

Determination of the detection-threshold of human observers in acoustic drone detection

Samuel Huber^{*a}, Peter Wellig^b, Kurt Heutschi^c

^a Forventis GmbH, Predigergasse 19, 8008 Zurich, Switzerland

^b armasuisse, Science and Technology, Feuerwerkerstrasse 39, 3602 Thun, Switzerland

^c Empa, Swiss Federal Laboratories for Materials Science and Technology, Laboratory for Acoustics/Noise Control, Ueberlandstrasse 129, 8600 Dubendorf, Switzerland

ABSTRACT

Background: Nowadays, small drones are inexpensive and can be purchased and used very easily. Unfortunately, they are also relatively easy to convert to weapons. As they become more widespread, these drones may become a serious security risk. One possible way to address this threat could be the early detection of small drones by using acoustic cameras. However, the question arises as to how good the detection performance of such cameras is, compared to that of a human observer. The goal of this project was to determine the acoustic detection-threshold of human observers for drones in the presence of ambient noise. **Methods:** Nineteen subjects volunteered to take part in the study. The constancy method was used to determine the detection threshold. During the test, the study participants were presented with a recording of a DJI Phantom2 Vision+ drone that varied in level in steps of 1dB over a range of 27dB around the estimated threshold value. The signals were superimposed by three different kinds of ambient noise which were presented in three successive test-runs. The subjects wore headphones over which they heard the ongoing ambient noise while they were presented with the drone sound at random intervals and levels. The test signal was on for 2 seconds during which the trial subject had to confirm the detection of the drone sound by pressing an assigned key on a notebook. **Results:** We've found detection thresholds for white noise, water or highway noise at -17dB, -18dB and -17dB respectively, expressed as level differences between test signal and noise. Comparison of our results with the detection performance of human observers in a simulated drone detection scenario, reproduced by loudspeakers in an anechoic chamber, showed good agreement. Further, it seems possible to assess the detection performance of an acoustic camera using our results.

Keywords: Acoustic detection threshold, human factors, operator performance, drone detection

1. BACKGROUND

Small drones are now very cost-effective to purchase and easy to use in a wide range of commercial and hobby-like applications. However, the widespread use has also provoked criticism, particularly in the area of privacy, noise or security. It is possible to turn drones into an effective weapon with just a few minor modifications, for example by attaching a hand grenade. One area of the research program "Reconnaissance & Surveillance" at the Swiss Procurement Agency "armasuisse" deals with the detection of small drones by means of optical cameras. The question arises how good the detection performance of these cameras is compared to that of a human observer.

In this research project, the acoustic detection threshold of human observers for small drones is determined in the presence of ambient noise. Knowledge of the detection threshold allows to estimate human detection performance for existing and future drone models. armasuisse Science & Technology monitors the development of acoustic cameras. These consist of an array of several microphones. Due to their focusing possibilities, the sensitivity of these cameras usually exceeds the hearing capacity of humans [1]. However, the successful recording of a drone does not yet correspond to its detection. For this purpose, special sound analysis algorithms are required. In this respect, humans still outperform machines. The final detection performance results from microphone quality/hearing performance as well as detection quality. For this reason, it is interesting to know whether the detection performance of acoustic cameras exceeds that of human observers. Data on

*huber@forventis.ch

Target and Background Signatures V, edited by Karin U. Stein,
Ric Schleijsen, Proc. of SPIE Vol. 11158, 111580L · © 2019 SPIE
CCC code: 0277-786X/19/\$21 · doi: 10.1117/12.2533174

the human detection threshold will finally allow to compare man/machine performance on a theoretical basis. Furthermore, the data can be used to estimate the noise pollution from drones.

2. METHODS

The determination of human perception thresholds is a sub-discipline of psychophysics. Psychophysics is one of the oldest areas of psychological research. It refers to the interrelationship between subjective psychic (mental) experience and quantitatively measurable objective physical stimuli as the triggering processes. Psychophysics was introduced as a sub-discipline of experimental psychology by Gustav Theodor Fechner around 1860 and is based on the preliminary work of Ernst Heinrich Weber. It is sometimes considered as a sub-discipline of sensory physiology, perceptual psychology or psychophysiology.

There are basically three different methods for determining thresholds: 1. The method of adjustment requires test subjects to repeatedly adjust a variable stimulus so that it is just perceived (absolute threshold) or matches a predetermined standard stimulus (difference threshold). 2. The method of limits requires the investigator to vary the stimulus. In an ascending procedure, test subjects indicate when they first perceive the stimulus while, in a descending procedure, test subjects indicate when they just no longer perceive the stimulus. 3. In the method of constant stimuli, stimuli of different intensities are presented in random order so that no habituation or expectation errors may occur. The range of intensities is chosen, so that the lowest intensities are never, and the highest intensities are always detected.

In the present study, subjects had to detect drone sounds that were superimposed on ambient noise. We used the method of constant stimuli to determine the absolute detection threshold. In order to determine the threshold, a series of evenly graduated physical stimuli is presented to a subject several times in random sequence. The stimulus intensities chosen are distributed around the presumed threshold. The subjects assess whether to perceive the respective stimulus. The stimulus that is perceived in 50% of cases is the threshold stimulus. This method gives a sigmoid-like psychometric function, since clearly below-threshold stimuli are almost never, and clearly above-threshold stimuli are almost always recognized. The absolute detection threshold for this stimulus is – by convention – the stimulus intensity at which detection takes place in 50% of cases.

The test stimulus, in our case the drone sound, is presented in the presence of one of three ambient noises. This is a classic setting for the application of signal detection theory (SDT). The basics of the theory originate from radar applications and were first applied by Tanner to the perception of visual stimuli [2]. SDT is a model for measuring the quality of stimulus perception, considering sensitivity as well as specificity. In an SDT experiment, subjects are repeatedly offered stimuli (e.g. a sound) of different intensity in the presence of noise. However, there are also passages where no stimulus is present. The task of the test subjects is to indicate whether a stimulus has been presented. The background sound is referred to as noise (N) and the stimulus itself as signal (S). If the stimulus is weak, it may be confused with the noise. This results in four outcomes, namely “hits” (S detected), “correct rejections” (N correctly detected), “false alarms” (N confused with S) and “misses” (S confused with N). The question is, how well subjects can distinguish between N and S+N. The SDT assumes that the participants' answers in such an experiment are influenced by two parameters: sensitivity and the answer criterion. Sensitivity refers to the mean distance between the probability distributions for the detection of N and S+N. The answer criterion, on the other hand, denotes the subjective value that the perceptual effect must achieve for the subject to be indicated as perceived. For example, a person with a liberal criterion produces many false alarms with many hits at the same time, while a person with a conservative criterion hardly produces false alarms but produces more misses. The answer criterion can be directly influenced, e.g. by specifying examples of behavior such as “report a drone noise even if you only have a slight suspicion that a drone noise can be heard” or “do not report a drone noise until you are absolutely certain that it is actually there”. While the decision criterion may certainly influence the frequency of false alarms or missed signals, it doesn't change the detection quality as a whole. This means that the sensitivity measurement can be carried out independently of varying answer criteria. However, the assessment criterion must be taken into account for the determination of the threshold by correcting the absolute detection rates by means of the Abbott formula [3].

Stimuli

The stimuli used during the experiment and played by headphones involved a drone signal played at different levels (test signal) and three different background noise signals (masking signal). Two of the back-ground noise signals were real-world recordings of a creek and a highway, the third one was synthetically generated white noise. The test signal was a recording of a small drone (DJI Phantom2 Vision+) attached to a stand and operated at full throttle, taken in a distance of 1.5 m at an elevation angle of 30° below the horizontal plane.

Signal analysis

Table 1 shows the A-weighted signal levels of the stimuli with respect to a rectangular signal of amplitude 1 (full scale). Consequently, in the subsequent discussion, an A-weighted drone signal level of -x dB will represent -19.1-x dB(A). An A-weighted level corresponds to the rms-value (expressed in dB) of a signal that has been filtered with the A-weighting curve [4] to simulate the frequency dependent sensitivity of the human ear. Figure 2 shows the frequency contents of the stimuli in form of third octave band spectra.

Table 1: A-levels of the stimuli used in the listening tests with reference to full scale.

Signal	A-level re. full scale [dB(A)]
highway noise	-37.6
water sound	-32.0
white noise	-30.9
reference drone signal at 0 dB	-19.1

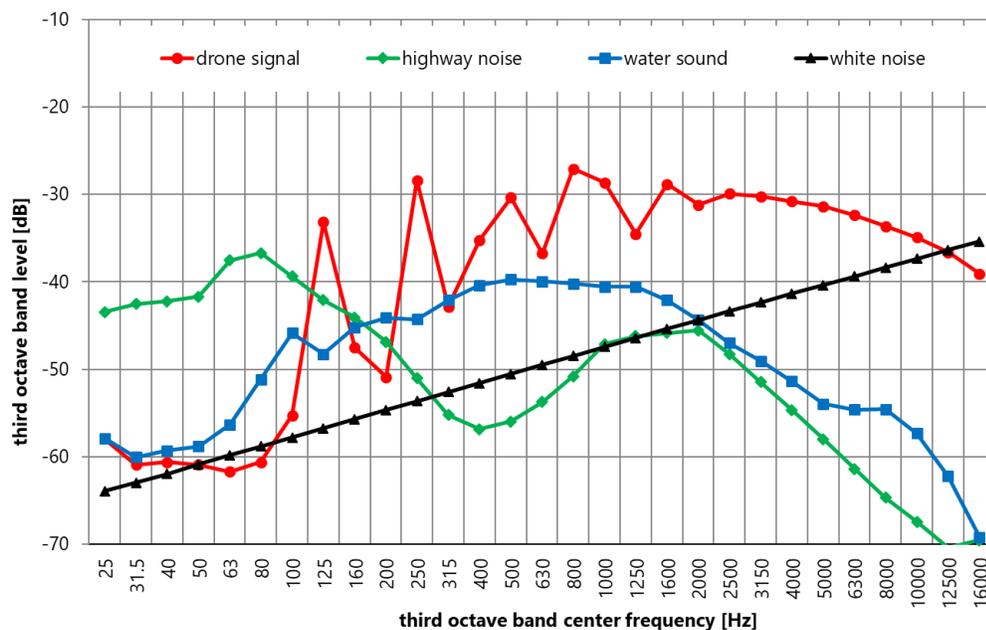


Figure 1: Third octave band spectra of the stimuli used in the listening tests.

Absolute level calibration

In the listening tests, the output voltage at the headphones output of the audio-interface was adjusted as follows (sequences with background noise only): a) water sound: 15 mV; b) highway noise: 15 mV and c) white noise: 18.3 mV. The same signal voltages were used in the calibration procedure in the lab. To this end, the signals were reproduced by the same

headphones and with the above voltage setting. The sound pressure at the position of the ear drums was measured with help of an artificial head manufactured by HEAD Acoustics (Figure 2). Subsequently the same signals were reproduced by a loudspeaker with the signal amplitudes adjusted for identical levels at the dummy head. Finally, after replacing the dummy head by a sound level meter, the free-field sound pressure levels were measured (Table 2).



Figure 2: Set-up in the lab showing the headphones mounted on the dummy head of HEAD Acoustics and the loudspeaker used for the free-field sound pressure measurement.

Table 2: Free-field A-weighted sound pressure levels of the stimuli used in the listening tests.

Signal	A-weighted sound pressure level re. 20 μ Pa [dB(A)]
highway noise	66
water sound	73
white noise	71
drone signal 0 dB	84
drone signal -10 dB	74
drone signal -20 dB	64
drone signal -30 dB	54

The acoustical emission measurement in 1.5 m distance of the drone *DJI Phantom 2 Vision+* resulted in an A-weighted sound pressure level of 80 dB(A) re. 20 μ Pa. Consequently, in 1 m distance a level of 83.5 dB(A) is to be expected. With a simplified propagation calculation that only considers geometrical spreading, the drone signal level L_{Drohne} in a distance r (in meter) can be estimated as:

$$(1) L_{Drohne} = 83.5 - 20 \text{Log}(r)$$

As an example, in a distance of 30 m, the sound pressure level of the drone signal is expected to drop by 30 dB, resulting in 54 dB(A). According to Table 2, this level just corresponds to the "drone signal -30 dB", that was played at 54 dB(A).

Study population

A total of 19 people with an average age of 34.9 years (range 14-59 years) took part in the study. 11 subjects were female and 8 were male. Participation in the study was voluntary and was compensated for. The only criterion for admission was normal hearing. Before the study, all applicants were subjected to a corresponding hearing test performed by the "Centre for Occupational Medicine, Ergonomics and Hygiene AG" (AEH) or the study manager. The recruiting of the subjects took place on the one hand among the employees at AEH as well as their relatives and acquaintances, on the other hand by the study manager. One person was rejected before the start of the study because of insufficient results from the hearing test. A second person with a pathological hearing test was admitted to the study, but the data was excluded from threshold determination (subject 15). Table 3 summarizes the characteristics of the subjects and gives an overview of the randomized order of the signal presentation as well as of the use of the data.

Table 3: Overview table Subject characteristic, interference signal sequence and data usage.

Nr.	Sex (F=female; M=male)	Age (years)	Normal hearing	Ambient noise: order of presentation (1=water; 2=highway; 3=white noise)	Data analysis
1	F	57	Yes	2 3 1	Ok
2	M	14	Yes	3 2 1	Ok
3	F	19	Yes	1 2 3	Ok
4	M	19	Yes	1 3 2	Ok
5	F	29	Yes	2 1 3	Ok
6	F	40	Yes	3 1 2	Ok
7	F	21	Yes	1 3 2	Ok
8	M	59	Yes	2 1 3	Ok
9	F	29	Yes	2 3 1	Ok
10	F	47	Yes	1 2 3	Ok
11	F	24	Yes	2 1 3	Ok
12	M	18	Yes	2 3 1	Ok
13	M	34	Yes	3 1 2	Ok
14	F	43	Yes	1 3 2	Ok
15	M	51	No	2 3 1	Data rejected
16	F	46	Yes	1 2 3	Ok
17	M	17	Yes	3 2 1	Ok
18	F	51	Yes	2 1 3	Ok
19	M	52	Yes	3 1 2	Ok
20	M	45	Yes	3 2 1	Ok

Study setup

The aim of the study was to determine the signal/noise ratio required to detect drone signals. For this purpose, an experiment-control software developed by Empa was used to present the continuous background noises together with the overlay of a two seconds lasting stationary DJI Phantom2 Vision+ drone signal of varying intensity. The level of the signal was varied and played in a sequence with random pauses of 3 to 10 seconds. The WAV files used have a sampling frequency of 44.1 kHz. The play-back sequence was controlled by an ASCII control file. For documentation of the experiment, an ASCII file was generated as the protocol of the hearing test. For each play-back segment (frame with or without drone signal) a new entry was generated, including ID, person details, date/time of start of the experiment, file name of the background noise and drone signal, running number, amplification of the drone signal, start and end time in seconds of the section, rating of the possible key press and time in seconds of the key press and response time.

Study procedure

For each participating subject the study took a total of 60 minutes to complete. Each subject went through a standardized study briefing and had the opportunity to ask questions and complete a trial run. The test signal presentation was carried out at 25 different signal levels between -25 dB and -49 dB. Each of the 25 signals was presented three times in a random order during every run with one of the three background noises. Hence, for each background noise, a total of 75 (3x25) test signals were presented. In total, the subjects were presented with 225 test signals and three different ambient noises within 45 minutes.

Study environment

The study was carried out in the premises of AEH and Forventis GmbH. A quiet but non-soundproofed room was provided for this purpose. The subjects used a pair of closed headphones (FOCAL Spirit Professional, frequency response 5 Hz - 22 kHz) for the study. Due to the insulation effect of the headphones, no noise from the study environment was perceived by the subjects. The study runs were conducted between 8am and 7pm.

Data collection & analysis

The data collection was carried out automatically by means of the hearing-test software described above. The data was imported into an Excel database (Excel 2016). The data processing as well as all evaluations and analyses were carried out in Excel.

3. RESULTS

First, the measured absolute detection probabilities were corrected for the assessment criterion by means of the Abbott formula:

$$(1) \ p^*(hit) = \frac{p(hit) - p(fa)}{(1 - p(fa))}$$

$p^*(hit)$ = corrected detection probability

$p(hit)$ = absolute detection probability

$p(fa)$ = probability for "false negative"

The detection threshold is determined graphically according to the constancy method and for each background noise separately. For the threshold determination, the corrected detection rates are used.

White noise

The absolute detection threshold of the drone sound in the presence of the background noise "White Noise" is -30 dB (50% of the signals detected). Figure 3 shows the uncorrected (Hit Rate) and the corrected detection rate (Hit Rate*) plotted against the False Negative Rate.

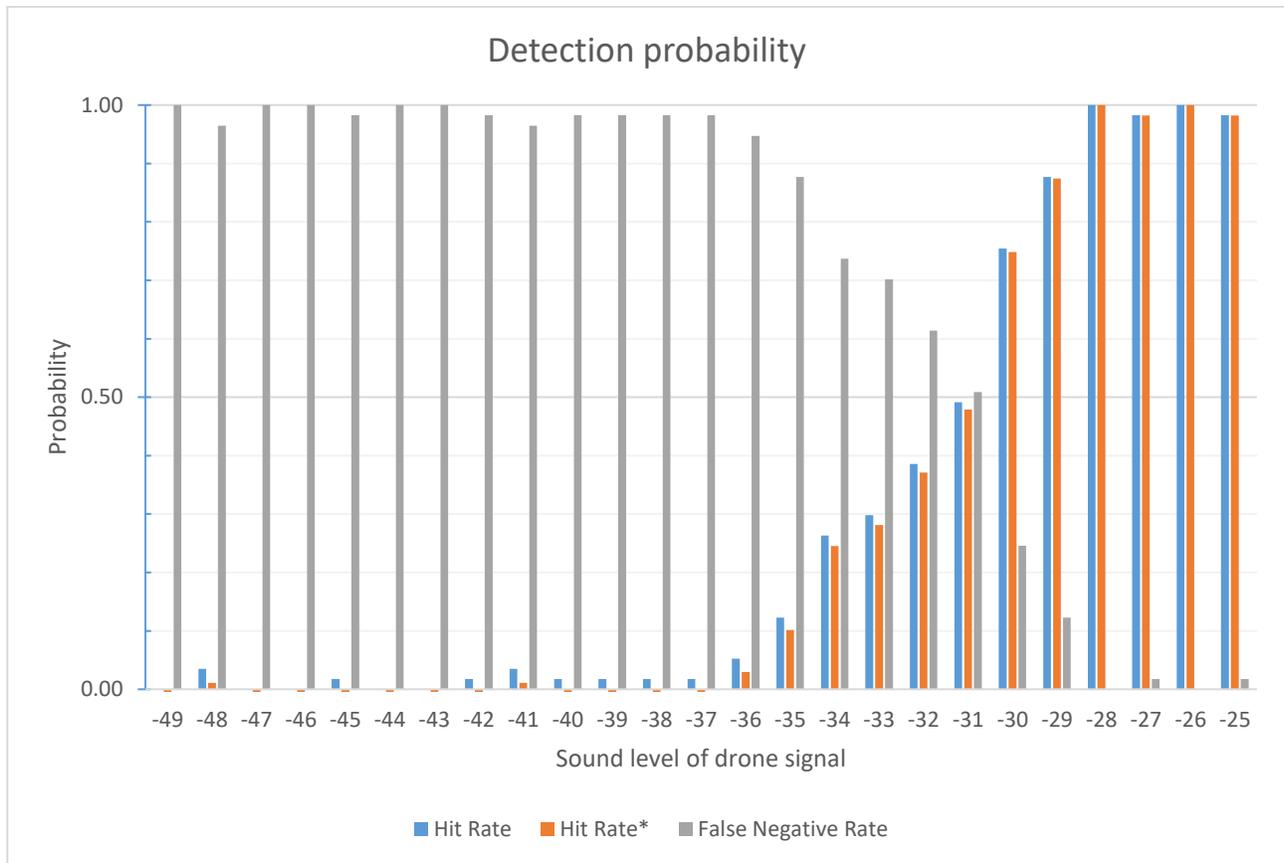


Figure 3: Detection probability for the background noise "white noise". Absolute detection threshold (50% detection probability) at -30 dB.

Water sound

The absolute detection threshold of the drone sound in the presence of the background noise "water sound" is -29 dB (50% of the signals detected). Figure 4 shows the uncorrected (Hit Rate) and the corrected detection rate (Hit Rate*) plotted against the False Negative Rate.

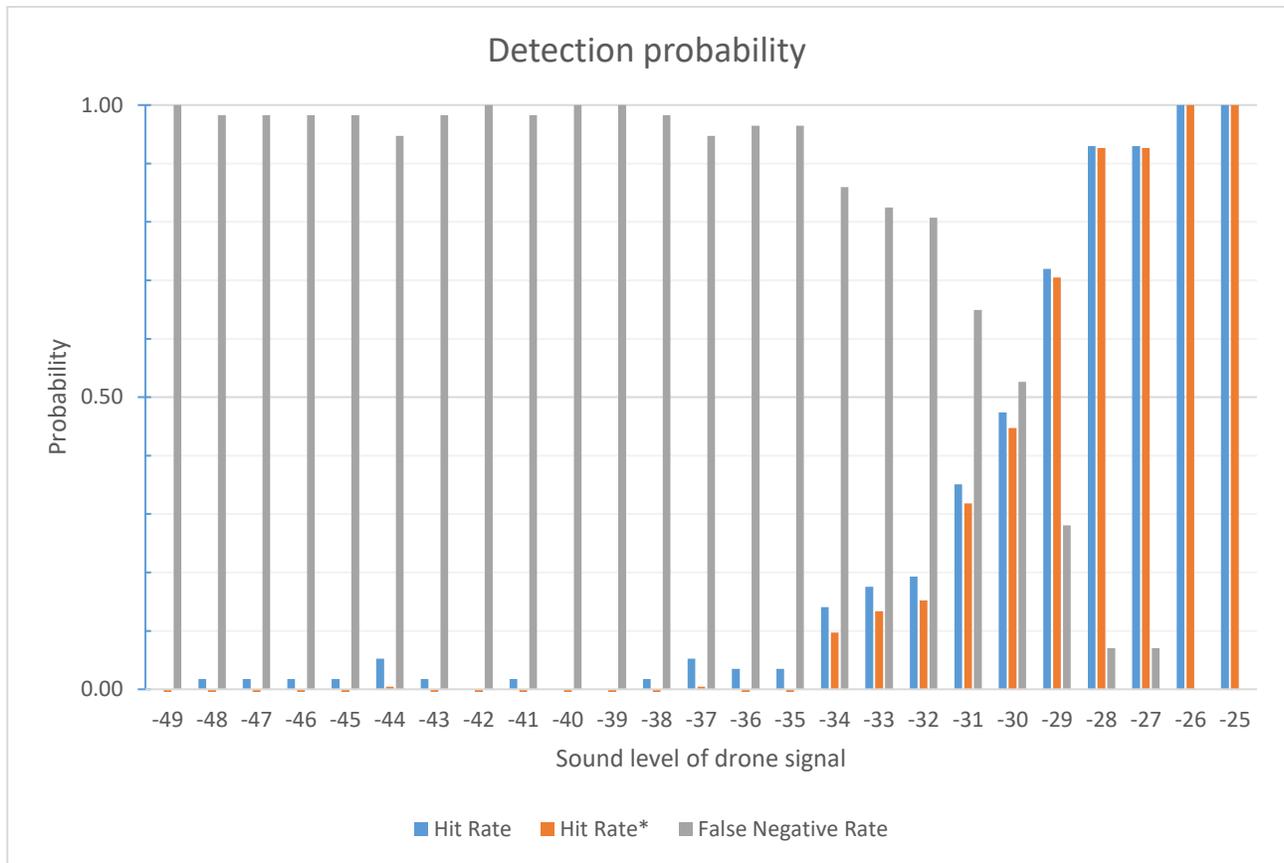


Figure 4: Detection probability for the background noise "water sound". Absolute detection threshold (50% detection probability) at -29 dB.

Highway noise

The absolute detection threshold of the drone sound in the presence of the background noise "highway noise" is -35 dB (50% of the signals detected). Figure 5 shows the uncorrected (Hit Rate) and the corrected detection rate (Hit Rate*) plotted against the False Negative Rate.

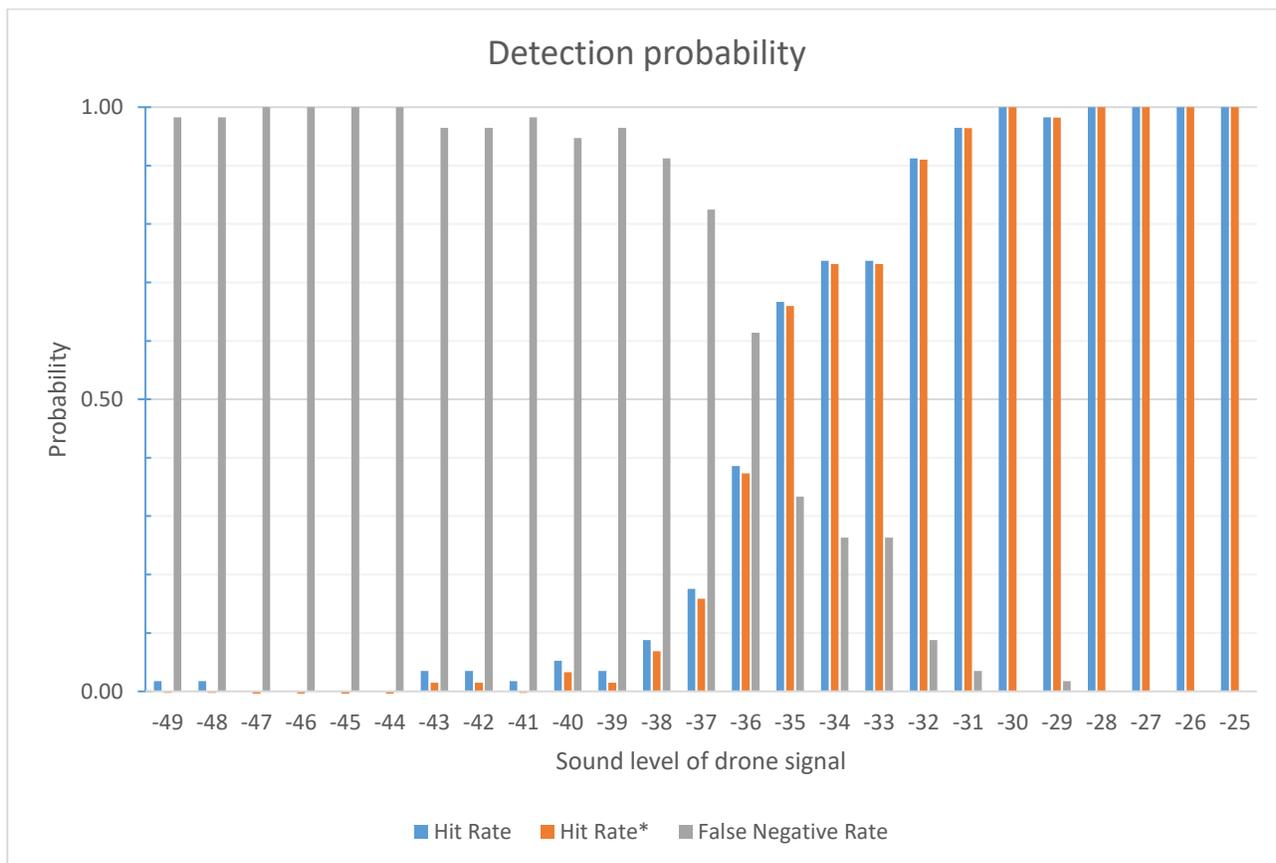


Figure 5: Detection probability for the background noise “highway noise”. Absolute detection threshold (50% detection probability) at -35 dB.

Table 4 gives an overview over the study results.

Table 4: Summary of the results.

Background noise	Detection threshold	Uncorrected detection probability (Hit Rate)	Corrected detection rate (Hit Rate*)
<i>White noise</i>	-30 dB	75%	75%
<i>Water sound</i>	-29 dB	72%	71%
<i>Highway noise</i>	-35 dB	67%	66%

4. CONCLUSION

The aim of this study was to determine the detection threshold of human observers for a drone signal in the presence of three different ambient noises. The goal of this setup was to determine human detection performance rather in a realistic noise environment than under laboratory conditions with the intention to enable a direct comparison of human detection performance with that of acoustic cameras in a real-life setting.

Three different sounds were used as background noise. The sounds used were chosen because of their (more or less) constant signal level. This guaranteed the test signals to be played in the presence of an approximately constant noise level so that the signal-to-noise ratio remained constant. However, on the one hand, natural sounds exhibit a certain level variance. While the water sound level remained more or less constant, the fluctuation of the highway noise was larger. The effect can be seen in the results, where the detection threshold for highway noise was significantly lower than for the noise of the creek or white noise, which can be explained by the nature of the highway noise (depending on the passage of a vehicle). Inevitably, some of the test signals were played during acoustic pauses between the passage of the vehicles, which might have supported their detection. The noise of the creek is much more constant in level, which also might explain the small difference to the results for the detection threshold for white noise (constant noise level).

Natural sounds have an inhomogeneous frequency spectrum, which can affect detectability, since the signal to noise ratio differs with frequency. This might explain the slightly higher detection threshold (=more difficult perceptibility) of the drone signal in overlay with the water noise compared to the one at overlay with the white noise. However, the difference in detection threshold of 1 dB is very small and has no practical relevance.

The perceived thresholds can be converted into detection distances. These distances are a practical measure for comparisons with acoustic cameras outside the laboratory in the open field. The detection thresholds listed in Table 4 refer to the sound level of the drone test signal. At an output level of 0 dB, the absolute A-rated free field level of the drone signal used in the hearing test is 84 dB(A). Based on this, the simulated observer drone distances can be determined via formula (1) as listed in Table 5.

Table 5: Perception threshold and simulated drone observer detection distance

<i>Background noise</i>	Absolute detection threshold	Absolute, A-rated open field level of the drone signal	Simulated detection distance
<i>White noise</i>	-30 dB	54 dB	30 m
<i>Water sound</i>	-29 dB	55 dB	27 m
<i>Highway noise</i>	-35 dB	49 dB	53 m

The distance values listed in Table 5 are to be understood as estimates. However, comparison of our results with the detection performance of human observers in a simulated drone detection scenario, reproduced by loudspeakers in an anechoic chamber, showed good agreement. The distances are quite small, which is due to the relatively high level of the background noise played during the experiment. Furthermore, in nature, the free-field conditions are usually not met, as the sound wave can often not spread undisturbed, but is subject to disturbances caused by e.g. reflections or damping. Also, the sound dynamics is largely disregarded in the present study, since the drone signal used lasted only two seconds and was almost constant. Sound dynamics is particularly present in drone sounds, both in terms of frequencies and volume. Thus, further studies are required to determine whether and to what extent sound dynamics affects the detection threshold of human observers. However, we believe the results of this study might still be useful when assessing the detection performance of acoustic cameras.

REFERENCES

- [1] Busset, J., Perrodin, F., Wellig, P., Ott, B., Heutschi, K., Rühl, Th., Nussbaumer, Th., "Detection and tracking of drones using advanced acoustic cameras", SPIE Security + Defence, 2015, Toulouse, France
- [2] Tanner, W. P., Jr., & Swets, J. A., "A decision-making theory of visual detection." *Psychological Review*, 61(6), 401-409 (1954)
- [3] Abbott, W.S., "A method of computing the effectiveness of an insecticide." *J. Econ. Entomol.*; 18 : 265-267 (1925)
- [4] DIN EN 61672-1:2014-07 Elektroakustik - Schallpegelmesser - Teil 1: Anforderungen (IEC 61672-1:2013).