Human response to eVTOL drone sound: an online listening experiment exploring the effects of operational and contextual factors

James Woodcock[1]*, Adam Thomas[1], David Hiller[1],
Ana Luisa Pereira Maldonado[1], Laura McLeod[1], Calum Sharp[1] and
Fiona Smith[2]

[1]Arup, London, United Kingdom, [2]AGS Airports, Glasgow, United Kingdom

Introduction: Project CAELUS is developing the United Kingdom's first national distribution network using drones to transport vital medical supplies throughout Scotland. Noise is a major barrier to public acceptance of drone networks, yet empirical data on the human response to drones used in a medical delivery context remains limited. This study addresses that gap by investigating the annoyance response to sounds from the eVTOL medical delivery drone used in Project CAELUS.

Methods: An online listening experiment was conducted to assess annoyance related to overflight (N − 425) and take-off (N − 278) operations. The experiment examined the effects of listener−drone distance, ambient soundscape (remote rural, rural village, urban), and contextual framing (medical delivery vs. no context) on annoyance. Data were analysed using aligned rank transform ANOVAs to test for main effects and interactions for each factor.

Results: Aligned rank transform ANOVAs revealed significant effects of listener−drone distance, ambient soundscape, and contextual framing on annoyance ($p < 0.01$ for all three factors). Annoyance decreased with increasing distance from the drone and was higher in quieter ambient soundscapes. Providing contextual information about the medical use of the drone significantly reduced annoyance.

Discussion: Findings indicate that both acoustic and non-acoustic factors influence perceived annoyance from drone operations. In particular, contextual information about medical use reduced annoyance, suggesting that effective community engagement may improve public acceptance of drone networks.

KEYWORDS

drone sound, perception, human response, non-acoustic factors, noise annoyance

1 Introduction

The use of drones and uncrewed aerial systems (UAS) for services such as deliveries, medical logistics, and infrastructure inspections is expected to grow significantly in the coming years (PwC, 2023; CAA, 2024). This will introduce a new source of noise into the environment, raising concerns about potential impacts on public health, including annoyance and sleep disturbance. Public perception and the development of regulatory frameworks have emerged as critical challenges for the implementation of advanced air mobility (AAM)[1] networks, with regulatory authorities in the United Kingdom, the United States, and Europe working to establish such frameworks.

Noise has been identified by the European Union Aviation Safety Agency (EASA) as the second most significant barrier to the adoption of UAM, following safety concerns (EASA, 2021). As a result, there is growing interest in understanding how people perceive and respond to the sound of AAM operations.

1.1 Project CAELUS

This paper presents the results of a listening experiment designed to investigate the annoyance response to the sound of an electric vertical take-off and landing (eVTOL) drone within different ambient soundscape environments. The study was conducted as part of the Care & Equity - Healthcare Logistics UAS Scotland project (Project CAELUS), which aims to develop a national drone network for transporting essential medical supplies (including medicines, blood, and diagnostic materials) across Scotland, particularly to remote communities.

Globally, approximately 20 documented drone trials for medical applications have been conducted in countries including Haiti, Bhutan, Papua New Guinea, Rwanda, Ghana, the United States, Switzerland, Germany, the Netherlands, and the United Kingdom (see, for example, (Choi-Fitzpatrick et al., 2013; DronePrep, 2021; Karim et al., 2021; Musanabaganwa et al., 2020; Scott and Scott, 2017; Swiss Foundation for Mine Action FSD, 2014; TU Delft, 2020; UPS, 2021; Zip line and Pfizer, 2021; Sylverken et al., 2022)). Most of these trials were temporary, conducted outside of airport environments, and nearly half focused on emergency response or COVID-19 vaccine delivery. Only two had formal approval from national regulatory bodies.

Unlike previous trials, Project CAELUS involves the creation of a full scale pilot for a national drone network in Scotland. As well as flying directly between healthcare facilities, drones will also fly to and from airports under regulated airspace. The entire ecosystem has been developed under the project, demonstrating the feasibility of using drones for medical deliveries, validating the necessary

FIGURE 1
CAELUS drone on site during the field measurements.

technological and regulatory frameworks, and laying the groundwork for the integration of UAS into the UK's airspace.

The impact of CAELUS has been substantial. The project delivered the first public business case for a medical drone service and presented an economic assessment of the impacts medical use drones could have on the Scottish economy, estimating a GVA contribution of £16 m, just under 700 additional jobs and over £3 m in tax revenue generation. It created the first movement of items between mainland health boards in Scotland by drone, and developed a new blueprint for NHS innovation projects.

As part of the project, a large scale online listening experiment has been conducted to investigate the annoyance response to the sound of the drone used in the project (see Figure 1). This study provides empirical data on perceived annoyance and contributes to the broader regulatory and community engagement efforts required to support the integration of UAS into everyday healthcare operations. A unique feature of this study is its comparison of the annoyance response to drones used for healthcare purposes versus those with unspecified applications (e.g., package or food delivery). This is an important factor for understanding the impacts of drone noise, as public acceptance may vary depending on the perceived social value of the operations.

1.2 Human response to drone sound

Exposure-response curves for different transportation noise sources have been developed through decades of research (Miedema and Oudshoorn, 2001). It is well established that, for equivalent levels of A-weighted ($L_{Aeq,T}$) noise exposure, the annoyance response to transportation noise varies between road, rail, and conventional aviation sources. Typically, aviation noise elicits the highest annoyance, followed by road and rail noise (Miedema and Oudshoorn, 2001). These differences arise from variations in acoustic characteristics, such as temporal variations and frequency content, which are not adequately captured by the $L_{Aeq,T}$ metric along with personal, attitudinal, and contextual factors (non-acoustic factors). Several studies have demonstrated that non-

---

1 Advanced Air Mobility (AAM) is an umbrella term that encompasses a range of emerging aviation technologies and operations, including uncrewed aerial systems (UAS), urban air mobility (UAM), and other drone-based platforms designed for passenger and cargo transport in both urban and regional settings.

acoustic factors can significantly influence the annoyance response to transportation noise. These factors include self-reported noise sensitivity, fear related to the noise source, existing environmental characteristics, and current activity (for a summary, see (Guski, 1999; Miedema and Vos, 1999)).

AAM vehicles exhibit a different sound signature to conventional aircraft, often characterised by prominent tonal components. These tonal components can consist of several closely spaced frequencies due to multiple rotors operating at slightly different rotation speeds (Schäffer et al., 2021; Lotinga et al., 2023). Psychoacoustic studies on AAM vehicle sound suggest that, similar to other transportation noise sources, overall loudness (expressed by psychoacoustic loudness, perceived noise level [PNL], or A-weighted sound levels) is the primary contributor to annoyance (Lotinga et al., 2023). Additionally, psychoacoustic metrics such as sharpness, fluctuation strength, roughness, and tonality have been shown to influence perception (Lotinga et al., 2023).

Studies comparing drone noise to other transportation noise sources have demonstrated that drone noise is perceived as more annoying than road traffic and conventional aviation noise for equivalent levels of noise exposure (Christian and Cabell, 2017; Gwak et al., 2020; Kawai et al., 2024). Differences in response have also been found for different drone operational modes after controlling for the sound level of the drone. Green et al. (2024) found that landing operations were perceived as the loudest and most annoying, followed by take-offs and hovering. Flyovers were perceived to be the least loud and annoying of the different operations. Kawai et al. (2024) found that drone landings and take-offs were perceived as more annoying than flyovers. Additionally, slow flyovers were rated as more annoying than fast flyovers. Lotinga et al. (2025) found that drone take-offs and landings were perceived as more annoying than flyovers. Additionally, hovering was found to be the most annoying of the three operational modes.

Torija et al. (2020) investigated the perception of drone noise within existing soundscapes and found that, in environments heavily impacted by road traffic noise, the presence of drone noise led to only minor changes in perceived loudness, annoyance, and pleasantness. Conversely, in quieter environments, participants reported significantly higher perceived loudness and annoyance and lower pleasantness. Aalmoes et al. (2023) found that drone flyovers were considerably more noticeable, loud, and annoying in quiet rural settings compared to urban environments. Lotinga et al. (2025) found that drone noise was perceived as more annoying in quieter environments compared to noisier ones, with the difference between the drone noise level and the ambient noise level being a key factor in determining the annoyance response.

There have been a limited number of studies exploring the effects of attitudinal and contextual factors on the response to AAM noise. Lotinga et al. (2025) found that personal attitudes towards drones significantly influenced the perception of drone noise. Individuals with positive attitudes towards drones, such as those who see them as beneficial for society, were less likely to find drone noise annoying. Conversely, those with negative attitudes, such as concerns about privacy or safety, were more likely to perceive drone noise as annoying. Kawai et al. (2024) investigated the influence of

context on annoyance by priming participants with either an instruction emphasising the medical usage of drones, the commercial use of drones, or a neutral instruction. However, no significant effects of the priming were found.

1.3 Research question and approach

The drone that is being used in the CAELUS project is a lift and cruise type eVTOL drone with a cruising speed of 55 kt and a maximum take-off weight of 17 kg. There have been no systematic studies of annoyance due to this type of drone. As highlighted in the literature discussed in this section, annoyance can be influenced by the operational mode (i.e., cruise/take-off/landing) and the effect of the ambient soundscape in which the drone is operating. There is also a lack of evidence on the influence of contextual factors (i.e., the drone being used in a medical delivery context) on annoyance. Considering this, this paper presents a study designed to collect data on annoyance related to overflight and take-off operations of the drone used in Project CAELUS. The study addresses the research question, "How does listener-drone distance, ambient soundscape type, and contextual information influence the annoyance response to drone overflight and take-off operations?"

2 Methodology

2.1 Study design

An online listening experiment was designed to address the research question set out in Section 1.3. The advantage of online delivery over in-person laboratory tests is the ability to reach a large and diverse sample of the population at the expense of diminished control over the test environment and overall calibration. Online remote listening tests have been demonstrated to replicate the results of in-person laboratory testing (Krishnamurthy et al., 2023).

Two separate experiments were conducted focusing on drone overflights and take-off operations respectively. The stimuli for the experiments consisted of a single overflight or take-off operation presented against an ambient soundscape. The decision to deploy two separate experiments was taken to minimise the length of the listening experiment and reduce dropout rate. Each experiment took around 15 min to complete and participants were able to choose to participate in one or both of the experiments. In total, 425 participants completed the overflight experiment and 278 participants completed the take-off experiment.

The following independent variables were investigated in the experiments:

- Listener-drone distance
- Type of ambient environment
- Context of use of the drone

Variations in listener-drone distance were achieved by adjusting the drone altitude (120, 90, and 60 m) for overflight operations and the distance between the listener and the final approach and take-off area (FATO) (30, 60 and 120 m) for take-off operations. The sound of the drone operations were presented against three distinct
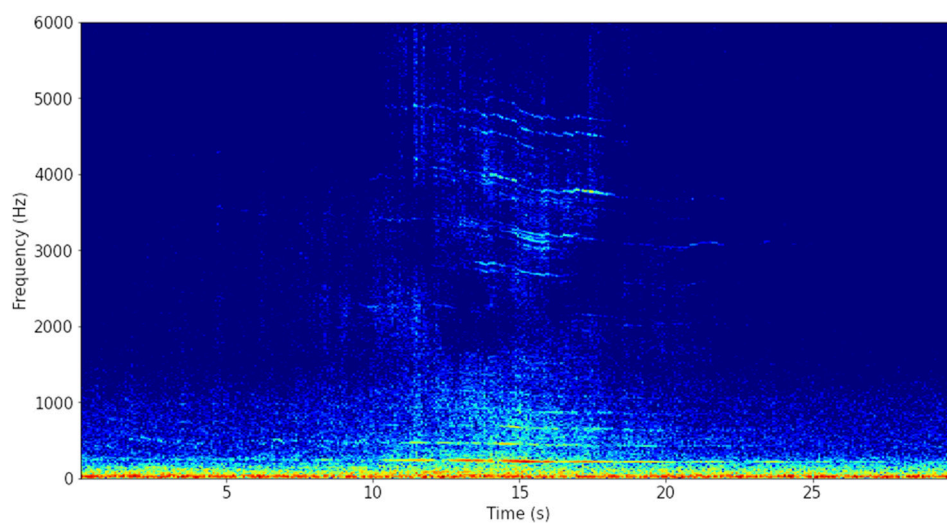
FIGURE 2
Spectrogram of a typical overflight event.

ambient soundscapes representing remote rural, rural village, and urban environments. A factorial design was employed for each of the experiments, resulting in nine stimuli per experiment (3 listener-drone distances × 3 ambient soundscapes).

The influence of contextual information on annoyance was explored as a between-participants factor. Two versions of the experiment were conducted concurrently with different groups of participants. One group received contextual information via an online portal regarding the intended use of the drone for medical deliveries, whilst the other group received no contextual information. All other experimental parameters and procedures remained consistent between the two groups.

2.2 Auralisation of the project CAELUS drone

2.2.1 Field measurements

Field measurements were conducted on 9th and 10th August 2022 at Westcott Innovation Centre in southern England to gather audio recordings and acoustic data for the type of drone used in Project CAELUS. Measurements were taken of overflight and take-off/hover/landing operations. The drone's position was tracked via GPS throughout the measurements, enabling synchronisation with the acoustic measurements. A photograph of the drone is shown in Figure 1.

The sound of the drone was captured using GRAS 46AE 1/2 inch microphones mounted on aluminium ground plates, with data acquired as raw 16-bit WAV files at a sampling frequency of 48 kHz. High-quality audio recordings were also conducted in ambisonic format using ST450 MKII Soundfield microphones and monaurally using DPA 4006 omnidirectional condenser microphones. Calibrated acoustic measurements were conducted using RION NL-52 Class 1 sound level meters. Meteorological conditions were logged throughout the measurements, with weather conditions remaining still and dry, and windspeeds below 5 m/s. Background noise on site was characterised by distant road traffic, other aircraft from a nearby airfield, industry, and wildlife.

Measurements of overflight operations were taken while the drone completed circuits at altitudes of 60 m and 100 m at a cruising speed of 55 kt. Measurements were undertaken at planimetric distances of 40, 60 and 140 m from the flight path (due to site constraints, it was not possible to measure directly beneath the flight path), with each measurement position consisting of a groundplate microphone, a SoundField microphone, a condenser microphone, and a sound level meter. A total of 40 individual overflight events were measured. The average L_{AFmax} of the overflight events at the closest measurement position was 48 dB at an altitude of 60 m and 44 dB at an altitude of 100 m.

Take-off and landing operations were measured at distances of 10 m and 60 m from the take-off point. As with the overflight operations, the measurement positions consisted of a groundplate microphone, a SoundField microphone, a condenser microphone, and a sound level meter. The drone completed seven take-off, hover, and landing cycles up to altitudes of 40, 60, or 100 m. The average L_{AFmax} of the take-off/hover/landing events was 82 dB 10 m from the take-off point and 66 dB 60 m from the take-off point.

Spectrograms of representative overflight and take-off/hover/landing events measured during the field survey are shown in Figures 2, 3 respectively. The overflight spectrogram exhibits tonal components related to the BPF and its harmonics, higher frequency components relating to motor noise, and a broadband noise component likely due to airframe noise. The spectrogram of take-off/hover/landing is dominated by components of the BPF of the four lift rotors.

2.2.2 Auralisation methodology

The level of background noise during the field measurements was such that the signal-to-noise level was too low to use the overflight recordings directly as stimuli in the listening experiment as the background noise in the recordings was clearly audible during the overflights. An auralisation method was required
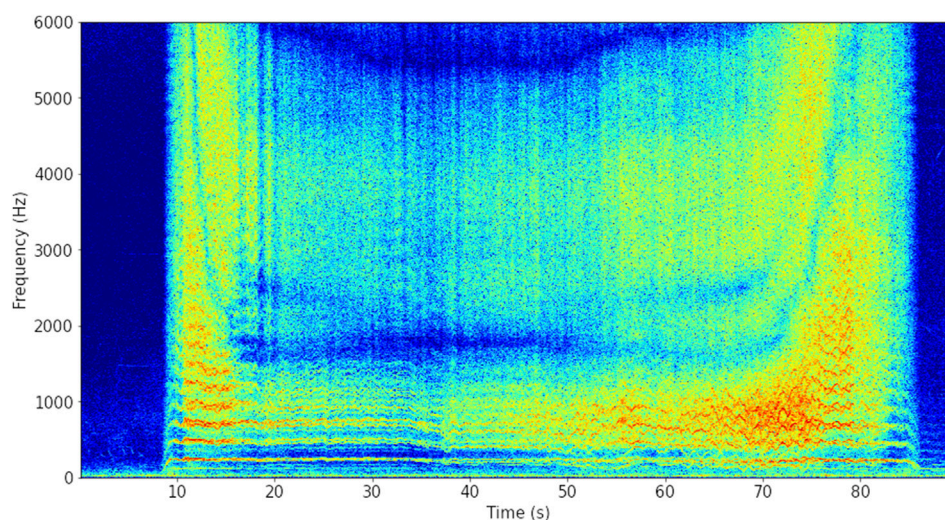
FIGURE 3
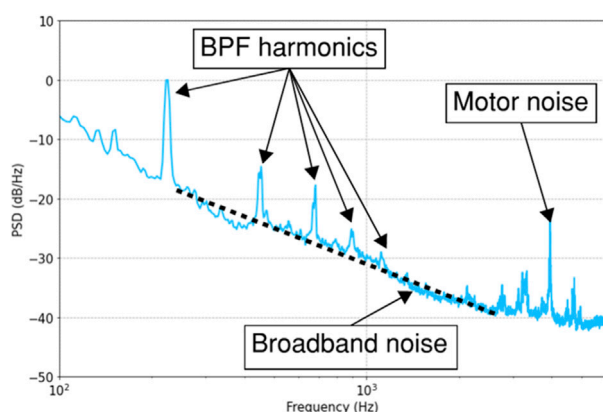Spectrogram of a typical take-off event.

FIGURE 4
Power spectral density of a typical de-Dopplerised
overflight signal.

to synthesise clean audio that could be presented with the different ambient sound environments used in the study.

The overflight auralisation methodology follows the principles of the virtual microphone signal approach proposed by Heutschi et al. (2020). In this approach, virtual microphone signals representing a microphone 1.5 m from the drone and adjusted for manoeuvre-specific and turbulence-induced rotational speed variations are used to simulate the sound pressure at a given listener position by applying frequency-dependent directivity, Doppler shift, and time-frequency dependent propagation effects. In the present study, the virtual microphone signals were synthesised based on an analysis of the groundplate microphone data taken in the field measurements.

The groundplate microphone data were de-Dopplerised in the time domain based on the time varying delay between the drone and the receiver position (calculated from the drone GPS data). The de-Dopplerised overflight signals were then used to identify the

frequency and amplitude of the BPF harmonic components, motor tonal components, and broadband components as shown in Figure 4. Virtual microphone signals were synthesised for each rotor and motor through additive synthesis of sinusoidal signals at the identified frequencies and amplitudes. The broadband noise component was generated using shaped noise fit by inspection to the spectrum of the de-Dopplerised signal. Directivity was incorporated using the angle and frequency-dependent directivity pattern of a multicopter drone reported in Heutschi et al. (2020), which was implemented as a second-order high-shelf filter.

To simulate the non-stationary operation of the rotors, a time-dependent delay $\Delta\tau(t)$ was applied to each of the synthesised rotor signals (Heutschi et al., 2020). The time derivative of $\Delta\tau$ is a signal, denoted as $s(t) = \frac{R_{ref} - R(t)}{R(t)}$, with time-varying rotational speed $R(t)$ and mean rotational speed R_{ref}. The statistical properties of the time-varying rotor speed $R(t)$ and R_{ref} were determined through an analysis of the variation of the BPF measured over an overflight event. This analysis involved tracking the variation in the fundamental frequency over the duration of the de-Dopplerised overflight signal and estimating the RPM according to the relationship $RPM = 30 f_0$. The histogram in Figure 5 shows the estimated variation in rotor RPM. The mean and standard deviation of this distribution were used to define the statistical characteristics of $s(t)$ and hence $\Delta\tau$.

Finally, time-varying filters calculated according to the spatial position of the drone on a given trajectory were applied to simulate geometric spreading (assuming point source propagation), Doppler shift, atmospheric absorption, and amplitude modulation due to air turbulence.

Unlike the overflight recordings, the measurements of take-off/hover/landing operations were made close to the FATO location and therefore had sufficient signal-to-noise level to use directly as stimuli in the experiment. To auralise these operations at different distances, the monoaural recordings taken at 10 m from the take-off point were adjusted using time varying filters to take into account geometric spreading (assuming point source propagation), Doppler shift, and air absorption.

FIGURE 5
Histogram showing the variation in rotor RPM over a typical overflight passby.
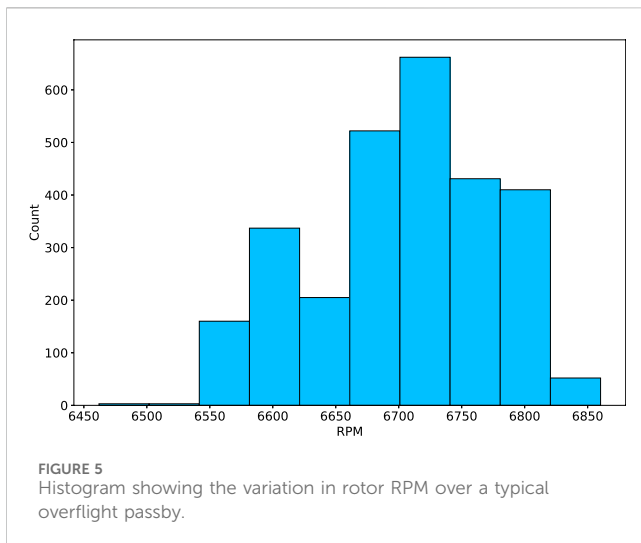
The auralisations were spatialised by encoding the signal as third-order ambisonic, taking into account the trajectory of the drone. The use of ambisonics allows for accurate spatial encoding of the drone, preserving directional cues and enabling realistic binaural rendering over headphones via HRTF-based decoding. For overflights the drone was simulated passing from left to right at a speed of 55 kt directly overhead. For the take-off operations, the drone began on the ground and ascended vertically to an altitude of 100 m, with no lateral movement.

The ambisonic drone auralisations were mixed with B-format ambisonic recordings of three different ambient environments (rural, rural village, and urban), which were captured using a ST450 MKII Soundfield microphone alongside a Class 1 sound level meter. The drone auralisations and soundscape recordings were calibrated to their measured on-site levels in Arup's SoundLab® (Hiller et al., 2021) (a 16-channel ambisonic listening room) prior to being mixed. The $L_{Aeq,T}$ of the combined drone and soundscape audio was measured in the sweetspot of the ambisonic array to allow subsequent calibration of the binaural renders.

A description of the three ambient environments, along with their $L_{Aeq,T}$, is provided below:

- Remote Rural, $L_{Aeq,T}$ = 35dB: This environment is representative of the wilder and more rural areas of Scotland, where nature serves as the dominant source of sound.
- Rural Village $L_{Aeq,T}$ = 45dB: This environment represents a typical village setting, characterised by a mix of sounds from anthropogenic sources (including a small amount of audible road traffic noise) and natural sounds.
- Urban $L_{Aeq,T}$ = 53dB: This environment reflects a more built-up setting, typical of densely populated areas in towns or urban locations, where road traffic becomes the dominant sound.

Finally, the ambisonic auralisations were converted to binaural using the Google Resonance VST plugin, which convolves the decoded ambisonic signals with HRTFs to allow for spatialised presentation over headphones. The final binaural renders were calibrated using a GRAS 43AA-S2 CCP artificial ear to the overall $L_{Aeq,T}$ of the combined drone and soundscape audio.

Animated visualisations were created to accompany the auralisations, indicating the type of ambient soundscape environment and the position of the drone in the sky. The visualisations of the three ambient soundscape environments are shown in Figure 6.

2.2.3 Validation

Figure 7 presents a comparison of the power spectral density (PSD) of an overflight event recorded during the field measurements and an auralisation at the same speed and altitude using the methods detailed in Section 2.2.2. The PSD has been calculated over a 45-s overflight event, with the peak sound pressure level centred at 22.5-s. It can be seen that the synthesised overflight faithfully captures the BPF and its harmonics, the dominant motor tone, and the broadband component.

Since the take-off auralisations were based on measured signals, all acoustic features present in the spectrogram shown in Figure 3 have been captured in the auralisation. The assumption of point source propagation in the geometric spreading correction has been verified from the field measurements to ensure that the relative level of the auralisation at different distances is valid. The measured propagation loss with distance from the $L_{AF,max}$ of a take-off event was −15.8dB between the 10 m and 60 m measurement positions. The theoretical point source propagation loss due to geometric spreading is calculated as $20lg\left(\frac{10}{60}\right) = -15.6$dB demonstrating that the assumption of point source propagation is considered valid.

In addition to objective validation, a series of expert listening sessions were conducted as a perceptual check to ensure that the auralisations did not sound overtly artificial or unnatural in comparison to the original field recordings. Three expert listeners with experience in acoustic modelling and auralisation compared the synthesised signals to field recordings in Arup's SoundLab®. The auralisations were judged to be sufficiently realistic to be interpreted by listeners as authentic drone sounds.

2.3 Stimuli

Binaural auralisations were generated for overflights passing directly overhead at altitudes of 60, 90, and 120 m, and for take-off operations at distances of 30, 60, and 120 m using the methods described in Section 2.2.2. The resulting auralisations were mixed with three different ambient soundscapes representing remote rural, rural village, and urban environments as described in Section 2.2.2. All stimuli were 45-s in length. This duration was selected as it allowed the approach and departure of the drone to become inaudible against the ambient background for all auralised altitudes.

2.4 Experimental procedure

A single stimulus magnitude rating method was employed, where participants rated their annoyance for individual stimuli on a 7-point numerical scale with end points "Not at all" and "Extremely". The scale end points and phrasing of the question mirrored that of ISO/TS 15666:2021 (ISO/TS 15666, 2021): *"To*

FIGURE 6
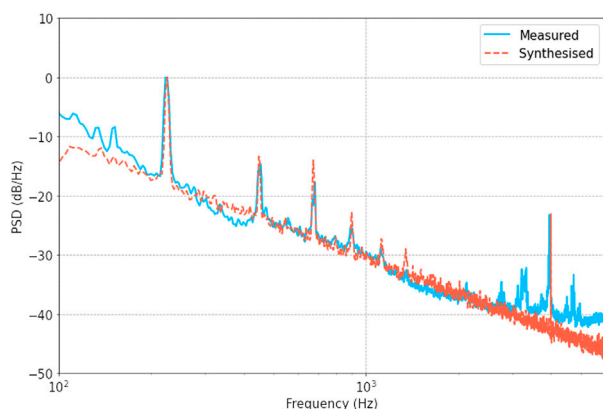Visualisations of the three ambient soundscape environments.

FIGURE 7
Comparison of the power spectral density of a measured and synthesised overflight signal.

what extent are you personally bothered, annoyed or disturbed by the sound of the drone?"

The experiment was hosted on the CommonPlace platform. Participants could choose to undertake the overflight experiment, the take-off experiment, or both. Links to both experiments were available on the CommonPlace landing page and the order in which the experiment links appeared was reversed halfway through the data collection period. The experiment launched with overflight experiment as the first link, which may explain the higher number of participants for that condition.

Stimuli were presented as embedded videos, as shown in Figure 8. The volume control for the videos was disabled to minimise the risk of participants adjusting the level of the stimuli during the experiment. Before beginning the experiment, participants were guided through a setup phase via a video instructing them to minimise any sources of noise in their environment, ensure they were wearing their headphones correctly, and to adjust the volume of their device so that a recording of speech was at a comfortable conversation level. Participants were asked not to adjust the level of their device following this setup phase.

Following this setup phase, participants were asked to rate the annoyance of each stimulus using the 7-point rating scale shown in Figure 8. Due to limitations of the delivery platform, it was not possible to randomise the stimulus presentation order for each participant. However, to minimise potential biases introduced by the fixed presentation order, all stimuli and response scales were accessible to the participant from the same screen (by scrolling up or down the webpage) to allow them to compare the stimuli before providing a response. This approach has been shown to provide similar results to a single stimulus rating method with randomised stimuli (Parizet et al., 2005).

Participants in the experimental condition where contextual information was provided entered the experiment via a portal containing information about the CAELUS project and the medical delivery use of the drone. Participants were introduced to the survey in the following way: *"Welcome to the CAELUS sound demonstrations. This area provides you with an opportunity to listen to recordings of drones that could be used to provide healthcare services, and to provide feedback on what you think of these sounds, when heard with different background ambient soundscapes that are typical across Scotland. Your feedback on these sounds is valued, and the data we collect will be used to assess the likely impact of the sound from drones on different communities. Please consider, when giving your answers, that these drones would be used to provide healthcare services."*

Participants in the no-context experimental condition were routed to the experiment directly, without accessing the information portal. All branding related to the CAELUS project, including the website URL, and mention of the use of the drone for medical delivery purposes was removed for this group. All other experimental parameters and procedures remained consistent between the two groups.

Demographic information including the age and gender of the participants was collected at the end of the experiment.

2.5 Participants

Participants for the experiment were recruited through the CommonPlace platform via email and social media. Participants provided informed consent through the platform and were able to drop out of the study at any time and without consequence. Participants were not paid for their involvement in the study. In total, 703 participants completed the experiments. The demographic breakdown of the sample in terms of age and gender is shown in Figures 9, 10.

3 Results

The statistical analysis of the data collected in the listening experiments was performed using R version 4.4.2. To allow for direct

FIGURE 8
Online experiment interface.

comparisons with studies using the 11-point ISO numerical scale, the data from the 7-point numerical scale used in the study have been linearly rescaled to a range of 0–10 for the analysis presented in this section.

A repeated measures ANOVA was initially conducted to investigate the effects of listener-drone distance (altitude for the overflight experiment and distance from FATO for the take-off experiment), soundscape, and context on annoyance ratings.

Listener-drone distance and soundscape were treated as within-participants factors, while context was treated as a between-participants factor. To assess the suitability of parametric analysis, a Shapiro-Wilk test was performed on the residuals of the ANOVA model. The test indicated a significant deviation from normality ($W = 0.95$, $p < .001$), violating the assumption of normally distributed residuals. Consequently, a non-parametric approach using the aligned rank transform (ART) was employed

FIGURE 9
Age distribution of sample.

FIGURE 10
Gender distribution of sample.
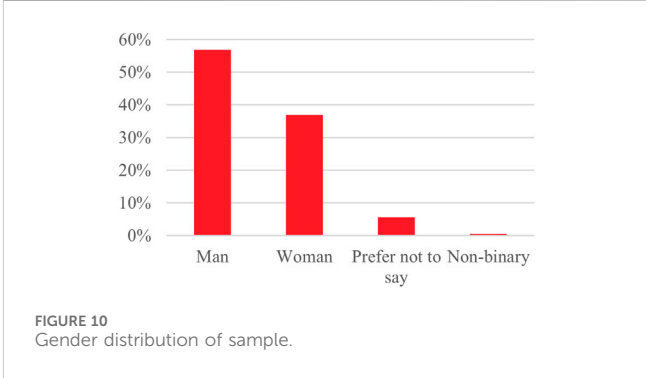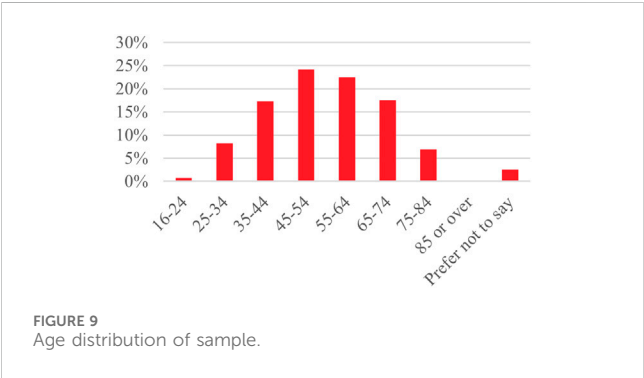
for all subsequent analyses using the ARTool package (Kay and Wobbrock, 2016). ART allows for the use of standard ANOVA procedures while accommodating non-normal data and preserving the ability to test interaction effects. Post-hoc comparisons were conducted using aligned rank contrasts with Tukey-adjusted p-values. Effect sizes were estimated using Cliff's Delta (δ), interpreted according to standard conventions (negligible < 0.147, small < 0.28, medium < 0.474, otherwise large) (Romano et al., 2006).

Although efforts were made to standardise the presentation level of the stimuli between participants by including a setup phase (see Section 2.4), the absolute level of the stimuli will vary depending on the participant's listening setup. Considering this, it is not possible to relate the annoyance ratings to absolute sound levels. However, it is valid to quantify the stimuli in terms of relative differences. Lotinga et al. (2025) used the difference between the drone and ambient L_{Aeq} to investigate the effect of the ambient environment on annoyance. This approach has been explored in the current analysis via a supplementary linear mixed-effects model (LMM), fitted using the difference between drone and ambient sound levels as a continuous predictor. An LLM was used for this analysis to allow for the continuous variable. This additional analysis allowed for the examination of whether this level difference could account for annoyance ratings across both overflight and take-off operations and different soundscape conditions.

3.1 Overflight operations

An ART ANOVA was conducted to examine the effects of altitude, soundscape, and context on annoyance for overflight operations. Altitude and soundscape were treated as within-participants factors, while context was treated as a between-participants factor and participant ID as a random factor. The analysis revealed significant main effects of soundscape, $F(2, 3352) = 378.46$, $p < .001$, altitude, $F(2, 3352) = 51.24$, $p < .001$, and context, $F(1, 419) = 59.38$, $p < .001$. Post-hoc aligned rank contrasts revealed significant differences in mean annoyance ratings between drone altitudes of 60 m and 120 m ($\delta = 0.12$, negligible effect size) and 90 m and 120 m ($\delta = 0.11$, negligible effect size). Furthermore, mean annoyance ratings differed significantly between the rural and village soundscapes ($\delta = 0.20$, small effect size), rural and urban soundscapes ($\delta = 0.37$, medium effect size), and village and urban soundscapes ($\delta = 0.18$, small effect size). Finally, significant differences were found between the context and no-context conditions ($\delta = 0.39$, medium effect size). Annoyance was observed to decrease with altitude. For the same altitude, annoyance was higher in the rural environment, followed by the village environment, then urban environment. Finally, for the same altitude annoyance was lower in the medical context scenario than the no-context scenario.

There were also significant two-way interactions between soundscape and altitude, $F(4, 3352) = 3.59$, $p = .006$, soundscape and context, $F(2, 3352) = 43.33$, $p < .001$, and altitude and context, $F(2, 3352) = 8.32$, $p < .001$. These interaction effects indicate that the impact of drone altitude on annoyance varies depending on contextual information and the type of soundscape in which the drone is operating. Annoyance decreased more gradually with

increasing altitude in the rural environment compared to the urban and village settings. Similarly, the reduction in annoyance with altitude was more gradual in the medical context than in the no-context condition. The three-way interaction between soundscape, altitude, and context was not significant, $F(4, 3352) = 1.31$, $p = .263$.

These results are illustrated in Figure 11, showing the mean annoyance ratings for the different experimental conditions alongside their 95% confidence intervals (calculated as 1.96 * standard error).

3.2 Take-off operations

An ART ANOVA was conducted to examine the effects of distance from the FATO, soundscape, and context on annoyance for take-off operations. Distance and soundscape were treated as within-participants factors, while context was treated as a between-participants factor and participant ID as a random factor. The analysis revealed significant main effects of soundscape, $F(2, 2208) = 293.35$, $p < .001$, distance, $F(2, 2208) = 273.51$, $p < .001$, and context, $F(1, 276) = 32.42$, $p < .001$. Post-hoc aligned rank contrasts revealed significant differences in mean annoyance ratings between distances of 30 m and 60 m ($\delta = 0.09$, negligible effect size), 30 m and 120 m ($\delta = 0.30$, small effect size), and 60 m and 120 m ($\delta = 0.23$, small effect size). Furthermore, mean annoyance ratings differed significantly between the rural and village soundscapes ($\delta = 0.09$, negligible effect size), rural and urban soundscapes ($\delta = 0.27$, small effect size), and village and urban soundscapes ($\delta = 0.18$, small effect size). Finally, significant differences were found between the context and no-context conditions ($\delta = 0.31$, small effect size). Annoyance was observed to decrease with distance. For the same distance, annoyance was higher in the rural environment, followed by the village environment, then urban environment. Finally, for the same distance annoyance was lower in the medical context scenario.

Significant two-way interactions were found between Soundscape and Distance, $F(4, 2208) = 25.07$, $p < .001$, Soundscape and Context, $F(2, 2208) = 33.37$, $p < .001$, and Distance and Context, $F(2, 2208) = 58.76$, $p < .001$. These interaction effects indicate that the impact of distance from the FATO on annoyance varies depending on the contextual information provided and the type of soundscape experienced. Annoyance decreased more gradually with increasing distance in the rural environment compared to the urban and village settings. Similarly, the reduction in annoyance with distance was more gradual in the medical context than in the no-context condition. The three-way interaction between soundscape, distance, and context was not significant, $F(4, 2208) = 1.80$, $p = .126$.

These results are illustrated in Figure 12, showing the mean annoyance ratings for the different experimental conditions alongside their 95% confidence intervals (calculated as 1.96 * standard error).

3.3 Level difference

As discussed in Section 3, although it is not possible to relate the results to absolute sound levels due to the
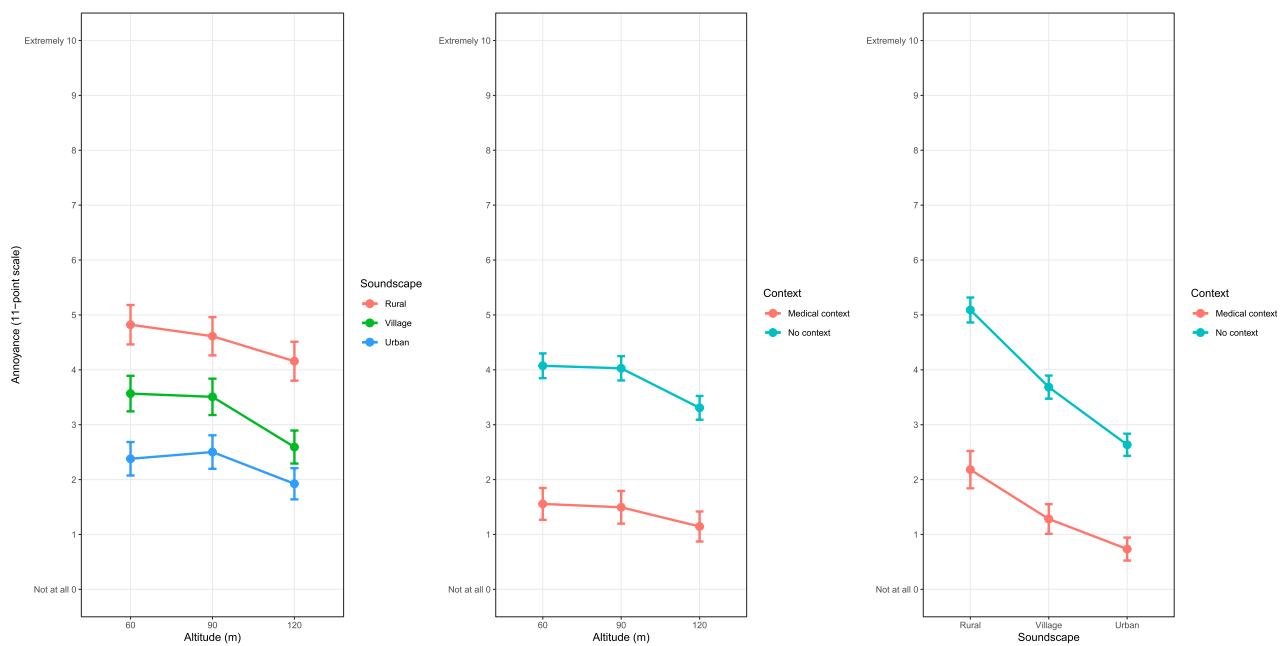
FIGURE 11
Mean annoyance ratings and 95% confidence intervals for overflight operations for different drone altitudes, soundscape types, and contexts.
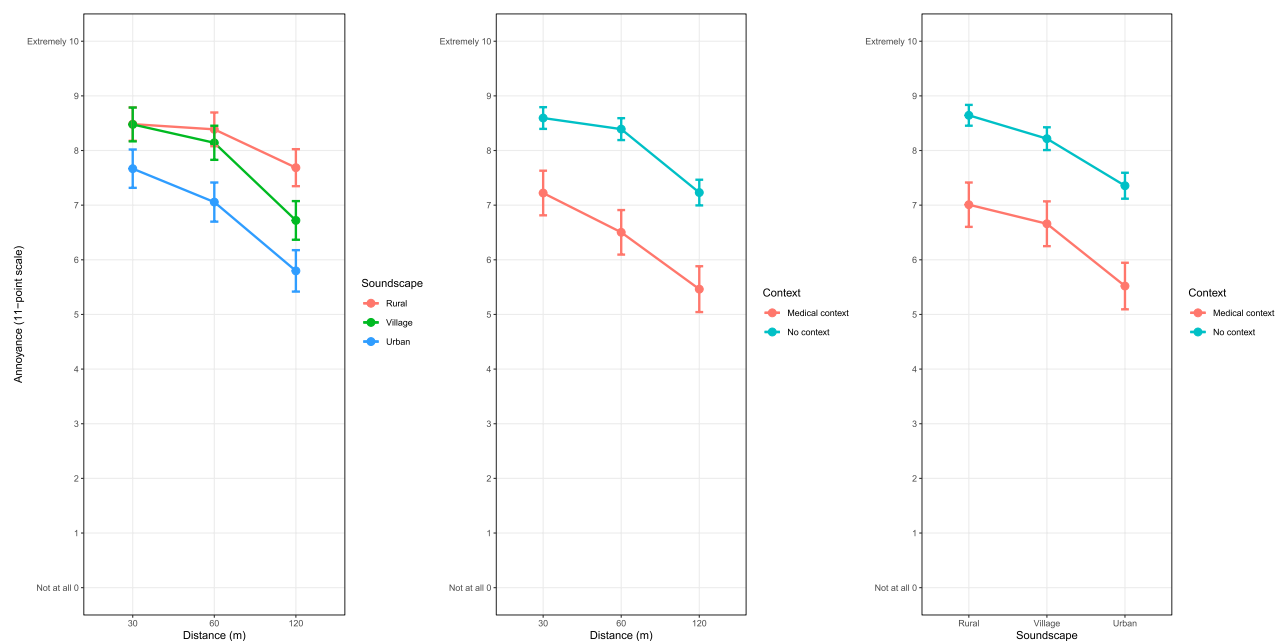

FIGURE 12
Mean annoyance ratings and 95% confidence intervals for take-off operations for different drone altitudes, soundscape types, and contexts.

uncontrolled presentation levels introduced by online testing, it is possible to quantify the stimuli in terms of relative differences. Using the approach taken by Lotinga et al. (2025), the difference between the drone and ambient L_{Aeq} over the full 45-s stimulus (ΔL_{Aeq}) was calculated for each experimental condition.

As ΔL_{Aeq} will take into account the differences in sound level between the overflight and take-off operations, the datasets for the two experiments have been combined. Figure 13 shows the mean annoyance ratings for all stimuli as a function of ΔL_{Aeq}. Annoyance appears to increase linearly up to a ΔL_{Aeq} of approximately 15 dB, beyond which the relationship begins to plateau.
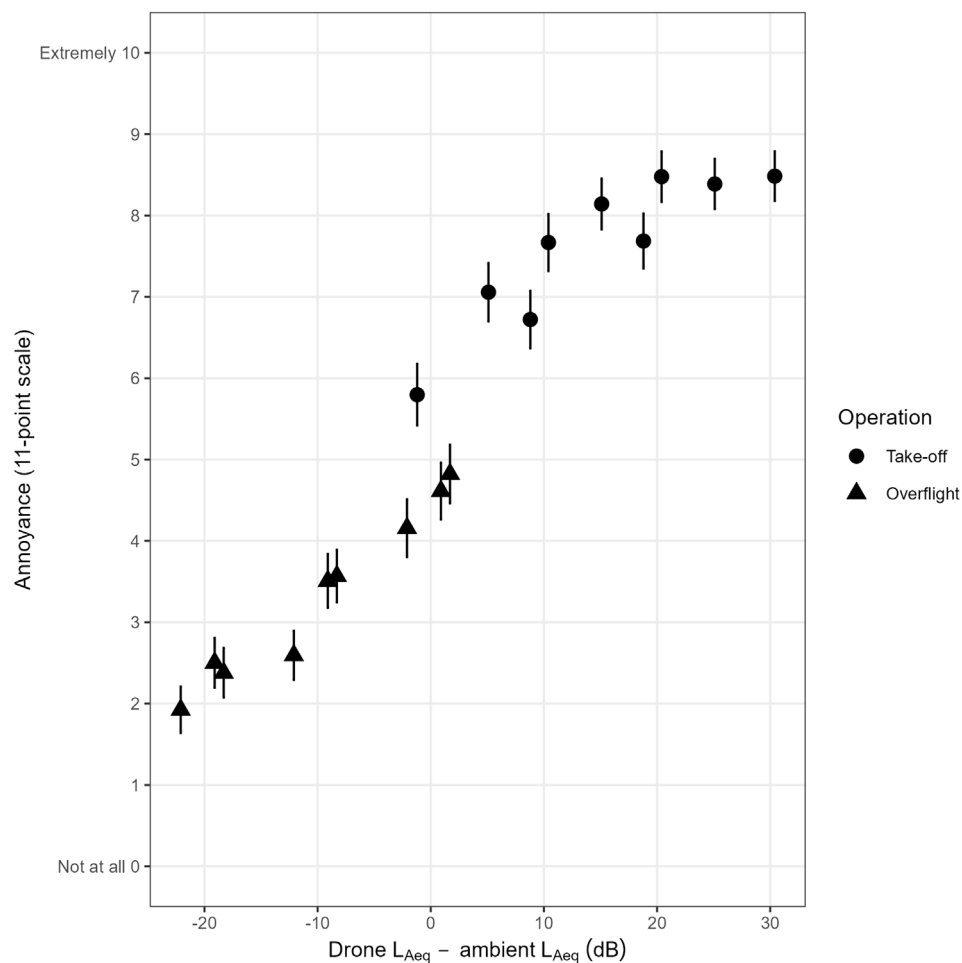
FIGURE 13
Mean annoyance ratings and 95% confidence intervals as a function of difference between the drone and ambient L_{Aeq}.

A linear mixed-effects model was fitted to the data using the `lmer` function from the `lme4` package in R (Bates et al., 2015). The model included ΔL_{Aeq}, operation type, and context as fixed effects, along with their two-way interactions. Participant ID was included as a random intercept to account for repeated measures within individuals. The model was fitted using the `bobyqa` optimiser, which is well-suited for complex random effects structures due to its robustness in achieving convergence. Parameters were estimated using restricted maximum likelihood (REML), which provides less biased estimates of variance components and is commonly used in mixed-effects modelling (Rameez et al., 2022).

The results of the model are presented in Table 1. The marginal R^2 of the model is 0.38, indicating that the fixed effects explain 38% of the variance in annoyance ratings. The conditional R^2 is 0.86, suggesting that the full model, including both fixed and random effects, accounts for 86% of the variance.

All main effects and two-way interactions were statistically significant predictors of annoyance. ΔL_{Aeq} had a positive effect, indicating that greater differences in sound level were associated with increased annoyance. Overflight operations were associated with lower annoyance while the no-context scenario was associated with higher annoyance. The interaction terms indicate that the effect

of ΔL_{Aeq} on annoyance varied depending on both context and operation type.

4 Discussion and conclusion

The study revealed that the listener-drone distance, ambient soundscape type, and contextual information all have a statistically significant effect on annoyance. Contextual information was found to have the largest effect, highlighting the importance of considering non-acoustical factors in the assessment of drone noise.

Annoyance was found to decrease with increasing distance from the drone. This relationship is expected, as listener-drone distance affects several acoustic parameters known to influence annoyance. As discussed in Section 2.2.3, field measurements confirmed that the drone behaves approximately as a point source, implying a 6 dB reduction in sound level for each doubling of distance. In addition, drones at lower altitudes exhibit less atmospheric absorption of high-frequency content. Both of these factors increase perceived loudness and sharpness, which have been found to influence annoyance. Considering this, the relationship between annoyance and distance from the drone is not surprising, but rather reflects well

TABLE 1 Results of the linear mixed effects model.

Predictor	Estimate	Std. Error	df	t-value	p-value
(Intercept)	5.352	0.321	765.8	16.678	< 0.001
ΔL_{Aeq}	0.0707	0.0047	5589	14.931	< 0.001
Context (No context)	1.429	0.376	751.2	3.797	< 0.001
Operation (Overflight)	−2.943	0.452	779.7	−6.517	< 0.001
ΔL_{Aeq}:Context	0.0167	0.0050	5589	3.353	< 0.001
ΔL_{Aeq}:Operation	0.0319	0.0042	5589	7.550	< 0.001
Context:Operation	1.139	0.518	779.7	2.198	< 0.05

established relationships between the characteristics of drone noise and annoyance.

Consistent with the findings of Torija et al. (2020) and Aalmoes et al. (2023), annoyance was significantly higher in the rural setting compared to the village and urban settings. This finding could be explained by both acoustic and non-acoustic factors. The higher annoyance observed in rural environments, followed by village and urban settings, is likely due to the greater level difference between the drone and background sound levels. This interpretation is supported by the analysis in Section 3.3, which showed that the difference between drone and ambient L_{Aeq} was a significant predictor of annoyance. Non-acoustic factors may also play a role. For example, in rural settings, the presence of drone noise may be perceived as more incongruent or intrusive in a soundscape dominated by natural sounds, potentially increasing annoyance. In contrast, in urban environments, drone noise may be more readily accepted as part of a complex urban acoustic environment already dominated by transportation noise sources. This suggests that understanding the existing soundscape and strategically planning drone routes, including take-off and landing locations, could help enhance public acceptance and minimise annoyance. For instance, avoiding flights over populated rural areas or scheduling operations during times when ambient noise levels are higher could be effective strategies.

Beyond the main effects, several interaction effects were observed that offer further insight into how contextual factors influence annoyance responses. The interaction between soundscape and altitude (for overflights) and between soundscape and distance (for take-offs) likely reflects the influence of ambient masking and the noticeability of the drone within the given soundscape. Annoyance decreased more rapidly with increasing altitude or distance in the village and urban soundscapes. This likely reflects greater masking by background noise in these environments at higher altitudes or distances, which reduced the noticability of the drone and hence annoyance. In contrast, in the rural soundscape, drone sounds were clearly audible across all altitudes and distances. This is consistent with the findings of Tracy et al. (2024), who observed that a masking discount effect exists for some individuals and can improve annoyance prediction models.

When information on the context of use of the drone was provided, there was a significant reduction in annoyance for the medical delivery use case compared to when no contextual information was provided. This is consistent with the findings of Krishnamurthy et al. (2023), where a significant effect of providing a contextual cue was found. However, a recent study by Kawai et al. (2024) found that providing the contextual cue of medical delivery did not have a significant effect on annoyance. The strong effect of context on annoyance found in the present study may be due to the level of information provided, as participants had access to a full information portal about the project prior to the experiment. This result indicates that socially valuable applications are less likely to be annoying and therefore more likely to be acceptable to individuals and communities, provided that the context of their operations is understood. This highlights the need for effective community engagement when planning drone networks. Additionally, visual cues such as the universally recognised red cross symbol could be prominently displayed on medical delivery drones to communicate their purpose.

Data availability statement

The datasets presented in this article are not readily available because The data is commercially sensitive. Requests to access the datasets should be directed to james-s.woodcock@arup.com.

Ethics statement

Ethical approval was not required for the studies involving humans because The study formed a part of the wider CAELUS project, within which a Data Impact Assessment was undertaken under the supervision of a Data Protection Officer. In our data analysis process, we ensured the highest standards of data privacy by anonymising and removing all personal data from the dataset before any analysis was conducted. A formal data privacy agreement was established between CommonPlace (the data owner) and Arup (the data processor) to maintain confidentiality and eliminate any ethical or other biases. This approach guarantees that individual identities are protected and that the data used is free from any personally identifiable information, thereby maintaining confidentiality and compliance with privacy regulations. Additionally: 1. Participants were provided with comprehensive details of the study prior to commencement 2. Informed consent was obtained from participants via the CommonPlace platform 3. Participants were given the option to withdraw from the study at any time without any consequences 4. Participants responses were anonymised, and no identifying personal data was stored. The studies were conducted in accordance with the local legislation and institutional requirements.

The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Acknowledgments

Conflict of interest

Authors JW, AT, DH, AM, LM, and CS were employed by Arup. Author FS was employed by AGS Airports.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aalmoes, R., de Bruijn, B., and Sieben, N. (2023). The influence of contextual non-auditory factors on drone sound perception. *INTER-NOISE NOISE-CON Congr. Conf. Proc. Inst. Noise Control Eng.* 266, 773–779. doi:10.3397/nc_2023_01_1105

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi:10.18637/jss.v067.i01

CAA (2024). New trials set to help unlock drone deliveries and inspections in the UK. Available online at: https://www.caa.co.uk/newsroom/news/new-trials-set-to-help-unlock-drone-deliveries-and-inspections-in-the-uk/(Accessed March 17, 2025).

Choi-Fitzpatrick, A., Chavarria, D., Cychosz, E., Dingens, J. P., Duffey, M., Koebel, K., et al. (2016). Up in the air: a global estimate of non-violent drone use 2009-2015.

Christian, A. W., and Cabell, R. (2017). "Initial investigation into the psychoacoustic properties of small unmanned aerial system noise," in *23rd AIAA/CEAS aeroacoustics conference*, 4051.

DronePrep (2021). NHS trials isles of scilly emergency drone deliveries. Available online at: https://droneprep.uk/news/nhs-trials-isles-of-scilly-emergency-drone-deliveries/(Accessed August 11, 2025).

EASA (2021). Study on the societal acceptance of urban air mobility in Europe.

Green, N., Torija, A. J., and Ramos-Romero, C. (2024). Perception of noise from unmanned aircraft systems: efficacy of metrics for indoor and outdoor listener positions. *J. Acoust. Soc. Am.* 155, 915–929. doi:10.1121/10.0024522

Guski, R. (1999). Personal and social variables as co-determinants of noise annoyance. *Noise health* 1, 45–56.

Gwak, D. Y., Han, D., and Lee, S. (2020). Sound quality factors influencing annoyance from hovering uav. *J. sound Vib.* 489, 115651. doi:10.1016/j.jsv.2020.115651

Heutschi, K., Ott, B., Nussbaumer, T., and Wellig, P. (2020). Synthesis of real world drone signals based on lab recordings. *Acta Acust.* 4, 24. doi:10.1051/aacus/2020023

Hiller, D., Arup, U., Maldonado, A. L., Jurdic, V., Harris, H., Heggie, C., et al. (2021). A 10 year retrospective and look ahead at the use of sound and visualisations for stakeholder engagement and understanding perceptions of disruptive technologies. *Present. A. T. 47th Eur. Rotorcr. Forum* 7, 9.

ISO/TS 15666 (2021). Assessment of noise annoyance by means of social and socio-acoustic surveys. Iso/Ts 15666.

Karim, N., Jing, L., Lee, J. A., Kharel, R., Lubetkin, D., Clancy, C. M., et al. (2021). Lessons learned from Rwanda: innovative strategies for prevention and containment of COVID-19. *Ann. Glob. Health* 87, 23. doi:10.5334/aogh.3172

Kawai, C., Jäggi, J., Georgiou, F., Meister, J., Pieren, R., and Schäffer, B. (2024). Short-term noise annoyance towards drones and other transportation noise sources: a laboratory study. *J. Acoust. Soc. Am.* 156, 2578–2595. doi:10.1121/10.0032386

Kay, M., Elkin, L., Higgins, J. J., and Wobbrock, J. O. (2025). ARTool: aligned rank transform for nonparametric factorial ANOVAs. R package version 0.11.2. Available online at: https://cran.r-project.org/web/packages/ARTool/ARTool.pdf (Accessed August 20, 2025).

Krishnamurthy, S., Rizzi, S., Biziorek, R., Czech, J., Berg, J., Tannler, D., et al. (2023). Remotely administered psychoacoustic test for suas noise to gauge feasibility of remote uam noise study. *INTER-NOISE NOISE-CON Congr. Conf. Proc. Inst. Noise Control Eng.* 266, 780–793. doi:10.3397/nc_2023_01_1106

Lotinga, M. J., Green, M. C., and Torija, A. J. (2025). Human perception and response to sound from unmanned aircraft systems within ambient acoustic environments. *npj Acoust.* 1, 2. doi:10.1038/s44384-024-00001-6

Lotinga, M. J., Ramos-Romero, C., Green, N., and Torija, A. J. (2023). Noise from unconventional aircraft: a review of current measurement techniques, psychoacoustics, metrics and regulation. *Curr. Pollut. Rep.* 9, 724–745. doi:10.1007/s40726-023-00285-4

Miedema, H. M., and Vos, H. (1999). Demographic and attitudinal factors that modify annoyance from transportation noise. *J. Acoust. Soc. Am.* 105, 3336–3344. doi:10.1121/1.424662

Miedema, H. M., and Oudshoorn, C. G. (2001). Annoyance from transportation noise: relationships with exposure metrics dnl and denl and their confidence intervals. *Environ. health Perspect.* 109, 409–416. doi:10.2307/3454901

Musanabaganwa, C., Semakula, M., Mazarati, J., et al. (2020). Use of technologies in Covid-19 containment in Rwanda. *Rwanda Public Health Bull.* 2, 7–12.

Parizet, E., Hamzaoui, N., and Sabatie, G. (2005). Comparison of some listening test methods: a case study. *Acta Acustica united Acustica* 91, 356–364.

PwC (2023). Advanced air mobility: uk economic impact study.

Rameez, R., Jahageerdar, S., Jayaraman, J., Chanu, T. I., Bangera, R., and Gilmour, A. (2022). Evaluation of alternative methods for estimating the precision of reml-based estimates of variance components and heritability. *Heredity* 128, 197–208. doi:10.1038/s41437-022-00509-1

Romano, J., Kromrey, J. D., Coraggio, J., and Skowronek, J. (2006). Appropriate statistics for ordinal level data: should we really be using t-test and cohen'sd for evaluating group differences on the nsse and other surveys. *Annu. Meet. Fla. Assoc. Institutional Res.* 177.

Schäffer, B., Pieren, R., Heutschi, K., Wunderli, J. M., and Becker, S. (2021). Drone noise emission characteristics and noise effects on Humans—A systematic review. *Int. J. Environ. Res. public health* 18, 5940. doi:10.3390/ijerph18115940

Scott, J., and Scott, C. (2017). "Drone delivery models for healthcare," in *50th Hawaii international conference on system sciences.*

Swiss Foundation for Mine Action (FSD) (2014). Using drones for medical payload delivery in Papua New Guinea.

Sylverken, A., Owusu, M., Agbavor, B., Kwarteng, A., Ayisi-Boateng, N. K., Ofori, P., et al. (2022). Using drones to transport suspected COVID-19 samples; experiences from the second largest testing centre in Ghana, West Africa. *PLoS ONE* 17, e0277057. doi:10.1371/journal.pone.0277057

Torija, A. J., Li, Z., and Self, R. H. (2020). Effects of a hovering unmanned aerial vehicle on urban soundscapes perception. *Transp. Res. Part D Transp. Environ.* 78, 102195. doi:10.1016/j.trd.2019.11.024

Tracy, T., Boucher, M., Christian, A., Krishnamurthy, S., Rizzi, S., Begault, D., et al. (2024). An annoyance model for urban air mobility vehicle noise in the presence of a masker. *INTER-NOISE NOISE-CON Congr. Conf. Proc. Inst. Noise Control Eng.* 269, 440–453. doi:10.3397/nc_2024_0053

TU Delft (2020). Ambulance drone. Available online at: https://www.tudelft.nl/io/onderzoek/research-labs/applied-labs/ambulance-drone (Accessed August 11, 2025).

UPS (2021). UPS drone COVID vaccine deliveries. Available online at: https://about.ups.com/ca/en/our-stories/innovation-driven/drone-covid-vaccine-deliveries.html (Accessed August 11, 2025).

Zipline and Pfizer (2021). End to end testing C19 through uas aircraft mode of transport. Available online at: https://assets.ctfassets.net/pbn2i2zbvp41/2zoeIUgBrPcYeRT3R2S4pK/4d65242647ccfac6ea9341644fa1f241/Zipline-Pfizer_Position_Paper.pdf (Accessed August 11, 2025).