

# Estimating cue rates for animals with silence-bout sound patterns

Tiago A. Marques, Sónia Coelho, Ana I. Leal

15 August 2023

## Contents

<b>Introduction</b>	<b>1</b>
<b>Methods</b>	<b>2</b>
Silence-bout sound patterns . . . . .	2
The proposed estimator . . . . .	3
Illustrative example . . . . .	5
<b>Results</b>	<b>5</b>
<b>Discussion</b>	<b>9</b>
<b>References</b>	<b>10</b>

## Introduction

Reliable methods for estimating animal abundance are fundamental for effective management and conservation of wildlife. Traditional survey methods to estimate animal abundance rely on humans to make detections, usually visually but also aurally in the case of sound producing animals, like forest birds. For inconspicuous animals for which visual detections might be rare, but for which sound production occurs, passive acoustics might be an alternative to traditional surveys. In recent years passive acoustic monitoring density estimation has seen a steep increase in use, given that under some settings, it might be more effective than traditional methods, allowing for continuous monitoring of wildlife under settings which human observers might not even be able to operate. Examples of the use of passive acoustic data to estimate animal abundance include a wide variety of sound producing taxa, namely fish Rowell *et al.* (2017), birds Dawson & Efford (2009), frogs Measey *et al.* (2016), gibbons Kidney (2016) and cetaceans Marques *et al.* (2009).

There are many ways one can use animal produced sounds to derive animal density, but perhaps one of the most appealing is cue counting. Cue counting is an indirect method of estimating abundance, where instead of focusing on the animals, we rely on the sounds detected on acoustic sensors. These might be hydrophones in the water and microphones in terrestrial environments, covering the area for which we intend to estimate density. To convert the number of detected sounds to animal density we need to account for (1) sensor performance, which might include both the effective area of detection of the sensors, or equivalently a detection probability, and their probability of generating false positives, and (2) an estimate of cue production rate, the number of sounds produce per animal per unit time. Here we focus on the latter, cue production. Dealing with the first component is not the focus of this paper, but Marques *et al.* (2013) provide an overview of what is involved in the former regarding sensor performance characterization.

When conducting a passive acoustic density estimation survey, reliable cue rates for the time and place a survey is being conducted are required (e.g. Marques *et al.* (2013)). However, perhaps surprising, little is known about the frequency of sound production for many species of animals. Therefore, to obtain cue rates for passive acoustic density estimation, dedicated data collection is typically required. At sea, cue rates are

often estimated by placing animal borne tags with acoustic sensors on individuals. By counting the number of cues produced over the tag duration for a sample of animals cue rates can be obtained (e.g. Warren *et al.* (2017)). However, on terrestrial animals, like birds, focal follows of animals are typically considered (e.g. Sebastián-González *et al.* (2018)). If the start and end of a focal follow are independent of the mechanism by which animals produce sounds, then we are home free. One just needs a random sample of focal animals and using these we can obtain an average cue rate for the population of interest. However, if the inclusion of an animal in the focal follow also depends on its acoustic state, in particular, if we start and/or end the focal follow at the time of sound emission, we might introduce bias in the sound production rate. As an example, it might be expected that when finding birds for focal follows, birds that are vocally active are easier to find than those that are silent. If that were the case, we would over sample vocally active birds, overestimating the cue rate. This would then have a known effect inducing negative bias in abundance or density estimates.

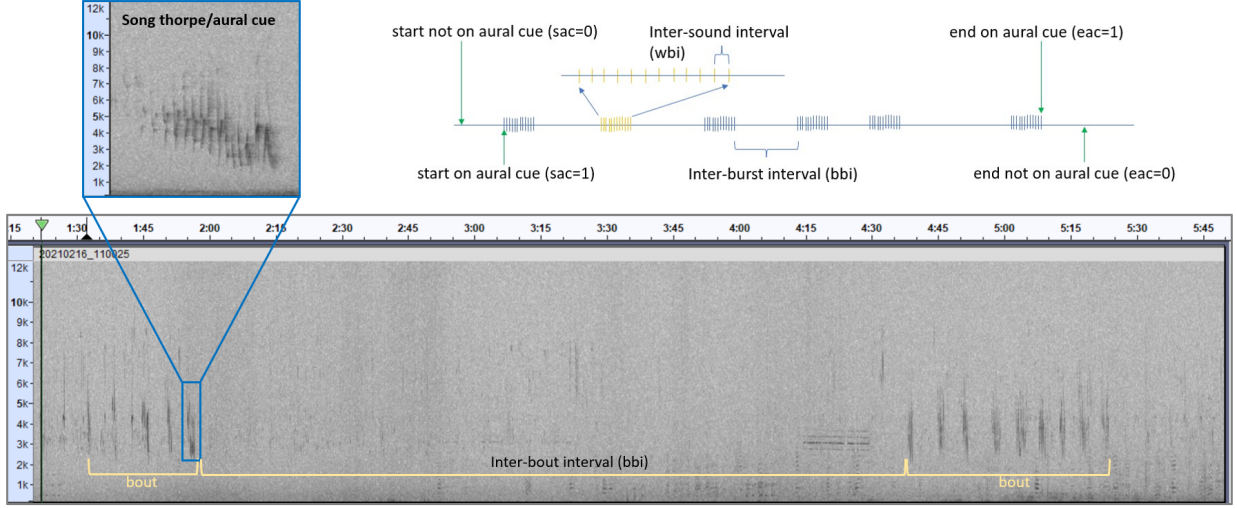
Here we present one such scenario, that might be plausible for a variety of taxa, where sounds are produced in pulses interspersed with silences. Animals will be therefore tendentially included in a focal follow when in a bout of sounds, and then the focal follow will stop when the animal is no longer detected, also typically acoustically. This will, unless corrected for, induce bias in cue rates. Inspired by a real case study of passive acoustic density estimation for the Common Chaffinch *Fringilla coelebs*, we propose a method to correct for the possible bias in start and end times of the focal follows. This leverages on a modified estimator for cue rates that adjusts the time considered as the recording period for each focal follow. We discuss the implications for estimates of cue rates from individual animals and the implications for passive acoustic density estimation via cue counting, which extends beyond the strict case of animals that produce sounds in bouts.

## Methods

### Silence-bout sound patterns

Consider a situation where animals tend to be silent for some periods, interspersed with periods where sounds are emitted in a relatively regular manner for a short period of time. We refer to such sound process describing the emission times of sounds of interest by a given species as a silence-bout sound pattern. We define the time between two successive sounds as an inter-sound-interval (*isi*), and we can distinguish two types of *isi*'s, (1) the time between successive sounds within a bout, referred as a within-bout-interval (*wbi*) and (2) the sound in between bouts, referred as a between-bout-interval (*bbi*). Figure 1 represents a spectrogram of a Common Chaffinch song (the blue window on the longer spectrogram) on the top left, with a diagram illustrating the concepts on the top right, and those same concepts overlaid on the real data spectrogram on the bottom panel.

We propose a novel way to estimate the mean *bbi* value based on focal follows, where the times of sound emission are recorded, and use that to correct the recording time associated with each focal follow. By doing so we avoid potential biases in defining the start and end of the recording period for each focal follow. We observe *isi*'s. Instead of defining strictly what are the observations that correspond to either *wbi*'s or *bbi*', we model the *isi*'s as a two-part finite mixture (e.g. Gamma mixtures, to account for the strictly non-negative support for cue rates), and assume that the component with the smaller mean corresponds to distribution of the *wbi*, while the component with the higher mean will correspond to the *bbi*. Therefore, all the data can be used in an integrated procedure to estimate the distribution of the *bbi*'s, and in particular their mean. The estimate of the mean *bbi* will be required for the cue rate estimator, as described below.



When we want to estimate cue production rates for terrestrial animals, which as noted above are required to convert estimates of sound density into estimates of animal density, we often select animals over which focal follows are conducted. During the focal follow a human observer will record the times of each sound emission, and at the end typically the number of sounds emitted divided by the focal follow duration are used as estimates of the individual cue production rate. A population averaged cue rate is obtained as an average of the individual cue rates.

Typically, however, two potential sources of bias might be at play:

1. if an animal is first detected by an aural cue, then the animal is most likely to be in a sound producing mode, here in particular, in a bout. Since this is the most likely way an animal will be found, we will tend to start recording during bouts, which means that the cue rate will be biased high;
2. if the animal is last detected exclusively aurally, then we are not sure whether the animal has left the area or not since it was last heard, and if it did, when it did, and so there is uncertainty about the time one should record to end the focal follow. The only natural choice is the last sound detected, but then this is likely to be the last sound in a bout, and again this will bias the cue rate estimate up.

Since both 1 and 2 are highly likely and both lead to overestimating cue rates, these will both contribute to underestimate animal density.

In this note we propose an approach to alleviate the problem, inspired by a real case study of Common Chaffinch in Portugal.

## The proposed estimator

Considering a focal follow, where the start  $T_{si}$  and end  $T_{ei}$  of the focal follow are independent with respect to the song production mechanism, an unbiased estimate of individual sound production rate for individual  $i$ , given a total observation period  $T_i = T_{ei} - T_{si}$  and  $n_{ci}$  detected cues will be, by definition of cue production rate

$$\hat{r}_i = \frac{n_{ci}}{T_{ei} - T_{si}} = \frac{n_{ci}}{T_i}.$$

Therefore, the overall cue rate based on a random sample of  $n$  focal follows will be given by

$$\hat{r} = \frac{\sum_{i=1}^n \hat{r}_i}{n}.$$

However, if the inclusion of the animal in a focal follow or the end of the focal follow are dependent on the sound production mechanism, as would be likely for sound producing species, also detected aurally, this cue rate estimator might be expected to be biased up.

Here we propose a more sensible estimator. It essentially adjusts the time of the focal follow duration based on the data collected. Two possible options are described below.

Consider  $T_{csi}$  and  $T_{cei}$  to be the times of the first sound detected and the time of the last sound detected, respectively. Additionally, consider two indicator variables, the Start on an Aural Cue ( $sac$ ) and the End on Aural Cue ( $eac$ ). Consider that  $sac=1$  if the start of a focal follow was on an aural cue,  $sac = 0$  otherwise. Therefore if  $sac = 1$ ,  $T_{cs} = T_s$ . On the other hand,  $eac=1$  if the end of the focal follow was defined at the last aural cue detected, otherwise  $eac = 0$ . Therefore if  $eac = 1$ ,  $T_{ce} = T_e$ . We suggest that if  $sac = 0$  the data are truncated such that the recording time starts at  $T_{cs}$ , and if  $eac = 0$  we truncate the data such that the recording time ends at  $T_{ce}$ , while adjusting for the average  $bbi$  at the start and end of the recording time, leading to the estimator

$$\hat{r}_i = \frac{n_{ci}}{\frac{\hat{bbi}}{2} + (T_{cei} - T_{csi}) + \frac{\hat{bbi}}{2}} = \frac{n_{ci}}{(T_{cei} - T_{csi}) + \hat{bbi}}.$$

The focal follow duration will then be assumed to correspond to the time difference between the first and the last cue detected to which we add an estimate of the  $bbi$ . This corresponds to half the  $bbi$  at the start and to half the  $bbi$  at the end of the focal follow. This assumes that, if the focal follow start and end coincided with a cue production time, or if we truncate the recording time to make it so, instead of taking these as the observed start and end of the focal follow, we must add to the recording time defined by the first and last sound detected, the average amount of time that an animal would be silent in between bouts.

Note that for any given focal follow, if  $sac = 0$  one might assume that the start was random with respect to the sound process. On the other hand if  $eac = 0$ , one might assume that the end of the focal follow was random with respect to the sound generating process. If the mean value for  $T_{c1} - T_s | sac = 0$  or  $T_e - T_{cn} | eac = 0$  is considerably different from  $bbi/2$  it is likely that the start and end of the follow are actually not independent with respect to the sound generating process. In such cases, the above procedure is recommended. If one is willing to assume that the start and end of a focal follow not coinciding with a sound are independent of the sound producing mechanism, then we could use the following estimator for animal's  $i$  cue rate

$$\hat{r}_i = \frac{n_c}{sac \frac{\hat{bbi}}{2} + (T_{cei} - T_{csi}) + eac \frac{\hat{bbi}}{2}}$$

where one-half of the  $bbi$  is added only if the start for a given animal focal follow was an acoustic cue, and one half of the  $bbi$  is added only if the end for a given animal follow was an acoustic cue. As before, considering a random sample of  $n$  animals for which focal follows are available, the empirical mean of the individual cue rates provides a population average.

To obtain an estimate of precision for the population's cue rate we propose to consider a non-parametric bootstrap, where the individual animals are taken as the independent sampling units. This allows to propagate into the cue rate precision the variability associated with estimating the mean  $bbi$ .

The proposed methods are illustrated below for a dataset of Common Chaffinch, which after a number of focal follows, we identified as having a clear silence-bout sound pattern. We compare the results obtained using a naive approach with our new estimator. We a priori hypothesized that the naive approach would be biased high for estimating cue rate and correspondingly biased low for estimating density. As can be seen below, our a priori hypothesis was wrong.

## Illustrative example

### Study area

The study was carried out in Companhia das Lezírias, a large farm in Santarém, Portugal, with around 11 000 ha of forest area. Cork oak woodlands are the most common habitat present, making up to 75% of the total forested area. Oak woodlands (Cork oak *Quercus suber* and Holm oak *Quercus rotundifolia*) are managed as silvo-pastoral systems, known as *montados* in Portugal and *dehesas* in Spain Pinto-Correia *et al.* (2011), and are recognized as an excellent example of balance between socio-economic development and biodiversity conservation Leal *et al.* (2016) Tellería (2007).

### Chosen Species

The Common Chaffinch is one of the most common species in these oak woodlands areas. Given its distinct and easily recognizable song compared to other species, which makes it easier to identify individuals of this species. This makes them suitable candidates for passive acoustic monitoring methods.

### Data collection using recorders

The fieldwork occurred between February and April of 2022, and data was collected using automatic sound recorders called AudioMoth (AM, Open Acoustics Devices). AudioMoths are small, light weighted and affordable recorders with an inbuilt microphone that can be programmed to record within specific filters and schedules, ranging in frequencies up to 348 kHz and recording throughout 24h, with or without duty cycling Hill *et al.* (2018) Hill *et al.* (2019).

For this study, AudioMoths were programmed to record at 48 kHz, which is the recommended frequency for birds, and they were set to start recording the first 4 hours after sunrise, for periods of 1 hour with a 5 seconds interval between each recording. No surveys were conducted under rainy or windy conditions.

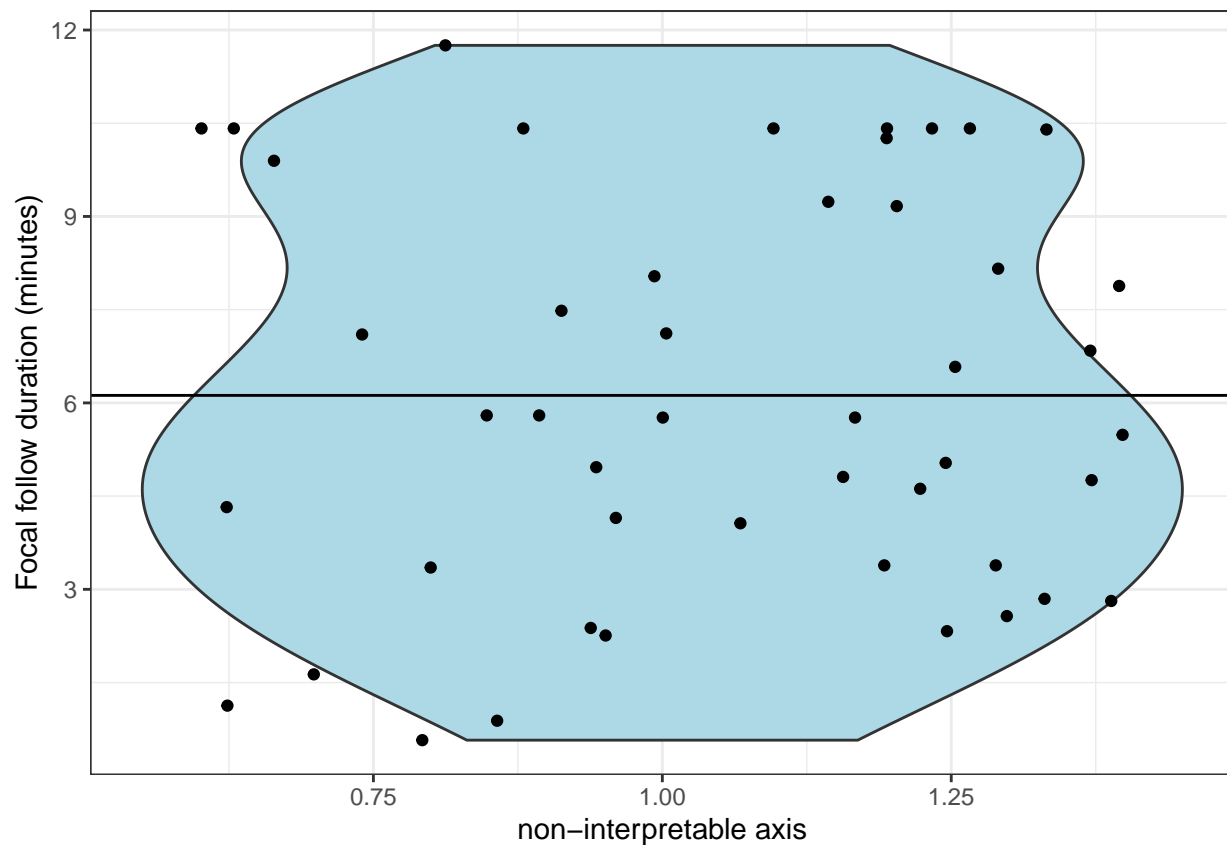
Birds produce two main types of vocalizations: songs and calls. Songs are usually loud and often long, more complex vocalizations that are mostly used by males to defend their territory, and they are formed by syllables (basic units), phrases (repeated syllables) and trills (rapid repetitions of 3 or more simple syllables) Gill (2006). Calls, on the other hand, are short and simple vocalizations, that are used by both males and females that can be warning calls, distress calls, flight calls, nest calls and flock calls Gill (2006). In this study, we used songs as our vocalization-type focus, given the fact that it will be a longer and easier to detect vocalization in an area with other vocalizing species. Chaffinches can have a repertoire of up to six song types Slater (1983), but we did not distinguish between them when counting aural cues. Regarding cue rate estimation, fieldwork was based on focal follows of individual chaffinches. The observer walked in the field and, once an individual was spotted, we registered every time it sang, up to a maximum of 12 minutes. During the data collection, if by any chance we would lose track of the individual we were following or we could not be certain anymore that it was the same individual, we would stop the timer and stop registering data for that individual. After each bird focal follow we would change location to reduce the likelihood of repeated inclusion in our sample of the same individuals.

### Statistical Analysis

We calculated cue rates as described above, considering the individual birds for which focal follows were conducted as a random sample of animals. To fit the Gamma mixture we used the R package moveHMM Michelot *et al.* (2016). All the code required to process the data and implement the methods is freely available at the following github repository: <https://github.com/TiagoAMarques/CueRates4SilenceBurstAnimals/>

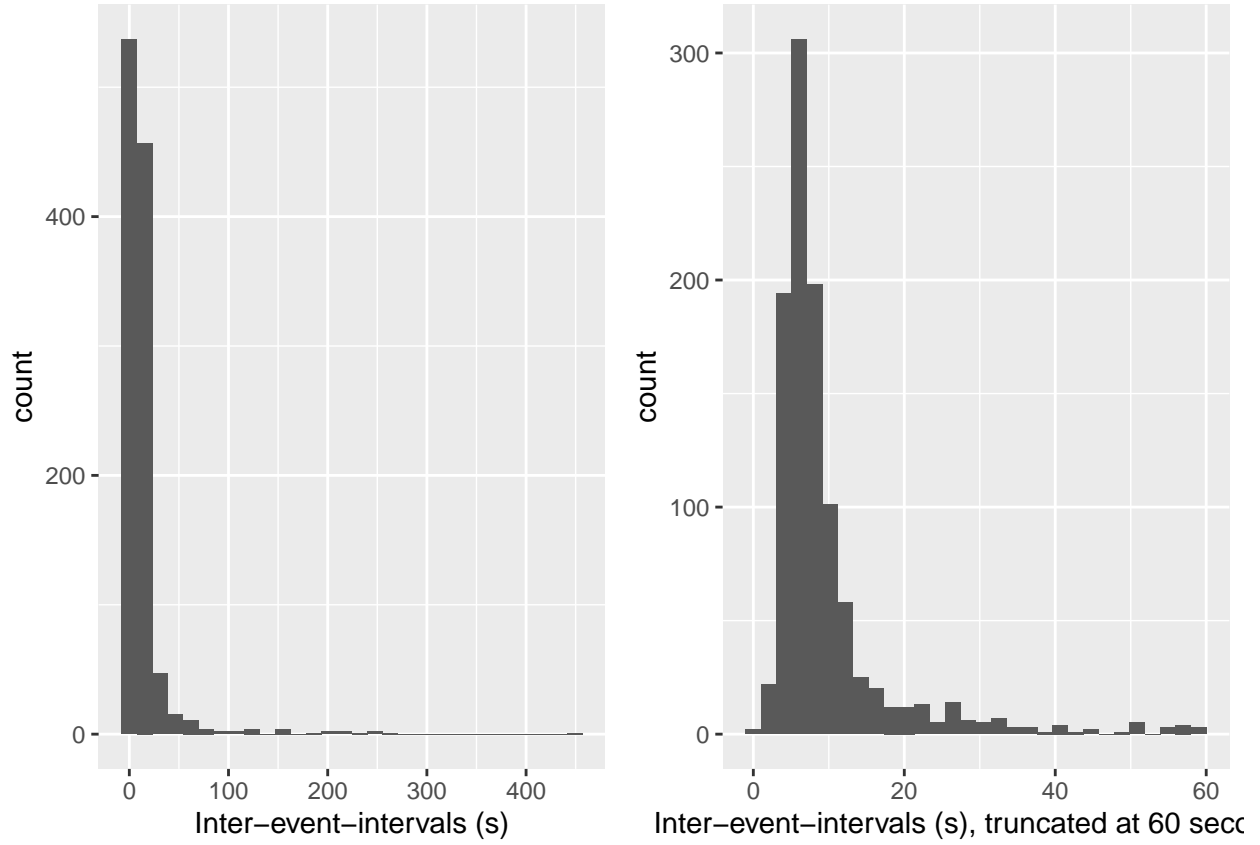
## Results

We have a total of 1057 times of sound production obtained from focal follows of 47 individuals. Individuals were followed on average for 6.12 minutes, ranging from 0.57 to 11.75 minutes. The distribution of these focal follow times is shown in Figure 1.

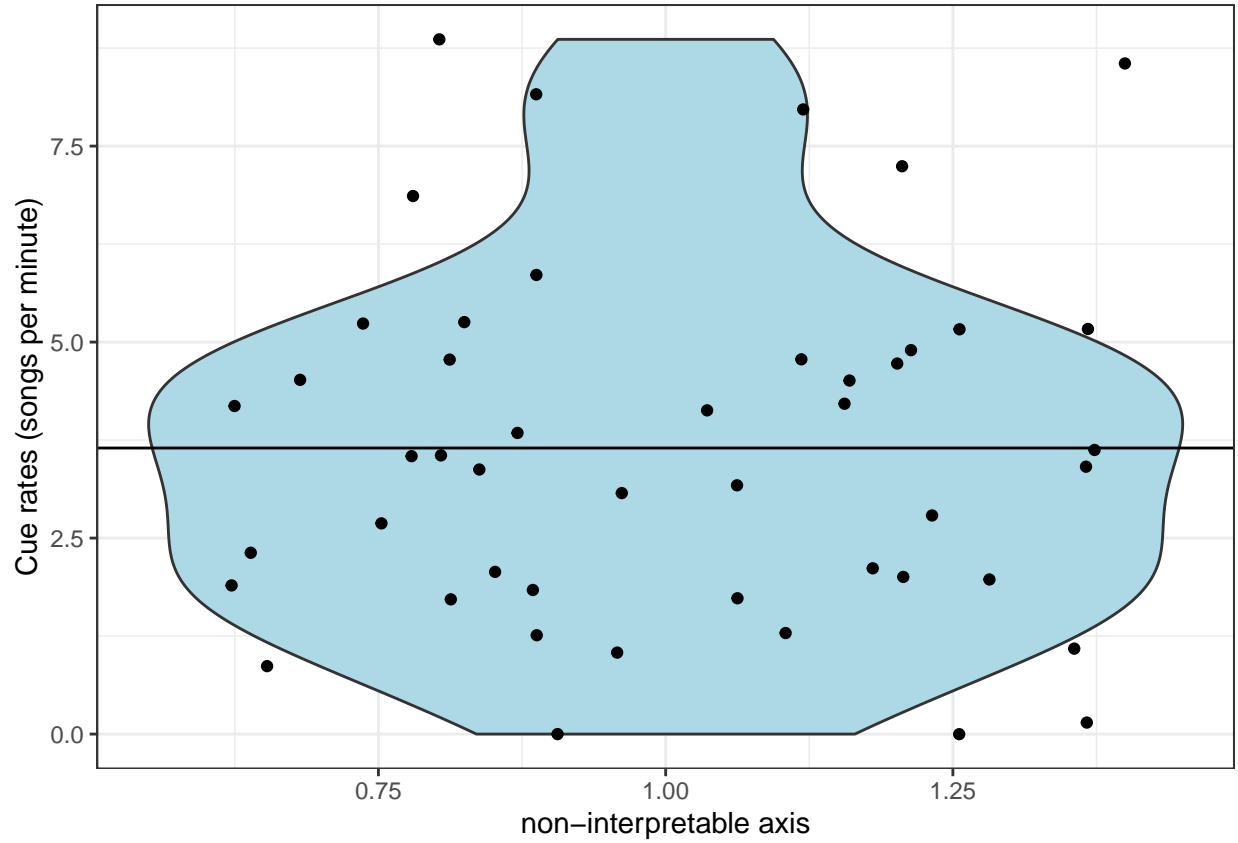


The times in between events are represented in Figure 2. Only 27 of the *isi*'s, corresponding to 0.03 % of the *isi*'s, are longer than 1 minute. We focus the attention on those *isi* that are under a minute (Figure 2, right panel).

```
## Warning: Removed 85 rows containing non-finite values (`stat_bin()`).
```

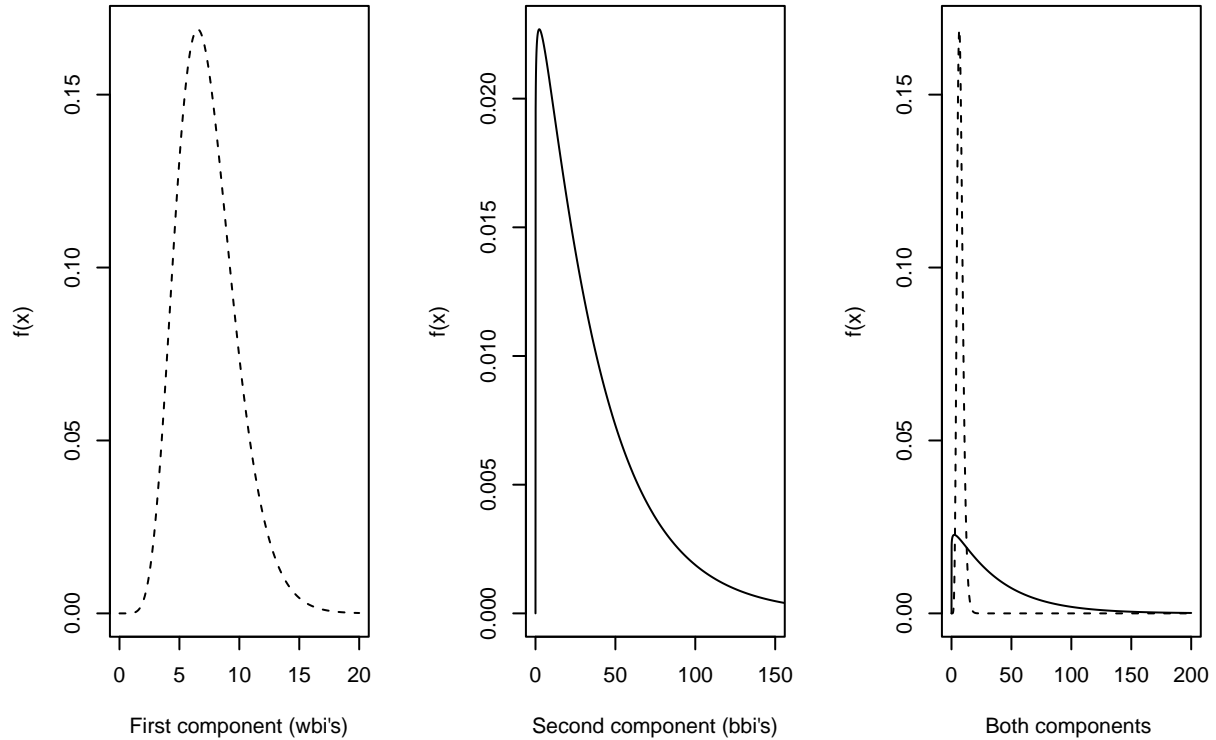


Based on the number of sounds recorded divided by the recording duration, conventional estimates of cue rates per animal are shown in Figure 3. The corresponding estimate of the average cue rate, estimated using a standard mean, is 3.65 cues per minute, with 95% CI of 2.98-4.31. The coefficient of variation is 9.04 %. As described in the methods, we suspect this might be a biased estimate of cue rate.



A graphical representation of the modelling of the *iei*'s as a two-part Gamma mixture is shown in Figure 4. Note this implies that the mean value of the first component, corresponding to the *wbi*, is 38.02 seconds, with a standard error of 36.82. On the other hand, the mean value of the second component, corresponding to the *bbi*, is 7.36 seconds, with a standard error of 2.48. This leads to an estimate of cue rate of 4.07, with 95% CI of 3.45-4.7. Incorporating a non-parametric bootstrap procedure, resampling the birds subjected to focal follows, leads to a 95% CV of 3.29-4.56 (CV of 7.81).





## Discussion

Cue rates are fundamental to implement passive acoustic cue counting density estimation exercises. Here we reported a cue rate that might be used elsewhere to estimate Common Chaffinch density: 4.07 sounds per minute (95% CI 3.29-4.56, % CV 7.81). This value is different from previous reported average values of 6 to 7 songs/minute Deoniziak & Osiejuk (2016) and Riebel & Slater (1999). Interestingly, these authors did not identify the silent-bout pattern we describe here. Therefore we note that this estimate should be valid for our study area, a Portuguese *montado*, during spring, during the species breeding season and for morning (x-y) surveys, consistent with the birds peak vocal activity. It stands on researchers wanting to use this cue rate estimate for Common Chaffinch density estimates from passive acoustics at other places and/or times the responsibility to assess and discuss whether our estimates might be sensible for them. Further studies are required to assess how cue rates for the Common Chaffinch change as a function of internal (sex, behaviour) or external (temperature, time of day, season) factors. A density estimate from cue counting based on a biased cue rate estimate will necessarily be biased Marques *et al.* (2013). Therefore further research is welcome in general looking further at the drivers of sound production of species for which passive acoustic density estimation might be considered.

We presented a way to correct for the duration of a focal follow to prevent biases in estimates of cue production rate estimates. We anticipated that since animals are more likely to be found - to be subsequently followed - aurally, a naive cue rate would produce biased high cue rates, oversampling individuals in a vocally active state. Nonetheless, the naive estimate of cue rate was 3.65 (95% CI 2.98-4.31, % CV 9.04), while the cue rate based on the proposed estimator was 4.07 (95% CI 3.29-4.56, % CV 7.81). Therefore, against our expectations, we obtained a corrected cue rate estimate that was higher than the naive estimate. This provides an excellent example of how two wrongs do not make a right, reflecting something we initially overlooked. Knowing *a priori* about the expected direction of the bias, the researcher responsible for the focal follows was actively trying to find birds non-acoustically. Instead of having a sample of animals that is somehow biased towards

animals producing sounds, we obtained the contrary. These animals were quite active during the focal follow periods, and more easily detected while flying. Therefore, animals would often be included in the sample as they came into the area of detection of the observer looking for potential individuals for focal follows. In other situations, the observer followed the sound of the bird to identify its exact location, but began the focal follow as soon as got sight of it, which often happened during a period of silence. Hence, the recording for most animals period started when silent, and hence we oversampled for the silence periods. In fact, of the 47 individuals, only 9 started to be followed by acoustics. On the other hand, only 8 animals were last recorded first as acoustics, showing that the researcher was quite effective at keep tracking of the animals being followed once they went silent. In total only 2 of the animals recording period started and ended at a sound event. This indicates that the desire to avoid biasing results by including only vocal birds might have had the unexpected outcome of biasing the cue rate low, by mostly picking up birds during silent periods. This reflects, once again, the dangers of trying to do random sampling in a scenario where the ability to truly randomize sampling is not possible. Trying to undo expected biases might lead to over-correction of such a possible expected bias.

While the estimator proposed here might be used to obtain cue rate estimates which will be more accurate than naive estimates, the case study also illustrates the dangers of assuming one has a random sample of individuals when estimating cue rates, even if not from focal follows. If the sampling procedure is not truly random, which is often hard to insure logistically under real life scenarios of wildlife sampling, we are very prone to over sample somehow STRANGE (*sensu* Webster & Rutz (2020)) animals. This could lead to severely biased estimates, with bias of unknown direction. A good example comes from the cetacean world, where cue rates are mostly estimated based on animal borne tags with acoustic sensors. Tagging cetaceans is extremely difficult, and therefore animals that are somehow easy to tag might be over-sampled. If there is any correlation between the probability of an animal being tagged and its cue production, as is likely given that both might be dependent on, say, behavioral state, bias in cue rate estimates might be introduced. Some of the tagging induced bias are often considered, like the fact that a tagged animal might not behave the same as if it had not been tagged, as a consequence of the tagging process itself, often referred as tag-on effect Warren *et al.* (2020) Nielsen *et al.* (2023). However, whether the sample of tagged animals is indeed random with respect to the population of interest has been often ignored. We urge researchers working with animals sampled “at random” but without a formal mechanism to enforce said randomness to consider carefully and discuss possible selection biases in their samples and implications in the inferences made based on said data.

## References

- Dawson, D.K. & Efford, M.G. (2009). Bird population density estimated from acoustic signals. *Journal of Applied Ecology*, **46**, 1201–1209.
- Deoniziak, K. & Osiejuk, T.S. (2016). Disentangling relations among repertoire size, song rate, signal redundancy and ambient noise level in european songbird (M. Manser, Ed.). *Ethology*, **122**, 734–744.
- Gill, F.B. (2006). *Ornithology*. W. H. Freeman.
- Hill, A.P., Prince, P., Covarrubias, E.P., Doncaster, C.P., Snaddon, J.L. & Rogers, A. (2018). AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. *Methods in Ecology and Evolution*, **9**, 1199–1211.
- Hill, A.P., Prince, P., Snaddon, J.L., Doncaster, C.P. & Rogers, A. (2019). AudioMoth: A low-cost acoustic device for monitoring biodiversity and the environment. *HardwareX*, **6**, e00073.
- Kidney, B.M.A.B., Darren AND Rawson. (2016). An efficient acoustic density estimation method with human detectors applied to gibbons in Cambodia. *PLoS ONE*, **11**, e0155066.
- Leal, A.I., Rainho, A., Martins, R.C., Granadeiro, J.P. & Palmeirim, J.M. (2016). Modelling future scenarios to improve woodland landscapes for birds in the Mediterranean. *Journal for Nature Conservation*, **30**, 103–112.
- Marques, T.A., Thomas, L., Martin, S.W., Mellinger, D.K., Ward, J.A., Moretti, D.J., Harris, D. & Tyack, P.L. (2013). Estimating animal population density using passive acoustics. *Biological Reviews*, **88**, 287–309.
- Marques, T.A., Thomas, L., Ward, J., DiMarzio, N. & Tyack, P.L. (2009). Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville’s beaked whales. *The Journal of*

- the Acoustical Society of America*, **125**, 1982–1994.
- Measey, G.J., Stevenson, B.C., Scott, T., Altwegg, R. & Borchers, D.L. (2016). Counting chirps: Acoustic monitoring of cryptic frogs. *Journal of Applied Ecology*, **56**, 894–902.
- Michelot, T., Langrock, R. & Patterson, T.A. (2016). moveHMM: An R package for the statistical modelling of animal movement data using hidden markov models. *Methods in Ecology and Evolution*, **7**, 1308–1315. Retrieved from <http://dx.doi.org/10.1111/2041-210X.12578>
- Nielsen, L.R., Tervo, O.M., B.Blackwell, S., Heide-Jørgensen, M.P. & SusanneDitlevsen. (2023). Using quantile regression and relative entropy to assess the period of anomalous behaviour of marine mammals following tagging. *Ecology and Evolution*.
- Pinto-Correia, T., Ribeiro, N. & Sá-Sousa, P. (2011). Introducing the montado, the cork and holm oak agroforestry system of Southern Portugal. *Agroforestry Systems*, **82**, 99–104.
- Riebel, K. & Slater, P.J.B. (1999). Song type switching in the chaffinch, *fringilla coelebs*: Timing or counting? *Animal Behaviour*, **57**, 655–661.
- Rowell, T.J., Demer, D.A., Aburto-Oropeza, O., Cota-Nieto, J.J., Hyde, J.R. & Erisman, B.E. (2017). Estimating fish abundance at spawning aggregations from courtship sound levels. *Scientific Reports*, **7**, 3340. Retrieved from <https://doi.org/10.1038/s41598-017-03383-8>
- Sebastián-González, E., Camp, R.J., Tanimoto, A.M., Oliveira, P.M. de, Lima, B.B., Marques, T.A. & Hart, P.J. (2018). Density estimation of sound-producing terrestrial animals using single automatic acoustic recorders and distance sampling. *Avian Conservation and Ecology*, **13**, 7.
- Slater, P.J.B. (1983). Sequences of song in chaffinches. *Animal Behaviour*, **31**, 272–281.
- Tellería, J.L. (2007). Passerine bird communities of iberian dehesas: A review. *Animal Biodiversity and Conservation*, **24**, 67–78.
- Warren, V.E., Marques, T.A., Harris, D., Tyack, P.L., Thomas, L., Soto, N.A. de, Hickmott, L. & Johnson, M.P. (2017). Spatio-temporal variation in click production rates of beaked whales: Implications for passive acoustic density estimation. *The Journal of the Acoustical Society of America*, **141**, 1962–1974.
- Warren, V., Miller, P. & Tyack, P. (2020). Short-term responses of sperm whales *physeter macrocephalus* to the attachment of suction cup tags. *Marine Ecology Progress Series*, **645**, 219–234.
- Webster, M.M. & Rutz, C. (2020). How STRANGE are your study animals? *Nature*, **582**, 337–340.