TURUN YLIOPISTO
UNIVERSITY OF TURKU

DEVELOPING ACOUSTIC RATING QUANTITIES USING EXPERIMENTAL PSYCHOACOUSTIC DATA

Petra Virjonen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. F OSA – TOM. 2 | TECHNICA – INFORMATICA | TURKU 2021

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

DEVELOPING ACOUSTIC RATING QUANTITIES USING EXPERIMENTAL PSYCHOACOUSTIC DATA

Petra Virjonen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. F OSA – TOM. 2 | TECHNICA – INFORMATICA | TURKU 2021

University of Turku

Faculty of Technology
Department of Computing
Computer Science
Doctoral Programme in Mathematics and Computer Sciences (MATTI)

Supervised by

Research Group Leader, Valtteri Hongisto
Turku University of Applied Sciences
Adjunct Professor of Environmental Psychology (Indoor environment)
University of Turku
Adjunct Professor of Acoustic Noise Control
Aalto University

Professor, Marko Mäkelä
Department of Mathematics and Statistics
University of Turku

Associate Professor, Tapio Pahikkala
Department of Computing
University of Turku

Reviewed by

Professor, Jin Yong Jeon
Department of Architectural Engineering
Hanyang University

Head of Group, Dr. Beat Schäffer
Laboratory for Acoustics/Noise Control
Empa – Swiss Federal Laboratories for Materials Science and Technology

Opponent

Head of Working Group, Dr.-Ing. Volker Wittstock
Applied Acoustics
Physikalisch-Technische Bundesanstalt PTB –
National Metrology Institute of Germany

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-8536-4 (PRINT)
ISBN 978-951-29-8537-1 (PDF)
ISSN 2736-9390 (PRINT)
ISSN 2736-9684 (ONLINE)
Painosalama, Turku, Finland, 2021

To my family

UNIVERSITY OF TURKU
Faculty of Technology
Department of Computing
Computer Science
VIRJONEN, PETRA: Developing acoustic rating quantities using experimental psychoacoustic data
Doctoral dissertation, 123 pp.
Doctoral Programme in Mathematics and Computer Sciences (MATTI)
August 2021

ABSTRACT

Noise is frequently present in our daily lives. Noise is encountered both indoors and outdoors, in various environments such as at home, in traffic and at work place. Excessive noise causes annoyance and other negative health effects. Acoustical quantities are needed to quantify the properties of noise as well as to determine sound insulation capabilities of structures. Psychoacoustic experiments have revealed that the existing quantities do not always correspond well to subjective perception. The objective of this thesis was to develop acoustical quantities that correspond better to subjective perception than the existing quantities by deploying psychoacoustic data. This thesis contains four publications. Three publications consider developing quantities for rating sound insulation, and the fourth deals with annoyance penalty for periodic amplitude modulated sounds.

Sound insulation of a structure (e.g. a wall, window, floor, or façade) is measured in several one-third octave bands. To simplify the information gathered with sound insulation measurements, and to enable easier comparison between structures as well as to facilitate the imposition of legislative requirements, single-number quantities have been developed for rating sound insulation of different structures, such as described in standards ISO 717-1 and ISO 717-2. It is important that these simplified quantities are based on psychoacoustical experiments to ensure adequate correspondence to subjective perception. The development of the existing quantities has not taken this aspect fully into account.

In this thesis, single-number quantities for rating sound insulation were developed using mathematical optimization and psychoacoustic data from three already published experiments. New reference spectra were derived for various living sounds for airborne sound insulation, typical domestic impact sounds for impact sound insulation as well as for several road traffic spectra for façade sound insulation. The results showed that an optimized reference spectrum resulting in well-performing single-number quantity could be derived for each studied sound type. The developed mathematical optimization scheme enables systematic development of new reference spectra, which are based on scientific evidence.

Amplitude modulated sounds (e.g. road traffic noise and wind turbine noise) have been found to cause more annoyance than steady-state sounds. A penalty can be added to the measured sound level to take into account the increase in annoyance due to modulation in sound. In this thesis, it was studied if a penalty is needed for periodically amplitude modulated wide-band general sounds having a low sound level by conducting a psychoacoustical experiment with 40 participants. The results showed the need for penalty for certain ranges of modulation frequencies and modulation depths for the studied sounds. The penalties increased with increasing modulation frequency and increasing modulation depth, which does not support using a constant penalty.

The results of this thesis can be exploited in the improvement of future academic research, acoustical standards, and building regulations.

KEYWORDS: psychoacoustics, sound insulation, annoyance, mathematical optimization, amplitude-modulation, penalty

TURUN YLIOPISTO
Teknillinen tiedekunta
Tietotekniikan laitos
Tietojenkäsittelytiede
VIRJONEN, PETRA: Developing acoustic rating quantities using experimental psychoacoustic data
Väitöskirja , 123 s.
Matemaattis-tietotekninen tohtoriohjelma (MATTI)
Elokuu 2021

TIIVISTELMÄ

Melu on usein läsnä päivittäisessä elämässämme. Kohtaamme melua erilaisissa ympäristöissä kuten kotona, liikenteessä ja työpaikoilla. Liiallinen melu häiritsee, ja melun on todettu voivan vaikuttaa myös muilla tavoin negatiivisesti terveyteemme. Akustisten suureiden avulla voidaan määrittää melun ominaisuuksia sekä erilaisten rakenteiden ääneneristävyyskykyä. Psykoakustiset kokeet ovat tuoneet ilmi, että olemassa olevat suureet eivät aina vastaa riittävän hyvin subjektiivista kokemusta melun häiritsevyydestä. Tämän väitöskirjan tarkoituksena oli psykoakustista dataa hyödyntäen kehittää akustisia suureita, jotka vastaavat subjektiivista kokemusta nykyisiä suureita paremmin. Väitöskirja koostuu neljästä osajulkaisusta. Kolme julkaisuista käsittelee ääneneristävyyden yksilukuarvojen kehitystä, ja neljäs jaksollisesti moduloitujen äänten häiritsevyyssanktion määrittämistä.

Rakenteen (esimerkiksi seinä, ikkuna, lattia tai julkisivu) ääneneristävyys mitataan kolmasosaoktaaveittain useammalla taajuuskaistalla. Monimutkaisen taajuuskaistainformaation yksinkertaistamiseksi on kehitetty yksilukuarvoja (esimerkiksi standardit ISO 717-1 ja ISO 717-2) eri rakenteiden vertailuun sekä mahdollistamaan rakennusmääräysten asettamisen. On tärkeää, että kehitetyt yksilukuarvot perustuvat psykoakustiseen tutkimusnäyttöön, ja vastaavat riittävän hyvin subjektiivista häiritsevyyttä. Nykyisten käytössä olevien standaroitujen suureiden kehityksessä tätä ei ole täysin huomioitu.

Tässä väitöskirjassa ääneneristävyyden yksilukuarvoja kehitettiin matemaattisen optimoinnin ja kolmessa jo aiemmin julkaistussa tutkimuksessa kerätyn psykoakustisen datan avulla. Ilmaäänen- ja askelääneneristävyyden referenssispektrit johdettiin erilaisille asuinäänille. Julkisivun ääneneristävyydelle johdettiin referenssispektrit usealle eri tieliikennemeluspektrille. Tulokset osoittivat, että optimoitujen referenssispektrien avulla saadut yksilukuarvot suoriutuivat hyvin kullakin tutkitulla äänityypillä. Kehitetty matemaattinen optimointimalli mahdollistaa tieteelliseen näyttöön perustuvien referenssispektrien ja täten uusien yksilukuarvojen systemaattisen kehityksen.

Amplitudimoduloitujen äänien kuten liikennemelun ja tuulivoimalamelun on havaittu häiritsevän enemmän kuin tasaisten äänien. Moduloinnin aiheuttama häiritsevyyden lisäys voidaan huomioida lisäämällä mitattuun keskiäänitasoon sanktio. Tässä väitöskirjassa tutkittiin, tarvitaanko sanktiota jaksollisesti vaihtelevalle

laajakaistaiselle melulle, jonka äänitaso on matala. Tätä tutkittiin toteuttamalla psykoakustinen kuuntelukoe, johon osallistui 40 koehenkilöä. Tulokset osoittivat, että sanktio voi olla tarpeen tietyllä modulaatiotaajuuden ja modulaatiosyvyyden arvoalueella. Sanktion suuruus kasvoi modulaatiotaajuuden ja modulaatiosyvyyden kasvaessa, mikä ei tue vakiosanktion käyttöä.

Tämän väitöskirjan tuloksia voidaan hyödyntää tieteellisessä tutkimuksessa sekä akustisten standardien ja rakennusmääräysten kehittämisessä.

ASIASANAT: psykoakustiikka, ääneneristävyys, häiritsevyys, matemaattinen optimointi, amplitudimodulaatio, sanktio

Acknowledgements

My journey in acoustics began in spring 2002 when I started as a trainee in the Laboratory of Ventilation and Acoustics at Finnish Institute of Occupational Health. My first assignment was to learn how to use the newly arrived Head and Torso Simulator (aka Keino, see Figure 1 in Section 2.2.2). Keino played a part also on my Master's Thesis, which dealt with measuring and predicting speech intelligibility. After graduating in 2003, I continued working as a researcher in the laboratory. During the following years, I worked with several topics of acoustics, such as office acoustics and speech privacy, building acoustics, and psychoacoustic experiments.

During my study leave in 2014–2015, I studied several courses of applied mathematics at University of Turku. I attended the Optimization algorithms course lectured by Professor Marko Mäkelä. By then, a psychoacoustic experiment studying the annoyance towards living sounds through wall structures had been conducted in the laboratory (at that time called Indoor environment) and the need for a better single-number-quantities had been recognized. My supervisor, Research Group Leader Valtteri Hongisto suggested to study this further. Optimizing the reference spectrum for airborne sound insulation using the data from the psychoacoustic experiment seemed like a perfect topic for the optimization course exercise. The exercise work led finally to the first publication included in this present thesis, Publication I. It was natural to continue the work to optimizing the reference spectra for impact sound and façade sound insulation as well (Publications II and III).

In 2016, the business of the laboratory was transferred to Turku University of Applied Sciences, and the work continued there in the Built environment laboratory. The psychoacoustic experiment of Publication IV was conducted in 2016. It studied the need of a penalty for amplitude modulated sounds, and thus, dealt also with developing acoustic quantities to correspond better with subjective experience.

In 2017, I started working at the department of Future Technologies in University of Turku, and became familiar with data analysis and machine learning, which helped me especially with model selection and model performance evaluation aspects of the thesis. From July 2019 forward, I was granted a 12 month grant for my doctoral studies by MATTI, Doctoral Programme in Mathematics and Computer Sciences in University of Turku. This helped me to concentrate more efficiently to the studies.

Completing the thesis has taken a rather long time. However, it is nice to realize that it combines several topics from my career, namely building acoustics, psychoacoustical experiments, mathematical optimization, and data analysis and machine learning into a rather coherent work. And of course, Keino was along again, in the calibration process of the sounds in the psychoacoustic experiment!

First I want to express my appreciation to Valtteri, you have provided invaluable support, thorough guidance, and lasting patience throughout this project. I am impressed by your ambition and open-minded vision. Among many things, you have taught me to strive for simple, understandable and target-friendly outcome, which is very important as the results are not solely meant for scientific community.

I would like to thank my supervisors and co-authors Professor Marko Mäkelä and Associate Professor Tapio Pahikkala for all your guidance and smooth co-authoring. Your support has been essential especially in questions concerning mathematical optimization and model performance.

I am grateful to my thesis director Professor Jukka Heikkonen. You have continuously encouraged and supported me in every step. I greatly appreciate your guidance in writing the synopsis.

I have been honored to have Professor Jin Yong Jeon from Hanyang University and Head of Group, Doctor Beat Schäffer from Swiss Federal Laboratories for Materials Science and Technology as the preliminary examiners of this thesis. I highly appreciate your constructive comments and feedback.

My special thanks go to my co-authors Mikko Kylliäinen, Jenni Radun and David Oliva for fluent and successful cooperation. I am grateful for the opportunity to learn from your expertise.

I sincerely thank my competent and helpful colleagues Jarkko Hakala, Jukka Keränen, Henna Maula, Pekka Saarinen, and the rest of the research group at Turku University of Applied Sciences. You have always been there to help me with smaller problems as well as putting out fires.

I would like to recognize the help I received from my colleagues at University of Turku: Paavo Nevalainen, Markus Viljanen, Yury Nikulin, and Antti Airola, you have each supported me in special aspects of my thesis. I want to thank Alaleh Maskooki, Maria Jaakkola, Parisa Movahedi, Ileana Montoya Perez, Riikka Numminen and Elise Syrjälä for long walks and invigorating discussions, not just on the thesis matters but on everything else as well.

Thank you Pia Lindroth for good friendship and for being always there for me.

I thank Finnish Institute of Occupational Health and Turku University of Applied Sciences for enabling the research, and permitting the use of their data for this thesis. I also want to acknowledge Business Finland and participating companies, and University of Turku for the financial support.

I want to thank my parents and parents-in-law. You have truly made this possible by supporting our family in every way. And lastly, I thank my beloved family, my husband Tero: it seems you have never doubted whether I am able finish this thesis, you have pushed me forward, even on times when the work seemed difficult or tedious. My dear children Enni, Voitto, Vilho: you have shown me what is most important in life.

July 28, 2021
Petra Virjonen

Table of Contents

Abbreviations

A	Measured equivalent absorption area of the receiving room [m^2]
A_0	Reference absorption area, 10 [m^2]
A_c	Amplitude of the carrier sound [Pa]
AM sound	Amplitude modulated sound
C	Spectrum adaptation term (A-weighted pink noise) [dB]
c	Speed of sound [m/s]
C_I	Spectrum adaptation term (unweighted linear impact sound level) [dB]
C_{tr}	Spectrum adaptation term (A-weighted urban traffic noise) [dB]
CV	Cross validation
LOOCV	Leave-one-out cross validation
D	Level difference [dB]
D_m	Modulation depth [dB]
f_c	Frequency of the carrier sound [Hz]
f_m	Modulation frequency [Hz]
L_1	Spatially averaged sound pressure level in source room [dB]
L_2	Spatially averaged sound pressure level in receiving room [dB]
L_{Aeq}	A-weighted equivalent continuous sound pressure level [dB]
L_{AF}	A-weighted sound pressure level with Fast (F) time weighting [dB]
L_{den}	Day-evening-night noise level [dB]
L_{eq}	Equivalent continuous sound pressure level [dB]

L_{i}	Impact sound pressure level [dB]
L_{i} or L_{j}	Reference spectrum value at frequency band i or j [dB]
L_{impact}	Impact sound power level of the tapping machine [dB]
L_{n}	Normalized impact sound pressure level [dB]
$L_{\mathrm{n,w}}$	Weighted normalized impact sound pressure level [dB]
$L_{\mathrm{nT,w}}$	Weighted standardized impact sound pressure level [dB]
m	Modulation index
P	Sound power [W]
p	Sound pressure [Pa]
p_0	Reference sound pressure, 20 μ[Pa]
R	Sound reduction index [dB]
r	Pearson's correlation coefficient
R_{w}	Weighted sound reduction index [dB]
S	Area of the test element [m^2]
SNQ	Single-number quantity
SPL	Sound pressure level [dB]
T	Modulation period [s]
t	Time [s]
T_0	Reference reverberation time, 0.5 [s]
T_2	Reverberation time of the receiving room [s]
V_2	Volume of the receiving room [m^3]

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I   Virjonen P., Hongisto V. and Oliva D. Optimized single-number quantity for rating the airborne sound insulation of constructions: Living sounds. *The Journal of the Acoustical Society of America*, 2016; **140**(6): 4428–4436.

II   Kylliäinen M., Virjonen P. and Hongisto V. Optimized reference spectrum for rating the impact sound insulation of concrete floors. *The Journal of the Acoustical Society of America*, 2019; **145**(1): 407–416.

III   Virjonen P., Hongisto V., Mäkelä M., and Pahikkala T. Optimized single-number quantity for rating the façade sound insulation. *The Journal of the Acoustical Society of America*, 2020; **148**(5): 3107–3116.

IV   Virjonen P., Hongisto V. and Radun J. Annoyance penalty of periodically amplitude-modulated wide-band sound. *The Journal of the Acoustical Society of America*, 2019; **146**(6): 4159–4170.

The list of original publications have been reproduced with the permission of the copyright holders.

1 Introduction

1.1 General introduction

Noise is an unwanted sound that annoys or disturbs. On our everyday life, we encounter various types of noise, for example environmental noise due to different transports (road, rail and air traffic) as well as industrial activity [1] or wind turbine noise [2]. What is experienced as noise, depends on the situation: what is beneficial or delightful sound for some, may be annoying noise for others.

Apart from feeling annoyed, noise may cause various other health effects [3]. The most obvious effect of exposing to high levels of noise is a hearing damage [4] but also lower noise levels may have a negative effect on well-being. According to Haapakangas et al. [5], people's conversations may cause dissatisfaction and impair cognitive performance at the office. Jensen et al. [6] found that neighbour noise annoyance may be strongly associated with physical and mental health symptoms, such as shoulder or neck-pain. Maschke and Niemann [7] found a connection between severe annoyance by neighbour noise and increased health risk in the cardio-vascular system, and increased risk of depression and migraine. Halperin [8] reviewed studies concerning the negative effects of sleep disturbances to health, and highlighted the potential of nocturnal environmental noise to disturb sleep quality.

The emergence of the effects depend on the sound levels and daily dosage of noise which people are exposed to. Brink et al. [9] revealed that the percentage of highly annoyed persons due to road traffic noise indoors increased from 3% to 46% as L_{den} (day-evening-night level) outdoors grew from 30–35 to 75–80 dB. World Health Organization, WHO, recommends for example reducing the noise levels produced by road traffic below 53 dB level (day-evening-night-weighted sound pressure level) as higher levels have been associated with adverse health effects [10]. However, sound level is not the only factor affecting the experienced annoyance. Radun et al. [11] showed that the sound level explained only a fraction of the experienced annoyance caused by wind turbine noise. Other factors were the concern for health effects as well as noise sensitivity, and general attitude towards the sound source.

Countries need to ensure healthy indoor and outdoor environments, and protect their inhabitants from excessive noise. It can be implemented by preventing noise production, blocking noise from spreading, and restricting the land use by zoning areas for specific use, as well as by guiding the building and renovation of apartments and offices. Each country sets its own regulations. A research network report COST Action TU0901 [12] gathered the main requirements for airborne and impact sound insulation rating quantities in 35 European countries (status of June 2013). The report revealed a large diverse of practices, and emphasized the need of harmonization of the quantities. Also the target levels have a large variation: for example the requirements for weighted standardized impact sound pressure level $L'_{nT,w}$ varied from 48 to 65 dB in Europe.

Physical measurable quantities are needed to determine the properties of noise as well as the noise reduction capabilities of different structures. It is important that these quantities do not only depict purely physical aspects of noise or a structure but also correlate well with the subjective perception of noise, and hence, they should be based on scientific evidence gathered with psychoacoustical experiments.

1.2 Background

1.2.1 Sound insulation

Sound insulation of a structure (e.g. a wall, window, floor, or façade) is measured in several one-third octave bands. Standards guide the measurement procedure. ASTM International (formerly the American Society for Testing and Materials) standards are in use in North America, whereas ISO (The International Organization for Standardization) standards are widely used in European and Asian countries. There are some differences between the standards, reviewed by Höller [13]. To simplify the information gathered from several frequency bands, and to enable easier comparison between structures as well as to facilitate the imposition of legislative requirements, single-number quantities (SNQs) have been developed for measuring the sound insulation capability of different structures [14–18].

Psychoacoustic experiments have revealed that the existing SNQs do not always correspond well to the subjective perception. Thus, the order of different structures may not be the same if they are ordered according to their physical SNQ values or the subjective perception towards a sound transmitted through the structures. Several studies have gathered subjective responses towards different living sounds transmitted through various structures, and compared the performances of the existing SNQs such as Vian et al. [19], Tachibana et al. [20], Mortensen [21], Park et al. [22], Hongisto et al. [23], Bailhache et al. [24], Rychtarikova et al. [25], and Monteiro et al. [26]. Also tapping machine noise, Japanese rubber ball and impact type of living sounds have been studied by Jeon et al. [27], Gover et al. [28], Späh et al.

[29], Kylliäinen et al.[30], and Frescura et al. [31]. Vardaxis and Bard conducted reviews on the laboratory experiments studying the subjective responses towards airborne sounds [32] and impact sounds [33] in dwellings. Traffic noise transmitted through façade structures has been studied by Hongisto et al. [34], Bailhache et al. [24], Torija et al. [35], and de la Prida et al. [36].

Reference spectrum takes into account the spectrum of different sound sources. According to Rindel [37], the development of the ISO 717-1 reference spectrum C_{tr} has not been straightforward. The spectrum was combined from two physical measurement sets which were not especially aimed for this purpose. Psychoacoustic experimental evidence was not exploited in the derivation process. This is a serious disadvantage since the standardized SNQs are globally used in such a belief that they link physical one-third octave band measurement data to the subjective perception. However, this is necessarily not the case. The SNQs should be based on reference spectra which rank the structures according to subjective perception. There is very little research on this issue.

Scholl et al. [38] presented the proposal for revision of standard ISO 717, and suggested reducing the number of the standardized SNQs. The frequency ranges of the suggested SNQs have been under debate. The main question has been whether to include the frequency range from 50 to 80 Hz. [32, 33]. For airborne living sounds, the discussion has seen arguments for and against the inclusion, as for impact sounds the general opinion has been more united for the inclusion.

Only a few studies have tried to improve the existing reference spectra for living sounds, such as Park et al. [22], Park and Bradley [39], and for impact sound by Bodlund [40] and Ljungren et al. [41]. Chmelik et al. [42] suggested two new SNQs by extending the reference curve for R_{w} to cover also one-third octave frequency bands from 50 to 80 Hz using 20 dB and 40 dB per octave slopes. It is surprising that mathematical optimization has not been utilized in the selection of the shape of the reference spectra.

The above mentioned studies showed that there is room for improving the existing SNQs. While there exist SNQs that perform well for some sound types, some sound types remain troublesome. Also the question of the relevant frequency range remains. There is a need for a method that enables developing better SNQs that correlate well with the subjective perception, as well as help to find the relevant frequency range.

1.2.2 Amplitude modulation

Sound is called amplitude modulated, if the amplitude of the carrier sound varies with time. Ocean waves reaching the beach is an example of amplitude modulated sounds (AM sounds). Also many living sounds are amplitude modulated (such as walking sounds and speech) rather than being steady-state. The peaks of AM sounds make the sound distinguishable and the peaks may be better associated with the annoyance than the equivalent level of the sound.

Amplitude modulated sounds have been found to cause more annoyance than steady-state sounds. The annoyance caused by AM sounds has been studied for many sound types: wind turbines by Lee et al. [43], Hünerbein et al. [44], Hafke-Dys et al. [45], and Ioannidou et al. [46], wind turbine and generic sounds by Schäffer et al. [47], low-speed engines by Kantarelis and Walker [48], Bradley et al. [49], ventilation by Wang and Novak [50] and unmanned aerial vehicles by Gwak et al. (UAV, drones) [51]. Also traffic flow can be amplitude modulated, studied by Schäffer et al. [52].

National regulations specify the limit values for environmental noise levels. A penalty is added to the measured sound level to take into account the increase in annoyance that a specific sound feature causes. In Finland, tonal and impulsive sounds are sanctioned [53–55], but there does not exist any regulatory penalty for AM sounds. McKenzie et al. [56] have suggested a penalty for wind turbine sounds.

There are several physical properties, that depict the AM sound. Depending on the origin of the sound, and how it is generated, amplitude modulation may be periodic or randomly changing. This thesis focuses on periodic sounds. For periodic AM sounds, the modulation frequency f_{m} determines the fluctuation period. The modulation depth D_{m} depicts the strength of the modulation. In real sounds, the modulation content is complex and sound specific. There may exist several modulation frequencies, and the modulation depth varies according to the modulation frequency and carrier frequency band. To be able to determine the need of a penalty for a periodic AM sound, there is a need for general research that studies the annoyance caused by periodic AM sounds by changing the parameter values of f_{m}, D_{m}, and spectrum of the carrier sound systematically.

1.3 Research questions and contributions

The goal of this thesis was to develop acoustic quantities that correlate well with the subjective perception of noise. Research question Q1 focuses on developing rating quantities for sound insulation, and research question Q2 on studying the need of a penalty for periodically amplitude-modulated wide-band sounds. The detailed research questions are listed below.

1.3.1 Research questions

- Q1 Can physical SNQs for rating sound insulation be developed by mathematical optimization using data from psychoacoustical experiments? This question was divided into sub questions:
  - Q1.1 Is there a reference spectrum, which results in a better correlation between the resulting SNQ and subjective ratings of annoyance against *living sounds transmitted through wall partitions* when compared with the existing SNQs?
  - Q1.2 Is there a reference spectrum, which results in a better correlation between the resulting SNQ and subjective ratings of annoyance against *impact sounds transmitted through concrete floors with various coverings* when compared with the existing SNQs?
  - Q1.3 Is there a reference spectrum, which results in a better correlation between the resulting SNQ and subjective ratings of annoyance against *road traffic sounds transmitted through a façade* when compared with the existing SNQs?
  - Q1.4 Out of the four frequency ranges presented by the standard ISO 717-1, can the most suitable frequency range be selected for a reference spectrum?
  - Q1.5 Can the performance of the resulting optimized SNQ be evaluated?
- Q2 Is there a need for a penalty for periodically amplitude modulated wide-band low level sounds? If so, could a constant penalty be used?

Publication I addresses the research question Q1.1, Publication II the research question Q1.2, Publication III the research questions Q1.3, Q1.4, and Q1.5, and Publication IV the research question Q2.

1.3.2 Publications

A short introduction of the publications included in this thesis is given below:

Publication I focuses on living sounds and airborne sound insulation of partitions between dwellings. Optimal reference spectra were sought for six common living sound types using data from a psychoacoustic experiment [23]. The data included loudness, disturbance and acceptability ratings from 59 participants for the sounds, filtered through nine common wall structure filters with varying airborne sound insulation capabilities. A general reference spectrum was also suggested.

Publication II focuses on impact sounds and impact sound insulation of floors between dwellings. Optimal reference spectra were sought for five common impact sound types using data from a psychoacoustic experiment [30]. The data included loudness, annoyance and acceptability ratings from 55 participants for the sounds, filtered through nine common floor type filters with varying impact sound insulation capabilities. A general reference spectrum was also suggested.

Publication III focuses on road traffic noise and façade sound insulation. Optimal reference spectra were sought for road traffic sounds with five spectrum alternatives using data from a psychoacoustic experiment [34]. The data included loudness and annoyance ratings from 43 participants for the sounds, filtered through twelve common façade structure type filters with varying airborne sound insulation capabilities. A general reference spectrum was also sought.

In Publication IV, a psychoacoustic experiment was conducted to study the possible need of a penalty for periodic AM sounds with two different spectra. The artificial test sounds included both steady sounds as reference sounds and AM sounds. Forty participants rated the loudness and annoyance of the experimental sounds. Penalties were determined at different modulation frequencies and modulation depths.

1.3.3 Contributions

Authors' contributions are specified according to CRediT – Contributor Roles Taxonomy [57].

Publication I

Petra Virjonen: Conceptualization (equal), Methodology (lead), Software (lead), Validation (lead), Formal Analysis (lead), Data curation (lead), Writing – Original Draft (equal), Writing – Review & Editing (equal), Visualization (lead)

Valtteri Hongisto: Conceptualization (equal), Formal Analysis (supporting), Writing – Original Draft (equal), Writing – Review & Editing (equal), Supervision (lead), Project Administration (lead), Funding acquisition (lead)

David Oliva: Writing – Original Draft (supporting)

Publication II

Mikko Kylliäinen: Conceptualization (supporting), Formal Analysis (supporting), Writing – Original Draft (equal), Writing – Review & Editing (lead), Visualization (equal), Funding acquisition (supporting)

Petra Virjonen: Conceptualization (equal), Methodology (lead), Software (lead), Validation (lead), Formal Analysis (lead), Data curation (lead), Writing – Original Draft (equal), Writing – Review & Editing (supporting), Visualization (equal)

Valtteri Hongisto: Conceptualization (equal), Methodology (supporting), Formal Analysis (supporting), Writing – Original Draft (equal), Writing – Review & Editing (supporting), Supervision (lead), Project Administration (lead), Funding acquisition (lead)

Publication III

Petra Virjonen: Conceptualization (lead), Methodology (lead), Software (lead), Validation (lead), Formal Analysis (lead), Data curation (lead), Writing – Original Draft (lead), Writing – Review & Editing (equal), Visualization (lead)

Valtteri Hongisto: Formal Analysis (supporting), Writing – Original Draft (supporting), Writing – Review & Editing (equal), Supervision (lead), Project Administration (lead), Funding acquisition (lead)

Marko Mäkelä: Methodology (supporting), Writing – Original Draft (supporting), Writing – Review & Editing (supporting), Funding acquisition (supporting)

Tapio Pahikkala: Methodology (supporting), Writing – Original Draft (supporting), Writing – Review & Editing (supporting), Funding acquisition (supporting)

Publication IV

Petra Virjonen: Conceptualization (supporting), Methodology (equal), Software (lead), Validation (lead), Formal Analysis (equal), Investigation (lead), Writing – Original Draft (equal), Writing – Review & Editing (lead), Visualization (lead)

Valtteri Hongisto: Conceptualization (lead), Formal Analysis (equal), Writing – Original Draft (equal), Writing – Review & Editing (supporting), Supervision (lead), Project Administration (lead), Funding acquisition (lead)

Jenni Radun: Conceptualization (supporting), Methodology (equal), Formal Analysis (equal), Writing – Original Draft (supporting), Writing – Review & Editing (supporting)

1.4 Organization of the thesis

This thesis consist of two parts. The first part (Chapters 1–6) introduces the topic, its background and motivation, and provides information on the main concepts needed to understand the work. The main findings of the research questions and the limitations are discussed. The second part gathers the four original publications included in this thesis.

The subsequent chapters are organized in the following way: Chapter 2 deals with psychoacoustic experiments, and the procedures involved in their conduction. Chapter 3 explains the concepts of airborne, impact and façade sound insulation as well as the procedure of optimizing a reference spectrum of a SNQ for rating sound insulation. Chapter 4 deals with periodic amplitude modulation of wide-band sounds. Chapter 5 presents the results concerning the research questions. In Chapter 6, the results and their applicability are discussed. Chapter 7 concludes the thesis.

2 Psychoacoustic experiments

Psychoacoustics studies the relationship between auditory perception and physical variables [58]. One of the earliest publications on psychoacoustics was made by Gustav Fechner [59]. He introduced psycho-physical methods and relationships. Nowadays psychoacoustics can be exploited in various applications within acoustics, for example for evaluation of sound quality by rating the pleasantness of a sound [60]. For the present study, the scope lies in the applications of evaluating noise with psychoacoustic experiments, for example by rating the sound by how loud, annoying or disturbing it is.

Publications I–III exploited data from published psychoacoustic experiments (Hongisto et al. [23], Hongisto et al. [34], Kylliäinen et al. [30]). In Publication IV, a psychoacoustic experiment was conducted and analyzed to derive penalties for periodic amplitude modulated wide-band sound. It is essential to understand, how the data was gathered in all of these experiments to interpret the results as well as to understand their applicability. There are various ways to conduct a psychoacoustic experiment, and the choices made in the experiment procedure affect the results and ultimately the generalization power of the results to a larger population outside the group of participants. This chapter explains these procedures. The goal is not to list all available methods but rather to illuminate the choices made in the experiment of Publication IV and the above mentioned studies whose data was exploited in this thesis. The choices have mainly conformed with procedures utilized in other psychoacoustic experiments found in literature. Examples of differing methods are also presented.

2.1 Participants

2.1.1 Recruiting

In all of the experiments [23, 30, 34] and Publication IV, the participants were voluntary, sought by e-mailing lists and advertisements on news for students' web pages. The participants were mainly students from University of Turku and Turku University of Applied Sciences. The participants came from varying disciplines and were likely unfamiliar with the psychoacoustic experiment procedures. Bailhache et al. [24] used workers from the same institute, most of them unfamiliar with acoustics and psychoacoustic experiments. Ioannidou et al. [46] recruited students from Technical University. Vian et al. [19] recruited participants via a newspaper advertisement. Monteiro et al. [26] conducted a psychoacoustic experiment in two countries. They had both inexperienced participants, participants with background in acoustics as well as participants, who had participated listening tests before.

2.1.2 Age

According to Morrell et al. [61], hearing starts to slowly degrade approximately after 40 years of age. The effects of aging on hearing vary largely between individuals. The sounds may become attenuated, it may be difficult to perceive the direction of the sound, and speech recognition and speech discrimination against background noise may become difficult. The degradation effect is more prominent for high frequencies, whereas the changes of hearing at low frequencies are more subtle. The age range of the participants was 18–59 in the experiment of [23], 20–57 in the experiment of [30], 21–50 in the experiment of [34], and 20–39 in the experiment of Publication IV. This is a rather limited selection of ages, when compared to the whole population, which does produce some bias. However, the frequency content of the studied sounds (living sounds, road traffic sounds heard through a wall or floor) is focused on low and middle frequencies where the difference between the hearing is not that prominent between young and old people. The same kind of age range has been applied in most psychoacoustic experiments, e.g. Bailhache et al. [24] (age range 20–59). Lee et al. [43] had somewhat smaller age range, 20–30 years.

2.1.3 Sample size

Psychoacoustic experiments are laborious and time-consuming to arrange, which limits the number of included participants. If the experimental sounds are not loud, only one participant can take part in the experiment in the same test space at a time as other participants might cause some background noise. It is beneficial to have a large number of participants to get a descriptive mean value of the response as the variation is typically rather large, due to noise sensitivity, and other subjective features. If some descriptive model is to be derived from the data, accurate performance estimation of the model also requires a sufficient number of data. (Model performance estimation is explained more detail in Chapter 3.) The total number of participants was 43–59 in the psychoacoustic experiments, whose data were used in Publications I–III. In Publication IV, 40 participants took part in the experiment.

Women were on average more keen to attend the tests, on average 64% of the participants were women. In recent studies, the typical number of participants has been approximately between 20–40. There are also studies having a fewer number of participants: Tachibana et al. [20] had only eight participants. Also larger samples have been used: de la Prida [36] had 119 participants, and Torija et al. [35] had 100 participants in their experimental session.

2.1.4 Preconditions

In the experiments of [23], [30], and [34], the participants were sought with normal hearing ability, Finnish native language, and currently residing in a multi-storey apartment. The residing condition was not included in the Publication IV as the appearance of AM sounds is not restricted to multi-storey apartments. Sometimes the preconditions are not revealed in publications, even though they do confine the group of people for whom the results are applicable at some level. Monteiro et al. [26] conducted the experiments in two countries, as mentioned before. In Spain, the guidance was given with native language Spanish for all participants. In Belgium, English was used as there were several native languages among the participants.

2.1.5 Hearing

Normal hearing was one of the requirements set for the participants, and it was mentioned in the recruitment advertisement. The hearing ability of each participant was tested before attending the experiment to check if there existed any hearing loss the participant might be unaware. All of the participants had a normal hearing, that is, the reference equivalent threshold sound pressure level was below 20 dB in each octave frequency band studied (250–4000 Hz). Normal hearing has been required in most other studies as well. A typical requirement of normal hearing has been within 15–20 dB from the reference equivalent threshold sound pressure level. Rychtarikova et al. [25] did not require normal hearing, and did not conduct a hearing ability test but relied on the knowledge of the participants on their hearing. They included a few participants with hearing aids as well.

2.2 Conducting the experiment

2.2.1 Training

The participants were untrained, and thus, an orientation phase and a rehearsal phase were conducted prior to the experiment. A subgroup of the experimental sounds were presented in the orientation phase to familiarize the participants with the range and type of the experimental sounds. In the rehearsal phase the participants were instructed how to rate the sounds, and the procedure was rehearsed. In most of the recent studies, a training session prior to the actual experiment sound rating has been performed, but also the phase has been excluded for example by Bailhache et al. [24].

2.2.2 Sound reproduction

It is essential to be able to accurately reproduce the experimental sounds and their target levels. Also it is important to ensure that the reproduced sound spectrum and levels are equal for each participant. Thus, sound reproduction and thorough verification of the desired levels are an important part of a psychoacoustic experiment set-up. With loudspeaker reproduction, the listening situation is more natural compared to headphones. To ensure the same sound level and spectrum for each participant, the listening location has to be well defined and stationary. The location of the head varies due to different heights, seating postures and movements of the participants. With headphones, this does not pose a problem, but the listening situation may seem less natural. Also different shapes of ears and adjustment of the headphones color the incoming sound in a different way, which causes some differences between the participants. Individual equalization of the headphones would be preferable [62], but infeasible for such a large group of participants. The identification of the spacial characteristics may be less accurate for loudspeaker than headphone listening [63]. A possible difference between the reproduction methods is not crucial as such as the purpose was to observe the differences between the experimental sounds and structures.

In each of the psychoacoustic experiments, the sound level and spectrum were adjusted with an iterative process to ensure as accurate output as possible. In experiments [23, 30, 34], the reproduction of the experimental sounds was made using loudspeakers. In Publication IV, headphones were used (Figure 1 and Figure 2). Loudspeakers have been the choice in most of the recent psychoacoustic experiments, but also headphone reproduction has been used by for example Rychtarikova et al., Monteiro et al., and de la Prida et al. [25, 26, 36].

Figure 1. Participants performed the psychoacoustic experiment in a special listening room. The sound reproduction was made through open back headphones in the AM experiment (Publication IV). Picture courtesy of Turku University of Applied Sciences/Laboratory of Built Environment.

2.2.3 Rating

The choice of the rating parameter is not necessarily obvious. How should the subjective perception of a sound be rated? Sound can cause annoyance, disturbance, or it can be perceived loud or it can impede concentration. How acceptable is the sound, in regards with the listening situation? Hongisto et al. [23] rated the loudness, disturbance, and acceptability against airborne living sounds. Kylliäinen et al. [30] studied the loudness, annoyance, and acceptability of impact sounds. Hongisto et al. [34] studied the loudness and annoyance of road traffic sounds. In each of the optimization tasks in Publications I–III, the rated annoyance (or disturbance) was chosen as the parameter for optimizing the reference spectra, because it is the most common health effect from noise.

There are also different methods of rating. There is no perfect method for every task, but the selection of the deployed method depends on the aim of the study. In experiments [23, 30, 34] and Publication IV, annoyance and loudness ratings were made as direct rating (Figure 3). The participant was to select a number corresponding the experienced annoyance towards the ongoing sound from integer values ranging from 0 to 10 with extremes labeled as Not at all – Very much. This means that the participant should rate the sound in proportion to the sounds included in the experiment, that is, to a predefined range of sounds. For this reason, the rehearsal phase is essential before the actual rating phase so that the participant is familiar with the range when the experiment begins. The downside of direct rating is that the extreme values are more rarely used than the center ones. Each participant uses the rating range in their own way.

Mortensen [21] used a line with the end points named as Not annoying / Very

Figure 2. The verification of the experiment sound levels was made with Brüel&Kjær Head and Torso Simulator 4100 in the AM experiment (Publication IV). Picture courtesy of Turku University of Applied Sciences/Laboratory of Built Environment.

annoying, and the participant was to set a mark on the line according to his/her evaluation.

Torija et al. [35] used a free-number magnitude estimation. They wanted to avoid predefined answering score, so the participants were asked to assess a number corresponding their annoyance, relative to the previous sound.

Forced paired comparison is a rating method, where two sounds are presented one after another, and the participant needs to choose which of them is more annoying. For example Rychtarikova et al. [25] used this rating by asking which one of the played sounds were perceived louder. In paired comparison, participant does not need to remember any other sounds or sound range to rate the sounds. However, with paired comparison, ordering of the sounds may not be systematic, and the number of different sounds listened is more limited compared to direct rating, as the pairs are usually listened in both orders, sound1-sound2, and sound2-sound1. There are different versions of paired comparison, varying on the evaluation method. A compromise can be found between the accuracy and the time needed to perform the experiment [64].

Tachibana et al. [20] used the method of adjusting by subject. In this method, the participant adjusts the magnitude of the experiment sound to match with another experiment sound called standard stimulus until the sounds are perceived equally loud.

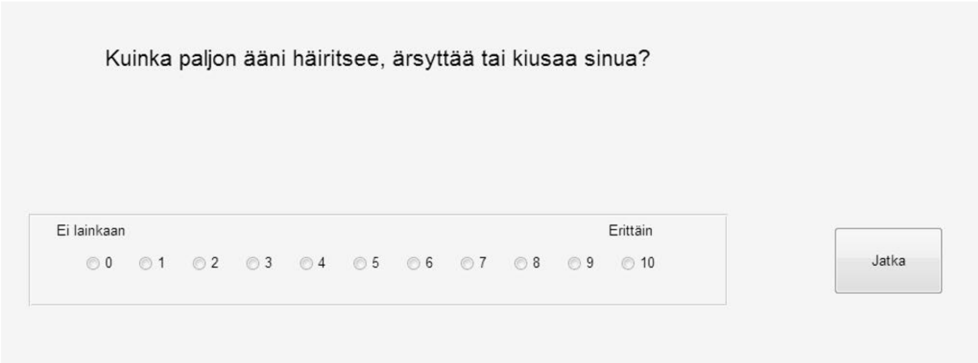

Figure 3. The scale for rating the annoyance in the AM experiment (Publication IV). The question was: How much does the sound bother, disturb, or annoy you? The extremes were labeled as Not at all / Very much. Picture courtesy of Turku University of Applied Sciences/Laboratory of Built Environment.

2.2.4 Listening order

To remove any order effect, the experimental sounds were played in different orders to the participants. However, in each of the experiments [23, 30, 34] and Publication IV, loudness was rated first, and annoyance after that. For annoyance rating, an imaginary listening situation was given (Figure 4). The participants were instructed to imagine being at home, relaxing, eating or reading a newspaper or surfing the Internet, while hearing the sound. This cue would have been difficult to ignore if it were given before the loudness phase, were the instruction was plainly to rate the loudness of the sound without any associations to a specific listening situation.

2.2.5 Stimulus time

In experiments [23, 30, 34], the participants were forced to listen to the experiment sound 15–20 seconds before the rating was enabled. After enabling the rating, the sound continued playing until the rating was finished. In Publication IV, the participants were forced to listen to the sound eight seconds before enabling the rating window. In this case the sounds were periodic and several periods were played within eight seconds so a shorter stimulus time was deemed adequate. In practice, the participants rated each sound in 30 seconds. Mortensen and Torija et al. [21, 35] had a distinctly longer stimulus time, 2 and 10 minutes, respectively. Bailhache et al. [24] permitted the repetition of the sound samples as many times as the participant needed. Their sound samples were 5–38 seconds long.

Figure 4. The example picture of a relaxing place shown to the participants to orientate for the AM experiment (Publication IV) in the annoyance rating phase. Picture courtesy of Juhani Helavuori.

2.2.6 Level of the experimental sounds and background noise

The level of the experimental sounds should correspond to the normal level as Rychtarikova et al. [25] pointed out. In their study, the adequacy of the studied SNQ depended on the experiment sound level, and the studied SNQ performed well only when the experiment sound levels were adjusted unrealistically loud. In the experiments whose data were exploited for optimization [23, 30, 34], the experiment sound levels were adjusted to normal levels, that is, reasonable for each sound type. In Publication IV, the A-weighted equivalent sound pressure level (SPL) was set to 35 dB, which is within scope when considering the limit values for environmental noise both indoors and in residential yards [53, 54]. (A-weighting is explained in Chapter 4.3). It is also important that the range of the sounds corresponds to the range that is of interest, as the least annoying and most annoying (or whatever is the studied parameter) sounds set the limits for all sounds to be rated, when direct rating is deployed. This means that if the experimental sounds included many low-level sounds and one high-level sound, it would be difficult to find a difference between the low-level sounds as the high-level sound sets the rating range so wide.

Subjective experience depends also on how well the experimental sounds can be perceived. If there is a high background noise level, the experiment sound will be masked by the background noise. With low background noise level, the sound can be perceived more clearly. The background noise level in the experiment room should correspond to ordinary levels for the actual situation. In experiments [23, 30, 34], the A-weighted background noise level was 20–23 dB (Figure 5). According to Kylliäinen et al. [65, see Figure 8], this roughly corresponds to the normal sound level for Finnish living rooms. In other studies, the A-weighted background noise level has been a bit higher, 35 dB [22, 24]. On the other hand, Monteiro et al. [26], conducted their experiments in semi-anechoic rooms with very low background noise conditions (Spain <17 dB, Belgium <0 dB). Rychtarikova et al. [25] conducted their experiment in anechoic room with low background noise level (exact level not reported). This facilitates more accurate comparison between structures as the experimental sounds are perceived more clearly but such low background noise levels do not correspond normal living environments.

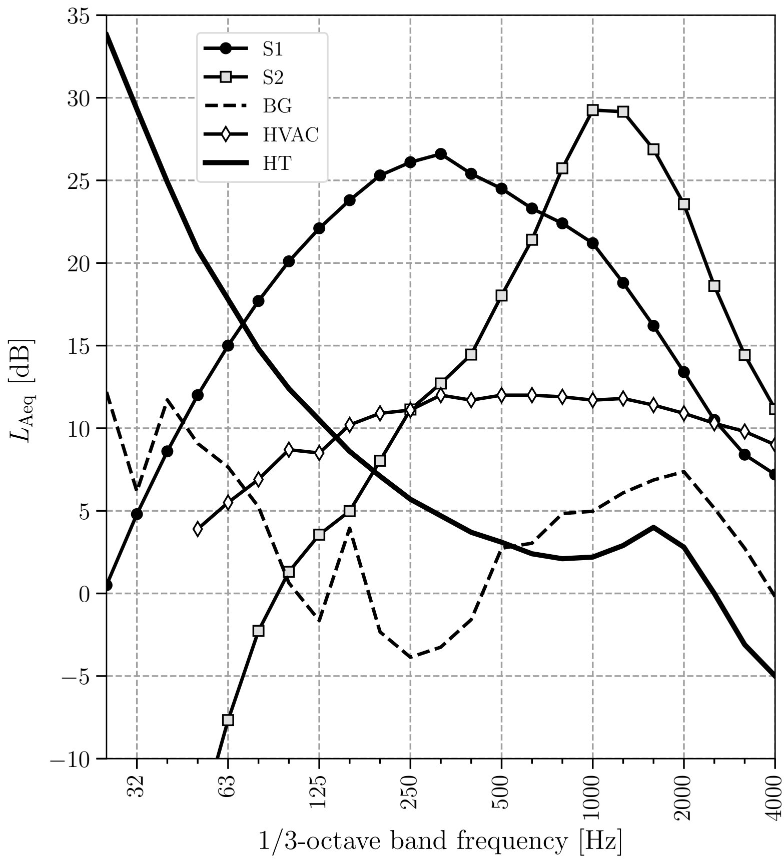

Figure 5. A-weighted equivalent continuous SPLs in the AM experiment (Publication IV, see Figure 2). S1 = experimental sounds, wind turbine spectrum L_{Aeq} = 35 dB, S2 = experimental sounds, road traffic spectrum L_{Aeq} = 35 dB, BG = background noise in the experiment L_{Aeq} = 21 dB, HVAC = average background noise level from HVAC systems in Finnish dwellings [65] L_{Aeq} = 23 dB, HT = hearing threshold according to ISO 226 [66].

3 Optimizing a reference spectrum for sound insulation

This chapter introduces the concepts of sound insulation, and describes the methodology of developing the reference spectra for different sound types using mathematical optimization and psychoacoustic experiment data. Also the uncertainty estimation of the results is explained.

Sound insulation depicts, how well a building element (e.g. wall, floor, ceiling, window, door, shutter, some technical element or a structure consisting of several parts such as façade) attenuates sound. An element with a high sound insulation capability transmits only a small proportion of the sound incident to the surface of the element. Sound can be transmitted into an element by air (airborne sound insulation) or mechanically (impact sound insulation).

3.1 Measuring sound insulation

Sound insulation of a building element can be measured in a laboratory or in a real building. In laboratory, the structure under test is mounted in a special measurement opening. The mounting and the surrounding structures are constructed in such way that the flanking sound is minimized, and the measured sound insulation accurately depicts the sound insulation capability of the tested element alone. In field conditions, flanking sound influences the result. Flanking sound is sound that enters the receiving room through other paths than straight through the element under test, and thus, the measurement indicates lower sound insulation for the element, compared with a sound insulation measurement in a laboratory.

3.1.1 Airborne sound insulation

Airborne sound insulation depicts the capability of a structure to dampen sound that propagates through air. Typical airborne sounds in domestic environments are speech, television and music sounds, traffic noise, and animal sounds [67, 68].

Sound reduction index R of an element can be measured using two adjacent rooms ("source room" and "receiving room"), separated by the element under test. Sound reduction index R is determined as

$$R = 10 \lg \frac{P_1}{P_2} \text{ dB}, \tag{1}$$

where P_1 is the sound power incident on the element and P_2 is the sound power transmitted into the receiving room. In a diffuse sound field, the distribution of sound energy is uniform. If a simplifying assumption of diffuse sound field is made in both source room and receiving room, sound reduction index R can be presented by

$$R = L_1 - L_2 + 10 \lg \frac{S}{A_2} \text{ dB}, \tag{2}$$

where L_1 is the spatially averaged SPL in the source room and L_2 the spatially averaged SPL in the receiving room, and S is the area of the test element. Absorption area of the receiving room, A_2, can be determined from

$$A_2 = \frac{55.3 V_2}{c T_2}, \tag{3}$$

where V_2 is the volume of the receiving room, c is the speed of sound and T_2 is the reverberation time of the receiving room. The measurement is conducted by producing a pink noise signal into the source room through loudspeakers, and measuring the resulting energy-average SPL in the source and the receiving room.

ISO standard 10140-2 guides the airborne sound insulation measurement procedure for laboratory measurements [69], and ISO 16283-1 and ISO 16283-3 for field measurements [70, 71]. Sound reduction index is measured in one-third octave bands from 100 to 3150 Hz. The measurement frequency range can be enlarged to cover the frequency range from 50 to 5000 Hz. Measuring the low frequencies requires a different measurement procedure as the variation at low frequencies is very high due to low number of room nodes [72, 73].

In field conditions, the area of the common separating element may be difficult to determine, for example with staggered rooms. Instead of sound reduction index R, level difference values may be used $D = L_1 - L_2$. Level difference may be standardized to a reference value corresponding a typical dwelling room ($T_0 = 0.5$ s) or normalized to a typical reference absorption area ($A_0 = 10$ m^2).

3.1.2 Impact sound insulation

Impact sound insulation depicts the capability of an element to dampen the noise caused by structure-borne sound source, for example a walking person. Impact SPL L_i is the energy-average SPL measured in the receiving room when the floor under test is excited by the standardized impact source. Standard ISO 10140-3 [74] defines the procedure for laboratory measurements and standard ISO 16283-3 [71] for field measurements. The standard tapping machine is used as the impact sound source, and it is supposed to mimic a walking person with shoes. The machine drops five hammers one at a time with 10 Hz frequency. The standards mention also a heavy rubber ball, which may be used to mimic impact sources with strong low-frequency content, such as a bare-foot walking person or children jumping.

Normalized impact SPL L_n can be calculated from

$$L_{\mathrm{n}} = L_i + 10\lg\frac{A}{A_0}\ \mathrm{dB}, \tag{4}$$

where A is the measured equivalent absorption area of the receiving room, and A_0 is the reference absorption area (10 m^2). Several impact sound source positions as well as receiver positions are used to attain a reliable mean value of the impact SPLs at the receiver room. A large value of L_{n} corresponds low impact sound insulation, in contrary to airborne sound insulation, where a large value of sound reduction index R corresponds to high sound insulation.

According to standard ISO 10140-3, L_n is measured in one-third octave bands from 100 to 5000 Hz, and optionally down to 50 Hz.

3.1.3 Single-number quantities

A single-number quantity simplifies the sound insulation information from several frequency bands into a single number, which characterizes the acoustic performance of the element under study. This enables easier comparison between different structures and facilitates posing of legal requirements. There are different ways of calculating SNQs.

Comparing to a reference curve

Standard ISO 717-1 [14] introduces a method for determining SNQs for rating airborne sound insulation in buildings and of building elements. The sound insulation values acquired with measurements according to ISO 10140-2, ISO 16283-1 or ISO

16283-3 at one-third octave bands from 100 to 3150 Hz are compared with reference values. The reference curve is shifted towards the measurement curve with 1-dB steps until the sum of unfavourable deviations is as high as possible but at most 32 dB. Unfavourable deviation means that the measured value is less than the reference value at that particular frequency band. The value of the SNQ is the value of the reference curve at 500 Hz frequency band. The resulting SNQ calculated using the measured sound reduction indices R is called the weighted sound reduction index R_w. An example of determination of R_w is shown in Figure 6.

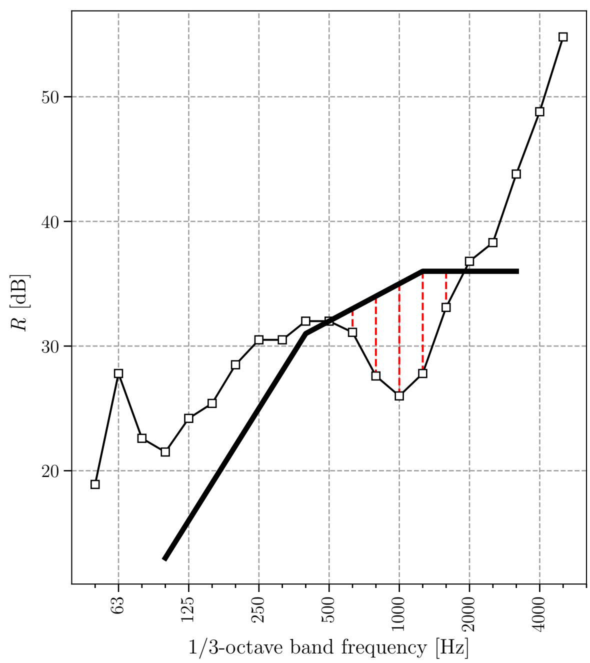

Figure 6. An example of airborne sound insulation measurement result and determination of the weighted sound reduction index R_w. The measured sound reduction indices at 1/3-octave frequency bands are marked with squares. The reference curve from ISO 717-1 is marked with a solid line. The unfavourable deviations are marked with red dashed line. The sum of the unfavourable deviations is 28.4 dB $\leqslant$ 32.0 dB. The weighted sound reduction index R_w is the value of the reference curve at 500 Hz frequency band (R_w = 32 dB).

Standard ISO 717-2 [15] introduces a similar method for impact sound insulation in buildings and of building elements. The impact SPLs acquired with measurements according to ISO 10140-3 or ISO 16283-2 at one-third octave bands from 100 to 3150 Hz are compared with reference values. The reference curve is shifted towards the measurement curve with 1-dB steps until the sum of unfavourable deviations is as high as possible but at most 32 dB. Unfavourable deviation means that the measured value is higher than the reference value at that particular frequency band. Again, the value of the SNQ is the value of the reference curve at 500 Hz frequency band.

Spectrum adaptation terms

A spectrum adaptation term can be added to the SNQ calculated with the comparison method to take into account the characteristics of a particular sound spectrum.

Standard ISO 717-1 supplies two A-weighted reference spectra for calculating the spectrum adaptation terms. The first one, C, is guided to be used with living activities (talking, music, radio, TV, children playing), and highway road traffic at speeds higher than 80 km/h and factory emission noise (medium and high frequency noise emissions). The second one, C_{tr}, is intended for use with urban road traffic, disco music, and factory emission noise (low and medium frequency noise emissions). The SNQ value is determined by calculating the difference between the A-weighted SPLs on the source side and on the receiving side of the separating element:

$$SNQ = 10\lg \frac{\sum_{j=K_1}^{K_2} 10^{L_j/10}}{\sum_{j=K_1}^{K_2} 10^{(L_j-R_j)/10}}, \tag{5}$$

where L_j is the reference spectrum value at frequency band j, R_j is the sound reduction index at frequency band j, and K_1, and K_2 determine the deployed frequency band range. The frequency range can be the normal range, 100–3150 Hz, or an extended range: 50–3150 Hz, 50–5000 Hz or 100–5000 Hz. The reference spectrum is normalized, that is, $\lg \sum_{j=K_1}^{K_2} 10^{L_j/10} = 0$ dB.

According to ISO 717-2 [15], weighted normalized impact SPL $L_{\mathrm{n,w}}$ does not work as well with all floor types, such as timber joist floors. Due to this, standard ISO 717-2 introduces an adaptation term C_{I} in an informative Annex. Adaptation term C_{I} can be determined from

$$C_{\mathrm{I}} = 10\lg \sum_{j=K_1}^{K_2} 10^{L_{\mathrm{n},j}/10} - 15\ \mathrm{dB} - L_{\mathrm{n,w}}. \tag{6}$$

Adaptation term C_{I} is defined using frequency range 100–2500 Hz. It is noted in the standard that the calculations of the spectrum adaptation term may additionally be carried out for an enlarged frequency range including also one-third octave band frequencies 50–80 Hz. This quantity, $C_{\mathrm{I,50\text{–}2500}}$, has been adopted in legal requirements in Sweden and Finland [12, 54].

Scholl et al. [38] suggested impact sound reduction index, which is calculated in the same way as airborne sound reduction index in Equation (5). In this case the reference spectrum values are given as

$$L_{\text{impact},j} = 82.1 + 10\lg\frac{f_j}{1\ \text{Hz}}, \tag{7}$$

where f_j is the centre frequency for the one-third octave band j. The reference spectrum is normalized to the total impact power of the tapping machine $\lg\sum_{j=K_1}^{K_2} 10^{L_{\text{impact},j}/10} = 122.9$ dB (for frequency range 50–2500 Hz). Indices K_1, and K_2 determine the deployed frequency band range.

3.2 Using mathematical optimization to derive a reference spectrum

The sound reduction index curve may be differently shaped for two structures, and yet they can possess the same SNQ. Physical SNQs should be able to order the structures in a meaningful way. Often this is not the case as discovered by for example Rychtarikova et al. [25]. The order of the structures should be the same regardless whether they are ordered according to their physical SNQ or the subjective perception towards the sounds transmitting through the structures. To reach a better conformance between the SNQs and the subjective perception, the reference spectrum can be adjusted accordingly. Earlier attempts to improve the reference spectra [22, 39–42] have relied on predefined alternatives derived using experience and small modifications on the present system. Rather to test some predefined alternatives for the reference spectrum, mathematical optimization can be deployed. Equation (5) is in explicit form and thus, it is more suitable for optimization purposes than the procedure of shifting the reference curve. To deploy mathematical optimization, psychoacoustic data are needed, more precisely, a representative number of subjective ratings for sounds transmitted through various structures. Optimization of the reference spectrum resulted in a nonlinear optimization problem with constraints. The optimization procedure is depicted in detail in Publications I, II and III. The choices made in the procedure are discussed below.

3.2.1 Objective function

In mathematical optimization problem setting, the first thing is to determine an objective function, which is to be minimized or maximized. For the case of optimizing the reference spectrum, it was assumed that the mean value of the subjective annoyance ratings decreases linearly as the SNQ value increases. This seemed fair according to the subjective data available [23, 30, 34]. The objective was to minimize the sum of the squared residuals (i.e. the differences between the actual mean subjective ratings and the predicted values for the SNQ). This resulted in nonlinear optimization problem, as the SNQ is a nonlinear function of the reference spectrum L according to Equation (5). Figure 7 shows an example of the outcome of the optimization process.

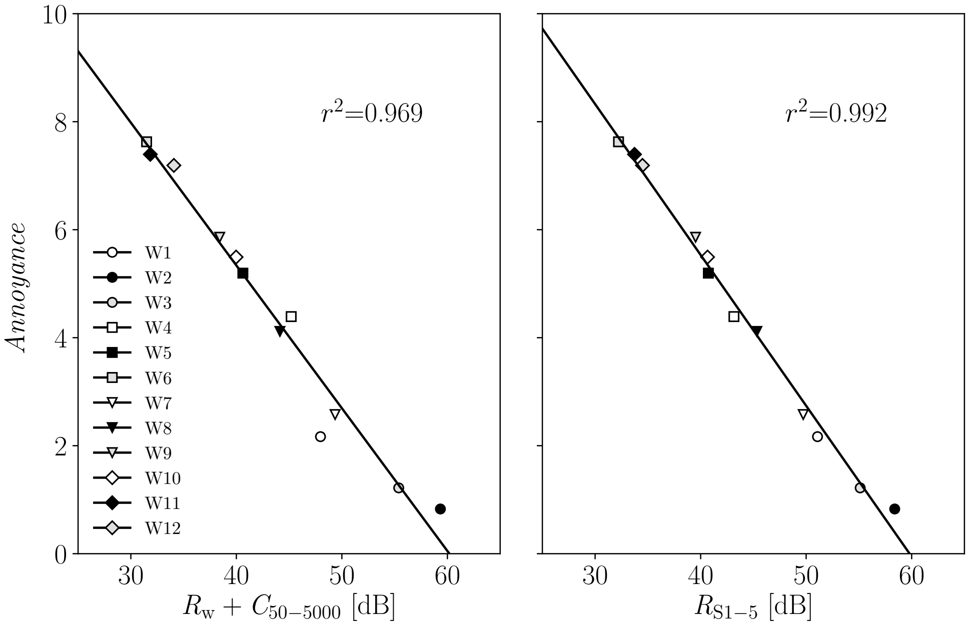

Figure 7. The algorithm finds suitable values for the reference spectrum so that the fit between the mean annoyance ratings and resulting SNQs are optimal. Example of the outcome for the aggregate sound type S1–5 (derived using the mean annoyance ratings over all sound types S1 ... S5) from Publication III (see Figure 5). On left: The best performing existing SNQ for the aggregate sound type S1–5, R_w+$C_{50–5000}$. On the right: Optimized SNQ $R_{S1–5}$. The optimized SNQ orders façades W1 and W4 in a different order than R_w+$C_{50–5000}$.

Hongisto et al. [34] presented the squared Pearson's correlation coefficient between the subjective ratings and different SNQs in two ways: by using all ratings given by the participants as well as by using the mean value of the ratings for each façade structure. In Publication III, the reference spectrum values were also optimized using the former way, which produced in practice the same results as using the latter way.

3.2.2 Constraints

The reference spectrum values should be bounded to some reasonable range (called feasible region), and thus, constraints were introduced. Also it is preferable to pursue a rather smooth solution, which will probably generalize better than a strongly varying reference spectrum. A condition was set to limit the difference between the adjacent one-third octave bands to 5 dB at maximum in Publications I, II, and to 3 dB in Publication III. The decision of the maximum limit was somewhat arbitrary, but the value of the limit was of same order of magnitude as in reference spectra for C and C_{tr} from ISO 717-1. Yet another constraint was introduced, specifically normalization of the reference spectrum to 0 dB for airborne sound insulation and to 123 dB for impact sound insulation. This resulted in a nonlinear constraint.

3.2.3 Solving the constrained nonlinear optimization problem

Nonlinear optimization problems are often impossible to solve analytically, so numerical methods have been developed to approximate the solution. Sequential Quadratic Programming, SQP, is meant for solving constrained nonlinear optimization problems, whose objective function and the constraints are twice continuously differentiable [75, p. 576–589]. The algorithm starts from some initial point and iteratively improves the solution. The algorithm solves the search direction by converting the problem into quadratic subproblem with linear constraints. Then the step size is chosen using a line search procedure. The algorithm proceeds as long as the step produces a solution that is improved more than the predefined tolerance value or the number of the maximum iterations is reached. SQP-algorithm operates on the feasible area, which means that the solution is in every step within the boundaries set by the constraints. SQP-algorithm is one of the most popular methods for the numerical solution of constrained nonlinear optimization problems. It has been implemented in various environments including Matlab [76] and Python [77], which were both utilized in this thesis.

3.3 Selecting the best model

A model is the learning outcome when an algorithm has been applied to data. A model can be used to make predictions for future inputs. Here, we wanted to develop a model, that accurately predicts the SNQ values for structures if their (airborne or impact) sound insulation values are known, so that the resulting SNQs will order the structures in a reasonable way according to the subjective annoyance towards sounds transmitted through the structures. If there are different parameter choices available in the model derivation, the best model should be selected based on its performance. The performance of a model can be estimated by dividing the data into separate data sets: training data to train the model and test data to test the trained model. If the data is not very large, a good way is to use cross validation (CV) [78]. In CV, the data is divided into several folds, and the performance estimation is repeated so many times that each fold has acted once as the test data, and the remaining folds as the training data. The final estimation of the performance is given as the average of the performance estimations of the folds. The most accurate way is to have only one datum left as the test data, and use the rest as the training data. This is called leave-one-out cross validation (LOOCV) (Figure 8). This way, most of the data can be used for training, and the resulting performance value does not depend on the split of the data but it is deterministic. The downside is that the calculation time increases significantly as the size of the data increases.

The squared Pearson's coefficient was used as the measure of prediction performance in Publication III. If the model predicted such SNQ values that the ratings of the participant forming the test data would fit well with the linear line fitted to the mean annoyance ratings given by the participants included in the training data, and the predicted SNQs, the model received high performance estimation.

In addition to finding the best values for the reference spectrum, the best frequency range is also of interest. This question was tackled in Publication III. The choice of frequency range was included as a hyper parameter in the model selection. It means that the optimal reference spectrum was calculated using several frequency ranges, and the performance of each resulting model was estimated using LOOCV. The frequency range that the best performing model used, was selected as the best. In this case, the main interest was to offer knowledge on the differences between the frequency ranges included in ISO 717-1 [14]: 100–3150 Hz, 100–5000 Hz, 50–3150 Hz, and 50–5000 Hz.

3.4 Estimating the uncertainty of the model

How well can we trust the model that was derived from the data available? There are several aspects that influence the ”correctness” of the resulting model. Has the optimization algorithm converged to a local minimum? How sensitive is the result to changes in the subjective data? Can we say something about the generalization power of the model outside this data set?

3.4.1 Initial value

The process of finding the optimum is numerical, and ceases after no improvement cannot be found. Also mathematically, it cannot be guaranteed that the result is even a local optimum, when the objective function and/or the constraint functions are non-convex. Due to these reasons, to strengthen the reliability of the result, several initial values, where the algorithm starts to proceed, were used. They all converged closely

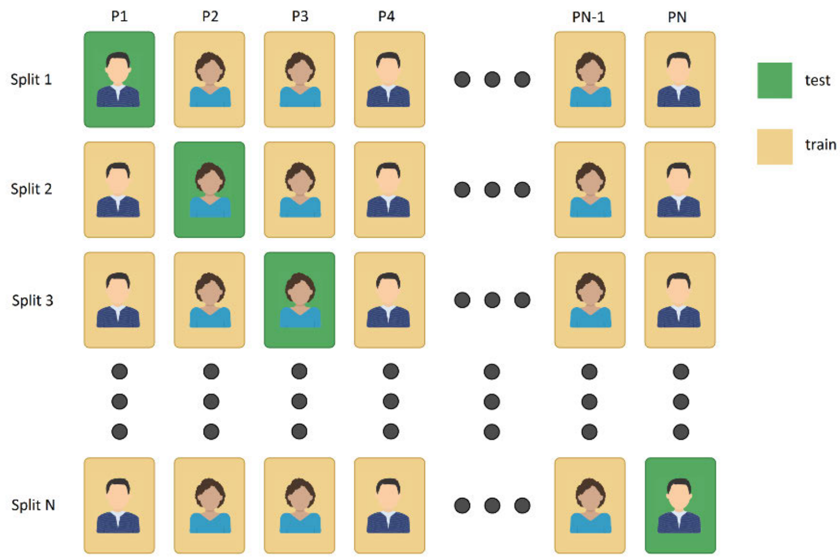

Figure 8. Leave-one-out cross-validation (LOOCV) illustrated with a data set containing ratings from N participants (P1, P2, P3, ..., PN-1, PN). In each round, the ratings given by one participant are left as the test data (green) and the ratings given by all of the remaining participants are taken as the training data (yellow). Thus, N models are derived and tested, and the resulting model performance is given as the mean performance of the N models.

to the same result. The natural choice for initial value was the current reference spectrum to improve. In Publication I, reference spectrum for $C_{50\text{–}5000}$ was the obvious choice as the initial value, as it is recommended for living sounds by ISO 717-1 [14]. In Publication II, the impact source power level of the tapping machine, L_{impact}, was chosen as the initial value, and in Publication III, the reference spectrum C_{tr} was used as it is recommended for urban road traffic noise.

3.4.2 Sensitivity analysis

As the data is subjective, there is a lot of variation in the ratings due to for example personal preferences and subjective noise sensitivity. With empirical bootstrap technique, the variation of any estimate calculated from the data, due to a bit different sample from the participants, can be estimated. In this case it is of interest, how the optimized reference spectrum changes, if the mean subjective ratings change by some small amount. In empirical bootstrapping, samples with replacement are drawn from the data. Samples are the same size as the data itself. Empirical confidence intervals can be determined by taking adequate number of bootstrap samples, determining the optimal reference spectrum based on each sample, and comparing them to the original optimized spectrum calculated using the whole data. The confidence intervals also enable comparison between the existing reference spectra, whether the optimized results differ from the existing spectra or not. If the existing spectrum is contained by the confidence interval, no difference can be stated, with the confidence level in question.

The basic version of calculating the bootstrap empirical confidence intervals (the ordered differences between the reference spectrum values acquired with the bootstrap sample and the original data were used to determine upper and lower bounds for the intervals deploying the 2.5% and 97.5% quantiles) was used. There exist also more accurate versions of bootstrap confidence intervals. Puth et al. [79] recommended the use of a more accurate method and stated the unreliability of bootstrapping with small samples. The adequate number of bootstrap samples was investigated in Publication III to include enough samples that the confidence interval did not practically change even though more samples were added.

3.4.3 Generalization of the model

It is important to evaluate the performance of the resulting reference spectrum. This aspect was taken into consideration in Publication III. The reference spectrum predicts SNQ values for the different structures according to their sound reduction spectrum. How well will the ratings of one participant fit with the predicted SNQ values, if his/her ratings have not taken part in the model derivation process? Generally, model generalization depends on the data that the model is trained with. A model will probably not perform well for data that is from a different distribution than the data that it was trained with. In this context, we can expect that the model gives reasonable results for structures with similar frequency behaviour for sound reduction, participants with similar hearing capabilities and subjective preferences, and same kind of noise types.

The most accurate answer to the question of performance would be attained if the experiment was repeated with new participants and the model performance was tested with the ratings given by the participants of this new experiment. But as the realization of a psychoacoustic experiment is laborious, an easier way is to estimate the performance using only the data gathered in the original experiment. The estimation was made using nested LOOCV, which means that the model selection (optimized reference spectrum with frequency range selection) and testing the selected model against the left-out data is repeated as many times as there are participants in the data, and the resulting performance estimation is given as the average performance of these models.

4 Amplitude modulation

As introduced in Chapter 1, amplitude modulation can be encountered in various environmental sounds. In the present chapter, some properties of amplitude modulated sounds are explained, and ways to determine these properties from a sound are reviewed. The penalty scheme is also explained.

4.1 What is AM sound?

A steady sound is a sound, whose SPL stays quite the same over time. In AM sound, the amplitude of the carrier sound varies with time. In periodic AM, the amplitude of the carrier sound varies periodically. The length of the modulation period T depends on the AM sound source. For example in sea waves the period is several seconds, and in a modern 3.3 MW wind turbine rotating at full speed, the period is around one second between the blade passes, whereas the hits of a pile-driving machine may occur every half a second. The inverse of the modulation period is called modulation frequency f_{m} [Hz]. In random AM, the amplitude varies randomly, and no period can be determined. Road traffic noise, speech and impact sounds from jumping children are examples of sounds, which can exhibit random AM.

Figure 9 illustrates a sinusoidally modulated tone. In this case, the sound pressure p of the modulated sound can be presented as:

$$p(t) = (1 + m \cos 2\pi f_{\mathrm{m}} t) A_{\mathrm{c}} \cos 2\pi f_{\mathrm{c}} t, \tag{8}$$

where m is the modulation index (ratio between the amplitude of the modulating signal and the amplitude of the carrier signal), f_{m} is the modulation frequency, A_{c} is the amplitude of the carrier signal, and f_{c} is the frequency of the carrier signal. In real sounds the modulation wave form can be other than sinusoid as well, for example triangle wave has been used for wind turbine noise auralization [80].

Modulation depth D_{m} is defined as the difference between the maximum and minimum SPL of AM sound. It can be expressed using the modulation index m as:

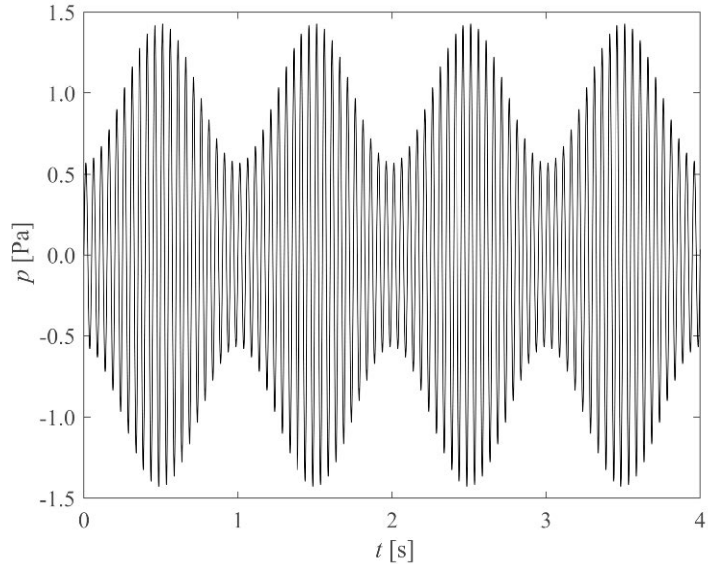

Figure 9. Sound pressure p as a function of time t of an amplitude modulated tone. The amplitude of the carrier sound was sinusoidally modulated (carrier sound frequency f_c = 20 Hz, modulation frequency f_m = 1 Hz).

$$D_m = 20\lg(\frac{1+m}{1-m}) \tag{9}$$

Figure 10a shows an example of an AM sound, and Figure 10b of a steady noise sound. Figure 10c illustrates the combination of the AM sound and the steady noise sound, which shows that background noise reduces the effective modulation depth.

Natural AM sounds can have several modulation frequencies whose modulation depth may vary on different frequency bands depending on the sound production mechanism. For example in wind turbines, so-called "normal AM" is caused by the rotating blade (a swishing sound), and is prominent in the 400–1000 Hz frequency range, whereas in certain circumstances "other AM" (a thump sound) causes an increase in low frequency content and modulation depth. The exact origin of "other AM" has not been understood yet. [81]

Sound is impulsive, if it contains rapid audible level changes. This means that AM sound having a large modulation depth is actually impulsive noise for modulation frequencies above some threshold frequency. Rajala and Hongisto [82] determined penalties for periodic impulsive sounds presented at L_{Aeq} = 55 dB. They discovered that the penalty values increased as the onset rate and level difference increased.

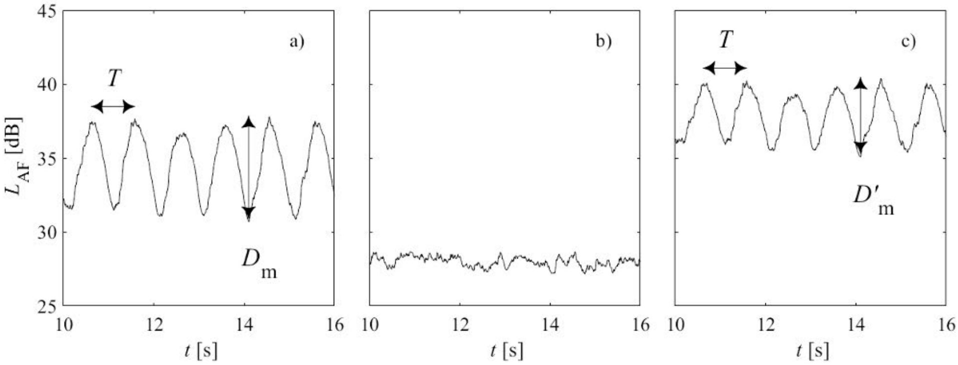

Figure 10. **a)** Amplitude modulated sound (modulation period T = 1 s, modulation depth $D_m \approx 7$ dB, L_{Aeq} = 35 dB). **b)** A steady noise sound (filtered pink noise, L_{Aeq} = 28 dB). **c)** Sum of the AM sound (a) and steady noise sound (b) (resulting modulation depth $D'_m \approx 5$ dB, L_{Aeq} = 38 dB). L_{AF} is the Fast (F) time-weighted, A-weighted SPL, sampled with 100 Hz sampling frequency.

4.2 Measurement of the properties of an AM sound

There does not yet exist standardized measurement methods to determine the modulation depth or the modulation frequency of an AM sound. Different ways have been deployed in various studies: Lee et al. [43] assumed a sinusoidally amplitude modulated wind turbine signal, and used Fourier transform to find the modulation depth at different frequency bands. Fukushima et al. [83] determined the level difference between the A-weighted SPLs measured through Fast and Slow dynamic characteristics of a sound level meter (Fast and Slow indicate the time weightings used in sound level meters, Fast corresponding to 125 ms time constant, and Slow to 1000 ms time constant). They defined the AM depth as the 90% range of the level difference. Pieren et al. [80] used auto-correlation function of level fluctuations to determine the wind turbine blade passing frequency and standard deviation of the level fluctuations. Alamir et al. [84] reviewed current allowable limits and penalties for wind farm noise as well as the methods for quantifying AM. All in all, the chosen method deployed should always take into account the properties of the signal being measured to be able to reveal the true modulation. For example the method deployed by Fukushima et al. is sufficient when the modulation period is much more than 125 ms, which implies that the modulation frequency should be much smaller than 8 Hz.

4.3 Penalty

Equivalent continuous SPL, L_{eq}, is a noise quantity, which determines the steady sound level which over a given period of time has the same total energy as the fluctuating noise being measured:

$$L_{\mathrm{eq}} = 10 \lg \left[\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \frac{p^2}{p_0^2} \, dt \right] \tag{10}$$

In the equation above, t_1 and t_2 are the start and end times of the event being measured, p is the sound pressure, and p_0 is the reference sound pressure (20 μPa). A-weighted equivalent continuous SPL, L_{Aeq}, is widely used in decrees and measurement guidelines. A-weighting is a spectral weighting applied to sound level measurements, defined for example in [85]. It takes into account the fact that the sensitivity of ear is not constant over the whole hearing range but it is most sensitive at 1000 Hz and less sensitive on low-frequencies and very high frequencies. As mentioned in Chapter 1, it has been found in several studies that AM sounds cause higher annoyance than steady sounds with corresponding sound level and spectrum. In such cases, L_{Aeq} is not an adequate quantity for the annoyance of an environmental sound. Thus, when measuring the SPL of a sound, which possesses a modulating feature, a penalty value could be added to the measured equivalent SPL to take into account the excessive annoyance caused by AM. There exist legislative penalties for tonal and impulsive sounds (for example in Finland [53–55]) but not for AM sounds yet. The discussion of AM penalties has mostly been concerning wind turbines. McKenzie et al. [56] have suggested a penalty for medium to large scale wind turbines with rotational speed up to 32 rpm. The suggested penalties having values of 3–5 dB depending on the IOA metric value (based on the noise level exceeded for 90% of the time during a measurement period, L_{A90}, according to [86]). However, amplitude modulation is present in various other sound types as well, and a more general research is needed for this topic.

Before specifying an AM penalty, research is needed to answer questions like: Is a constant penalty adequate or should it depend on the modulation frequency and/or modulation depth or other features of the sounds? Is there a threshold for modulation depth above which the penalty should be used? Is there a need for sound type dependent penalties or could the penalty be set depending on the general properties of a sound?

In Publication IV the penalties were determined for synthetic periodic wide-band AM sounds. The experimental sounds included AM sounds with L_{Aeq} = 35 dB, as well as steady reference sounds with the same spectrum with various values of L_{Aeq}. The reference sounds were used to find a linear relationship between the annoyance ratings and L_{Aeq}. For each AM sound, it was determined which L_{Aeq} its annoyance rating corresponded on the line fitted with the reference sounds. This was called apparent SPL. The penalty was found as the difference between the apparent SPL and the true equivalent level of the AM sound (L_{Aeq} = 35 dB) (see Figure 3 in Publication IV).

5 Results

In this chapter, the results based on Publications I, II, III, and IV are shortly reviewed and considered on the basis of research questions set in Chapter 1.

5.1 Deriving the reference spectra for single-number quantities rating sound insulation

5.1.1 Research question Q1.1

Is there a reference spectrum, which results in a better correlation between the resulting SNQ and subjective ratings of annoyance against living sounds transmitted through wall partitions when compared with the existing SNQs?

Reference spectra for six common living sound types (S1: guitar, S2: music–traffic [1], S3: music–living [2], S4: baby cry, S5: loud speech, S6: dog bark) were determined in Publication I. The optimized reference spectra are shown in Figure 11. Also an averaged reference spectrum was derived, shown in Figure 12.

According to ISO 717-1 [14], reference spectrum for C is recommended for living sounds. For sound types S1, S3, S4, S5, and S6, the resulting optimized reference spectrum had lower values than the reference spectrum for C at low frequency bands from 50 Hz at least to 125 Hz, and depending on the sound type, even up to 400 Hz. For disco music, standard ISO 717-1 recommends the reference spectrum for C_{tr}, and the spectrum of the sound type S2 was weighed to correspond with the reference spectrum for C_{tr}. However, the optimized reference spectrum for sound type S2 corresponded with the reference spectrum for C.

All of the optimized reference spectra resulted in better correlation (squared Pearson's correlation coefficient r^2) between the mean disturbance ratings and the resulting SNQs than the existing standardized SNQs, within this data set (Table V in Publication I). The improvement was largest for sound type S4; none of the existing standardized SNQs were able to predict that well, even though it is a very common

[1] spectrum weighted to correspond with the reference spectrum for C_{tr} of ISO 717-1 [14]

[2] spectrum weighted to correspond with the reference spectrum for C of ISO 717-1 [14]

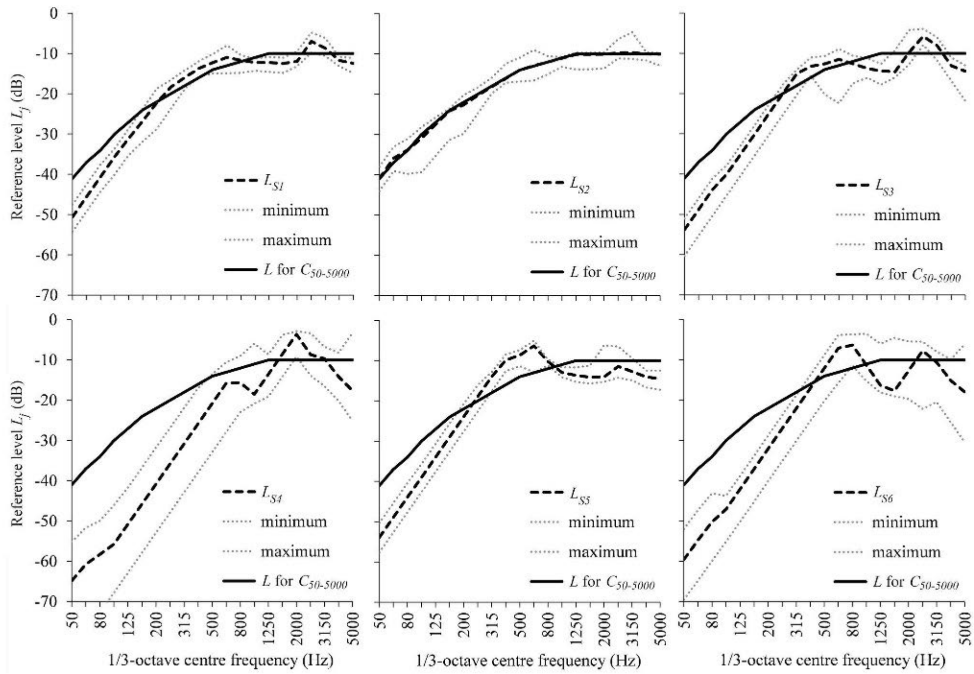

Figure 11. Optimized reference spectra L_{S1}, ..., L_{S6} for living sound types S1, ..., S6. Also the reference spectrum for calculating $C_{50—5000}$ according to ISO 717-1 [14] is shown (Figure 4 in Publication I). Reproduced from The Journal of the Acoustical Society of America with the permission of AIP Publishing.

living sound type. The averaged reference spectrum performed better than the existing standardized SNQs in five cases out of six. Only the sound type S2 required the use of either the optimized reference spectrum for that particular sound type or the reference spectrum for *C*.

5.1.2 Research question Q1.2

Is there a reference spectrum, which results in a better correlation between the resulting SNQ and subjective ratings of annoyance against impact sounds transmitted through concrete floors with various coverings when compared with the existing SNQs?

A reference spectrum for five different impact sound types (S1: walking with hard shoes, S2: walking with socks, S3: walking with soft shoes, S4: super ball bouncing, S5: chair moving) were determined in Publication II. Also an optimized reference spectrum over all sound types was derived. All of the optimized reference spectra are shown in Figure 13.

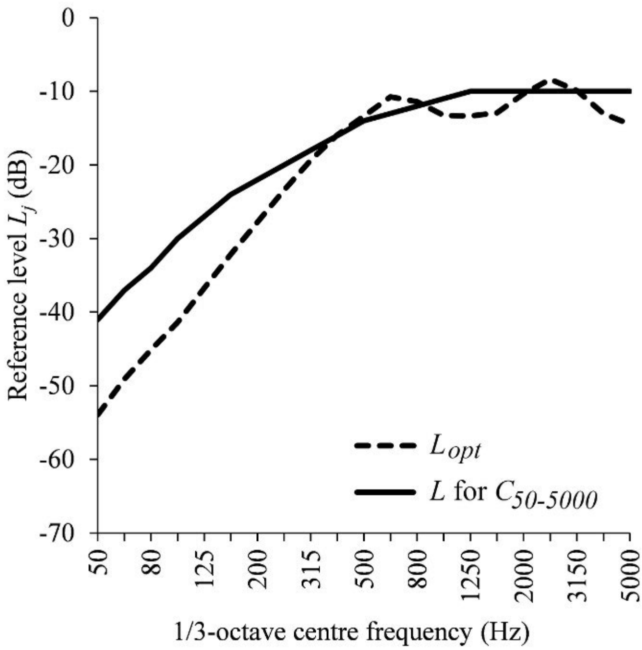

Figure 12. Averaged reference spectrum L_{opt} over all living sound types S1—S6. Also the reference spectrum for calculating $C_{50—5000}$ according to ISO 717-1 [14] is shown (Figure 5 in Publication I). Reproduced from The Journal of the Acoustical Society of America with the permission of AIP Publishing.

The confidence intervals for sound type S3 and the reference spectrum over all sound types mostly included the reference spectrum for C_{I} from standard ISO 717-2 [15]. For sound types S1, S4, and S5, the differences were somewhat larger: there were some frequency areas, where the reference spectrum for C_{I} was not contained within the confidence intervals. For sound type S2, the optimized reference spectrum was distinctly different from the reference spectrum for C_{I}.

All of the optimized reference spectra resulted in better correlation (squared Pearson's correlation coefficient r^2) between the mean annoyance ratings and the resulting SNQs than the existing standardized SNQs, within this data set (Table IV in Publication II). The improvement was largest with the sound type S2, which is a common impact sound type as well. The correlations between the mean annoyance ratings and the existing standardized SNQs were exceptionally low for the sound type S4. The correlation was improved with the optimized reference spectrum also for that sound type, although the result remained clearly lower than for the SNQs optimized for other sound types. This may be due to the fact that the sound type S4 was clearly modulating or even an impulsive sound which creates excessive annoyance.

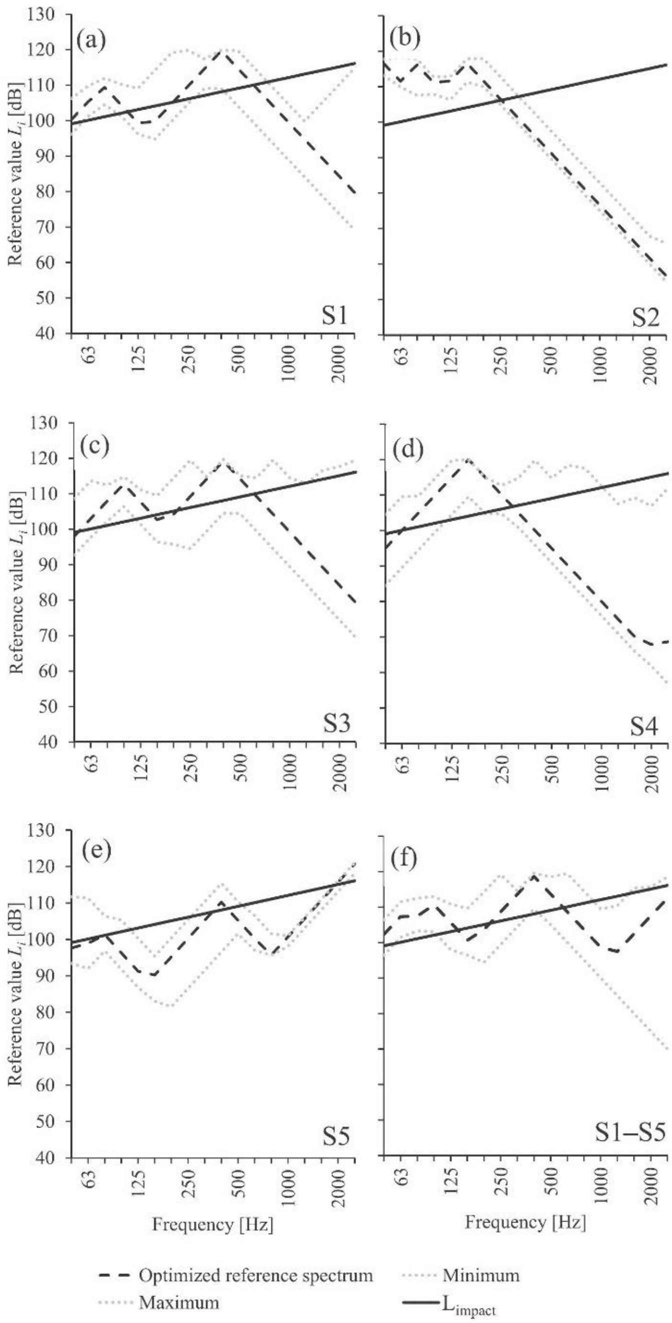

Figure 13. Optimized reference spectra L_{S1}, ..., L_{S5} for impact sound types S1, ..., S5 and the optimized reference spectrum $L_{S1\text{-}S5}$ over all impact sound types S1–S5. Also the impact source power level of the tapping machine, L_{impact}, defined by Equation (7), is shown (Figure 6 in Publication II). Reproduced from The Journal of the Acoustical Society of America with the permission of AIP Publishing.

5.1.3 Research question Q1.3

Is there a reference spectrum, which results in a better correlation between the resulting SNQ and subjective ratings of annoyance against road traffic sounds transmitted through a façade when compared with the existing SNQs?

A reference spectrum for five spectrally different road traffic sounds (S1: Light vehicles, urban street, 50 km/h, S2: Light vehicles, motorway, 80 km/h, S3: Light vehicles, motorway, 100 km/h, S4: Only heavy vehicles, urban street, 60 km/h, and S5: Both heavy and light vehicles, urban street, 60 km/h [3]) was derived in Publication III. Also an aggregate reference spectrum over all sound types was derived. The optimized reference spectra are shown in Figure 14.

Reference spectrum *C* is recommended in ISO 717-1 [14] for highway road traffic at speeds higher than 80 km/h, whereas C_{tr} is recommended to be used for urban road traffic. That is, C_{tr} is recommended for sound types S1, S4 and S5. The optimized reference spectrum for sound type S4 did conform with the reference spectrum for C_{tr} but for S1 and S5, the optimized reference spectrum conformed better with the reference spectrum for *C*, even though the spectrum of S5 was weighted to correspond with C_{tr}. For sound types S2 and S3, the optimized reference spectrum conformed better with the reference spectrum for *C* as recommended by ISO 717-1. Also the aggregate spectrum conformed with the *C* spectrum.

The correlations (squared Pearson's correlation coefficient r^2) were already rather good between the mean annoyance ratings and the existing standardized SNQs (Table II in Publication III). Yet, the SNQs acquired with the optimized reference spectra performed better than the existing SNQs for all sound types. The improvements were statistically significant ($p<0.05$) for sound types S2 and S3, and the aggregate reference spectrum.

[3] spectrum weighted to correspond with the reference spectrum for C_{tr} of ISO 717-1 [14]

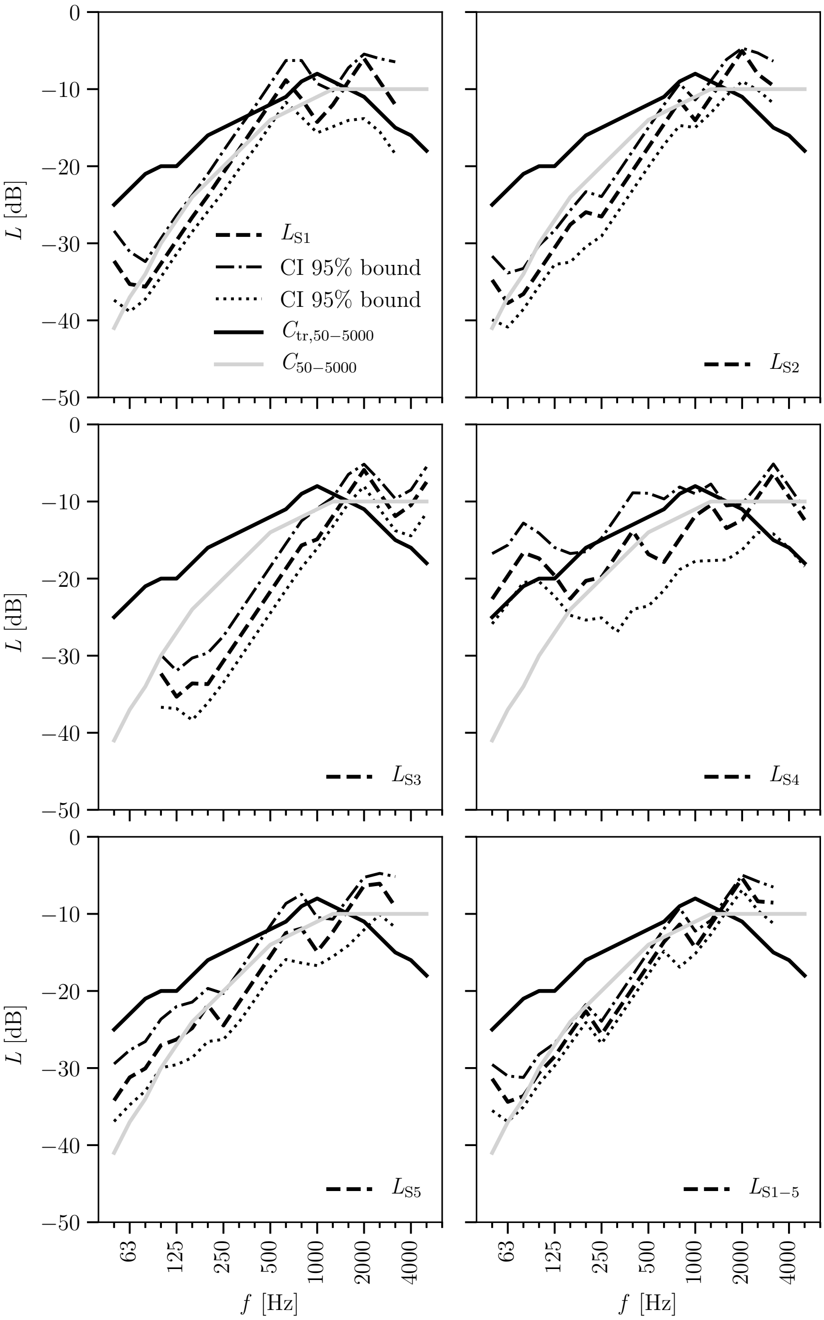

Figure 14. Optimized reference spectra L_{S1}, ..., L_{S5} for road traffic sound types S1, ..., S5 and the aggregate reference spectrum $L_{S1\text{-}5}$ over all road traffic sound types S1–5. Also the reference spectra for calculating $C_{50–5000}$ and $C_{tr,\ 50–5000}$ according to ISO 717-1 [14] are shown (Figure 6 in Publication III). Reproduced from The Journal of the Acoustical Society of America with the permission of AIP Publishing.

5.1.4 Research question Q1.4

Out of the four frequency ranges presented by the standard ISO 717-1, can the most suitable frequency range be selected for a reference spectrum?

In Publication III, it was tested for each sound type, which of the frequency ranges stated in the ISO 717-1[14] produced the best results. Thus, this research question can be answered for road traffic sounds transmitted through a façade. In this case, the optimal frequency range depended on the sound type. The frequency range 50–3150 Hz was selected as the best for sound types S1, S2, S5 and the aggregate spectrum. For sound type S3, the selected frequency range was 100–5000 Hz, and for sound type S4 it was 50–5000 Hz. Sound type S4 was a special case containing only heavy vehicles on urban street, which may be a rare situation in reality. For each sound type, the resulting frequency range was selected as the best in majority of the calculation rounds of the performance estimation. Thus, the selection of the best frequency range was rather stable.

5.1.5 Research question Q1.5

Can the performance of the resulting optimized SNQ be evaluated?

In Publication III, the performance of each optimized reference spectra was evaluated using nested LOOCV. The results were presented in Table II in Publication III. The resulting performance values were compared with the existing SNQs. The results showed that the optimized reference spectra led to statistically significant improvements in the correlations between the annoyance ratings and the resulting SNQs when comparing with the existing SNQs for two road traffic sound types out of five as well as for the aggregate reference spectrum.

5.2 Determining a penalty for periodic wide-band AM sounds

5.2.1 Research question Q2

Is there a need for a penalty for periodically amplitude modulated wide-band low level sounds? If so, could a constant penalty be used?

In Publication IV, penalties were determined for synthetic periodically amplitude modulated wide-band low level sounds, having two different spectra (S1: low frequency sound similar to wind turbine noise, and S2: high frequency sound similar to road traffic noise). The resulting penalties are presented in Figure 15 (and Table III in Publication IV) for different modulation depths and modulation frequencies. According to the results, for these test sounds, a penalty can be suggested if the modulation depth is 4 dB or higher, and the modulation frequency is 1 Hz or higher. This resulted in penalties varying from 3.9 to 12.3 dB depending on the modulation depth and frequency.

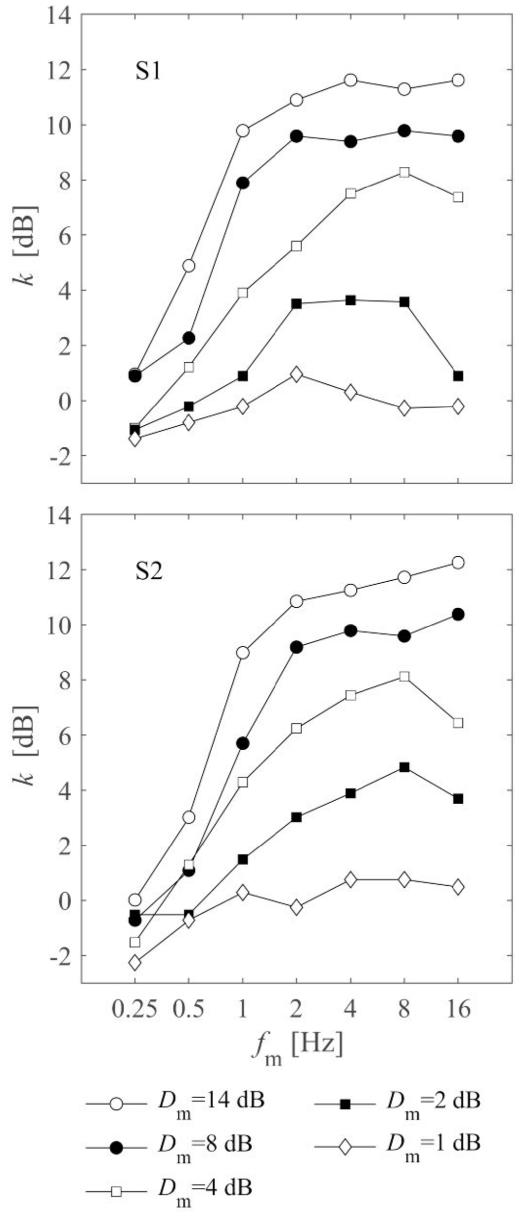

Figure 15. The mean annoyance penalty k as a function of modulation frequency, f_m, for five modulation depths, D_m and two sound spectra S1 and S2. (Figure 6 in Publication IV). Reproduced from The Journal of the Acoustical Society of America with the permission of AIP Publishing.

6 Discussion

The results presented in the previous chapter are discussed in this chapter. First, the specific research questions are considered, and thereafter, the applicability of the results is pondered.

6.1 Single-number quantities for sound insulation

6.1.1 Main results

The goal was to produce such SNQs that correlate well with the subjective annoyance towards sounds transmitted through various structures (Q1.1–Q1.3). The results showed that an optimized reference spectrum resulting in well-performing SNQ could be derived using mathematical optimization and psychoacoustic experimental data for each studied sound type. The results also showed that there is room for improvement in regards with the existing standardized SNQs.

The resulting reference spectra were sound dependent. For some sounds the gained improvement was greater than for the others: for example the SNQs acquired with the optimized reference spectra derived for the sound type S4 (baby cry) in airborne living sounds as well as for S2 (walking with socks) in impact sounds produced clearly better correlations compared with the existing standardized SNQs. The mathematical optimization scheme was well-suited for this kind of task. In cases, where the existing SNQs were already rather good, the resulting optimized SNQs produced still small improvements: for example the road traffic sounds S2 (Light vehicles, motorway, 80 km/h) and S3 (Light vehicles, motorway, 100 km/h).

The existing standardized calculation system for sound insulation needs to be improved. Adjusting the values of a reference spectrum is a rather simple way to improve the existing system as it is possible to apply to the existing calculation scheme without making radical changes.

6.1.2 Frequency range selection

Research question Q1.4 dealt with the selection of the relevant frequency range. The frequency range selection was taken as a part of the model selection in Publication III. The performance estimation of the optimized reference spectrum with selected frequency range was made with nested LOOCV, and it showed that the deployed method was robust for each road traffic sound type. For most of the studied road traffic sounds, 50–3150 Hz was selected as the best frequency range. It should be noted, that there may be even a better frequency range available, as the frequency range selection was restricted within the choices given by ISO 717-1 [14], which have been under a debate in the recent literature.

6.1.3 Performance of the resulting reference spectra

It is important to estimate, whether the attained improvements are significant when compared with the existing SNQs. In Publications I and II this was considered by calculating the confidence intervals for the reference spectrum. If the reference spectrum of the existing SNQ was included within the confidence intervals, it was assumed that there is no significant difference between the optimized and existing reference spectrum. In Publication III, the evaluation scheme was further developed to also estimate the performance of the resulting reference spectrum for a participant, whose ratings had not been part of the optimization process as well as whether the correlations acquired for the optimized SNQs differed statistically from the existing SNQs (Q1.5). This more detailed evaluation showed that even though the existing SNQs for determining the sound insulation of façade against road traffic noise were already rather good, statistically significant improvements were possible to attain by optimizing the reference spectra for certain sound types. This underlines that it is possible to estimate the performance of the optimized model for persons whose annoyance ratings have not been a part of the model selection process. To take the step even further, is it possible to estimate the model generalization, to persons who have not taken part in the experiment? This will be discussed in the upcoming Section 6.1.4.

6.1.4 Applicability and future work

The variation is rather large within psychoacoustic data (e.g. Figure 3 in Publication III). The number of participants should be large enough to be able to evaluate the sensitivity and the performance of the optimized reference spectrum. The results in this thesis were derived from psychoacoustic experiments having number of participants between 40–59, which is rather adequate for this purpose. A large number of participants does not guarantee reliable results as such. To improve the reliability of the results, it would beneficial to exploit results from several psychoacoustical experiments.

In Publications I, II and III, in addition to optimizing the reference spectra for individual sound types, more general reference spectra were sought for airborne living sounds, impact living sounds and road traffic noise. The general reference spectra performed quite well but naturally not as well as the individually optimized reference spectra. The general reference spectra did not perform better than an existing SNQ for sound type S2 (music–traffic) in Publication I, and sound type S5 (chair moving) in Publication II. In each of the Publications, the general reference spectrum was a kind of average reference spectrum for the sounds included in the experiment. The sounds included in the experiments were rather representative of normal sounds encountered in living environments, however the sound types were not weighted in any way but affected the resulting general reference spectrum equally. The construction of the general spectrum is not evident: should the reference spectrum be affected more by the most annoying sounds or by incidence of the sounds? In the psychoacoustic experiments utilized in this study, the experimental sounds were short and included only one sound type at a time. It would be interesting to derive reference spectra from annoyance ratings given for experimental sounds including several sounds in a more natural sound scheme. The construction of the general reference spectrum is a topic to be studied more in future.

The experimental setting in the psychoacoustic experiment affect the generalization power of the results. All of the choices made in the setup have their own effect. For example the selection of the structures should correspond the variability of the structures applied in actual buildings. The sound types and their implementation should correspond real sounds encountered in normal housing. A valid set of sounds may depend on the country at some extent. Living sounds are generally the same overall but cultural differences and acceptable noise level may vary [87]. Traffic noise spectrum varies from region to region depending on the speed limits, types of vehicles and their incidence. The selection of the participants assign for which group of people the results are applicable. The age span in the psychoacoustic data exploited in these studies was 18–59 years, and normal hearing was assumed, which does not represent the whole population. The selected presentation levels of the sound (resulting from the choice of the structures, their sound insulation capabilities and the sound level of the original sounds) set the scale within which the subjective scaling is made. In Publication III the sound level variation within the experimental sounds was small for each sound type. A different temporal variation is also a factor affecting annoyance as Brink et al. found [9].

The results acquired from a psychoacoustic experiment are directly applicable only for the specific sounds, the structures used, and for the same distribution of people, and thus, the results should be generalized with care. To acquire a more general reference spectrum for example for standardization, it would be beneficial to optimize the reference spectra with data acquired from psychoacoustic experiments conducted with several different settings.

De la Prida et al. [36] implemented a psychoacoustic experiment with 119 participants, who rated the annoyance (two-alternative choice method) towards five different sound types filtered through six different façade structure filters. The sounds were urban traffic sounds representative of traffic noise in Madrid with different proportions of vehicle types and road speeds, reproduced with headphones. They derived reference spectra from these ratings utilizing the mathematical optimization scheme presented in Publication I. It is intriguing to compare their results with the results acquired for urban road traffic noise types S1 and S5 in the Publication III. The reference spectra from these two studies are depicted in Figure 16. The results are in good conformance with each other.

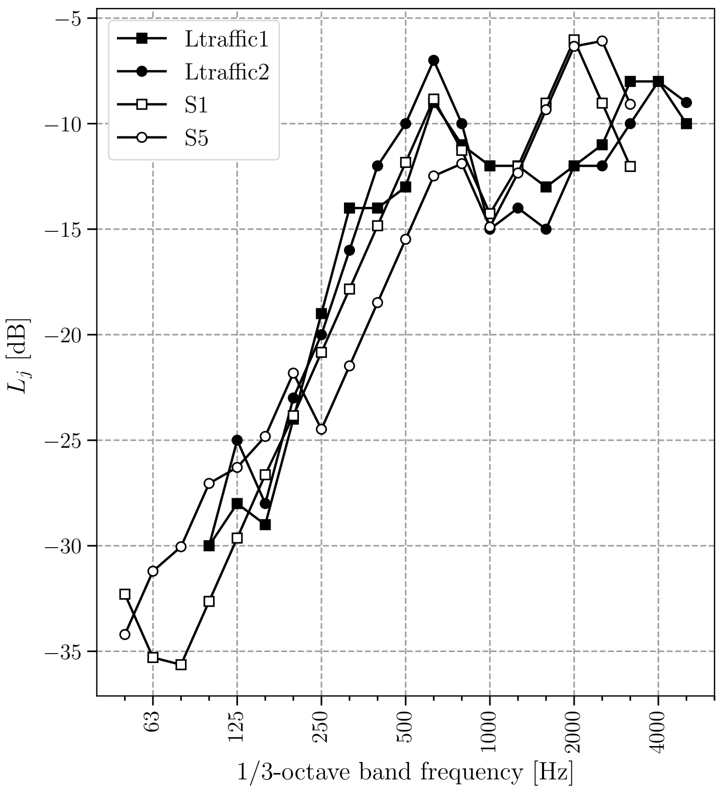

Figure 16. Reference spectrum values L_j acquired for urban road traffic sound types S1 and S5 (Publication III) and for urban traffic noise samples derived by de la Prida et al. [36].

6.2 Penalty for periodic AM broad-band sounds

6.2.1 Results

The results showed that there is a need for penalty for certain ranges of modulation frequencies and modulation depths for periodic wide-band AM sounds with low sound level. The penalties increased with increasing modulation frequency and increasing modulation depth, which does not support using a constant penalty.

6.2.2 Applicability and future work

The goal of this experiment was to derive penalties at different modulation frequencies and depths for a general low level sound without any specific cues to its origin. Penalties for sounds with different sound and modulation spectrum may differ from the penalties derived in this experiment. Yet, the results did conform reasonably well with the previous studies found in literature with wind turbine sound [43, 46, 52]. In future, it would be beneficial to widen the choices made in this experiment: to use a different total sound level, to increase the range of the modulation frequencies above 16 Hz, alter the carrier sound spectra, as well as to enable other than flat modulation spectrum. There is also a need for developing convergent measurement methods to determine the modulation properties of a sound for an adequate range of modulation frequencies and depths if penalties are to be given based on these parameters.

7 Conclusions

It is important that the physical quantities used to assess the properties of noise as well as sound insulation capabilities of structures correspond to subjective perception. In literature, it has been shown that this is not always the case.

Reference spectra were derived using mathematical optimization and annoyance ratings from published, comprehensive psychoacoustic experiments in order to develop optimized single-number quantities for rating airborne and impact sound insulation. The derivation of the reference spectra has not been made using mathematical optimization and psychoacoustic data before. The results showed that the developed calculation scheme enabled systematic development of well-performing SNQs for rating sound insulation. This thesis covers all three types of standardized sound insulation: airborne, impact and façade sound insulation. The results as well as the developed research methodology are worth taking into account when developing the rating quantities included in the standards ISO 717-1 [14] and ISO 717-2 [15]. The decisions of the most important sound types and their weighting in the derivation process of a reference spectrum are selected by political grounds, and are thus, out of scope of this thesis.

Designating a penalty for amplitude modulated sounds is also a political decision. The decisions, however, should be based on strong scientific evidence. This thesis offers unique information on the order of magnitude of a potential penalty for periodic wide-band AM sounds with a wide range of modulation frequencies and modulation depths. The results are not restricted for example to wind turbine sounds having specific modulation frequency range or modulation depth but aim at more generic understanding.

The results of this thesis can be exploited in the improvement of acoustical standards and building regulations, as well as in developing better building structures with improved sound-proof capabilities. This will hopefully eventually lead to healthier outdoor and indoor environments.

List of References

[1] European Environment Agency. Environmental noise in Europe—2020, 2020. https:doi://10.2800/686249.

[2] J. H. Schmidt and M. Klokker. Health effects related to wind turbine noise exposure: A systematic review. *PLOS ONE*, 9(12):1–28, 12 2014.

[3] M. Basner, W. Babisch, A. Davis, M. Brink, C. Clark, A. Janssen, and S. Stansfeld. Auditory and non-auditory effects of noise on health. *Lancet*, 383(9925):1325–1332, 2014.

[4] T. N. Le, L. V. Straatman, J. Lea, and B. Westerberg. Current insights in noise-induced hearing loss: a literature review of the underlying mechanism, pathophysiology, asymmetry, and management options. *Otolaryngol. Head Neck Surg.*, 46(1):41, 2017.

[5] A. Haapakangas, V. Hongisto, and A. Liebl. The relation between the intelligibility of irrelevant speech and cognitive performance – a revised model based on laboratory studies. *Indoor Air*, 20 (6):1130–1146, 2020.

[6] H. A. R. Jensen, B. Rasmussen, and O. Ekholm. Neighbour noise annoyance is associated with various mental and physical health symptoms: Results from a nationwide study among individuals living in multi-storey housing. *BMC Public Health*, 19(1):1508, 2019.

[7] C. Maschke and H. Niemann. Health effects of annoyance induced by neighbour noise. *Noise Control Eng. J.*, 55(3):348–356, 2007.

[8] D. Halperin. Environmental noise and sleep disturbances: A threat to health? *Sleep Sci.*, 7(4), 2014.

[9] M. Brink, B. Schäffer, D. Vienneau, M. Foraster, R. Pieren, I. C. Eze, C. Cajochen, N. Probst-Hensch, M. Röösli, and J.-M. Wunderli. A survey on exposure-response relationships for road, rail, and aircraft noise annoyance: Differences between continuous and intermittent noise. *Environ. Int.*, 125:277–290, 2019.

[10] World Health Organization. Environmental noise guidelines for the European region, 2018. https://www.euro.who.int/__data/assets/pdf_file/0008/383921/noise-guidelines-eng.pdf, accessed in Feb 19, 2021.

[11] J. Radun, V. Hongisto, and M. Suokas. Variables associated with wind turbine noise annoyance and sleep disturbance. *Build Environ*, 150:339–348, 2019.

[12] B. Rasmussen and M. Machimbarrena (editors). COST Action TU0901 – Building acoustics throughout Europe. Volume 1: Towards a common framework in building acoustics throughout Europe, 2014. https://www.cost.eu/publications/building-acoustics-throughout-europe-volume-1-towards-a-common-framework-in-building-acoustics-throughout-europe, accessed in Feb 19, 2021.

[13] C. Höller. Review and comparison of ASTM and ISO standards on sound transmission in buildings. In *Proceedings of 47th International Congress and Exposition on Noise Control Engineering: Impact of Noise Control Engineering, INTER-NOISE 2018, August 26–29, 2018, Chicago, USA*, 2018.

[14] ISO 717-1. *Acoustics — Rating of sound insulation in buildings and of building elements — Part 1: Airborne sound insulation*. International Organization for Standardization, Geneva, 2013.

[15] ISO 717-2. *Acoustics — Rating of sound insulation in buildings and of building elements — Part 2: Impact sound insulation*. International Organization for Standardization, Geneva, 2013.

[16] ASTM E413-10. *Classification for rating sound insulation*. ASTM International, West Conshohocken, PA, 2016.

[17] ASTM E989-18. *Standard Classification for Determination of Single-Number Metrics for Impact Noise*. ASTM International, West Conshohocken, PA, 2018.

[18] ASTM E1332-10a. *Standard Classification for Rating Outdoor-Indoor Sound Attenuation*. ASTM International, West Conshohocken, PA, 2010.

[19] J.-P. Vian, W. F. Danner, and J. W. Bauer. Assessment of significant acoustical parameters for rating sound insulation of party walls. *J. Acoust. Soc. Am.*, 73:1236–1243, 1983.

[20] H. Tachibana, Y. Hamada, and F. Sato. Loudness evaluation of sounds transmitted through walls—basic experiment with artificial sounds. *J. Sound Vib.*, 127(3):499–506, 1988.

[21] F. R. Mortensen. Subjective evaluation of noise from neighbours with focus on low frequencies. Technical report, Technical University of Denmark, 1999.

[22] H. K. Park, J. S. Bradley, and B. N. Gover. Evaluating airborne sound insulation in terms of speech intelligibility. *J. Acoust. Soc. Am.*, 123(3):458–1471, 2008.

[23] V. Hongisto, D. Oliva, and J. Keränen. Subjective and objective rating of airborne sound insulation – living sounds. *Acta Acust united Ac*, 100:848–863, 2014.

[24] S. Bailhache, J. Jagla, and C. Guigou-Carter. Environnement et ambiances, effet des basses frequences sur le confort acoustique–tests psychoacoustiques. Technical report, Centre Scientifique et Technique du Batiment CSTB, 2014. English translation available `https://hal-cstb.archives-ouvertes.fr/hal-01045056`.

[25] M. Rychtáriková, H. Müllner, V. Chmelík, B. Roozen, D. Urbán, D. Pelegrin-Garcia, and C. Glorieux. Perceived loudness of neighbour sounds heard through heavy and light-weight walls with equal Rw+C50-5000. *Acta Acust united Ac*, 102(1):58–66, 2016.

[26] C. Monteiro, M. Machimbarrena, D. de la Prida, and M. Rychtarikova. Subjective and objective acoustic performance ranking of heavy and light weight walls. *Appl. Acoust.*, 110:268–279, 2016.

[27] J. Y. Jeon, J. K. Ryu, and H. Tachibana. Review of the impact ball in evaluating floor impact sound. *Acta Acust united Ac*, 92(5):777–786, 2006.

[28] B. N. Gover, J. S. Bradley, B. Zeitler, and S. Schoenwald. Objective and subjective assessment of lightweight wood-framed floor assemblies in response to footstep and low-frequency impact sounds. In *Proceedings of 40th International Congress and Exposition on Noise Control Engineering 2011, INTER-NOISE 2011, September 4–7, 2011, Osaka, Japan*, 2011.

[29] M. Späh, K. Hagberg, O. Bartlomé, L. Weber, P. Leistner, and A. Liebl. Subjective and objective evaluation of impact noise sources in wooden buildings. *Build. Acoust.*, 20(3):193–213, 2013.

[30] M. Kylliäinen, V. Hongisto, D. Oliva, and L. Rekola. Subjective and objective rating of impact sound insulation of a concrete floor with various coverings. *Acta Acust united Ac*, 103(2):236–251, 2017.

[31] A. Frescura, P. J. Lee, F. Schöpfer, and U. Schanda. Correlations between standardised and real impact sound sources in lightweight wooden structures. *Appl. Acoust.*, 173:107690, 2021.

[32] N.-G. Vardaxis and D. Bard. Review of acoustic comfort evaluation in dwellings: Part III–airborne sound data associated with subjective responses in laboratory tests. *Build. Acoust.*, 25:289–305, 2018.

[33] N.-G. Vardaxis and D. Bard. Review of acoustic comfort evaluation in dwellings: Part II—impact sound data associated with subjective responses in laboratory tests. *Build. Acoust.*, 25:171–192, 2018.

[34] V. Hongisto, D. Oliva, and L. Rekola. Subjective and objective rating of the sound insulation of residential building façades against road traffic noise. *J. Acoust. Soc. Am.*, 144(2):1100–1112, 2018.

[35] A. J. Torija, D. P. Ruiz, B. De Coensel, D. Botteldooren, B. Berglund, and Á. Ramos-Ridao. Relationship between road and railway noise annoyance and overall indoor sound exposure. *Transport. Res. Part D Trans. Environ.*, 16(1):15–22, 2011.

[36] D. de la Prida, A. Pedrero, M. A. Navacerrada, and A. Díaz-Chyla. An annoyance-related SNQ for the assessment of airborne sound insulation for urban-type sounds. *Appl. Acoust.*, 168:107432, 2020.

[37] J. H. Rindel. *Sound Insulation in Buildings*. Taylor & Francis Group, Boca Raton, 2017.

[38] W. Scholl, J. Lang, and Wittstock V. Rating of sound insulation at present and in future. The revision of ISO 717. *Acta Acust united Ac*, 97:686–698, 2011.

[39] H. K. Park and J. S. Bradley. Evaluating standard airborne sound insulation measures in terms of annoyance, loudness, and audibility ratings. *J. Acoust. Soc. Am.*, 126(1):208–219, 2009.

[40] K. Bodlund. Alternative reference curves for evaluation of the impact sound insulation between dwellings. *J. Sound Vib.*, 102(3):381–402, 1985.

[41] F. Ljunggren, C. Simmons, and K. Hagberg. Correlation between sound insulation and occupants' perception – proposal of alternative single number rating of impact sound. *Appl. Acoust.*, 85:57–68, 2014.

[42] 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. *Appl. Acoust.*, 160:107144, 2020.

[43] S. Lee, K. Kim, W. Choi, and S. Lee. Annoyance caused by amplitude modulation of wind turbine noise. *Noise Control Eng. J.*, 59(1):38–46, 2011.

[44] S. von Hünerbein, A. King, B. Piper, and M. Cand. Wind turbine amplitude modulation: Research to improve understanding as to its cause & effect, Work package B(2): Development of an AM dose-response relationship. Technical report, Renewable UK, 2013.

[45] H. Hafke-Dys, A. Preis, T. Kaczmarek, A. Biniakowski, and P. Kleka. Noise annoyance caused by amplitude modulated sounds resembling the main characteristics of temporal wind turbine noise. *Arch. Acoust.*, 41(2):221–232, 2016.

[46] C. Ioannidou, S. Santurette, and C.-H. Jeong. Effect of modulation depth, frequency, and intermittence on wind turbine noise annoyance. *J. Acoust. Soc. Am.*, 139(3):1241–1251, 2016.

[47] B. Schäffer, R. Pieren, S. J. Schlittmeier, and M. Brink. Effects of different spectral shapes and amplitude modulation of broadband noise on annoyance reactions in a controlled listening experiment. *Int. J. Environ. Res. Public Health*, 15(5):1029, 2018.

[48] C. Kantarelis and J.G. Walker. The identification and subjective effect of amplitude modulation in diesel engine exhaust noise. *J. Sound Vib.*, 120(2):297–302, 1988.

[49] J. Bradley. Annoyance caused by constant-amplitude and amplitude-modulated sounds containing rumble. *Noise Control Eng. J.*, 42:203, 1994.

[50] L. M. Wang and C. C. Novak. Human performance and perception-based evaluations of indoor noise criteria for rating mechanical system noise with time-varying fluctuations. In *ASHRAE Transactions*, volume 116, pages 553–568, 2010. 2010 ASHRAE Annual Conference June 26–30, 2010.

[51] D. Y. Gwak, D. Han, and S. Lee. Sound quality factors influencing annoyance from hovering UAV. *J. Sound Vib.*, 489:115651, 2020.

[52] B. Schäffer, S. J. Schlittmeier, R. Pieren, K. Heutschi, M. Brink, R. Graf, and J. Hellbrück. Short-term annoyance reactions to stationary and time-varying wind turbine and road traffic noise: A laboratory study. *J. Acoust. Soc. Am.*, 139(5):2949–2963, 2016.

[53] Ministry of the Environment. The Government's decision 993/92 on the regulated noise levels, 1992. Helsinki, Finland. Available at: `https://www.finlex.fi/fi/laki/alkup/1992/19920993`, accessed Feb 19, 2021.

[54] Ministry of the Environment. Decree of the Ministry of the Environment on the Acoustic Environment of Buildings, 2017. Helsinki, Finland. Available at: `https:/finlex.fi/fi/laki/alkup/2017/20170796`, accessed Feb 19, 2021.

[55] Ministry of social affairs and health. Decree 545/2015 of the ministry of social affairs and health on health-related conditions of housing and other residential buildings and qualification requirements for third-party experts, 2015. Helsinki, Finland. Available at: `https://finlex.fi/en/laki/kaannokset/2015/en20150545`, accessed Feb 19, 2021.

[56] A. McKenzie, M. Cand, D. Bowdler, M. Jiggins, G. Irvine, M. Reid, R. Perkins, M. Lotinga, M. Hayes, and A. Bullmore. A planning condition for wind turbines. *Acoust. Bull.*, 42:56–60, 2017.

[57] A. Brand, L. Allen, M. Altman, M. Hlava, and J. Scott. Beyond authorship: attribution, contribution, collaboration, and credit. *Learn. Publ.*, 28(2):151–155, 2015.

[58] W. A. Yost. Psychoacoustics: A brief historical overview. *Acoust. Today*, 11(3):46–53, 2015.

[59] G. Fechner. *Elemente der Psychophysik: Band 1*. Breitkopf und Härtel, 1860. *Elements of psychophysics. 1*. New York, 1966. Translated by H. E. Adler. Edited by D. H. Howes and E. G. Boring.

[60] R. Guski. Psychological methods for evaluating sound quality and assessing acoustic information. *Acta Acust united Ac*, 83:765–774, 1997.

[61] C. H. Morrell, S. Gordon-Salant, J. D. Pearson, L. J. Brant, and J. L. Fozard. Age- and gender-specific reference ranges for hearing level and longitudinal changes in hearing level. *J. Acoust. Soc. Am.*, 100(4):1949–1967, 1996.

[62] H. Møller, D. Hammershøi, C. B. Jensen, and M. F. Sørensen. Transfer characteristics of headphones measured on human ears. *J. Audio Eng. Soc.*, 43:203–217, 1995.

[63] A. Fischetti, J. Jouhaneau, and Y. Hemim. Differences between headphones and loudspeakers listening in spatial properties of sound perception. *Appl. Acoust.*, 39(4):291–305, 1993.

[64] E. Parizet, N. Hamzaoui, and G. Sabatie. Comparison of some listening test methods : A case study. *Acta Acust united Ac*, 2(91):356–364, 2005.

[65] M. Kylliäinen, J. Lietzén, V. Kovalainen, and V. Hongisto. Correlation between single-number-quantities of impact sound insulation and various noise ratings of walking on concrete floors. *Acta Acust united Ac*, 101:975–985, 9 2015.

[66] ISO 226. *Acoustics — Normal equal-loudness-level contours*. International Organization for Standardization, Geneva, Switzerland, 2003.

[67] V. Hongisto, M. Mäkilä, A. Haapakangas, J. Hakala, J. Hyönä, and M. Kylliäinen. Acoustic satisfaction in multi-storey buildings built after 1950 : preliminary results of a field survey. In *Proceedings of Inter-Noise 2013, the 42nd International Congress and Exposition on Noise Control Engineering, Innsbruck, Austria, 15.-18. September 2013*, number 835, pages 1–8, 2013.

[68] C. Grimwood. Complaints about poor sound insulation between dwellings in England and Wales. *Appl. Acoust.*, 52(3):211–223, 1997.

[69] ISO 10140-2. *Acoustics — Laboratory measurement of sound insulation of building elements — Part 2: Measurement of airborne sound insulation*. International Organization for Standardization, Geneva, Switzerland, 2010.

[70] ISO 16283-1. *Acoustics — Field measurement of sound insulation in buildings and of building elements — Part 1: Airborne sound insulation*. International Organization for Standardization, Geneva, Switzerland, 2014.

[71] ISO 16283-3. *Acoustics — Field measurement of sound insulation in buildings and of building elements — Part 3: Façade sound insulation*. International Organization for Standardization, Geneva, Switzerland, 2016.

[72] C. Scrosati, F. Scamoni, A. Prato, S. Secchi, P. Fausti, A. Astolfi, L. Barbaresi, F. D'Alessandro, A. Di Bella, C. Schenone, and G. Zambon. Uncertainty of facade sound insulation by a round robin test. evaluations of low-frequency procedure and single numbers. *Build. Environ.*, 105: 253–266, 2016.

[73] J. Keränen, J. Hakala, and V. Hongisto. The sound insulation of façades at frequencies 50–5000 Hz. *Build. Environ.*, 156:12–20, 2019.

[74] ISO 10140-3. *Acoustics — Laboratory measurement of sound insulation of building elements — Part 3: Measurement of impact sound insulation*. International Organization for Standardization, Geneva, Switzerland, 2010.

[75] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear programming: Theory and algorithms*. John Wiley & Sons Inc., Hoboken, New Jersey, 3. edition, 2013.

[76] Mathworks. *User's guide: Optimization toolbox*. 2020. Retrieved December 2020 from `https://se.mathworks.com/help/optim/index.html`.

[77] Scipy.org. *SciPy v1 1.5.4 Reference Guide: Optimization*. 2020. Retrieved December 2020 from `https://docs.scipy.org/doc/scipy/reference/tutorial/optimize.html`.

[78] S. Varma and R. Simon. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, 7(91), 2006.

[79] M.-T. Puth, M. Neuhäuser, and G. D. Ruxton. On the variety of methods for calculating confidence intervals by bootstrapping. *J. Anim. Ecol.*, 84(4):892–897, 2015.

[80] R. Pieren, K. Heutschi, Menno Müller, M. Manyoky, and K. Eggenschwiler. Auralization of wind turbine noise: Emission synthesis. *Acta Acust united Ac*, 100:25–33, 2014.

[81] M. G. Smith. Wind turbine amplitude modulation: Research to improve understanding as to its cause & effect, Work package A(2): Fundamental research into possible causes of amplitude modulation. Technical report, Renewable UK, 2013.

[82] V. Rajala and V. Hongisto. Annoyance penalty of impulsive noise – the effect of impulse onset. *Build. Environ.*, 168:106539, 2020.

[83] A. Fukushima, K. Yamamoto, H. Uchida, S. Sueoka, T. Kobayashi, and H. Tachibana. Study on the amplitude modulation of wind turbine noise: Part 1 – Physical investigation. In *Proceedings of InterNoise 2013*, 2013. September 15–18, Innsbruck, Austria.

[84] M. A. Alamir, K. L. Hansen, and P. Catcheside. Penalties applied to wind farm noise: Current allowable limits, influencing factors, and their development. *J. Clean. Prod.*, page 126393, 2021.

[85] IEC 61672-1:2013. *Electroacoustics – Sound level meters – Part 1: Specifications*. International Electrotechnical Commission (IEC), Geneva, 2013.

[86] M. Cand, R. Davis, C. Jordan, M. Hayes, and R. Perkins. *A good practice guide to the application of ETSU-R-97 for the assessment and rating of wind turbine noise*. Institute of Acoustics, United Kingdom, 2013.

[87] L.-A. Gille, C. Marquis-Favre, and J. Morel. Testing of the European Union exposure-response relationships and annoyance equivalents model for annoyance due to transportation noises: The need of revised exposure-response relationships and annoyance equivalents model. *Environ. Int.*, 94:83–94, 2016.

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama, Turku, Finland 2021

ISBN 978-951-29-8536-4 (PRINT)
ISBN 978-951-29-8537-1 (PDF)
ISSN 2736-9390 (PRINT)
ISSN 2736-9684 (ONLINE)