forum acusticum 2026
8-12 September · Graz, Austria

Exploring the Roles of Loudness and Sound Character in Heat Pump Noise Annoyance

J. Becker[1], V. Acun[2], C. Kasess[3], L. Stürenburg[4], S. Graetzer[2], N. Amiryarahmadi[5], J. Fels[4], A. J. Torija[2]

[1] *HEAD acoustics GmbH, Ebertstr. 30a, 52134 Herzogenrath, Germany*
[2] *Salford Acoustics Innovation Institute, University of Salford, UK*
[3] *Acoustic Research Institute, Wien, Austria*
[4] *Institute for Hearing Technology and Acoustics, RWTH Aachen University, Germany*
[5] *RISE Research Institutes of Sweden, Borås, Sweden*

Abstract

This study presents a detailed analysis of how sound character influences the annoyance ratings of air source heat pump (ASHP) noise, based on data from a listening experiment conducted as part of the IEA/HPT Annex 63 on Placement Impact on Heat Pump Acoustics. This large-scale listening experiment was originally designed to investigate the impact of ASHP placement in residential environments, considering factors such as the number of units, their spatial arrangement, operating conditions, background noise, and partition types.

Using this dataset, the present work extends the analysis to investigate key perceptual sound characteristics in ASHP noise. In particular, it examines the hypothesis that loudness provides a more suitable description of annoyance than conventional sound level metrics, including A-weighted sound pressure level. In addition, the study explores whether further psychoacoustic attributes, such as tonality, roughness, and sharpness, contribute to annoyance beyond what can be explained by level or loudness alone.

To this end, a selected sample of ASHP test sounds are systematically compared under controlled conditions of similar level or loudness, allowing the influence of specific sound characteristics to be investigated. The dataset, while originally developed for a different research focus, provides a suitable basis for an initial assessment of these perceptual factors.

The results support both hypotheses: perceived loudness appears to be more closely related to annoyance than level-based measures, and additional sound characteristics further influence listener responses. These findings highlight the importance of considering both the intensity and the character of sound when investigating heat pump noise annoyance, while also pointing to the need for dedicated studies to confirm these effects.

Keywords: *Heat pump noise, Psychoacoustics, Noise Annoyance, Sound Character*

1. Introduction

Air source heat pumps (ASHPs) are considered a key technology for decarbonizing residential heating, and their deployment has accelerated in recent years [1]. As outdoor units become more common in residential areas, acoustic acceptability is gaining importance. However, environmental noise is still assessed mainly using level-based descriptors, such as A-weighted sound pressure levels and related long-term indicators [2]. While essential for regulation and compliance, these measures reduce complex acoustic exposure to a single energetic quantity.
rom a psychoacoustic perspective, this reduction is limited, as sound perception is multidimensional. The evaluation of technical sounds depends not only on overall intensity but also on loudness, tonality, roughness, fluctuation strength, and sharpness [3], [4]. This is particularly relevant for tonal sounds, spectral imbalances, or temporal modulations. Thus, sounds with similar A-weighted levels may be perceived differently and evoke different annoyance responses [5].

Corresponding author: julian.becker@head-acoustics.com.

For ASHP noise, however, the available evidence remains limited. While recent studies have begun to investigate the role of psychoacoustic parameters in the evaluation of heat pump sounds [6], the relative roles of A-weighted level, perceived loudness, and additional sound-character descriptors remain unclear [7]. In particular, it is unknown whether annoyance is more closely related to loudness than to sound level, and whether further psychoacoustic characteristics help to explain annoyance differences once overall loudness has been approximately controlled.

This study addresses this gap by analysing selected listening-test data from IEA HPT Annex 63 [8]. First, it tests whether recordings with higher loudness are judged more annoying at comparable A-weighted levels. Second, it examines whether additional sound-character metrics explain annoyance beyond level or loudness alone. In this way, the study aims to clarify to what extent ASHP noise annoyance is governed by overall intensity and to what extent it depends on the specific character of the emitted sound, as reflected in psychoacoustic metrics.

2. Background

The analyses are based on a multi-laboratory listening experiment conducted within IEA HPT Annex 63 on ASHP noise annoyance under realistic residential conditions [8]. Ten organisations across seven European countries tested 273 participants, of whom 239 were included in the analysis after screening for hearing status and language proficiency. Binaurally auralised residential sound scenes were used to represent everyday outdoor environments.

The simulated setting represented opposing terraced-house gardens, with ASHP sounds rendered at different positions and in various configurations. The stimuli were based on recordings of a Mitsubishi Electric PUZ-WM112YAA outdoor unit [9]. Playback was standardised across all participating institutions using a HEAD acoustics binaural reproduction chain with Sennheiser HD 650 headphones. For the analysis, all stimuli were equalised using ID (independent of direction) equalisation, appropriate for the simulated sound field conditions. The experiment included a suburban background recording from Zoetermeer, the Netherlands, adjusted to 39 or 45 dB(A) and added to the auralised ASHP sounds. Thus, one background sound was presented at two levels rather than using two sound types. Background-only controls were also included, one of which serves as the reference condition for the pairwise analyses in Section 4.

The listening experiment comprised three parts: Part 1 examined a single ASHP at different positions, background-noise levels, and partition types; Part 2 investigated two simultaneously operating units in various arrangements; and Part 3 considered complex residential layouts with three to six units. Annoyance was rated on a standardised 0–10 scale according to ISO/TS 15666 [10].

This dataset forms the basis for the present analyses, which examine how annoyance relates not only to sound level but also to loudness and additional psychoacoustic parameters.

The primary analysis of the listening experiment indicated that annoyance was strongly correlated with several acoustic metrics, including A-weighted level, loudness, tonality, and roughness. At the same time, these metrics were themselves highly intercorrelated across the stimulus set, which limited a clear interpretation of their individual roles. For this reason, the present paper focuses on a secondary analysis based on targeted pairwise comparisons, with the aim of separating the effects of loudness and sound character more clearly.

3. Comparison of Level vs. Loudness

3.1 Data Selection and Statistical Analysis

To examine whether annoyance is more closely related to loudness than to A-weighted level, pairs of recordings were selected that were closely matched in A-weighted level but differed clearly in loudness. A-weighted level refers to the equivalent continuous A-weighted sound pressure level $L_{\mathrm{Aeq},T}$, calculated over the duration of each stimulus. The A-weighted levels of the left and right channels were combined into a single value by energy-based averaging. For psychoacoustic metrics, the two channels were combined using quadratic averaging. as recommended in ECMA-418-2 for loudness, roughness, and fluctuation strength and was applied consistently across all psychoacoustic metrics considered in this study. The analysis was performed twice, using two alternative loudness metrics: loudness according to ECMA-418-2 [4] and loudness according to ISO 532-1 [11]. Only pairs from the same experimental part were considered in order to avoid cross-part comparisons between different stimulus sets. Candidate pairs were retained if the A-weighted sound pressure level difference, ΔL_A, did not exceed $0.50\ \mathrm{dB}$ and the loudness difference, ΔN, was at least 0.50 in the respective metric, i.e. $0.50\ \mathrm{sone_{HMS}}$ for ECMA-418-2 and $0.50\ \mathrm{sone}$ for ISO 532-1. The higher-loudness recording was defined as the target and the lower-loudness recording as the reference. The selection thresholds were chosen as pragmatic but psychoacoustically informed criteria. The limit of $\Delta L_A \leq 0.5\ \mathrm{dB}$ kept residual level differences below typical discrimination thresholds of about 1 dB, depending on frequency and sensation level [12], and was therefore treated as negligible despite differences between controlled discrimination tasks and complex environmental sounds. By contrast, the requirement $\Delta N \geq 0.5$ in the respective loudness metric was not considered a universal just-noticeable difference, but an operational threshold ensuring a meaningful contrast between target and reference..

For each selected pair, the annoyance ratings of target and reference were compared within participants. As a descriptive first step, pair-specific annoyance differences were calculated and inspected individually. For each selected pair, the annoyance difference was first defined as

$$\Delta A_{i,j} = A_{i,j}^{\text{target}} - A_{i,j}^{\text{reference}}, \quad (1)$$

where $A_{i,j}^{\text{target}}$ and $A_{i,j}^{\text{reference}}$ denote the annoyance ratings given by participant i for pair j. Positive values therefore indicate that the acoustical scene containing the higher loudness was judged as more annoying than the scene with similar A-weighted sound pressure level but lower loudness.

To test the overall hypothesis that, at comparable A-weighted levels, the louder signal is perceived as more annoying, participant-level mean annoyance differences were then derived across all available pairs:

$$\overline{\Delta A}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \Delta A_{i,j}\,, \quad (2)$$

where n_i is the number of selected pairs rated by participant i. This averaging ensures that the analysis reflects participant-level rather than pair-level differences.

The distribution of $\overline{\Delta A}_i$ is shown in the histograms in Figure 1 and Figure 2. For testing statistical significance these results were tested against zero using a one-sided Wilcoxon signed-rank test. This test was chosen because the annoyance ratings were collected on an ordinal response scale and because it does not require normally distributed difference scores.

3.2 Results

The results of both analyses (using ECMA-418-2 and ISO 532-1) consistently showed that, at closely comparable A-weighted sound levels, the recordings with higher loudness were rated as more annoying. The corresponding distributions of participant-level mean annoyance differences, $\overline{\Delta A}_i$, are shown in Figure 1 for ECMA-418-2 and in Figure 2 for ISO 532-1.

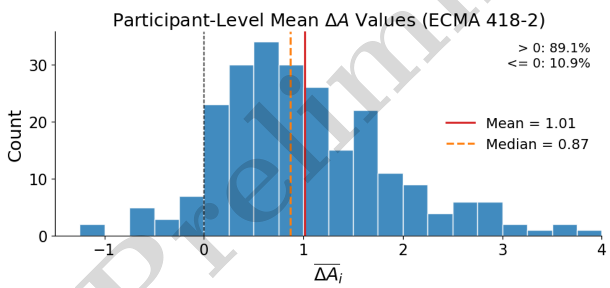

Figure 1. Histogram of participant-level mean annoyance differences of ECMA-418-2 loudness-based pairs.

For ECMA-418-2 loudness, 15 recording pairs met the selection criteria. The mean A-weighted level difference was $\Delta L_A = 0.03$ dB, while the mean loudness difference was $\Delta N = 0.854\ \text{sone}_{\text{HMS}}$. The participant-level mean annoyance difference $\overline{\Delta A}$ was 1.01 (median 0.87) corresponding to approximately one point on the 0 - 10 annoyance scale. Moreover, 89.1% of participants showed a positive mean annoyance difference, indicating that the louder target signal was perceived as more annoying. This effect was highly significant in the one-sided Wilcoxon signed-rank test ($p < 0.001$).

The same overall pattern was obtained when loudness was evaluated according to ISO 532-1 (see Figure 2), where 17 recording pairs satisfied the selection criteria. The mean A-weighted level difference across these pairs was $\Delta L_A = -0.03$ dB, and the mean loudness difference was $\Delta N = 1.91$ sone. The participant-level mean annoyance difference was 0.94, with a median of 0.82. Here, 91.6% of participants showed a positive effect. Again, the result was highly significant ($p < 0.001$).

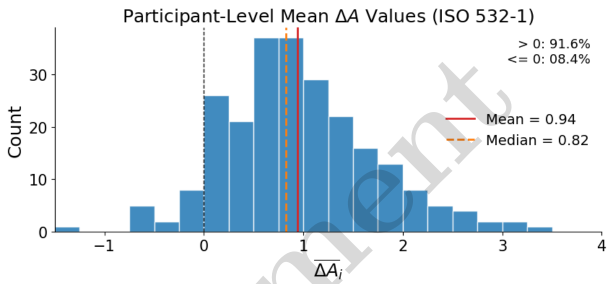

Figure 2. Histogram of participant-level mean annoyance differences of ISO 532-1 loudness-based pairs.

4. Sound Character at Comparable Loudness

4.1 Data Selection and Pair Definition

To examine the influence of sound character on annoyance at comparable loudness, a targeted pairwise analysis was conducted. The primary aim was to introduce meaningful differences in sound character while controlling for overall perceived loudness.

Direct comparisons between ASHP signals provided little variation in sound character when loudness was matched. Therefore, ASHP-containing signals were compared with a background-only reference, producing clearer and more systematic psychoacoustic differences. The reference was the background noise with an A-weighted sound pressure level of $L_A = 45$ dB, representing an acoustic scene without a ASHP.

As before, only signals from the same experimental part were compared to ensure a consistent context. While all signal pairs were formed within the same experimental part, the selected pairs originate from all three parts of the experiment.

Loudness matching was based on ECMA-418-2 since this standard also contains methods for the evaluation of sound character (tonality, roughness, fluctuation strength), and only pairs with an absolute loudness difference of $|\Delta N| \leq 0.20\ \text{sone}_{\text{HMS}}$ were retained. ΔN was defined as the difference of the loudness of the target signal and the reference signal. The threshold was chosen as a pragmatic criterion to ensure that residual loudness differences remained small relative to the overall loudness range of the stimuli (all above approximately $2\ \text{sone}_{\text{HMS}}$). Based on the approximate relationship between loudness and annoyance observed in the dataset (about 1.2 rating-scale points per

sone$_{HMS}$), a residual loudness difference of 0.2 sone$_{HMS}$ would correspond to an expected annoyance difference of at most about 0.24 scale points. This estimate serves only to illustrate that the potential contribution of residual loudness differences under the chosen threshold is small. In practice, the mean loudness difference across the selected pairs was substantially smaller ($\Delta N = 0.046$ sone$_{HMS}$), indicating that the signals were closely matched in terms of perceived intensity.
These criteria yielded 14 eligible reference pairs across Parts 1-3. Table 1 lists all selected pairs with their residual loudness differences and pair-mean annoyance differences. All target sounds included ASHP noise combined with the background sound at a level of 39 dB. Most of the selected configurations correspond to minimum-capacity operation of the ASHP, and thus represent relatively low-level ASHP noise conditions within the overall dataset. This selection procedure ensured that the subsequent analysis focused on differences in annoyance and sound character under conditions of closely comparable loudness.

Table 1. Selected reference pairs, residual loudness differences, and pair-mean annoyance differences.

Part	Target configuration	ΔN	$\overline{\Delta A_j}$
1	Single unit, min capacity, wire fence (Aura28)	-0.19	0.98
1	Single unit, min capacity, wooden fence (Aura28)	-0.16	1.16
1	Single unit, min capacity, wire fence (Aura27)	-0.12	1.28
1	Single unit, min capacity, wooden fence (Aura27)	-0.03	1.59
2	Two units, min/min operation, max distance, wooden fence	0.04	1.62
2	Two units, min/min operation, max distance, brick wall	0.08	1.72
2	Two units, min/min operation, medium distance, wooden fence	0.14	1.84
2	Two units, min/min operation, medium distance, brick wall	0.17	1.91
3	Scene 2, three units, medium capacity, wooden fence	-0.06	1.74
3	Scene 4, four units, min capacity, wooden fence	0.12	2.15
3	Scene 6, six units, min capacity, wooden fence	0.14	2.10
3	Scene 1, three units, min capacity, wooden fence	0.16	2.08
3	Scene 3, four units, min capacity, wooden fence	0.17	2.08
3	Scene 5, four units, min capacity, wooden fence	0.18	2.23

4.2 Annoyance Differences at Comparable Loudness

To assess how differences in sound character influence annoyance under conditions of comparable loudness, annoyance differences $\Delta A_{i,j}$ were again calculated as described in Equation (1) such that positive values indicate higher annoyance for the heat-pump-related signal than for the reference signal without heat pump noise. At the pair level, all 14 pairs showed positive mean annoyance differences $\overline{\Delta A_j}$ (see Table 1). A mean annoyance difference $\overline{\Delta A_i}$ was calculated for each participant as described in Equation (2) in order to summarize the effect at the participant level.
Figure 3 shows a clear shift toward positive participant-level mean annoyance differences, indicating greater annoyance for ASHP-related scenes than for loudness-matched background-only references. The pooled participant-pair differences $\Delta A_{i,j}$ showed a similar tendency: across all individual paired comparisons, the ASHP-related signal was rated as more annoying in 78.4% of cases, equally annoying in 15.5%, and less annoying in 6.1% of cases. This distribution already indicates a clear tendency toward higher annoyance for the signals containing ASHP noise, even though the selected comparisons were restricted to small loudness differences ($\leq$ 0.20 sone$_{HMS}$).

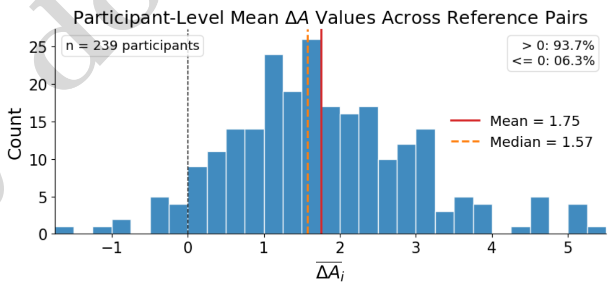

Figure 3. Histogram of participant-level mean annoyance differences across all selected reference pairs.

The mean participant-level annoyance difference was 1.75 rating-scale points (median 1.79; 95% confidence interval (CI): 1.52-1.97). To test the hypothesis that ASHP-related signals are judged as more annoying than the reference signal even at comparable loudness, the participant-mean annoyance differences were tested against zero using a Wilcoxon signed-rank test. The observed annoyance difference was statistically significant ($p < 0.001$). Thus, even at closely matched loudness, scenes containing ASHP noise were consistently rated as more annoying than the background-only reference.
Notably, Table 1 also shows that positive annoyance differences remained even in pairs with negative residual loudness differences, i.e., cases in which the ASHP-related signal was slightly less loud than the reference. In this subset of 5 pairs, the mean annoyance differences still ranged from 0.98 to 1.74 rating-scale points. This finding further supports the interpretation that the observed annoyance differences cannot be attributed solely to residual loudness differences.

4.3 Sound-Character Differences Across the Selected Pairs

To examine whether the selected pairs also differed systematically in sound character, signed pairwise differences (ΔT, $\Delta R, \Delta F, \Delta S$) were calculated for Tonality, Roughness and Fluctuation Strength according to ECMA-418-2, and Sharpness according to DIN 45692, using the same 14 pairs. Positive values indicate higher metric values for the ASHP-related signal The results show that the ASHP-related signals differed systematically from the reference signal in Tonality, Roughness, and Sharpness.

Tonality differences were positive for all 14 pairs with a mean pairwise difference of $0.134\ \mathrm{tu_{HMS}}$ (median $0.140\ \mathrm{tu_{HMS}}$; 95% CI: 0.108- $0.160\ \mathrm{tu_{HMS}}$). Statistical significance of the pairwise differences was assessed using Wilcoxon signed-rank tests against zero, yielding a significant effect for Tonality ($p < 0.001$). Roughness differences were also predominantly positive, with 13 out of 14 pairs showing higher values for the ASHP-related signal. The mean difference was 0.109 asper (median 0.134 asper; 95% CI: 0.074-0.144 asper). This effect was likewise statistically significant ($p = 0.0015$). Sharpness differences were positive for all 14 pairs, with a mean difference of 0.336 acum (median 0.375 acum; 95% CI: 0.284-0.387 acum). Again, the effect was statistically significant ($p < 0.001$). By contrast, Fluctuation Strength did not show a similarly systematic difference between target and reference. Across the 14 selected pairs, the mean difference was 0.001 vacil, (median -0.002 vacil; 95% CI: -0.005-0.007 vacil). Only 5 of the 14 pairs showed positive ΔF values, and the Wilcoxon signed-rank test did not indicate a significant shift away from zero ($p = 0.135$). Thus, Fluctuation Strength provided little evidence of a sound-character contrast within the selected pairs.

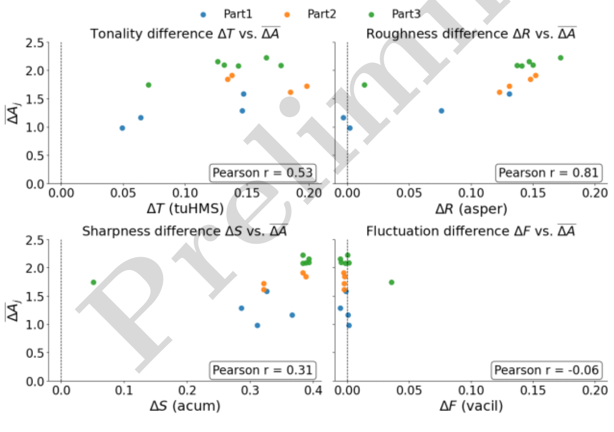

Figure 4. Pair-mean annoyance differences plotted against pairwise differences in tonality, roughness, and sharpness and fluctuation strength.

Figure 4 shows the pairwise relationships and Pearson correlations between annoyance difference and the four psychoacoustic metrics. Because annoyance is expected to depend on several perceptual attributes at the same time, very high correlations with any single psychoacoustic metric are not necessarily expected. The correlations should therefore be interpreted as descriptive indicators of association, not as a measure of the full explanatory power of each metric.

For tonality, the positive correlation ($r = 0.53$) indicates a meaningful tendency for larger tonality differences to be associated with larger annoyance differences. In the present context, this can be regarded as a relevant association, as annoyance is not expected to be determined by tonality alone. For roughness, the correlation was even stronger ($r = 0.81$), which is remarkably high for a single metric and suggests that roughness aligned particularly closely with the pairwise annoyance differences in the selected comparisons. For sharpness, the correlation was weaker but still positive ($r = 0.31$), indicating that sharpness may also contribute to annoyance, although probably not as the dominant factor in this subset of pairs. At the same time, the interpretability of sharpness is limited by the comparatively small spread of ΔS, which reduces its ability to differentiate strongly between pairs. By contrast, fluctuation strength showed virtually no association with annoyance difference ($r = -0.06$), consistent with the observation that ΔF remained close to zero across the selected pairs.

To visualize the joint pattern of the two sound-character dimensions with the highest correlation to perceived annoyance, Figure 5 shows ΔT against ΔR, with the pair-mean annoyance difference $\overline{\Delta A}$ encoded by color. The highest annoyance differences tend to occur when both metrics are high. Non-negligible annoyance differences also occur at low values, indicating that these two descriptors alone do not fully explain annoyance differences. A further contributing factor may be the consistently positive sharpness difference between target and reference (see Figure 4).

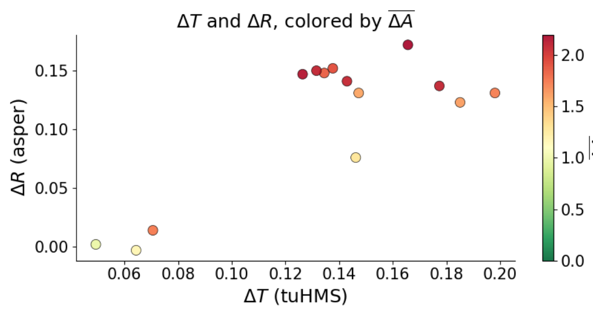

Figure 5. Pairwise differences in tonality and roughness, with pair-mean annoyance difference encoded by color.

While multiple linear regression could be applied, the dataset was not designed to isolate individual psychoacoustic dimensions, and the metrics covary. Regression coefficients would therefore be specific to the present dataset and not generally applicable. The present analysis therefore focuses on controlled pairwise comparisons and descriptive relationships.

Overall, the selected pairs differed in annoyance at comparable loudness, but also showed systematic and statistically significant differences in sound character.

5. Discussion and Conclusions

The present analyses indicate that annoyance caused by residential heat pump noise cannot be described adequately by A-weighted level alone (Section 3). At closely matched A-weighted levels, signals with higher loudness were consistently rated as more annoying using both ECMA-418-2 and ISO 532-1 metrics.. Zhe effect was statistically significant and consistent across participants, indicating that loudness might provide a more perceptually relevant descriptor of annoyance than A-weighted level in the present dataset.

Section 4 shows that loudness alone does not fully explain annoyance. At comparable loudness, heat-pump-related signals were consistently rated as more annoying than background-only conditions. The results therefore suggest that participants preferred a background-only sound over a sound of comparable overall loudness containing heat pump noise. Notably, most selected configurations correspond to minimum-capacity operation, indicating that these effects already occur under relatively low-level conditions.

The sound-character analysis supports this interpretation. Heat-pump-related signals showed greater tonality, roughness, and sharpness. Tonality and roughness were associated with larger annoyance differences, whereas sharpness was consistently elevated and may contribute to an overall increase in annoyance without explaining variation between pairs.

Overall, significant annoyance differences between heat-pump signals with similar A-weighted levels show that level alone is insufficient to describe annoyance. Within this dataset, corresponding loudness differences suggest that loudness captures aspects of annoyance not reflected by A-weighted level alone. At the same time, annoyance differences remained after approximately controlling for loudness, suggesting that additional sound-character information is required. These results agree with the findings reported in [6].

At the same time, the results should be interpreted with some caution. This is an analysis of subsets of a larger dataset that was originally designed to study placement-related effects rather than to isolate individual psychoacoustic dimensions experimentally. The sounds are from a single heat pump and mainly differ in operating state, number of units, and spatial position as described in Section 2. As a result, the psychoacoustic metrics covary across the available stimuli, and the present analyses do not permit a clear quantitative or causal attribution of annoyance to individual sound-character descriptors.

Nevertheless, the results consistently suggest that perceptual assessment of heat pump noise should extend beyond level-based metrics and should include both loudness and additional sound-character information.

6. Acknowledgments

Antonio J. Torija, Volkan Acun, and Simone Graetzer would like to acknowledge the funding provided by the Innovation Accelerator programme of Innovate UK for the Future Homes Project (Project number 10054845); and by the UK Department of Energy Security and Net Zero for the Participation in the IEA HPC Annex 63 on “Placement Impact on Heat Pump Acoustics”.

Christian H. Kasess would like to acknowledge the funding provided by the Austrian Research Promotion Agency (FFG) (Project number 898198).

Nata Amiryarahmadi would like to acknowledge the funding provided by the Swedish Energy Agency (Energimyndigheten) (Project number P2023-01510).

References

[1] J. Rosenow, D. Gibb, T. Nowak *et al.*, “Heating up the global heat pump market,” *Nat Energy*, vol. 7, pp. 901–904, 2022.

[2] European Parliament and Council, “Directive 2002/49/EC relating to the assessment and management of environmental noise,” 2002.

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

[4] Ecma International: *ECMA-418-2: Psychoacoustic metrics for ITT equipment (methods for describing human perception based on the Sottek Hearing Model)*, 2025.

[5] L. Stürenburg, L. Aspöck, S. J. Schlittmeier, and J. Fels, “Pleasantness judgements of air diffuser noise and their relation to psychoacoustic parameters,” *Acta Acustica*, vol. 10, p. 44, 2026.

[6] L. Stürenburg, H. Braren, L. Aspöck, and J. Fels, “Heat pump noise: Determination and modelling of preference-equivalent levels,” *Acta Acustica*, vol. 10, p. 20, 2026.

[7] V. Acun, S. Graetzer, M. Radivan and A.J. Torija, “Human response to air source heat pump noise: influence of background noise, operating conditions and acoustic characteristics,” *Acta Acustica*, vol. 10, p. 36, 2026.

[8] L. Stürenburg *et al.*, “Study on human response to the cumulative impact of air source heat pump noise,” in *Proc. Heat Pump Conference*, 2026.

[9] L. Stürenburg, H. Braren, L. Aspöck and J. Fels, Recordings of an Air-to-Water Heat Pump [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.13365535, 2024

[10] International Organization for Standardization: ISO/TS 15666:2021, *Acoustics — Assessment of noise annoyance by means of social and socio-acoustic surveys*, 2021.

[11] International Organization for Standardization: *ISO 532-1:2017, Acoustics — Methods for calculating loudness Part 1: Zwicker method*, 2017.

[12] W. Jesteadt, C. C. Wier, and D. M. Green, “Intensity discrimination as a function of frequency and sensation level,” *J. Acoust. Soc. Am.*, vol. 61, no. 1, pp. 169–177, 1977.