https://doi.org/10.5194/wes-2024-125
Preprint. Discussion started: 5 November 2024

eawe european academy of wind energy
Open Access
WIND ENERGY SCIENCE
DISCUSSIONS

A listening experiment exploring the relationship between noise annoyance and sound quality metrics for airborne energy systems

Helena Schmidt[1], Renatto. M. Yupa-Villanueva[2], Daniele Ragni[1], Roberto Merino-Martínez[2], Piet van Gool[3], Roland Schmehl[1]

[1]Department of Flow Physics and Technology, Delft University of Technology, Delft, 2629 HS, The Netherlands
[2]Department of Aircraft Noise and Climate Effects, Delft University of Technology, Delft, 2629 HS, The Netherlands
[3]Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Eindhoven, 5612 AE, The Netherlands

Correspondence to: Helena Schmidt (h.s.schmidt@tudelft.nl)

Abstract. The present study investigates the relationship between sound quality metrics (SQMs) and noise annoyance for airborne wind energy systems (AWESs) reported in a listening experiment. A convenience sample of 75 participants rated their annoyance on the International Commission on Biological Effects of Noise (ICBEN) scale in response to nine recordings from in-field measurements of two different fixed-wing and one soft-wing ground-generation AWES. All recordings were normalized to have the same A-weighted equivalent sound pressure level. The acoustical analyses showed that the fixed-wing kites presented a more tonal and narrowband sound signature than the soft-wing kite. Linear-mixed effects models indicated that sharpness was the only SQM predicting participants' annoyance ratings and that the fixed-wing kites were rated as more annoying than the soft-wing kite. In addition, the effect of some SQMs on annoyance depended on participant characteristics, with loudness having a weaker impact on annoyance for participants familiar with AWESs and tonality having a weaker effect on annoyance for older participants. However, these moderation effects could be random due to the non-probability sampling used.

1 Introduction

Wind energy is one of the most widely available renewable energy sources, and its capacity must increase by 320 GW by 2030 to meet the climate goals of the Paris Agreement (IEA, 2023; UNFCCC, 2016). One promising yet unexploited novel renewable energy technology is airborne wind energy (AWE) (BVG Associates, 2022; Vos et al., 2024). AWE uses tethered flying devices, called kites, to harness higher-altitude winds. AWE can complement conventional wind energy by accessing wind resources above 200 m altitude and providing power supply at inaccessible or only temporarily used sites, such as during natural disasters. Research indicates that substantial mass savings translate to a lower environmental impact (Hagen et al., 2023). While AWE is regarded as a potential game-changer for the energy transition (IRENA, 2021), the technology has not yet converged towards a single standard configuration. The existing prototypes can be divided into two main categories: ground-generation and fly-generation concepts, as shown in Fig. 1 (Cherubini et al., 2015). The former concept alternates between energy-generating reel-out and energy-consuming reel-in phases. During the reel-out phases, the kite is flown in a loop or figure of eight maneuvers, generating more energy than is used during the reel-in phases, resulting in a positive net power output (Vermillion et al., 2021). Ground-generation AWES commonly use soft-wing kites based on flexible membrane wings or fixed-wing kites typically made from carbon fiber-reinforced polymers. The latter concept employs small wind turbines onboard a fixed-wing kite.

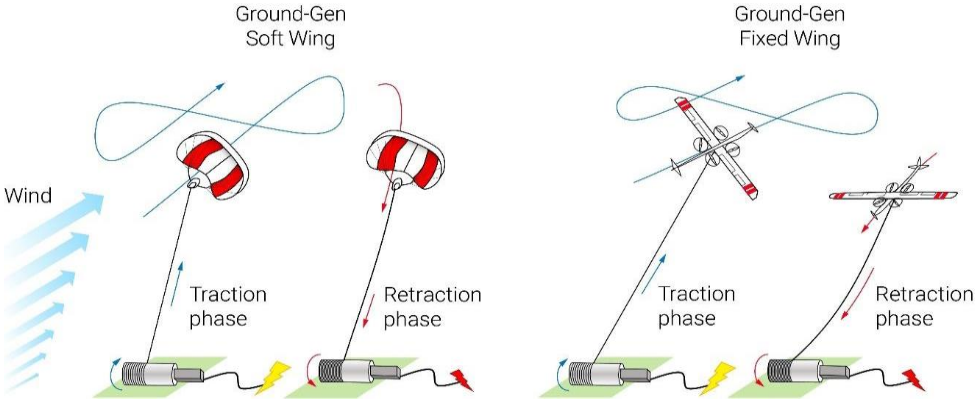


Figure 1: Schematic representation of ground-generation airborne wind energy systems employing a soft-wing and a fixed-wing kite, respectively (based on Fagiano et al., 2022)**.**

As with any other wind energy technology, AWE systems (AWESs) must comply with environmental regulations on sound emissions to limit the impact on surrounding residents (van Kamp and van den Berg, 2021). Extensive research on wind turbines shows that their sound emissions can be related to noise annoyance and stress symptoms (Bakker et al., 2012; Hübner et al., 2019; Pohl et al., 2018; Turunen et al., 2021). Noise annoyance refers to an individual's negative evaluation of wind turbine sound emissions, sometimes in combination with self-reported psychological or physical stress complaints, such as sleep disturbances, irritability, or lacking concentration (Pohl et al., 2018). Although AWESs are said to be a quiet technology due to their high operating altitude (for a review, see Schmidt et al., 2022), this assumption disregards several non-technical factors influencing the human perception of wind energy sounds and technological aspects of the systems and their operation. Non-technical factors include individual dispositions (e.g., noise sensitivity, especially to low-frequency sounds) (Haac et al., 2019; Michaud et al., 2016; Pedersen et al., 2010; Pedersen and Persson Waye, 2007; Schutte et al., 2007), perceptions (e.g., aesthetics of the technology and the fairness of the planning process) (Haac et al., 2019; Hübner et al., 2019; Pedersen and Larsman, 2008), and attitudes (e.g., towards a specific wind energy project or wind energy in general) (Hoen et al., 2019; Hübner et al., 2019; Ki et al., 2022; Pawlaczyk-Łuszczyńska et al., 2018; Pedersen and Persson Waye, 2007; Schäffer et al., 2019). Technological aspects refer to tethers, rotating components onboard, and the relatively high speeds at which the kites operate while flying, which enhances tonal components and modulation of the sound emitted, severely affecting subjective human perception (Hansen et al., 2021; Lee et al., 2011; Schäffer et al., 2018; Torija et al., 2019; Yokoyama and Tachibana, 2016; Yonemura et al., 2021). A preliminary study from Bouman (2023) on an existing fixed-wing kite found a narrowband spectral distribution of the emitted noise, enhanced by laminar flow regimes on the suction side of the wing with a relatively short chord. In the same study, a larger soft-wing kite showed a broadband distribution mostly determined by turbulent boundary-layer trailing-edge noise. To the best of the authors' knowledge, how these noise sources relate to potential noise annoyance has not been studied yet. In fact, due to the limited deployment of AWESs to date, only one field study has investigated to what extent residents near an AWE test site are annoyed by its sound emissions (Schmidt et al., 2024). The study found that 35.2% of respondents living on average 2 km from the AWES perceived sounds at home, with 13.1% being annoyed (score of at least 2 on a scale from 0 to 4) and 7.5% highly annoyed (score of at least 3 on the same scale). However, the study did not investigate the relationship between the AWES sound emissions and the reported annoyance. Hence, it must be better understood how the sound produced by AWESs relates to residents' annoyance to mitigate sound emissions effectively in the manufacturing and operation of AWESs.

The main objective of the present study is to investigate to what extent perception-based sound quality metrics (SQMs) can predict the noise annoyance caused by AWESs. This study combined five SQMs into global Psychoacoustic Annoyance (PA) metrics. The benefit of the PA metric is that it provides a quick estimate of the noise annoyance perceived for a given sound without having to measure respondents' annoyance levels. Additionally, the performance of PA metrics for estimating annoyance levels will be compared to conventional metrics (i.e., the effective perceived noise level EPNL). This study focuses on sound emissions of soft-wing and fixed-wing kites recorded during in-field measurements. A laboratory listening experiment was conducted with 75 participants who rated their annoyance in response to the different AWES recordings.

Section 2 describes the applied study design, procedure, and materials. Section 3 presents the results from the acoustical analyses of the sound recordings and the statistical analyses of the reported annoyance. Finally, Section 4 offers the most relevant conclusions from the study.

2 Methodology

In the following, the methodologies employed to record the sound samples and the laboratory listening experiment are explained in detail, including characteristics of the sound samples and participants, annoyance ratings, and laboratory procedures.

2.1 Sound recordings

Nine sound recordings from three different AWESs (i.e., three recordings from each prototype) were used for the listening experiment. All recordings were taken during the energy-harvesting reel-out phase. A total signal length of 25 s per recording was extracted from longer, more complex audio footage that included additional non-relevant preparation phases for the three AWESs. All three AWESs implement ground-based electricity generation (see Section 1). One is a soft-wing kite (AWES A), and the other two are fixed-wing kites (AWESs B and C). Table 1 provides more information about the AWESs and the sound measurement campaigns. Given that there are currently no specific sound regulations for AWESs, the sound pressure levels of the recordings were normalized to an equivalent A-weighted sound pressure level value of 45 dBA to align with European regulations for wind turbines, which commonly range between 35 dBA and 55 dBA during the day (Solman and Mattijs, 2021). Normalizing the sound pressure levels helped to evaluate aspects of sound quality other than loudness (Boucher et al., 2024).

Table 1 Overview of the investigated airborne wind energy systems (AWESs) and the corresponding sound measurement campaigns.

	AWES A	AWES B	AWES C
Kite type	Soft-wing	Fixed-wing	Fixed-wing
VTOL propellers	None	Present, inactive during the measurements	Present, inactive during the measurements
Ram-air turbine	Present, tied down during the measurements to prevent free-spinning	None	Present, active during the measurements
Flight pattern	Figure of eight	Circle	Circle
Wind speed (m/s)	5 – 10	9	8 – 9[a]
Max relative flying airspeed (m/s)	38	42	43
Max kite altitude during experiment (m)	253	231	150
Distance to microphone (m)	428 – 620	305 – 689	Approximately 100 – 700
Test location and type	Field; standard flight test	Inoperative airfield; standard flight test	Inoperative airfield; tow test (i.e., the ground

			station was on the back of a truck driving straight to create an artificial wind field, while the kite was flying crosswind loops of about 60 – 70 m diameter)
Recording instrumentation	Brüel & Kjær 4189 microphone at 1 m height and 650 m downwind from the winch of the ground station; Brüel & Kjær UA-650 windscreen over the microphone to reduce wind sounds; Brüel & Kjær sound level meter 2250	Brüel & Kjær 4189 microphone at 1 m height and 679 m downwind from the winch of the ground station; Brüel & Kjær UA-1650 windscreen over the microphone to reduce wind sounds; Brüel & Kjær sound level meter 2250	Three Brüel & Kjær 4189 microphones were positioned at equal distances along the driving route; The vehicle sounds were mainly emitted at the ground level and absorbed by padded microphone covers

Note. VTOL: vertical take-off and landing. [a]The values refer to the ambient wind speed, but the towing speed was higher.

2.2 Listening experiment

2.2.1 Psychoacoustic Listening Laboratory

The listening experiment was conducted in the Psychoacoustic Listening Laboratory (PALILA) at the Faculty of Aerospace Engineering of Delft University of Technology. PALILA is a soundproof booth inside a separate room specifically designed to research the human perception of aeroacoustics sound sources, including aircraft, drones, and wind turbines. The booth is 2.32 m long, 2.32 m wide, and 2.04 m tall inside. The background noise level inside the room is 13.4 dBA. Merino-Martínez et al. (2023) describe the design and acoustic characterization in detail. PALILA's audio reproduction system is a Dell Latitude 7420 laptop (with an Intel® CoreTM i5-1145G7 vPro® processor and 16 GB of RAM) connected through a universal audio jack connector to a set of Sony WH-1000XM4 over-ear, closed-back headphones. The headphones allow for binaural hearing and have a 40 mm diameter dome-type driver unit, a frequency response between 4 Hz and 40 kHz, and a sensitivity of 105 dB/mW at 1 kHz. The audio reproduction system had been calibrated with a G.R.A.S. 45BB-14 KEMAR head and torso simulator. Participants are seated in the booth's center, and the laptop is placed on a table in front of them, as shown in Fig. 2.

Figure 2: Exemplary photograph of the laboratory setup used for the listening experiment.

2.2.2 Participant recruitment and procedure

Participants were recruited using convenience and snowball sampling (Passer, 2014), mainly targeting students and employees. Participants were eligible to participate if they reported no hearing impairment and felt physically well on the day of the experiment. The study was conducted between June and September 2023. A trained experimenter instructed each participant individually, after which they completed the experiment independently. In the first part of the questionnaire, participants were asked to self-report their hearing ability, hearing-affecting incidents (e.g., ear diseases, accidents, loud work environments), and well-being to establish their eligibility for participation. The second part of the questionnaire, the listening experiment, started with a practice round to get familiar with the process and the scales. It was followed by two counterbalanced blocks separated by an automatic and mandatory one-minute break: one block on AWES sounds and another on wind turbine sounds (the latter are not reported here). The sequence of the sound recordings within each block was randomized to minimize order and learning effects on participants' annoyance ratings (Passer, 2014). Participants listened to and evaluated one recording at a time. The recordings could not be replayed. The final part of the questionnaire asked about participants' noise sensitivity, familiarity with AWE, and demographic information. At the end of the experiment, the experimenter debriefed the participant and handed over a 20-euro voucher as a participation reward.

2.2.3 Annoyance ratings and questionnaire

Annoyance was defined according to the ISO 15666 standard as "one person's individual adverse reaction to noise. For example, dissatisfaction, bother, annoyance and disturbance due to noise" (International Organization for Standardization, 2021). In line with the definition and recommended practice for psychoacoustic research (Alamir et al., 2019), annoyance levels were measured using the International Commission on Biological Effects of Noise (ICBEN) scale. For each sound recording, participants were asked to rate their experienced annoyance on both the verbal and numerical scale, and the average was calculated to increase measurement reliability (International Organization for Standardization, 2021). The 5-point verbal scale, ranging from "not at all" (0) to "extremely" (4), asked: "Imagine you are at home and hearing the noise at home; how much does the noise bother, disturb, or annoy you?" The 11-point numerical scale, ranging from 0 ("not at all") through 10 ("extremely"), asked: "Imagine you are at home and hearing the noise at home; what number from 0 to 10 best shows how much you are bothered, disturbed, or annoyed by the noise?". The wording of the scales was slightly adapted to acknowledge the laboratory setting. To establish whether participants were eligible to partake in the study, their hearing ability was self-reported using a 5-point scale (from "poor" to "excellent"). Since an audiometric test was unavailable, self-evaluations were used, which are reasonably valid in assessing an individual's hearing (Hong et al., 2011). The occurrence of hearing-affecting conditions and incidents was also self-reported (e.g., hearing aid usage, ear diseases, accidents, tinnitus, loud work environments), and participants' well-being was queried (e.g., common cold, fatigue). Noise sensitivity was assessed with the condensed version of the 4-point NoiSeQ scale, ranging from "strongly disagree" (0) to "strongly agree" (3) and consisting of 12 items covering noise sensitivity in different situations (Griefahn, 2008). This scale has been shown to have high internal consistency ($\alpha = 0.87$) (ibid.). Furthermore, whether participants were familiar with AWESs and had ever listened to one before was also assessed. Finally, information about the participants' age, gender, and education level was gathered. A graphical user interface (GUI) that guided participants through the entire questionnaire, including the listening experiment, was specifically developed for this experiment using MATLAB R2021b (see supplementary materials).

2.3 Participant characteristics

Of the 75 participants, 73.3% were male, 24% female, and 2.7% non-binary. The proportion of men was higher because participants were mainly recruited from a technical university. The age ranged from 18 to 66 years, with an average of 28 years and a standard deviation of 9.57 years. The sample was overall highly educated, with 74.7% holding a Bachelor's or Master's degree, 16% currently or previously enrolled in university, and 8% having a doctoral degree. On average, participants took 22 minutes to complete the experiment, excluding the examiner's briefings. The average reported hearing ability was very good [Mean (M) = 4.07, standard deviation (SD) = 0.64, scale: 1-5], and the mean noise sensitivity was medium (M = 1.56, SD = 0.38, scale: 0-3). About half of the participants reported being familiar with AWE (n = 37), but only 17.3% (n = 13) had listened to an AWES prior to the experiment. Because so few participants had heard an AWES before, the analyses did not consider this variable as a confounding factor.

2.4 Post-processing of the results

2.4.1 Acoustic analyses

The EPNL metric (Kephalopoulos et al., 2014; Pieren et al., 2019) was used to explore how well conventional acoustic metrics can describe annoyance to AWESs. Furthermore, the following five SQMs (Merino-Martínez et al., 2021) were calculated for each considered sound wave of every recording:

- Loudness (N): the perception of the sound magnitude corresponding to the overall sound intensity. Based on Zwicker's method, loudness was calculated using the ISO norm 532–1 (ISO/TC 43, 2017).
- Tonality (K): the perceived strength of unmasked tonal energy within a complex sound. Tonality was computed using Aures' method (Aures, 1985).
- Sharpness (S): the high-frequency sound content. The DIN 45692:2009's (Deutsches Institut für Normung, 2009) method was used here.
- Roughness (R): the hearing sensation caused by modulation frequencies between 15 Hz and 300 Hz. Roughness was calculated according to Daniel and Weber (1997).
- Fluctuation strength (FS): assessment of slow fluctuations in loudness with modulation frequencies up to 20 Hz, with maximum sensitivity for modulation frequencies around 4 Hz. The method by Osses Vecchi et al. (2016) was used.

The five SQMs were evaluated over time using a subset of the full sound recordings to assess the repeatability of the metrics in the 25 s full-time span. To evaluate the sound quality through single quantities, the 5th percentile values were used, which represent the level of each SQM exceeded during 5% of the total recording time (indicated henceforth by the subindex 5). From the SQMs, the PA metrics were calculated according to the models by Zwicker and Fastl (1999), More (2010), and Di et al. (2016). The general expression for the PA metric is

$$PA = N\left(1 + \sqrt{C_0 + C_1\omega_S^2 + C_2\omega_{FR}^2 + C_3\omega_T^2}\right), \tag{1}$$

where the term ω_S contains the sharpness S (and loudness N) contribution:

$$\omega_S = \begin{cases} 0.25(S - 1.75)\log_{10}(N + 10), & \text{for } S \geq 1.75, \\ 0, & \text{for } S < 1.75. \end{cases} \tag{2}$$

The term ω_{FR} contains the contributions of the roughness R and fluctuation strength FS (and loudness N),

$$\omega_{FR} = \frac{2.18}{N^{0.4}}(0.4FS + 0.6R), \tag{3}$$

and the term ω_T contains the tonality K (and loudness N) contribution,

$$\omega_T \begin{cases} 0, \\ (1 - e^{-0.29N})(1 - e^{-5.49K}) \\ \frac{6.41}{N^{0.52}}K, & \text{for the model by Zwicker and Fastl (1999)} \end{cases} \tag{4}$$

for the model by More (2010)

for the model by Di et al. (2016).

Lastly, the coefficients C_0 to C_3 of Eq. (1) for each PA model are listed in Table 2. The conventional sound metrics, SQMs, and PA metrics were computed using the open-source MATLAB toolbox SQAT (Sound Quality Analysis Toolbox) v1.1 (Greco et al., 2023).

Table 2 Coefficients for Eq.(1) for each considered psychoacoustic (PA) model.

PA model	C_0	C_1	C_2	C_3
Zwicker and Fastl (1999)	0	1	1	0
More (2010)	-0.16	11.48	0.84	1.25
Di et al. (2016)	0	1	1	1

2.4.2 Annoyance ratings and percentage of highly annoyed respondents

Following Brink and colleagues' approach (2016), verbal and numerical scale responses were linearly transformed to a 0-100 scale and averaged to obtain a total annoyance score per participant for each recording. The verbal and numerical scales were strongly correlated in the present data, justifying calculating average scores (Tau-b item correlations were between 0.75 and 0.88). The average scores were used to determine the percentage of highly annoyed (%HA) participants for each recording. Following Miedema and Vos (1998) and the ISO standard (International Organization for Standardization, 2021), the top 28% of the scale were considered highly annoyed. That is, participants whose score was 72 or higher on the 100-point scale were classified as highly annoyed.

2.4.3 Linear-mixed effects models

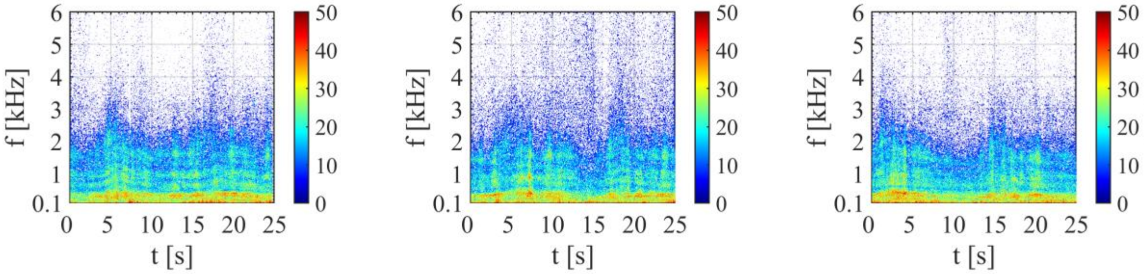
Linear-mixed effects models were applied to identify significant predictors and to examine whether significant differences existed in the annoyance ratings across the three AWESs. Linear-mixed effects models can separate fixed effects (in this case, the acoustic predictors) from random effects (the participants with their individual characteristics). This type of hierarchical analysis has been successfully employed in past research on wind turbine noise annoyance (Merino-Martínez et al., 2021; Schäffer et al., 2016, 2019).

In this study, the sound recordings were nested within AWES types as each participant rated every recording that belonged to one of the three AWES types. Additionally, participants served as another level of nesting, as each participant contributed multiple ratings across the different AWESs. Following Judd et al. (2017), the nested structure was addressed by employing linear mixed-effects models with random effects for participants and AWES types. The conditions were contrast-coded to aid interpretation and included as random effects (ibid.). This approach allowed modeling the variability in annoyance ratings attributable to individual participants and differences between AWES types. Following Aguinis and colleagues' step-wise approach (2013), participant characteristics were first included as fixed effects to determine their predictive value on annoyance ratings. Second, the SQMs were added as fixed effects, assessing each characteristic in separate models to avoid multicollinearity. Third, the impact of the SQMs was randomized to examine whether these effects varied between individuals. Fourth, interaction terms were included between participant characteristics and SQMs to explore whether the participant characteristics could explain individual differences in the impact of SQMs on annoyance ratings. Finally, using the -2-log likelihood ratio, the goodness-of-fit for the final linear mixed-effects models was assessed to quantify the variance explained by the fixed factors alone and by both fixed and random factors. Separate linear-mixed effects models, including EPNL or the PA models as predictors, evaluated how effectively these (psycho)acoustic metrics could predict the annoyance ratings. All statistical analyses were performed using the software R version 4.4.0 (R Core Team, 2023), and linear mixed-effects models were fitted using the 'lme4' package (Bates et al., 2024).

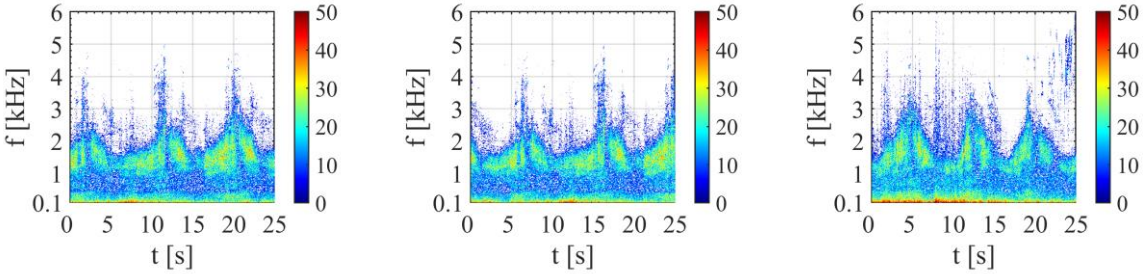
3. Results

3.1 Acoustic results

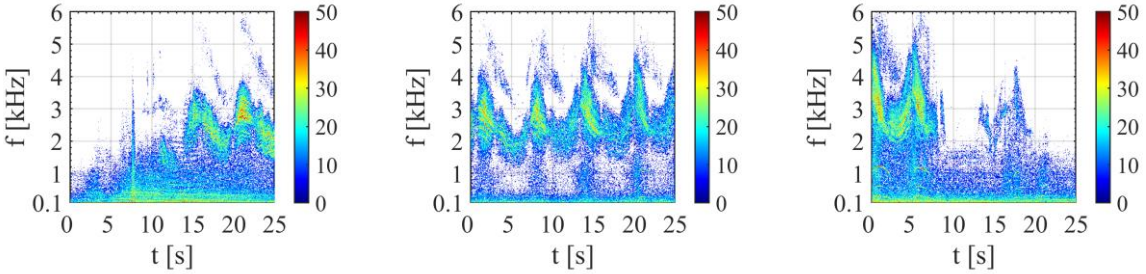
The time-frequency sound levels were represented as spectrograms (see Fig. 3). The spectrograms were calculated with a sampling frequency of 48000 Hz for every audio sample using 4800 samples per time block (i.e., 0.1s) with Hanning windowing and a 50% data overlap. These parameters provided a frequency resolution (Δf) of 10 Hz. For AWES A, the lower frequencies (0-1 kHz) exhibited higher sound levels, which decreased as the frequency increased. Examining the spectrograms demonstrated that the recordings were representative. For AWES B, the highest sound levels were found at extremely low frequencies, up to approximately 200 Hz. Sound levels decreased between 200 Hz and 1 kHz but increased again in the frequency range between 1 and 3 kHz. AWES B exhibited a periodic sound pattern over time, likely due to its circular flight trajectory. A periodic sound pattern was observed for recordings corresponding to AWES C, characterized by a significant absence of sound levels in the frequency range between 200 Hz and 1.2 kHz (C1 and C2) and between 200 Hz and 2 kHz (C3). These periodic behaviors are again attributed to the circular flight trajectory. For C1, the acoustic energy was predominantly concentrated between 15-25 s and in the frequency range between 1.2 and 4 kHz. C2 showed consistent sound levels, peaking between 1.2 and 5 kHz. Conversely, C3 displayed higher levels within the first 8 s at 2-5 kHz.



(**a**) Soft-wing kite (A1) (**b**) Soft-wing kite (A2) (**c**) Soft-wing kite (A3)



(**d**) Fixed-wing kite (B1) (**e**) Fixed-wing kite (B2) (**f**) Fixed -wing kite (B3)

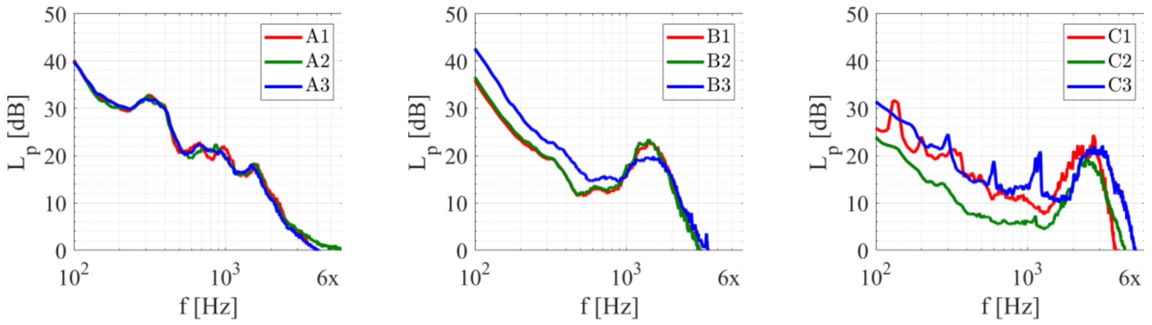


(**g**) Fixed-wing kite (C1) (**h**) Fixed-wing kite (C2) (**i**) Fixed-wing kite (C3)

Figure 3: Spectrograms corresponding to each recording.

Time-averaged sound pressure levels (SPL) were computed, as shown in Fig. 4, to compare the sound levels produced by each type of AWES. For AWES A, AWES B, and the second recording of AWES C (C2), SPLs were averaged over the whole 25 s recording duration. In contrast, for the first and third recordings of AWES C (i.e., C1 and C3), the averages considered the last 10 s and first 8 s, respectively, when the kite noise was perceivable. In the listening experiment, the full recordings were used. Only slight variations were observed when considering the entire recording. SPLs were virtually the same across the entire frequency range for AWES A, displaying a bump in the 200 Hz to 2 kHz range. AWES B showed similar trends and sound levels across the recordings, although there was a difference of approximately 4 dB between B1, B2, and B3 for frequencies up to 1 kHz. For frequencies higher than 1 kHz, the SPLs were nearly identical across all B-recordings, with a bump pattern observed between 1 kHz and 2 kHz.

Regarding AWES C, time-averaged SPLs showed more significant differences for frequencies below 2 kHz, with C3 having the highest sound levels, followed by C2 and C1. On the other hand, the frequencies above 2 kHz were similar among the recordings, though C3 exhibited sound levels approximately 6 dB higher than C1 and C2 in the 3 to 5 kHz range. Additionally, it was observed that AWES A and AWES B had higher sound levels than AWES C, particularly for frequencies below 100 Hz, see Fig. 5. The SPLs in C1 and C3 exhibited a tonal behavior in the frequency range of 60 to 2000 Hz, which is believed to be related to the ram-air turbine. The flight patterns for both AWESs B and C are circular, which could cause the bump sound pattern for frequencies higher than 1 kHz. In contrast, AWES A, which follows a figure eight flight pattern, did not show this acoustic behavior.



(**a**) AWES A. (**b**) AWES B. (**c**) AWES C.

Figure 4: Time-averaged sound pressure levels for each airborne wind energy system (AWES).

Fig. 5 shows the comparison of time-averaged SPLs between each AWES. For this purpose, one representative case of each AWES configuration (i.e., A2, B3, and C3) has been selected. The soft-wing (A2) and fixed-wing (B3 and C3) kites exhibited a broadband acoustic trait. However, the fixed-wing kites showed an acoustic bump at high frequencies (950 to 3420 Hz for B3 and 1910 to 5180 Hz for C3) that the spectrum of the soft-wing kite did not. Additionally, the spectrum of C3 revealed narrowband peaks around 300 Hz, 600 Hz, and 1200 Hz, which could be related to the ram-air turbine. The 600 Hz and 1200 Hz peaks also seem to be harmonics of the rotations of the ram-air turbine (300 Hz), as they were equally spaced. The broadband acoustic nature of the soft-wing kite is believed to arise from its flexible, deformable structure and complex, turbulent aerodynamic interactions. This acoustic component was also higher than the broadband acoustic signature found in fixed-wing kites (i.e., 180 to 1000 Hz for B3 and 100 to 1900 Hz for C3). This may be related to the soft-wing kites' fabric-based material, which promotes constant deformation and fluttering, creating turbulence that produces a stronger broadband noise component than fixed-wing kites. This turbulence-induced acoustic trait was spread over a broad frequency range, contributing to the broadband nature of the noise. In contrast, besides the broadband noise component observed in fixed-wing kites, the acoustic bump in the high-frequency range is likely due to the efficient transmission of vibrations through the rigid materials, which is not as pronounced in the more dampened, flexible soft-wing kite.

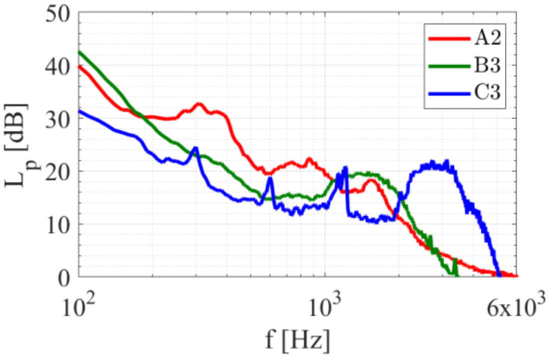


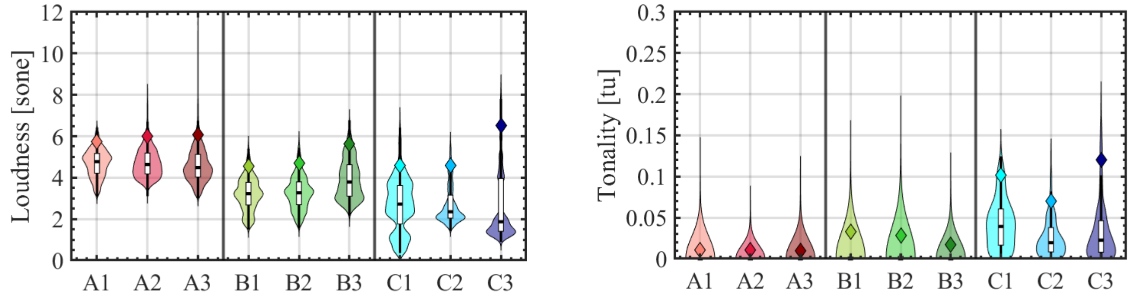
Figure 5: Comparison of time-averaged sound pressure levels of one representative recording for each airborne wind energy system (AWES).

3.2 Psychoacoustic sound quality metrics of AWESs and their relation to annoyance

An analysis of the SQMs (Table 3) revealed differences across the three AWESs, as illustrated in the violin plots in Fig. 6. Regarding loudness (Fig. 6a), AWES A recordings exhibited nearly identical values, consistent with the spectra shown in Fig. 4a. AWES B recordings showed slight variations, with B3 displaying higher loudness levels than B1 and B2. This difference can be attributed to higher noise levels in the 100 to 1000 Hz range for B3 (Fig. 4b), likely due to its closer proximity to the microphone. For AWES C, C3 exhibited higher loudness values than C1, and C1 had higher values than C2. This pattern aligns with the spectra depicted in Fig. 4c. Although the spectra for C1 and C3 appear similar, the higher loudness values of C3 compared to C1 can be explained by the ISO 532-1 loudness calculation accounting for the frequency-dependent human ear sensitivity, particularly in the 500 Hz to 6 kHz range. Among all the recordings, C3 showed the highest 5% percentile loudness values, potentially related to the sudden increase in sound levels around 1200 Hz. This sudden sound increase could be attributed to the vibration of the rigid structure of the fixed-wing kite compared to the soft-wing kite (i.e., inflatable kite made from fabric) or to the ram-air turbine on the fixed-wing kite. Regarding tonality (Fig. 6b), both AWES A and AWES B showed relatively low values compared to AWES C. This behavior can be explained by the narrowband peaks in the sound spectra observed in C1 and C3, as shown in Fig. 4c. The soft-wing kite generally exhibited the lowest tonality values, which can be explained by its tendency to produce more broadband and less tonal sound. Most of the noise from soft-wing kites is due to fabric flutter and aerodynamic noise. Regarding sharpness (Fig. 6c), AWES C notably showed higher values than AWES A and B, consistent with the sound spectra (Fig. 4) since the sharpness calculation emphasizes frequencies for critical bands above 15 Bark (corresponding to approximately f = 2700 Hz). Additionally, C3 presented the sharpest sound, which aligns with the definition of sharpness since this kite reported higher sound values than the other kites for frequencies above 2700 Hz (Fig. 5). Roughness (Fig. 6d) and fluctuation strength (Fig. 6e) quantify the perception of modulated sounds with a modulation

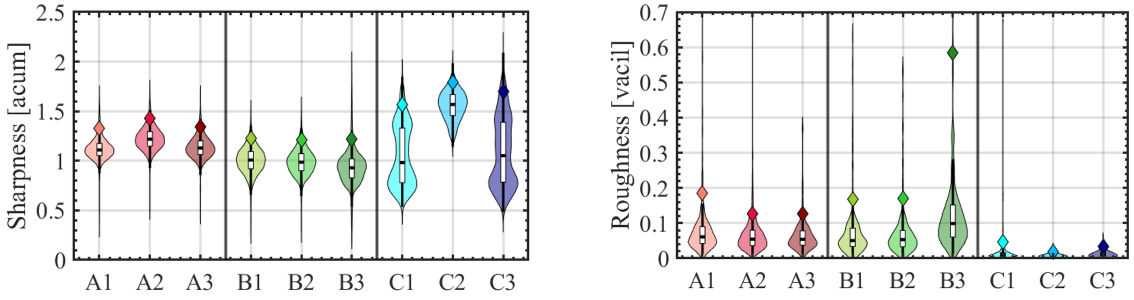
https://doi.org/10.5194/wes-2024-125
Preprint. Discussion started: 5 November 2024

frequency between 15 Hz and 300 Hz and below 20 Hz, respectively. Regarding roughness, B3 was perceived as the 'harshest' compared to all other recordings, while the AWES C recordings were the 'least harsh.' Regarding fluctuation strength, AWES B was observed to have the 'strongest beating' effect, whereas AWES C was 'less pulsating.'



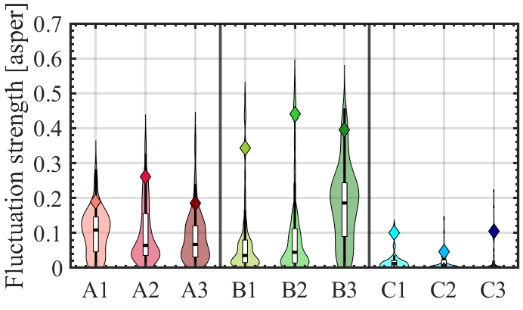
(**a**) Loudness.

(**b**) Tonality.



(**c**) Sharpness.

(**d**) Roughness.



(**e**) Fluctuation strength.

Figure 6: Violin plots of sound quality metrics for all recordings. Plot widths represent the probability density at given values in the y-axis. Diamond markers indicate the 5th percentile values. In each boxplot, the central horizontal line denotes the median values, and the edges of the white box plot represent the 25th and the 75th percentiles.

Table 3 The 5th percentile values of the five sound quality metrics per recording.

Recording	L5 (sone)	K5 (tu)	S5 (acum)	R5 (vacil)	FS5 (asper)
A1	5.75	0.011	1.32	0.18	0.19
A2	5.99	0.011	1.42	0.13	0.26
A3	6.90	0.010	1.34	0.13	0.19
B1	4.55	0.033	1.22	0.17	0.34
B2	4.71	0.028	1.21	0.17	0.44
B3	5.62	0.018	1.22	0.58	0.39
C1	4.57	0.102	1.57	0.05	0.10
C2	4.57	0.071	1.79	0.02	0.05
C3	6.53	0.121	1.70	0.03	0.10

3.2.1 Analysis of annoyance ratings

305 The percentage of highly annoyed participants (%HA) per AWES ranged from around 7 to 33%, as seen in Table 4. Overall, %HA was highest for AWES C, followed by B and then A, corresponding with the previously reported higher tonality and sharpness values for AWES C compared to B and A.

Table 4 Percentage and frequency of highly annoyed participants (%HA) per airborne wind energy system (AWES).

AWES	%HA
A (soft-wing)	2.7 (5)
B (fixed-wing)	6.7 (8)
C (fixed-wing)	18.7 (14)

https://doi.org/10.5194/wes-2024-125
Preprint. Discussion started: 5 November 2024

Additionally, the distribution of annoyance ratings obtained from the listening experiments for all AWESs is shown in Fig. 7. Pairwise comparisons between AWESs were conducted using the linear mixed-effects model. The model revealed significant differences across all three AWESs (all p-values < 0.05). In line with the previous results on the percentage of highly annoyed participants, fixed-wing kite C was, on average, rated as the most annoying [Mean (M) = 54.39, standard deviation (SD) = 22.91], followed by fixed-wing kite B (M = 39.78, SD = 22.04) and soft-wing kite A (M = 33.98, SD = 20.47). A separate linear mixed-effects model was calculated to examine whether noise annoyance depended on participant characteristics. Noise sensitivity was significantly related to annoyance [t-statistic (t) = 2.035, $p < 0.050$], indicating that individuals more sensitive to noise generally rated the recordings as more annoying. Age (t = 1.332, p = 0.187) and familiarity with AWE (t = 0.056, p = 0.956) were not significantly related to annoyance ratings.

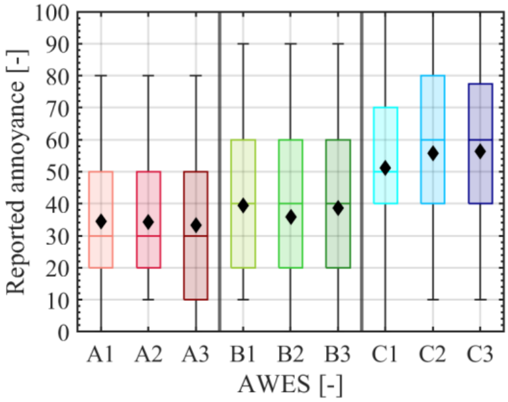


Figure 7: Distribution of annoyance ratings per recording. In each boxplot, the diamond marker denotes the mean value; the central horizontal line denotes the median values; the edges of the box are the 25th and the 75th percentiles; and the whiskers extend to the most extreme data points.

A linear mixed-effects model of the relation between annoyance ratings and SQMs showed that sharpness significantly predicted annoyance (t = 2.285, p = 0.023), while tonality (t = 0.933, p = 0.393), loudness (t = 0.416, p = 0.695), roughness (t = -0.601, p = 0.574), and fluctuation strength (t = 0.676, p = 0.529) did not. The relationship between sharpness and annoyance is visualized in Fig. 8. The results align with the finding that the annoyance ratings were significantly higher for AWES C, which exhibited higher sharpness values than AWESs A and B. To evaluate whether the impact of the SQMs on annoyance ratings varied across participants, models incorporating SQMs as fixed effects were compared with those treating them as random effects, computing the -2-log likelihood ratio between these models. The models treating all SQMs except fluctuation strength as random effects—loudness ($\chi^2(1) = 18.725$, $p < 0.001$), sharpness ($\chi^2(1) = 9.121$, $p = 0.003$), tonality ($\chi^2(1) = 7.146$, $p = 0.008$), and roughness ($\chi^2(1) = 8.723$, $p = 0.003$)—showed a significantly improved fit compared to the models treating them as fixed effects. This suggests that all tested SQMs, except for fluctuation strength, showed varying effects on annoyance ratings among individuals. To investigate whether these between-individual differences in the effects of SQMs on annoyance ratings could be explained by participant characteristics (i.e., age, AWE familiarity, and noise sensitivity), the interaction effects of the SQMs with the participant characteristics were included in the models with the random SQM effects. The results revealed that the interaction effect of participant characteristics and loudness was significant for AWE familiarity (t = -2.902, p = 0.005) but not for age (t = 0.988, p = 0.327) nor noise sensitivity (t = 0.699, p = 0.049). That is, the effect of loudness on annoyance was weaker for those more familiar with AWE. Furthermore, the interaction effect of participant characteristics and tonality was significant for age (t = -2.233, p = 0.028) but not for AWE familiarity (t = -0.452, p = 0.652) nor noise sensitivity (t = 0.045, p = 0.964). This suggests that the effect of tonality on annoyance was weaker for older individuals, also independent of participants' self-reported hearing ability. The interaction effects of participant characteristics and sharpness, roughness, and fluctuation strength were not significant for any of the included participant characteristics (with p-values ranging from 0.139 to 0.915). The full model, including all interactions between participant characteristics and SQMs, explained 19% of the variance in annoyance scores due to the fixed effects alone and 82% of the variance when both fixed and random effects were considered.

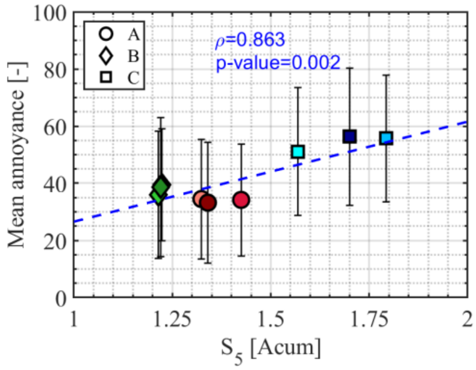


Figure 8: Linear correlation between the average annoyance rating per recording and sharpness across the airborne wind energy systems (AWESs).

3.2.2 Validity of conventional and psychoacoustic metrics in predicting annoyance ratings for AWESs

It was explored with linear mixed-effects models to what extent EPNL as a conventional metric and the Psychoacoustic Annoyance (PA) (i.e., Zwicker, More, and Di et al.) models predict the annoyance ratings reported in the experiment. Table 5 presents the values used to perform these analyses. EPNL (t = 0.700, p = 0.515) did not significantly predict the annoyance ratings. Linear mixed-effects models comparing the annoyance ratings with the estimated annoyance scores (5th percentile values) for each PA metric separately (Zwicker, More, and Di) showed that the PA metrics did not significantly predict the

https://doi.org/10.5194/wes-2024-125
Preprint. Discussion started: 5 November 2024

annoyance ratings: Zwicker (t = 0.117, p = 0.911), More (t = 0.541, p = 0.612), and Di (t = 0.466, p = 0.661). Because PA metrics heavily depend on loudness, the aforementioned normalization of all recordings might explain why the PA metrics were not significant predictors.

Table 5 EPNL and 5th percentile values of the three psychoacoustic (PA) models per recording.

	Conventional metric	PA models		
Recording	EPNL [EPNLdB]	Zwicker	More	Di et al.
A1	61.2146	6.6344	6.2142	6.6352
A2	60.3708	6.6663	6.3786	6.6832
A3	59.7796	6.9043	6.5423	6.9138
B1	59.2138	5.1527	4.9296	5.2142
B2	58.8830	5.2055	4.6478	5.2324
B3	60.8026	7.5308	6.5923	7.5316
C1	61.6614	4.8561	4.9263	5.3365
C2	61.9017	4.7786	4.8967	5.1501
C3	63.7321	7.0838	7.4605	8.2037

4 Conclusion

To characterize the relation between sound quality metrics (SQMs) and noise annoyance for airborne wind energy systems (AWESs), a laboratory listening experiment was conducted, including acoustic analyses of the used, scaled sound recordings from three AWESs. The findings indicate differences in the SQMs across the three AWESs, with the fixed-wing kites showing higher loudness and tonality values than the soft-wing kite. This is mostly due to the former's aerodynamic characteristics and the use of a ram-air turbine for the C fixed-wing kites. The lower values for the soft-wing kite can be explained by its tendency to produce more broadband and less tonal sound. In addition, the fixed-wing kites presented relatively higher sharpness values than the soft-wing kite. When using the results from the SQM analyses to predict participants' annoyance ratings for the recordings from the three AWESs, it was found that only sharpness significantly predicted annoyance. This corresponds with the finding that participants were significantly more annoyed by recordings from the fixed-wing kites than the soft-wing kite. Interestingly, the effect of some SQMs on annoyance depended on participant characteristics, with loudness having a weaker impact on annoyance for those more familiar with AWE and tonality having a weaker effect on annoyance for older individuals. These moderation effects should be cautiously interpreted because they could be random due to the non-probability sampling and the lack of representativeness of the sample in this study. Neither the conventional metric (i.e., EPNL) nor the PA metrics significantly predicted the annoyance ratings. However, this is most likely due to scaling the recordings to the same dBA value.

Data availability

The data that support the findings of this study are openly available in 4TU.ResearchData at https://data.4tu.nl/private_datasets/11wCOgnga-GfjzL4oufzdd1UPRPP2YVFVnjRaKcHJhE, reference number [doi: 10.4121/2716b49f-b44c-400a-a873-eea276b081f6].

Author contributions

HS, RMYV, DR, RMM, and RS planned the campaign; HS performed the measurements; HS, RMYV, and PvG analyzed the data; HS and RMYV wrote the manuscript draft; DR, RMM, PvG, and RS reviewed and edited the manuscript.

Competing interests

At least one of the authors is a member of the editorial board of Wind Energy Science. Roland Schmehl is a co-founder of and advisor for the start-up company Kitepower B.V.

Acknowledgments

We thank the three airborne wind energy developers and Naomi Bouman for providing the sound measurements and acoustic data. We are grateful to Niels Adema with helping with planning the campaign and supporting the measurements.

Financial support

This work was supported by the Dutch Research Council (NWO) and Kitepower B.V. [grant number 17628]. The research work of Roberto Merino-Martínez is funded by the project Listen to the future [project number 20247] of the research program Veni 2022 (Domain Applied and Engineering Sciences), which NWO partly finances.

References

Aguinis, H., Gottfredson, R. K., and Culpepper, S. A.: Recommendations for estimating cross-level interaction effects using multilevel modeling, Academy of Management Proceedings, 2013, 10839, https://doi.org/10.5465/ambpp.2013.10839abstract, 2013.

Alamir, M. A., Hansen, K. L., Zajamsek, B., and Catcheside, P.: Subjective responses to wind farm noise: A review of laboratory listening test methods, Renewable and Sustainable Energy Reviews, 114, 109317, https://doi.org/10.1016/j.rser.2019.109317, 2019.

Bakker, R. H., Pedersen, E., van den Berg, G. P., Stewart, R. E., Lok, W., and Bouma, J.: Impact of wind turbine sound on annoyance, self-reported sleep disturbance and psychological distress, Science of The Total Environment, 425, 42–51, https://doi.org/10.1016/J.SCITOTENV.2012.03.005, 2012.

Bates, D., Maechler, M., Bolker, B., Walker, S., Bojesen Christensen, R. H., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., Bauer, A., Krivitsky, P. N., Tanaka, E., and Jagan, M.: Package ‘lme4’ - Linear Mixed-Effects Models using “Eigen” and S4, http://dk.archive.ubuntu.com/pub/pub/cran/web/packages/lme4/lme4.pdf, 2024.

Boucher, M. A., Christian, A. W., Krishnamurthy, S., Tracy, T., Begault, D. R., Shepherd, K., and Rizzi, S. A.: Toward a Psychoacoustic Annoyance Model for Urban Air Mobility Vehicle Noise, 2024.

Bouman, N.: Aeroacoustics of airborne wind energy systems, Master thesis, Delft University of Technology, Delft, 2023.

Brink, M., Schreckenberg, D., Vienneau, D., Cajochen, C., Wunderli, J.-M., Probst-Hensch, N., and Röösli, M.: Effects of scale, question location, order of response alternatives, and season on self-reported noise annoyance using ICBEN scales: A field experiment, Int J Environ Res Public Health, 13, 1163, https://doi.org/10.3390/ijerph13111163, 2016.

BVG Associates: Getting airborne - the need to realise the benefits of airborne wind energy for net zero/White Paper for Airborne Wind Europe, https://doi.org/10.5281/zenodo.7809185, 2022.

Cherubini, A., Papini, A., Vertechy, R., and Fontana, M.: Airborne Wind Energy Systems: A review of the technologies, Renewable and Sustainable Energy Reviews, 51, 1461–1476, https://doi.org/10.1016/J.RSER.2015.07.053, 2015.

Daniel, P. and Weber, R.: Psychoacoustical roughness: Implementation of an optimized model, Acta Acustica united with Acustica, 83, 113–123, 1997.

Deutsches Institut für Normung: DIN 45692: Measurement technique for the simulation of the auditory sensation of sharpness, 2009.

Fagiano, L., Quack, M., Bauer, F., Carnel, L., and Oland, E.: Autonomous Airborne Wind Energy Systems: Accomplishments and Challenges, Annual Review of Control, Robotics, and Autonomous Systems , 5, 603–631, https://doi.org/10.1146/ANNUREV-CONTROL-042820-124658, 2022.

Greco, G. F., Merino-Martínez, R., Osses, A., and Langer, S. C.: SQAT: a MATLAB-based toolbox for quantitative sound quality analysis, in: INTER-NOISE and NOISE-CON Congress and Conference Proceedings, 7172–7183, https://doi.org/10.3397/IN_2023_1075, 2023.

Griefahn, B.: Determination of noise sensitivity within an internet survey using a reduced version of the Noise Sensitivity Questionnaire, in: The Journal of the Acoustical Society of America, 3449–3449, https://doi.org/10.1121/1.2934269, 2008.

Haac, T. R., Kaliski, K., Landis, M., Hoen, B., Rand, J., Firestone, J., Elliott, D., Hübner, G., and Pohl, J.: Wind turbine audibility and noise annoyance in a national U.S. survey: Individual perception and influencing factors, J Acoust Soc Am, 146, 1124–1141, https://doi.org/10.1121/1.5121309, 2019.

Hagen, L. van, Petrick, K., Wilhelm, S., and Schmehl, R.: Life-Cycle Assessment of a Multi-Megawatt Airborne Wind Energy System, Energies (Basel), 16, 1750, https://doi.org/10.3390/en16041750, 2023.

Hansen, K. L., Nguyen, P., Micic, G., Lechat, B., Catcheside, P., and Zajamšek, B.: Amplitude modulated wind farm noise relationship with annoyance: A year-long field study, J Acoust Soc Am, 150, 1198–1208, https://doi.org/10.1121/10.0005849, 2021.

Hoen, B., Firestone, J., Rand, J., Elliot, D., Hübner, G., Pohl, J., Wiser, R., Lantz, E., Haac, T. R., and Kaliski, K.: Attitudes of U.S. Wind Turbine Neighbors: Analysis of a Nationwide Survey, Energy Policy, 134, 110981, https://doi.org/10.1016/j.enpol.2019.110981, 2019.

Hong, O., Ronis, D. L., and Antonakos, C. L.: Validity of Self-Rated Hearing Compared With Audiometric Measurement Among Construction Workers, Nurs Res, 60, 326–332, https://doi.org/10.1097/NNR.0b013e3182281ca0, 2011.

Hübner, G., Pohl, J., Hoen, B., Firestone, J., Rand, J., Elliott, D., and Haac, R.: Monitoring annoyance and stress effects of wind turbines on nearby residents: A comparison of U.S. and European samples, Environ Int, 132, 105090, https://doi.org/10.1016/J.ENVINT.2019.105090, 2019.

IEA: Net Zero Roadmap - A Global Pathway to Keep the 1.5 °C Goal in Reach, 2023.

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

IRENA: Offshore renewables: An action agenda for deployment, Abu Dhabi, 2021.

ISO/TC 43: ISO 532-1:2017 Acoustics Methods for calculating loudness Part 1: Zwicker method, 2017.

Judd, C. M., Westfall, J., and Kenny, D. A.: Experiments with more than one random factor: designs, analytic models, and statistical power, Annu Rev Psychol, 68, 601–625, https://doi.org/10.1146/annurev-psych-122414-033702, 2017.

van Kamp, I. and van den Berg, F.: Health Effects Related to Wind Turbine Sound: An Update, Int J Environ Res Public Health, 18, 9133, https://doi.org/10.3390/ijerph18179133, 2021.

Kephalopoulos, S., Paviotti, M., Anfosso-Lédée, F., Van Maercke, D., Shilton, S., and Jones, N.: Advances in the development of common noise assessment methods in Europe: The CNOSSOS-EU framework for strategic environmental noise mapping, Science of The Total Environment, 482–483, 400–410, https://doi.org/10.1016/j.scitotenv.2014.02.031, 2014.

Ki, J., Yun, S.-J., Kim, W.-C., Oh, S., Ha, J., Hwangbo, E., Lee, H., Shin, S., Yoon, S., and Youn, H.: Local residents' attitudes about wind farms and associated noise annoyance in South Korea, Energy Policy, 163, 112847, https://doi.org/10.1016/j.enpol.2022.112847, 2022.

Lee, S., Kim, K., Choi, W., and Lee, S.: Annoyance caused by amplitude modulation of wind turbine noise, Noise Control Eng J, 59, 38–46, https://doi.org/10.3397/1.3531797, 2011.

Merino-Martínez, R., Pieren, R., and Schäffer, B.: Holistic approach to wind turbine noise: From blade trailing-edge modifications to annoyance estimation, Renewable and Sustainable Energy Reviews, 148, 111285, https://doi.org/10.1016/j.rser.2021.111285, 2021.

Merino-Martínez, R., von den Hoff, B., and Simons, D. G.: Design and acoustic characterization of a psycho-acoustic listening facility, in: Proceedings of the 29th International Congress on Sound and Vibration, 2023.

Michaud, D. S., Keith, S. E., Feder, K., Voicescu, S. A., Marro, L., Than, J., Guay, M., Bower, T., Denning, A., Lavigne, E., Whelan, C., Janssen, S. A., Leroux, T., and van den Berg, F.: Personal and situational variables associated with wind turbine noise annoyance, J Acoust Soc Am, 139, 1455–1466, https://doi.org/10.1121/1.4942390, 2016.

Miedema, H. M. E. and Vos, H.: Exposure-response relationships for transportation noise, J Acoust Soc Am, 104, 3432–3445, https://doi.org/10.1121/1.423927, 1998.

More, S. R.: Aircraft noise characteristics and metrics, Purdue University, West Lafayette, 2010.

Osses Vecchi, A., García León, R., and Kohlrausch, A.: Modelling the sensation of fluctuation strength, in: Proceedings of Meetings on Acoustics, 050005, https://doi.org/10.1121/2.0000410, 2016.

Passer, M. W.: Research Methods: Concepts and Connections, 1st ed., Worth Publishers, New York, 2014.

Pawlaczyk-Łuszczyńska, M., Zaborowski, K., Dudarewicz, A., Zamojska-Daniszewska, M., and Waszkowska, M.: Response to noise emitted by wind farms in people living in nearby areas, Int J Environ Res Public Health, 15, 1575, https://doi.org/10.3390/ijerph15081575, 2018.

Pedersen, E. and Larsman, P.: The impact of visual factors on noise annoyance among people living in the vicinity of wind turbines, J Environ Psychol, 28, 379–389, https://doi.org/10.1016/j.jenvp.2008.02.009, 2008.

Pedersen, E. and Persson Waye, K.: Wind turbine noise, annoyance and self-reported health and well-being in different living environments, Occup Environ Med, 64, 480–486, https://doi.org/10.1136/oem.2006.031039, 2007.

Pedersen, E., van den Berg, F., Bakker, R., and Bouma, J.: Can road traffic mask sound from wind turbines? Response to wind turbine sound at different levels of road traffic sound, Energy Policy, 38, 2520–2527, https://doi.org/10.1016/j.enpol.2010.01.001, 2010.

Pieren, R., Bertsch, L., Lauper, D., and Schäffer, B.: Improving future low-noise aircraft technologies using experimental perception-based evaluation of synthetic flyovers, Science of The Total Environment, 692, 68–81, https://doi.org/10.1016/j.scitotenv.2019.07.253, 2019.

Pohl, J., Gabriel, J., and Hübner, G.: Understanding stress effects of wind turbine noise – The integrated approach, Energy Policy, 112, 119–128, https://doi.org/10.1016/j.enpol.2017.10.007, 2018.

R Core Team: R: A Language and Environment for Statistical Computing, 2023.

Schäffer, B., Schlittmeier, S. J., Pieren, R., Heutschi, K., Brink, M., Graf, R., and Hellbrück, J.: Short-term annoyance reactions to stationary and time-varying wind turbine and road traffic noise: A laboratory study, J Acoust Soc Am, 139, 2949–2963, https://doi.org/10.1121/1.4949566, 2016.

Schäffer, B., Pieren, R., Schlittmeier, S., and Brink, M.: 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, 1029, https://doi.org/10.3390/ijerph15051029, 2018.

Schäffer, B., Pieren, R., Wissen Hayek, U., Biver, N., and Grêt-Regamey, A.: Influence of visibility of wind farms on noise annoyance – A laboratory experiment with audio-visual simulations, Landsc Urban Plan, 186, 67–78, https://doi.org/10.1016/J.LANDURBPLAN.2019.01.014, 2019.

Schmidt, H., de Vries, G., Renes, R. J., and Schmehl, R.: The Social Acceptance of Airborne Wind Energy: A Literature Review, Energies (Basel), 15, 1384, https://doi.org/10.3390/EN15041384, 2022.

Schmidt, H., Leschinger, V., Müller, F. J. Y., de Vries, G., Renes, R. J., Schmehl, R., and Hübner, G.: How do residents perceive energy-producing kites? Comparing the community acceptance of an airborne wind energy system and a wind farm in Germany, Energy Res Soc Sci, 110, 103447, https://doi.org/10.1016/j.erss.2024.103447, 2024.

Schutte, M., Marks, A., Wenning, E., and Griefahn, B.: The development of the noise sensitivity questionnaire, Noise Health, 9, 15–24, https://doi.org/10.4103/1463-1741.34700, 2007.

Solman, H. and Mattijs, S.: UPWARDS - Project deliverable D7.2 - Reports with data from on and offline panels, 2021.

Torija, A. J., Roberts, S., Woodward, R., Flindell, I. H., McKenzie, A. R., and Self, R. H.: On the assessment of subjective response to tonal content of contemporary aircraft noise, Applied Acoustics, 146, 190–203, https://doi.org/10.1016/j.apacoust.2018.11.015, 2019.

Turunen, A. W., Tiittanen, P., Yli-Tuomi, T., Taimisto, P., and Lanki, T.: Self-reported health in the vicinity of five wind power production areas in Finland, Environ Int, 151, 106419, https://doi.org/10.1016/j.envint.2021.106419, 2021.

UNFCCC: Paris Agreement to the United Nations Framework Convention on Climate Change 2015, 2016.

Vermillion, C., Cobb, M., Fagiano, L., Leuthold, R., Diehl, M., Smith, R. S., Wood, T. A., Rapp, S., Schmehl, R., Olinger, D., and Demetriou, M.: Electricity in the air: Insights from two decades of advanced control research and experimental flight testing of airborne wind energy systems, Annu Rev Control, 52, 330–357, https://doi.org/10.1016/j.arcontrol.2021.03.002, 2021.

Vos, H., Lombardi, F., Joshi, R., Schmehl, R., and Pfenninger, S.: The potential role of airborne and floating wind in the North Sea region, Environmental Research: Energy, 1, 025002, https://doi.org/10.1088/2753-3751/ad3fbc, 2024.

Yokoyama, S. and Tachibana, H.: Perception of tonal components contained in wind turbine noise, in: INTER-NOISE and NOISE-CON Congress and Conference Proceedings, 2435–2446, 2016.

Yonemura, M., Lee, H., and Sakamoto, S.: Subjective Evaluation on the Annoyance of Environmental Noise Containing Low-Frequency Tonal Components, Int J Environ Res Public Health, 18, 7127, https://doi.org/10.3390/ijerph18137127, 2021.

Zwicker, E. and Fastl, H.: Psychacoustics: Facts and models, 2nd ed., Springer, Berlin, 1999.