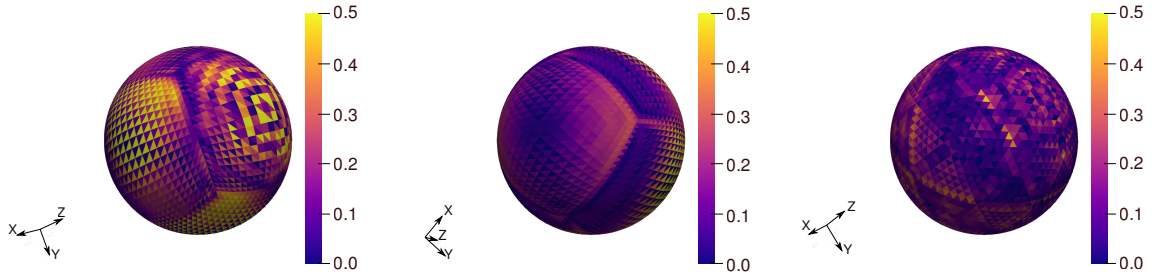


Figure 4: Difference between the dB sound pressures for the BEM calculation and the analytical solution at $f_2 = 1000$ Hz. Left: "Sunny"side, discretization based on a projected cube. Middle: "Shadow" side, discretization based on a projected cube. Right: Discretization based on an icosphere, "sunny" side.

Dependence on the number of elements

For most problems in practice, 10 elements per wavelength provide users with a solution that is sufficiently accurate. To illustrate the accuracy to be expected when using NumCalc, Fig.5 shows the mean relative error $\frac{|p-p_a|}{|p_a|}$ as a function of the number of triangular elements N. As a triangularization based on an icosphere is very restrictive in terms of number of elements, the discretization in this subsection is based on projecting a cube onto the sphere, which has the drawback, that the triangles are not as regular as for an icosphere. p is the

sound pressure calculated with NumCalc, p_a is the sound pressure calculated using an analytic expression, cf. [15, Chapter 6.10.3]. The error is given as the mean error over all collocation nodes on the sphere (continuous line) and as the mean sound pressure over evaluation points around the sphere (dashed line). These points are the same 24874 points that were used for depicting the field outside the sphere in Fig. 3. The dotted lines depict the function $g = \frac{1}{\sqrt{N}}$. From Fig. 5 it becomes clear that the simplicity of the collocation method with constant

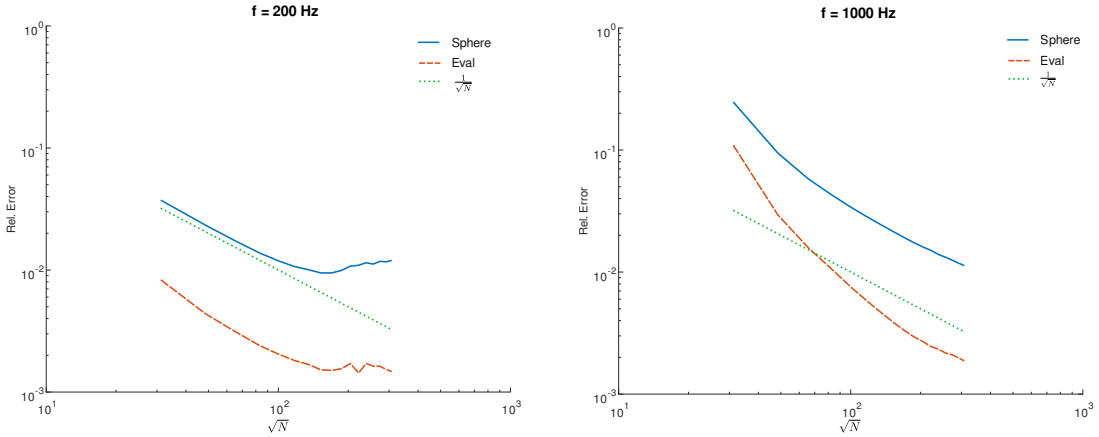


Figure 5: Mean relative error $\frac{|p-p_a|}{|p_a|}$ between the calculated sound pressure p and the analytic solution p_a on the surface of the sphere (continuous line) and on the evaluation grid (dashed line) as a function of the square root of the number of elements N. On the left side the frequency was set to $f_1 = 200\,\text{Hz}$, on the right side $f_2 = 1000\,\text{Hz}$. The average edge length of the elements used in the BEM are proportional to $\sqrt{N}$. The dotted line depicts the function $g = \frac{1}{\sqrt{N}}$.

elements has to be paid with a low convergence order of the error as a function of the number of elements. Also, the rise in the mean error in Fig. 5 for the case $f_1 = 200\,\text{Hz}$ shows the known fact, that the accuracy of the BEM solution suffers at low frequency when the mesh is too fine.

The rise in the error is also caused by the fact, that in the default setting, some numerical parameters, e.g. the truncation parameter of the multipole expansion or the accuracy of the quadrature methods for integrals over elements, are set rather small. This low setting on the one hand means, that the accuracy suffers, but on the other hand, the calculations become faster. Using a higher order in the FMM and a higher quadrature order for the integrals over the elements can stop this rise in the relative error.

5 Conclusion

In this paper, the open source software project Mesh2HRTF and the BEM code NumCalc have been presented briefly. Mesh2HRTF and NumCalc are aimed at providing a simple and easy-to-use method to solve acoustic wave propagation problems in 3D using the boundary element method. Input parameters have been described briefly and an example was given to illustrate the accuracy of the code. For in-depth tutorials and examples we refer to the homepage of Mesh2HRTF [16], where also the development version of the upcoming 1.0 release can be found, or the Wiki page for Mesh2HRTF [8]. There, also a list with the most common errors used in connection with Mesh2HRTF and NumCalc can be found.

Acknowledgements

We want to express our deepest thanks to Z.S. Chen, who was the original programmer behind NumCalc.

This work was supported by the European Union (EU) within the project SONICOM (grant number: 101017743, RIA action of Horizon 2020).

References

[1] C. H. Kasess, H. Waubke, M. Conter, C. Kirisits, R. Wehr, and H. Ziegelwanger. The effect of railway platforms and platform canopies on sound propagation. *Appl. Acoust.*, 151:137–152, 2019.

[2] W. Kreuzer, P. Majdak, and Z. Chen. Fast multipole boundary element method to calculate head-related transfer functions for a wide frequency range. *J. Acoust .Soc. Am.*, 126:1280–1290, 2009.

[3] N. H. Fletcher and T. D. Rossing. *The physics of musical instruments*. Springer, New York, 1991.

[4] Blender Online Community. *Blender - a 3D modelling and rendering package*. Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018. URL http://www.blender.org.

[5] P. Majdak, F. Zotter, F. Brinkmann, J. De Muynke, M. Mihocic, and M. Noisternig. Spatially Oriented Format for Acoustics 2.0: Introduction and Recent Advances. *J Audio Eng Soc*, accepted, 2022. URL https://projects.ari.oeaw.ac.at/research/Publications/Articles/2022/Majdak%202022%20SOFA%203A.pdf.

[6] Tutorials for Mesh2HRTF. https://sourceforge.net/p/mesh2hrtf/wiki/Tutorials, last visited 19.04.2022.

[7] ARI Mesh2HRTF Homepage. https://www.oeaw.ac.at/isf/das-institut/software/mesh2hrtf, last visited 19.04.2022.

[8] Wikipage for Mesh2HRTF. https://sourceforge.net/p/mesh2hrtf/wiki/Home, last visited 19.04.2022.

[9] Z.-S. Chen, H. Waubke, and W. Kreuzer. A formulation of the fast multipole boundary element method (FMBEM) for acoustic radiation and scattering from three-dimensional structures. *J. Comput. Acoust.*, 16 (2):303–320, 2008.

[10] S. Marburg. Six boundary elements per wavelength: Is that enough? *J. Comput. Acoust.*, 10:25–51, 2002.

[11] R. Coifman, V. Rokhlin, and S. Wandzura. The fast multipole method for the wave equation: a pedestrian prescription. *IEEE Antennas Propag. Mag.*, 35(3):7–12, 1993.

[12] Yousef Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, second edition, 2003.

[13] A. J. Burton and G. F. Miller. The application of integral equation methods to the solution of exterior boundary-value problems. *Proc. R. Soc. London Ser. A*, 323:201–210, 1971.

[14] H. A. Schenck. Improved integral formulation for acoustic radiation problems. *J. Acoust. Soc. Am.*, 44 (1):41–58, 1968.

[15] E.G. Williams. *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. Academic Press, London, 1999.

[16] Any2HRTF/Mesh2HRTF. https://github.com/Any2HRTF/Mesh2HRTF, last visited 21.04.2022.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

A Digital Twin for full-field monitoring in multi-channel control with applications to direct field acoustic testing

A.G. de Miguel[1,*], O. Atak[2], M. Alvarez Blanco[1]

[1]Siemens Industry Software NV, Research Park 1237, Interleuvenlaan 68, 3001 Leuven, Belgium.

[2]Siemens Digital Industries Software, Francis House, 112 Hills Rd, Cambridge CB2 1PH, United Kingdom.

[*]alberto.garcia_de_miguel@siemens.com

Abstract

Direct field acoustic excitation (DFAX) is an increasingly adopted method to conduct environmental acoustic tests of space hardware which avoids the need of large reverberant rooms. Such tests are conducted by exciting the payload via direct acoustic radiations generated by a set of high-powered loudspeakers. Based on previous research, a comprehensive Digital Twin of the electro-acoustic plant is utilized to simulate the frequency responses of all the sub-systems in the plant, including digital audio processing, loudspeaker performance, sound propagation and structural response. In this paper, a model update approach is proposed to tune the electro-acoustic contributions in the simulation based on experimental data obtained during the test. Such model can be subsequently linked to the multiple-input multiple-output (MIMO) control loop to complement the physical responses obtained from the field sensors with simulated data. Consequently, the resultant Digital Twin enables test engineers to obtain the system responses at locations where no physical sensors are available, thus improving the control test performance via full-field monitoring.

Keywords: Direct Field Acoustic eXcitation, Digital Twin, Environmental testing, Multi-Input-Multi-Output, Vibro-acoustic Simulation.

1 Introduction

Environmental acoustic tests are conducted to assess the survivability of payloads against harsh noise conditions. In the space industry, the performance of such tests is part of the qualification campaign for satellites and spacecrafts which is carried out to ensure that the high sound pressure levels (SPL) generated during launch do not compromise their structural integrity or the operation of their components. The traditional method to test against such conditions is the use of large reverberant chambers in which the payload is excited by diffuse sound fields [1]. Although this approach has set the standards for environmental acoustic tests, the space industry is looking for more efficient techniques to alleviate the need of transporting the payload to such facilities. In this context, direct field acoustic excitation (DFAX) is gaining attraction in the recent years as a more flexible and cost-efficient approach to perform acoustic qualification tests of spacecraft without the necessity of an ad-hoc facility [2].

In DFAX tests, the payload is excited by the direct sound waves generated by arrays of high-powered loudspeakers installed around it. The electro-acoustic plant can be built at any conventional room large enough to house the set-up, with overall sound pressure levels over 145 dB being reported [3]. In order to achieve similar acoustic fields as in reverberant chambers, advanced multiple-input multiple-output (MIMO) control systems are usually put in place to replicate the test references provided by the launcher authority, as shown by Larkin [4]. Indeed, the design of the test set-up and the selection of control parameters play a huge role in the ability of

the control system to represent efficiently the acoustic references, thus much care must be put to avoid costly, and risky, trial and error iterations. To mitigate such issues, different methods have been proposed in the past years to automatically define key parameters of the MIMO control system, such as the pre-test analysis proposed by Alvarez Blanco et al. [5].

Given all these challenges associated to the realization of efficient DFAX tests, different simulation approaches are being introduced to support test engineers in the design and dimensioning of the electro-acoustic plant, see for instance [6] or [7]. Depending on the detail of the (vibro-)acoustic model, engineers can get from a rough estimation of the structural responses in the payload subjected to the acoustic excitation, to a full simulation of all the system responses in the MIMO control test. In order to achieve the latter, a Digital Twin containing all the transfer functions of the system must be generated, i.e. including the responses of the electro-acoustic actuators as well as a proper description of the waves propagating, scattering and reflecting in between the loudspeaker stacks and off the payload, as described for instance in a previous work by de Miguel et al. [8].

Based on the models proposed in the latter work, in this research we generated a full Digital Twin of a small-scale DFAX set-up which has been built in the facilities of Siemens in Leuven (Belgium). The simulation model accounts for the dynamics of the electro-acoustic transducers through well-known lumped-parameter models [9], and the wave propagation in the test volume is numerically calculated via a state-of-the-art vibro-acoustic solver based on adaptive order finite elements, described in the work of Bériot et al. [10]. This paper proposes to reuse such model to estimate full-field responses in a physical DFAX test. By complementing the experimentally measured transfer functions with numerical ones coming from the Digital Twin, test engineers can visualize all the system responses, e.g. the pressure fields, in locations were no physical sensors are placed. Such technology has the potential change the way this type of MIMO control tests are performed, as much more information can be make available to evaluate the performance of the acoustic test. In addition, in order to increase the confidence on the model accuracy, a model update technique is introduced to reduce the error of the numerical model with respect to the measured acoustic responses on-site.

The paper is structure as follows: first, the small-scall DFAX plant used in the current research is presented in Section 2. Then, Section 3 introduces the Digital Twin of the physical set-up, including a brief description of the models employed. Section 4 aims at describing the approach followed to generate full-field solutions during the MIMO control test. The results are discussed in Section 5, focusing mainly on the correlation levels achieved by the Digital Twin. Finally, the conclusions are drawn in Section 6.

2 Electro-acoustic system

The proposed DFAX set-up is shown in Fig. 1. It consisted of a total of 45 small-size loudspeakers stacked in a cylindrical configuration of 9 columns. Such arrangement was introduced to mimic the typical layout of the electro-acoustic plants utilized in full-scale DFAX tests for qualification and acceptance of space components. The test set-up had a inner diameter of 66 cm and a maximum height of 76 cm at the top of the speaker stack. For each column, the loudspeakers were stacked on top of each other and attached to a metallic beam structure to ensure they stayed aligned in the vertical direction.

The loudspeaker model was the Control 1 Pro by JBL, a compact 2-way cabinet with a specified frequency range of 80 - 20000 Hz. This model's cabinet contains a 135mm diameter low frequency woofer and a 19 mm tweeter, and features a crossover frequency of 4.2 kHz. The input signal to the loudspeakers was delivered by an Audio Toolbox by Auvitran [11], which provides signal processing and power amplification.

The DFAX system was instrumented with 24 free-field array 1/4-inch microphones to acquire the acoustic responses inside the test volume. The microphones were placed randomly at different locations in between the loudspeaker stacks, hanging from a metallic grid which is placed on top of the beams that sustain the speakers, as depicted in Fig. 1. Such configuration was implemented to avoid introducing any stands in the test volume, as the microphones were hanging from their SMB cables directly.

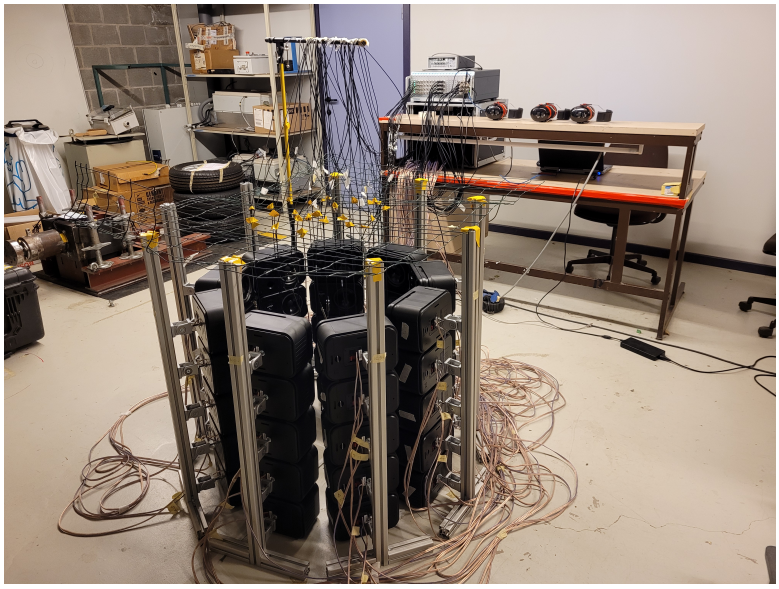


Figure 1: Picture of the small-scale DFAX setup.

In addition, a rectangular plate of dimensions 27.3 $\times$ 32.8 cm was installed at the center of the test volume with the aim of simulating the payload in real DFAX tests. The structure was made of a thin layer (approximately 1 mm) of Aluminum and was hanging with elastic cords from two large stands placed on opposite sides of the set-up.

The signal generation and data acquisition in the test was done with a Simcenter SCADAS Lab hardware [12] and the control algorithm was provided by the software Simcenter Testlab MIMO Random Control [13]. The test specification was introduced in the software in the form of a third-octave band spectrum profile of constant value with an overall SPL (OASPL) of 90 dB between 50 and 10000 Hz. 9 independent drive signals (one per column) were set to control the acoustic fields, which were characterized by the 24 microphones aforementioned. An automatic selection of 12 control sensors was performed based on the pre-test analysis [5], leaving another 12 microphones for monitoring purposes.

3 Digital Twin

The Digital Twin was generated in the simulation platform Simcenter 3D [14]. The aim of the model was to simulate as accurately as possible the system responses in the MIMO test, thus all the main components of the electro-acoustic plant were modelled, namely: signal processor, electro-dynamic transducers, the acoustic media in which the sound waves propagate and the plate structure. An illustration of the 3D vibro-acoustic model created in this work is included in Fig. 2.

In a DFAX configuration, the pressure responses measured at the field sensors are defined by the direct acoustic radiation from the loudspeakers as well as the reflections off the stacks and the structural item. Indeed, it has been observed that standing waves appear due to wave interference patterns in the test volume as well as the room geometry [4]. In order to capture such complex responses, a full 3D numerical model of the acoustic media was generated using state-of-the-art acoustic solvers. The acoustic domain covered the entire test set-up from the ground and it is discretized using the adaptive order finite element method, also known as FEMAO [10]. This method solves numerically the acoustic propagation problem described by the Helmholtz equation in a very efficient manner by adjusting locally the interpolation order of the finite elements for different frequencies and element sizes.

Free-field conditions were applied over the air external surfaces via the so-called Automatically Marched Layer

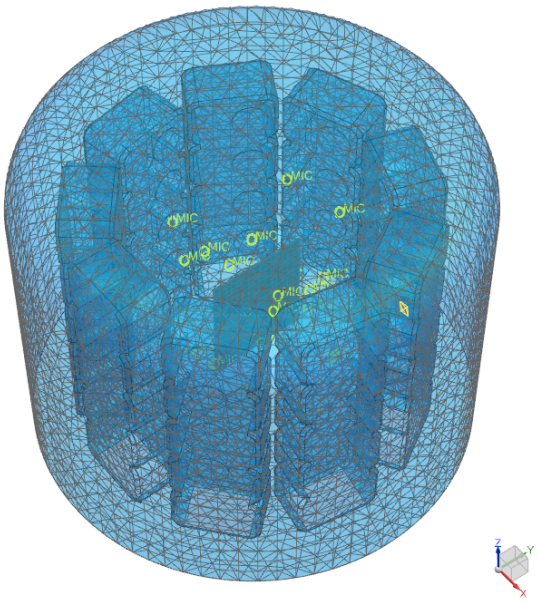


Figure 2: Illustration of the vibro-acoustic model of the DFAX plant generated in Simcenter 3D.

(AML), which efficiently absorbs outgoing waves in the convex boundary [15]. The ground was modelled as an acoustically rigid surface and the fluid-structure interaction with the structural component is represented with two-way coupling. The targeted frequency range was 50 - 1000 Hz, with a resolution of 3.125 Hz. The maximum number of degrees of freedom of the acoustic model were 113,334, and the computation lasted slightly less than 5 min in a laptop of 64 Gb of RAM using 4 parallel processes.

The radiation from the loudspeakers was introduced as acoustic velocity loads applied on the driver's faces. The response of the electro-acoustic transducers was estimated via lumped-parameter models, which provided an analytical representation of the actuator response based on the Thiele-Small parameters [16]. The level of correlation achieved by the loudspeaker model is depicted in Fig. 3, which shows a comparison of the velocity measured with a laser Doppler vibrometer (LDV) at the center of the membrane against the estimation of the lumped-parameter model. Finally, the voltage amplification functions were also measured and added to the electro-dynamic response in the model.

The resulting model can be used to perform virtual testing activities in the early stages of the test campaign, which can be extremely useful for test engineers to design of the acoustic set-up according to the target references, perform what-if scenarios and optimize various parameters of the MIMO test via numerical analyses. In this context, recent works [7, 17] have proposed the application of pre-test algorithms in the Digital Twin to simulate the behavior of the control system and optimize relevant parameters such as the selection of control sensors.

4 Full-field monitoring in MIMO control acoustic tests

Another use which can be given to the Digital Twin presented in the previous section is to augment the space of physical responses available in the test with virtual solutions. Indeed, the FRF matrix generated by the simulation has the same form as that generated by the control software from the non-parametric model of the

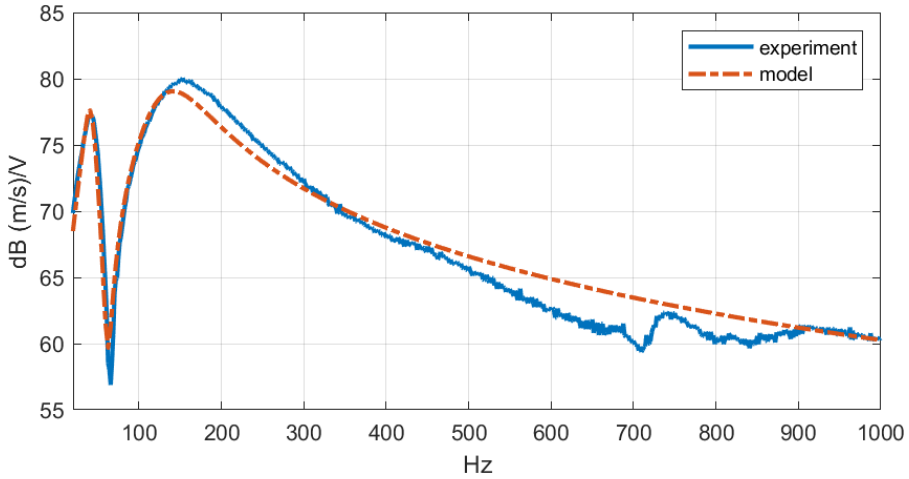


Figure 3: Acoustic velocity measured at the center of the woofer (blue line) and estimated by the lumped-parameter model (dashed red line).

system, i.e.

$$\boldsymbol{p} = \boldsymbol{H}\,\boldsymbol{v} \tag{1}$$

where $\boldsymbol{H} \in \mathbb{C}^{m\times d}$ is the transfer matrix between the drive signals $\boldsymbol{v} \in \mathbb{C}^{d}$ and the pressure responses $\boldsymbol{p} \in \mathbb{C}^{m}$. Note that all the operations hereinafter are performed per frequency line. By substituting the experimental transfer matrix with the one obtained from the simulation model, for any given set of values of the drive signals one can estimate the full-field responses as

$$\boldsymbol{p_s} = \boldsymbol{H_s}\,\boldsymbol{v} \tag{2}$$

where $\boldsymbol{H_s} \in \mathbb{C}^{n\times d}$ contains the simulated FRF at n virtual senors and $\boldsymbol{p_s} \in \mathbb{C}^{n}$ are the corresponding complex pressures. Note that in the model, one can introduce as many virtual sensors as desired, see Fig. 4, thus enabling the computation of full-field responses during the test.

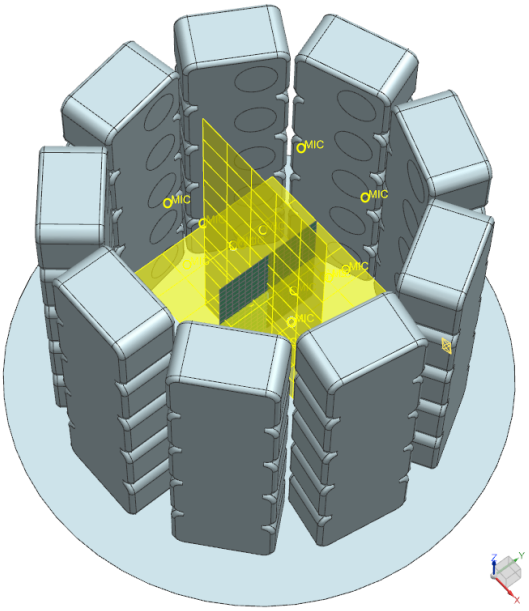


Figure 4: 3D model of the DFAX set-up including the modelled microphone sensors in yellow.

In closed-loop MIMO random tests, the role of the controller is to compute the voltage signal values that minimize the error between the test targets and the measured acoustic responses. The estimated drive signals are in a first step provided in the form of a spectral density matrix (SDM), $\boldsymbol{S_{vv}} \in C^{d \times d}$, which contains the power spectral densities (PSD) and cross-power spectral densities (CSD). Once available, this matrix can be retrieved and used to estimate the responses at the virtual sensors as follows

$$\boldsymbol{S_{pp}} = \boldsymbol{H_s}\, \boldsymbol{S_{vv}}\, \boldsymbol{H_s^H}, \tag{3}$$

where $\boldsymbol{S_{pp}} \in C^{n \times n}$ contains the spectral terms of the virtual responses. Such approach allows test engineers to visualize the pressure values achieved at the test in locations where no physical sensors are placed. In DFAX tests, such capabilities can be used to get a complete view of the properties of the acoustic fields generated by the control system at any desired location of the test volume.

4.1. FRF update

Due to various uncertainties and assumptions in the model, one should expect a certain level of error when correlating the simulated FRFs with those measured at the physical plant. If this error is too high, the confidence in the virtual solutions computed via Eq. 3 would be reduced and the value of the full-field monitoring to extract quantitative metrics might be heavily compromised. In this paper, we propose a simple method to improve the correlation of the model outputs with respect to the experimentally identified FRFs of the system. The underlying idea is to adjust the contribution of the boundary conditions, which in the current modeling framework correspond to the electro-acoustic functions described in Section 3, based on a comparison of the simulated and measured functions at a subset of observation microphones, o, for which there is geometrical correspondence. Mathematically, such operation can be written as

$$\boldsymbol{H^o} = \boldsymbol{H_s^o}\, \boldsymbol{E} \tag{4}$$

where $\boldsymbol{H^o}$ and $\boldsymbol{H_s^o} \in \mathbb{C}^{o \times d}$ are the experimental and simulated transfer matrices, respectively, at the observation sensors, and $\boldsymbol{E} \in \mathbb{C}^{d \times d}$ is a complex matrix containing the correction factors of the electro-acoustic loads. The values of $\boldsymbol{E}$ can be computed in different manners based on the model update choice, but, in general, one can assume that the number of observation sensors is greater than the columns of $\boldsymbol{H_s^o}$, i.e. the number of drive signals, leading to an over-determined system of linear equations which can be solved via least-squares. Once $\boldsymbol{E}$ is computed, the full updated matrix of FRFs can be recovered as

$$\boldsymbol{H_{upd}} = \boldsymbol{H_s}\, \boldsymbol{E} \tag{5}$$

The performance of the model update can be checked by evaluating the correlation between the responses obtained at the remaining physical sensors, $k = m - o$, and their virtual twins. Such checks provide the test engineers with very important information on how well the improved model is representing the physical system in locations which are not utilized during the model update process.

5 Results

The results included in the present section follow two purposes: first, the validation of the numerical models via correlation analysis with the measured FRF values; and secondly, to show the visualization of virtual data at measured and unmeasured locations in the MIMO random control test.

5.1. FRF correlation

The experimental transfer functions were obtained from a low-level open-loop test using 9 uncorrelated pseudo random signals. Figure 5 shows the FRF values [Pa/V] in dB obtained at one of the microphone sensors from:

the test (solid blue line), the original model described in Section 3 (dashed red line), and the updated model based on the method described in Section 4 (dash-dotted yellow line). The selected sensor, microphone 24, is located at the point 15.3 cm $\times$ 10.65 cm $\times$ 55 cm with respect to a coordinate system placed on the ground at the center of the set-up, and the FRF is calculated with respect to the drive signal number 5 (fifth column of the $\boldsymbol{H}$ matrices).

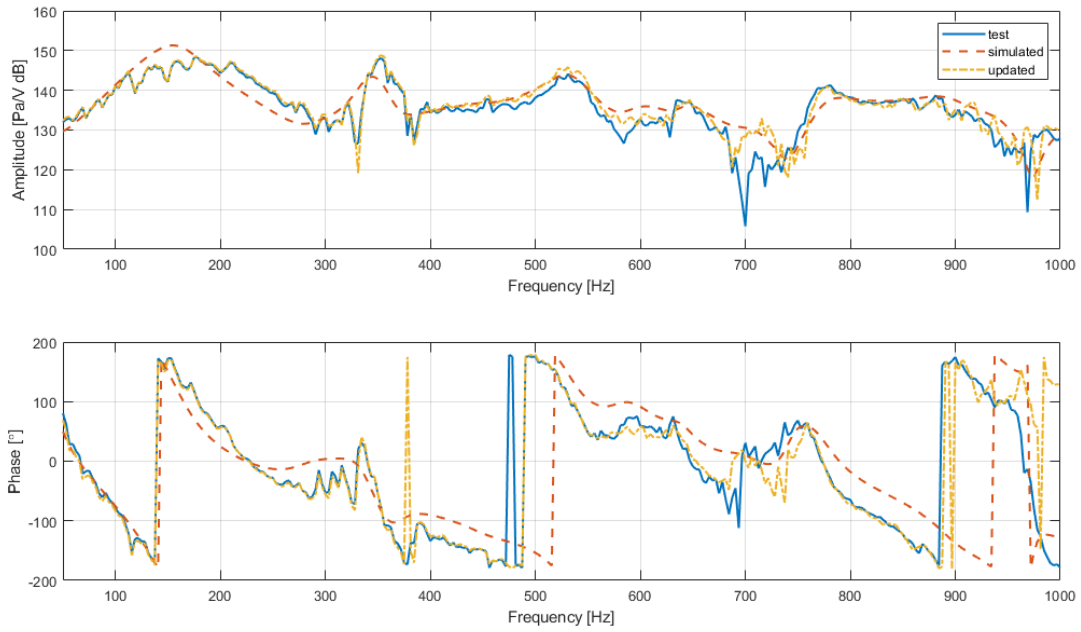


Figure 5: Amplitude and phase values of the FRF at microphone 24 obtained at the test and from the simulated and updated models.

Regarding the original model, one can observe that the direct frequency simulation estimates fairly well the transfer function identified in the test, both in amplitude (top graph) and phase (bottom graph). As expected in this type of acoustic measurements, the experimental curves are noisier than that provided by the model, which does not simulate certain effects such as reflections from the room and surrounding objects. However, it is clear that the main dynamics of the system are well captured. Also, a small delay is noticed in the phase response, which is probably due to a latency in the electro-acoustic system not properly represented in the lumped-parameter model.

The updated FRFs are computed using a total of 15 microphones as observation sensors (microphones 1 to 15), thus leaving 9 sensors for validation purposes. It is noted that the selection of the observation sensors for the model update is independent of the choice of control/monitoring microphones for the MIMO test. The results shown in Fig. 5 correspond to one of the latter set. It is clear from the plot that the FRF values are considerably improved over the entire frequency range considered (50 Hz - 1000 Hz), especially at low frequencies where the values are almost on top of the experimental ones. Another aspect to note is that the phase delay observed in the original model is effectively corrected. Figure 6 includes the correlation with respect to the experimental values based on the root mean square error (RMSE) for all 24 microphones. As expected, the error is reduced at the observation sensors when the FRF update is performed. Also, the correlation in all validation microphones (microphones 16 to 24) improves considerably, although slightly less than in the observation sensors which are used to update the model. These results indicate that there is a good level of confidence in the virtual responses provided by the updated Digital Twin at any other point within the test volume.

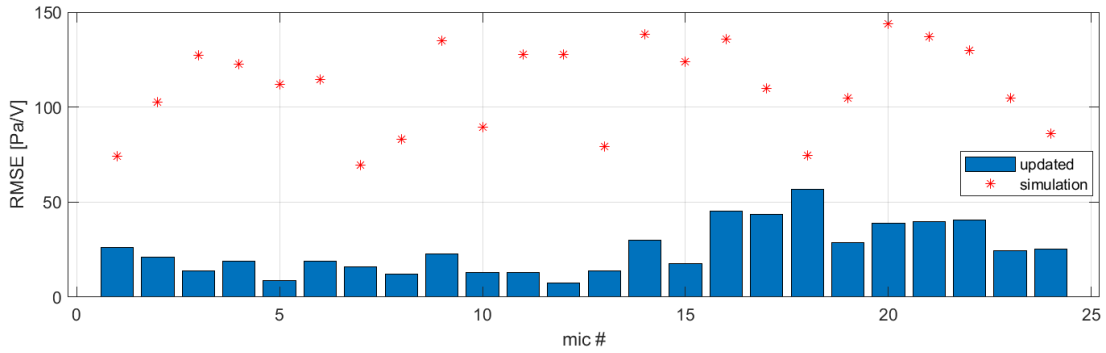


Figure 6: Comparison of the correlation error (RMSE) with respect the test references at all 24 microphones for one of the drive signals.

5.2. Full-field monitoring

The MIMO control test was performed using the SPL profile depicted in Section 2, a flat third-octave band spectrum of 90 dB OASPL between 50 and 10000 Hz, as reference at all the control sensors. Once the SDM of the drive signals were computed by the controller, the matrix of simulated FRFs can be used to obtain the acoustic responses at all virtual channels via Eq. 3. Figure 7 shows a plot of the PSD data obtained in Simcenter Testlab at one of the control sensors (microphone 20), which is used in this exercise for validation of the model update. The curves correspond to the targeted pressure values (green line), the measured response (red line) and the virtual response at the same location (blue line). As before, the agreement between physical and simulated solutions is very good, in particular at low frequencies, which gives us further confidence in the full-field data.

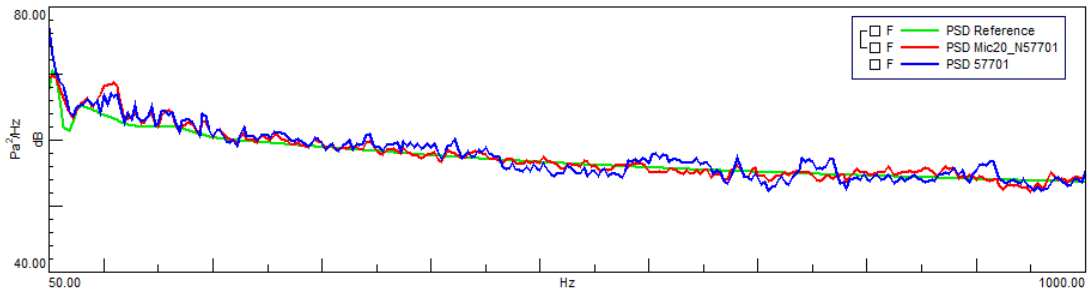


Figure 7: PSD values measured (label Mic20_N57701) and predicted (label 57701) at one of the control sensors during the MIMO control test. The acoustic reference is also included.

Other than the virtual measurements at the observation and validation microphones, the main goal of the proposed Digital Twin is to provide an estimation of the system responses in unmeasured locations, anywhere in the test volume were there are no physical sensors in place. As an example, Fig. 8 illustrates these full-field monitoring capabilities of the proposed approach for MIMO control tests. The image shows the contour plots of the PSD curve at 50 Hz over the microphone planes included in the model of Fig. 4. These virtual solutions provide the test engineers with highly useful information regarding for instance the capabilities of the control system to achieve the acoustic references, as well as the eventual presence of high pressure areas or reflection patterns which might be uncontrolled. Corrective actions to improve the environmental acoustic test campaign might be taken onsite based on the extra information provided by the Digital Twin.

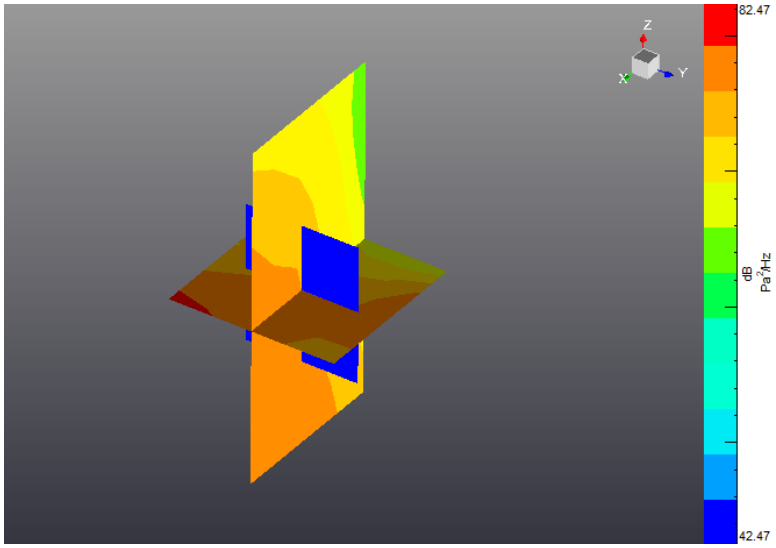


Figure 8: Visualization in Simcenter Testlab of full-field PSD solutions in two planes intersecting the plate used as test specimen.

6 Conclusions

The present work introduces a comprehensive Digital Twin of a small-scale DFAX for the estimation of full-field responses in MIMO control tests. The electro-acoustic plant under study consisted of a reduced size DFAX set-up with 45 small loudspeakers pointed at the center of the cylindrical configuration and driven by 9 independent signals routed to each speaker column. A full simulation of the MIMO set-up was created by modelling all the relevant transfer functions of the control system in between the signal generation and the sensor measurements. Such model is then used during the test to evaluate the acoustic fields in locations where no physical microphones are available.

The possibility of visualizing full-field data on-site has the potential to change the current procedure to conduct environmental tests as far more information is available through the Digital Twin for engineers to evaluate the performance and efficiency of the control strategy. In order to increase the confidence in the simulation, a simple model update procedure is introduced to improve on-site the correlation of the numerically computed FRFs, once the system transfer functions have been identified in the test. The proposed method, which is based on the correction of the electro-acoustic loadings in the model, has proven to be very effective in reducing the error of the virtual estimations for the problem under study. Future works will be focused in validating the method in more realistic environmental acoustic tests. Furthermore, on-going research is dedicated to the implementation of more comprehensive model update methods which go beyond the limiting assumptions of the current approach, for instance based on machine learning techniques or Kalman filtering methods.

Acknowledgements

The authors gratefully acknowledge the Belgian Federal Science Policy Office (BELSPO) and the European Space Agency (ESA) for their support through the General Support Technology Programme (No. 4000127640/19/NL/GLC/vr).

The authors would like to also acknowledge the contribution of KU Leuven University in providing part of the electro-acoustic set-up used in the test.

References

[1] R. K. Cook, R. V. Waterhouse, R. D. Berendt, S. Edelman, and M. C. Thompson. Measurement of correlation coefficients in reverberant sound fields. *The Journal of the Acoustical Society of America*, 27(6): 1072–1077, 1955. doi: 10.1121/1.1908122. URL https://doi.org/10.1121/1.1908122.

[2] NASA-HDBK-7010 Direct Field Acoustic Testing (DFAT). Technical report, NASA, 2016.

[3] C. Fabries, B. Brevart, A. Carrella, M. Alvarez Blanco, D. Dal Fitto, and S. Charfenberg. Experimental validation of direct field acoustic testing. In *Proceedings of 14th European Conference on Spacecraft Structures, Materials and Environmental Testing*, 2016.

[4] P.A. Larkin. Developments in direct-field acoustic testing. *Sound & Vibration*, 48:6–10, 2014.

[5] M. Alvarez Blanco, P. Van Vlierberghe, M. Rossetti, K. Janssens, B. Peeters, and W. Desmet. Pre-test analysis to reproduce random pressure fields with multi-channel acoustic control. *Mechanical Systems and Signal Processing*, 163:108103, 2022. ISSN 0888-3270. doi: https://doi.org/10.1016/j.ymssp.2021.108103. URL https://www.sciencedirect.com/science/article/pii/S0888327021004878.

[6] V. Cotoni, B. Gardner, and A. Kolaini. Numerical simulation of pressure field in a direct field acoustic test setup. In *Proceedings of 2012 Spacecraft and Launch Vehicle Workshop, El Segundo, CA*, 2012.

[7] M. Alvarez Blanco, E. Matas, H. Bériot, B. Peeters, and W. Desmet. Frequency dependent selection of control sensors in multi-channel acoustic control. *CEAS Space Journal*, 13:119 – 131, 2021. ISSN 1868-2510. doi: https://doi.org/10.1007/s12567-020-00329-1.

[8] A.G. de Miguel, M. Alvarez Blanco, E. Matas, H. Bériot, J. Cuenca, I.C.S. Ngan, and B. Peeters. Numerical pre-test analysis for multi-channel control strategies in environmental acoustic tests. In *Proceedings of 16th European Conference on Spacecraft Structures, Materials and Environmental Testing*, 2021.

[9] L. L. Beranek and T. J. Mellow. *Acoustics: Sound Fields and Transducers*. Oxford: Academic Press, 2012.

[10] H. Bériot, A. Prinn, and G. Gabard. Efficient implementation of high-order finite elements for helmholtz problems. *International Journal for Numerical Methods in Engineering*, 106(3):213–240, 2016. doi: https://doi.org/10.1002/nme.5172. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.5172.

[11] Auvitran. Auvitran Audio Toolbox. https://www.auvitran.com/audio-toolbox-platforms/.

[12] Siemens Digital Industries Software. Simcenter SCADAS Lab. https://www.plm.automation.siemens.com/global/en/products/simcenter/scadas-lab.html, .

[13] Siemens Digital Industries Software. Simcenter Testlab. https://www.plm.automation.siemens.com/global/en/products/simulation-test/vibration-control-environmental-testing.html, .

[14] Siemens Digital Industries Software. Simcenter 3D. https://www.plm.automation.siemens.com/global/en/products/simcenter/simcenter-3d.html, .

[15] H. Bériot and A. Modave. An automatic perfectly matched layer for acoustic finite element simulations in convex domains of general shape. *International Journal for Numerical Methods in Engineering*, 122 (5):1239–1261, 2021. doi: https://doi.org/10.1002/nme.6560. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.6560.

[16] N. Thiele. Loudspeakers in vented boxes: part 1. *Journal of the Audio Engineering Society*, 19(5):382–392, may 1971.

[17] A.G. de Miguel, M. Alvarez Blanco, E. Matas, H. Bériot, J. Cuenca, O. Atak, K. Janssesn, and B. Peeters. Virtual pre-test analysis for optimization of multi-channel control strategies in direct field acoustic testing. 2022. Submitted.

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Spatial Transformations Effect to Soundscape: The Case of Istanbul Land Walls

Zeynep Sena AKDEMİR[1,*], Esin Özlem AKTUGLU AKTAN[2]

[1]Department of Urban and Regional Planning, Architectural Faculty, Yildiz Technical University, Istanbul, Turkey.

[2] Department of Urban and Regional Planning, Architectural Faculty, Yildiz Technical University, Istanbul, Turkey.

*zeynepsenaakdemir@gmail.com

Abstract

Sound atmosphere gives cues about environmental perceptions and memories. Istanbul's land walls play an essential role in Istanbul's historical, cultural, sensory, and spatial memory. With the establishment of the Ottoman Empire, land walls began to lose their defensive function and became a part of civil life.
The transformation of the land wall has been detached from the positive perception and cultural context of open green spaces. This perception replaced them with unplanned and uncontrolled spatial transformation. As a result of highways and uncontrolled constructions, the sound atmosphere of the district around the Istanbul land wall was adversely affected. Especially in the region between Yedikule and Topkapı, it intensified the population and transformed the demographic structure.
In this article, the contemporary sonic perceptual environment around the Istanbul land walls is studied. This research aims to contribute sonic perception and experience of the people in the field. The effect of the determined spatial transformation on the sonic landscape comfort in the adult user group is investigated. While defining sonic landscape comfort and perception, the standard methodologies described in ISO/TS 12913-2 were used. Methods of surveys, on-site sound recordings, and listening to volunteers under special conditions were used. Many types of data gathering methods were used. The obtained data were separated with the help of analysis programs. The spatial acoustic comfort was obtained; hence the outcome was evaluated by classifying the survey data. The result was evaluated based on the identified sound source, affective quality, perception and overall sound environment.

Keywords: Soundscape, Istanbul Landwalls, Space Perception, Outdoor Perception

1 Introduction

Individuals' relationships with the environment are shaped by their impressions, experiences, and perceptions. People's environmental perceptions and experiences in the outdoors are shaped by their hearing senses. The concept of soundscape is used to identify and systematize the auditory perceptual relationship that individuals establish with the environment. The soundscape is a concept that has been investigated and discussed by many disciplines from the past to the present[1]. Sound influences the livability of people through disciplines such as psychology, engineering, and architecture.

Schafer[2] developed the earliest studies about the soundscape. He categorized the soundscape into three categories: natural sounds, human sounds and technological sounds. According to Zhang and Kang's model, four essential components determine the soundscape of the outdoors; the sound source, space, person and environmental factors[3]. Characteristics define the sound source; it is determined by its location, sound pressure level, social or psychological features and whether the source is mobile or static. Humans are considered as both the source and the receiver of sound. Temporary sound source factors include the time of day, season and procedure[4].

The soundscape studies aim to establish how the auditory landscape is perceived and generate soundscape data suitable for the spaces by generating varied data for desired or undesirable soundscape design. Nowadays, soundscape studies focus on creating artificial or unique themes, identifying symbolic sounds specific to urban spaces, the documentation of symbolic sounds and the classification of sounds [5,6,7]. Many methods are used to determine, register and archive the quantitative and qualitative data of the soundscape in the outdoor environment. Binaural audio recordings, qualitative and quantitative sound tests, soundwalks, laboratory research, behavioral observations and simulative sound investigations supported by virtual technologies are the various methodologies of soundscape [7,8,9]. Many investigations are handled on the perceptual terms about the sonic environment to specify the environmental perception [10,11,12]. These terms are defined by their emotional and sensory connotations [12,13].

Istanbul's land walls and surroundings have a significant role in the city's historical, sensory, and cultural memories from the past to the present. The land walls were built 7.5 kilometers long in the reign of II. Theodosius. The purpose of constructing walls was to protect the territorial integrity of the city, determining the borders and administration [14, 15]. The walls lost their defensive function and became a part of civilian life with the establishment of the Ottoman State. Over time, gardens, farming areas, new neighborhoods and ports were established in Ottoman times. This transformation, which was realized with the construction of 10th street in the 1980s, accelerated the transformation in the surrounding area. This transformation has mostly resulted in the separation of open green spaces from their positive perception and cultural context. Unplanned and uncontrolled spatial transformations took place in place of the destroyed open green spaces [16]. In recent years, the highway and construction activities developed in this region have transformed the functions of outdoor green spaces by affecting the soundscape. This region, which has become an undefined urban rift [17], which contains many functions out of context within the land walls and its surroundings, has been transformed by many large-scale projects in the last decade. The increasing presence of highways has caused the destruction of natural and cultural landscapes in the region, and the transformation of the soundscape.
Although people have historical knowledge of the area, few studies have been done on its sensory heritage. From past to present, urban gardens, cemeteries and historical neighborhoods play an essential role in sensory memories of the city. This area was chosen for two main reasons. The first reason is that the sensory studies developed on the region are insufficient, and the second reason is that the region is transforming rapidly. The main reason for research is contributing to livability and sensory awareness of research fields. This study focuses on existing sonic data. Based on existing sound references, perceptions of public green spaces in this region were examined.
The paper focused on the region between Yedikule and Topkapı. The used methodologies are binaural sound recordings, classifying them, questionnaires and analysing. This study has been discussed through binaural records and tests by contributing to the auditory comfort and livability of the region.

2 Methodologies

In this paper, a field study was conducted in the Land wall region of Istanbul. The methodology consists of three main steps. The first step is to record sound from field areas public green spaces. The second step is to do a questionnaire with 37 participants. The third step is classifying and analysing test outputs.

2.1 Voice Records

In the scope of the survey, the Zoom H6 voice recorder was used for binaural recordings, and 13 different cases were recorded. Every single record has been five minutes long and recorded from one point of selected locations. At the same time, weather conditions are considered in the recording process.
The Brüel Kjaer type 2236 -sound pressure level (SPL) meter simultaneously measured the average SPL dB. Out of the 13 recordings, eight were selected according to their quality and variety of sound sources. The audio recordings were converted into two minute recordings with the Adobe Audition program.

2.2 Questionnaire

Eight recordings were listened to by 37 subjects with special headphones. During the listening test, three conditions were applied: silence conditions for the environment, each subject should not have hearing loss and every person should be over 18. A sound source, definitive sound, voice level, disturbing voice level and sound perception are asked in questions. Sound source and definitive sound were asked with the same metrics to subjects **(**Table 1). The list that occurred depends on Schafer's theory[2](Table 1). Sound volume and disturbing level asked to subject with metrics from 1 to 5. Subjects' perceptual conditions are specified with thirty different psychoacoustic metrics [10,13]. According to their perception, every person listens to eight different records and replies to metrics from 1 to 5. Overall, subjects' response to definition, sound level, disturbing level and perception tests for every record. The whole process takes 40 minutes for every subject.

Table 1: Definitive and Sound Source List (The list created according to the Schafer [2] criteria)

Nature Sound	Human Sound	Technological Sound
Dog Sound	Child Sound	Klaxon Sound
Cat Sound	Azan Sound	Train-Subway Sound
Bird Sound	Speech Sound	Ticket Sound
Leaf Sound	Other Human Sound	Bus Sound
Water Sound		Motorcycle Sound
Wind Sound		Construction Sound
		Music Sound
		Swing Sound
		Plane Sound
		Traffic Sound
		Other

2.3 Analysing

As a result of the survey, the spatial acoustic comfort survey data obtained were classified and seperated in the Excel program. The separated data were analysed with average numbers comparatively by Excel. The result of the analysis was evaluated with the positive, negative and spatial aspects of the perception of the soundscape in and around the Istanbul Land Walls.

2.4 Case study site

The line between Yedikule to Topkapı is focused on the scope of research. Overall, Land walls (Figure 1; a) include 12 gates and 7.5 km. In this study, the research work on the environment of six historical gates which are Topkapı, Mevlanakapı, Silivrikapı, Belgradkapı, Yedikule, and Mermerkapı. Yellow lines show the border of the site study (Figure 1; b). In the detailed study, the site study focuses on selected 13 different points, which are red and orange (Figure c-d).

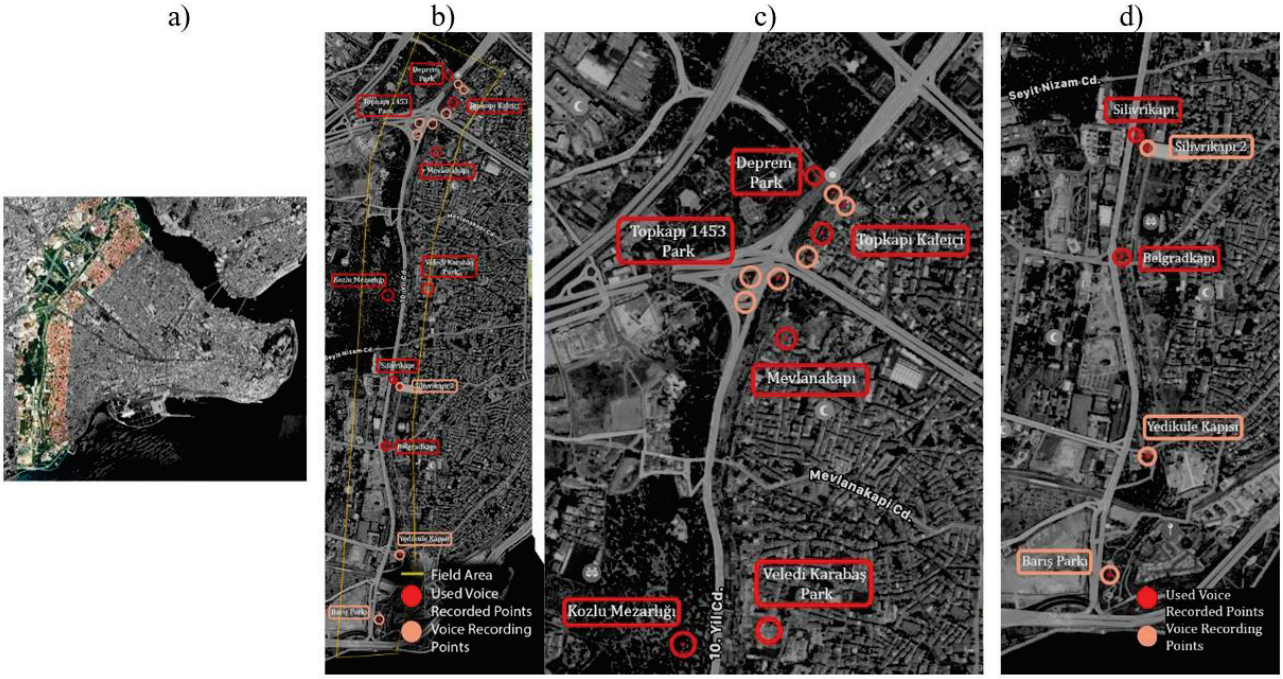


Figure 1: a) The location of Land walls in Historical Peninsula **b)** Field Area General Map **c)** Area's Map between Vatan Street to Silivrikapı **d)** Field map's Silivrikapı to Altın Kapı

The criteria for site selection were based on these two components. Firstly, the function variety happens in these environments, such as existing urban gardens, cemeteries, public parks and roads. The second reason is the accessibility of that environment.

The scope of the research site area includes 7 neighborhoods and 31 urban gardens. These areas' owners are mostly municipalities, governments and foundations; for this reason, there is not any limitation about use by pedestrians and citizens. Voice recordings were made in 13 different places and these places belong to the municipality and government. When choosing these regions, it was preferred that they be areas where pedestrian access could be provided continuously along the line of the Land Walls and have public functions. (Figure 2). After the obtained sound data, the spaces were separated. The functions of the areas decided to focus on are monumental, urban park, and railroad (Table 2).

Table 2: Functions of Recording Area's

Location	Function
Eski Kozlu Cemetery	Cemetery, Roads, Urban Park
Silivrikapı	Bus Station, Urban Garden, Pedestrian way, University, Roads
Belgradkapı	Urban Garden, Traffic Lights, Roads
Mevlanakapı Park	Urban Park, Cultural Centre
Topkapı Park	Urban Park, Café, Library
Deprem Park	İnterchange way of transportations, Railroad, Urban Park
1453 Park	Museum, Urban Park
Karabaş Park	Neighbourhood Park, Cafe

3 Findings

The obtained data shows that five of the eight regions are parks and three are used as public spaces. In the light of the data, the sound sources, descriptive sounds, sound levels, sound comfort level and people's perceptions of these sounds were calculated with the determined adjective pairs. In addition, function, perception and sound comfort data were evaluated by examining 30 adjective pairs. Twenty of the thirty-seven participants were female; seventeen participants were male. Twelve of them are between the ages of 18-25, twenty-four of them are between the ages of 26-35 and two of them are between the ages of 36-45. The average age of participants is 27.

3.1 Subjective Data

The obtained data illustrates the existing vehicle, traffic, and klaxon sounds, mostly considered in the category of technological sounds in the recordings. In all venues and recordings, these sounds were evaluated in "annoying" and "extremely annoying". Speech and child sound are selected as existing and descriptive categories in many places (Table 3). These sound level of annoyance varies according to the category, location and sound recordings. Although sound recording places have three different functions, the vehicle sounds and discomfort levels do not differ significantly and influence the subjects in a negative way.

Table 3: Eski Kozlu Cemetery, Silivrikapı, Belgradkapı, Mevlanakapı Park, Topkapı Park, Deprem Park 1453 Park and Karabaş Park's definitive and existing voices average outputs depends on Questionnaire

Location	Data of Questionnaire
Eski Kozlu Cemetery	Human sound, traffic and klaxon sound selected as existing sound. Klaxon sound selected as extremely annoying.
Silivrikapı	Traffic and human sound selected as existing sounds. Traffic and klaxon sound preffered as definable sound of that environment at the some time these vocices found extremely annoying.
Belgradkapı	Speech, Traffic, Klaxon, bus, motorcycle sounds are existing sounds. Klaxon and traffic sounds are the defining sounds of that environment. Traffik and Klaxon sound prefered extremelly annoying and annoying with 95 percent.
Mevlanakapı Park	Speech sound, child sound, bird sound and wind sound selected as existing sound. Human sounds are the definable sounds of that place. Human sound and bird sounds selected as pleasant.
Topkapı Park	Wind sound, human sound and machine sounds are the existing sounds of place. The definitive sound is wind sound. Wind sound and human sounds are preffered as pleasant and unsure sounds.
Deprem Park	The definable sound is traffic and motorcyle sound. Motorcycle, subway, bird, speech, construction and music sounds are the existing sound of place. Traffic, klaxon, train, motorcycle and construction sound selected as annoying and extremely annoying.
1453 Park	Mostly animal sound, wild, child are existed sound and selected as pleasant.
Karabaş Park	Bird, wind, speech, child, traffic and klaxon sound are existed sound. Child and traffic sound are definable sound of Park. All selected sound mostly preferred as pleasant or highly pleasant except traffic and klaxon sound.

Surveys on the study area's sound levels and disturbance levels revealed that Silivrikapı and Belgradkapı regions are significantly different from other regions. It has been concluded that Karabaş Park is louder than other parks in the park region, but the disturbance level of this noise level varies more than in other regions.

Table 4: Average Sound Levels of Locations according to subjects' perception

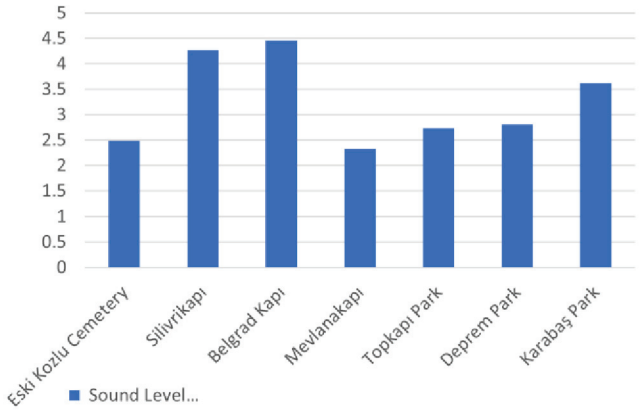


Table 5: Recorded Points dB according to SPL meter data

Location	Eski Kozlu Cemetery	Silivrikapı	Belgradkapı	Mevlanakapı Park	Topkapı Park	Deprem Park	1453 Park	Karabaş Park
Decibel	52	73	73	59	61	58	56.8	69

When selecting perception metrics, academic outputs checked and selected used with 30 metrics [7,10,12]. While metrics were classified and evaluated, they were evaluated into two basic classes. It is classified into two assessment pairs: positive-negative and neutral evaluation pairs. Neutral metrics include common-strange, heavy-light, sharp-dull, far-nearby, lively-calming, not sharp-sharp, unclear-distinct, rough-smooth metrics. The results were gained from the average of the data. Overall, common metrics are chosen in every location at neutral metrics (Table 6). Karabaş Park, Belgradkapı and Silivrikapı regions are selected by lively and sharp.

Table 6: The result of the Questionnaire Neutral metric test from Recorded Points

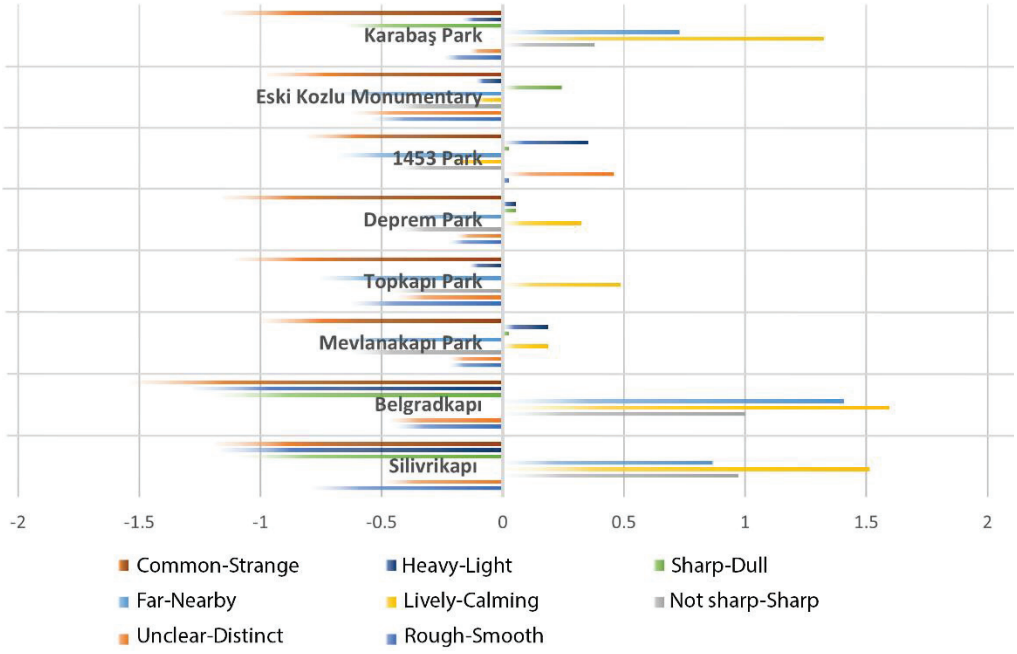
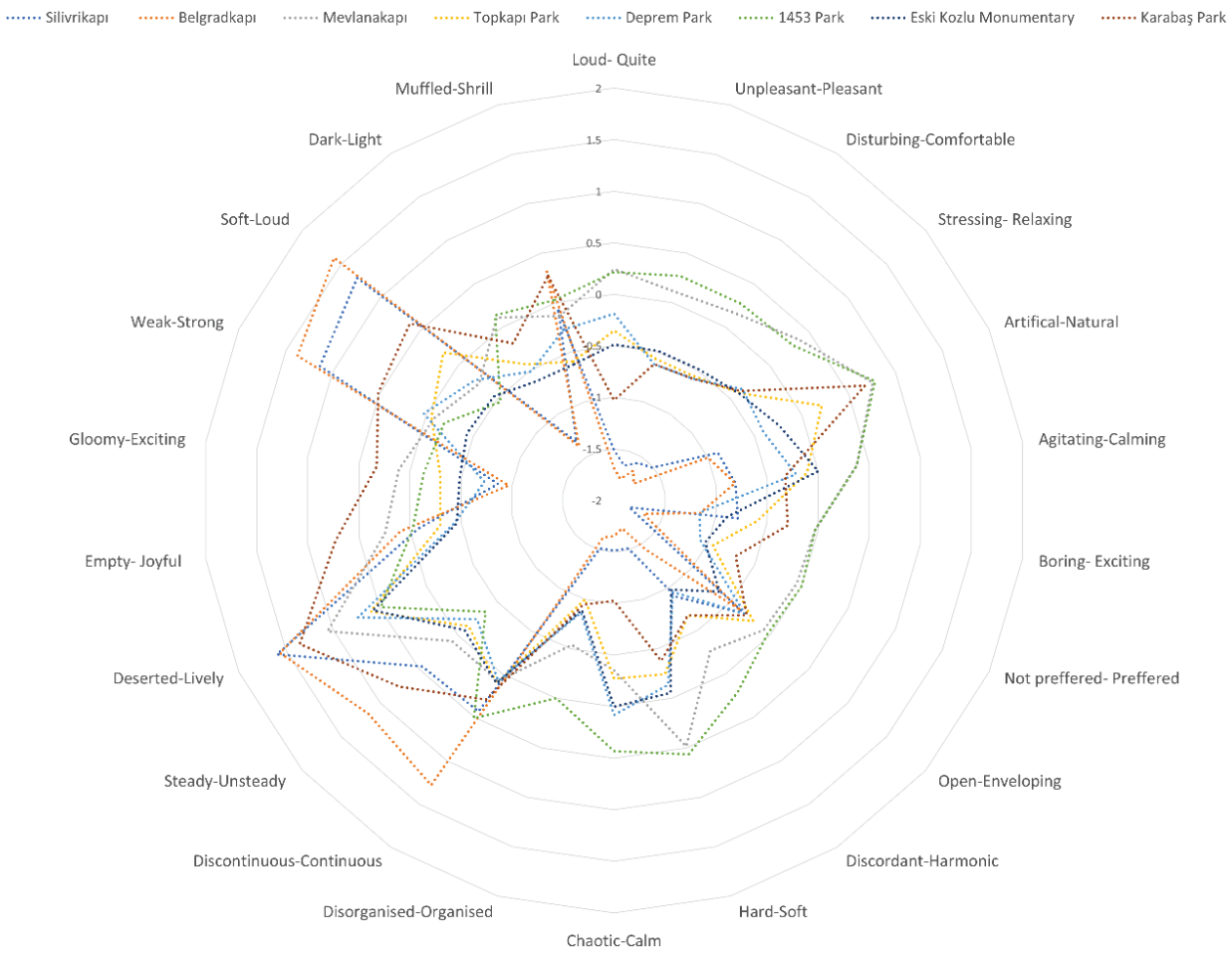


Table 7: The result of the Questionnaire Subjective metric test Depends on Recorded Points



Across eight regions, a total of 22 metrics were contrasted positively and negatively (Table 7). Lively and continuous metrics show the result on the positive side. Boring, not preferred, disorganized and open metrics are mostly chosen as negative metrics.

In a positive majority, Karabaş park, 1453 park and Mevlanakapı park were selected. The data was arranged similarly across the table. On the other hand, regional similarities could not be found in neutral metrics. Belgradkapı and Silivrikapı regions mostly preferred negative metrics. Also, almost all metrics were followed by each other with similar data. This similarity continued to confirm neutrality in the qualifying metrics. The data was arranged in similarly across the table.

On the other hand, regional similarities could not be found in neutral metrics. Belgradkapı and Silivrikapı regions mostly preferred negative metrics. Also, almost all metrics were followed by each other with similar data. This similarity continued to confirm neutrality in the qualifying metrics.

Table 8: Objective Data of Field Area

Location	Decibel	Weather	Date
Eski Kozlu Cemetery	52	+13 Celsius, %70 Humidity 7.4 km/h wind, Sunny	28.11.2021
Silivrikapı	73	+12 Celsius, %75 Humidity 7.1 km/h wind, Cloudy	24.11.2021
Belgradkapı	73	+12 Celsius, %75 Humidity 7.1 km/h wind, Cloudy	24.11.2021
Mevlanakapı Park	59	+13 Celsius, %70 Humidity 7.4 km/h wind, Sunny	28.11.2021
Topkapı Park	61	+13 Celsius, %70 Humidity 7.4 km/h wind, Sunny	28.11.2021
Deprem Park	58	+13 Celsius, %70 7.4 km/h wind, Sunny	28.11.2021
1453 Park	56.8	+13 Celsius, %70 7.4 km/h wind, Sunny	28.11.2021
Karabaş Park	69	+13 Celsius, %70 7.4 km/h wind, Sunny	28.11.2021

During the study, the weather conditions of the places are sunny or cloudy; the humidity is between %70 -%75 and the wind speed is between 7.1-7.4 km/h. At the same time, the decibel range of binaural sounds is in the range of 52-73 decibels (Table 8).

4 Conclusions-Discussion

The study examined the area between the districts of Yedikule and Topkapı. The study focuses on the urban gardens, parks and cemeteries. The main purpose of this study is the fill the literature gap by considering that no studies developed specifically for this region have focused on the sense of sound and the existing data is changing rapidly. The findings show that the spatial transformation experienced in recent years has also affected the soundscape. In this research, three main methodologies, which are binaural recordings, questionnaires and analysis used. In this study, the perceptions of the people about the public green spaces in this region were examined based on the following results:

- It is concluded that there is a serious relationship between the acoustic comfort of the regions and the function usage purpose of the places.
- It is deduced that the acoustic comfort and perception data of the parks built by the municipality in the last 30 years are better than the historical gardens and the surroundings of the gates.
- The historical urban gardens' acoustic comforts are worse than urban parks because of the highway distance and traffic density.
- In negative parameters, all places in the questionnaire are obtained lively and continuous metrics.
- In negative parameters, most places are chosen as boring, not preferred, disorganised and with open metrics.
- Overall, common metrics are chosen in every location.
- Karabaş Park, Belgradkapı and Silivrikapı regions are ticked by, lively and sharp.

Acknowledgements

This research is supported by the institute of YTU graduate school of Science and Engineering. YTU Department of Building Physics and Dr. Hasan Baran FIRAT have supported about acoustic vehicles. The

author acknowledges the department of Naval Architecture and Marine Engineering, the department of Urban and Regional Planning, and Dr. Hasan Baran FIRAT for all their efforts.

References

[1] BSI (The British Standards Institution), "BSI Standards Publication Acoustics — Soundscape Part 1 : Definition and conceptual framework," *BSI Standards Publication*, 2014.
[2] R. Murray. Schafer, "The New Soundscape," *Berandol Music Limited*, 1969.
[3] J. Kang and M. Zhang, "Semantic differential analysis of the soundscape in urban open public spaces," *Building and Environment*, vol. 45, no. 1, pp. 150–157, 2010, doi: 10.1016/j.buildenv.2009.05.014
[4] J. Kang, *Urban sound environment*. London and New York: Taylor & Francis Group, 2006. doi: 10.1201/9781482265613
[5] A. Lex Brown, "A review of progress in soundscapes and an approach to soundscape planning," *International Journal of Acoustics and Vibrations*, vol. 17, no. 2, pp. 73–81, 2012, doi: 10.20855/ijav.2012.17.2302
[6] M. Southworth, "The sonic environment of cities," *Environment and Behavior*, vol. 1, no. 1, pp. 49–70, 1969, doi: 10.1177/001391656900100104
[7] M. S. Engel, A. Fiebig, C. Pfaffenbach, and J. Fels, "A Review of the Use of Psychoacoustic Indicators on Soundscape Studies," *Current Pollution Reports*, vol. 7, no. 3, pp. 359–378, 2021, doi: 10.1007/s40726-021-00197-1
[8] BSI (The British Standards Institution), "PD ISO / TS 12913 - 3:2019 BSI Standards Publication Acoustics — Soundscape - Part 3: Data analysis," pp. 1–30, 2019.
[9] A. Lex Brown, "A review of progress in soundscapes and an approach to soundscape planning," *International Journal of Acoustics and Vibrations*, vol. 17, no. 2, pp. 73–81, 2012, doi: 10.20855/ijav.2012.17.2302
[10] A. Özçevik, "İşitsel Peyzaj – Soundscape' Kavramı İle Kentsel Akustik Konforun İrdelenmesinde Yeni Bir Yaklaşım," 2012.
[11] L. M. Aiello, R. Schifanella, D. Quercia, and F. Aletta, "Chatty maps: Constructing sound maps of urban areas from social media data," *Royal Society Open Science*, vol. 3, no. 3, 2016, doi: 10.1098/rsos.150690
[12] Ö. Axelsson, M. E. Nilsson, and B. Berglund, "A principal components model of soundscape perception," *The Journal of the Acoustical Society of America*, vol. 128, no. 5, pp. 2836–2846, 2010, doi: 10.1121/1.3493436
[13] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980, doi: 10.1037/h0077714
[14] E. Durusoy Özmen, "Değer Olarak Mekânsal Boşluk Ve Korunmasına Yönelik Bir Yöntem Önerisi: İstanbul Kara Surları Dünya Miras Alanı Örneği," 2019.
[15] A. Gide, "İstanbul Historical Peninsula Management Plan," 2018.
[16] A. İ. Kıran, "İstanbul Kara Surları Dünya Miras Alanındaki Tarihi Bostanların Kültürel Peyzaj Kapsamında Değerlendirilmesi," 2019.
[17] E.Durusoy Özmen, "Tarihsel Süreç İçerisinde Değişen Ve Dönüşen bir Mekânsal Boşluk:İstanbul Kara Surlari Dünya Miras Alanı," *Türkiye Bilimler Akademisi Kültür Envanteri Dergisi*, pp. 0–2, 2021, doi: 10.22520/tubaked.2021.23.010

EUROREGIO BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Heart rate-based dynamic sound intervention – a pilot study

Sissel Raahede[1,2,*], Henriette Aa. Holm[1,2], Søren H. Nielsen[1,2]

[1]Aarhus University, Aarhus, Denmark.

[2]SoundFocus ApS, Aarhus, Denmark.

* sr@soundfocus.dk

Abstract
In this paper we present a pilot study on the prospects of working with dynamic sound intervention that adapts to physiological data. The study concerns an investigation of whether dynamic synth-based soundscape compositions, that adapt in tempo to the heart rate of a user, can lower and stabilize the heart rate faster than a 60 beats per minute (BPM) static synth-based soundscape composition. The assumption of the dynamic soundscapes is that if the tempo constantly decreases with either an adaptation of 10% or 2% below the actual heart rate, this will aid in bringing the user from a stressed to a more relaxed state (after 1 minute of physical/psychological exercise). To test the assumption, we included 46 test subjects in a repeated measures test with four different conditions of silence, static, and two variations of dynamic sound stimuli. The study furthermore includes short interviews on how the test subjects experienced the different conditions. Using the velocity of the decreasing heart rate as a measure, the influence of the four different conditions was estimated. The study cannot present any results of statistical significance between how fast the static and dynamic soundscapes decrease the heart rate. Qualitative insights from the interviews reveal how a majority of the test subjects found the characteristics of the soundscape composition relaxing and comfortable, while more varied responses were found in assessing the composition as exciting, boring, and vibrant. The paper discusses the findings and presents critical thoughts on what could be considered in a further development of a concept like this. Even though the study cannot reject the null hypothesis, the paper brings a new perspective to the area of working with sound intervention in health.

Keywords: Sound intervention, Physiological data, Generative soundscape composition, Adaptive sound

1 Introduction

This paper is an investigation on the possible effect of a heart rate-based dynamic sound intervention. Throughout the last decades thorough research has been conducted within music intervention in the health care sector, and this amongst other suggests an advantageous use of music when it comes to releasing stress amongst critically ill patients [1]. The hypotheses in this study is primarily shaped around questions regarding tempo of a soundscape composition. The idea of a correlation between the tempo of music and the heart rate is not a new scientific focus. Former studies have explored this with a clear confirmation on the close relation between the tempo of music and the heart rate [2,3,4,5]. A new perspective to add to this is a dynamic aspect of the development of tempo in the soundscape. One thing is to state that slow music will decrease the heart rate and in contrast fast music will increase the heart rate, but what about a soundscape that continuously adapts to meet the heart rate in trying to aid it decreasing? Could this have an even more fortunate role in bringing a patient from a stressed to a more relaxed state? The agenda in addressing this area of research is to explore whether you can detect any significant difference in static versus dynamic sound intervention. Preliminary studies with 8 test subjects on hypothesis 1 revealed indications about the conditions of soundscape intervention having an

advantageous effect in decreasing the heart rate faster than no intervention. There was no tendency in relation to a difference between the static and dynamic sound intervention. These initial tests gave rise to a further investigation with an increased number of test subjects and a second hypothesis proposing an increased percentage gap between the heart rate and the tempo of the soundscape composition.

H_1: With a generative synth-based soundscape composition, that adapts to the heart rate of a user by setting the Tempo (BPM) 2 % below the measured heart rate, it is possible to lower and stabilize the heart rate faster than with a static synth-based soundscape composition or no intervention.

H_2: This hypothesis is similar to H_1 with exception of the tempo of the soundscape composition being 10% below the measured heart rate instead of 2%.

H_0: There is no difference to detect in how fast the heart rate is lowered and stabilized when you listen to the dynamic or static synth-based soundscape composition.

1.1 Project context

The project is a sub-project with connection to a larger context of investigating interactive sound zones in domestic and health care settings in a four-year long research project named Interactive Sound Zones for Better Living (ISOBEL) [6]. ISOBEL is a project in collaboration of the Danish companies Bang&Olufsen, SoundFocus, WaveCare and two departments of Aalborg University (Department of Electronic Systems/ Department of Computer Science). Due to an overall interest in exploring and developing new interaction techniques enabling dynamic sound zones, this minor project erupts as a pilot study to detect what could be an interesting field of heart rate-based sound intervention and possible inspiration for content within these sound zones. The purpose thus is to kickstart the process of exploring ways to work with dynamic sound content and how that might affect the user in both a physiological and perceptual manner.

2 Soundscape Composition

The synth-based soundscape composition is designed in the software programs Ableton and Max MSP specifically for the purpose of this context. It is characterized by self-generating elements that evolves on the basis of both fixed and random parameters. Due to studies about the pentatonic major scale being able to increase the parasympathetic tone [7], a fixed parameter is for the composition to stay within the major pentatonic C-scale. Due to the focus on detecting a correlation between sound and heart rate, the rhythmical aspect of the composition plays an important role in creating the perceptual synchronization between the current heart rate and musical tempo. A semi-random drum pattern is divided into three intervals with slight differences in the rhythm and number of subdivisions of the beat. This is to accommodate a pleasant and non-stressful feel in 130 as well as 50 BPM. The self-generating and random parameters in the composition, are an attempt to mimic characteristics found in nature sounds. *The Attention Restorative Theory* [8] suggest that nature can help us improve our mental health and one of the reasons has to do with the concept of *Soft Fascination,* which is about attending to inspiring patterns effortlessly while still being able to focus on other things. Nature sounds can at the same time possess captivating qualities and soothing effects [8]. The synth-based soundscape composition aims at balancing the expected and unexpected in how the evolvement of the sounds behave. This is to evoke some connotations to that same feeling of being both inspired and soothed.

Regarding the different functions of the soundscape being both static and dynamic, there are different conditions for the tempo. The static soundscape has a fixed tempo of 60 BPM, since findings indicate that music around 60 BPM can cause a synchronization between brain and music creating alpha brainwaves, which characterizes a relaxed state [9]. When the soundscape is dynamic, the composition evolves in accordance with the heart rate, by continuously being 2% or 10% lower than the real-time measured pulse. This adaptation of 2% or 10% is based on another study investigating the relation between increased acoustic tempo and elevation

in the heart rate [3]. Here the increment of tempo was based on heart rate measurements from a baseline period and not continuous measurements of the current heart rate. Even though we apply this in a different real-time context, these findings have functioned as inspiration for determining the conditions for the adaptation between soundscape and heart rate. Since results from the abovementioned study indicate that the tempo of the music should neither be too close nor too far from the heart rate [3], the current study included two dynamic sound interventions with the different adaption conditions of 2% and 10%. In this way this study not only tested for an effect of the dynamic conditions compared to no intervention and the static condition, but also for how closely the tempo of the soundscape composition should follow the heart rate.

3 Experiment

An experiment was conducted to test the different hypotheses. Physiological data on the heart rate development during each phase was gathered. Furthermore, short interviews were carried out with the aim of addressing the perception of the four different conditions.

3.1 Test conditions

The test setup consisted of four different phases each including one of the four stimulus conditions: no intervention, static soundscape intervention and dynamic soundscape intervention with an adaption rate of 2% or 10% respectively. The sensor used to adapt the dynamic sound stimuli to the heart rate and to monitor the evolvement of the pulse throughout the different conditions, was a Garmin HRM-Pro breast strap, measuring the heart rate through an electrocardiogram (ECG). The test was conducted on 46 subjects in an age ranging from 20 to 28. 19 men and 27 women who was all told that they participated in a listening experiment with four different soundscape conditions each of five minutes but not the characteristics of these conditions. Prior to each phase, each test subject was exposed to a one-minute long combined physical and psychological stress test, to increase the heart rate and emulate a stressed body state. This involved stepping up and down a step bench in the tempo of a 115 BPM, while performing a *Stroop Color test* [10]. After the one-minute stress test, the subject was without further instructions placed in a comfortable chair wearing headphones and then one of the four stimuli would play for five minutes. The playback level was adjusted to the individual preference. To avoid any unfortunate effects from a fixed stimuli order, the order of the different conditions was randomized. Furthermore, the soundscape composition played through speakers during the introduction of the test for the test subjects to habituate to its characteristics. When all four phases had been conducted the test was wrapped up with a short interview to collect demographic data and to address the perceptual experience of the stress test and the soundscape characteristics.

3.2 Data processing

The method in analyzing the collected data was directed at determining the (quite stable) heart rate reached during the five-minute resting period, and the velocity of how fast this value was reached. The heart rate time series data was modelled using a piecewise linear shape: A slope describing the change of heart rate from high to low, and a flat section describing the stabilized (low) heart rate. In order to find the value of the stabilized heart rate the heart rate dataset was divided into two distributions corresponding to two groups of high and low heart rates, respectively. Using a Gaussian mixture model (GMM) [11] the datasets were split into two clusters with mean and variances as output for each cluster. The lowest mean is the stabilization value. The stabilization time was found as the time it takes the heart rate to reach the lower mean from the maximum value. The stabilization velocity is the slope of the linear regression applied to the heart rate data from t = 0 (the starting time of the five-minute intervention after the one-minute stress test) to t = stabilization time. An example of these calculations and the development of the heart rate for one of the test subjects is shown in Figure 1. These values enable us to talk about at what level the heart rate of a test subject can be considered to be stabilized

and the velocity of how fast they reach that level. In correlation with the different stimulus conditions, these values are what form the basis of the statistical test.

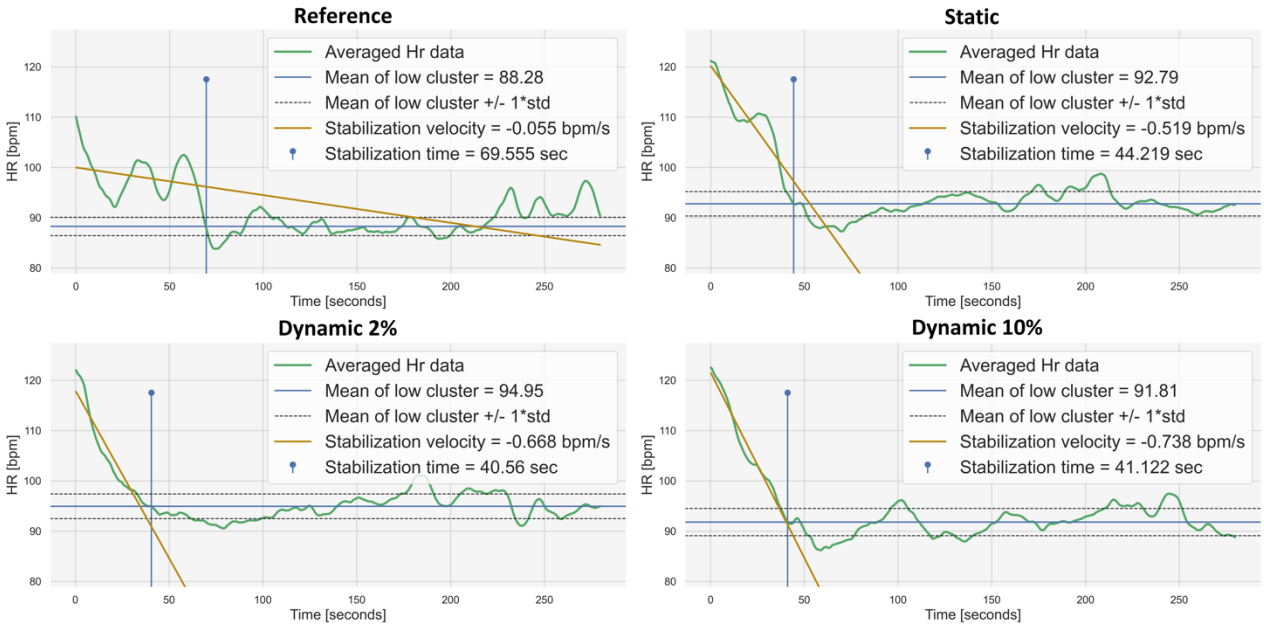


Figure 1: Calculation of heart rate stabilization and velocity of decrement. T = 0 is the starting time of the five-minute intervention after the one-minute stress test.

3.3 Statistical method

To test for an effect on the velocity in the different conditions, a one-way repeated measures ANOVA within subjects was conducted using the *ezANOVA()* function in RStudio with *Condition* being the predictor variable and *Velocity* the outcome variable [12].

3.4 Results

3.4.1 Statistical findings

With the statistical analysis it has been found that the null hypothesis cannot be rejected, since the calculated p-value is 0.8 and thus there is an 80% chance of the result being arbitrary. The study therefore indicated no significant difference between how fast the dynamic and the static soundscape decreases the heart rate. Figure 2 illustrates how the four conditions have a wide distribution of velocity values and a small effect size. With the velocity values distributed on both sites of zero, it indicates that the effect of the conditions sometimes opposed what was assumed. The characteristics of the data distribution corresponds to the lack of statistical significance. Nor does Figure 3 present any convincing results. If we were to extract any indications in favor of the study though, a feeble inclination of the expected tendency, about the dynamic soundscapes being faster in decreasing the heart rate than the static soundscape or no intervention, can be found in the mean values of the measured data. However, with all the mean values being within the overlapping confidence intervals, the final excerpt is still that the results might as well be arbitrary.

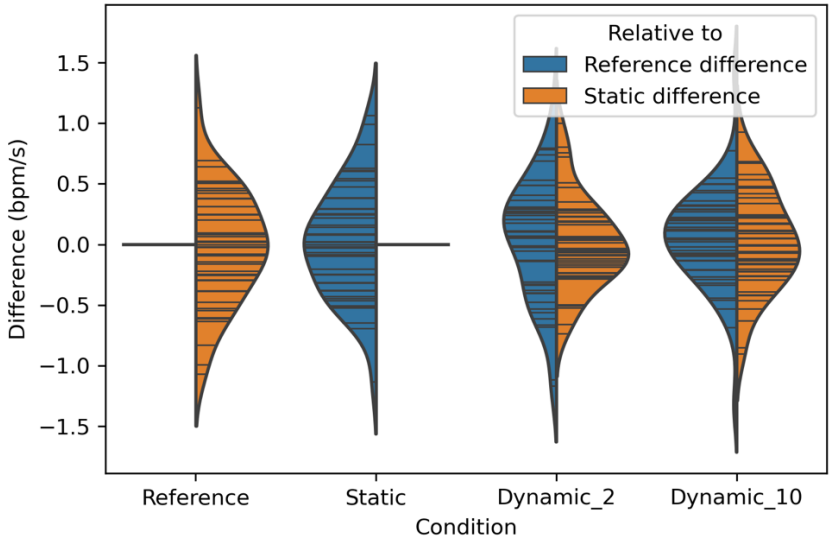


Figure 2: Violin plot illustrating the data distribution and the conditions and their difference in velocity relatively to each other.

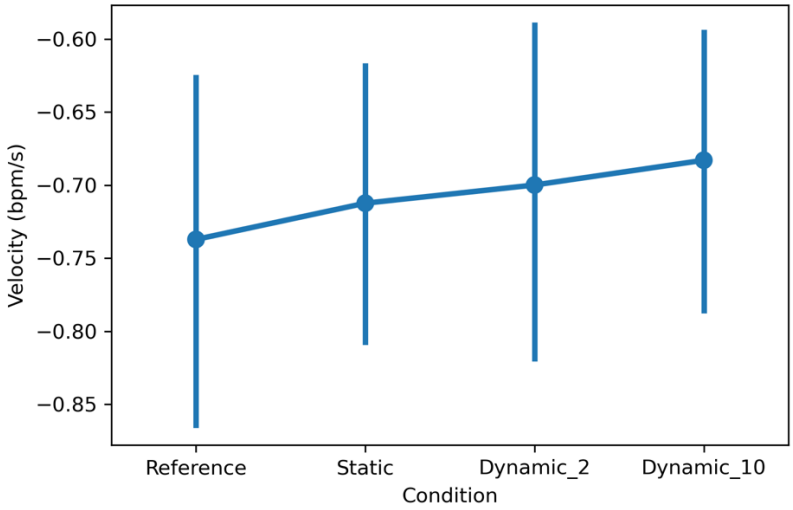


Figure 3: Plot of mean values and confidence intervals.

3.4.2 Demographic data and the perceptual experience

Regarding the collected demographic data, the key study was that the group of test subjects appeared relatively uniform. All test subjects were somehow related to programs at Aarhus University, which reflects the age ranging from 20 to 28. Furthermore, we asked for a self-assessment on physical shape which varied over a scale from one to ten, but with the majority assessing themselves above average. The results and any possible tendencies can therefore only assert itself for this specific demography.

According to how the test subjects perceived the overall characteristics of the soundscape composition, they were asked to what degree they agreed on the soundscape possessing eight predetermined characteristics. Figure 4 presents a broad consensus about the soundscape characteristics being relaxing and comfortable. In contrast, Figure 4 reveals inconsistencies when it comes to the perception of the soundscape being exciting, boring, and vibrant. In more open questions about a comparison between the different stimulus, the majority commented on how they felt that the sound stimulus helped them direct focus towards the soundscape and their breathing. It slowed down their racing thoughts, which instead was dominant in the silent stimuli. Furthermore, the interview did not disclose any prevalent pattern in whether the subjects perceived any difference between the soundscape tempo being static or dynamic. Some perceived the dynamic evolvement of the tempo being a guide in decreasing the pulse and reaching a more relaxed state, while others found the change of tempo less relaxing. Several subjects found it hard to tell the three sound conditions apart.

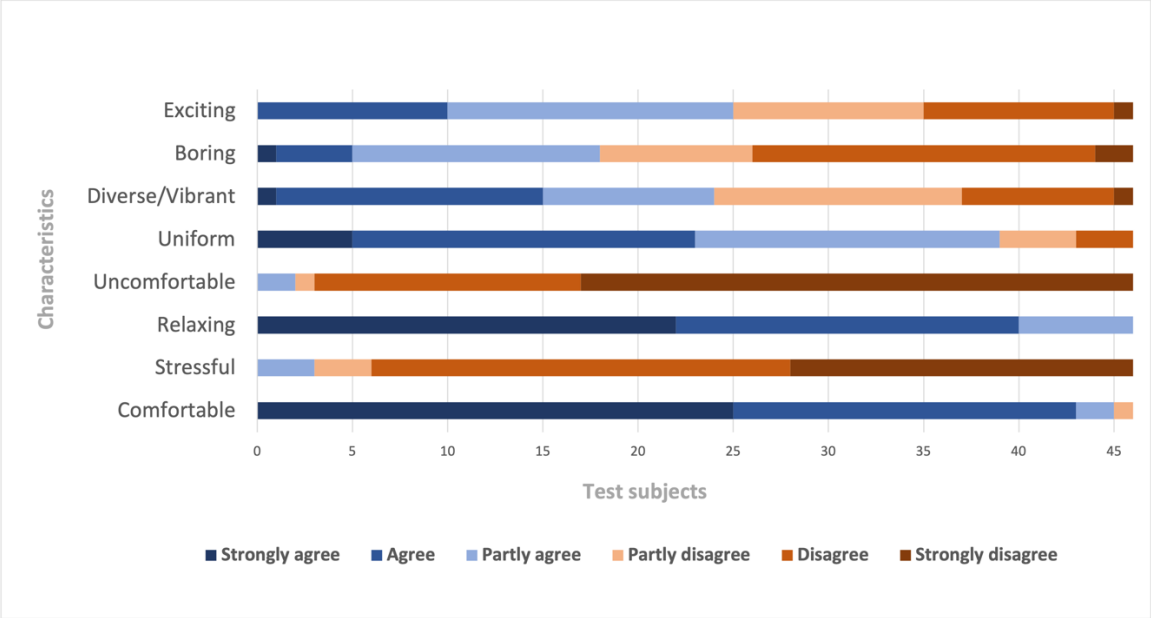


Figure 4: Evaluation on the experience of the soundscape within predetermined characteristics.

4 Discussion

The results from the statistical model do not allow us to reject our null hypothesis, since there is no significant difference between any of the four conditions. This could be due to many factors.

4.1 Data processing methods

In considerations of what could have had an impact on the results not achieving statistical significance, one reason could be in the chosen method for processing the physiological measurements. Perhaps the simple data model consisting of a constant stabilized heart rate in combination with a linear decrease in heart rate was not the ideal method. Other data models which can better accommodate the more complex evolution of heart rate as function of time could be investigated.

4.2 Characteristics of soundscape composition

Regarding the characteristics of the soundscape composition, this too could have an impact on the decrement velocity. Figure 4 indicates an overall agreement on the soundscape characteristics being relaxing and non-stressful, but even more efficient approaches in composing soundscapes for the purpose of relaxation and decreasing the heart rate could exist. Though tempo could seem like the obvious parameter to control the heart rate by, other parameters such as frequencies or amplitude levels could be of relevance to consider. Furthermore, one could ask why the composition consisted of synthetic sounds and not nature sounds, when these are the ones, the composition tries to mimic. An answer to this could be the advantage of flexibility in the synthetic elements and the fact that the dynamic tempo changes could add unwanted and unnatural distorted effects to nature sounds. Though it might be more convenient to use synthetic sounds in a dynamic context, it does not necessarily mean that synthetic sounds are the best choice for every occasion. Exposure to nature sounds has proven to be associated with reduced stress, and the reason why the natural elements are calming to us, is asserted to have linkage to human biology and our survival in the past [13]. It could be essential to investigate a comparison between an actual natural soundscape and a synthetic (dynamic) soundscape simulating the behavior of nature sounds, to see if the relaxation response mainly depends on the acoustic qualities of the sounds or the referentiality in the behavior of a nature soundscape.

To address the more opposing responses in determining the sound characteristics as boring, exciting, or vibrant, this could be due to individual preferences of music. Music-genre preferences depend on both intrinsic properties of the music and external associations. Therefore, the affective reaction to music is not only shaped around musical qualities, but also factors such as psychological dispositions, social interactions, generation, culture, etc. [14]. With the test subjects agreeing on the sound stimuli being relaxing and comfortable, but disagreeing in it being exciting and vibrant, it could suggest that, whether the soundscape succeeds in reaching a perceived level of *Soft Fascination* varies from individual to individual. Another interesting observation was the variation in whether the test subjects perceived any change of tempo in the dynamic soundscape compositions. The fact that many test subjects could not tell the three sound interventions apart, could indicate a positive effect of the three intervals in the drum pattern trying to accommodate a pleasant and non-stressful feel of the soundscape across a wide range of BPM values. Those test subjects who noticed the tempo change in the dynamic conditions varied in whether they found it soothing or not, which is in consistence with the results about the effect of the conditions sometimes opposing what was assumed.

4.3 Related work

As stated, this is not the first research project investigating correlations between a musical tempo and the heart rate, but there are not many with a focus of dynamic sound adaptation to physiological parameters in health care. There is a commercial system though, that works with personalized soundscapes. The company *Endel* has created an app, for mobile devices, based on a technology that adjusts soundscapes in real-time to physiological and environmental parameters such as location, weather, and heart rate [15]. It could be of inspiration to take more than just one factor into account, when deciding how to adapt a soundscape to a specific user. The increment of heart rate is not necessarily equal to an unfortunate mental state since the heart rate also increases due to many other factors and feelings [16]. Stress is a complex phenomenon and cannot be assessed reliably through standard heart rate measures only. The physiological level of stress is for example tightly coupled to the heart rate variability (HRV), indicating which part of the autonomic nervous system is currently active. Moreover, this is just one of many factors suitable to clarify a more precise stress assessment [17]. Therefore, it seems appropriate to take other factors and other sensory inputs into account in an adaptive environment. In this way we may move closer to presenting a soundscape solution as accurately as possible in responding reliably to the actual mental or physiological state of a user or patient.

4.4 Reflections on future work

Future work could revolve around trying to meet the assumptions of the study and obtain statistical evidence. This might be achieved through adjusting the data processing methods, possible improvements on the accuracy in assessing the stress level of a patient, or different approaches to composing soundscapes. Another essential perspective to add to this relatively new area of research could be how the concept of dynamic sound intervention could take shape in an actual health care context. Future investigations could potentially benefit from an expanded interdisciplinary approach involving a technical perspective and a more holistic approach in assessing not only the isolated effect and experience of the soundscape composition in itself, but also the soundscape composition in the context of the hospital environment. Following the International Organization for Standardization (ISO) on acoustics and soundscapes one must apply mixed methods to fully grasp the complexity of an acoustic environment and hence reach a comprehensive understanding of the context in which intervention is made [18]. The interviews might give us some insights on how the different conditions of stimuli and the characteristics of the soundscape composition are experienced, however we have no knowledge on the usage scenarios. How will patients experience interacting passively with a soundscape system while being in a perhaps critical physical or mental state? Could the system in any way be of disturbance to the work of the healthcare staff? How is the soundscape composition experienced in the context of other environmental sounds? Within the tradition of Human Computer Interaction (HCI), designing interactive systems does not only require considerations on the system being usable, but also about who the users are, how the system is going to be used, and what the context of the system is [19]. With an intention of investigating the application of dynamic soundscape intervention in hospitals, one could advocate a more ecological approach to not only consider justified or non-justified outcomes, but also the complexity of the context for intervention [20,21]. This could expand the study in a contextual matter, which might also inspire for solutions on how to improve technical matters.

5 Conclusions

In conclusion, this pilot study, and the investigation on the effect of a heart rate-based dynamic sound intervention, reveals that the null hypothesis cannot be rejected. This is due to the lack of significant difference in the velocity of how fast a static and the dynamic soundscapes decrease the heart rate. Besides justifying an effect of working dynamically with sound intervention, the purpose of this study was also to begin a process of exploring ways to work with dynamic sound content in health care. The study gave rise to some interesting considerations within the field of dynamic sound intervention. The approach presented in this study is not the only way of working adaptively with soundscapes. Heart rate is just one way of applying physiological sensory inputs for dynamic sound intervention, and likewise are physiological inputs just some of many possible aspects to adapt soundscape composition around.

Acknowledgements

Thanks to Shagen Djanian and Rune Møberg Jacobsen for assistance in statistical processing. The ISOBEL project is partly funded by the Danish Innovation Fund case number 9069-00038B.

References

[1] Heslet, L. and Dirckinck-Holmfeld. (2007). Musik som medicin. In Sansernes hospital. Arkitektens Forlag.

[2] Bora, B., Krishna, M. and Phukan, K.D. 2017. The Effects of Tempo of Music on Hear Rate, Blood Pressure and Respiratory Rate – A Study in Gauhati Medical College. *Indian J Physiol Pharmacol.* 61(4), (2017), 445-448.

[3] Watanabe, K. et al. 2017. Heart rate responses induced by acoustic tempo and its interaction with basal heart rate. *Scientific Reports*. 7, 1 (2017).

[4] Agrawal, A., Makhijani, N. and Valentini, P. 2013. The effect of Music on Heart. *Journal of emerging investigators.* (2013).

[5] Suguna, S. and Deepika, K. 2017. The effects of music on pulse rate and blood pressure in healthy young adults. *International Journal of Research in Medical Sciences*. 5, 12 (2017).

[6] About ISOBEL: 2021. *https://isobel.dk/index.php/about-isobel/*. Accessed: 2021- 11- 16.

[7] Ranger, A. et al. 2018. Physiological and emotional effects of pentatonic live music played for preterm neonates and their mothers in the Newborn Intensive Care Unit: A randomized controlled trial. *Complementary Therapies in Medicine*. 41, (2018), 240-246.

[8] Kaplan, S. 1995. The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology*. 15, 3 (1995), 169-182.

[9] Releasing Stress Through the Power of Music | Counseling Services: 2021. *https://www.unr.edu/counseling/virtual-relaxation-room/releasing-stress-through-the-power-of-music*. Accessed: 2021- 11- 16.

[10] Pelivanoglu, B. et al 2005. Computer Adapted Stroop Colour-word Conflict Test As a Laboratory Stress Model. *Erciyes Tip Dergisi* (2005).

[11] Vanderplas, J. 2016. Machine Learning. In *Python Data Science handbook: essential tools for working with data. O'Reilly Media,* Sebastopol, CA.

[12] Field, A., Miles, J. and Field, Z. 2012 *Discovering Statistics using R.* SAGE Publications.

[13] Largo-Wight, E. et al. 2016. The Efficacy of a Brief Nature Sound Intervention on Muscle Tension, Pulse Rate, and Self-Reported Stress. *HERD: Health Environments Research & Design Journal*. 10, 1 (2016), 45-51.

[14] Rentfrow, P. et al. 2011. The structure of musical preferences: A five-factor model. *Journal of Personality and Social Psychology*. 100, 6 (2011), 1139-1157.

[15] Endel: 2021. *https://endel.io*. Accessed: 2021- 11- 25.

[16] Ekman, P. et al. 1983. Autonomic Nervous System Activity Distinguishes Among Emotions. *Science*. 221, 4616 (1983), 1208-1210.

[17] Wu, W. et al. 2019. Quantitative Assessment for Self-Tracking of Acute Stress Based on Triangulation Principle in a Wearable Sensor System. *IEEE Journal of Biomedical and Health Informatics*. 23, 2 (2019), 703-713.

[18] International Organization for Standardization. 2018. *Acoustics – Soundscapes – Part 2: Data collection and reporting requirements* (ISO standard no. 12913-2:2018). https://www.iso.org/standard/75267.html

[19] Preece, J., Sharp, H. and Rogers, Y. 2015. *Interaction design*. Wiley.

[20] Becker, F. et al. 2011. Integrated Healthscape Strategies: An Ecological Approach to Evidence-Based Design. *HERD: Health Environments Research & Design Journal*. 4, 4 (2011), 114-129.

[21] Højlund, M. 2017. Overhearing: An Attuning Approach to Noise in Danish Hospitals. (2017).

AUTHOR INDEX