Technology Collaboration Programme
by iea

Technology Collaboration Programme on Heat Pumping Technologies (HPT TCP)

Annex 51

Acoustic Signatures of Heat Pumps

Final Report – Part 12

6 Annoyance rating and psychoacoustical analysis of heat pump sound

Editors:
Henrik Hellgren, RISE, Sweden
Christian Kaseß, ARI, ÖAW Austrian Academy of Sciences, Austria

April 2020

Report no. HPT-AN51-12

Published by Heat Pump Centre
c/o RISE – Research Institutes of Sweden
Box 857, SE-501 15 Borås
Sweden
Phone +46 10 16 53 42

Website https://heatpumpingtechnologies.org

Legal Notice

Neither the Heat Pump Centre nor any person acting on its behalf:
(a) makes any warranty or representation, express or implied, with respect to the information contained in this report; or
(b) assumes liabilities with respect to the use of, or damages, resulting from, the use of this information.
Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement recommendation or favoring.
The views and opinions of authors expressed herein do not necessarily state or reflect those of the Heat Pump Centre, or any of its employees. The information herein is presented in the authors' own words.

Production Heat Pump Centre, Borås, Sweden

ISBN 978-91-89561-68-7
Report No. HPT-AN51-12

Preface

This project was carried out within the Technology Collaboration Programme on Heat Pumping Technologies (HPT TCP), which is a Technology Collaboration Programme within the International Energy Agency, IEA.

The IEA

The IEA was established in 1974 within the framework of the Organization for Economic Cooperation and Development (OECD) to implement an International Energy Programme. A basic aim of the IEA is to foster cooperation among the IEA participating countries to increase energy security through energy conservation, development of alternative energy sources, new energy technology and research and development (R&D). This is achieved, in part, through a programme of energy technology and R&D collaboration, currently within the framework of nearly 40 Technology Collaboration Programmes.

The Technology Collaboration Programme on Heat Pumping Technologies (HPT TCP)

The Technology Collaboration Programme on Heat Pumping Technologies (HPT TCP) forms the legal basis for the implementing agreement for a programme of research, development, demonstration, and promotion of heat pumping technologies. Signatories of the TCP are either governments or organizations designated by their respective governments to conduct programmes in the field of energy conservation.

Under the TCP, collaborative tasks, or "Annexes", in the field of heat pumps are undertaken. These tasks are conducted on a cost-sharing and/or task-sharing basis by the participating countries. An Annex is in general coordinated by one country which acts as the Operating Agent (manager). Annexes have specific topics and work plans and operate for a specified period, usually several years. The objectives vary from information exchange to the development and implementation of technology. This report presents the results of one Annex.

The Programme is governed by an Executive Committee, which monitors existing projects and identifies new areas where collaborative effort may be beneficial.

Disclaimer

The HPT TCP is part of a network of autonomous collaborative partnerships focused on a wide range of energy technologies known as Technology Collaboration Programmes or TCPs. The TCPs are organized under the auspices of the International Energy Agency (IEA), but the TCPs are functionally and legally autonomous. Views, findings and publications of the HPT TCP do not necessarily represent the views or policies of the IEA Secretariat or its individual member countries.

The Heat Pump Centre

A central role within the HPT TCP is played by the Heat Pump Centre (HPC).

Consistent with the overall objective of the HPT TCP, the HPC seeks to accelerate the implementation of heat pump technologies and thereby optimize the use of energy resources for the benefit of the environment. This is achieved by offering a worldwide information service to support all those who can play a part in the implementation of heat pumping technology including researchers, engineers, manufacturers, installers, equipment users, and energy policy makers in utilities, government offices and other organizations. Activities of the HPC include the production of a Magazine with an additional newsletter 3 times per year, the HPT TCP webpage, the organization of workshops, an inquiry service and a promotion programme. The HPC also publishes selected results from other Annexes, and this publication is one result of this activity.

For further information about the Technology Collaboration Programme on Heat Pumping Technologies (HPT TCP) and for inquiries on heat pump issues in general contact the Heat Pump Centre at the following address:
Heat Pump Centre
c/o RISE - Research Institutes of Sweden
Box 857, SE-501 15 BORÅS, Sweden
Phone: +46 10 516 53 42
Website: https://heatpumpingtechnologies.org

Acoustic Signatures of Heat Pumps

IEA HPT Annex 51

6: Annoyance rating and psychoacoustical analysis of heat pump sound

Henrik Hellgren, RISE, Sweden
Christian Kaseß, ARI, ÖAW Austrian Academy of Sciences, Austria

Date: April 2020	Final Version, Review: 04

1 Introduction

Noise from heat pumps has a potential to cause annoyance of owners and neighbours, the degree of which is influenced by several factors. Factors like the acoustic characteristics of the noise, the installation and placements of the heat pump unit and individual's sensitivity to noise. To further increase the acceptance of heat pumps it is important to reduce this noise induced annoyance. In order to achieve this, detailed knowledge of how acoustic parameters influence the annoyance is necessary. An important topic in environmental noise research is to find proper ways to assess and quantify the changes in sound caused by such measures and to understand the relation to noise perception. Most commonly the A-weighted level is used to relate sound to annoyance. But, other acoustic parameters may better explain the level of annoyance. These parameters could be the presence of low frequency noise and tonality, which the A-weighted level inefficiently assesses. Common parameters used to assess the subjective perception of noise is loudness, sharpness, roughness, and tonality. From traffic noise, it seems clear that other noise descriptors more related to human perception such as loudness can yield a better description of the annoyance experienced by environmental noise e.g. [1] [2].

A way to assess the annoyance of noise sources is to perform listening tests. In this way it is possible to gain knowledge of the acoustic parameters that influence the annoyance. A possible drawback of these tests is that it often requires to use short sound stimuli, making it difficult to assess long term annoyance. A desired result is an annoyance index that show how different acoustic parameters explain the assessed annoyance response. Development of an annoyance index of heat pump noise could be beneficial when setting regulatory demands for heat pump noise.

This report summarize as a study of annoyance related to air source heat pump noise, were a listening test was performed on an Austrian and Swedish listening panel. The report additionally present results from previous studies dealing with different aspects of heat pump noise: an Austrian study investigating the effects of different noise mitigation measures on perception and also a Swedish study investigating the noise perception of ground source heat pumps.

2 Annoyance rating of air source heat pump sound

2.1 Listeners

2.1.1 Austria

20 normal hearing listeners (10 female) were tested. The mean age was 29.7±6.8 years. All but one listener had hearing thresholds less than 20 dB higher than normal thresholds for all frequencies tested. A single listener had a single sided increase in hearing threshold of 30 dB at 8000 Hz but had otherwise normal hearing.

2.1.2 Sweden

20 normal hearing listeners (10 female). The mean age was 46.0 ± 9.5 years. Two listeners self-reported a small general hearing deficiency.

2.2 Recording procedure

The recordings were made in an hemi-anechoic room. The unit was an air-to-air heat pump with a heating capacity of approx. 6 kW at nominal condition. The recordings were made with free field microphones and a sampling frequency 51 200 Hz. The operation was controlled by adjusting the setting of the indoor unit fan speed. The recordings were made at five different operating conditions summarized by Table 2-1. The sound power level was determined according to ISO 3744. Each operating condition was recorded simultaneously at four microphone positions. The position at the right side was closest to the location of the compressor. The distance between the microphone and the unit was 1 meter. The microphone setup is shown in Figure 2-1. The recordings were 30 seconds long from which 5-second long sound samples were extracted and equalized to 40 dB(A) to be used in the experiment.

Setting	**Compressor speed [Hz]**	**Fan speed [rpm]**	**Input power [kW]**	**A-weighted sound power level, L_{WA} [dB]**
Low	34	610	0.78	52.7
Medium	48	770	1.09	56.5
High	73	770	1.76	59.1
Super high	79	770	1.9	58.2
Emergency *	58	770	1.28	57.6

Table 2-1: List of recorded heat pump settings including fan and compressor speed and the measured A-weighted sound power level (according to ISO 3744).

** Emergency setting is a pre-defined program for test operation*

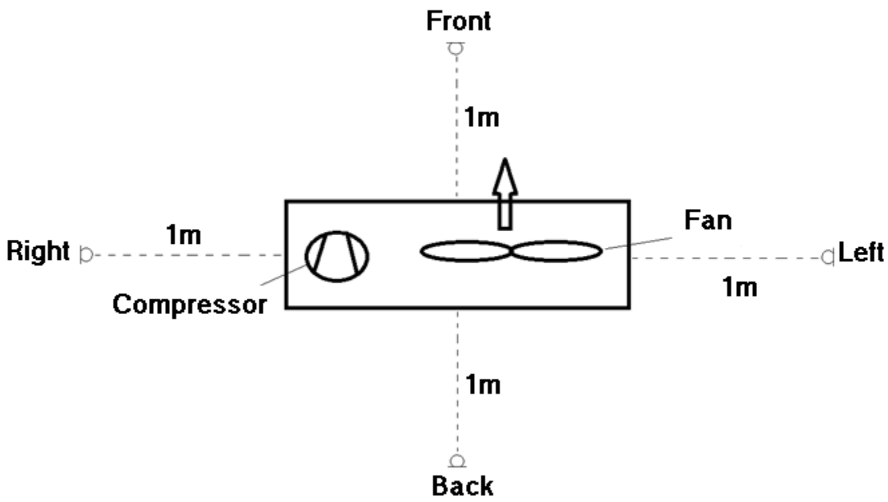

Figure 2-1: Microphone setup for acoustic measurements.

2.3 Annoyance rating

A free magnitude estimation was performed to determine the annoyance ratings of the different heat pump noises [1] [2], thus the listeners judged the relative annoyance rather than an absolute impression which is highly context dependent [3]. After listening to the stimulus via headphones, listeners were asked to input a numerical rating corresponding to the perceived annoyance. While listeners were free in choosing their starting value, they were instructed to avoid extremely high or low starting values in order to stay within a comfortable range of numbers. Listeners were asked to perform a proportional rating, i.e. double the annoyance should result in doubling the value. Listeners were also instructed not to use 0 or negative numbers. They were also explicitly told to keep their rating scale constant within and across all runs. Once the rating was entered, listeners continued by pressing a key.

Before the main test, listeners received written instructions containing the definition of annoyance and a description of the procedure in the respective language. For this the instructions were first derived in English and then translated into German and Swedish.

Annoyance was defined as a feeling of discomfort, caused by noise or a feeling of aversion, discomfort, or irritation if the current activity is disturbed or affected by noise. Listeners were also asked to base their annoyance rating on imagining how annoying and distracting they would find the noise, if they were subjected to it on a regular basis [4] [5]. After reading the instructions, listeners performed a training covering a range of stimuli. The training consisted of a few trials, after which listeners were allowed to adapt their rating range in the case they felt uncomfortable with their initial choice. After the training, listeners had the opportunity to clarify open issues.

The experiment was performed in three runs in which each stimulus appeared three times. For each listener and run, stimuli order was randomized. Between runs a break of at least 5 minutes was enforced.

2.4 Data analysis

2.4.1 Psychoacoustic and acoustic parameters

Acoustic as well as psycho-acoustic parameters of the 5-second long sound samples were calculated using the Matlab-toolbox psysound3 [6]. These quantities encompassed loudness based on the Glassberg und Moore model [7], psychoacoustic roughness [8], tonality [9], sharpness, and loudness fluctuation [10]. Furthermore, C-weighted sound pressure levels (time-weighting fast) were calculated. The median as well as the 5%-percentile (the value that is exceeded 5% of the time) were calculated, denoted e.g. as S_{50} and S_5 for the median and the 5% sharpness. The loudness level in phon was also determined. For all segments, the first 500 ms were discarded to avoid systematic errors due to transient response of the models.

2.4.2 Preprocessing

Three listeners reported a total of four input typos all of which were reproducible and could be corrected. As the magnitude estimation leads to a ratio scale, we applied the logarithm of base 2 on the data. Thus an increase by 1 in the log-ratings implies a doubling of the perceived annoyance. No datapoints were detected outside the 3-fold standard deviation across the subject data. The overall consistency of the ratings was good.

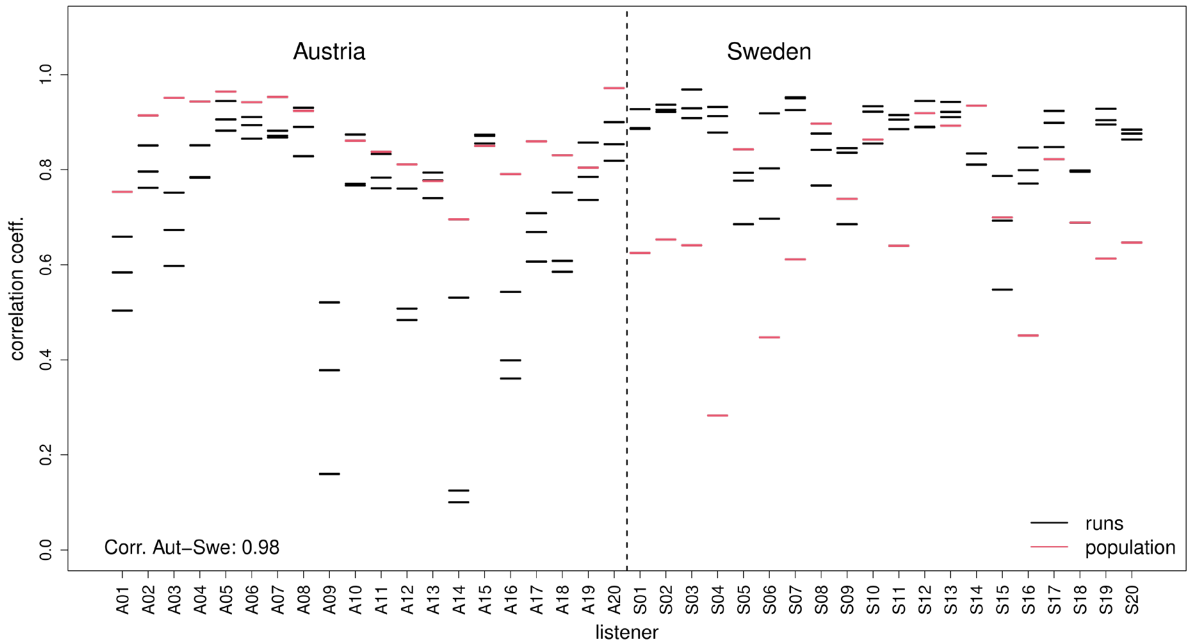

Figure 2-2: Correlation coefficient per listener between runs (black) and for the average population rating (red)

2.4.3 Consistency of the ratings

Figure 2-2 shows the correlation between the runs per listener (3 combinations, black symbols) as well as the correlation of the average response per listener to the full population response (pooled data from both sites, red symbols). It is clear that most listeners were able to produce consistent ratings across runs. Notably, in the Austrian data at least two listeners seemed to have some difficulties exhibiting very low correlations. For listener A14 the ratings in the third run did not correspond well to the first and second run which may be a sign of fatigue. The mean rating was still somewhat related to the group average. Listener A09 also showed low between-run correlations and a low negative correlation to the population. In the Swedish data, on average a high intra-listener consistency was observed. Compared to the overall mean, however, a number of listeners showed correlations of 0.6 and lower. When comparing the grand mean over all subjects from Sweden and Austria, a high correlation of 0.98 was observed.

For this and further analysis, mean log ratings per listener and condition were calculated and the grand mean per listener was subtracted in order to normalize the data. (see [1] [2]). For the group mean the average across all listeners were calculated per condition.

2.4.4 Statistical analysis

Statistical analysis was performed using the software R [11]. The mean log-ratings per listener and condition were the input for a repeated-measures-analysis-of variance (RM-ANOVA) with operating condition and direction as factors. Furthermore, the site of the experiment (Austria, Sweden) was included as a between-subject factor. The R-package *afex* was used for this purpose [12]. For significant effects omnibus post-hoc tests were performed using multivariate testing using *emmeans* (Estimated Marginal Means (Least-Squares Means)) [13].

Furthermore, the relation between acoustical properties and the annoyance rating was investigated using a linear regression. As the use of a stepwise model selection [14] on the pooled data based on the Bayes Information Criterion (BIC, [15]) lead to relatively poor models all possible model permutations for up to 5 explanatory variables were estimated and compared (see Section 2.5.3).

2.5 Results

2.5.1 Psychoacoustical and acoustical quantities

Figure 2-3 shows the median and inter-quartile-range of the different acoustical quantities over time as a function of position after equalization to 40 dB(A). For the L_C and to a lower degree also the loudness (N) and loudness level (L_N) the different operating condition show most clearly up in the right position, where the compressor was located. There is also a visible effect of position for these two quantities, whereas for sharpness (S) the fluctuations are in the range of the effect. Roughness (R) is slightly elevated for the low condition and in the right position high and superhigh mode produce elevated roughness. For tonality only two conditions lead to non-zero peak values and only one condition produced non-zero values for at least a quarter of the time. Figure 2-4 shows the same data as Figure 2-3 arranged as a function of operating condition.

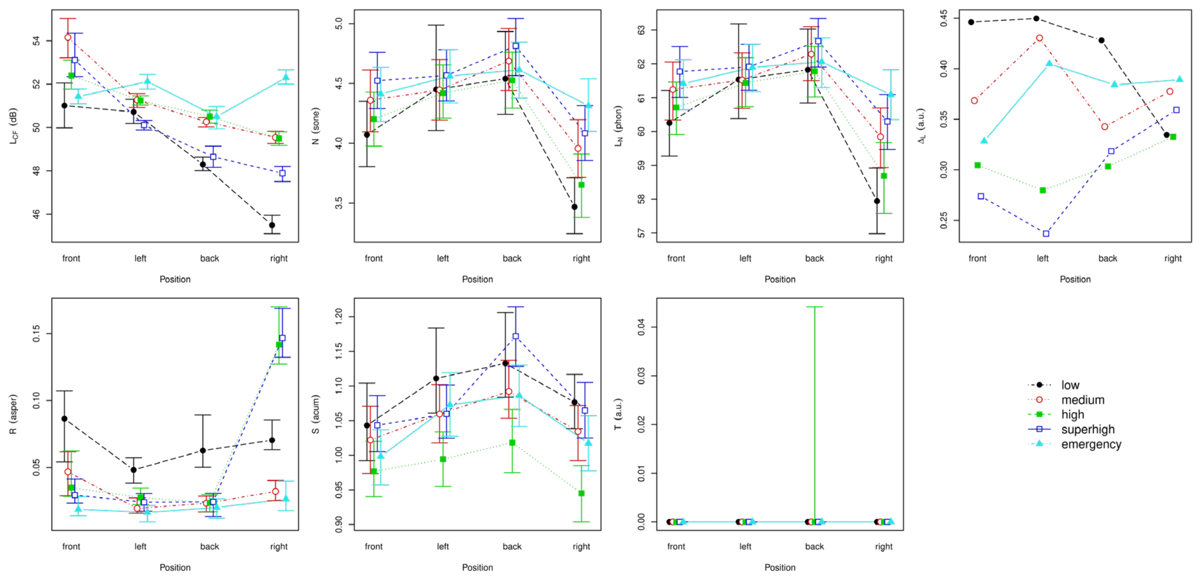

Figure 2-3: Acoustic descriptors as a function of position. Operating condition is shown as different colors.

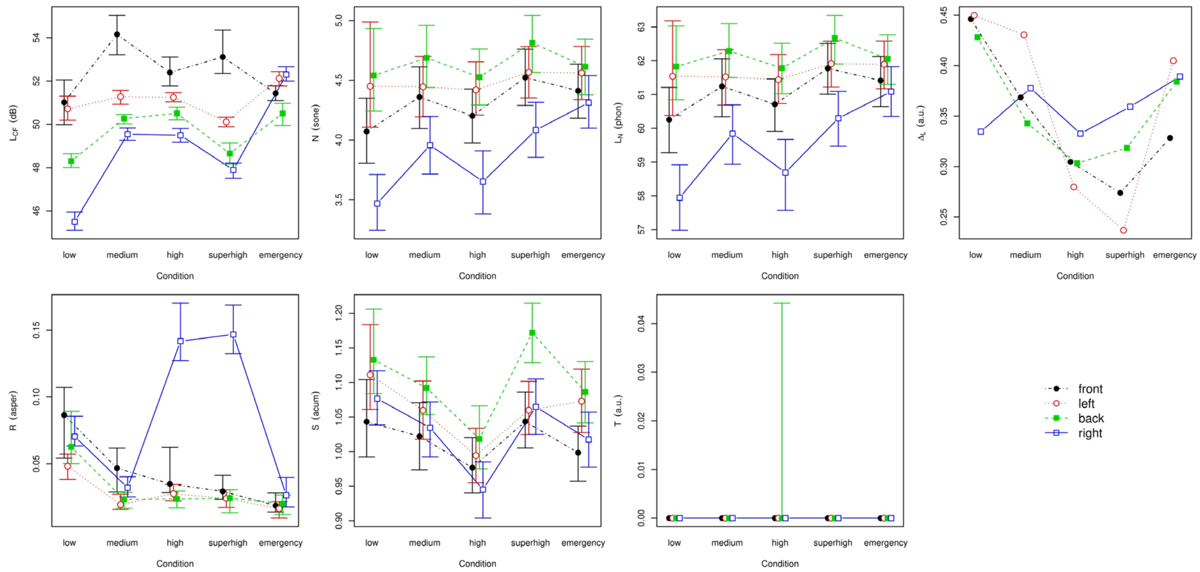

Figure 2-4: Acoustic descriptors as a function of operating condition. Recording position is shown as different colors.

2.5.2 Annoyance

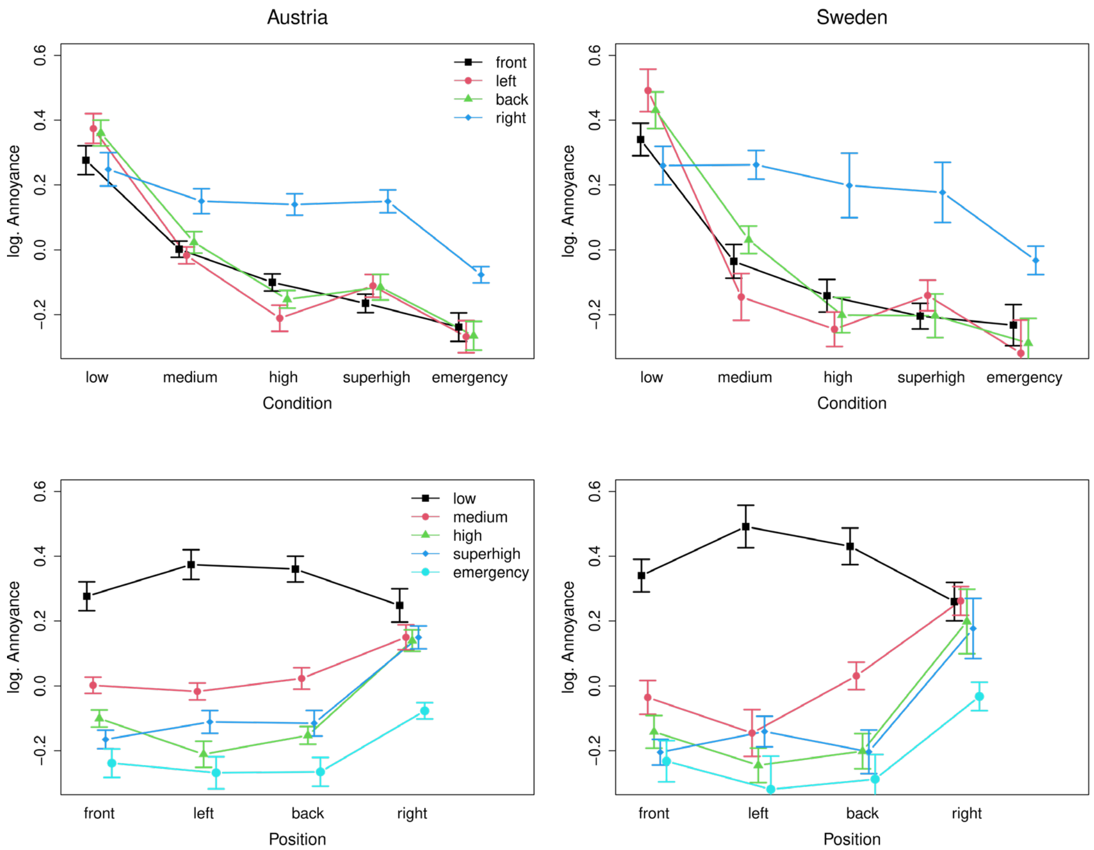

Figure 2-5: Mean and standard error of the annoyance as a function of operating condition and position. Left panels show the Austrian data, the right panels the Swedish data.

Figure 2-5 shows mean and standard error of the annoyance ratings across the Austrian study population (left panels) and the Swedish data (right panels). Comparing the two sites, some differences can be seen (e.g. left position). However, the general trends are similar enough for the factor site to yield no significant interaction effects with either position, condition, or both. The main effect of site is also not significant which is a consequence of normalizing the log-ratings of each listener to zero mean as the absolute scaling (a shift for the log-ratings) in a free magnitude estimation is of no consequence.

The ANOVA yielded significant main effects of position and condition as well as a significant interaction between the two factors ($p<0.0001$ for all effects). Mauchly's test for sphericity showed a significant deviation from the sphericity assumption, thus a Greenhouse-Geisser correction was applied [16]. After correction, all effects were still significant with $p_{GG}<0.0001$.

For four levels in position and five levels in condition a total of 60 possible pairwise interactions exist for which a post-hoc analysis was performed. P-values were Bonferroni-corrected, i.e. with the number of post-hoc test performed. The main results of this analysis is that all 16 significant interactions include either the right position or the low condition or both. Thus, main effects containing either of these levels have to be treated with caution.

2.5.2.1 Main effect position

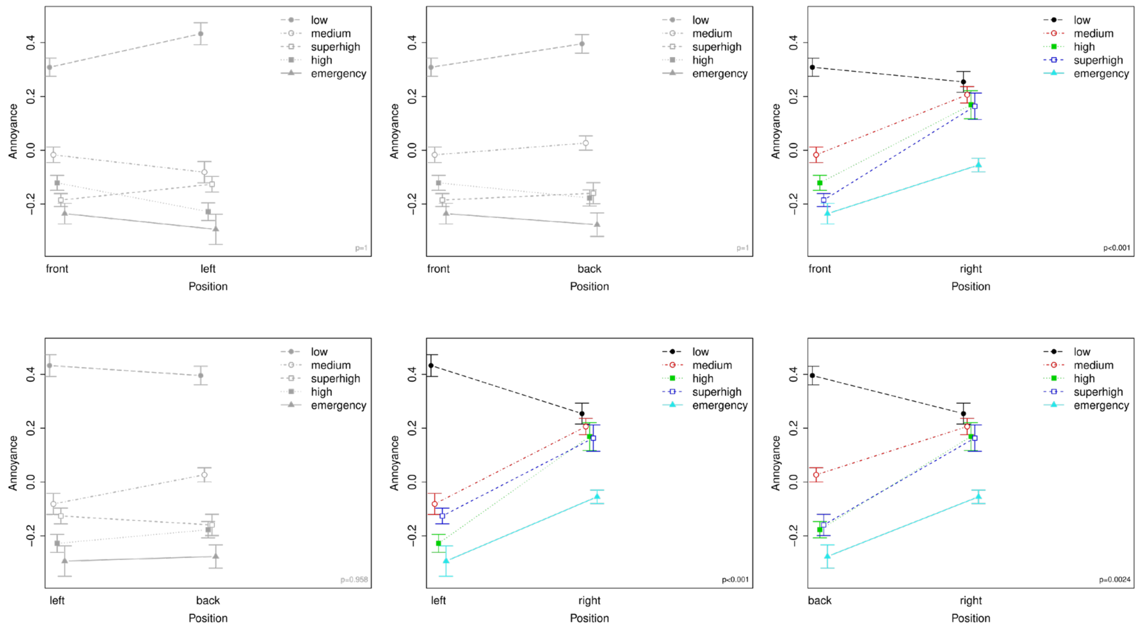

Figure 2-6: Pairwise post-hoc results for the main effect position. Gray plots are not significant after correction (p>0.05).

A post-hoc test on all possible main effects between 2 positions shows, that the recording from the right position is significantly more annoying than all other positions. However, clearly when looking at the different contributions of the condition (Figure 2-6), the low condition has the opposite effect which is also significant for all pairwise interactions between the respective positions and the remaining conditions. Due to this significant qualitative interaction effect the position effect cannot be properly interpreted as such.

2.5.2.2 Main effect conditions

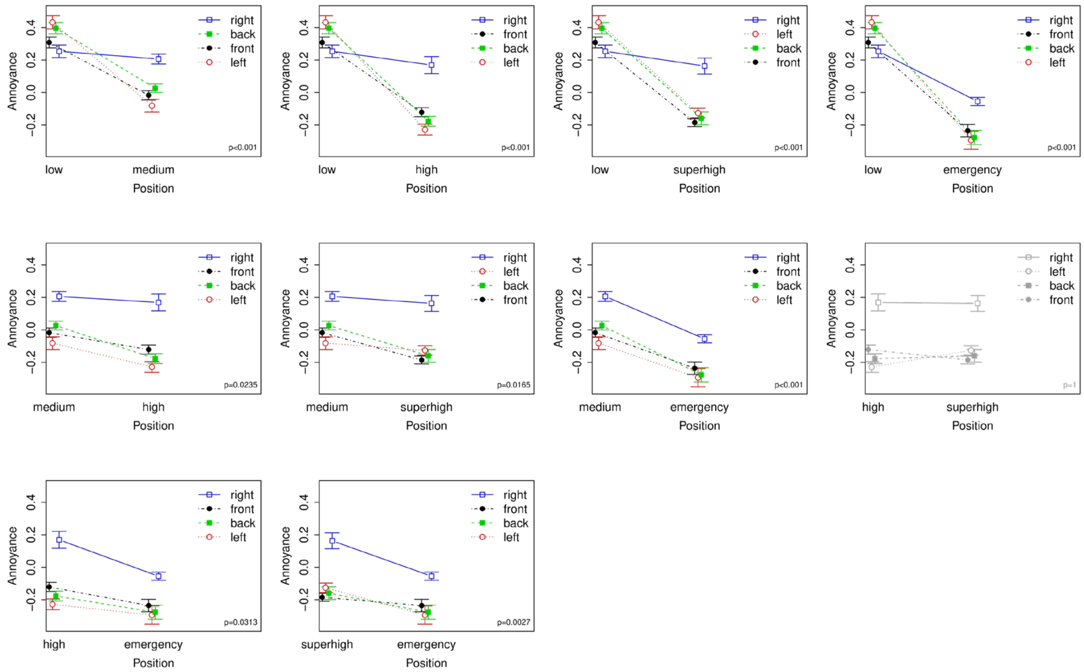

Figure 2-7: Pairwise post-hoc results for the main effect condition. Gray plots are not significant after correction (p>0.05).

Figure 2-7 illustrates the post-hoc results for the main effect condition. Most conditions are significantly different from each other with the exception of superhigh and high condition. Although some significant interaction effects are also significant, all these interactions are of a quantitative nature, i.e. they do not alter the direction of the effect. For example, for low-medium the right position shows an interaction with all other positions, however still the low condition is more annoying than the medium condition for all positions. Some of the effects (low-medium, medium-high, and medium-superhigh) could be considered borderline as the annoyance is almost constant for the right position.

2.5.3 Annoyance index

The statistical analysis showed significant differences between the conditions and positions. An approach was used where all possible models with up to 5 acoustical parameters were estimated. As no significant effect of the experimental site was detected, this analysis is based on the pooled data.

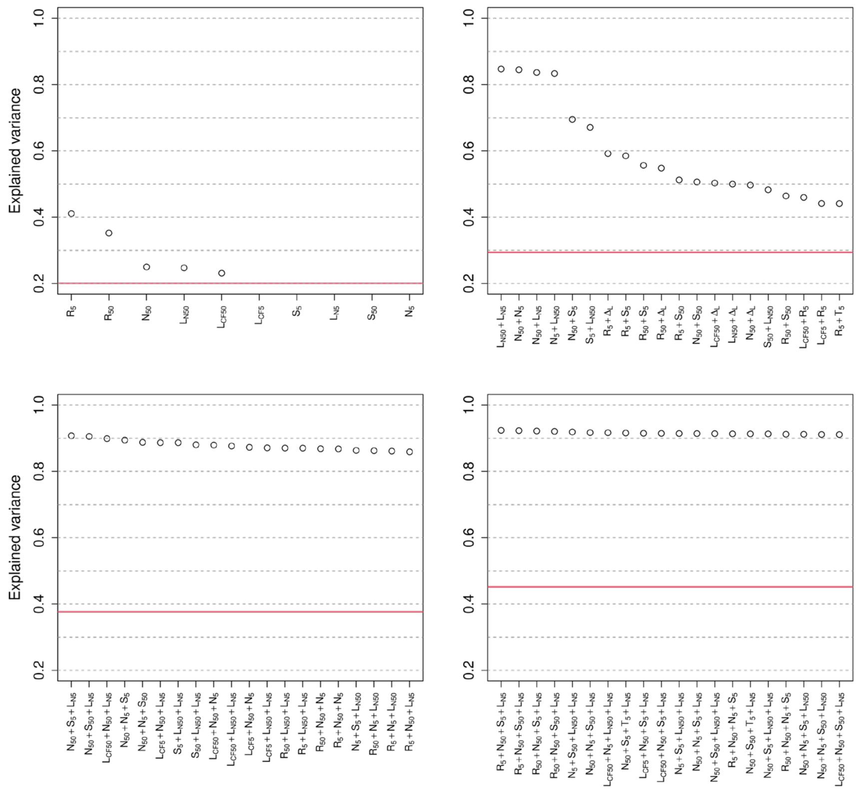

Figure 2-8: Best models for combinations of up to 4 acoustical parameters. The red line indicates the 95% value for the explained variance if purely random parameters were used.

Figure 2-8 shows, that no single variable explains more than about 40 % of the variance whereas a combination of peak and median loudness or loudness level already explain more than 80 %. Using 4 parameters, peak sharpness as well as peak roughness together with peak loudness level and median loudness lead to the best results, although differences to the next best models are relatively minor. Using a fifth variable, adding the peak tonality yields the best model. [1]

[1] In a preliminary analysis of the Austrian data only, these same variables were found using a stepwise model selection. On the pooled data, however, the stepwise selection ran into a suboptimal local optimum

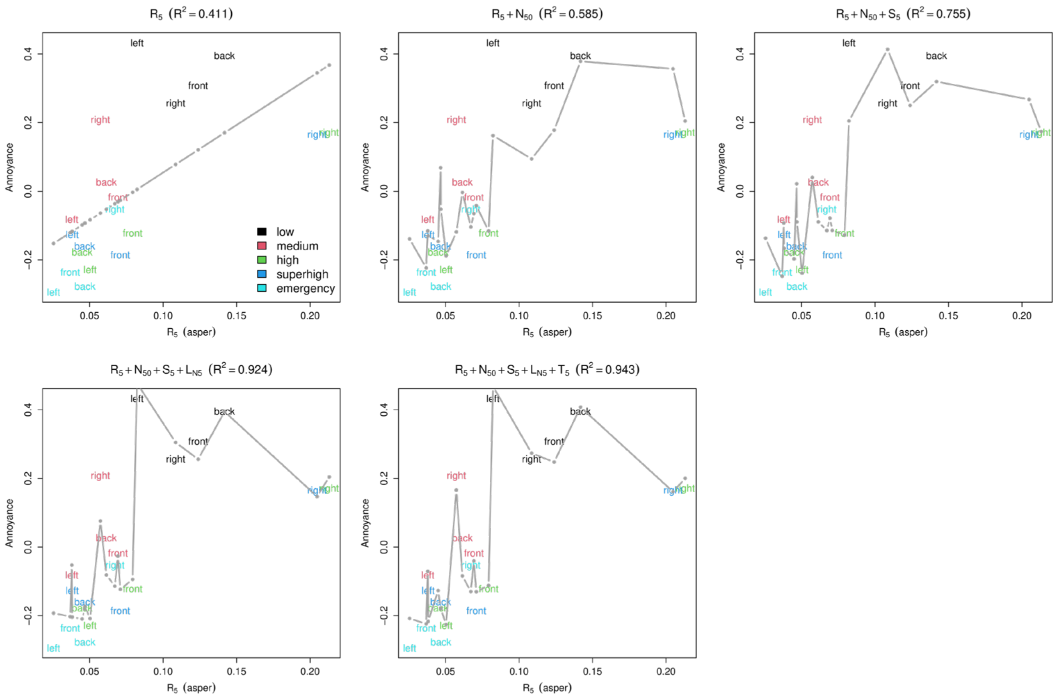

Figure 2-9: Step-wise parameter addition. The response is plotted vs. the single best descriptor.

Figure 2-9 shows the stepwise addition of these variables in order of the highest additional contribution. The best model consisted of the following 5 descriptors:

$$\textbf{log annoyance} = \textbf{const} + 1.0153 \cdot R_5 + 1.2272 \cdot S_5 - 1.3313 \cdot N_{50} + 0.2985 \cdot L_{N5} + 1.5560 \cdot T_5$$

The single best descriptor is the peak psychoacoustic roughness which explained about 40% of the variance. Peak sharpness, median loudness and peak loudness level explain roughly an additional 20%, 17% and 15%, respectively. Peak tonality had only a minor effect on the explained variance.

2.6 Summary

The differences of annoyance of different operating conditions and recording directions were investigated.

An important result is that no significant effects of the tested population was observed. This indicates that results from different sites can be compared and potentially generalized. However, data from more different sites would be necessary to further support this conclusion. Initially this was planned within the Annex 51, however, the Covid-19 pandemic measures in various countries made it impossible for other sites to collect experimental data as planned.

A main effect of the operating condition on the annoyance was observed. In particular the low-compressor speed condition was judged the most annoying whereas the emergency condition was judged less annoying than any other condition. In between the effects were minor. As the

A-level was equalized to 40 dB for all conditions, this implies that the sound characteristics of the low setting were most annoying. However, the sound power level is also the lowest, with up to about 6 dB lower than for other conditions. Still, the implication is that the emission level of operating modes with similar characteristics as the low-condition needs particular attention.

Position does not lead to a consistent main effect since interactions between position and condition are present. In comparison to all other operating conditions the low-condition produces comparatively low annoyance ratings for the right measurement position and thus result in this interaction effect.

From the annoyance index the low annoyance of the emergency mode could be explained by a low psychoacoustic roughness. However, for the low-condition roughness as well as sharpness do not seem to be the main contributing factor. Adding loudness seems to improve in particular on the fit for this condition.

3 Annoyance rating of air source heat pump sound with different mitigation measures applied

Noise mitigation measures that only by a small degree affect energy efficiency are thus an active field of research, e.g. [17] [18]. An important topic in environmental noise research is to find proper ways to assess and quantify the changes in sound caused by such measures and to understand the relation to noise perception. Most commonly the A-weighted level is used to relate sound to annoyance. However, from traffic noise, it seems clear that other noise descriptors more related to human perception such as loudness can yield a better description of the annoyance experienced by environmental noise e.g. [1] [2]. In a previous study the noise emission of an air-to-water heat pump was investigated for four variants and four directions. The variants under investigation comprised the heat pump without any modifications, a diffuser attached to the fan outlet and an acoustic deflection with and without a splitter-type silencer. Using emission recordings from four different directions the effect of the variant and the directivity were investigated by means of various acoustic quantities and a perception experiment in the lab. This data has been published in [19].

3.1 Listeners

20 normal hearing listeners (10 female) were tested. The mean age was 28.6±6.6 years. All but one listener had hearing thresholds less than 20 dB higher than normal thresholds for all frequencies tested. A single listener had a single sided increase in hearing threshold of 30 dB at 8000 Hz but had otherwise normal hearing.

3.2 Recording procedure

Measurements were performed in a climate chamber with absorbing walls and a reflecting ground. Four microphone positions were chosen around the heat pump in a height of 127 cm above the floor. One recording position was located at the fan axis inlet (0°). The second position was located at the fan outlet (180°). Two further positions were located along an axis perpendicular to the fan axis on either side of the head pump.

Stationary noise samples were extracted from the recordings and investigated as a function of the four different heat pump variants. To guarantee comparable operating states for the different mitigation measures, sound samples were extracted about 60 seconds after the end of a defrosting cycle. The duration of the audio segments used for analysis and psychoacoustical testing was chosen to be 5 s to allow the listeners to properly assess the sound and at the same time avoid a lengthy investigation that might impact on the listeners focus. Sounds were presented via headphones.

All 16 conditions (four variants, four directions) were used in the test and repeated 9 times. Furthermore, 8 samples containing pink noise at different A-levels within the range of the heat pump noise were included. The pink noise should allow to compare the results with those from other studies (cf. [20]).

Determination of psychoacoustic and acoustic parameters as well as the perception test were performed as described above.

3.3 Results

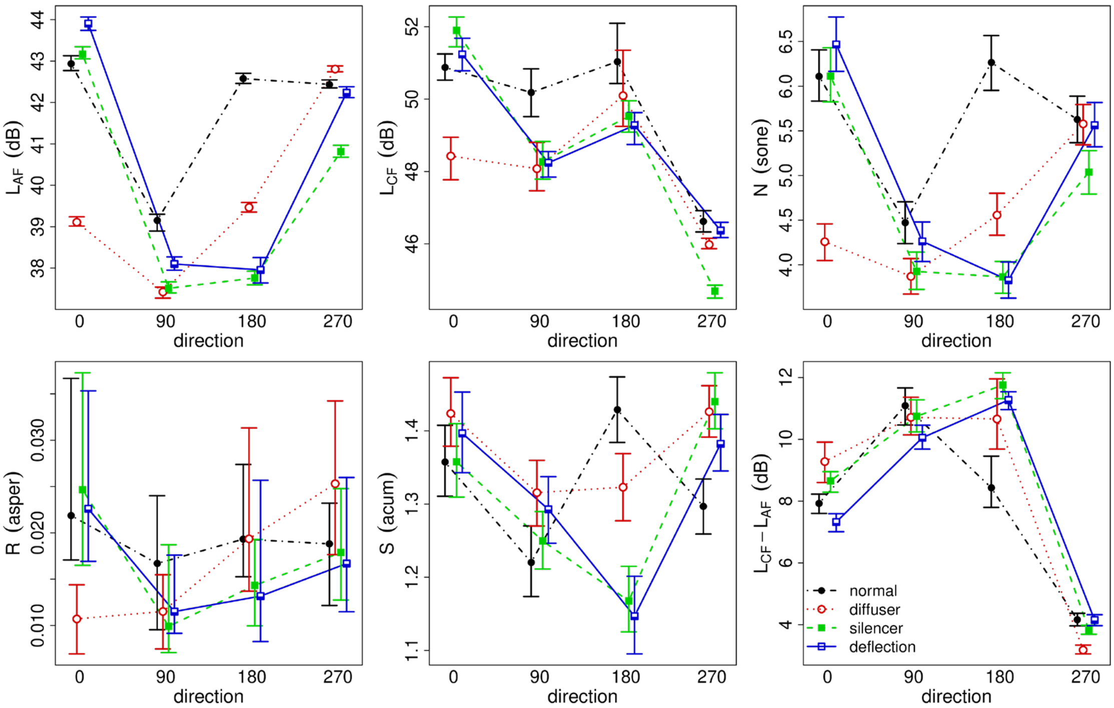

Figure 3-1: Acoustical and psychoacoustical parameters as a function of direction and variant (taken from [19])

The highest values for A-level and loudness (Figure 3-1, median and inter-quartile range) at the fan outlet were observed in the reference variant (i.e. no measures applied). At the fan inlet only the diffuser caused a difference in level. At the sides the different variants only resulted in minor changes, however, at 90° the A-weighted levels and loudness were much lower than at 270°. The smaller difference between C- and A-level at 270° indicates a small low-frequency contribution for that direction.

A change in sharpness was observed at the fan outlet with the reference variant and the diffuser exhibiting higher values. Differences for the roughness were mostly minor at 0° for the variant with the diffuser. Overall, the effect of the condition seems smaller for the side positions than for the positions at the fan inlet and outlet.

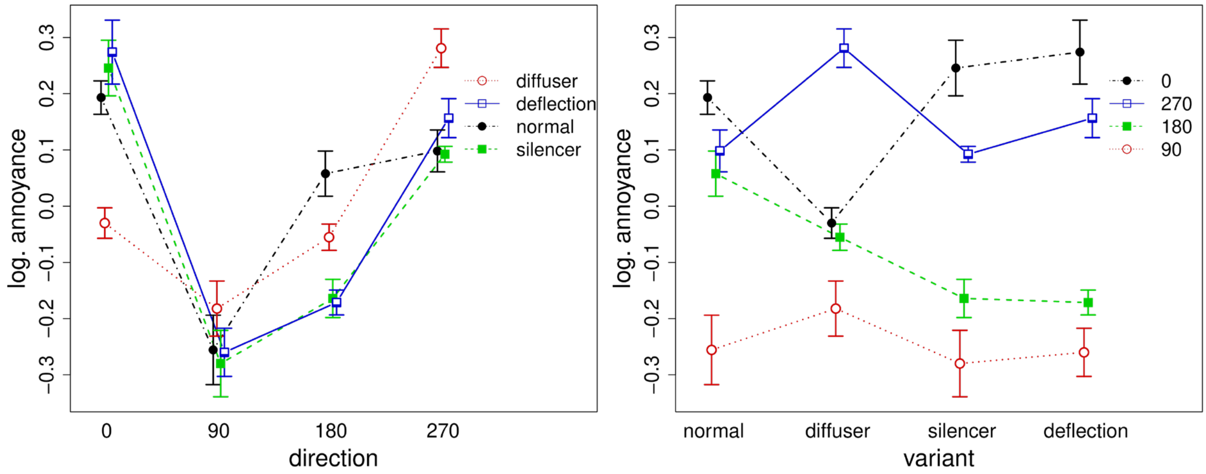

Figure 3-2: Logarithmic annoyance as a function of direction and variant (taken from [19])

Concerning the annoyance ratings (Figure 3-2) a repeated measures ANOVA was performed. A significant main effect of direction ($p<0.0001$) as well as a significant interaction for variant and direction ($p<0.0001$) was observed. No significant main effect of the variant was observed which seems to be a consequence of the interaction with the measurement direction. For details see [19] .

For the direction effect all directions pairs were significantly different except 0° (fan inlet) and 270° (heat pump side). The 90° direction (opposite side) was rated less annoying than any other direction whereas the fan inlet was rated to be more annoying than 90° and 180°.

The question that remains is, which acoustical quantities explain the differences in ratings. A stepwise linear regression yielded the model considering the median A-level L_{AF50},the peak loudness level L_{N5}, the peak sharpness S_5, the peak loudness N_5, and the median psychoacoustic roughness R_{50}. All loudness related quantities (N_5, N_{50}, L_{N5}, L_{N50}) explain between 75 % and 78 % of the variance whereas the L_{AF50} explained 88 %.

3.4 Summary

Summarizing, the acoustic effects of the different noise mitigation measures was heavily direction dependent. This was also observed in psychoacoustical parameters, in particular in sound pressure levels, loudness, as well as sharpness. This dependency between variant and emission direction was, however, difficult to interpret.

The directivity was also observed for the annoyance, although caution is recommended when interpreting these differences due to the significant interaction with the factor variant. No clear consistent effect of the variant was observed in the annoyance ratings. Rather, different mitigation measures resulted in various effects, depending on the microphone position. This indicates that for noise measures to be effective, care needs to be taken about the positioning of the heat pump. However, the recordings were done close to the heat pump. The direction dependence will most likely be less pronounced, when considering a realistic setting with reflections from surrounding structures. Additional tests are necessary to investigate this issue in more detail.

The A-level and the loudness level explained the annoyance ratings to a high degree. In particular sharpness showed some improvement for the explained variance of the model.

4 Multidimensional scaling of experiences from geothermal heat pumps

To determine the most salient parameters influencing perception of geothermal heat pumps and the corresponding level of annoyance a dissimilarity rating was conducted along with a preference mapping. Dissimilarity ratings are powerful tools to obtain a multidimensional scaling of the stimuli, free of the restrictions imposed by predetermined scales or response criteria. It builds on the limited ability of the listener to only focus on a set of varying parameters [21]. To determine the prevalent or dominant perceptual features in different geothermal heat pumps the dissimilarity rating conducted included three different models and 10 different recordings of varying situations (3+3+4 of the three models). The corresponding multidimensional map was compared with specific psychoacoustic parameters as well as rated level of annoyance.

Few studies have been conducted on the perception and experience from geothermal heat pumps. A study by Persson Waye and Rylander compared heat pumps and ventilation systems dominated by lower frequencies (<200Hz) and heat pumps and ventilation systems dominated by mid frequencies [22]. The results showed that people exposed to low frequency noise from heat pumps were more annoyed and had a higher level of disturbed concentration than those exposed to the noise of mid-frequency character. Wang and Novak analysed several different heating, ventilation and air-conditioning systems, they determined that high sound levels (>50 dBA), excessive low frequency rumble and larger timescale fluctuations (e.g., a heat pump cycling on and off every 30 seconds) were the dominating characteristics influencing levels of annoyance [23].

Annoyance to mechanical systems related to heating appear to often be related to the dominance of low frequency content. Broner and Leventhall proposed using the difference between A-weighted SPL and C-weighted SPL that values greater than 20 dB would signify a low frequency noise problem [24]. Holmberg et al suggested that the problem would occur already at 15 dB. In the present listening test three stimuli had a greater difference than 15 dB (a1, b1, and c2) whereof one had a greater difference than 20 dB (c2) [25].

4.1 Listeners

In the listening test 14 people participated, 4 women and 10 men (M= 40 years old, s.d. = 9 years). 1 participant did not comply with the instructions and was removed from further analysis. 1 participant reported hearing problems, but that did not affect the results.

4.2 Recording procedure

Three different geothermal heat pumps were used in the experiment. Each heat pump was represented by three or four different recordings. In total 10 stimuli were utilized. All stimuli were 3 seconds long and presented at 42 dB(A). The sound pressure level choice was made as the current labelling is done using dB(A) levels instead of loudness measures.

Model	C-weighted level (dB(C))	Loudness (sone)	Roughness (asper)	Sharpness (acum)	Compressor speed (Hz)	Rated preference
a1	60.21	3.78	0.066	1.85	45	3.9
a2	55.78	3.63	0.125	1.51	50	2.8
a3	48.49	4.44	0.056	1.71	57	6.3
b1	61.85	3.65	0.018	1.25	41	4.9
b2	53.73	3.77	0.071	1.47	52	5.4
b3	40.71	4.14	0.030	1.33	110	7.0
c1	57.38	3.85	0.096	1.17	58	3.6
c2	63.49	3.92	0.062	0.88	69	6.5
c3	56.75	3.78	0.080	0.88	83	4.8
c4	48.87	3.73	0.094	1.19	100	6.3

Table 4-1: Psychoacoustical properties and settings for the different heat pumps.

4.3 Dissimilarity and preference rating

The listening test took place in a 3rd order ambisonics lab with little visual distraction. The sounds presented were mono sounds presented using the two front speakers. Each participant performed pairwise ratings of dissimilarities between the different sounds using a sliding scale. In addition the participants marked which of the sound in the pair s/he preferred. Using a half-matrix design (testing all possible pairs in one direction) this resulted in 45 pairs. To be noted: the participants were aware of the sounds coming from different geothermal heat pump systems as it could affect their choice of preference.

4.4 Results

The dissimilarity ratings were analysed using the individual difference scaling (INDSCAL) model. The INDSCAL model assumes that all participants share the same psychological scale but attends differently to the underlying psychological dimensions (Ashby et al, 1994). An advantage with the INDSCAL model is that it provides a unique configuration solution that requires no further rotation of the model [26]. The analysis resulted in a 2-dimensional model (Stress=.131). Stress values <.133 were considered acceptable as determined by Sturrock and Rocha [27]. The MDS solution is presented in Figure 4-1 labelled by their model (a-c).

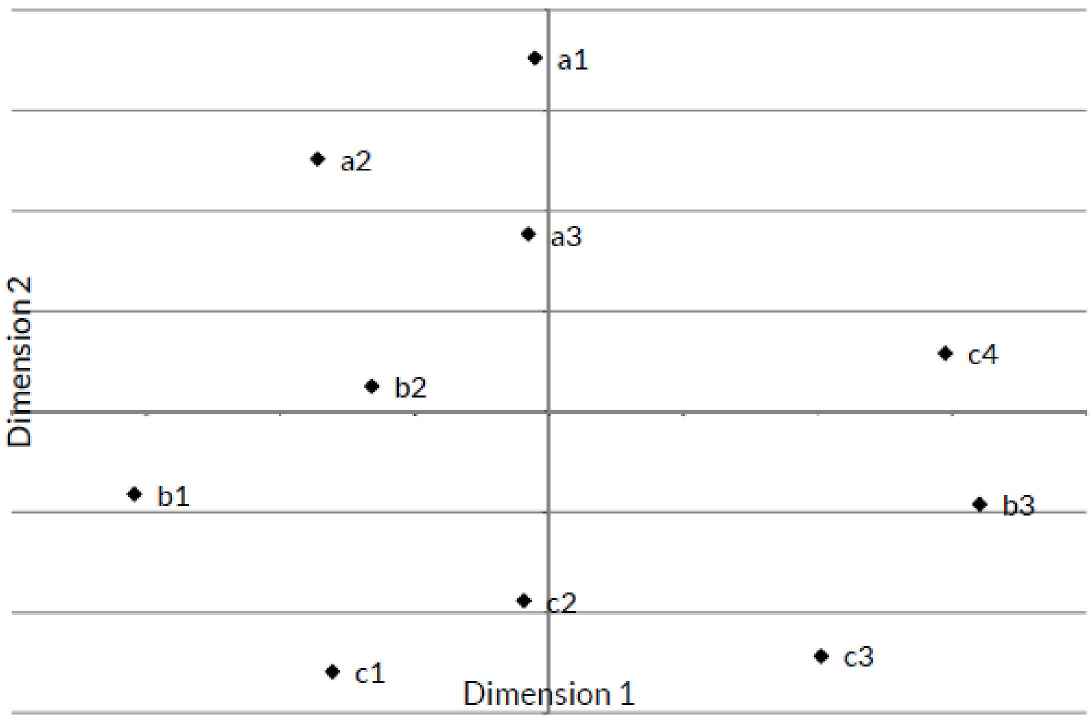

Figure 4-1: The MDS solution labelled by their brand.

The two dimensions were analysed using the preference ratings of the listening test, the psychoacoustic parameters and the compressor speed. The preference ratings are listed in Table 4-1. The results showed that Dimension 1 is partly explained by the preference mapping (*R^2 adj.=.32, F=5.2, p<.05)* but mostly by the compressor speed (*R^2 adj.=.83, F=46.2, p<.001*). Dimension 2 is explained by the variance in sharpness (*R^2 adj.=.74, F=27.0, p<.001*). The other psychoacoustic parameters showed no significant relationship with either dimension. Regression analyses further showed that the preference mapping could be explained by both compressor speed and sharpness (*R^2 adj.=.67, F=10.0, p<.01*), the participants preferred sounds with less sharpness and a compressor speed at higher frequency.

4.5 Discussion

Little of previous research on the sounds of heat pumps have focused on other aspects than low frequency content and tonality. This experiment is a first step to further distinguish the dominating parameters to explain perception of ground source heat pumps. Creation of perceptual maps require an inter comparison between the specific stimuli used in the experiment. The result will thus depend on which stimuli are used. The aim of the experiment was to use as different heat pump sounds as possible to set a ground work for later experiments on finding the parameters explaining annoyance for heat pumps. The experiment was limited to ground source heat pumps, as we believe that air source heat pumps has a distinct different sounds, the latter will instead be evaluated in a later experiment.

The low frequency content did not influence the level of annoyance. This might seem surprising, but Holmberg et al [25] proposed that the difference between the C-weighted and the A-weighted SPL may be limited as predictor of annoyance when the overall noise level is too low. This could be a reason to the lack of connection between annoyance and dB(C) in the present study. However, most ground source heat pumps hold a relatively low sound pressure level, indicating that dominating low frequency character might not influence the annoyance level to a higher degree.

Dissimilarity ratings require the use of shorter sound stimuli to enable comparison between presented pairs. This makes it difficult to discern whether fluctuations in the heat pumps influence level of annoyance. For future experiments longer stimuli are needed.

The results showed that the most salient parameters are compressor speed and the sharpness level. Both have a significant impact on annoyance responses to the ground source heat pumps. To further evaluate whether fluctuations also influence annoyance longer stimuli are needed.

5 Conclusions

Overall, the results show that, in addition to the A-weighted level other acoustical parameters such as loudness, roughness, and sharpness may help to better model the perception of heat pump noise. But, it is difficult to find a single descriptor to explain the variance in the annoyance response. To improve the models obtained more descriptors needs to be added. The most important seem to be loudness, sharpness and roughness from the results presented in this report. Furthermore, the directional effects observed indicate, that the placement of heat pumps could have a relevant effect on how annoying people perceive the unit.

6 FIGURES INDEX

7 TABLES INDEX

8 REFERENCES

[1] M. Nilsson, M. Andehn and P. Lesna, "Evaluating roadside noise barriers using an annoyance-reduction criterion," *J. Acoust. Soc. Am.,* vol. 124, pp. 3561-3567, 2008.

[2] C. Kasess, P. Majdak, T. Maly and H. Waubke, "The relation between psychoacoustical factors and annoyance under different noise reduction conditions for railway noise," *J. Acoust. Soc. Am.,* vol. 141, pp. 3151-3163, 2017.

[3] S. Fidell, "Noise-Induced Annoyance," in *Handbook of Noise and Vibration Control*, Hoboken, New Jersey, John Wiley and Sons, 2007, pp. 316-319.

[4] J. Vos, "Annoyance caused by the sounds of a magnetic levitation train," *The Journal of the Acoustical Society of America,* vol. 115, no. 4, pp. 1597-1608, 2004.

[5] C. H. Kasess, A. Noll, P. Majdak and H. Waubke, "Effect of train type on annoyance and acoustic features of the rolling noise," *The Journal of the Acoustical Society of America,* vol. 134, no. 2, pp. 1071-1081, 2013.

[6] D. Cabrera, S. Ferguson, F. Rizwi and E. Schubert, "PsySound3: software for acoustical and psychoacoustical analysis of sound recordings," in *Proceedings of the 13th International Conference on Auditory Display*, Montreal, Canada, 2007.

[7] B. Glasberg and B. Moore, "Derivation of Auditory Filter Shapes from Notched Noise Data," *Hearing Research,* vol. 47, pp. 103-137, 1990.

[8] P. Daniel and R. Weber, "Psychoacoustical roughness: implementation of an optimized model," *Acustica,* vol. 83, pp. 113-123, 1997.

[9] E. Terhardt, G. Stoll and M. Seewan, "Algorithm for extraction of pitch and pitch salience from complex tonal signals," *J. Acoust. Soc. Am.,* vol. 71, pp. 679-688, 1982.

[10] J. Chalupper and H. Fastl, "Dynamic Loudness Model (DLM) for Normal and Hearing-Impaired Listerners," *Acta Acustica United with Acustica,* vol. 88, pp. 378-386, 2002.

[11] R Core Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing, 2018.

[12] H. Singmann, B. Bolker, J. Westfall, F. Aust and M. S. Ben-Shachar, "afex: Analysis of Factorial Experiments," 2020.

[13] R. Lenth, "emmeans (Estimated Marginal Means (Least-Squares Means))," 2020.

[14] W. Venables and B. Ripley, Modern Applied Statistics with S Fourth Edition, New York: Springer, 2002.

[15] G. Schwarz, "Estimating the dimension of a model," *Annals of Statistics,* vol. 6, pp. 461-464, 1978.

[16] S. W. Greenhouse and S. Geisser, "On methods in the analysis of profile data," *Psychometrika,* vol. 24, pp. 95-112, 1959.

[17] O. Gustafsson, H. Hellgren, C. Haglund Stignor, M. Axell, K. Larsson and C. Teuillieres, "Flat tube heat exchangers - Direct and indirect noise levels in heat pump applications," *Applied Thremal Engineering,* vol. 66, no. 1-2, pp. 104-112, 2014.

[18] O. Gustafsson, C. Teuillieres, H. Hellgren, M. Axell and J. O. Dalenbäck, “Reversing air-source heat pumps - Noise at defrost initiation and a noise reducing strategy,” *International Journal of Refrigeration,* vol. 62, pp. 137-144, 2016.

[19] C. H. Kasess, C. Reichl, H. Waubke and P. Majdak, “Perception Rating of Acoustic Emissions of Heat Pumps,” in *Forum Acusticum*, Lyon, 2020.

[20] G. Di, K. Lu and X. Shi, “An optimization study on listening experiments to improve the comparability of annoyance ratings of noise samples from different experimental sample sets,” *International Journal of Environmental Research and Public Health,* vol. 15, no. 3, pp. 474-486, 2018.

[21] J. R. Miller and E. C. Carterette, “Perceptual space for musical structures,” *J. Acoust. Soc. Am.,* vol. 58, pp. 711-720, 1975.

[22] K. Person Waye and R. Rylander, “The prevalance of annoyance and effects after long-term exposure to low-frequency noise,” *Journal of Sound and Vibration,* vol. 240, pp. 483-497, 2001.

[23] 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,” *ASHRAE Transactions,* vol. 116, pp. 553-568, 2010.

[24] N. Broner and H. G. Leventhall, “Low frequency noise annoyance assessment by low frequency noise rating (LFNR) curves,” *Journal of Low Frequency Noise, Vibration and Active Control,* vol. 2, pp. 20-28, 1983.

[25] K. Holmberg, U. Landström and A. Kjellberg, “Low frequency noise level variations and annoyance in working environments,” *Journal of Low Frequency Noise, Vibration and Active Control,* vol. 16, pp. 81-87, 1997.

[26] W. L. Martens and N. Zacharov, “Multidimensional perceptual unfolding of sptially processed speech I: Deriving space using INDSCAL,” in *Proceeedings AES 109th Convention*, Los Angeles, CA, 2000.

[27] K. Sturrock and J. Rocha, “A multidimensional scaling stress evaluation table,” *Field methods,* vol. 12, no. 1, pp. 49-60, 2000.

www.heatpumpingtechnologies.org

Heat Pump Centre
c/o RISE - Research Institutes of Sweden
PO Box 857
SE-501 15 BORÅS
Sweden
Tel: +46 10 516 5512
E-mail: hpc@heatpumpcentre.org

www.heatpumpingtechnologies.org

Report no. HPT-AN51-12