An automatic classifier of bat sonotypes around the world

Charlotte Roemer, Jean-François Julien, Yves Bas

An automatic classifier of bat sonotypes around the world

Charlotte Roemer[1,2], Jean-François Julien[1] & Yves Bas[1,2]

[1] Centre d'Ecologie et des Sciences de la Conservation, (Muséum national d'Histoire naturelle, CNRS, Sorbonne Université), Paris, France

[2] CEFE, Univ Montpellier, CNRS, EPHE, IRD, Univ Paul Valéry Montpellier 3, Montpellier, France

Corresponding author:

Charlotte Roemer, Centre d'Ecologie et des Sciences de la Conservation, Muséum National Histoire Naturelle, 43 rue Buffon, 75005 Paris, France, charlotte.roemer1@mnhn.fr

Abstract

1. Bioacoustics is one of the most popular methods in bat research. Bat species are identifiable through their echolocation call features (e.g. peak frequency, duration, bandwidth) but the amounts of recordings to process generally requires the help of machine learning algorithms. Yet, classifiers are only developed in some areas of the world and it may take dozens of years before they are available everywhere because reference calls are still lacking for numerous species. Our goal was to develop a universal classifier that would classify bat sonotypes according to call shape and peak frequency.
2. To achieve this, we first defined eight sonotype categories that cover all bat echolocation shapes worldwide. We then trained a classifier using random forest decision trees with

a database of 1,154,835 labelled sound events containing bat and non-bat sounds from four continents. After classification, we developed a process to group detected sound events according to the probability scores of their predicted sonotype category and their peak frequency. We then tested the performance of our classifier on a different set of recordings originating from five continents.

3. Depending on the bat sonotype tested, the performance (area under ROC curve) of our classifier ranged between 0.77 and 0.99 for low-quality calls (SNR < 25 dB). Performance ranged between 0.89 and 1 for middle or high-quality calls (SNR ≥ 25 dB). The performance for bat feeding buzz classification ranged between 0.93 and 0.98 depending on the SNR. The classifier was not developed to classify bat social calls; the majority of them were classified as a bat sonotype.
4. The classifier is an open data format and can be used by anyone to study bats around the world. It can be used to spot acoustically described species but for which a classifier was not developed, and even to detect species that were not acoustically described yet. The grouping of sound events according to call sonotype and peak frequency may be used to describe bat communities and compare the composition of acoustic niches across time and space. This allows the monitoring of bats and the assessment of bat conservation issues in any region of the world.

Résumé

1. La bioacoustique est l'une des méthodes les plus populaires pour l'étude des Chiroptères. Les espèces de chauves-souris sont identifiables grâce aux propriétés de leurs cris d'écholocation (ex. fréquence dominante, durée, largeur de bande) mais les quantités d'enregistrements à analyser demandent généralement l'aide d'algorithmes en apprentissage machine. Cependant, les classificateurs ne sont développés que dans

certaines régions du monde, et cela pourrait durer des dizaines d'années avant qu'ils ne soient disponibles partout, car les cris de référence manquent encore chez de nombreuses espèces. Notre but était de développer un classificateur universel qui classifierait les sonotypes de chauves-souris en fonction de la forme des cris et des fréquences dominantes.

2. Afin de parvenir à cette fin, nous avons d'abord défini huit catégories de sonotypes qui couvrent toutes les formes d'écholocation à travers le monde. Nous avons ensuite entraîné un classificateur en utilisant des forêts d'arbres décisionnels avec une base de données de 1 154 835 évènements sonores étiquetés contenant des sons de chauves-souris et de non-chauves-souris issus de quatre continents. Après classification, nous avons développé un procédé pour regrouper les évènements sonores détectés en fonction des scores de probabilité de leur catégorie de sonotype prédite et de leur fréquence dominante. Nous avons ensuite testé la performance de notre classificateur sur un jeu différent d'enregistrements issus de cinq continents.
3. Selon le sonotype testé, la performance (aire sous la courbe ROC) de notre classificateur variait entre 0,77 et 0,99 pour les cris de basse qualité (SNR < 25 dB). La performance variait entre 0,89 et 1 pour les cris de qualité moyenne ou haute (SNR > 25 dB). La performance pour les buzz de capture de proie variait entre 0,93 et 0,98 en fonction du SNR. Le classificateur n'a pas été développé pour classer les cris sociaux ; la majorité d'entre eux ont été classés en sonotype de chauve-souris.
4. Le classificateur est une donnée ouverte et peut être utilisé par quiconque souhaitant étudier les chauves-souris à travers le monde. Il peut être utilisé pour repérer des espèces décrites acoustiquement mais pour lesquelles il n'existe pas encore de classificateur, et même pour détecter des espèces encore non décrites acoustiquement. Le groupement d'évènement sonores en fonction du sonotype et de la fréquence dominante peut être utilisé pour décrire des communautés de chauves-souris et comparer la composition des

niches acoustiques à travers le temps et l’espace. Cela permet le suivi et l’évaluation de problématiques de conservation des Chiroptères dans n’importe quelle région du monde.

Keywords

Africa; Asia; Automatic ID; bioacoustics; call; classification; Neotropics; Passive Acoustic Monitoring (PAM);

Introduction

Bioacoustics, a growing field in science, brings together a wide range of disciplines, from the inventory of animal species to the characterisation of soundscapes, including sound source tracking and the study of social interactions. Most of these research domains are made possible by the existence of a specific signature inherent in bioacoustic signals, which identifies a species or a group of species uniquely. In fact, animal vocalisations are designed to fulfil different functions, which may be classified into two categories: social communication on one hand (e.g. territorial marking, courtship, group gathering) and echolocation on the other hand (Obrist et al., 2010). For emitters to target conspecific receivers, social vocalisations necessarily bear features shared among individuals of the same species, even if their structure may show individual signatures, for example in mother-pup interactions (Sauvé et al., 2015; Wiley, 2006). This means that one expects at least as many different vocalisations as there is of vocally active species within an ecosystem (Sueur et al., 2012). On the other hand, echolocation signals, because of their primary function of object location and recognition for orientation or foraging, are much more subject to evolutionary convergences, and thus display much less diversity (Jones and Holderied, 2007; Schnitzler et al., 2003). For example, cetacean clicks resemble feeding buzzes of bats capturing prey (Jones, 2005; Madsen and Surlykke, 2013).

Nonetheless, unlike in dolphins, the numerous foraging strategies extant in bats today, tied to specific echolocation call structures, make the unique identification of most bat species in a community possible thanks to the differences in the frequency, duration and shape of their echolocation calls (Au, 1997; Walters et al., 2013). Indeed, among the more than 1,400 extant bat species, food resources include fruits, insects, nectar, vertebrates, fish and blood (Wilson and Mittermeier, 2019). Even among species exploiting a similar trophic resource (e.g. insectivorous bats), adaptation to prey led to different echolocation and foraging strategies.

Three distinct signal structures are used by bats, each suited for a specific task: narrowband signals for long-range detection of the target, broadband signals for target localisation and classification, and long constant frequency signals with Doppler-shift compensation for detection and classification of fluttering insects (Schnitzler and Kalko, 2001). From these three categories, a multitude of combinations are used by bats (Collen, 2012; Jones and Teeling, 2006). These different combinations can be referred to as 'sonotypes' (Fidelino and Gan, 2019; Fraser et al., 2020; López-Baucells et al., 2019; Ochoa et al., 2000), independently of signal frequency.

Acoustic monitoring of bats, thanks to its high cost-effectiveness, is gaining popularity all over the world, among scientists, conservation organisations, land managers or private consultancies. Bat inventories are carried out to study species richness and abundance, which necessitates several nights to obtain a near to exhaustive assessment (Fraser et al., 2020; Richardson et al., 2019). With dozens of nights of data accumulated across different study sites, the help of automated acoustic identification becomes necessary.

Several tools using machine learning were developed in the last years to detect and identify species of a given country or a biogeographical area (Bas et al., 2017; Chen et al., 2020; Kobayashi et al., 2021; Mac Aodha et al., 2018; Nocera et al., 2019; Obrist and Boesch, 2018; Rydell et al., 2017; Zamora-Gutierrez et al., 2016). This automated identification process can be used with success when combined with either manual validation or a statistical sorting based on the associated confidence indexes, depending on the objectives of the study (Barré et al., 2019; López-Baucells et al., 2019). Obviously, it is only possible to identify a species if its reference calls are present in the training set of the classifier, which restricts the usage of these softwares to specific areas. Free or commercially available bat classifiers currently only cover the Neotropics, North America and/or Europe. These developments are correlated with the degree of knowledge available in different regions of the world, and it might take several decades before auto-ID softwares are made available for Africa, Asia, South-America or Australia (Walters et al., 2013). Yet, conservation issues are such that acoustic monitoring is utterly needed to assess the state of bat populations and design conservation plans accordingly.

Our goal was to build a universal bat classifier that could be used anywhere in the world. To train a taxonomic classifier for all extant species, it is necessary to possess a sample of all combinations of call sonotypes and frequencies used by bats around the world. Lacking this resource, we chose to approach this task in two steps: first training a classifier with bat sonotypes independently of call frequency, and then grouping these sonotypes after classification according to their frequency.

Choosing a universal definition of bat call types – or sonotypes – is a very difficult enterprise, although some attempts have been carried out, such as in Jones and Teeling (2006). However, this classification arbitrarily distinguishes long from short calls, when call duration is greatly influenced by the echolocation task performed (e.g. foraging vs. commuting) (Holderied, 2006), which may lead to confusions. Moreover, this basic classification does not represent the full diversity of echolocation calls. For instance, it is not clear how the echolocation calls of *Pteronotus davyi* (constant frequency followed by frequency modulation and constant frequency again) should be classified according to this study. A second attempt by Collen (2012) started from the latter study and added more classes. Here again, this classification is not satisfactory, because it also uses the criterion of call duration, and because some species may be classified in several of those categories depending on the echolocation task they perform.

Therefore, a novel approach is needed to guarantee exhaustiveness and avoid any confusion.

Gathering sound references covering all combinations between acoustic types and frequency domains existing in the world is not possible. However, gathering a sufficiently large diversity of acoustic types so that bat calls could be robustly identified independently from absolute frequency is a feasible task. Our objectives were thus (1) to develop a comprehensive framework to classify the different bat sonotypes occurring worldwide, (2) to build a classifier of bat sonotypes, (3) after classification, to group calls of similar frequency inside the same acoustic sequence to isolate species recorded simultaneously and displaying the same sonotype, and (4) to test the efficiency of this classifier. The purpose of our tool was to be used for passive or active acoustic monitoring, in which it is common practice to count the number of sequences (i.e. recordings of a certain time interval) containing one or more bat calls of a given species (Fraser et al., 2020).

Material and methods

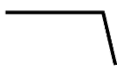
Definition of bat sonotypes

We first conducted a literature review of the diversity of bat sonotypes occurring in the world (Arias-Aguilar et al., 2018; Barataud, 2015; Barataud et al., 2013; Collen, 2012; Fenton and Bell, 1981; Lopez-Baucells et al., 2016) and completed this literature review with our own bat acoustic surveys in Europe, South-East Asia, Central Africa, South-America, the Neotropics and North-America. It appeared that any bat echolocation call may be conveniently divided into a maximum of three consecutive elements, where the main element can be preceded by a prefix and followed by a suffix (see Fig. 1). Each of those elements may contain one of the structures described by Schnitzler and Kalko, (2001), namely a narrowband (quasi-constant frequency, QCF), a broadband (frequency modulation, FM) or a constant frequency (CF) signal, produced by bats to achieve different echolocation tasks. FM may be upward (FMu) or downward (FMd). We thus chose to describe the diversity of sonotypes based on this method and found the occurrence of eight different sonotypes (Table 1 and supplementary file 1). We chose to not create sonotype classes based on the presence of multiharmonics, because harmonics are more or less perceivable depending on recording quality, which may lead to confusions. We however quantified the intensity of potential harmonics with ancillary variables so that users can access this information: we selected the maximum value among the ratios of the average amplitude between the elements whose frequency is 1/2, 1/3, 2/3, 4/3, or twice that of the DSE and the amplitude of the DSE (these ratios are named Ramp); we used the 90 % percentile of this value among the calls of the same groups (see section post-classification grouping). Positive values are usually associated with harmonics.

Figure 1 here

Fig. 1: Illustration of the method for the definition of bat sonotypes with three examples on a sonogram (time as a function of frequency). The upper sonotype is a call divided into a frequency modulated (FM) prefix and a main quasi-constant frequency (QCF) element. The sonotype in the middle is a call containing only a main FM element. The lower sonotype is a call divided into an FM prefix, a main constant frequency (CF) element and an FM suffix.

Table 1: Description of bat sonotypes. FM: Frequency modulated; CF: Constant frequency; QCF: Quasi-constant frequency; d: downward; u: upward. See Fig. 1 for the definition of prefix, main element and suffix.

Sonotype	Prefix	Main element	Suffix	Sonogram	Example species
FMd-QCF	Downward FM or none	QCF	-		*Pipistrellus pipistrellus, Lasurius borealis*
FMu-QCF	Upward FM	QCF	-		*Promops centralis*
QCF-FMd	-	QCF	Downward FM or none		*Peropteryx macrotis, Molossus molossus*
CF-FMd	-	CF	Downward FM		*Hiposideros commersoni, Noctilio leporinus*
FMu-QCF-FMd	Upward FM	QCF	Downward FM		*Cormura brevirostris*

FMu-CF-FMd	Upward FM	CF	Downward FM		*Rhinolophus ferrumequinum, Pteronotus cf. parnellii*
FMd	-	Downward FM	-		*Myotis nattereri, Carollia perspicillata*
CF-FMd-CF	CF	Downward FM	CF		*Pteronotus personatus*

Species used most of the time a single sonotype. If more than one sonotype was displayed by a species – which was the case in *Promops centralis*, *Molossops* sp., *Chaerephon* sp. and *Molossus* sp. – we only labelled the dominant sonotype to build the classifier (see supplementary file 1). Within a sonotype, shapes could vary significantly due to changes in call duration and bandwidth (see example in Fig. 2), but the curvature still corresponded to the description of Table 1. When call duration was extremely short, e.g. in FMd-QCF inferior to 3 ms, call shape necessarily resembled a FMd (Fig. 2), but we still labelled it as FMd-QCF.

Figure 2 here

Fig. 2: Example of shape variation for sonotype "FMd-QCF" represented as sonogram.

Call labelling

Our sound database contains passive and active recordings of free-flying bats as well as recordings of released individuals after capture (individuals were measured and identified in hand). Different acoustic recorders were used for these recordings and they are listed in supplementary file 1. This table also lists the country in which recordings were made. 90 % of the sounds labelled to build the classifier originated from Europe (France, Spain, Croatia, Lithuania, Poland and the United Kingdom) but also from other regions of the world, such as South-America (Uruguay, French Guiana), Central America (Costa Rica), Africa (Benin) and Middle-East (Turkey) (see Fig. 3).

Figure 3 here

Fig. 3: Map of the origin of the call library used to build the classifier. Numbers indicate the number of DSE (detected sound events) labelled and used. Black circles: bat DSE. White circles: non-bat DSE.

We used the Tadarida-L 2.1 software (https://github.com/YvesBas/Tadarida-L) to detect and label reference calls. This software includes a detection function to isolate detected sound events (DSE), originating from a single acoustic source, in both frequency and time (see Bas et al. (2017) for more details). Each species name was then associated with a sonotype in a separate table.

Bat feeding buzzes (Griffin et al., 1960) – a series of more than five calls of very short intervals (less than 10 milliseconds) produced by bats at an attempt of prey capture – were also labelled and constituted an additional acoustic class (different from a sonotype). These sequences are usually preceded by a gradual acceleration of rhythm and followed by a sudden resumption of a similar rhythm to that before the acceleration. Non-bat sounds were also labelled as additional acoustic classes. They include ground-crickets, bush-crickets, grasshoppers, bees, beetles, cicadas, flies, frogs, moths, other insects, other mammals, birds, and noise (electrical or mechanical).

In total, we labelled 321,132 DSE belonging to 9,245 recordings of 121 bat species or groups. We also labelled 833,703 DSE belonging to 13,625 recordings of 153 non-bat species or noise types (see supplementary file 1 in which the column N_DSE indicates the number of DSE labelled).

Building of the classifier

We used Tadarida-C (https://github.com/YvesBas/Tadarida-C) on R (R Core Team, 2013) to assemble the table containing the acoustic parameters of all labelled DSE and to build the classifier. Tadarida-C builds classifiers based on the random forest method for machine learning (see Bas et al. (2017) for more details). We used all of the acoustic features measured by Tadarida-L to build the classification trees (see the list at https://github.com/YvesBas/Tadarida-L/blob/master/Manual_Tadarida-L.odt), except for features directly related to absolute frequency, because we wanted to define sonotypes independently of frequency. Nonetheless, to help distinguish birds from bats in the lowest frequencies, we added a binary feature, which took the value of 1 if the frequency of the master point (the highest amplitude value among the elements within the DSE defines the master point) was superior to 17 kHz where most bat calls and only harmonics of bird calls occur, or 0 in the other case. We kept features related to relative frequency (e.g. bandwidth, which is maximal frequency minus minimum frequency) to build the classification trees.

Classification

We modified Tadarida-C (see Ta_Tc_Sonotype.R and ClassifC1_Sonotype.R) to discard DSE below 8 kHz, which is the lowest peak frequency known to be emitted by a bat (Leonard and Fenton, 1984). We also discarded DSE suspected to be from the same bat call as the previous DSE, but separated by a short silence due to heterogeneous sound propagation. For this, we removed all DSE that were separated from the previous DSE by less than 5 ms.

Before classification, calls suspected to be higher harmonics of another DSE were discarded by excluding calls starting and ending simultaneously to DSE of a lower frequency and a higher amplitude. Nonetheless, the information of a call having a harmonic still exists in Tadarida features. Therefore, bat calls were classified taking into account the presence of their harmonics. To discard DSE suspected to be harmonics, we isolated DSE that occurred simultaneously and only kept the one with the highest amplitude.

We used Tadarida-C to obtain predictions of the acoustic identity (i.e. the bat and non-bat acoustic classes) of each DSE. Each DSE prediction is accompanied by a prediction probability for each of the possible acoustic classes present in the table containing the acoustic parameters of all labelled DSE.

Our R scripts for this section and the next one can be found at [https://doi.org/10.5281/zenodo.5483030] (folder 'Sonotypes').

Post-classification grouping of detected sound events

Our goal was to build a ready-to-use classifier for bat activity surveys. Since the majority of them are based on the quantification of sequences (or files) containing one or more bat calls of the same species (Fraser et al., 2020), we followed the same process. We processed each wav file separately. We modified Tadarida-C (see AggContacts_Sonotype.R) to group DSE after classification. This section aimed to group DSE belonging to the same species according to their classification probability score, following the algorithm of Tadarida-C (Bas et al., 2017), but also according to their peak frequencies (F_{peak}), i.e. each group of DSE in a file will eventually be identified uniquely by an acoustic class combined with a peak frequency. Thus, if several species are present, several groups of DSE are expected.

For this, several rounds were conducted until each DSE were attributed a group (see Fig. 4). At each round, the most probable acoustic class in the file was identified by selecting the best prediction probability score. The acoustic class containing this best score was defined as "dominant" for the current round.

At each round we applied the density function of the stats package of R (with 30% of the default bandwidth) to the F_{peak} of all DSE within the file to obtain their probability distribution and isolate their modes. The presence of different DSE of different frequencies in a file translates into the presence of different modes. For instance, if three species produce calls in three different frequency ranges, three peaks (i.e. modes) will appear in the density plot (see chart in Fig. 4). We then selected the mode closest to the F_{peak} of the DSE with the best prediction probability score and called it "dominant mode". If only one DSE remained per file, the dominant mode took the value of the F_{peak} of the remaining DSE. All DSE with a F_{peak} within a 5 kHz range of the dominant mode and with a probability score in the dominant acoustic class superior to 0.05 were attributed a final ID corresponding to the dominant acoustic class and stored for output. All other DSE were processed in the next round until there was no DSE left. We chose to use this conservative approach to avoid false positives, i.e. sonotype identification supported only by inconsistent probabilities. Before grouping, each DSE was associated to a prediction probability score for each acoustic class; after grouping, the final prediction probability score of a group of DSE of a given acoustic class is the highest score among the DSE of this group for this acoustic class.

Figure 4 here

Fig. 4: Example of the post-classification grouping of detected sound events. All files are processed at each round. DSE: Detected Sound Event. AC: probability score of the Acoustic Class. F_{peak}: frequency at the maximum energy.

Instructions on how to download and how to use the classifier are available in the README file at [https://doi.org/10.5281/zenodo.5483030] (folder 'Sonotypes').

Classifier performance

We tested the efficiency of the classifier on recordings from study sites that were not used to build the classifier. These new recordings originated from six regions of the world, namely Europe (France), North-America (United States of America), Central-America (Costa Rica), South-America (French Guiana), Asia (Cambodia) and Africa (Benin) (see supplementary file 2 and Fig. 5). For each region, we used recordings originating from three different locations. The mean distance between locations within the same country was 241 km (min = 14 km, max = 944 km). On each location, we used full-night or partial-night recordings (i.e. first hours of the night), in which files were cut to have a maximum duration of 5 seconds.

Figure 5 here

Fig. 5: Map of the origin of study sites sampled to assess the performance of the classifier.

The last output of the classifier is a table (see supplementary table 3 at [https://figshare.com/articles/dataset/Validation_table_for_the_bat_sonotype_classifier/15149523] and its column description in supplementary table 4) in which each line corresponds to a group of DSE of the same file which were grouped together according to their acoustic class and their peak frequency (see previous section). To test the efficiency of the classifier for each bat sonotype, for feeding buzzes, and for the most common non-bat classes (bush-crickets, noise and bird), we did a stratified random selection of 5 files per detected acoustic class at each location. The random selection was stratified according to the time of the night and the probability scores to ensure a representativity of the variety of sounds analysed and of the efficiency of the classifier. It could happen that less than 5 files per acoustic class were available. For each file checked, we browsed all acoustic classes detected and noted the occurrence of false positives, true positives, and false negatives. Files could contain the same acoustic class several times but in different frequency modes and we checked each of them. For this, we visualised the file sonograms on Syrinx (John Burt, USA). If the true nature of the acoustic class was ambiguous (e.g. in the case of low signal to noise ratios), to avoid confirmation bias resulting from a personal interpretation of the call identity, we did not classify it as a positive or a negative and left it unchecked. If the file was too noisy to check it without ambiguity from the checker's perspective, or if there were many overlapping calls, we left the file apart and did not check it (see Figure A1 in the supplementary information file for the distribution of SNR across checked and unchecked groups of DSE and see Figure A2 in the supplementary information file for examples of sonograms). We checked files containing obvious bat social calls (see Chaverri et al., 2018) apart and results are presented separately since our sonotype classification was not designed to cover all the complexity of those social calls. For this analysis, we only checked the groups of calls in the file that corresponded to social calls. We checked them as if they were echolocation calls, i.e. if the call had an FMd shape and was classified as such, we considered it a true positive. We classified calls whose shape was not listed in Table 1 as “complex”.

All checked sound sequences are available at [https://figshare.com/articles/media/Sounds_used_to_validate_an_automatic_classifier_of_bat_sonotypes/15141201].

We used receiver operator characteristic (ROC) curves to assess the efficiency of the classifier for the different acoustic classes. These curves are created using the rate of true and false positives. Since our classifier is probabilistic, the ROC curves take the probability of classification into account. As explained in the previous section, the probability of classification of a group is the highest score among the DSE of this group for the predicted acoustic class. We made one ROC curve for each of three different classes of signal-to-noise ratios (SNR) to take into account the recording quality. More precisely, for each acoustic class identified by the classifier within a file, we calculated the median value of calls maximum amplitudes. The median value gives a less important weight to outliers and is thus more representative of the majority of the calls in a file. We then created three amplitude classes: <25 dB, 25-75 dB, and >75 dB SNR. We calculated the area under the curve (AUC) to provide a numerical summary of the performance of the classifier.

The efficiency of the segregation of species according to the frequency modes (see Fig. 4) could not be assessed quantitatively since we do not have a perfect knowledge of bat acoustic identification in all countries sampled. Therefore, we could not assess whether two different species were put in the same frequency mode or not. We thus described whether all calls originating from the same individual – based on the similarity of calls and on the inter-call duration – were classified in the same group (i.e. one species) or if they were classified in several groups of frequency modes (i.e. several species). If calls were seemingly produced by the same individual and yet segregated in two frequency modes, we noted in what circumstances this occurred.

Results

Efficiency of the classifier

The percentage of files in which not a single DSE was found by Tadarida – or which only contained DSE below 8 kHz and were thus discarded by Tadarida – is equal to 7.1 %. In the rest of the dataset, we considered 715 files according to our stratified random sampling design. 54 files were considered separately as they contained social calls. 47 files (444 groups of calls) were too noisy to be checked. In the remaining 614 files, we checked 3,575 groups of calls classified by the classifier (see Table A3 in the supplementary information file). In these files, we noticed 3 false negative groups: three FMd-QCF. In the 614 files that were checked, 239 groups of calls were left unchecked because the true nature of the acoustic class was ambiguous (e.g. in the case of low signal to noise ratios).

ROC curves and their AUC (area under the curve) show that the highest performance of the classifier is for the sonotype FMu-CF-FMd, and that this performance is very little affected by call amplitude (Fig. 6). The classifier has a similar performance for CF-FMd-CF calls, however, the sample size is very low for this sonotype (n=21). The classifier shows the worst performance for the sonotype QCF-FMd when calls have an amplitude lower than 25 dB (AUC = 0.78), but the AUC for this sonotype varies between 0.92 and 0.94 when call amplitude is 25 dB or louder. The classifier has a very high performance for buzzes, and this performance is very little affected by call amplitude (AUC between 0.93 and 0.98).

Figure 6 here

Fig. 6: Receiver operating characteristic (ROC) curves between the confidence score of the false positive rate (FPR) and the true positive rate (TPR) for each acoustic class. Grey shades

represent the median of the maximal amplitude among the calls classified as the sonotype that was checked: light grey <25 dB, dark grey 25-75 dB, black >75 dB. AUC (Area under the curve) is a proxy of the performance of the classifier. N = number of groups of calls. Detailed ROC curves for each country are also provided in the appendix (Figures A3 – A8 in the supplementary information file). Micro- and macro-averaged ROC curves are shown in Figure A9 in the supplementary information file. FM = frequency modulated. QCF = quasi-constant frequency. CF = constant frequency. d = downward. u = upward. The summary of this figure can be found in Table A1 in the supplementary information file.

The confusion matrix (Table 2) shows confusions between acoustic classes. Bush-crickets were classified 47 % of the time as a non-bat (other than bush-cricket), 7 % of the time as a bat sonotype, and 10 % of the time as a buzz. FMd-QCF were classified 19 % of the time as FMd, which happened very often when individuals produced very short calls. QCF-FMd were classified 21 % of the time as FMd-QCF, which happened often for calls with extremely small FMd at the end (short bandwidth). Non-bat DSE were classified 11 % of the time as bat or buzz; bat or buzz DSE were classified 4.6 % of the time as non-bat.

Performance of the frequency-based grouping of detected sound events within acoustic classes

DSE emitted by one bat species were most of the time grouped in one unique frequency group except for three cases: species producing calls with alternating frequencies with a difference of more than 10 kHz in peak frequency (e.g. *Molossus molossus*) appeared in two different frequency groups; buzzes emitted by one species were grouped on average in 1.6 different groups (maximum = 4); FMd calls emitted by one species were grouped on average in 3 different groups (maximum = 7).

Bat social calls

80 % of the 111 checked social calls were classified as a bat sonotype that matched their shape (Table A2 in the supplementary information file). Among the remaining 20 %, 31 calls could not be assigned to one of the classes of Table 1 and we thus described them as “complex calls”. 55 % of these 31 complex calls were classified by the classifier as FMd-QCF. It must be noted

that 67 % of the files containing social calls that were checked belonged to the same study site (Valley of fire, USA) and were social calls of *Tadarida brasiliensis*. If this site is removed, 54 % of the 37 checked social calls were classified by the classifier as a bat sonotype that matched their shape (results not shown). Among the remaining 46 %, 17 calls were “complex calls”. 40 % of these 17 complex calls were classified by the classifier as FMd-QCF.

Table 2: Confusion matrix between acoustic classes identified by the classifier and after manual checking for groups of calls with an SNR (signal to noise ratio) superior or equal to 25 dB. These results do not take the probability index associated with the automated classification into account. FM = frequency modulated. QCF = quasi-constant frequency. CF = constant frequency. d = downward. u = upward. For the confusion matrix of all calls (all SNR) see Table A3 in the supplementary information file.

SNR ≥ 25 dB

Identification from automated classification	***Identification after manual checking***													
	bird	bush-cricket	buzz	CF-FMd	CF-FMd-CF	FMd	FMd-QCF	FMu-CF-FMd	FMu-QCF	FMu-QCF-FMd	noise	other insect or frog	QCF-FMd	**Total groups of calls**
bird	**2**	2	0	0	0	0	4	0	0	0	0	0	1	9
bush-cricket	0	**93**	1	0	0	2	2	0	0	0	5	0	2	105
buzz	0	26	**63**	0	0	1	3	0	0	0	13	0	0	106
CF-FMd	0	1	0	**8**	0	0	0	0	0	0	0	0	1	10
CF-FMd-CF	0	0	0	0	**1**	0	0	0	0	0	0	0	0	1
FMd	0	4	6	4	0	**134**	56	0	0	0	28	0	6	238
FMd-QCF	0	6	0	1	3	7	**192**	0	0	3	6	1	17	236
FMu-CF-FMd	0	2	0	2	0	1	2	**6**	0	0	6	0	0	19
FMu-QCF	1	1	0	0	0	0	3	0	**4**	0	0	0	1	10
FMu-QCF-FMd	0	3	0	0	0	0	6	0	3	**15**	1	0	5	33
noise	0	116	1	3	0	3	7	0	0	1	**563**	0	2	696
other insect or frog	0	2	0	0	0	0	1	0	0	0	0	**0**	0	3
QCF-FMd	1	0	0	0	1	0	15	0	0	2	1	0	**45**	65
Total groups of calls	4	256	71	18	5	148	291	6	7	21	623	1	80	1531

Discussion

The framework that we developed to classify the different bat sonotypes is a quick and easy approach to distinguish the main bat acoustic strategies occurring worldwide, without having to collect an exhaustive sound reference database of the local species calls. Because the definition of sonotypes does not take call duration into account, this objective tool avoids the classification of the same species in different sonotypes, excepted in the very few species emitting alternating call shapes (e.g. *Promops centralis*).

Classifier performance

The performance of our classifier was tested with success on recordings from bat communities from five different continents (Fig. 5). The rate of false negatives (i.e. calls that were not at all classified) was close to zero; a similar result was obtained in another assessment of the performance of Tadarida for another classifier (Barré et al., 2019). Thus, users can be confident that they will not miss sound events of interest.

It must be noted that most bat classifiers have a relatively high SNR threshold for classification, below which classification is not provided (Obrist and Boesch, 2018; Stahlschmidt and Brühl, 2012). This is not the case with Tadarida that aimed at detecting almost all hearable bat calls, only 3 bat sequences not being detected on 646 files. Here, bat sonotypes were classified with an AUC superior to 0.9 for signals of good quality (SNR > 25 dB), and with an AUC superior to 0.7 for signals of low quality (SNR < 25 dB). Moreover, the performance was tolerant to low SNR for the three sonotypes containing a CF element and for feeding buzzes. Although call duration and bandwidth influenced the success of bat sonotypes classification, since for instance short FMd-QCF were often classified as FMd, confusions between bats and non-bats were close to insignificant.

The sonotype FMd led to multiple groups although produced by only one individual, which is due to high variability in the peak frequency. Additionally, species using alternated frequencies led to multiple groups. These results can be adjusted by users by changing the tolerance to frequency input used to make groups of calls.

Among species using FMd, important differences exist in the way how energy is distributed among call harmonics; in the family of Phyllostomidae and in the genus *Plecotus* (Vespertilionidae), the same amount of energy is emitted in the different harmonics, while in the genus *Myotis* (Vespertilionidae), energy is stronger on the fundamental frequencies (Jones and Teeling, 2006). These differences are not accounted for by our classifier. However, to help users in this task, the potential presence of harmonics detected by Tadarida-L is shown in a dedicated column of the output (i.e. Ramp90, for which positive values are usually associated with harmonics).

A possible amelioration of sonotype classification lies in deep learning methods, that showed encouraging results in bat acoustic identification (Chen et al., 2020; Kobayashi et al., 2021; Mac Aodha et al., 2018). However, to our knowledge, there is still no extensive comparison between deep learning and random forest approaches for the classification of bat echolocation calls.

Usage perspectives

Although our method cannot approach the performance of a classifier trained to identify calls at the species level in specific bat communities (e.g. Ayala-Berdon et al., 2020; Barré et al., 2019; Obrist and Boesch, 2018), it is very convenient to discriminate different bat guilds (Denzinger and Schnitzler, 2013) in any geographical context. First, sonotypes can be used to separate the main acoustic strategies, such as flutter detecting foragers (FMu-CF-FMd, CF-FMd or CF-FMd-CF structures) versus gleaning foragers (FMd structure) versus aerial foragers (FMd-QCF, FMu-QCF, QCF-FMd or FMu-QCF-FMd structures) (Denzinger and Schnitzler, 2013). Then, frequencies may be used to separate species according to their spatial niche, such as open space foragers (< 30 kHz) versus edge space foragers (between 30 and 60 kHz) (Denzinger and Schnitzler, 2013; Kalko et al., 2008; Roemer et al., 2019). If, as we expect, species diversity matches the diversity of sonotype-frequency combinations, our classifier might be used to assess the state of bat communities anywhere in the world (e.g. Fidelino and Gan, 2019).

Another use of our classifier is to detect species that were described acoustically in geographical areas for which no specific classifier was developed and reference recordings are still scarce. Using our classifier, it is possible to target the species sonotype and frequency range in large amounts of recordings. The classifier can even be used for species that have not been described acoustically yet since they will be classified according to their sonotype and peak frequency. Detecting undocumented species will be especially facilitated in areas where few sympatric species are extant, e.g. on islands, deserts, or high latitudes (Barataud and Giosa, 2013; Mifsud and Vella, 2019; Walters et al., 2013; Ziegler et al., 2016).

Finally, it is possible to specifically study bat foraging behaviour using the "buzz" acoustic class of our classifier. Possible applications are the comparison of foraging activity across habitats, management practices and seasons (Ancillotto et al., 2021; Froidevaux et al., 2017; Toffoli and Rughetti, 2020; Weier et al., 2018), the study of pest regulation by bats (Charbonnier et al., 2021, 2014; Rodríguez-San Pedro et al., 2020; Salvarina et al., 2018), or of group foraging and competition (Gager, 2019; Lewanzik et al., 2019; Roeleke et al., 2020).

Acknowledgements

This work was supported by the Office Français de la Biodiversité and Veolia. We thank IN2P3 Computing Centre for providing facilities to process and archive in the long-term all the recordings used in this study, and Didier Bas for his help in this process. We thank all the contributors to the reference sound database; they are listed in supplementary file 1. We are grateful to two anonymous reviewers for their comments, which improved the quality of the

manuscript.

Conflict of interest

The authors declare no conflict of interest.

Authors' contributions

CR and YB conceived and designed the study. YB collected and labelled the majority of the sound files. CR labelled some sound files, adapted the scripts, checked the performance of the classifier and wrote the manuscript under the supervision of YB. CR, YB and JFJ interpreted the results and revised the manuscript critically. All authors gave final approval for publication.

Data availability

The R scripts and the tables associated with the findings of this study are openly available at GitHub [https://doi.org/10.5281/zenodo.5483030] (folder 'Sonotypes').

The classifier is available at [https://doi.org/10.6084/m9.figshare.14340341.v1].

The validation table for the test of the performance of the classifier is supplementary table 3, available at [https://figshare.com/articles/dataset/Validation_table_for_the_bat_sonotype_classifier/15149523].

All checked sound sequences to test the performance of the classifier are available at [https://figshare.com/articles/media/Sounds_used_to_validate_an_automatic_classifier_of_bat_sonotypes/15141201].

References

Ancillotto, L., Festa, F., De Benedetta, F., Cosentino, F., Pejic, B., Russo, D., 2021. Free-ranging livestock and a diverse landscape structure increase bat foraging in mountainous landscapes. Agrofor. Syst. 1–12. https://doi.org/10.1007/s10457-021-00591-0

Arias-Aguilar, A., Hintze, F., Aguiar, L.M., Rufray, V., Bernard, E., Pereira, M.J.R., 2018. Who’s calling? Acoustic identification of Brazilian bats. Mammal Res. 63, 231–253.

Au, W.W., 1997. Echolocation in dolphins with a dolphin-bat comparison. Bioacoustics 8, 137–162.

Ayala-Berdon, J., Medina-Bello, K.I., López-Cuamatzi, I.L., Vázquez-Fuerte, R., MacSwiney G, M.C., Orozco-Lugo, L., Iñiguez-Dávalos, I., Guillén-Servent, A., Martínez-Gómez, M., 2020. Random forest is the best species predictor for a community of insectivorous bats inhabiting a mountain ecosystem of central Mexico. Bioacoustics 1–21.

Barataud, M., 2015. Acoustic ecology of European bats: species identification, study of their habitats and foraging behaviour. Biotope éditions.

Barataud, M., Giosa, S., 2013. Identification et écologie acoustique des chiroptères de La Réunion. Le Rhinolophe 19, 147–175.

Barataud, M., Giosa, S., Leblanc, F., Rufray, V., Disca, T., Tillon, L., Delaval, M., Haquart, A., Dewynter, M., 2013. Identification et écologie acoustique des chiroptères de Guyane française. Le Rhinolophe 19, 103–145.

Barré, K., Viol, I.L., Julliard, R., Pauwels, J., Newson, S.E., Julien, J.-F., Claireau, F., Kerbiriou, C., Bas, Y., 2019. Accounting for automated identification errors in acoustic surveys. Methods Ecol. Evol. 10, 1171–1188. https://doi.org/10.1111/2041-210X.13198

Bas, Y., Bas, D., Julien, J.-F., 2017. Tadarida: A Toolbox for Animal Detection on Acoustic Recordings. J. Open Res. Softw. 5. https://doi.org/10.5334/jors.154

Charbonnier, Y., Barbaro, L., Theillout, A., Jactel, H., 2014. Numerical and functional responses of forest bats to a major insect pest in pine plantations. PLoS One 9, e109488.

Charbonnier, Y., Papura, D., Touzot, O., Rhouy, N., Sentenac, G., Rusch, A., 2021. Pest control services provided by bats in vineyard landscapes. Agric. Ecosyst. Environ. 306, 107207.

Chen, X., Zhao, J., Chen, Y., Zhou, W., Hughes, A.C., 2020. Automatic standardized processing and identification of tropical bat calls using deep learning approaches. Biol. Conserv. 241, 108269.

Collen, A., 2012. The evolution of echolocation in bats: a comparative approach (Doctoral dissertation). University College London.

Denzinger, A., Schnitzler, H.-U., 2013. Bat guilds, a concept to classify the highly diverse foraging and echolocation behaviors of microchiropteran bats. Front. Physiol. 4. https://doi.org/10.3389/fphys.2013.00164

Fenton, M.B., Bell, G.P., 1981. Recognition of Species of Insectivorous Bats by Their Echolocation Calls. J. Mammal. 62, 233–243. https://doi.org/10.2307/1380701

Fidelino, J.S., Gan, J.L., 2019. The Influence of Vegetation and Insect Abundance on Insectivorous Bat Activity during Dusk Emergence in an Urban Space in Metro Manila, Philippines. Sci. Diliman 31.

Fraser, E.E., Silvis, A., Brigham, R.M., Czenze, Z.J., 2020. Bat Echolocation Research: A handbook for planning and conducting acoustic studies, Second edition. ed. Bat Conservation International, Austin, Texas, USA.

Froidevaux, J., Louboutin, B., Jones, G., 2017. Does organic farming enhance biodiversity in Mediterranean vineyards? A case study with bats and arachnids. Agric. Ecosyst. Environ. 249, 112–122. https://doi.org/10.1016/j.agee.2017.08.012

Gager, Y., 2019. Information transfer about food as a reason for sociality in bats. Mammal Rev. 49, 113–120.

Goëau, H., Kahl, S., Glotin, H., Planqué, R., Vellinga, W.-P., Joly, A., 2018. Overview of BirdCLEF 2018: monospecies vs. soundscape bird identification, in: CLEF: Conference and Labs of the Evaluation Forum.

Griffin, D.R., Webster, F.A., Michael, C.R., 1960. The echolocation of flying insects by bats. Anim. Behav. 8, 141–154.

Holderied, M.W., 2006. Flight and echolocation behaviour of whiskered bats commuting along a hedgerow: range-dependent sonar signal design, Doppler tolerance and evidence for `acoustic focussing'. J. Exp. Biol. 209, 1816–1826. https://doi.org/10.1242/jeb.02194

Jones, G., 2005. Echolocation. Curr. Biol. 15, R484–R488.

Jones, G., Holderied, M.W., 2007. Bat echolocation calls: adaptation and convergent evolution. Proc. R. Soc. B Biol. Sci. 274, 905–912. https://doi.org/10.1098/rspb.2006.0200

Jones, G., Teeling, E., 2006. The evolution of echolocation in bats. Trends Ecol. Evol. 21, 149–156. https://doi.org/10.1016/j.tree.2006.01.001

Kahl, S., Wood, C.M., Eibl, M., Klinck, H., 2021. BirdNET: A deep learning solution for avian diversity monitoring. Ecol. Inform. 61, 101236.

Kalko, E.K., Estrada Villegas, S., Schmidt, M., Wegmann, M., Meyer, C.F., 2008. Flying high—assessing the use of the aerosphere by bats. Integr. Comp. Biol. 48, 60–73.

Kobayashi, K., Masuda, K., Haga, C., Matsui, T., Fukui, D., Machimura, T., 2021. Development of a species identification system of Japanese bats from echolocation calls using convolutional neural networks. Ecol. Inform. 62, 101253.

Leonard, M.L., Fenton, M.B., 1984. Echolocation calls of Euderma maculatum (Vespertilionidae): use in orientation and communication. J. Mammal. 65, 122–126.

Lewanzik, D., Sundaramurthy, A.K., Goerlitz, H.R., 2019. Insectivorous bats integrate social information about species identity, conspecific activity and prey abundance to estimate cost–benefit ratio of interactions. J. Anim. Ecol. 88, 1462–1473.

Lopez-Baucells, A., Rocha, R., Bobrowiec, P.E.D., Palmeirim, J.M., Meyer, C.F.J., 2016. Field guide to Amazonian bats. National Institute of Amazonian Research (INPA).

López-Baucells, A., Torrent, L., Rocha, R., Bobrowiec, P.E., Palmeirim, J.M., Meyer, C.F., 2019. Stronger together: Combining automated classifiers with manual post-validation optimizes the workload vs reliability trade-off of species identification in bat acoustic surveys. Ecol. Inform. 49, 45–53.

Mac Aodha, O., Gibb, R., Barlow, K.E., Browning, E., Firman, M., Freeman, R., Harder, B., Kinsey, L., Mead, G.R., Newson, S.E., 2018. Bat detective—Deep learning tools for bat acoustic signal detection. PLoS Comput. Biol. 14, e1005995.

Madsen, P.T., Surlykke, A., 2013. Functional convergence in bat and toothed whale biosonars. Physiology 28, 276–283.

Mifsud, C.M., Vella, A., 2019. Mitochondrial genetic diversity of bat species from the Maltese Islands and applications for their conservation. Nat. Eng. Sci. 4, 276–292.

Nocera, T., Ford, W.M., Silvis, A., Dobony, C.A., 2019. Let's Agree to Disagree: Comparing Auto-Acoustic Identification Programs for Northeastern Bats. J. Fish Wildl. Manag. 10, 346–361.

Obrist, M.K., Boesch, R., 2018. BatScope manages acoustic recordings, analyses calls, and classifies bat species automatically. Can. J. Zool. 96, 939–954.

Obrist, M.K., Pavan, G., Sueur, J., Riede, K., Llusia, D., Márquez, R., 2010. Bioacoustics approaches in biodiversity inventories. Abc Taxa 8, 68–99.

Ochoa, J., O'Farrell, M.J., Miller, B.W., 2000. Contribution of acoustic methods to the study of insectivorous bat diversity in protected areas from northern Venezuela. Acta Chiropterologica 2, 171–183.

R Core Team, 2013. R: A language and environment for statistical computing. Vienna, Austria.

Richardson, S.M., Lintott, P.R., Hosken, D.J., Mathews, F., 2019. An evidence-based approach to specifying survey effort in ecological assessments of bat activity. Biol. Conserv. 231, 98–102.

Rodríguez-San Pedro, A., Allendes, J.L., Beltrán, C.A., Chaperon, P.N., Saldarriaga-Córdoba, M.M., Silva, A.X., Grez, A.A., 2020. Quantifying ecological and economic value of pest control services provided by bats in a vineyard landscape of central Chile. Agric. Ecosyst. Environ. 302, 107063.

Roeleke, M., Blohm, T., Hoffmeister, U., Marggraf, L., Schlägel, U.E., Teige, T., Voigt, C.C., 2020. Landscape structure influences the use of social information in an insectivorous bat. Oikos.

Roemer, C., Coulon, A., Disca, T., Bas, Y., 2019. Bat sonar and wing morphology predict species vertical niche. J. Acoust. Soc. Am. 145, 3242–3251.

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. https://doi.org/10.1016/j.ecolind.2017.03.023

Salvarina, I., Gravier, D., Rothhaupt, K.-O., 2018. Seasonal bat activity related to insect emergence at three temperate lakes. Ecol. Evol. 8, 3738–3750.

Sauvé, C.C., Beauplet, G., Hammill, M.O., Charrier, I., 2015. Mother–pup vocal recognition in harbour seals: influence of maternal behaviour, pup voice and habitat sound properties. Anim. Behav. 105, 109–120.

Schnitzler, H.-U., Kalko, E.K.V., 2001. Echolocation by Insect-Eating Bats. BioScience 51, 557. https://doi.org/10.1641/0006-3568(2001)051[0557:EBIEB]2.0.CO;2

Schnitzler, H.-U., Moss, C.F., Denzinger, A., 2003. From spatial orientation to food acquisition in echolocating bats. Trends Ecol. Evol. 18, 386–394. https://doi.org/10.1016/S0169-5347(03)00185-X

Stahlschmidt, P., Brühl, C.A., 2012. Bats as bioindicators - the need of a standardized method for acoustic bat activity surveys: *Standardized bat surveys*. Methods Ecol. Evol. 3, 503–508. https://doi.org/10.1111/j.2041-210X.2012.00188.x

Sueur, J., Gasc, A., Grandcolas, P., Pavoine, S., 2012. Global estimation of animal diversity using automatic acoustic sensors. Sens. Ecol. Paris CNRS 99–117.

Toffoli, R., Rughetti, M., 2020. Effect of water management on bat activity in rice paddies. Paddy Water Environ. 18, 687–695.

Walters, C.L., Collen, A., Lucas, T., Mroz, K., Sayer, C.A., Jones, K.E., 2013. Challenges of using bioacoustics to globally monitor bats, in: Bat Evolution, Ecology, and Conservation. Springer, pp. 479–499.

Weier, S.M., Grass, I., Linden, V.M., Tscharntke, T., Taylor, P.J., 2018. Natural vegetation and bug abundance promote insectivorous bat activity in macadamia orchards, South Africa. Biol. Conserv. 226, 16–23.

Wiley, R.H., 2006. Signal detection and animal communication. Adv. Study Behav. 36, 217–247.

Wilson, D.E., Mittermeier, R.A., 2019. Handbook of the Mammals of the World, Vol. 9: Bats. Lynx Edicions Barc. Spain 1008pp.

Zamora-Gutierrez, V., Lopez-Gonzalez, C., MacSwiney Gonzalez, M.C., Fenton, B., Jones, G., Kalko, E.K., 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.

Ziegler, A.C., Howarth, F.G., Simmons, N.B., 2016. A second endemic land mammal for the Hawaiian Islands: a new genus and species of fossil bat (Chiroptera: Vespertilionidae). Am. Mus. Novit. 2016, 1–52.

Supplementary Information File

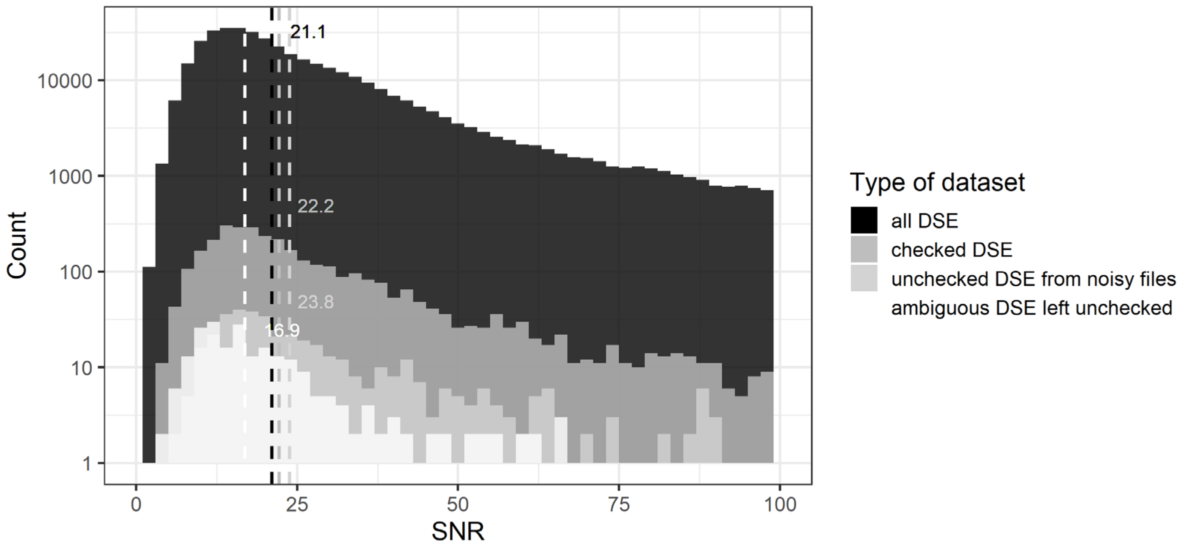
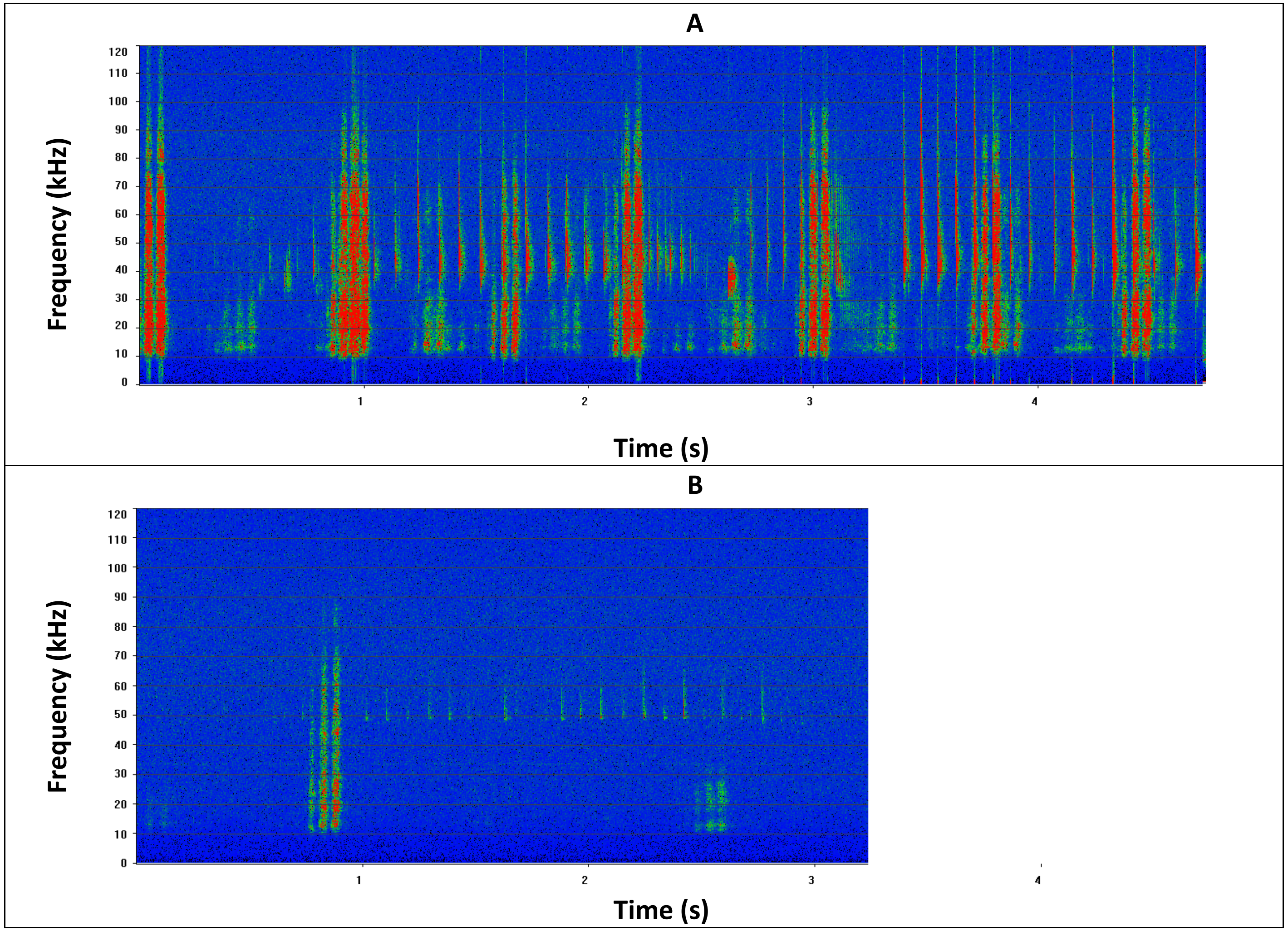
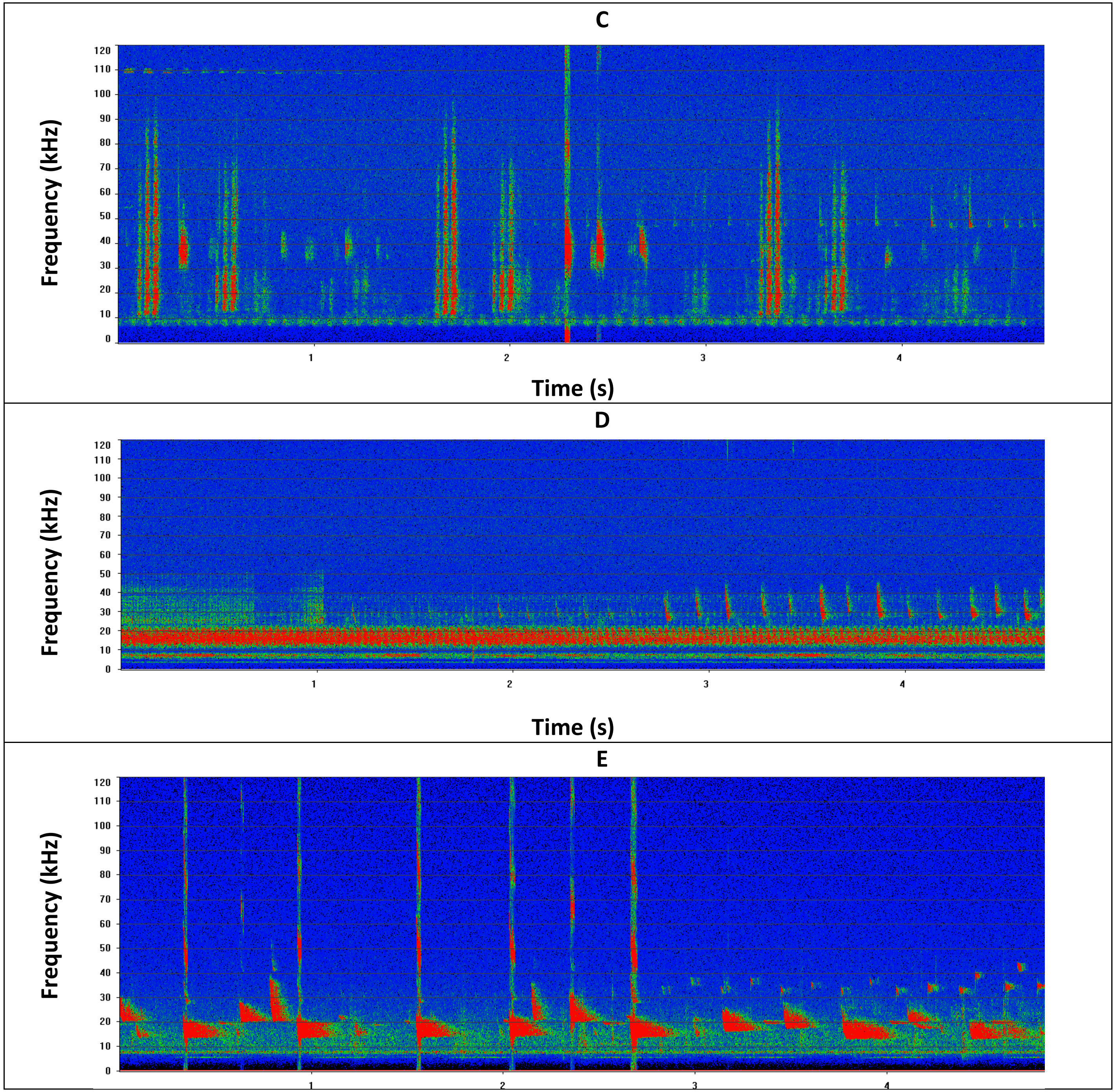


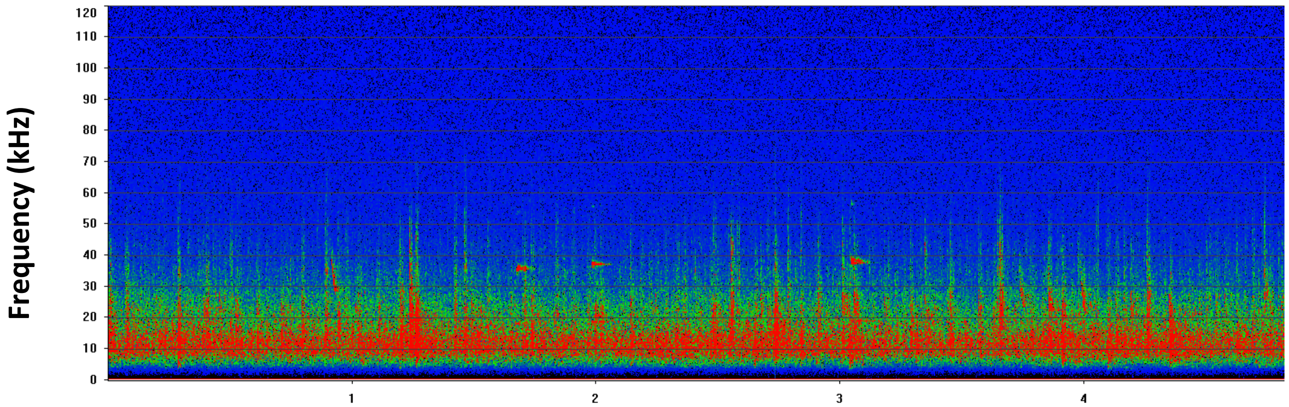
Figure A 1: Distribution of the SNR (signal-to-noise ratio) of the groups of DSE (detected sound events) of the different datasets; the x axis shows values between 0 and 100 and the y axis is in the logarithmic scale. The vertical lines and their associated values show the median value of each dataset. Ambiguous DSE were left unchecked but originate from checked files (i.e. the other DSE of this file were checked), whereas no DSE originating from noisy files were checked.


A
120
110
100
90
80
70
60
50
40
30
20
10
0
Frequency (kHz)
1
2
3
4
Time (s)
B
120
110
100
90
80
70
60
50
40
30
20
10
0
Frequency (kHz)
1
2
3
4
Time (s)


C
Frequency (kHz)
Time (s)
D
Frequency (kHz)
Time (s)
E
Frequency (kHz)

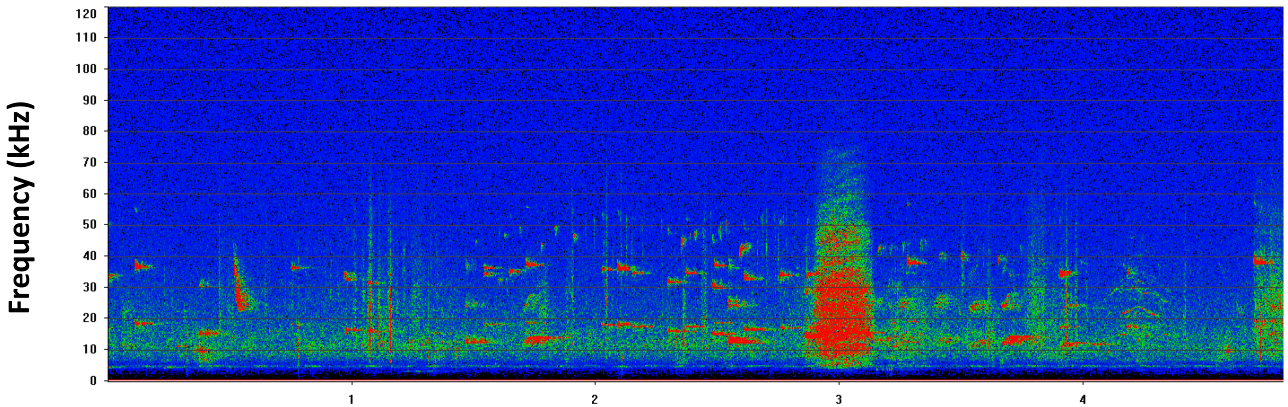
Time (s)

F

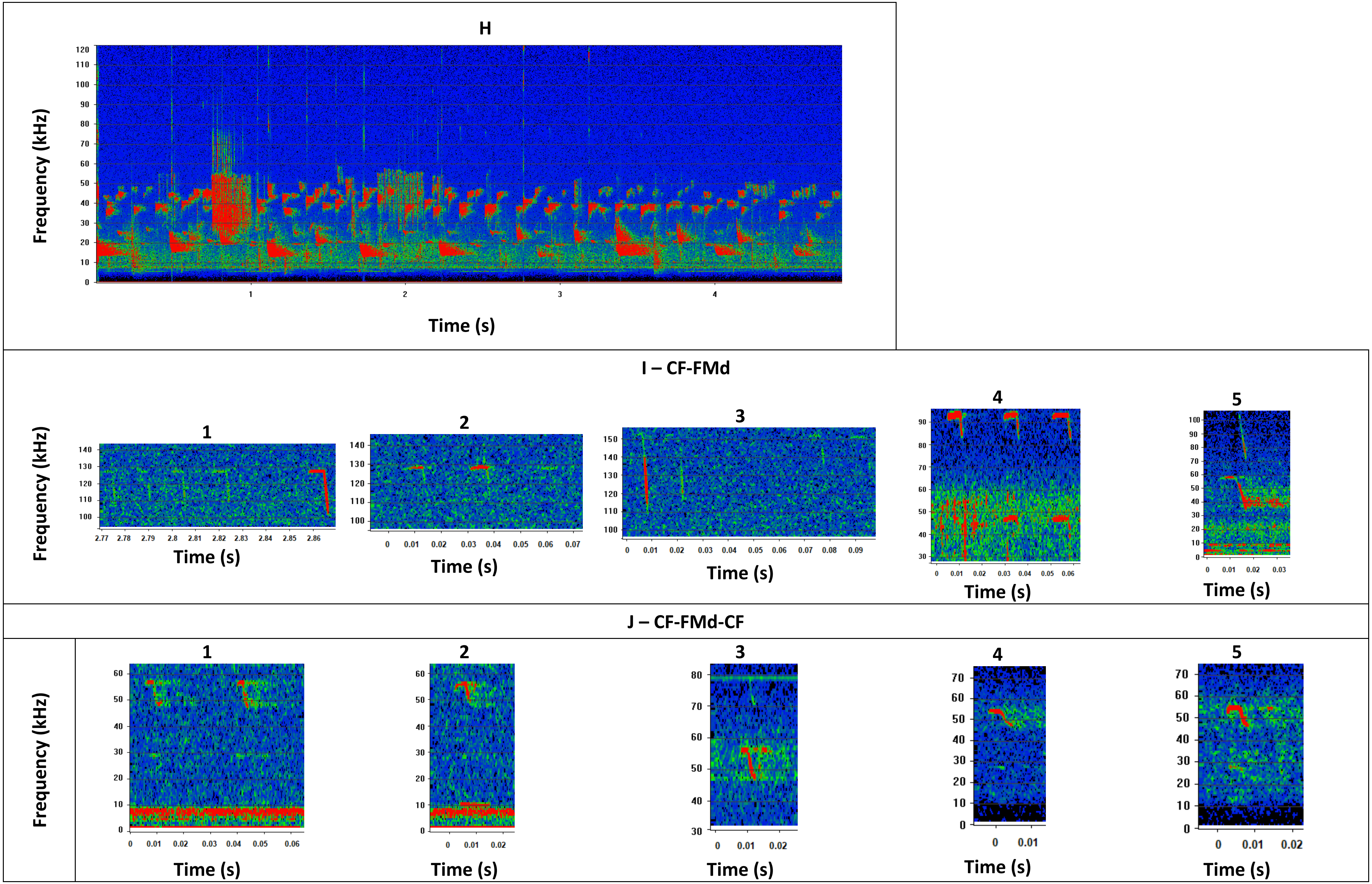


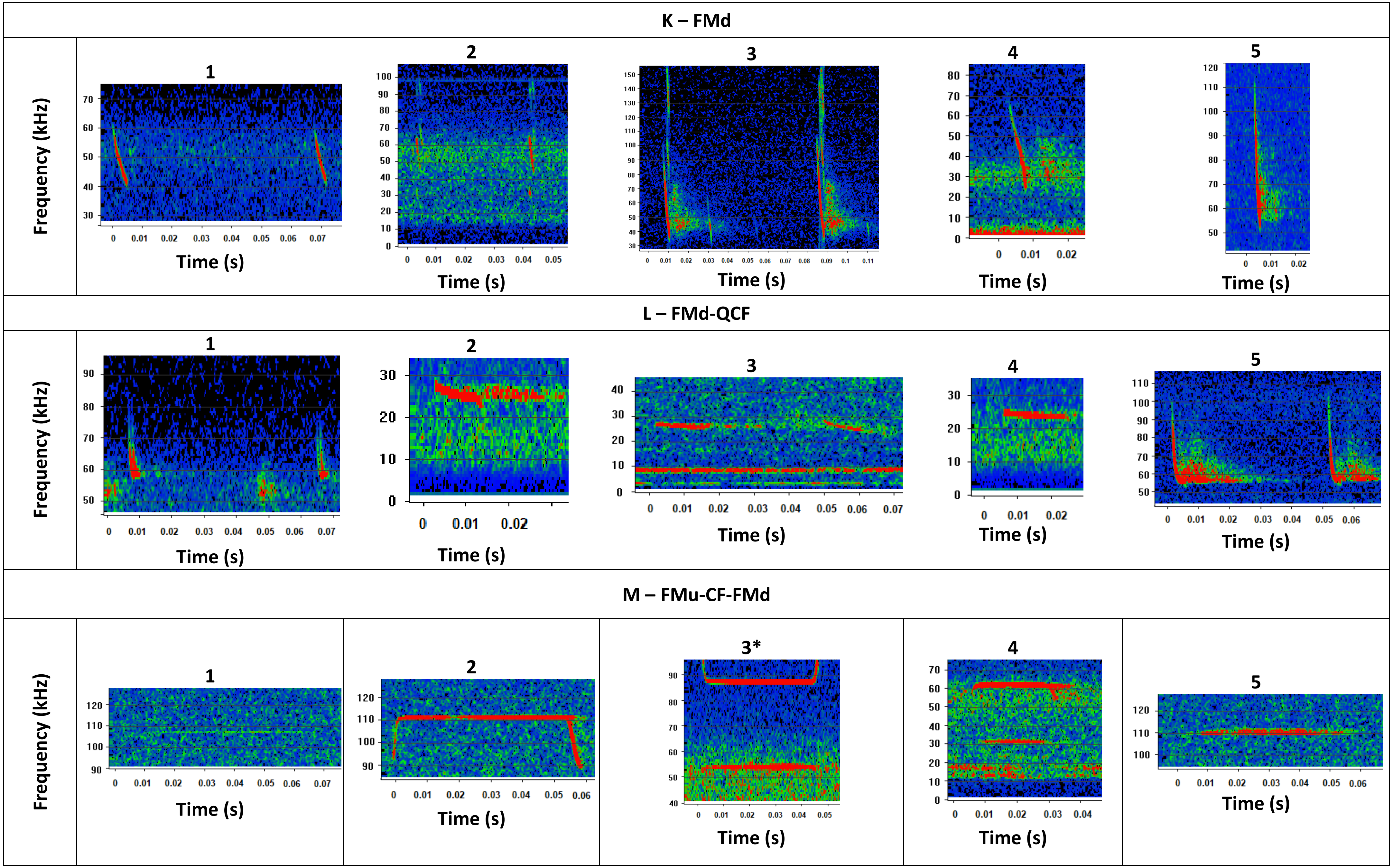
Time (s)

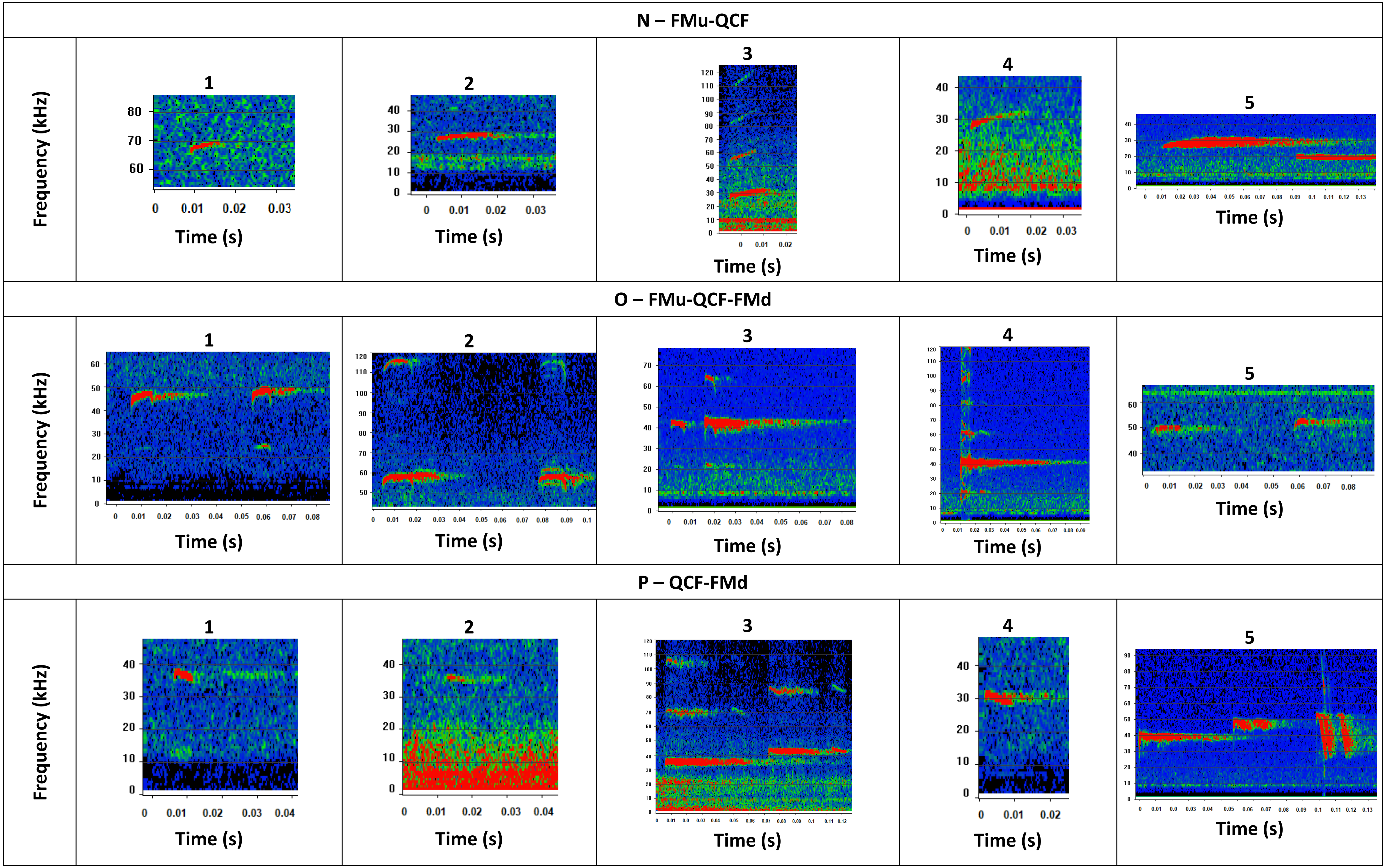
G



Time (s)


H
Frequency (kHz)
Time (s)
I – CF-FMd
1
2
3
4
5
Frequency (kHz)
Time (s)
J – CF-FMd-CF
1
2
3
4
5
Frequency (kHz)
Time (s)


K – FMd
1
2
3
4
5
Frequency (kHz)
Time (s)
L – FMd-QCF
1
2
3
4
5
Frequency (kHz)
Time (s)
M – FMu-CF-FMd
1
2
3*
4
5
Frequency (kHz)
Time (s)


N – FMu-QCF
O – FMu-QCF-FMd
P – QCF-FMd
Frequency (kHz)
Time (s)
1
2
3
4
5

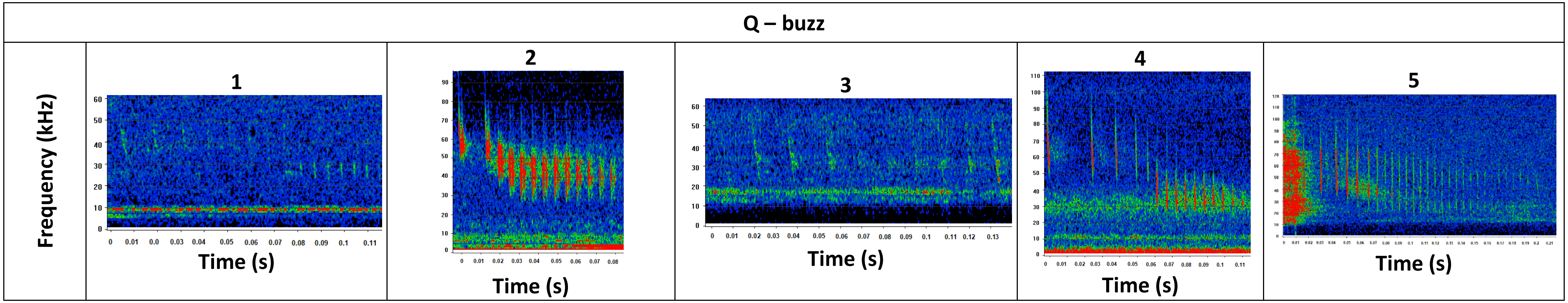


Figure A 2: Example of different sonograms of the dataset used to test the performance of the classifier. A-F: full-length files checked during the test. G-H: full-length files rejected by the checker for the test because they were too noisy or had too many overlapping calls. I-Q: isolated example calls retrieved from the checked files for the sonotypes CF-FMd (I), CF-FMd-CF (J), FMd (K), FMd-QCF (L), FMu-CF-FMd (M), FMu-QCF (N), FMu-QCF-FMd (O), QCF-FMd (P) and for the buzz (Q). *: The sonogram in M3 is the result when a call is produced with a frequency greater than the sample rate; the result is an inversed acoustic signal; the classifier was built using similar sounds for training. Sonograms were produced using Syrinx (John Burt, USA; FFT window type Hanning with transform size adapted to sample rate to obtain comparable graphical settings).

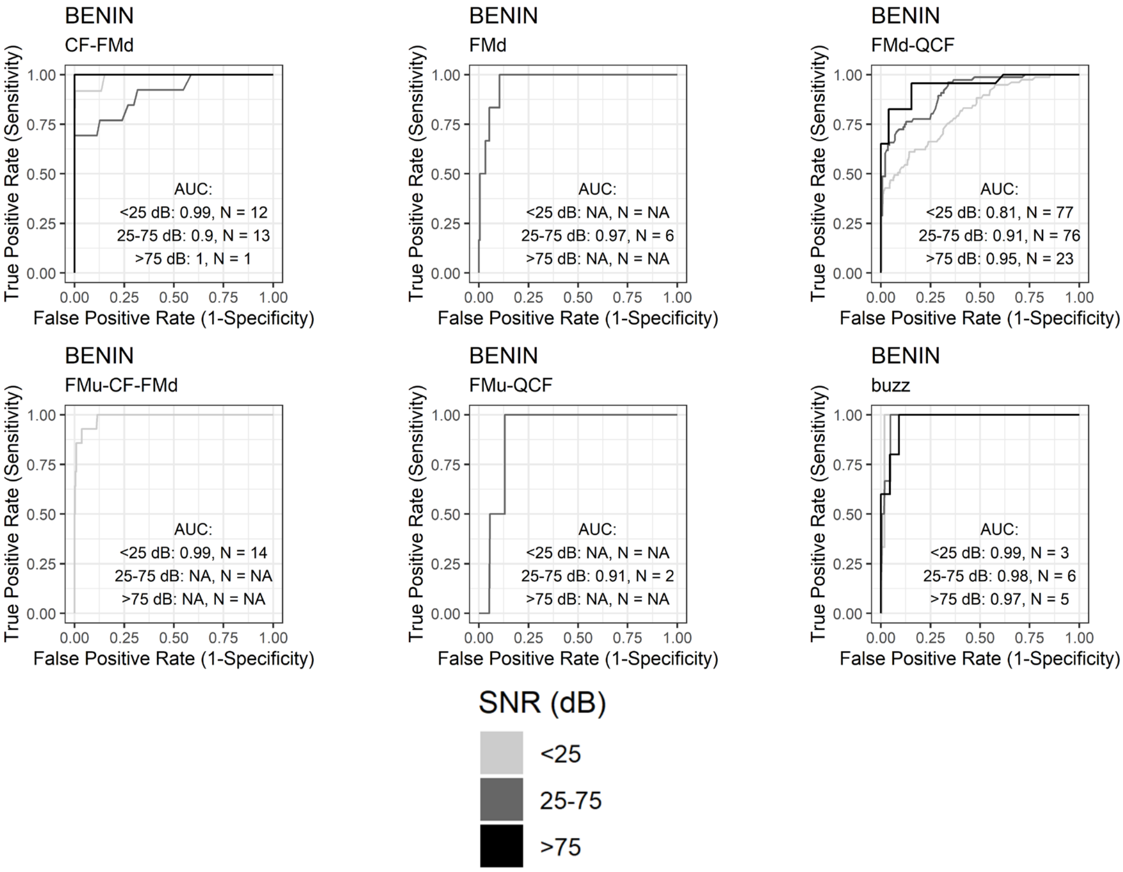


Figure A 3: Receiver operating characteristic (ROC) curves between the confidence score of false positive rate (FPR) and true positive rate (TPR) for each bat sonotype for Benin. Grey shades represent the median of the maximal amplitude among the calls classified as the sonotype that was checked: light grey <25 dB, dark grey 25-75 dB, black >75 dB. AUC (Area under curve) is a proxy of the performance of the classifier. N = number of groups of calls. FM = frequency modulated. QCF = quasi-constant frequency. CF = constant frequency. d = downward. u = upward.

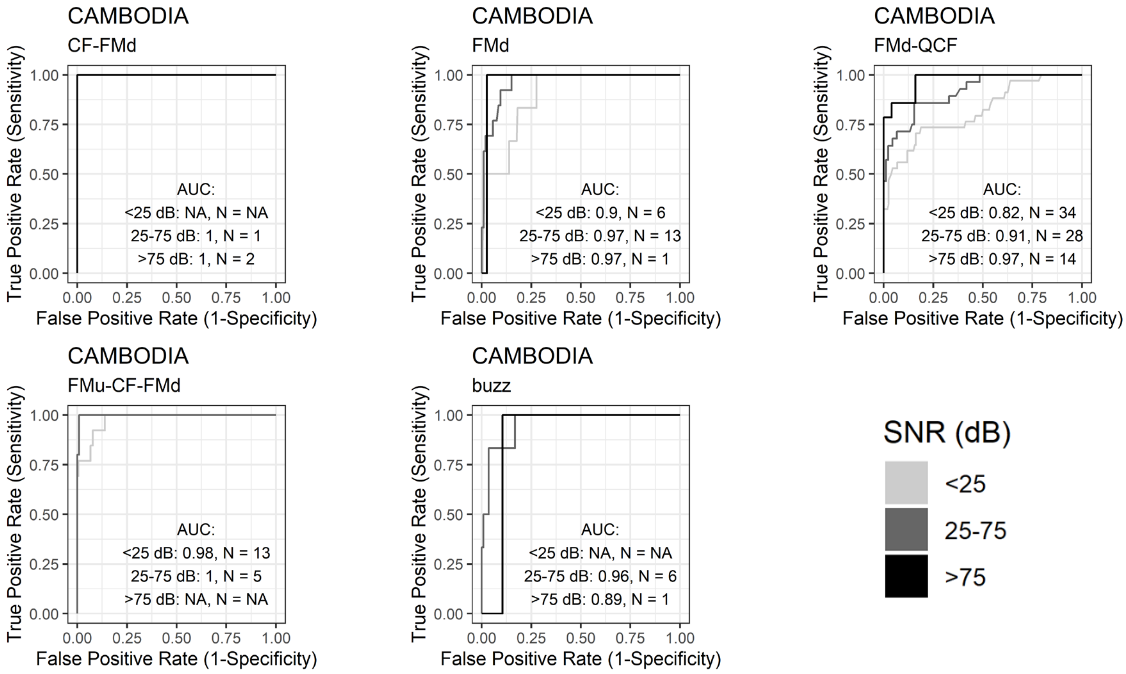


Figure A 4: Receiver operating characteristic (ROC) curves between the confidence score of false positive rate (FPR) and true positive rate (TPR) for each bat sonotype for Cambodia. Grey shades represent the median of the maximal amplitude among the calls classified as the sonotype that was checked: light grey <25 dB, dark grey 25-75 dB, black >75 dB. AUC (Area under curve) is a proxy of the performance of the classifier. N = number of groups of calls. FM = frequency modulated. QCF = quasi-constant frequency. CF = constant frequency. d = downward. u = upward.

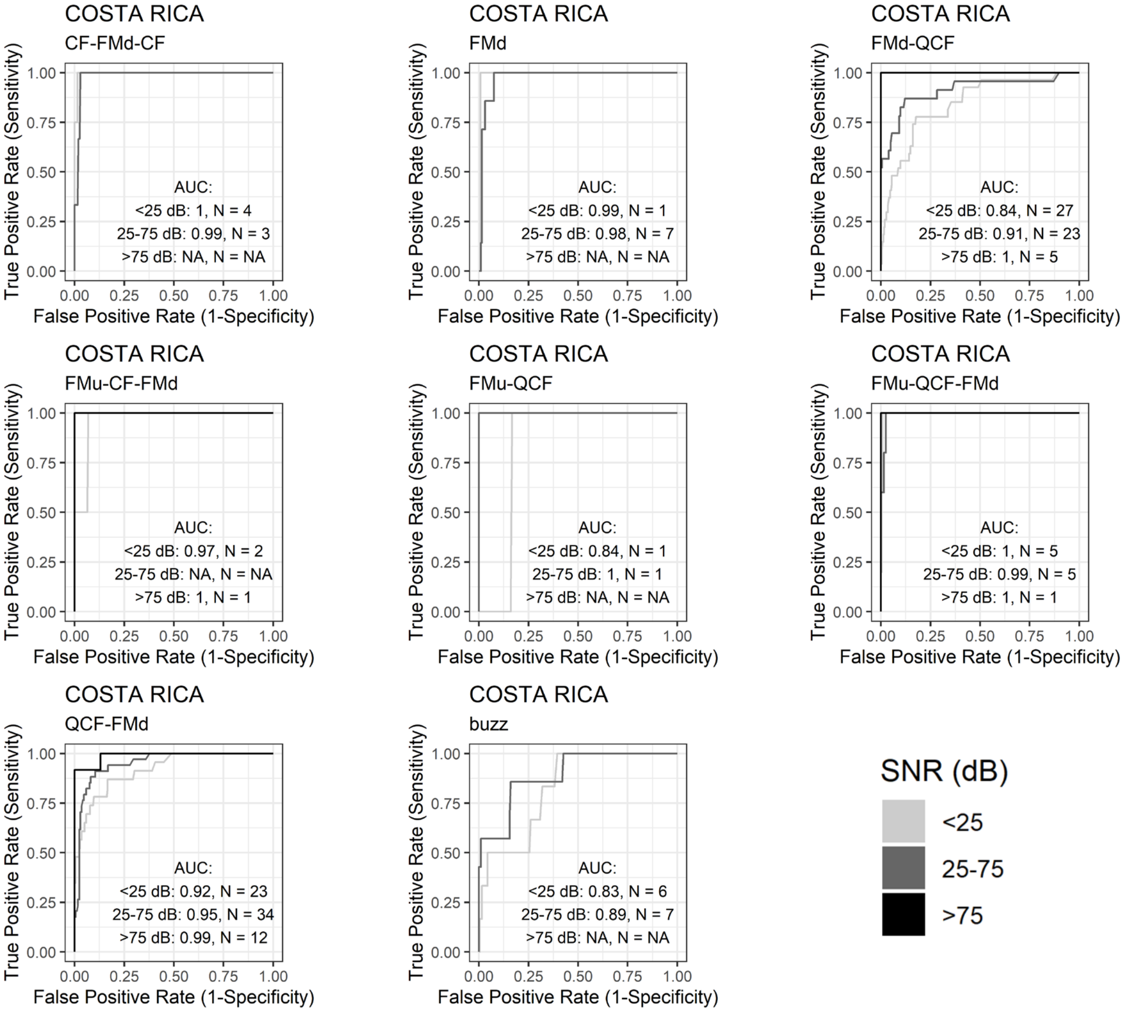


Figure A 5: Receiver operating characteristic (ROC) curves between the confidence score of false positive rate (FPR) and true positive rate (TPR) for each bat sonotype for Costa Rica. Grey shades represent the median of the maximal amplitude among the calls classified as the sonotype that was checked: light grey <25 dB, dark grey 25-75 dB, black >75 dB. AUC (Area under curve) is a proxy of the performance of the classifier. N = number of groups of calls. FM = frequency modulated. QCF = quasi-constant frequency. CF = constant frequency. d = downward. u = upward.

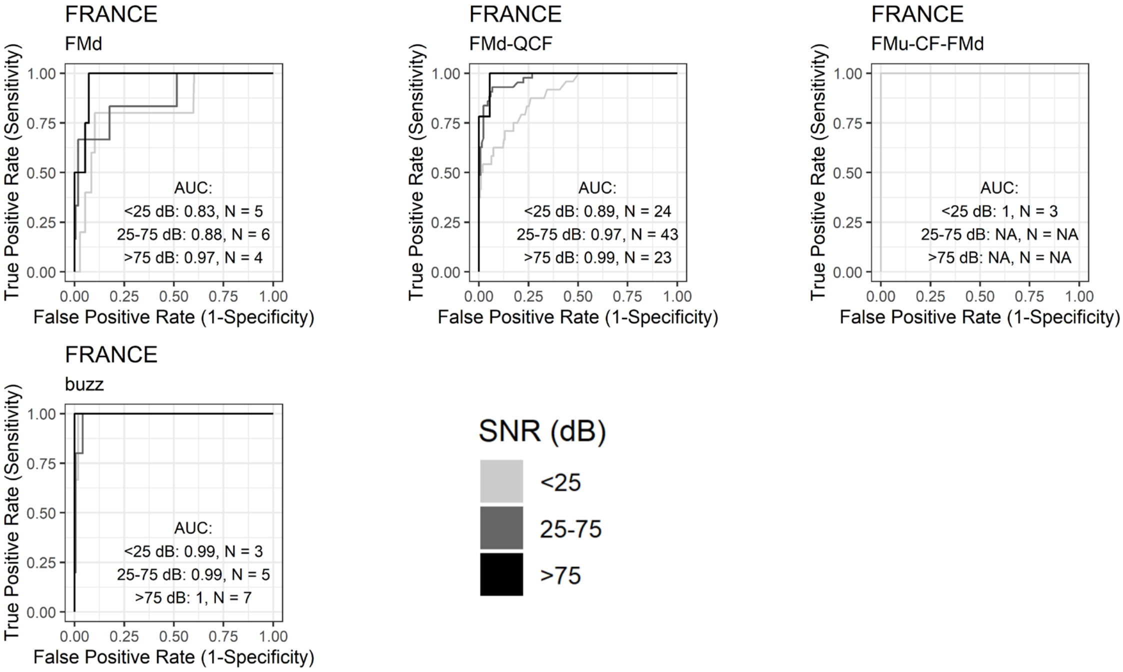


Figure A 6: Receiver operating characteristic (ROC) curves between the confidence score of false positive rate (FPR) and true positive rate (TPR) for each bat sonotype for France. Grey shades represent the median of the maximal amplitude among the calls classified as the sonotype that was checked: light grey <25 dB, dark grey 25-75 dB, black >75 dB. AUC (Area under curve) is a proxy of the performance of the classifier. N = number of groups of calls. FM = frequency modulated. QCF = quasi-constant frequency. CF = constant frequency. d = downward. u = upward.

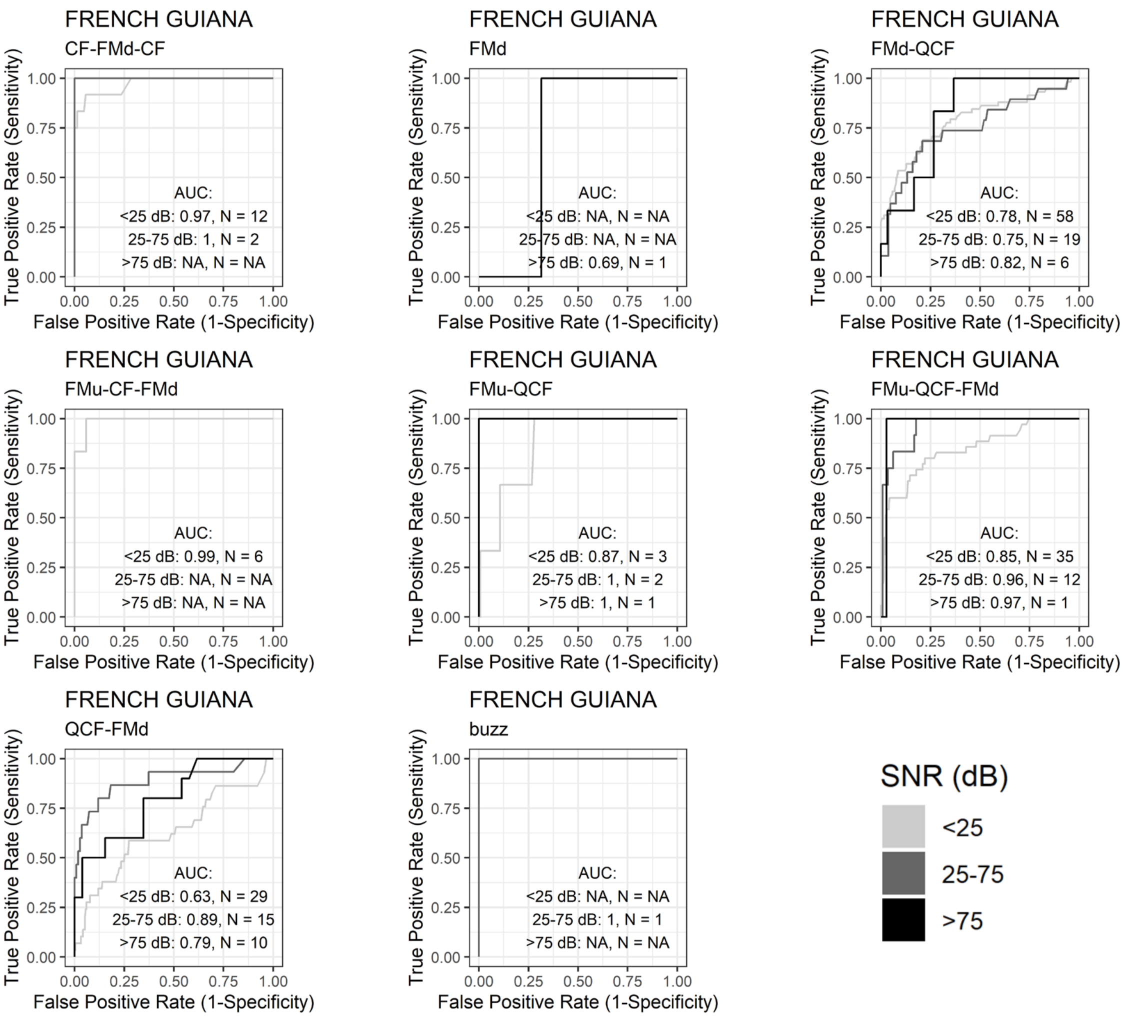


Figure A 7: Receiver operating characteristic (ROC) curves between the confidence score of false positive rate (FPR) and true positive rate (TPR) for each bat sonotype for French Guiana. Grey shades represent the median of the maximal amplitude among the calls classified as the sonotype that was checked: light grey <25 dB, dark grey 25-75 dB, black >75 dB. AUC (Area under curve) is a proxy of the performance of the classifier. N = number of groups of calls. FM = frequency modulated. QCF = quasi-constant frequency. CF = constant frequency. d = downward. u = upward.

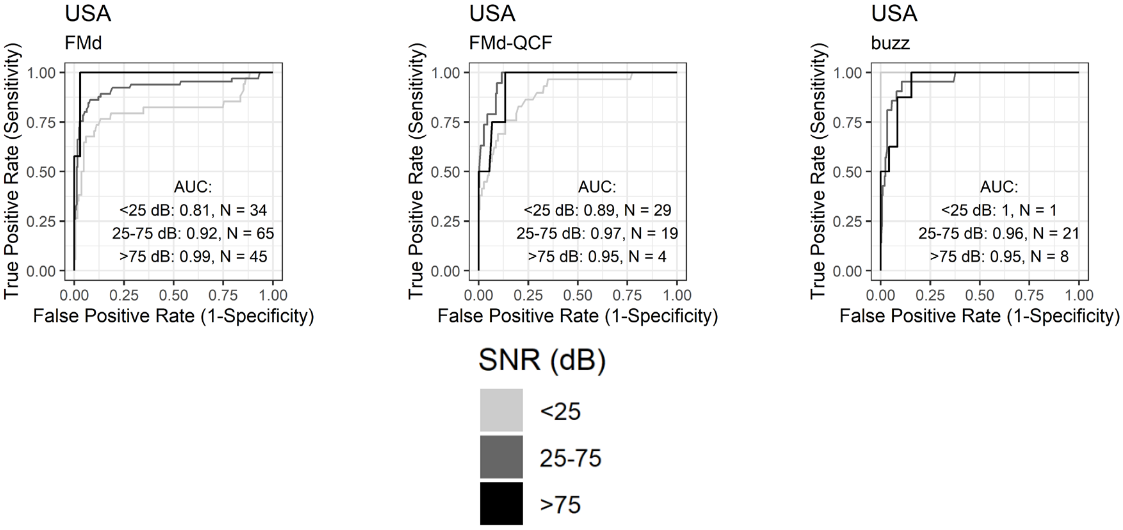


Figure A 8: Receiver operating characteristic (ROC) curves between the confidence score of false positive rate (FPR) and true positive rate (TPR) for each bat sonotype for the United States of America. Grey shades represent the median of the maximal amplitude among the calls classified as the sonotype that was checked: light grey <25 dB, dark grey 25-75 dB, black >75 dB. AUC (Area under curve) is a proxy of the performance of the classifier. N = number of groups of calls. FM = frequency modulated. QCF = quasi-constant frequency. CF = constant frequency. d = downward. u = upward.

a b

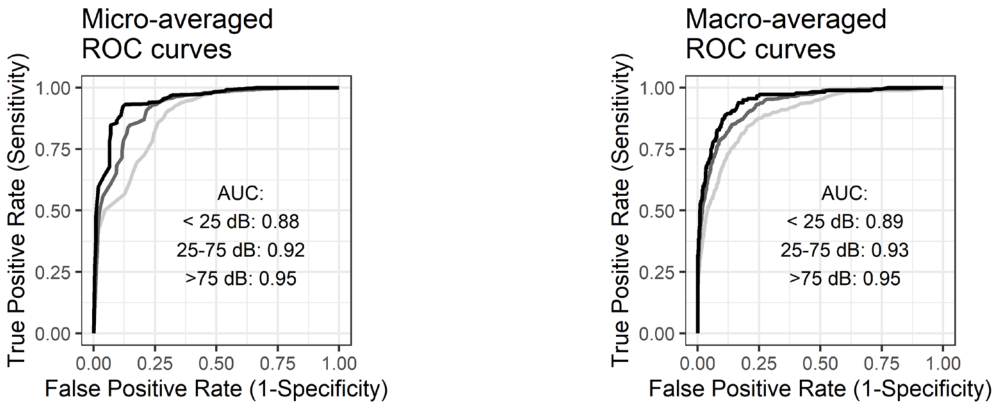


Figure A 9: Micro-averaged (a) and macro-averaged (b) receiver operating characteristic (ROC) curves between the confidence score of false positive rate (FPR) and true positive rate (TPR) for all bat sonotypes. Grey shades represent the median of the maximal amplitude among the calls classified as the sonotype that was checked: light grey <25 dB, dark grey 25-75 dB, black >75 dB. AUC (Area under curve) is a proxy of the performance of the classifier.

Table A 1: Summary table of the AUC (area under curve) for each acoustic class and each SNR (signal-to-noise ratio).

SNR (dB)	***Acoustic class***	***AUC***
<25	bird	0.93
	bush-cricket	0.73
	buzz	0.93
	CF-FMd	0.98
	CF-FMd-CF	0.98
	FMd	0.84
	FMd-QCF	0.83
	FMu-CF-FMd	0.98
	FMu-QCF	0.9
	FMu-QCF-FMd	0.93
	moth	0.73
	noise	0.83
25-75	other-mammal	0.85
	QCF-FMd	0.78
	bird	0.85
	bush-cricket	0.77
	buzz	0.95
	CF-FMd	0.92
	CF-FMd-CF	0.99
	FMd	0.93
	FMd-QCF	0.9
	FMu-CF-FMd	1
	FMu-QCF	0.98
	FMu-QCF-FMd	0.94
	moth	NA
	noise	0.89
	other-mammal	0.99
>75	QCF-FMd	0.94
	bird	NA
	bush-cricket	0.82
	buzz	0.98
	CF-FMd	1
	CF-FMd-CF	NA
	FMd	0.97
	FMd-QCF	0.89
	FMu-CF-FMd	1
	FMu-QCF	1
	FMu-QCF-FMd	0.99
	moth	NA
	noise	0.89
	other-mammal	NA
	QCF-FMd	0.92

Table A 2: Confusion matrix of social calls between acoustic classes identified by the classifier and after manual checking for groups of calls of any SNR. These results do not take the probability index associated with the automated classification into account. FM = frequency modulated. QCF = quasi-constant frequency. CF = constant frequency. d = downward. u = upward.

All SNR – social calls

Identification from automated classification	***Identification after manual checking***						
	buzz	CF-FMd	complex	FMd	FMd-QCF	FMu-QCF-FMd	***Total groups of calls***
bird	0	0	1	0	0	0	1
bush-cricket	5	1	5	2	0	0	13
buzz	17	0	0	0	0	0	17
CF-FMd	0	4	1	0	0	0	5
FMd	3	0	1	25	1	0	30
FMd-QCF	0	0	17	9	8	0	34
FMu-CF-FMd	0	0	1	0	0	0	1
FMu-QCF	0	0	2	0	0	0	2
FMu-QCF-FMd	0	0	0	0	0	1	1
noise	1	0	0	2	0	0	3
QCF-FMd	1	0	3	0	0	0	4
Total groups of calls	27	5	31	38	9	1	111

Table A 3: Confusion matrix between acoustic classes identified by the classifier and after manual checking for groups of calls of any SNR. These results do not take the probability index associated with the automated classification into account. FM = frequency modulated. QCF = quasi-constant frequency. CF = constant frequency. d = downward. u = upward. The three false negatives (FMd-QCF) are not shown.

All SNR

Identification from automated classification	*Identification after manual checking*													
	bird	bush-cricket	buzz	CF-FMd	CF-FMd-CF	FMd	FMd-QCF	FMu-CF-FMd	FMu-QCF	FMu-QCF-FMd	noise	other insect or frog	QCF-FMd	***Total groups of calls***
bird	**36**	69	0	0	0	0	27	0	0	0	19	3	4	158
bush-cricket	5	**291**	3	0	0	3	8	0	0	2	13	1	6	332
buzz	0	47	**72**	0	0	1	4	0	0	0	17	0	0	141
CF-FMd	0	1	0	**14**	0	0	0	0	0	0	0	0	1	16
CF-FMd-CF	0	0	0	0	**10**	0	0	0	0	0	0	0	0	10
FMd	0	48	6	5	0	**171**	99	2	0	1	77	0	8	417
FMd-QCF	3	39	1	2	8	9	**324**	2	2	12	19	2	34	457
FMu-CF-FMd	0	18	1	7	1	2	5	**39**	0	3	15	1	5	97
FMu-QCF	1	2	0	0	0	0	11	0	**5**	0	0	0	3	22
FMu-QCF-FMd	1	16	0	0	1	3	24	0	4	**40**	10	3	12	114
noise	2	314	1	4	0	5	21	1	0	1	**1329**	1	12	1691
other insect or frog	0	9	0	0	0	0	2	0	0	0	0	**3**	1	15
QCF-FMd	3	5	0	0	1	1	29	0	0	3	3	0	**60**	105
Total groups of calls	51	859	84	32	21	195	554	44	11	62	1502	14	146	3575