Applied Acoustics 254 (2026) 111474

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Joint sound source localisation of drones and other background sources

Anique Altena [a,*], Mirjam Snellen [a], Salil Luesutthiviboon [a], Guido de Croon [a], Mark Voskuijl [b]

[a] Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, Delft, 2629HS, Zuid-Holland, the Netherlands
[b] Faculty of Military Sciences, Netherlands Defence Academy, Het Nieuwe Diep 8, Den Helder, 1781AC, Noord-Holland, the Netherlands

HIGHLIGHTS

- We propose an effective method to localise multiple drones and other background sources using microphone array measurements.
- Frequency-domain beamforming is applied using single frequencies within a frequency band of 140 to 1200 Hz.
- The proposed method is validated on real-life drone measurements including background sources such as aircraft flyovers.
- Quadcopters could be localised up to 712 m in azimuth and 593 m in elevation direction.
- Fixed-wing drones could be localised up to 317 m in azimuth and 305 m in elevation direction.

ARTICLE INFO

Keywords:
Drones
Sound source localisation
Microphone array
Beamforming

ABSTRACT

It remains a challenge to acoustically localise drones, especially when there are multiple sound sources in the environment. The aim of this paper is to propose and validate an approach for joint acoustic localisation of drones and other background sources using microphone array measurements. This approach consists of a frequency-domain beamforming step, where a discrete number of frequencies in the range of 140 to 1200 Hz are beamformed individually. The location of the maximum output is identified and plotted in the azimuth and elevation direction per frequency and time step. The next step is to extract and separate the trajectories from this plot.

The validation shows the correctness of this approach, together with the remaining challenges. Although sources are extracted and separated correctly, overlapping sources are sometimes considered as only one source. Moreover, in low-level drone noise or high-level background noise measurements, erroneous sound sources are identified. Nevertheless, the results show that drones can be localised over long distances, up to 712 m and 593 m for a quadcopter drone in azimuth and elevation direction respectively, and 317 m and 305 m for a fixed-wing drone in azimuth and elevation direction respectively.

1. Introduction

A challenging aspect in drone localisation using acoustics is the detection and separation of different sound sources, including multiple drones and background sources [1]. This is because various sources transmit sound at similar frequencies, despite the fact that each source produces its own distinct frequency spectrum. This makes it difficult to discern sound sources in a spectrogram, thereby complicating the process of identifying which frequencies to use for localisation of a specific source.

Existing literature confirms that distinguishing between different drones, or between a drone and another source in the environment, remains challenging. Only a few studies on drone localisation have addressed the topic of multiple drones and background sources.

An early study from 2015 uses a spherical microphone array for drone localisation [2]. The authors investigated single drone scenarios, with three different drones. They speculated that the method could be employed in noisy environments with multiple drones; however, this still needs to be verified. One study applied time-domain delay-and-sum beamforming in combination with a peak finding algorithm to extract harmonics of tonal drone noise [3]. The algorithm successfully distinguished two sources in a simulation environment. A recent study uses phase-shifted signals obtained by beamforming as input for an extensive neural network with an attention module to predict the azimuthal

* Corresponding author.
Email address: A.Altena@tudelft.nl (A. Altena).

https://doi.org/10.1016/j.apacoust.2026.111474
Received 20 October 2025; Received in revised form 15 January 2026; Accepted 12 July 2026
Available online 22 July 2026

direction of two identical drones [4]. The approach was evaluated in a simulation using measured ego-noise of a drone from an open-access dataset [5]. The performance was promising; however it is unknown if the approach can be generalised for different drones and used in realistic, outdoor scenarios.

Some studies include real-world tests to validate their multi-drone localisation method. One study uses functional beamforming in combination with a local maxima detection step to obtain all local maxima in the beamformer output [6]. This method was validated using data obtained in an anechoic chamber with a maximum of four drones. Lastly, Blass et al. [7] use a combination of steered response power, Gaussian mixture model and delay-and-sum beamforming to obtain the location of a predefined number of sources. The authors tested the method on an outdoor measurement with two quadcopter drones flying almost the same path. This scenario was too challenging for the algorithm to separate the drones from each other, due to the drones' nearly identical sound.

Although these studies attempted to localise multiple drones simultaneously, no study has yet been published that not only localises multiple drones but also localises background sources in a real-life, uncontrolled scenario.

The research presented here takes on the challenge of obtaining the trajectories of multiple sound sources in a realistic, uncontrolled, outdoor environment, including one or multiple drones. The trajectories are obtained using frequency-domain beamforming, an often used robust acoustic imaging technique [8]. This method is chosen as it has shown promising results for acoustic localisation of drones in earlier studies [3,4,6,7,9–11]. The novelty compared to these studies is the use of single frequencies in the frequency band of 140 Hz to 1200 Hz for beamforming. For frequencies below 140 Hz, wind noise pollutes the measurements. The upper limit of 1200 Hz is the frequency up to which the dominant drone noise is typically measurable by the array, and attenuation by atmospheric absorption is still relatively low [12]. This beamforming step is essential to obtain trajectory information from various sources rather than just from drones. On top of this, the novelty lies not only in obtaining the trajectories but also in recovering dominant frequencies associated with these sound sources in this range of 140 Hz to 1200 Hz. This overview of dominant frequencies is interesting information for identification of noise sources.

The purpose of this manuscript is to provide and demonstrate the proposed methodology for localisation of drones and other background sources. This paper shows the results of two measurements in which sound sources can be separated successfully, and a more complex measurement in which a combination of lower noise levels of drones and high background noise levels makes the separation process more challenging.

The subsequent sections of this work are ordered as follows. Section 2 describes the measurement campaign performed to obtain the data used in this work. Then, Section 3 provides the methodology to obtain trajectories of multiple sound sources, to separate them and to recover the frequency spectrum associated with the trajectory. In Section 4, the results are summarised, which demonstrate the working of the proposed method. Lastly, Section 5 presents the conclusion.

2. Experimental setup

To test the proposed method, acoustic data of drones and background sources were measured. The measurement campaign was organised at the Dutch military base 'Luitenant-Generaal Bestkazerne'. A planar array was placed flat on the ground in an open area on a grass field, next to a runway which is out of use. An overview of the measurement location is given in Fig. 1(a).

The measurements were performed with a microphone array consisting of 64 microphones (PUI Audio 665-POM-2735P-R). The array is configured in an Underbrink distribution [13,14] and has a diameter of 3.4 meters. The spacing between the microphones varies from 0.11 to 3.4 meters. The array was placed on top of acoustic absorbing foam to minimise ground reflections. The microphones are covered by windshields to minimise the impact of wind on the measurements. The sample frequency of the microphones is 50 kHz. An image of the array is included in Fig. 1(b).

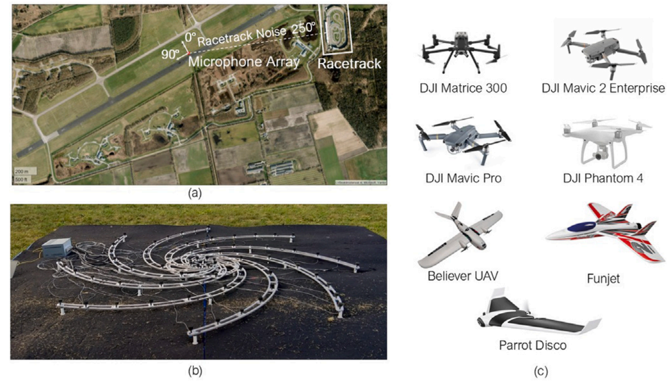

Fig. 1. Measurement information: (a) Map of the measurement location, (b) microphone array, (c) drones used during this research.

Table 1
Drones used during the measurement campaign, their take-off mass and propeller diameter.

Drone Name	Drone Type	Take-off Mass [g]	Propeller Diameter
DJI Matrice 300	Quadcopter	6300	21 inch
DJI Mavic 2 Enterprise	Quadcopter	899	8.7 inch
DJI Mavic Pro	Quadcopter	734	8.3 inch
DJI Phantom 4	Quadcopter	1380	9.4 inch
Believer UAV	Fixed Wing	4200	11 inch
Funjet	Fixed Wing	875	6 inch
Parrot Disco	Fixed Wing	750	8 inch

Throughout the campaign, a variety of drones were used, including both fixed-wing and quadcopter drones. Most drones were equipped with an external commercial GPS sensor, without an improved GPS method. This setup was beneficial for validating the proposed method. Information about the drones used during the campaign is provided in Table 1 and images of the drones are provided in Fig. 1(c).

The drones flew various scenarios on and around the military base. It should be noted that the scenarios were not designed specifically for the acoustic measurements and therefore it was an uncontrolled measurement campaign. During each measurement, multiple drones were flying, often a combination of fixed-wing and quadcopter drones. Measurements were taken with drones flying at different distances from the array. This resulted in measurements with loud and dominant drone sources in the vicinity of the array and measurements with no dominant source and silent drones flying in the far distance. Besides this, the measurements contain a variety of background noise, in particular aircraft flyovers and low frequency noise from a racetrack indicated in Fig. 1.

3. Methodology

The proposed approach consists of two steps. First, beamforming single frequencies in a predefined frequency range and obtaining the maximum per frequency and time stamp in azimuth and elevation directions. Second, separating these maxima into trajectories and linking the trajectories to their dominant frequencies. This section discusses the details of both steps.

3.1. Beamforming single frequencies

In previous research [12], we concluded that it is essential to investigate a wide range of the frequency spectrum to localise different sound sources, as different sources emit sound in different frequency bands.

The frequency spectrum was divided into 12 bands of 200 Hz width, and in each band 2 frequencies were selected for frequency-domain beamforming [15]. This beamforming method was chosen as it allows for a detailed analysis of the data, by focusing on relevant frequencies only. In this research, we use all frequencies obtained after the Fourier transform individually in a range of 140 Hz to 1200 Hz. The beamform output is computed for these frequencies f as follows [8,16]:

$$B_j(f) = \frac{\boldsymbol{g}_j^H \boldsymbol{C} \boldsymbol{g}_j}{||\boldsymbol{g}_j||^4}, \tag{1}$$

where $(\cdot)^H$ is the Hermitian transpose, or complex conjugate transpose, $\boldsymbol{g}$ is the steering vector and $\boldsymbol{C}$ is the cross-spectral matrix (CSM). The index j indicates a grid point. In this research, the grid is defined in azimuth ϕ and elevation θ directions. The azimuth angle ranges from 0° to 360° and the elevation angle ranges from 0° on the horizontal to 90° above the array. The grid steps are chosen to be 1° for both directions. These angles are defined with respect to the centre of the microphone array. A fine grid with small steps in azimuth and elevation directions is chosen such that a clear match with the continuous GPS can be made and the proposed method can be validated.

The cross-spectral matrix is defined as:

$$\boldsymbol{C} = \mathbb{E}[\boldsymbol{p}\boldsymbol{p}^H], \tag{2}$$

where $\mathbb{E}[\cdot]$ is the expectation operator and $\boldsymbol{p}$ is the pressure vector. This vector is the output of the fast Fourier transform which is used to transform the measured pressure measurements to the frequency domain. A block of 2500 samples is used for this, which is zero-padded to 8192 samples. An overlap of 50% is used. These settings ensure a good frequency resolution and do not neglect useful information.

Lastly, the steering vector is calculated for frequency f as:

$$g_{n,j}(f) = e^{-2\pi i f \frac{r_{n,j}}{c}}, \tag{3}$$

where c is the speed of sound and $r_{n,j}$ is the distance from microphone n to grid point j. To compute this distance, it is assumed that all sources are in the far field. Under this assumption, the sound waves arrive at the microphone array as a plane wave and therefore the absolute distance from the centre of the array to grid point j is not relevant in this computation. A detailed argumentation for this can be found in Ref. [12].

Beamforming is applied to each frequency in the range of 140 to 1200 Hz. The settings of the Fourier transform in this research lead to a frequency resolution of 6.1 Hz and therefore 174 frequencies that are evaluated at each time stamp. For each frequency, the grid point (ϕ, θ) with the maximum beamform output is saved. All maxima are plotted over time, which results in a point cloud where some points compose a path and some are random noise. At this stage of the research, only the global maximum is used and other local maxima are not included. This could result in non-dominant sources being overlooked. However the use of many single frequencies in a broad range can counteract this effect. It also avoids the potential for side lobes to be incorrectly identified as a source.

3.2. Separation of trajectories

The next step is to separate these points into different sources. During this process, all points are added to a source individually based on a threshold and their similarity to points in an already identified source. This is first of all done for points in the azimuth direction. If a point does not fall within the azimuth threshold of another source, a new source is created. Furthermore, if an associated elevation point exists, i.e., a point in elevation direction at the same time instant t and with the same frequency f, it is also added to this source. If the azimuth direction does not provide a clear source to which a point should be added, threshold and similarity are checked in the elevation direction. If it remains undecided, a point is added to the longest existing source. This is chosen as tests showed that it is not possible to state whether a single frequency fits within a trend of frequencies.

Lastly, it should be noted that there is no limit set on the number of sources present at a certain time step.

To summarize, points are added to a source according to the following procedure:

1. **Azimuth Threshold**: A point is added to a source if it falls within a threshold of 2° compared to the last 0.5s azimuth points of a source.
2. **Elevation Threshold**: If a point belongs to multiple sources based on the previous criterion, a point is added to a source if it falls within a threshold of 2° compared to the last 0.5s elevation points of a source.
3. **Non-exclusive Result**: If a point belongs to multiple sources based on previous criteria, a point is added to the longest source out of the possible sources it could belong to.

The azimuth and elevation thresholds are set to 2° to account for a gradual change in angles over time. Next to this, a threshold is necessary to account for deviations in the beamformer output. Different frequencies of the same source can be mapped to slightly different angles instead of a single angle, for example due to noise in the data and the finite resolution of the array. The threshold of 2° was a trade-off between these requirements and not including too many noisy points into a source.

Lastly, as a post-processing step only sources with more than 200 points in the azimuth direction are kept. Sources with fewer points are removed, as results showed that these sources often only contained noisy points.

The order of the azimuth and elevation criteria was determined based on the clearest results over the longest period. This means accurate and closely spaced location estimates, while providing these estimates as early and late as possible. In this research, data from a planar array facing upwards is considered. Rayleigh's criterion for a steered beam width [17] shows that the effective aperture of the microphone array reduces particularly for the elevation direction when a sound wave comes from low angles:

$$\theta_{B,s} = 1.22 \frac{\lambda}{L \sin(\theta_s)}. \tag{4}$$

Here, $\theta_{B,s}$ is the steered beam width, λ is the wavelength and L is the array aperture. θ_s is the steering angle, where 0° is defined on the horizontal, and 90° is defined above the array. When θ_s moves to 0°, close to the horizontal, the effective aperture $L \sin(\theta_s)$ decreases, resulting in a large beam width for the frequency range of interest. The beamform output will be less accurate and noisier in the elevation direction, resulting in incorrect estimates. The results showed that this occurred at elevation angles below 20° to 30°. Meanwhile, there was no degradation in performance in the azimuth direction yet. As a result of this, the azimuth angle will be assessed first.

The final step involves identifying the dominant frequencies of each localised source in the frequency range of 140 Hz to 1200 Hz. The azimuth and elevation points in a source can be associated with a single frequency through the beamformer step described in Section 3.1. This analysis provides a frequency profile for each source, facilitating subsequent sound source identification.

3.3. Challenges and limitations

Using the described procedure, most trajectories can be separated from each other and an associated frequency spectrum can be obtained. However, there are several use cases that pose challenges. The most prominent example is when paths overlap or when the last points of one path connect to the first points of a new one. In these scenarios, it is possible that two or more overlapping or consecutive sources will be considered as one source.

Table 2
Measurement summary. FW: Fixed-Wing, QC: Quadcopter.

Measurement	Drones Used	Background Sources
Measurement 1 *Clear*	Parrot Disco (FW) Funjet (FW) DJI Phantom 4 (QC)	Racetrack noise Aircraft Fly-Over (0 - 80 s)
Measurement 2 *Clear*	Parrot Disco (FW) Funjet (FW) DJI Mavic Pro (QC) DJI Matrice 300 (QC)	Racetrack noise Aircraft Fly-Over (150 - 250 s)
Measurement 3 *Challenging*	Believer UAV (FW) Parrot Disco (FW) DJI Mavic Pro (QC) DJI Phantom 4 (QC)	Racetrack noise Aircraft Fly-Over (150 - 250 s)

Another challenge occurs when there is no dominant source or when a loud (broadband) background source is present. This happens for example when drones are flying far away. In these scenarios, the beamform results are noisier. It often happens that a collection of noisy points falls within the prescribed criteria and forms an erroneous sound source that is not actually present.

One way to mitigate this is to add a filtering step in between the beamform step and the separation step. A possibility is to only keep points which have at least one neighbouring point in azimuth direction and time. However, this filtering also removes results aligned with an actual sound source, especially at the beginning and end of the interval in which the sound source can be detected and localised. To avoid losing valuable information, it was decided not to implement any form of filtering.

4. Results

This section shows the results of three measurements. Two measurements provide clear results and validate the proposed approach. The third measurement shows the challenges and limitations of this approach.

4.1. Measurement details

All chosen measurements have a duration of five minutes. During each measurement, a variety of drones were flown and different background noises were present. These details are summarised in Table 2. The racetrack noise is localised at an azimuth angle of 250° to 255°, as indicated in Fig. 1. In the elevation direction, this noise could not be localised as clearly. It is assumed that this is due to its relatively low elevation angle compared to the microphone array.

4.2. Measurement 1

The result of the beamforming process for measurement 1 can be found in Fig. 2, where the top plot shows the azimuth results and the bottom plot shows the elevation results. Both plots display numerous paths, even at same time stamps. This demonstrates the power of beamforming individual frequencies at each time stamp.

The next step is to extract and separate the paths and connect them in azimuth and elevation directions to the same source. The result of this process is shown in Fig. 3, in which the azimuth direction, elevation direction and distance from the array to the drone are plotted from top to bottom respectively. The black lines with various markers indicate the GPS of the drones. This figure clearly shows that paths are extracted correctly from the previous plots and that most paths can be separated. The method for extracting and separating paths of different sound sources works well.

Nevertheless, Fig. 3 also shows the challenging case of overlapping and connecting paths, starting at approximately 90 s in the azimuth direction for the purple source. At this moment in time, three sources come together: the aircraft flyover comes to an end, the racetrack noise becomes more pronounced due to the end of this flyover and a drone flyover passes these paths at the same time. Although three separate sources are present, the proposed approach considers them as one.

Lastly, the dominant frequencies of the tracked sources can be found in Fig. 4. The colours in this figure correspond to the colours in Fig. 3. The frequency profile emphasises another strength of the proposed method. By following this procedure, it is possible to get detailed frequency content of the different sources in the environment. The profile shows distinct characteristics, such as tonal sound for rotor drones, a broadband profile for the aircraft flyover and the low frequency noise from the racetrack. This information is of interest for e.g., identification of sound sources.

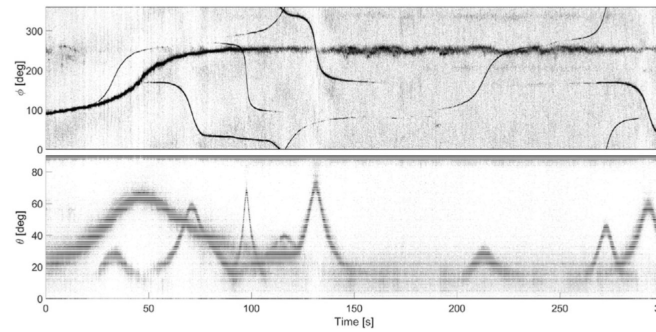

Fig. 2. Results of plotting the maxima of beamforming individual frequencies in the range of 140 to 1200 Hz at each time stamp in the azimuth (top) and elevation (bottom) directions for the first example measurement.

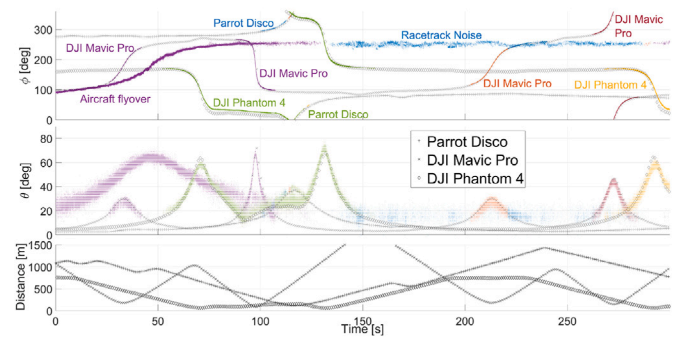

Fig. 3. Separated flight paths of the first example measurement with drones and other noise sources, including an aircraft flyover and racetrack noise. Same colours at the same time indicate the same source in azimuth and elevation directions. The GPS positions of the drones are plotted in black.

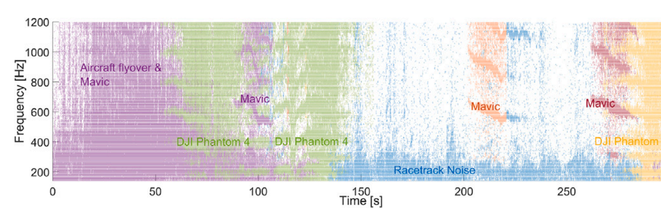

Fig. 4. The frequency spectrum of the first example measurement. The colours of the spectra correspond to the colours of the paths in Fig. 3.

4.3. Measurement 2

The results of the beamforming process of measurement 2 can be found in Fig. 5 for the azimuth direction in the top plot and for the elevation direction in the bottom plot.

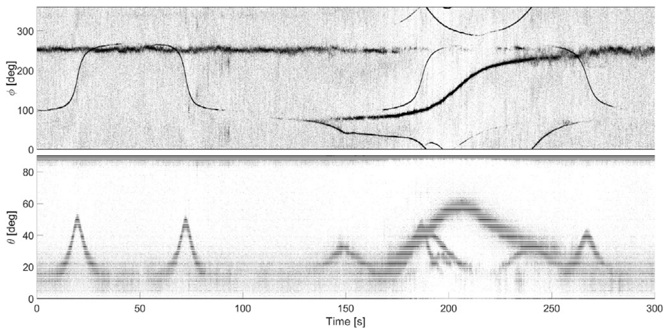

Fig. 5. Results of plotting the maxima of beamforming individual frequencies in the range of 140 to 1200 Hz at each time stamp in the azimuth (top) and elevation (bottom) directions for the second example measurement.

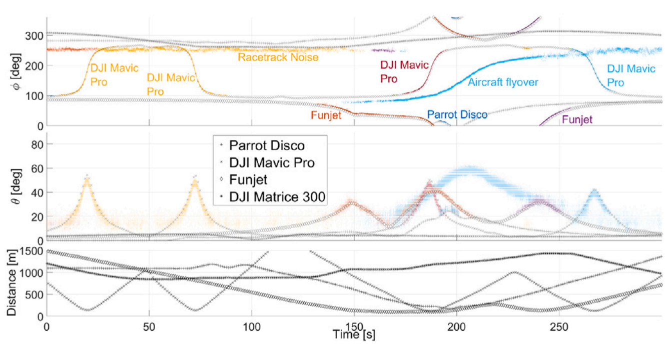

Fig. 6. Separated flight paths of the second example measurement with drones and other noise sources, including an aircraft flyover and racetrack noise. Same colours at the same time indicate the same source in azimuth and elevation direction. The GPS positions of the drones are plotted in black.

As for measurement 1, multiple paths can be identified at the same time, including the racetrack noise and an aircraft flyover between 150 and 250 s.

The results of the separation step can be found in Fig. 6. In this figure, similar challenges as in the previous measurement are visible. Some overlap is occurring, e.g., when a drone is interfering with the racetrack noise after 67 s.

4.4. Measurement 3

When drones are flying farther away, quieter drones are used, or a higher level of (broadband) noise is present, the beamforming step results in a more scattered result. The paths are less pronounced, making the separation step more challenging. To show this, measurement 3 is included, which contains a higher level of background noise that can partly be attributed to the racetrack noise.

The beamform output of measurement 3 is shown in Fig. 7 for azimuth and elevation directions in the top and bottom plots, respectively. In this figure, there are more noisy background points compared to Figs. 2 and 5.

The results of the separation process are presented in Fig. 8. Compared to the results of measurements 1 and 2 in Figs. 3 and 6 respectively, the results are noisier, particularly in the azimuth direction. Due to the more scattered output of the beamforming step, erroneous sound sources are obtained during the separation step. The majority of these erroneous sources are point clouds with a lower density compared to the actual sound sources. Therefore, an additional filtering step between the beamforming step and the separation step could reduce the number of erroneous sources.

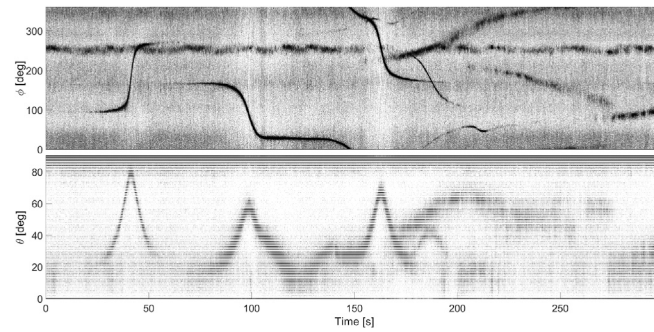

Fig. 7. Results of plotting the maxima of beamforming individual frequencies in the range of 140 to 1200 Hz at each time stamp in the azimuth (top) and elevation (bottom) directions for the second example measurement.

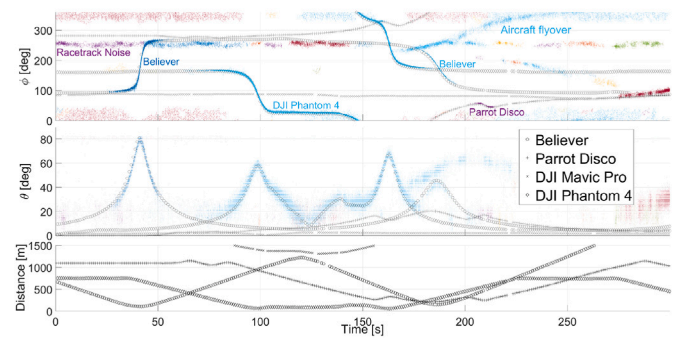

Fig. 8. Separated flight paths of the challenging example measurement with drones and other noise sources, including an aircraft flyover and racetrack noise. Same colours at the same time indicate the same source in azimuth and elevation direction. The GPS positions of the drones are plotted in black.

4.5. Localisation distance overview and comparison with literature

In addition to presenting the proposed procedure, this research also contributes to identifying realistic localisation ranges that are possible for drone localisation using acoustics. The range for which a drone can be localised depends on many factors, such as the array design, the actual drone, the weather conditions and the level of background noise. Nevertheless, the proposed localisation and separation process allows for simultaneous localisation of multiple sources at greater distances than achieved in our prior study [12].

Table 3 summarises the maximum localisation distances for all three measurements, all drones, and both azimuth and elevation directions. The distance at which a drone can be localised is defined as the distance at which the path of a source begins or ends. This definition was chosen after consulting Figs. 3, 6 and 8, which show a good alignment between the GPS of the drones and the paths of the sources. Note that for the DJI Matrice 300 in measurement 2 and the DJI Mavic Pro in measurement 3 no distance is indicated, as they flew too far away to be localised at all. They flew at distances of at least 844 m and 1314 m respectively.

This table provides two interesting insights. First, it gives an indication of the localisation distances that can be achieved in azimuth and elevation directions. The DJI Mavic Pro can be localised at an impressive distance of 712 m in azimuth direction in the second measurement under low background noise conditions. However, in elevation direction this is reduced to 582 m. This example clearly shows that when the elevation angle drops below 20° to 30°, the elevation angle becomes harder to estimate, which is as expected from Eq. (4). Another interesting result regarding the difference between azimuth and elevation directions can be seen in measurement 3. This measurement is the noisiest of the three and results in a different maximum distance for azimuth and elevation

Table 3
Overview of the maximum localisation distance per measurement and drone.

Measurement	Drone	Distance Azimuth [m]	Distance Elevation [m]
Measurement 1	Parrot Disco	278	278
	DJI Mavic Pro	593	593
	DJI Phantom 4	305	305
Measurement 2	Parrot Disco	218	218
	DJI Mavic Pro	712	582
	Funjet	246	246
	DJI Matrice 300	–	–
Measurement 3	Believer	270	254
	Parrot Disco	317	290
	DJI Mavic Pro	–	–
	DJI Phantom 4	294	258

angles for most drones. For quieter measurements 1 and 2, this occurred only once. The estimates in elevation direction are more impacted by higher levels of background noise than the estimates in azimuth direction.

Second, the maximum distances achieved per drone are different for different measurements. This can be attributed to the presence of background noise. This happens for all drones that are used more than once, which are in particular the Parrot Disco, DJI Mavic Pro and DJI Phantom 4. However, it should be noted that it is more clear when comparing the same drones within one measurement. In measurement 1, the maximum localisation distance for the DJI Mavic Pro varied from 593 m without background noise to 288 m during an aircraft flyover. In the same measurement, the maximum distance for the DJI Phantom 4 dropped from 305 m to 230 m during the aircraft flyover. The maximum distance thus drops approximately by 30 to 50% due to background noise. This emphasises the importance of real-world measurements such as those done in this study.

To get an indication on how the proposed procedure performs compared to state-of-the-art acoustic drone localisation research, Tab. 4 is included. This table summarises the achieved maximum distances of this study and 5 other relevant studies on this topic. The table includes the types of drones used and their masses. The masses are either taken from the reference or estimated based on the drone type. It should be noted that the achieved distances usually scale with the weight of the drone, i.e., the heavier the drone, the longer the distance.

The table shows that the achieved maxima vary widely among different studies. One reason for this is the use of different measurement equipment, as well as varying measurement conditions. One example of the impact of microphone array design is the study presented in Ref. [11]. This study uses a genetic algorithm to optimise the array design for drone localisation and achieves a promising localisation distance of 340 m for typical quadcopters.

Another factor which influences the maximum distance is the localisation method. The study presented in this manuscript, as well as [11,12], uses a beamforming approach. On the other hand, [18] uses a direct steered power response method to estimate the location of the drone. Lastly, Ref. [19] processes the data with a deep neural network. The study presented in Ref. [2] did not elaborate on the used localisation method. The interesting outcome when comparing these methods is that the neural network approach results in the lowest localisation distances. The beamforming approaches presented in this study and in Ref. [11] result in the highest distances.

5. Conclusion

The aim of this paper is to validate the proposed approach for acoustic localisation of drones and other sound sources in the environment. First, a beamforming step is applied, in which individual frequencies in the range of 140 to 1200 Hz are beamformed. This step results in a scatter plot in azimuth and elevation directions in which clear trajectories of different sound sources in the environment are visible. This result is as desired and it can be concluded that it is essential to beamform over a large range of frequencies to obtain a full acoustic image of the environment.

Second, a separation step is implemented, which obtains the individual trajectories from the output of the beamform step. The results show that the separation method is successful. Full trajectories are extracted from the beamform output, and azimuth and elevation points are attributed to the same source correctly. One limitation of the approach is that sources can be seen as one when they overlap, instead of separate sources. However, even though they are seen as one, the trajectories are still extracted correctly. Finally, drones can be localised over a long distance, up to 712 m for a quadcopter drone and 317 m for a fixed-wing drone during these specific measurements.

Nevertheless, it should be concluded that the more complex the measurement, the more challenging the separation process. First, increased levels of background noise reduce the range for which drones can be localised. Second, measurements which contain low-level drone noise or high-level (broadband) background noise result in a more scattered output of the beamform step. This noise can accidentally be seen as a sound source during the separation process, leading to erroneous sources.

CRediT authorship contribution statement

Anique Altena: Writing – original draft, Visualization, Software, Methodology, Investigation, Conceptualization. **Mirjam Snellen:** Writing – review & editing, Supervision. **Salil Luesutthiviboon:**

Table 4
Overview of maximum reported localisation ranges in literature on acoustic drone localisation.

Reference	Reported Distance(s)	Drone Type(s)	(Estimated) Mass
This study	- Fixed Wing 250 to 300 m - Typical Quadcopter 300 to 700 m m	- Believer, Funjet, Parrot Disco - DJI Phantom 4, DJI Mavic Pro	- 750 to 4200 gr - 734 to 1380 gr
Altena 2025 [12]	- Quadplane 360 m - Typical Quadcopter 130 m - Lightweight Quadcopter 90 m	- Avy Aera - DJI Phantom 3 & 4, DJI Mavic 3, Autel EVO II - DJI Mini 2	- 18,250 gr - 895 to 1380 gr - 242 gr
Blass 2024 [18]	- Large Multicopter 500 m - Typical Quadcopter 200 m - Lightweight Quadcopter 100 m	- twinFOLD KAT - DJI Mavic 2 Pro - DJI Mini 2 SE	- 6000 gr - 907 gr - 249 gr
Busset 2015 [2]	- Typical Quadcopter 150 to 290 m	- DJI Phantom 2, Parrot AR Drone 2.0, DJI Flamewheel F450	- 420 to 1200 gr
Hengy 2024 [19]	- Large Octocopter 180 m - Typical Quadcopter 120 m - Lightweight Quadcopter 90 m	- DJI S1000 - DJI Mavic Pro - DJI Spark	- 4000 gr - 734 gr - 300 gr
Itare 2025 [11]	- Typical Quadcopter 340 m	- DJI Phantom 4	- 1380 gr

Writing – review & editing, Investigation. **Guido de Croon:** Writing – review & editing, Supervision. **Mark Voskuijl:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Anique Altena reports that financial support was provided by Netherlands Ministry of Defence. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is part of the ACTION (ACoustic detecTION of class I (<20 kg) unmanned aircraft systems supported by optical sensors) project, funded by the Dutch Ministry of Defence under Nationaal Technologie Project (NTP) N21/005.

Data availability

Data will be made available on request.

References

[1] CWA 18150:2024. Unmanned aircraft systems - counter UAS - testing methodology. Standard European Committee for Standardization; 2024.

[2] Busset J, Perrodin F, Wellig P, Ott B, Heutschi K, Rühl T, Nussbaumer T. Detection and tracking of drones using advanced acoustic cameras. In: Unmanned/Unattended sensors and sensor networks XI; and advanced free-space optical communication techniques and applications, 96470F; 2015. p. 1–9. https://doi.org/10.1117/12.2194309

[3] Blanchard T, Thomas J-H, Raoof K. UAV's localization from a microphone array by exploiting the harmonic structure of the sound produced. In: Quiet drones 2020 - international e-Symposium on UAV / UAS noise; 2020. p. 1–10.

[4] Wu L, Fu Y, Yang X, Xu L, Chen S, Zhang Y, Zhang J. Research on the multi-signal DOA estimation based on ResNet with the attention module combined with beamforming (RAB-DOA). Appl Acoust 2025;231:110541. https://doi.org/10.1016/J.APACOUST.2025.110541

[5] Strauss M, Mordel P, Miguet V, Deleforge A. DREGON: dataset and methods for UAV-embedded sound source localization. In: IEEE/RSJ international conference on intelligent robots and systems (IROS 2018). IEEE; 2018. p. 5735–42. https://doi.org/10.1109/IROS.2018.8593581

[6] Herold G, Kujawski A, Strümpfel C, Huschbeck S, de Haag MU, Sarradj E. Flight path tracking and acoustic signature separation of swarm quadcopter drones using microphone array measurements. In: Quiet drones 2020 - international e-Symposium on UAV / UAS noise; 2020. p. 1–19.

[7] Blass M, Maly A, Graf F, Wellig P, Ott B. Towards mobile microphone array based UAV tracking. In: Quiet drones 2022 - international e-Symposium on UAV / UAS noise; 2022. p. 1–18.

[8] Merino-Martínez R, Sijtsma P, Snellen M, Ahlefeldt T, Antoni J, Bahr C, Blacodon D, Ernst D, Finez A, Funke S, Geyer T, Haxter S, Herold G, Huang X, Humphreys W, Leclère Q, Malgoezar A, Michel U, Padois T, Pereira A, Picard C, Sarradj E, Siller H, Simons D, Spehr C. A review of acoustic imaging methods using phased microphone arrays. CEAS Aeronaut J 2019;10:197–230. https://doi.org/10.1007/s13272-019-00383-4

[9] Ding S, Guo X, Peng T, Huang X, Hong X. Drone detection and tracking system based on fused acoustical and optical approaches. Adv Intell Syst 2023;5:2300251. https://doi.org/10.1002/AISY.202300251

[10] Blass M, Graf F. Design and analysis of mobile microphone arrays for acoustic drone tracking. In: Forum acousticum euronoise 2025; 2025. p. 1–8.

[11] Itare N, Thomas J-H, Raoof K. Genetic algorithm-based acoustic array optimization for estimating UAV DOA using beamforming. Drones 2025;9:149. https://doi.org/10.3390/DRONES9020149

[12] Altena A, Snellen M, Luesutthiviboon S, de Croon G, Voskuijl M. Frequency band analysis and comparison of localisation techniques for drones using microphone array measurements. JASA Express Lett 2025;5:24802. https://doi.org/10.1121/10.0035915

[13] James R. Underbrink, circularly symmetric, zero redundancy, planar array having broad frequency range applications. U.S. Patent Number 6, 205, 224 B1 2001.

[14] Vieira A, Snellen M, Simons DG. Experimental assessment of sound quality metrics for takeoff and landing aircraft. AIAA J 2021;59(1):240–9. https://doi.org/10.2514/1.J059633

[15] Van Veen BD, Buckley KM. Beamforming: a versatile approach to spatial filtering. IEEE ASSP Mag 1988;5:4–24. https://doi.org/10.1109/53.665

[16] Pieter S. Phased array beamforming applied to wind tunnel and fly-over tests. Tech. Rep. NLR-TP-2010-549 National Aerospace Laboratory (NLR); 2010.

[17] Lord Rayleigh FRS. XXXI. investigations in optics, with special reference to the spectroscope. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 1879;8:261–74. https://doi.org/10.1080/14786447908639684

[18] Blass M, Grebien S, Graf F. Experimental evaluation of acoustic drone tracking using mobile microphone arrays. In: Quiet drones 2024. University of Salford; 2024. https://doi.org/10.17866/RD.SALFORD.27697134.V1

[19] Hengy S, Mezzo SD, Sangwa-Simba M, Matwyschuk A, Pujol H, Bavu E. Drone intrusion detection using a network of acoustic arrays running a neural network for real-time drone detection, localization and identification. In: 30th international congress on sound and vibration (ICSV); 2024.