High microphone signal-to-noise ratio enhances acoustic sampling of wildlife

Kevin Felix Arno Darras Corresp., 1 , **Franziska Deppe** [1] , **Yvonne Fabian** [1] , **Agus Priyono Kartono** [2] , **Andres Angulo Rubiano** [1] , **Bjørn Kolbrek** [3] , **Yeni Aryati Mulyani** [2] , **Dewi Malia Prawiradilaga** [4]

[1] Department of Crop Sciences, University of Göttingen, Göttingen, Niedersachsen, Germany

[2] Department of Forest Resources, Conservation and Ecotourism, Bogor Institute of Agriculture, Bogor, Indonesia

[3] Celestion, Ipswich, Suffolk, United Kingdom

[4] Cibinong Science Centre, Research Centre for Biology LIPI, Bogor, Indonesia

Corresponding Author: Kevin Felix Arno Darras
Email address: kdarras@gwdg.de

Background. Automated sound recorders are a popular sampling tool in ecology. However so far, the microphones themselves received little attention. Specifications that determine the recordings' sound quality are seldom mentioned. Here, we demonstrate the importance of microphone signal-to-noise ratio for sampling sonant animals.

Methods. We tested 24 different microphones in the field and measured their signal-to-noise ratios and detection ranges. We also measured the vocalisation activity of birds and bats that they recorded, the bird species richness, the bat call types richness, as well as the automated detection accuracy of bat echolocation calls.

Results. We provide the first measurements of a range of microphone models in the ultrasound range. Microphone signal-to-noise ratio positively affects the sound detection spaces and consequently, the sampled vocalisation activity and richness of birds and bats, as well as the automated detection accuracy of bat echolocation calls.

Discussion. Microphone signal-to-noise ratio is a crucial characteristic of a sound recording system. It should be maximised by choosing appropriate microphones, and be quantified independently, especially in the ultrasound range.

High microphone signal-to-noise ratio enhances acoustic sampling of wildlife

Authors

Kevin Darras[1], Franziska Deppe[1], Yvonne Fabian[1], Agus Priyono Kartono[2], Andres Angulo Rubiano[1], Bjørn Kolbrek[3], Yeni A. Mulyani[3], Dewi M. Prawiradilaga[4]

[1] Agroecology, Department of Crop Sciences, University of Göttingen

[2] Wildlife Ecology, Department of Forest Resources, Conservation and Ecotourism, Faculty of Forestry, Bogor Agricultural University

[3] Celestion, Ipswich, Suffolk, IP6 0NL, UK

[4] Research Centre for Biology LIPI, Cibinong Science Centre, Jalan Raya Jakarta Bogor KM 46, Bogor 16911, Indonesia

Corresponding author:

Kevin Darras

Agroecology, Grisebachstrasse 6, Department of Crop Sciences, University of Göttingen, 37077 Göttingen, Germany

kdarras@gwdg.de

Abstract

Background. Automated sound recorders are a popular sampling tool in ecology. However so far, the microphones themselves received little attention. Specifications that determine the recordings' sound quality are seldom mentioned. Here, we demonstrate the importance of microphone signal-to-noise ratio for sampling sonant animals.

Methods. We tested 24 different microphones in the field and measured their signal-to-noise ratios and detection ranges. We also measured the vocalisation activity of birds and bats that they recorded, the bird species richness, the bat call types richness, as well as the automated detection accuracy of bat echolocation calls.

Results. ~~We provide the first measurements of a range of microphone models in the ultrasound range.~~ Microphone signal-to-noise ratio positively affects the sound detection spaces and consequently, the sampled vocalisation activity and richness of birds and bats, as well as the automated detection accuracy of bat echolocation calls

Discussion. Microphone signal-to-noise ratio is a crucial characteristic of a sound recording system. It should be maximised by choosing appropriate microphones, and be quantified independently, especially in the ultrasound range.

Keywords

automated sound recorders, bats, birds, sound detection spaces, detection range, autonomous recording units, signal-to-noise ratio, ecoacoustics, microphone self-noise, soundscape

Introduction

Acoustic recording of wildlife is a popular sampling method for birds, bats, and other sonant animals (Gibb et al., 2018, Darras et al., 2019). ~~In a sound recording system, the recorder converts the electrical signal output by the microphones to a digital signal that is saved on storage media.~~ For the ecologist, microphone quality is essential as it determines whether the resulting recording's quality is sufficient to detect the animal sounds of interest. Microphone quality is commonly described by its self-noise and signal-to-noise ratio (commonly written SNR). Self-noise is the noise produced by the microphone in the absence of sound, and is typically given in dB SPL (decibel sound pressure level, defined as 20 times the logarithm of ratio of the sound pressure to the reference sound pressure of 20 µPa) A-weighted. It describes the equivalent background noise level that would be measured by a perfect (noiseless) microphone, and is ideally measured by placing the microphone in a sound proof container. Microphone self-noise defines the lowest sound pressure level the microphone can detect, and also the resulting signal-to-noise ratio of the recorded signals. Signal-to-noise ratio in dB is defined as ~~the~~ 10 times the logarithm of ratio of a standard signal's power to the noise power of the microphone created by its self-noise (Stewart & Lindsay, 1930). The standard signal is commonly generated by a sound calibrator with a 94 dB SPL tone at a 1 kHz sound frequency. Signal-to-noise ratio is a relative measure, valid only for a given signal level, while self-noise is an absolute measure of the microphone quality. Signal-to-noise ratio at a calibrated SPL will however give a measure of self-noise.

In the following, we will focus on the more commonly mentioned signal-to-noise ratios. Their importance is routinely implied in technical literature about microphones (Lewis & Schreier, 2013). In contrast, in ecoacoustics, even though some studies have evaluated the effectiveness of different recorder types for birds (Venier et al., 2012) and bats (Adams et al., 2012), little attention has been paid to the microphone. Indeed, out of 20 published studies used in a recent meta-analysis about autonomous sound recording (Darras et al., 2018 a), only six mentioned signal-to-noise ratios, and only two of those specified the signal-to-noise ratio of their own microphones. However, it is the first element in the signal processing chain and it determines the output recording's quality.

Technical specifications of microphones, including their signal-to-noise ratio, have an impact on the sampling effectiveness, probably through their impact on the detection ranges: microphones that have a low signal-to-noise ratio (a high self-noise) add too much noise to the recordings, so that the animal sounds - especially faint, distant ones - are not detectable anymore (Darras, et al., 2018 a). However, an experimental proof of the relationship between signal-to-noise ratio and detection ranges is still missing. Moreover, high signal-to-noise ratios should facilitate the automated detection of animal sounds (Kaplan, 1972), as well as their classification (Chen & Maher, 2006), both of which are commonly used in acoustic bat surveys.

We evaluated 24 microphones spanning a wide range of signal-to-noise ratios with respect to their effectiveness for sampling birds and bats. We recorded: 1) silence to determine the self-noise floor of our microphones, 2) sound transmission sequences to determine the microphones' signal-to-noise ratios and their detection ranges and 3) bats at night and birds during the morning to determine the sampling efficiency. We measured the microphones’ calibrated signal-to-noise ratios with test signals in the audible and ultrasound frequency range and compared them to manufacturer specifications to check their reliability. For the first time, we measured microphone signal-to-noise ratios in the ultrasound range. We test whether signal-to-noise ratios determine the detection ranges of the microphones at audible and ultrasonic frequencies. Since detection ranges determine the acoustic sampling areas, we can ultimately test whether signal-to-noise ratios affect the measured activity and richness of the sampled birds and bats. Additionally, we demonstrate how signal-to-noise ratios affect the proportion of bat calls detected with an automated bat call detection method, and also whether using automatic detection templates based on audio recorded with other microphones affects that measure.

Materials and methods

Study site and setup

We ~~purchased~~ omnidirectional microphone elements ~~(the actual sensor of microphones)~~ of different types and qualities. We used two units of each of 12 different models from six different manufacturers, with specified signal-to-noise ratios from 55 to 80 dB SPL, resulting in a total of 24 microphone elements (Table 1). Eight microphone element models were in the form of traditional cylindrical capsules, and two of them are used in commercial microphones (WM-61A in SMX-II, FG-23629-C36 in SMX-U1, Wildlife acoustics). Four models were Micro-Electro-Mechanical Systems (hereafter MEMS) chips that can be integrated on printed circuit boards, and four of them (plus one capsule) are part of an open-source microphone system Sonitor (Darras et al., 2018 c). Details about microphone assembly are in the supplementary materials. Our microphones were calibrated at 1 and 40 kHz using reference microphones for audible sound calibration, because not all had the same, standard format to fit a sound calibrator. For audible sound, we used the discontinued SMX-US (Wildlife Acoustics, Massachusetts, USA); it has a standard ¼ inch diameter, fitting into a class I sound calibrator (PCE-SC42, PCE instruments, Germany) that emits a 94 dB SPL tone of 1 kHz. For ultrasound, we used the reference microphone ICS-40720; it was calibrated with the ultrasound calibrator (Wildlife Acoustics, Massachusetts, USA) that emits a 48 dB SPL tone of 40 kHz (measured at a distance of 30 cm, because microphones are not plugged into it).

We set up the microphones in a research plot situated in an oil palm plantation (S01.70725, E103.39781, WGS84 datum) belonging to the PTPN6 state company in Sumatra, Indonesia. We installed 12 microphones, one of each model, simultaneously in a microphone holder consisting of a wooden pane (approximately 25 cm × 35 cm) with holes padded with foam, to equalise the soundscapes that they record. The microphones were oriented parallel to the ground, pointing to the horizon. We removed wind screens, which incur a slight loss in sound transmission, to measure the performance of the actual microphone elements. We ensured there was no rain as most microphones were not protected. The microphones were connected to six sound recorders (SM2Bat+, Wildlife Acoustics) with 5 m cables. We chose the SM2Bat+ because it was the only recorder that we possessed which was compatible with all microphones. The manufacturer specifies an electric noise floor of -115 dB V for 44.1 kHz recordings and -105 dB V for 192 kHz recordings. After completing all measurements with the first set of microphones, we repeated them with the second set.

We used recordings sampled at 96 kHz and generated their spectrograms with a Fast-Fourier-Transform with a Hanning window size of 1024. We measured relative sound levels in dB in the frequency bins containing the 1 kHz and at 40 kHz test frequencies using the "Plot spectrum" function in Audacity (Audacity Team, 2018). We chose to consistently use the 1024 window size as it is the default setting in Audacity that allowed the best trade-off between temporal and frequency resolutions for locating and measuring the short 40 kHz signal tone in space and time. Since we are dealing with field measurements with noise outside of the test frequencies, we did not use root-mean-squared values to measure sound pressure levels as they would cover the entire frequency spectrum. Moreover, we need to apply filtering in the following for aurally and visually detecting those signals, justifying the use of sound levels derived from specific frequency bins. Statistical tests were performed in R 3.6.1 (R Core Team, 2018).

Silent recordings

We measured microphone self-noise by recording sound in an environment that was as silent as possible. We did not have access to an anechoic chamber and preferred to record silence in the field, far from anthropogenic machinery noise rather than in the laboratory or other urban buildings. We used an isolating, large cylindrical ice box with a hole in its cover to pass the microphone cables. The box was padded inside with synthetic foam and surrounded by a thick polyester sleeping bag to prevent extraneous noise from reaching the microphones, resulting in a basic anechoic box. We started the recording on all recorders, knocked on the box to be able to synchronize them later, and recorded silence for one minute at a 96 kHz sampling rate.

We measured the relative sound pressure levels in dB – representing the uncalibrated self-noise of the microphones – inside the silent recording for all microphones 1 and 40 kHz. We used the same simultaneous 60 s of sound for all microphones of one set. To compute calibrated self-noise, we needed the relative sensitivity values of all microphones and the absolute sensitivity values of the reference microphones from the measurement detailed below.

Sound transmission sequences

We recorded sound transmission sequences (Darras et al., 2016) to determine relative microphone sensitivity, to compute signal-to-noise ratios, and to measure detection spaces. We generated an audio recording in Audacity, consisting of a sequence of 1 s long test tones at 1 kHz, repeated at 5 different sound levels to be able to choose the most appropriate sound level *a posteriori*. Audible test sounds came from a battery-powered one driver loudspeaker (SoundCore Anker) with the audio recording loaded on a Mini-SD card. We emitted ultrasound with the ultrasonic calibrator, which produces 40 kHz chirps of constant loudness in “Chirp” mode.

Sound transmission sequences are obtained by recording test tones emitted at different distances from the microphone: The loudspeaker and ultrasound calibrator were held at a height of 2 m and pointed to the microphones as they emitted test sounds at distances of 2, 4, 6, 8, 10, 15, 20, 25, 30, 35, 40, 45, and 50 m to the front, the left, the right, and the back of the microphones to determine the entire sound detection area.

At 1 kHz and 40 kHz, we used the nearest common distance at which none of the microphones recorded clipped (saturated) test tones to extract the test signal; it had the same source sound pressure level and distance to all microphones of one set. For every microphone, we measured the test sound level at 1 kHz with a 0.9 s audio selection, and at 40 kHz with the mean of ten 0.07 s audio selections (due to the short chirp emitted by the ultrasound calibrator). The relative sound pressure levels output by the recorder were calibrated by subtracting the amplification applied by the recorder and adding the frequency-specific calibration value obtained from the reference microphones. These values were used to calculate each microphone's sensitivity offset relative to the calibrated reference microphone. These offsets and the frequency-specific calibration values were used to calculate the self-noise of the microphones and compute signal-to-noise ratios relative to the absolute sound pressure level of the calibrators (94 dB SPL for 1 kHz, 48 dB SPL for 40 kHz).

We plotted the measured and specified (i.e. manufacturer given) microphone signal-to-noise ratio to check how consistent they are across manufacturers. Note that we did not expect our signal-to-noise ratios to absolutely equal manufacturer specifications due to the different procedure used for measuring them: we use frequency-specific measures of sound level while manufacturers use broad-band root-means-squared measures of sound level. Some microphones had acoustic vents (GAW112, Gore, USA) in front of them, which reduce sound transmission by <1 dB (specified by the manufacturer) while providing protection against water ingress; we corrected this by adding 1 dB to their measured signal-to-noise ratio, only when comparing them to manufacturer signal-to-noise ratios. We compared the corrected Akaike Information Criterion (Sugiura 1978) and adjusted R-squared of different models that use all combinations of specified signal-to-noise ratio and manufacturer variables, for predicting the measured signal-to-noise ratio as to find the one with the highest predictive power.

For measuring the microphones' standardized detection ranges, we chose the loudest of the five recorded sound levels at which none of the microphones recorded detectable (i.e. not visible on spectrogram and not audible in recording) test tones at 50 m by focusing on the relevant frequency (1 or 40 kHz). By doing this, we ensured that we measured maximal detection ranges for each microphone, and obtained the most accurate relative range differences. We determined the detection range as extinction distances: the distance at which the test tone was not detectable anymore (Darras et al., 2016). Since we only emitted test tones every 5 m, we estimated the detection range to the meter based on the experience of a single listener who analysed all sound transmission sequences and based on how loud the last detectable test sound was. For some combinations of microphone and direction however, detection ranges exceeded 50 m at 1 kHz, so we chose to measure sound pressure levels of the test tone (when audible or visible in the spectrogram) and the ambient sound directly afterwards at distances of 2, 4, 8, 15, 30, and 50 m. We fitted linear models of test tone sound pressure level against log-transformed distances to find the extinction distance (i.e., the detection range) for those combinations.

The detection ranges in the four different directions formed four quarter-ellipses which were used to calculate the sound detection space areas (Fig S2). We did not use the mean detection ranges in our analysis because our setup could have resulted in directional sound pickup patterns, and also because detection area is ultimately the measure that determines the sampling area for our organisms of interest. We tested for the relationship between the log-transformed microphone signal-to-noise ratio and detection area, depending on the frequency, using a linear regression model, based on the assumption that signal-to-noise ratios have a stronger positive effect at low values than at high values, where the detection area is increasingly limited by the ambient sound (Darras et al., 2018 a). We compared the AICc of that model against one with non-transformed signal-to-noise ratios. Note that we did not model more complex relationships between signal-to-noise ratio and detection area because a multitude of interacting, and sometimes opposed factors affect sound transmission: ~~amplification due to reflection between the canopy and ground, attenuation due to absorption by leaves and atmospheric absorption (for ultrasound), deviations due to directive animal vocalisations, refraction due to microclimatic conditions, etc.~~ Moreover, our detection ranges represent standardised ranges that are specific to the particular sound emitters we used, which are characterised by their sound frequency, amplitude, and directivity, in contrast to effective detection radii that are specific to particular species (Matsuoka et al., 2012).

Bird and bat recordings

After carrying out the sound transmission recordings, the microphones were left in place to record bats and birds. After retrieval, the recordings' sound levels were equalised by amplifying them to the sound level of the most sensitive microphone using the microphone-specific sensitivity offsets (see above).

We uploaded the amplified recordings to our online platform for ecoacoustics Biosounds (https://soundefforts.uni-goettingen.de/biosounds/collection/show/31/4/gallery) and screened the first 30 minutes after sunrise for birds and the first 30 minutes after sunset for bats. Synchronous recordings from every microphone were opened side-by-side. The recordings' spectrograms and audio enabled visual and aural filtering and detection of bird and bat vocalisations. The recordings that appeared to have the best quality (clearest spectrogram and noise-free audio) were used as reference (Knowles FG-23629-C36 for bats, Primo EM258 for birds). Clear differences were visible in the spectrograms between microphones (Fig S3). Whenever a vocalisation was found in the reference recording, it was tagged to extract its coordinates in time. Calls that were separated by less than 15 (for birds) or 5 seconds (for bats) were put in a common tag. Note that different thresholds could result in slightly different results. However, it is more important to use a consistent threshold across the variables of interest (here, the microphones) for unbiased results. All other non-reference recordings were checked at the same time point and if vocalisations could be found, they were also tagged. We checked whether other vocalisations were missed in the non-reference recordings that could not be found in the reference recording. We measured bird and bat activity and richness. We assigned bird vocalisations to species, and bat vocalisations to call types due to the lack of comprehensive reference libraries for South-East Asia, and also because bat species identity is not relevant here. We counted richness as the number of bird species or bat calls for each microphone. We computed the total duration of tagged vocalisations for birds and bats for each microphone, which yielded the vocalisation activity, in seconds. We tested the relationship between log-transformed microphone signal-to-noise ratio and bird and bat activity using a linear mixed effects model with the sampling day as a random intercept to account for day-to-day variations in animal activity. We compared the AICc of that model against one with non-transformed signal-to-noise ratio. Oil palm plantations have limited species pools: in our study area, we expect no more than four call types for echolocating bats and seven bird species (Darras et al., 2019). We graphically show how quickly the species pool was sampled by microphones of different measured signal-to-noise ratios. To do that, we plotted species accumulation curves of birds and bat calls against the sampling time, split in 20 discrete time steps. We also statistically analysed the influence of signal-to-noise ratio on the species richness at each time step. Due to the bounded distribution of the response variable (lower bound: 0, higher bound: 4 for bats, 7 for birds), we chose to express richness as a proportion of the maximum and arc-sine transformed it to model it with linear mixed-effects models with sampling date as a random intercept. ~~We did not use beta regression as values and 0 and 1 might occur.~~

For measuring automated detection accuracy, we chose one common bat call type. For each unamplified recording, we exported the corresponding bat passes along with two representative bat calls contained within them, which were the same across microphones. We used the latter as templates to detect all other calls in the exported bat passes with monitoR (Katz, Hafner, & Donovan, 2016). We counted the number of positive matches of calls that monitoR found inside each recording and divided it by the actual number of calls that we counted visually in the spectrogram of the bat passes. We obtained the proportion of correctly detected calls, which was our measure of automated detection accuracy, and modeled it against the signal-to-noise ratio with a beta regression model (Cribari-Neto et al., 2010, ~~0 and 1 values did not occur).~~

Additionally, to test how reference audio from different sources affects detection probability, we checked whether automated detection accuracy changes when using “external” templates, obtained from other microphones than those used for the recording that was analysed.

Results

Signal-to-noise ratios and detection areas

The linear model using specified signal-to-noise ratios and manufacturer identity as predictors had the highest predictive power (lowest AICc, ΔAICc: 5.92; Fig. S4). We detected large differences between manufacturers: for instance, PUI Audio indicated signal-to-noise ratio values that were on average 26 dB below those of Knowles microphone elements.

For several microphone and direction combinations, extinction distances had to be extrapolated (Fig S1), but we excluded measurements from 15 m distance for a particular direction from the analysis due to terrain irregularities. Detection areas increased significantly with log-transformed signal-to-noise ratio from 220 to 10750 m^2 at 1 kHz and from 118 to 4671 m^2 at 40 kHz (all $P < 0.01$, adjusted R^2: 0.96; Fig. 1). ~~Actual effect sizes are hard to interpret due to the log-transformation, and they are specific to our study site, so we do not report them~~. Detection ranges roughly corresponded to beams of omni-directional microphones (Fig S2).

Sampling effectiveness

The sampled vocalisation activities significantly increased with log-transformed signal-to-noise ratio from 0 to 1539 s for birds, and from 0 to 360 for bats (all $P < 0.01$, marginal R-squared: 0.95; Fig. 2). ~~Actual effect sizes are hard to interpret due to the log-transformation, and they are specific to our study site and animals, so we do not report them.~~

Maximal bird and bat species richness levels reached higher levels, at a higher rate, with increasing microphone signal-to-noise ratio (Fig 3). For all time steps when at least one species was detected over all microphones, the influence of the signal-to-noise ratio on the bird and bat arc-sine-transformed species richness proportion was statistically significant ($P<0.01$).

Detection accuracy

Bat call automated detection accuracy, measured by the proportion of automatically detected calls, was positively affected by microphone signal-to-noise ratio (pseudo-R-squared: 0.82, P = 0.015) and the detection accuracy was consistently lower when using external detection templates (from other microphones than those used for the analysed recording; Fig. 4). We only used the 12 microphones from the first night: Bat calls from the second night were distant, so that they were not picked up by two microphones, and with another microphone, MonitoR inexplicably led to thousands of false positives.

Discussion

Our measured signal-to-noise ratios showed large discrepancies between microphone element manufacturers. Signal-to-noise ratios considerably affected sound detection areas for audible sound and ultrasound. In turn, the sampled bird and bat activity and richness was largely enhanced by high microphone signal-to-noise ratios, and automated detection accuracy of bat calls also increased with microphone signal-to-noise ratio.

Measuring signal-to-noise ratios and detection ranges

For ecoacoustic studies, we need standardised microphone signal-to-noise values for specific frequencies of interest corresponding to different animal groups. Microphone manufacturer-provided signal-to-noise ratios did not correlate well with our standardised signal-to-noise ratio measurements because of strong differences between manufacturers. Microphone element manufacturers do not follow any standard certification for measuring microphone signal-to-noise ratio (pers. comm. with Vesper and PUI audio representatives). ~~The latter are usually specified in technical documentation for wildlife microphones, but rarely featured prominently in product brochures.~~ Moreover, signal-to-noise ratios are usually only specified for 1 kHz, a representative frequency for human speech (but see Knowles, 2014), so that we could not compare our signal-to-noise measurements at 40 kHz with any reference. However, for bat researchers, knowing ultrasound signal-to-noise ratio is imperative, and signal-to-noise ratios at 1 kHz are not indicative of microphone performance in the ultrasound range (Fig S5), even though some of the variation in our ultrasound signal-to-noise ratio values might have been caused by variable alignment of the microphones. Finally, our self-noise measurements were carried out for specific frequencies, but manufacturers usually measure signal-to-noise – which is based on the self-noise measure - with A-weighting (which weights human-audible frequencies more) over a 20

Hz to 20 kHz bandwidth that encompasses different groups of vocalising animals. Moreover, different weightings and bandwidths can lead to different signal-to-noise ratios. In future studies, signal-to-noise measurements in audible sound and ultrasound ranges should be compiled by researchers with standard measurement protocols for different microphone models to support microphone selection for ecoacoustic studies.

Using our sound detection area measurement approach, one could also directly benchmark microphones based on their detection ranges. It would be possible to devise protocols to evaluate sound recording setups - recorders with microphones, or only microphones - of different manufacturers in respect to their sampling effectiveness in a standardised way. Relative differences should be independent of the habitat in which recordings are made, but absolute ranges would vary between habitats. Possibly, results of previously published comparisons (Adams, Jantzen, Hamilton, & Fenton, 2012) could be explained better by simple differences in detection ranges caused by differing microphone signal-to-noise ratios.

Other microphone characteristics

We consider that sensitivity is secondary compared to self-noise or signal-to-noise ratios with respect to their impact on detection ranges. Microphones used for wildlife recordings usually have a sensitivity of -36 dBV, their levels are equalised with amplification, and the added signal-to-noise ratio that arises from this amplification is generally negligible. Even with our discontinued recorders, and only at ultrasound sampling frequencies, only the ICS-40720 and Primo EM258 microphone elements are notably limited by having higher electrical noise floors than the amplifier. Thus, we argue that for a broad range of commercially available microphone elements that come into question for wildlife recordings, signal-to-noise ratios are the limiting factor determining sound detection areas. This is also clearly supported by the statistically highly significant dependence of detection areas on signal-to-noise ratios we found, which were the only variable factor in our study setup.

Different microphone models have specific acoustic signatures. We recorded the exact same bat calls, but in spectrograms, they were visibly different between microphones. As a result, automated detection became less effective when we used external detection templates (from microphones other than the one used for recording). Thus on one hand, caution should be exercised when using reference material from online databases (such as Xeno-Canto for birds) to detect calls automatically. It seems preferable to extract detection templates directly from the analysed recordings - which is not always done (but see Ovaskainen et al., 2018) - to achieve maximal representativeness and detection accuracy. However on the other hand, for birds, it can be challenging to extract clean detection templates from field recordings (the ultrasound frequency range of bats is usually less noisy). Also, we could not directly test the performance of detection templates from online databases as we could not access other reference audio for the bats found in our recording.

Maximising recording quality

The inherent sampling effectiveness of sound recorders, in terms of sampling area, should be maximised by choosing microphones with high signal-to-noise ratios. In our case, we could reach half of the largest detection area with signal-to-noise ratios as low as 42 dB for birds and 81 dB for bats. However, this number depends on the range of signal-to-noise ratios covered by our selected microphones. In contrast, a previous meta-analytical approach showed that microphones for audible sound perform as well as human observers at signal-to-noise ratios of approximately 80 dB (Darras et al., 2018 a). Thus, more importantly, detectability increases with signal-to-noise ratio. Even though extremely low signal-to-noise ratios microphone elements are almost non-existent (~~such products do not have any application~~), it is worthwhile to search for the highest signal-to-noise ratios in the market of existing microphone elements: The detection areas of our wide range of microphone elements did not reach a clear saturation point, although detection spaces are eventually limited by the ambient sound (Apol et al., 2018, Darras et al., 2018 a). Moreover, the best-performance microphones can be obtained at little additional expense: As an example, the lowest signal-to-noise PUI Audio microphone element (55 dB SPL) cost us 3.10 EUR, only a little less (28 %) than the highest signal-to-noise PUI Audio microphone element (80 dB SPL) that cost 4.25 EUR. Compared to the total cost of microphones, these expenses are negligible, and even more so with even cheaper MEMS microphones, which can cost roughly half as much.

Low signal-to-noise microphones are not unusable, but they come with manageable drawbacks. Even if it is more efficient to sample birds and bats with high-quality ratio microphones, in some special cases where the sampling area should not be too large (e.g., when sampling needs to be limited to small habitat patches), having small microphone detection ranges can be an advantage. Also, sub-optimal recordings from low-quality microphones can be used too by sampling for longer durations to obtain higher vocalisation activities. For species richness, the results depend on the animals' mobility: when animals are territorial, species richness would always be lower with low-quality microphones; when animals are mobile within the sampled region, low-quality microphones could reach the same species richness values, albeit with longer sampling durations. Low-quality microphone recordings can also be used together with high-quality recordings when accounting explicitly for the different detectabilities (e.g., with occupancy modeling approaches). However, high signal-to-noise ratios are required for accurate localisation of sound sources (Good & Gilkey, 1996), which would support the estimation of bird detection distances for distance sampling (Darras et al., 2018 b). Finally, for all microphones, sensitivity and signal-to-noise ratios degrade with time, so that they should be regularly assessed to keep sampling effectiveness to a maximum or to account for their variable detection ranges (Darras et al., 2018 c, Turgeon et al., 2017).

Conclusions

We suggest that microphone signal-to-noise ratio or self-noise becomes a standard metric for assessing microphone quality in ecoacoustics. Microphone signal-to-noise ratio largely determines the sound detection space. Through this, it dictates how many individuals and species are recorded, and how accurately vocalisations can be automatically detected. Thus, high-quality microphones are paramount for achieving maximum detection ranges with accurately detectable sounds.

Acknowledgements

We declare that the authors have no conflicts of interest. We are grateful to our counterpart Damayanti Buchori for supporting this research project.

Authors' contributions

KD, FD, and YF conceived the study and designed the methodology; KD, FD, YF and AAR collected the data; KD, FD, AAR, and APK analysed the data; KD led the writing of the manuscript. All authors contributed to the drafts and gave final approval for publication.

Conflict of Interest

The authors declare that they have no conflict of interest.

Data availability

All data required to reproduce the results and figures will be provided in Dryad upon acceptance.

References

Adams, A. M., Jantzen, M. K., Hamilton, R. M., & Fenton, M. B. (2012). Do you hear what I hear? Implications of detector selection for acoustic monitoring of bats. *Methods in Ecology and Evolution*, *3*(6), 992–998. doi:10.1111/j.2041-210X.2012.00244.x

Audacity Team. (2018). Audacity(R): Free Audio Editor and Recorder (Version 2.2.2). Retrieved from Audacity Team

Apol CA, Valentine EC, Proppe DS (2019) Ambient noise decreases detectability of songbird vocalizations in passive acoustic recordings in a consistent pattern across species, frequency, and analysis method. Bioacoustics 0:1–15. https://doi.org/10.1080/09524622.2019.1605310

Chen Z, Maher RC (2006) Semi-automatic classification of bird vocalizations using spectral peak tracks. The Journal of the Acoustical Society of America 120:2974–2984. https://doi.org/10.1121/1.2345831

Cribari-Neto F, Zeileis A (2010) Beta Regression in R. Journal of Statistical Software 34:1–24. https://doi.org/10.18637/jss.v034.i02

Darras, K., Batáry, P., Furnas, B., Celis-Murillo, A., Wilgenburg, S. L. V., Mulyani, Y. A., & Tscharntke, T. (2018 a). Comparing the sampling performance of sound recorders versus point counts in bird surveys: A meta-analysis. *Journal of Applied Ecology*, *55*(6), 2575–2586. doi:10.1111/1365-2664.13229

Darras, K., Furnas, B., Fitriawan, I., Mulyani, Y., & Tscharntke, T. (2018 b). Estimating bird detection distances in sound recordings for standardizing detection ranges and distance sampling. *Methods in Ecology and Evolution*, *9*(9), 1928–1938. doi:10.1111/2041-210X.13031

Darras, K., Kolbrek, B., Knorr, A., & Meyer, V. (2018 c). Assembling cheap, high-performance microphones for recording terrestrial wildlife: the Sonitor system. *F1000Research*, *7*, 1984. doi:10.12688/f1000research.17511.1

Darras, K., Pütz, P., Fahrurrozi, Rembold, K., & Tscharntke, T. (2016). Measuring sound detection spaces for acoustic animal sampling and monitoring. *Biological Conservation*, *201*, 29–37. doi:10.1016/j.biocon.2016.06.021

Darras K, Batáry P, Furnas BJ, et al (2019) Autonomous sound recording outperforms human observation for sampling birds: a systematic map and user guide. Ecological Applications 29:e01954. https://doi.org/10.1002/eap.1954

Darras KFA, Corre MD, Formaglio G, et al (2019 b) Reducing Fertilizer and Avoiding Herbicides in Oil Palm Plantations—Ecological and Economic Valuations. Front For Glob Change 2:. https://doi.org/10.3389/ffgc.2019.00065

Gibb, R., Browning, E., Glover-Kapfer, P., & Jones, K. E. (n.d.). Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in Ecology and Evolution*, *0*(ja). doi:10.1111/2041-210X.13101

Good, M. D., & Gilkey, R. H. (1996). Sound localization in noise: The effect of signal-to-noise ratio. *The Journal of the Acoustical Society of America*, *99*(2), 1108–1117. doi:10.1121/1.415233

Kaplan, H. (1972). Detection of Sinusoids in Noise as a Function of the Signal-to-Noise Ratio on Individual Trials. *The Journal of the Acoustical Society of America*, *52*(1A), 142–142. doi:10.1121/1.1981903

Katz, J., Hafner, S. D., & Donovan, T. (2016). Tools for automated acoustic monitoring within the R package monitoR. *Bioacoustics*, *25*(2), 197–210. doi:10.1080/09524622.2016.1138415

Knowles. (2014). Ultrasonic Applications for Knowles Electret and MEMS Microphones. Knowles. Retrieved from https://www.digikey.com/en/pdf/k/knowles/ultrasonic-electret-mems-microphones

Lewis, J., & Schreier, P. (2013, December 18). Low Self Noise: The First Step to High-Performance MEMS Microphone Applications. InvenSense. Retrieved from https://www.invensense.com/wp-content/uploads/2015/03/Low-Self-Noise-The-First-Step-to-High-Performance-MEMS-Microphone-Applications.pdf

Matsuoka SM, Bayne EM, Sólymos P, et al (2012) Using Binomial Distance-Sampling Models to

Estimate the Effective Detection Radius of Point-Count Surveys Across Boreal CanadaUtilisation de modèles binomiaux d'échantillonnages basés sur la distance pour estimer le rayon de détection effectif lors d'inventaires par points d'écoute en région boréale canadienne. Auk 129:268–282. https://doi.org/10.1525/auk.2012.11190

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., … Wagner, H. (2015). *vegan: Community Ecology Package*. Retrieved from http://CRAN.R-project.org/package=vegan

Ovaskainen, O., Camargo, U. M. de, & Somervuo, P. (2018). Animal Sound Identifier (ASI): software for automated identification of vocal animals. *Ecology Letters*, *21*(8), 1244–1254. doi:10.1111/ele.13092

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Sugiura N (1978) Further analysts of the data by akaike' s information criterion and the finite corrections. Communications in Statistics - Theory and Methods 7:13–26. https://doi.org/10.1080/03610927808827599

Stewart, G. ., & Lindsay, R. . (1930). *Acoustics: A text on theory and applications*. New York, NY: D. Van Nostrand Company.

Turgeon P, Van Wilgenburg S, Drake K (2017) Microphone variability and degradation: implications for monitoring programs employing autonomous recording units. Avian Conservation and Ecology 12:. https://doi.org/10.5751/ACE-00958-120109

Venier LA, Holmes SB, Holborn GW, et al (2012) Evaluation of an Automated Recording Device for Monitoring Forest Birds. WILDLIFE SOCIETY BULLETIN 36:30–39. https://doi.org/10.1002/wsb.88

Table 1(on next page)

Microphone elements used in the study, along with letter codes used in Figure 1. MEMS: Microelectromechanical systems

<table>
<tr><th>Microphone element</th><th>Code</th><th>Type</th><th>Format</th><th>Manufacturer</th><th>Signal-to-noise ratio 1kHz (dB)</th><th>Sensitivity (dB SPL)</th></tr>
<tr><td>POM-1345P-C3310-R</td><td>A</td><td rowspan="7">electret condenser</td><td rowspan="5">capsule</td><td rowspan="5">PUI Audio</td><td>55</td><td>-45</td></tr>
<tr><td>POM-2735P-R</td><td>B</td><td>60</td><td>-35</td></tr>
<tr><td>ROM-2235P-HD-R</td><td>C</td><td>68</td><td>-35</td></tr>
<tr><td>POM-2730L-HD-R</td><td>D</td><td>74</td><td>-30</td></tr>
<tr><td>AOM-5024L-HD-R</td><td>E</td><td>80</td><td>-24</td></tr>
<tr><td>ICS-40720</td><td>F</td><td>MEMS</td><td>Invensense</td><td>70</td><td>-38</td></tr>
<tr><td>WM-61A</td><td>G</td><td>capsule</td><td>Panasonic</td><td>62</td><td>-35</td></tr>
<tr><td>PMM-3738-VM1000-R</td><td>H</td><td>piezoelectric</td><td rowspan="3">MEMS</td><td>Vesper</td><td>62</td><td>-38</td></tr>
<tr><td>SPM0404UD5</td><td>I</td><td rowspan="4">electret condenser</td><td rowspan="3">Knowles</td><td>59</td><td>-42</td></tr>
<tr><td>SPU0410LR5H-QB</td><td>J</td><td>63</td><td>-38</td></tr>
<tr><td>FG-23629-C36</td><td>K</td><td rowspan="2">capsule</td><td>66</td><td>-53</td></tr>
<tr><td>EM258</td><td>L</td><td>Primo</td><td>74</td><td>-32</td></tr>
</table>

Figure 1

The influence of microphone signal-to-noise ratio on detection space areas in the audible (1 kHz) and ultrasound (40 kHz) ranges.

Lines show predictions from a linear regression against log-transformed signal-to-noise ratio.

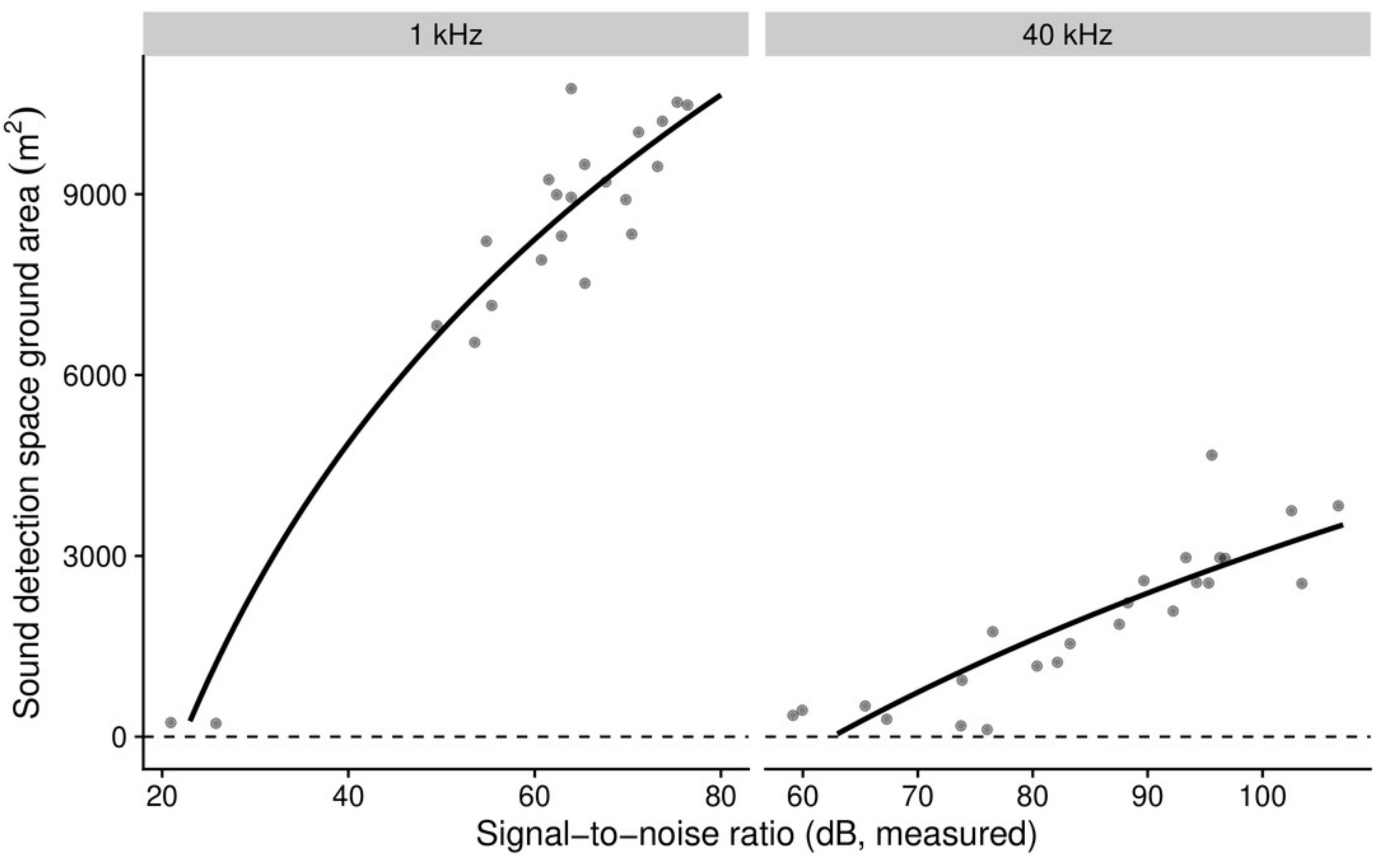


Figure 2

Vocalisation activity of birds and bats against microphone signal-to-noise ratio in the audible (1 kHz for birds) and ultrasound (40 kHz for bats) ranges.

Lines show predictions from a mixed-effects model with sampling day as random intercept.

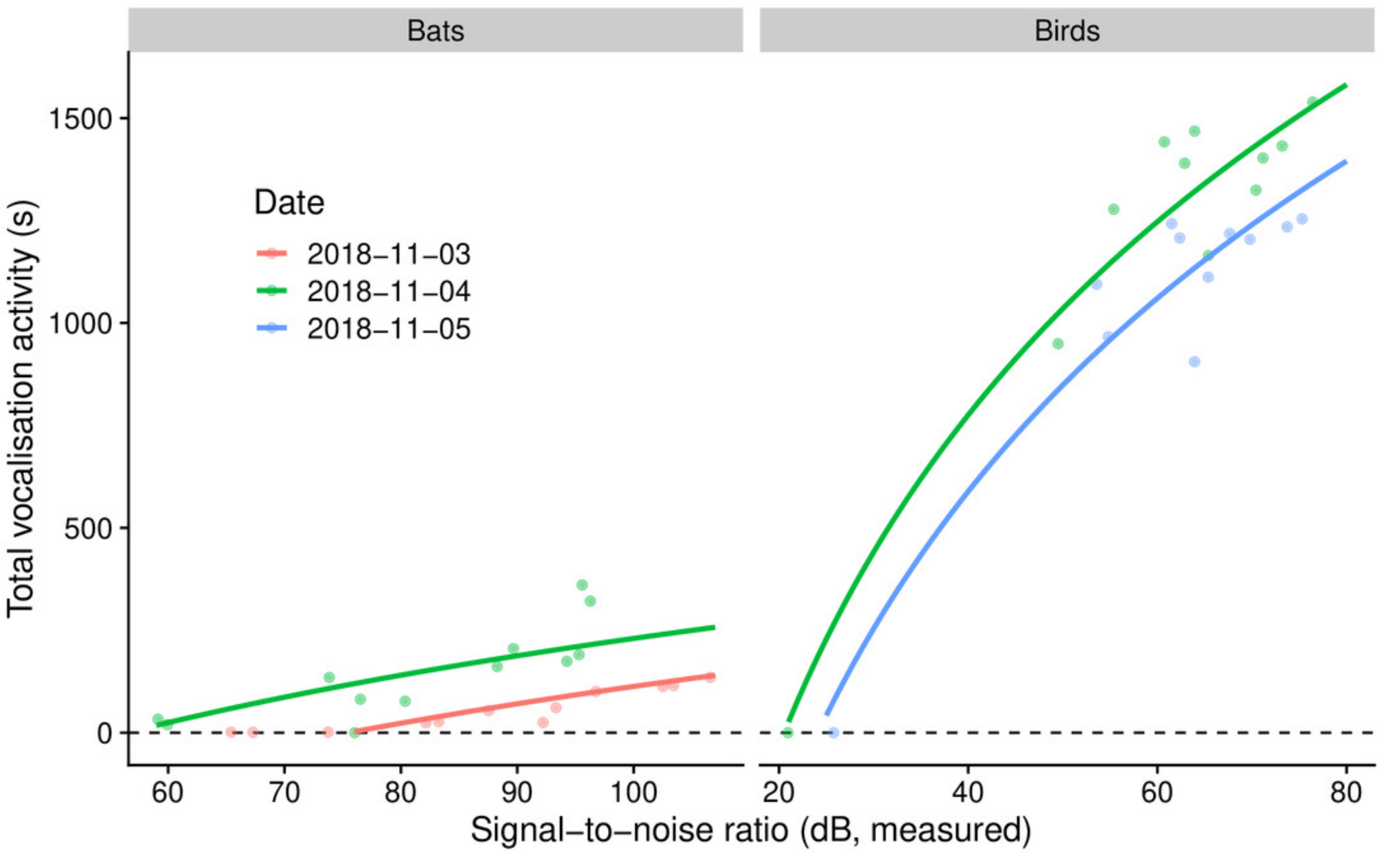


Figure 3

Species accumulation curves of each microphone, recording bird and bat species on different days, plotted against sampling time.

Signal-to-noise ratios are scaled within each taxon for achieving higher color contrast.

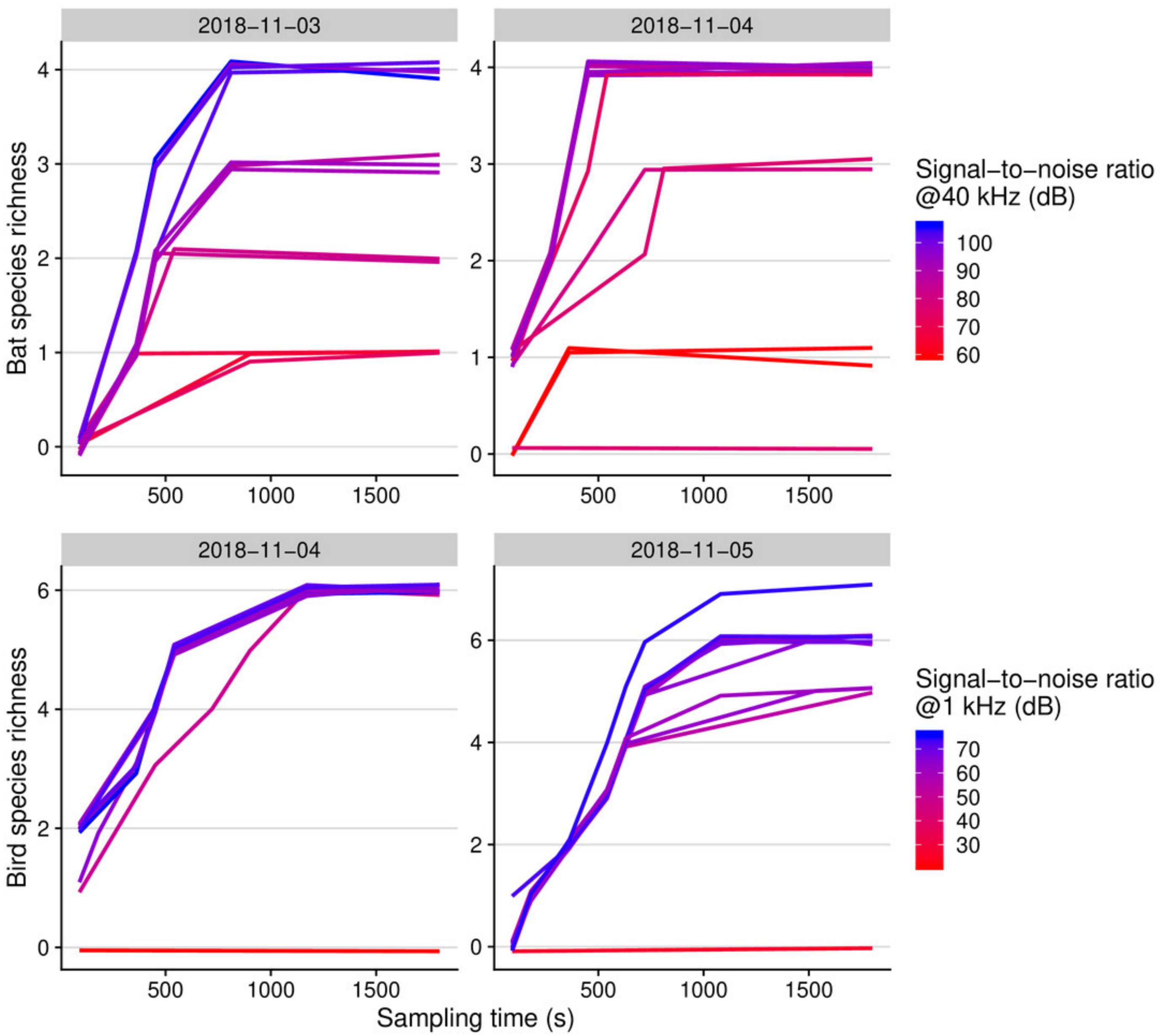
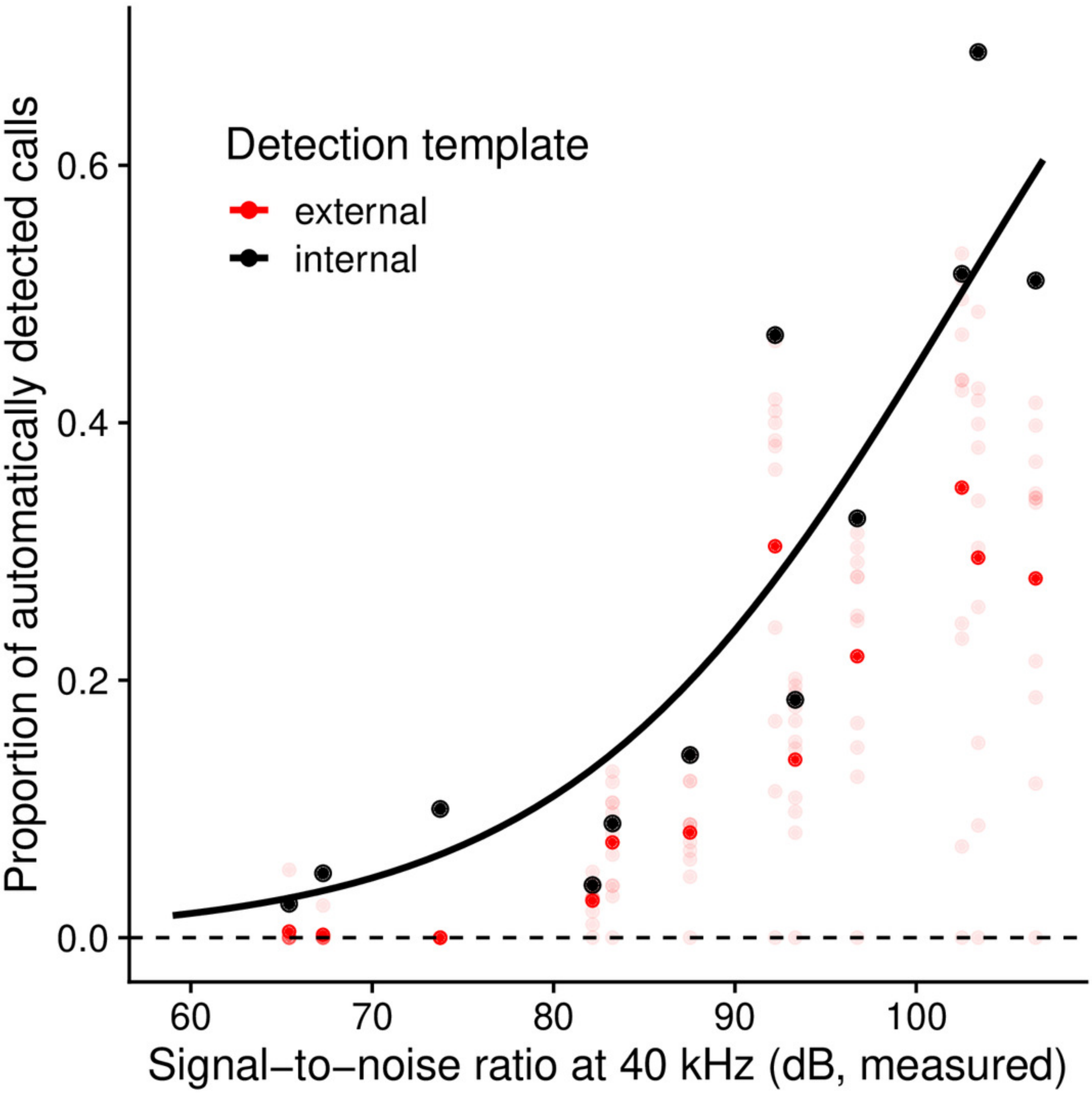


Figure 4

Bat call automated detection accuracy against microphone signal-to-noise ratio at 40 kHz, using a reference of 48 dB SPL.

The black dots are from internal detection templates (from within the actual analysed recording), and the line represents a beta regression over these points. The red transparent dots are from external detection templates (from recordings made with different microphones), and their means are shown without transparency.


Detection template
external
internal
Proportion of automatically detected calls
0.0
0.2
0.4
0.6
60
70
80
90
100
Signal–to–noise ratio at 40 kHz (dB, measured)