



# Uncertainty and ignored information in the analysis of bat ultrasound: Bayesian approximation to the rescue

Tore Christian Michaelsen<sup>a,\*</sup>, Jens Rydell<sup>b,1</sup>, Rasmus Bååth<sup>c</sup>, Knut Helge Jensen<sup>a</sup>

<sup>a</sup> Department of Biological Sciences, University of Bergen, 5020 Bergen, Norway

<sup>b</sup> Independent Researcher

<sup>c</sup> Department of Cognitive Science, Lund University, 221 00 Lund, Sweden

## ARTICLE INFO

### Keywords:

Bayesian statistics  
Bat ultrasound analysis  
Probability theory  
GNU R

## ABSTRACT

Bat ultrasound analysis has been around for several decades and it is one of the most important tools in studies of bat ecology. Discrimination between species is based on intra-specific features of echolocation calls. Identification of species and genera in audio files can be attempted either manually or through software which performs a fully automated discrimination between species. However, significant overlap in various features (e.g. frequencies of calls) exists between species and even genera. Species ID is therefore often not an absolute conclusion, but rather an opinion or best guess, as opposed to DNA tests or measurements on external characters of captured bats. To make things even worse, the probability of actually observing a bat of a given species in space and time is ignored when performing bat ultrasound analysis. This study introduces Bayesian approximation through a new method we have named Alternative Bayesian Bat Analysis (ABBA). We show, through a simple proof-of-concept example, the importance of adding information about the local composition of the bat community, hence making informed decisions regarding which species is most likely present in audio files. The superior performance of ABBA is also shown through an example using R code. Here, we use simulated data for three *Pipistrellus* spp., a genus with significant overlap in frequencies, but the code can easily be adapted to other bat species and genera worldwide. ABBA outperformed the non-Bayesian approach for all three species. The rare species in the simulated data set was super-inflated when using the non-Bayesian method. Further the results show, contrarily to common belief, that the frequency dominated by a given species in a data set, depends on the composition of the bat fauna and not just means and SDs reported in the literature. ABBA allows researchers to include all observations in statistical modeling, rather than excluding observations, an approach which can affect the reliability of studies. This study also, to a great extent, explains the poor performance of software attempting automated bat ID. Implementing Bayesian algorithms, and thereby allowing users to interact with the software, should significantly improve their performance.

## 1. Introduction

Ultrasound detectors or “bat detectors” of the heterodyne type have been around since the 1950s and could be used to monitor bat echolocation calls (e.g. Griffin, 1958; Pye, 2020; Sales and Pye, 1974). The English naturalist J. H. D. Hooper (1969) suggested that the audible output from his Holgate ultrasonic receiver (a portable bat detector) could be used to distinguish between different species of bats directly in the field, based on the tuned frequency content and time patterns of the audible output. Hooper’s idea slowly gained acceptance, and essentially the same technique is still used in all basic bat detectors. The heterodyne

technique has the advantage that it works in real time and therefore can be used in the field for instantaneous species recognition (Ahlén, 1981).

The frequency output from a heterodyne bat detector usually cannot be quantified and documented in a meaningful way, and this restricts its use to cases where the exact frequency content of the signal is not critical. However, this problem was overcome by the introduction of broadband bat detectors in the 1980s. Signals were recorded in real time and stored either on tape or digitally for later playback at reduced speed (time expansion). Eventually, full spectrum recordings could be analyzed easily for time- and frequency aspects using a sonograph or directly from the computer screen, using specific software (Ahlén and

\* Corresponding author.

E-mail address: [michaelsen@biometrika.no](mailto:michaelsen@biometrika.no) (T.C. Michaelsen).

<sup>1</sup> Deceased

Baagøe, 1999; Russo and Jones, 2002).

Hooper's idea and the introduction of the ultrasonic monitoring technique literally revolutionized the field of bat research (Parsons and Szwedczak, 2009), but there was, and still is, a fundamental problem. Bat calls, in contrast to e.g., bird-songs (Russo et al., 2018), are not aimed for species recognition at all, but designed to solve specific acoustic issues (Barclay, 1999; Russo et al., 2018; Voigt et al., 2021). This means that different species are likely to use similar calls to solve the same problem, and a single species will use many different calls, because there are many different problems to solve. Hence, there is a large intra-species variation in call frequency, and also a broad inter-specific overlap, which complicates the issue of species identification dramatically (Jones and Holderied, 2007; Obrist, 1995; Rydell, 1990). Because of the complexity of the analysis necessary, it was hard to avoid subjective aspects in the identification of species. To try to overcome this problem, species identification based on machine learning was introduced (Jennings et al., 2008; Parsons and Jones, 2000; Zamora-Gutierrez et al., 2016).

At present, bat species identification is still done manually through the inspection of spectrograms, but the use of automatic identification programs is increasing rapidly, because it speeds up the analysis and allows bat species identification on a large scale to be done by almost anyone (Lewandowski and Specht, 2015). Automated analysis first extracts features from echolocation calls and subsequently use these features in machine learning algorithms to identify species or genera. Tests of the performance of commercially available and frequently used automatic identification programs have revealed variable error rates, depending on species or species group (Russo and Voigt, 2016; Rydell et al., 2017). The error rate was high in some cases, and the overall result raised serious doubts about the reliability of acoustic identification of all except the most easy-to-recognize recordings, and this applies to both automatic and manual identification (Rydell et al., 2017). Hence, machine learning tools which can work exceptionally well when identifying signals from birds due to distinct features in their audio signals (Goëau et al., 2014), are less reliable when classifying bat echolocation calls which may be similar between species (Russo et al., 2018).

There are several guides aimed at assisting researchers and amateur bat-workers to identify bat echolocation calls (e.g. Skiba, 2003; Barataud, 2015; Dietz and Kiefer, 2016). They usually present spectrograms to show typical signals from individual species, and also inform about some of the variation that can be expected within species and genera. None of the guides provide the reader with a mathematical framework for species identification, and discriminant analysis is not applied. Therefore, any species identification based on echolocation calls should be considered an opinion or best guess without a known error rate. This is acknowledged by some authors, which calls for "much caution" when performing species identification based on bat echolocation (e.g. Dietz and Kiefer, 2016). Others are more confident (e.g. Barataud, 2015), but all acknowledge problems in the frequency zones where species overlap. A recent study (Montauban et al., 2021) shows a much greater plasticity in echolocation calls than assumed by the literature (e.g. Barataud, 2015). Hence, the general literature on bat species identification cannot be used with absolute confidence (e.g. Montauban et al., 2021).

Software performing automated identification of bats is in many ways the opposite and fully relies on automatic measurement, followed by discriminant analysis. However, developers of the machines often highlight the need for manual "verification" of any ID made by the software (e.g. Wildlife Acoustics, 2018). Therefore, in the end, the result from automated analysis needs to be "confirmed" through manual analysis, and by default inherits its drawbacks and uncertainties. A short overview of the two methods used to identify bats, with pros and cons, is presented in Table 1.

When publishing data in scientific journals, there are several options on how to deal with uncertainties related to manual ultrasound analysis, but authors, editors and reviewers may reach different conclusions. Whatever the final decision may be, it will affect the statistical analysis, and thereby the results and conclusion of a study. For this reason, a

**Table 1**

A short description of the two main approaches to identify bats in ultrasound recordings. Pros and cons are included.

Method	Main features	Pros	Cons
Manual analysis	Analysis is performed by inspecting signals presented in spectrograms, sometimes even measuring various parameters.	Fairly high accuracy if the analysis is performed by an expert.	Time consuming. Subjective conclusion which is difficult to replicate. Years of experience is needed to become an expert.
Automated analysis	Parameters of echolocation calls are automatically extracted and machine learning is used to ID bat species based on these parameters.	Time efficient and species ID does not require an expert level (in theory).	Relatively poor performance for some species, which also depends on manual verification by an expert.

scientific approach to ultrasound analysis is much needed.

The purpose of this work is to introduce simple Bayesian approximation (see Price, 1763) in the process of acoustic identification of bats, and the main focus is the application of such analysis in scientific work, including the development of automated bat identification software. The conclusions of this study (and the philosophy of Bayesian statistics) are also highly relevant to anyone working with bat ultrasound, such as amateur bat-workers. For illustration purposes, we use the relatively easy-to-recognize Scandinavian *Pipistrellus* spp. (Ahlén and Baagøe, 1999; Ahlén and Baagøe, 2004; Rydell et al., 2017), but the principles behind the method can be applied to any bat species and genera worldwide.

## 2. Materials and methods

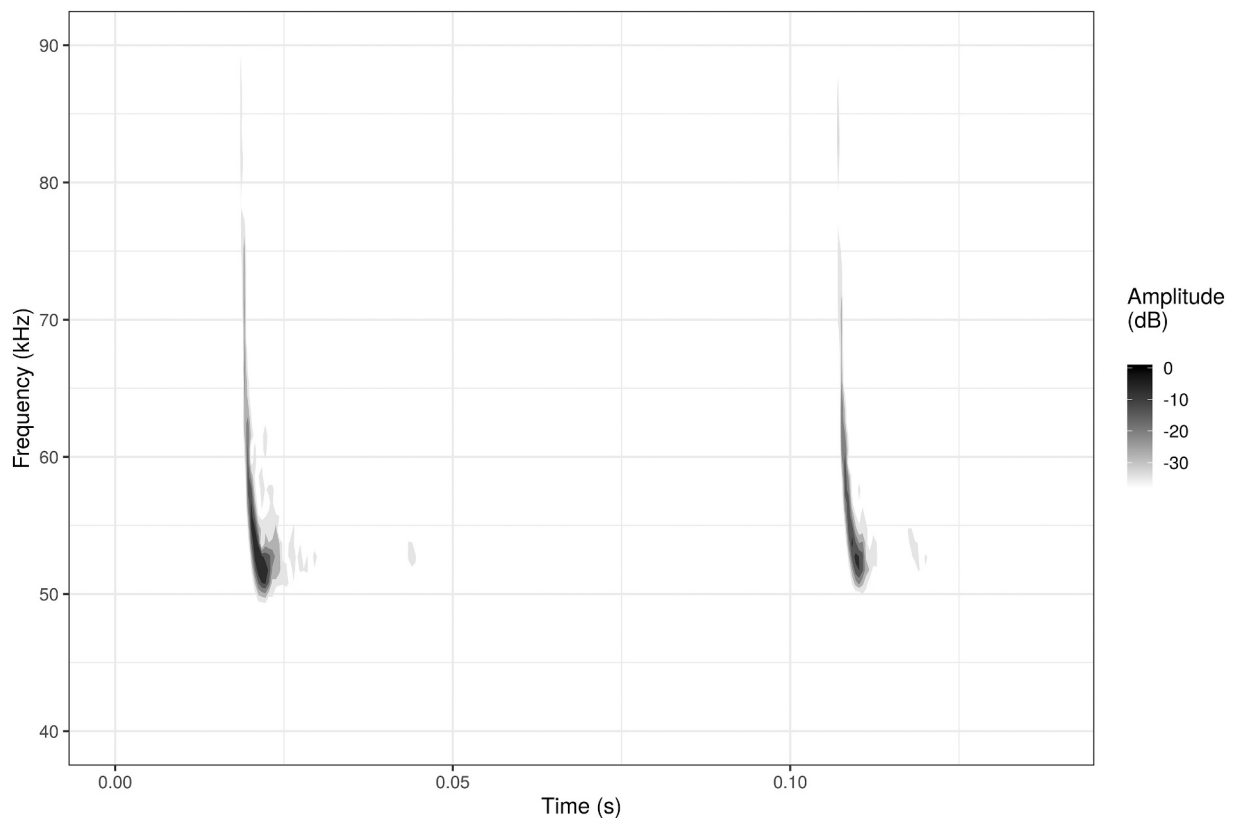
### 2.1. Features of ultrasound analysis

Identification of bat echolocation calls is generally based on one or more measurable characteristics of individual ultrasonic pulses. This can include parameters such as the frequency of maximum energy ( $f_{\max}$ ), start- and end frequencies (band-width), pulse intervals and pulse shape (e.g. Ahlén and Baagøe, 1999; Barataud, 2015; Skiba, 2003). Two pulses produced by a *Pipistrellus* sp. are shown in a spectrogram in Fig. 1.

Energy distribution in the signals is generally considered the most important feature when discriminating between echolocation calls of Scandinavian species in the *Pipistrellus* genus (e.g. Ahlén and Baagøe, 2004). Some authors even use  $f_{\max}$  when assigning popular names to these species (e.g. "the 45 kHz pipistrelle", Altringham, 2003). Other features could be used as supporting evidence, but these are less reliable. Barataud (2015) shows the relationship between various parameters in *Pipistrellus* spp. in scatter-plots, but it is only feasible to speak of a weak or modest correlation within a sample of several observations with a significant residual deviance. Correlations cannot be used to identify bats in individual recordings, as opposed to discriminant analysis. We will focus only on  $f_{\max}$  in this study, but acknowledge that other features may provide supporting evidence when analyzing bat ultrasound.

### 2.2. Bayesian statistics and ultrasound analysis

All statistics and production of figures were done in GNU R version 3.6.1 (R Core Team, 2019) or QGIS version 3.16 (QGIS Development Team, 2018). To show the obvious benefits of considering additional information about bats found outside the realm of the ultrasound itself, examples are given in the form of a simple artwork, and also through R-code. The simple example is self-explanatory, whereas the R-code is explained in some detail. The full R-code, written in Rmarkdown with user guidance, can be found in supplement S1. This supplement will also



**Fig. 1.** Two echolocation calls produced by a *Pipistrellus* sp. at 62 degrees north in western Norway. Only one distribution of frequency of maximum energy ( $f_{\max}$ ) has been observed in the *Pipistrellus* genus in this study area during the summer, with a mean of 54.3 kHz and SD = 2.7 (range approximately 46 kHz to 65 kHz, Michaelsen, 2016). Therefore, only *P. pygmaeus* is believed to be present at these latitudes this time of year. When assigning an ID to a bat in a recording as shown in the spectrogram through manual ultrasound analysis, a researcher would use one or several characteristics measurable in the calls to reach a conclusion. These signals are “club-shaped”, which is typical for *Pipistrellus* spp., and  $f_{\max}$  is around 52–53 kHz (amplitude close to 0). Both calls are believed to stem from *P. pygmaeus*. The spectrogram was made using the seewave package in R (Sueur, 2018).

reproduce data and figures presented in this study.

To produce a probability distribution (0–1) for  $f_{\max}$  of each of the three *Pipistrellus* spp. considered in this study, means and SDs from the literature were used. Here, published data from Michaelsen (2016, *P. pygmaeus*, mean = 54.3, SD = 2.7), Barlow and Jones (1999, *P. pipistrellus*, from the site Bleaton Hallet, mean = 45.3, SD = 1.7) and Barataud (2015, *P. nathusii*, mean = 40.3, SD = 1.2) was added to the R-code. This is the only information used to produce the probability distribution when applying the non-Bayesian method, and it is as far as manual ultrasound analysis will take its users. It will be referred to as the *uninformative prior* in this paper. With the Bayesian method, an additional *informative prior* is added, where the composition of the bat fauna is included. Here we will assume that bat capture data (or any other method applicable to estimate population sizes) suggest a composition of 20% *P. nathusii*, 1% *P. pipistrellus* and 79% *P. pygmaeus*, and in a hypothetical study area. The outcome when using uninformative and informative priors is presented in a figure illustrating the differences between the two approaches. The R-code in supplement S1 will also store the data in .csv format where it can be studied in detail.

Further, to test the performance of the two alternative approaches, populations of each species was simulated using a random normal sample through the `rnorm` function in R. This data set will from now on be referred to as the “*rnorm* sample”. A total of  $N = 10,000$  observations was created based on the means, SDs and the theoretical composition of each species listed above (20%, 1% and 79%), producing 2000 *P. nathusii*, 100 *P. pipistrellus* and 7900 *P. pygmaeus* observations. Simulated data for each species is stored in .csv format when running the code in supplement S1. The *rnorm* sample has been made fully reproducible through the `set.seed` function in R. Predictions were made on the

full sample ( $N = 10,000$ ) using both uninformative and informative priors through the R-code (supplement S1). The outcome is presented in a figure, but is also stored as a .csv file when running the R-code, hence, the number of errors at each frequency can be studied in detail.

Authors, reviewers and editors may consider removing observations which cannot be determined to species with a high degree of certainty as an acceptable solution. To compare the outcome on the full *rnorm* sample ( $N = 10,000$ ) if all observation in the frequency zone where *Pipistrellus* spp. overlap were to be excluded, we again used the same literature used to find estimates of means and SDs for the three species. Barataud (2015) was used to find the upper limit of *P. nathusii* (appr. 43 kHz), Barlow and Jones (1999, all locations) to find the range of *P. pipistrellus* (41–50 kHz), and Michaelsen (2016) to find the lower frequency of *P. pygmaeus* (46 kHz). We acknowledge that the true extremes are likely to be greater (e.g. Montauban et al., 2021), but for the purpose of this example, these data will be considered suitable.

### 2.3. Bayes' theorem and the R code

The simple example illustrating the importance of acknowledging population sizes, uses the standard Bayes formula;

$$P(H|E) = \frac{P(H) \cdot P(E|H)}{P(H) \cdot P(E|H) + P(-H) \cdot P(E|-H)}$$

which states that the probability of the hypothesis being true (H) given the evidence (E) or simply  $P(H|E)$ , depends on the probability of the hypothesis  $P(H)$  being true when observing the evidence ( $P(E|H)$ ), or simply  $P(H) \cdot P(E|H)$ , divided by the probability of seeing the evidence

when the hypothesis is true  $P(H) \cdot P(E|H)$  and when the hypothesis is not true  $P(\neg H) \cdot P(E|\neg H)$ . The *posterior belief*, or  $P(H|E)$ , therefore depends not only on the evidence supporting the hypothesis, but also on the evidence when the hypothesis is not true. The full model used in the R-code (supplement S1) can be described as follows;  $c_i$  is the  $f_{\max}$  of the  $i$ th call,  $s_i$  is the (unknown) species that made the  $i$ th call,  $\mu_j$  and  $\sigma_j$  are constants defining the  $f_{\max}$  distribution of the  $j$ th species, and  $p_j$  is the prior probability over what species that made the call.

$$c_i \sim \text{Normal}(\mu_{s_i}, \sigma_{s_i})$$

$$s_i \sim \text{Categorical}(p_1, p_2, p_3)$$

$$p_1 = 0.79, p_2 = 0.01, p_3 = 0.20$$

$$\mu_1 = 54.3, \mu_2 = 45.3, \mu_3 = 40.3$$

$$\sigma_1 = 2.7, \sigma_2 = 1.7, \sigma_3 = 1.2$$

1 : P.pygmaeus, 2 : P.pipistrellus, 3 : P.nathusii

When the R-code makes predictions using only the uninformative prior (for comparison),  $p_1 - p_3$  is set to 1, but is otherwise the same (supplement S1).

#### 2.4. A simple and illustrative example of Bayes' theorem

Imagine a recording of a *Pipistrellus* sp., where a researcher uses published means of  $f_{\max}$  to determine which species is most likely to have produced the echolocation calls. Based on this single piece of information, the researcher concludes that the calls are in the zone where  $f_{\max}$  of *P. pygmaeus* and *P. pipistrellus* overlap. However, the researcher observes that the recording may fit *P. pipistrellus* better due to relatively low frequencies (say somewhere in the range 48–51 kHz), and therefore estimates that approximately 30% of *P. pipistrellus* are likely to produce pulses at such frequencies during their nightly hunting bout, whereas only 10% of *P. pygmaeus* will do the same. Based on this limited piece of information, it is reasonable to assume that *P. pipistrellus* is the most

probable culprit behind the echolocation calls in the recording. This is as far as classical ultrasound analysis will take its users.

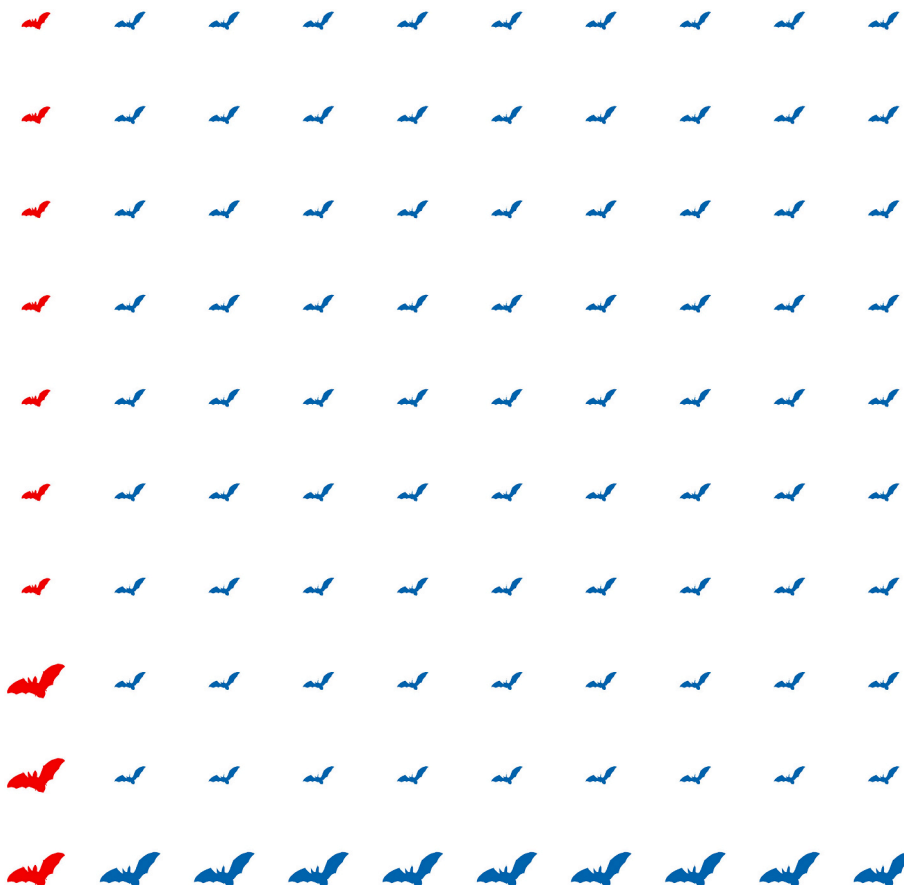
Imagine now that the researcher learns that the recording was made in a finite population comprised of 10 *P. pipistrellus* and 90 *P. pygmaeus*. The observer also assumes that each of the 100 bats are equally likely to be recorded, an assumption granted in this theoretical example. It now becomes obvious to the researcher, given this new evidence, that he/she has to update his/her belief of which species is more likely. The observer plots a scheme of the population, highlighting 30% of the  $n = 10$  *P. pipistrellus* and 10% of the  $n = 90$  *P. pygmaeus*, see Fig. 2. This shows that in this finite bat population, only  $n = 3$  *P. pipistrellus* are likely to use frequencies as found in the recording, whereas  $n = 9$  *P. pygmaeus* are likely to do the same (Fig. 2). The observer now concludes that the most likely species to have produced the echolocation calls in the recording is *P. pygmaeus*. In fact, according to Bayes' theorem, the probability of this recording belonging to *P. pipistrellus* is only 25% or

$$P(P.pipistrellus|E) = \frac{0.1 \cdot 0.3}{(0.1 \cdot 0.3) + (0.9 \cdot 0.1)} = 0.25$$

*P. pygmaeus* is the most likely bat to have produced the calls with an outcome of 75% or

$$P(P.pygmaeus|E) = \frac{0.9 \cdot 0.1}{(0.9 \cdot 0.1) + (0.1 \cdot 0.3)} = 0.75$$

The obvious takeaway from this simple example is that one should not only focus on a single ultrasound recording to reach a conclusion, but rather acknowledge that other relevant information in the real world can be decisive to the outcome. Current manual and automated methods used to identify bats completely ignores this important issue.



**Fig. 2.** The artwork represents a population of  $N = 100$  bats, where 10% ( $n = 10$ , the left column) are *P. pipistrellus* (all red bats) and 90% ( $n = 90$ ) are *P. pygmaeus* (all blue bats, the 9 columns to the right). In this sample, 30% of *P. pipistrellus* ( $n = 3$ , large red bats) are believed to fit a recording of a *Pipistrellus* sp. bat (see main text), whereas only 10% of *P. pygmaeus* ( $n = 9$ , large blue bats) will do the same. Due to the higher proportion of *P. pygmaeus* relative to *P. pipistrellus* in this population, *P. pygmaeus* is three times as likely to have produced the echolocation calls in the recording. The figure was made using QGIS. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

### 3. Results

#### 3.1. Assigning probabilities using both uninformative and informative priors

Fig. 3 shows the probability distribution for the three *Pipistrellus* spp., using only an uninformative prior (3A) and when adding an informative prior (3B). The R-code in supplement S1 will reproduce this figure and also tabulate and store the probabilities over the entire range of frequencies (with 0.1 kHz accuracy) in a .csv file.

#### 3.2. Predictions on a random sample

The Bayesian method, adding an informative prior considering population sizes, outperformed the approach using only an uninformative prior when assigning species ID at the 1 kHz level to the rnorm sample (Fig. 4). Supplement S1 will reproduce all data presented in Fig. 4. Table 2 shows predictions for each species when using uninformative and informative priors, and also the effect of excluding observations in the frequency zone where *Pipistrellus* spp. overlap.

#### 3.3. Sensitivity

To show the performance of both methods when incorrectly estimating population sizes using the same random normal data set, two tests were made. These are merely intended as examples, and the authors encourage readers to use the code in supplement S1 to investigate other problems and solutions. First, *P. pipistrellus* was overestimated up to 1000% ( $n = 1000$ ) of its true population size, in increments of 100.

These incorrect observations were all taken from *P. pygmaeus* (*P. nathusii* was held constant). Fig. 5 shows the total number of errors for all species combined, with or without an informative prior.

In the second test, the random normal data set was displayed in a histogram on a computer screen, and the size of each peak was measured using a simple ruler. To reduce the accuracy of the population estimates further, the effect of SDs was completely ignored, yielding an incorrect estimate of 30.8% *P. nathusii*, 3.8% *P. pipistrellus* and 65.6% *P. pygmaeus*. With the informative prior, there were a total of 194 errors, whereas the uninformative approach yielded 812. The outcome is shown in Fig. 6.

### 4. Discussion

This study indicates a much better performance when adding an informative prior in the decision-making process. Manual ultrasound analysis and currently available software are both locked in a vacuum, whereas Bayesian statistics is flexible and can include additional relevant information important to the outcome. This important feature of Bayesian statistics allows researchers to optimize species assignments based on local conditions. Indeed, the superior performance of Bayesian statistics was highly anticipated. We propose to call this method Alternative Bayesian Bat Analysis (ABBA), recognizing that the evidence can come from multiple sources of information and not just ultrasound.

Several other conclusions can be reached based on the outcome of this study, and they are relevant to both professionals and amateur batworkers alike. First, a keen reader will recognize, by observing Fig. 3A and B, that the more similar the proportion of different species are in the population, the less need there would be for Bayesian statistics. Ultimately, if population sizes were equal between the three species (an

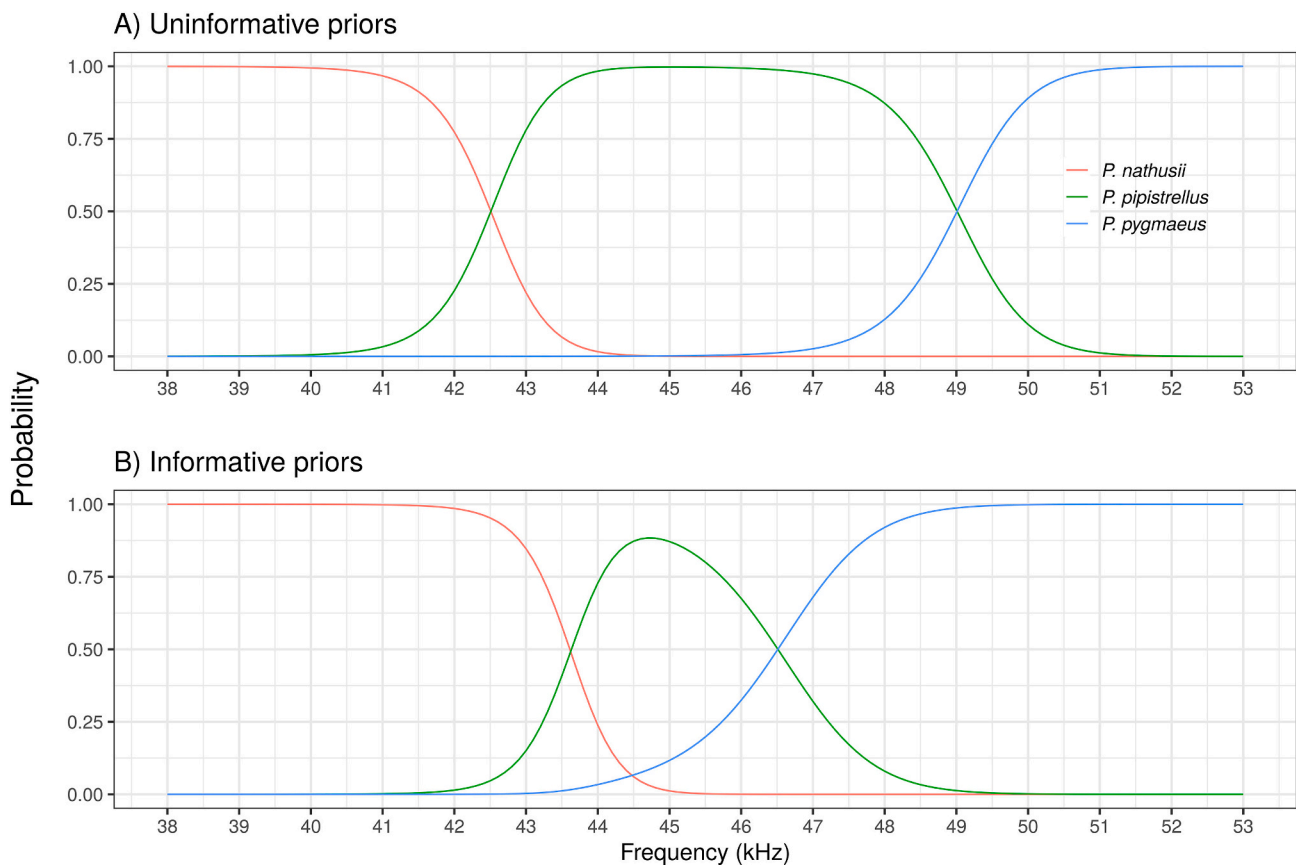
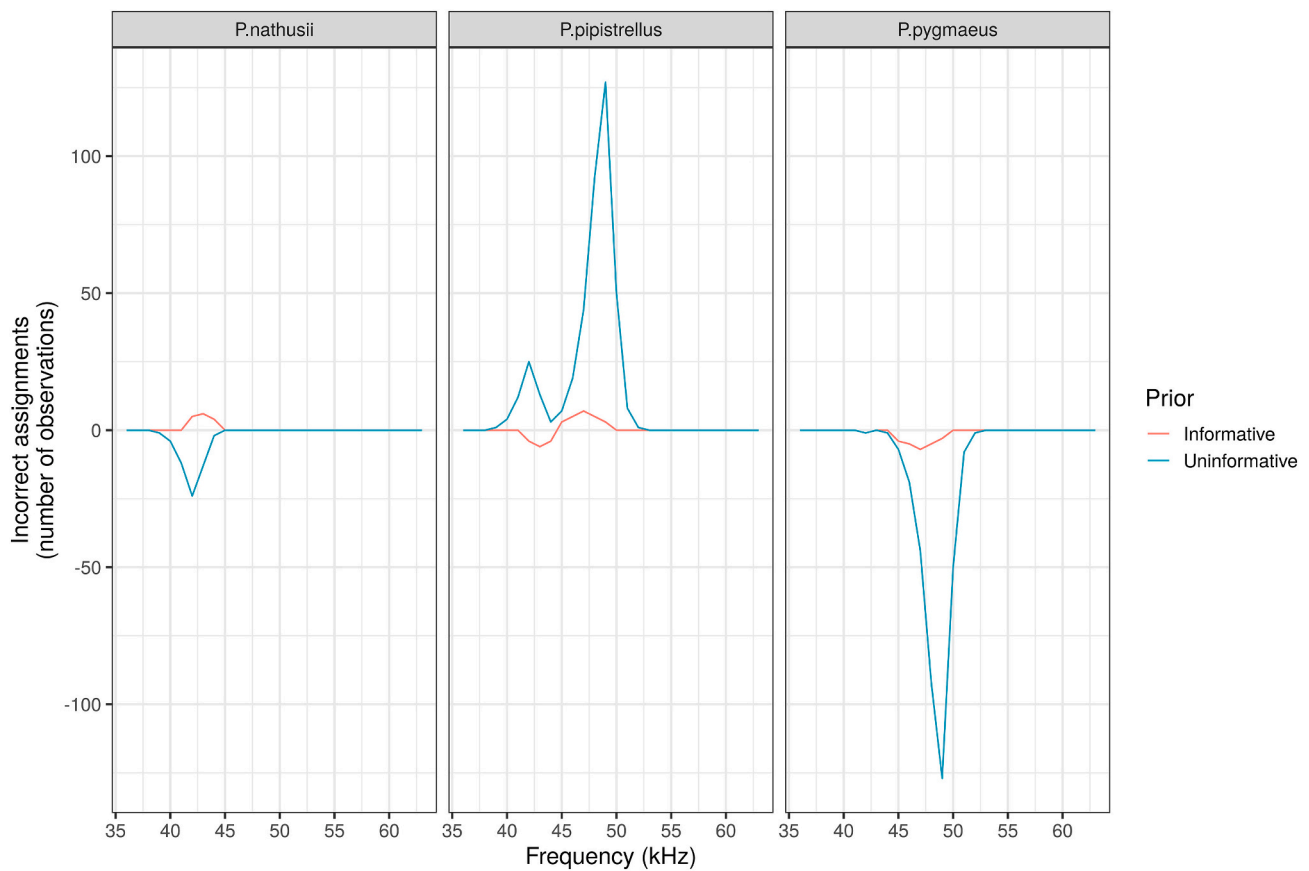


Fig. 3. The figure shows probabilities of three *Pipistrellus* spp. between 38 and 53 kHz when A) considering only means and SDs (uninformative prior) and B) when additional information about population sizes of the three species are considered (informative prior). Here, the informative prior assumes that *P. nathusii* comprise 20%, *P. pipistrellus* 1% and *P. pygmaeus* 79% of the total *Pipistrellus* population. One should note that the probabilities only approach 0 and 1 (not easily seen in the figure).



**Fig. 4.** Incorrect assignments of the three *Pipistrellus* spp. in the frequency range where they commonly overlap, when using only means and SDs to reach a conclusion (uninformative prior) and when adding information regarding population sizes (informative prior). The predictions are made on a simulated data set created using the `rnorm` function in R, where *P. nathusii* comprise 20% ( $n = 2000$ ), *P. pipistrellus* 1% ( $n = 100$ ) and *P. pygmaeus* 79% ( $n = 7900$ ) of the total *Pipistrellus* population ( $N = 10,000$ ).

**Table 2**

Predicted population sizes for all three *Pipistrellus* spp. when using an uninformative prior (means and SDs only), and when adding an informative prior (population sizes). The table also shows remaining observations of the `rnorm` sample when adopting a method of exclusion based on overlapping frequencies. Here we used available literature (see Materials and methods) to obtain the frequencies of each species, with an upper frequency for *P. nathusii* at 43 kHz, a range of 41–50 kHz for *P. pipistrellus* and a lower frequency of 46 kHz for *P. pygmaeus*. In parenthesis is the percentage of the original `rnorm` sample when applying these methods. True population sizes from the random normal sample are shown in parenthesis in the heading after the species names.

	<i>P. nathusii</i> (n = 2000)	<i>P. pipistrellus</i> (n = 100)	<i>P. pygmaeus</i> (n = 7900)
Uninformative	1944 (97.2%)	506 (506.0%)	7550 (95.6%)
Informative	2015 (100.8%)	109 (109.0%)	7875 (99.7%)
Exclusion	1453 (72.7%)	42 (42.0%)	7004 (88.7%)

unlikely scenario in most study areas), Fig. 3A and B would be identical. Further, the vigilant observer will recognize, by comparing Fig. 3A and B, that the frequency most likely to be dominated by a given species when using the Bayesian method, will depend on the composition of the bat fauna. For instance, the frequency dominated by *P. pipistrellus*, is below 45 kHz in this example, and does not coincide with the mean reported in the literature and used in the random normal sample (45.3 kHz). A similar conclusion is true for *P. nathusii*, where the highest probability is found at the lowest frequency observed in this 10,000 strong `rnorm` sample, with the opposite being true for *P. pygmaeus*. Although means and SDs of frequencies reported in the literature are useful in describing a distribution of frequencies of any given species,

the conclusion regarding the frequency dominated by a given species in a data set, can be influenced by features outside the realm of ultrasound analysis. This is true both in the example presented in Fig. 3 and in the real world.

Only a minute proportion of the different bat species have had their echolocation calls recorded and subsequently been identified through captures and/or DNA. The extreme frequency values that bats can use is unknown. Although some authors may attempt to set limits in features such as  $f_{maxe}$  for the *Pipistrellus* genus (Barataud, 2015), studies show that at least some individuals will by far exceed such tentative limits (e.g. Michaelsen, 2016; Montauban et al., 2021). Without knowledge of the most extreme values of  $f_{maxe}$ , the R-script has been coded to have no upper or lower frequency limits of  $f_{maxe}$  for any of the three species. The predicted probability only approaches 0 and 1, but never reaches this value in a situation with multiple species. It is highly likely that such an upper or lower endpoint of frequencies does exist, but (a lot) more data on the potential variation in these species must be examined to identify such approximate values (but see Montauban et al., 2021 regarding *P. pygmaeus*). That being said, never reaching the values 0 and 1 is a key feature of the R-code, and it underlines that the outcome is probabilities and not absolute IDs. It, therefore, is inherently different from general ultrasound analysis, where species ID is the ultimate goal.

In zoogeography, finding a number of “typical recordings of a species” may be sufficient to determine its presence, and all uncertain observations can be discarded. The same may also be acceptable in some binary models in scientific studies, although there will be caveats when sample sizes are very small. Assuming the same in behavioural studies, when the number of bats in space and time matters (e.g. any models with Poisson errors, proportion models with binary errors, or even classical

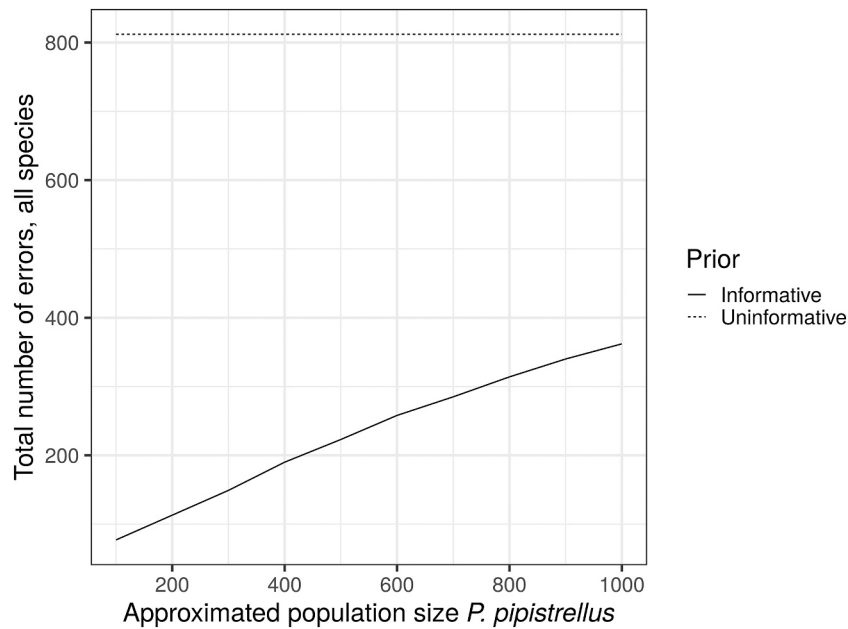


Fig. 5. The total number of errors for all three species combined when overestimating *P. pipistrellus* at the cost of *P. pygmaeus*. The x-axis shows the estimated numbers of *P. pipistrellus* ( $n = 100$  is a perfect estimate,  $n = 1000$  is an extreme overestimate) and the y-axis represents the total number of errors.

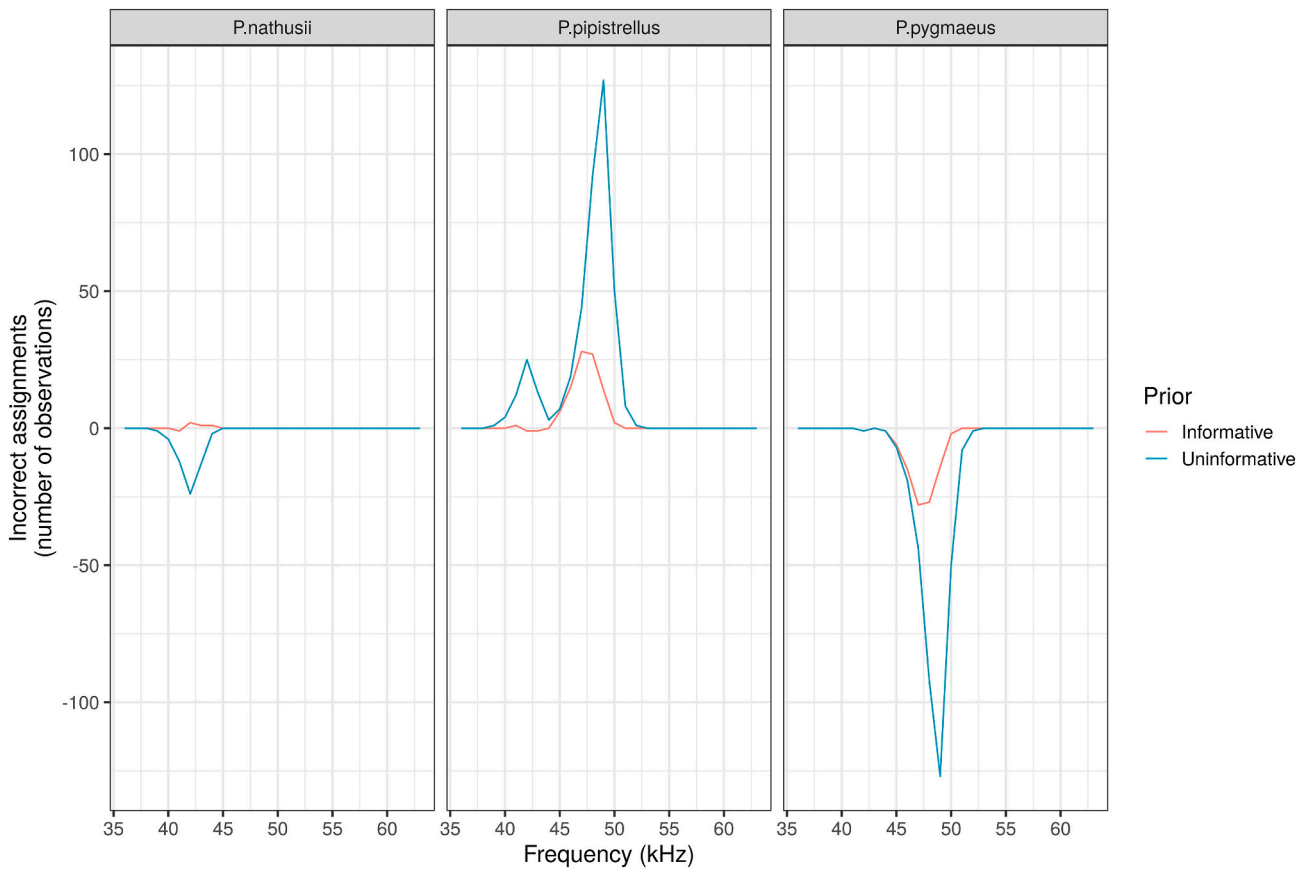


Fig. 6. Incorrect assignments of the three *Pipistrellus* spp. in the frequency range where they commonly overlap. In this example, population estimates, the informative prior, was obtained by measuring peaks in a histogram on a computer screen. The importance of SDs was completely ignored to further reduce the accuracy of the population predictions.

tests), would be hazardous, since the ultrasound bats produce, to some extent is guided by their behaviour (e.g. Ahlén and Baagøe, 2004; Barataud, 2015). Bats performing certain behaviours in certain habitats,

may be excluded from data sets more frequently, or even incorrectly classified, due to the adoption of higher or lower frequencies. The R-code (supplement S1) allows users to take such habitat-guided

behaviour into account by adopting different means and SDs (or even distributions by rewriting some R-code) depending on local relevant features of the study area.

The relatively rare species, *P. pipistrellus* in this example, was blown out of proportion (to around 500% of the actual  $n_{\text{norm}}$  sample) when using an uninformative prior only. An inflation was highly expected when ignoring highly relevant information, and easily predicted when observing the difference between Fig. 3A and B. Hence, this species would unjustifiably be assigned a large number of observations from *P. pygmaeus* (lower tail) and *P. nathusii* (upper tail) when applying manual ultrasound analysis or using automated software. In much of Scandinavia, *P. pipistrellus* is either void (Bjørge et al., 2021, Norway), rare (Baagøe, 2007, Denmark) or possibly an exception in the bat fauna (e.g. Ahlén, 2011, Sweden) and it comprises (at best) only a very small part of a percentage of the *Pipistrellus* individuals found in the region. In this scenario, we predict that *P. pipistrellus* is “super inflated” in publications and in zoogeography. The future may very well conclude that many, or perhaps even all of the “observations” of *P. pipistrellus*, simply represent echolocation rarely used by other members of the genus at northern latitudes in Scandinavia (see Montauban et al., 2021). Similar problems are likely to exist in other bat species and genera around the world, but will not be discussed here.

As shown in this study, when considering the exclusion of observations in the frequency zone where species overlap, the number of excluded observations of each species depends on the species position in the frequency range (and therefore by the occurrence of other similar species). *P. pipistrellus*, the species in the middle in this frequency range, lost more than half of its  $n_{\text{norm}}$  sample (58%), whereas the other two fared much better. Therefore, removing data in many scenarios in bat ecology, will only shift the problem from the ultrasound part of a study and onto the statistics. This will again affect the conclusion and finally the reliability of a study. In fact, if other publications had been used to consider the range of the *P. pygmaeus* (Montauban et al., 2021) all observations of *P. pipistrellus* would have to be removed. A stringent policy of exclusion would therefore make bat studies based on ultrasound unsuitable for some genera.

Adopting the laborious method of assigning probabilities to all observations is not always required. Simply plotting central measurements of ultrasounds, such as  $f_{\text{maxe}}$ , should reveal an occurrence of populations of the various species observed in e.g. the *Pipistrellus* genus. For instance, when plotting  $f_{\text{maxe}}$  of *Pipistrellus* spp. recorded in central Norway during summer, there is only one distribution with a mean around 54 kHz and a range from appr. 46–65 kHz (Michaelsen, 2016). There is no distribution of  $f_{\text{maxe}}$  around 45 kHz and no lower tail of  $f_{\text{maxe}}$  below 45 kHz has been observed. One should not assume that bats change their behaviour when recorded by an ultrasound detector, thus “hiding” in the upper or lower frequencies of similar overlapping species. This happens in sub-atomic particles in quantum physics, but would be scientific nonsense when considering bats. Providing scientific evidence, such as a distribution of frequencies, would certainly reduce type I errors, and thereby avoid “observing” species that are not there, and this simple method should be adopted whenever possible. The use of such proven scientific tools has largely been ignored in bat ultrasound analysis (but see Michaelsen, 2016).

The focus of this study was the *Pipistrellus* genus in Scandinavia, but the R-code can be adopted to any species and genera worldwide. This approach is not locked to one specific feature, and  $f_{\text{maxe}}$  can be replaced with any other measurements if deemed more appropriate to the task at hand (e.g. Hüpkes, 2016). We used a simple normal distribution when considering frequencies, but the code can be changed to acknowledge skew in the data (if relevant). Also, users can add more features to the code (e.g. pulse interval and band-width), which would probably increase correct classification at the cost of making the analysis more laborious. Further, ABBA can, in the future, be used in combination with less traditional, but certainly promising ways of dealing with bat ultrasound (e.g. Heim et al., 2020; Paumen et al., 2021; Tabak et al.,

2022). The most important feature of the code is that all parameters depend on important local conditions which cannot be ignored (e.g. Russo et al., 2018; Russo and Jones, 2002; Voigt et al., 2021). Acknowledging that no study areas are the same would demand more from a researcher in terms of understanding the local bat fauna. ABBA and the R-code in supplement S1 can be used to prepare data to be used in statistical modeling, although the researcher would have to acknowledge that the species assignments are probabilities and not absolutes. Manual species assignments may be considered by many researchers as being absolute IDs, but the evidence suggests otherwise (Rydell et al., 2017, Montauban et al., 2021, this study). Hence, Bayesian approximation should not be considered inferior to manual analysis in terms of uncertainty.

When working with multi-genera problems with several predictive features, a manual approach would be extremely time consuming and not a viable option for most researchers (Russo and Jones, 2002). To solve such problems, automation is required. Current algorithms used in software attempting automated species identification of bats assume that all study areas are the same in terms of species composition. The poor performance of such software (Russo and Voigt, 2016; Rydell et al., 2017) can probably to a great extent be explained by this feature alone. As shown in this study, significant improvements can be made if users are allowed to interact with algorithms through the introduction of informative priors determined by local conditions. Hence, Bayesian algorithms, with more user control, are likely to significantly increase the performance of such software tools in the future. Software could also aid in predicting means, SDs and finally population sizes based on the observed data (e.g. Bhattacharya, 1967), thus reducing manual labour to a minimum.

Ideally (and somewhat simplified), future software performing automated ID of bats should divide the process of identification into several steps after extracting relevant features from the ultrasounds. First, an algorithm should sort out species with distinct features which allows for certain identification, and pool together the remaining recordings into groups based on similarity (e.g. *Pipistrellus* spp., *Myotis* spp. or any other meaningful constellations). Second, the software should estimate means, SDs and finally population sizes for each species in each group based on user input, but always allow researchers to override any such machine-made estimates and use custom predictions. Third, the software should make predictions using classical machine learning tools and then make adjustments to probabilities based on the population sizes of the species in each group (the informative prior). Alternatively, developers could expand on the code in supplement S1 to implement more predictor variables and write a custom algorithm. Finally, the classification accuracy of software incorporating an informative prior in their algorithm should not be determined through training and testing data sets alone. The performance must also be evaluated through the number of correct classifications using real world examples. Current software attempting to ID bats are locked in a vacuum. Bats in the real world are not!

#### Authors' contributions

TCM general idea, R-code and first draft, RB original Bayesian R-code, KHJ R-code and supplement. TCM, RB and KHJ contributed critically to the drafts, including theory and practical implementation and gave their final approval for publication. JR made a significant contribution to the manuscript before passing away. None of the authors have any conflict of interest.

#### Declaration of Competing Interest

None.



## Acknowledgments

We would like to thank two anonymous reviewers for helpful comments and suggestions. Also many thanks to Lidia Kouzguinova and Yuri Makassiouk for checking our grammar. Only free software was used in this study, including GNU/Linux, GNU R, QGIS and LibreOffice. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

In loving memory of Jens Rydell (1953–2021).

## Appendix A. Supplementary data

All R-codes used in this publication can be found in supplement S1. Simulated data, predictions and figures/tables are fully reproducible. The authors are preparing an R-package which will implement the code in supplement S1. Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ecoinf.2022.101721>.

## References

- Ahlén, I., 1981. Identification of Scandinavian Bats by their Sounds. Swedish University of Agricultural Science, Department of Wildlife Ecology.
- Ahlén, I., 2011. The bat fauna of Sweden. Present knowledge on distribution and status. *Fauna och flora* 106, 2–19 (In Swedish with English summary).
- Ahlén, I., Baagøe, H.J., 1999. Use of ultrasound detectors for bat studies in Europe: experiences from field identification, surveys, and monitoring. *Acta Chiropterol.* 1, 137–150.
- Ahlén, I., Baagøe, H.J., 2004. Skillnader mellan Skandinavians tre *Pipistrellus*-arter. *Fauna och flora* 99, 18 (In Swedish with English summary).
- Altringham, J.D., 2003. *British Bats*. HarperCollins Publishers, London.
- Baagøe, H.J., 2007. *Pipistrellus pipistrellus* sensu stricto (Schreber, 1774). In: Baagøe, H.J., Jensen, T.S. *Dansk Pattedyratlas* (Eds.), Gyldendal, pp. 74–75 (In Danish).
- Barataud, M., 2015. Acoustic ecology of European bats. In: *Species Identification, Study of their Habitats and Foraging Behaviour*. Biotope - Muséum national d'Historie naturelle, Paris.
- Barclay, R.M.R., 1999. Bats are not birds – a cautionary note on using echolocation calls to identify bats. *J. Mammal.* 80, 290–296. <https://doi.org/10.2307/1383229>.
- Barlow, K.E., Jones, G., 1999. Roosts, echolocation calls and wing morphology of two phonic types of *Pipistrellus pipistrellus*. *Z. Säugetierkd.* 64, 257–268.
- Bhattacharya, C.G., 1967. A simple method of resolution of a distribution into Gaussian components. *Biometrics.* 23, 115–135.
- Bjorge, A., Eldegard, K., Kovacs, K., Støen, O.-G., Syvertsen, P.O., van der Kooij, J., 2021. *Pattedyr: Vurdering av tusseflaggermus Pipistrellus pipistrellus* for Norge. In: *Norsk rødliste for arter 2021*. Artsdatabanken, Trondheim (In Norwegian).
- Dietz, C., Kiefer, A., 2016. *Bats of Britain and Europe*. Bloomsbury, London.
- Goëau, H., Glotin, H., Vellinga, W.-P., Planqué, R., Rauber, A., Joly, A., et al., 2014. LifeCLEF Bird Identification Task 2014. CLEF2014. Sheffield, France. <https://hal.inria.fr/hal-01088829>.
- Griffin, D.R., 1958. *Listening in the Dark*. Yale University Press, New Haven.
- Heim, O., Heim, D.M., Marggraf, L., Voigt, C.C., Zhang, X., Luo, Y., Zheng, J., 2020. Variant maps for bat echolocation call identification algorithms. *Bioacoustics.* 29, 557–571.
- Hooper, J.H.D., 1969. Potential use of a portable receiver for the field identification of flying bats. *Ultrasonics* 7, 177–181.
- Hüpkens, M., 2016. *The Communication Value of Species-Specific Signatures in the Echo Location Signals of European Pipistrelle Bats*. Dissertation (Dr. rer. Nat.), Mathematisch-Naturwissenschaftlichen Fakultät. Eberhard Karls Universität, Tübingen.
- Jennings, N., Parsons, S., Pocock, M.J.O., 2008. Human vs. machine: identification of bat species from their echolocation calls by humans and by artificial neural networks. *Can. J. Zool.* 86, 371–377.
- Jones, G., Holderied, M.W., 2007. Bat echolocation calls: adaptation and convergent evolution. *Proc. Zool. Soc. B* 274, 905–912.
- Lewandowski, E., Specht, H., 2015. Influence of volunteer and project characteristics on data quality of biological surveys. *Conserv. Biol.* 29, 713–723.
- Michaelsen, T.C., 2016. Spatial and temporal distribution of bats (Chiroptera) in bright summer nights. *Anim. Biol.* 16, 65–80.
- Montauban, C., Mas, M., Tuneu-Corral, C., Wangenstein, O.S., Budinski, I., Martí-Carreras, J., Flaquer, C., Puig-Montserrat, X., López-Baucells, A., 2021. Bat echolocation plasticity in allopatry: a call for caution in acoustic identification of *Pipistrellus* sp. *Behav. Ecol. Sociobiol.* 75, 70. <https://doi.org/10.1007/s00265-021-03002-7>.
- Obrist, M., 1995. Flexible bat echolocation: the influence of individual, habitat and conspecifics on sonar signal design. *Behav. Ecol. Sociobiol.* 36, 207–219.
- Parsons, S., Jones, G., 2000. Acoustic identification of 12 species of echolocating bats by discriminant function analysis and artificial neural networks. *J. Exp. Biol.* 203, 2641–2656.
- Parsons, S., Szewczak, J.M., 2009. Detecting, recording, and analyzing the vocalizations of bats. In: Kunz, T.H., Parsons, S. (Eds.), *Ecological and Behavioral Methods for the Study of Bats*, 2nd ed. The Johns Hopkins University Press, Baltimore.
- Paumen, Y., Mälzer, M., Alipek, S., Moll, J., Lüdtké, B., Schauer-Weissahn, H., 2021. Development and test of a bat calls detection and classification method based on convolutional neural networks. <https://doi.org/10.1080/09524622.2021.1978863>.
- Price, R., 1763. An essay towards solving a problem in the doctrine of chances. In: *By the late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S. Phil. Trans.* 43, pp. 370–418.
- Pye, D., 2020. Echolocation signals of bats – a retrospective. *Bioacoustics.* <https://doi.org/10.1080/09524622.2020.1769323>.
- QGIS Development Team, 2018. QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>.
- R Core Team, 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>.
- Russo, D., Jones, G., 2002. Identification of twenty-two bat species (Mammalia: Chiroptera) from Italy by analysis of time-expanded recordings of echolocation calls. *J. Zool.* 258, 91–103.
- Russo, D., Voigt, C.C., 2016. The use of automated identification of bat echolocation calls in acoustic monitoring: a cautionary note for a sound analysis. *Ecol. Indic.* 66, 598–602.
- Russo, D., Ancillotto, L., Jones, G., 2018. Bats are still not birds in the digital era: echolocation call variation and why it matters for bat species identification. *Can. J. Zool.* 96, 63–78.
- Rydell, J., 1990. Behavioural variation in echolocation pulses of the northern bat *Eptesicus nilsoni*. *Ethology* 85, 103–113.
- Rydell, J., Nyman, S., Eklöf, J., Jones, G., Russo, D., 2017. Testing the performances of automated identification of bat echolocation calls: a request for prudence. *Ecol. Indic.* 78, 416–420.
- Sales, G., Pye, D., 1974. *Ultrasonic Communication by Animals*. Chapman and Hall, London.
- Skiba, R., 2003. *Europäische Fläckermause. Die Neue Brehm-Bücherei, Hohenwarsleben, Kennzeichen, Echoortung und Detektoranwendung*.
- Sueur, J., 2018. *Sound Analysis and Synthesis with R*. Springer International Publishing, Cham.
- Tabak, M.A., Murray, K.L., Reed, A.M., Lombardi, J.A., Bay, K.J., 2022. Automated classification of bat echolocation call recordings with artificial intelligence. *Ecol. Inform.* 68 <https://doi.org/10.1016/j.ecoinf.2021.101526>.
- Voigt, C.C., Russo, D., Runkel, V., Goerlitz, H.R., 2021. Limitations of acoustic monitoring at wind turbines to evaluate fatality risk of bats. *Mammal Rev.* 51, 559–570.
- Wildlife Acoustics, 2018. *Kaleidoscope Pro 5 user guide*. Wildlife Acoustics Inc., Maynard, USA. <http://condor.wildlifeacoustics.com/Kaleidoscope.pdf>.
- Zamora-Gutierrez, V., Lopez-Gonzalez, C., MacSwiney Gonzalez, M.C., Fenton, M.B., Jones, G., Kalko, E.K.V., Puechmaille, S.J., Stathopoulos, V., Jones, K.E., 2016. Acoustic identification of Mexican bats based on taxonomic and ecological constraints on call design. *Methods Ecol. Evol.* 7, 1082–1091. <https://doi.org/10.1111/2041-210X.12556>.