Large-scale automated acoustic monitoring of birds and the challenges of field data

ULISSES MOLITERNO DE CAMARGO

LUOVA

Finnish School of Wildlife Biology, Conservation and Management

Faculty of Biological and Environmental Sciences

University of Helsinki

ACADEMIC DISSERTATION

To be presented for public examination with the permission of the Faculty of Biological and Environmental Sciences of the University of Helsinki in Auditorium 2, Info Centre Korona (Viikinkaari 11) on November 15th 2019 at 12.00 o'clock noon.

Helsinki 2019

SUPERVISED BY: Prof. Otso Ovaskainen

Faculty of Biological and Environmental Sciences, University of Helsinki, Finland

REVIEWED BY: Professor Karl-L. Schuchmann
Federal University of Mato Grosso, Brazil and University of Bonn, Germany

Dr. Alison Johnston

Laboratory of Ornithology, Cornell University, USA

EXAMINED BY: Prof. David A. W. Miller

Department of Ecosystem Science and Management, College of Agricultural Sciences, The Pennsylvania State University, USA

CUSTOS: Prof. Otso Ovaskainen

Faculty of Biological and Environmental Sciences, University of Helsinki, Finland

MEMBERS OF THE THESIS ADVISORY BOARD COMMITTEE:

Dr. Patrik Byholm

Bioeconomy Research Team, Novia University of Applied Sciences, Finland

Prof. Jukka Corander

Department of Mathematics and Statistics, University of Helsinki, Finland

ISBN 978-951-51-5528-3 (paperback)

ISBN 978-951-51-5529-0 (PDF)

http://ethesis.helsinki.fi

Unigrafia

Helsinki 2019

“I took a walk in the woods and came out taller than the trees”

-Henry David Thoreau

Contents

ABSTRACT 06
TIIVISTELMÄ 07
SUMMARY 09
1 INTRODUCTION 09
1.1 Automated Acoustic Monitoring 10
1.2 Automated Sound Recording 10
1.3 Automated Species Identification 11
1.4 Ecological inference 17
1.5 Challenges for the automated acoustic monitoring of birds 19
2 THESIS OUTLINE 20
3 MATERIAL AND METHODS 21
3.1 Study area and sampling design 21
3.2 Audio databases and study species 22
3.3 Data pre-processing for automated acoustic identification 23
3.4 Audio classification 23
3.5 Data post-processing and ecological modelling 24
4 RESULTS AND DISCUSSION 25
4.1 A reliable method to search and retrieve query sounds against libraries of references . 25
4.2 The reliable identification of vocalizations in continuous data from the field 26
4.3 Ecological inference from extensive acoustic datasets 27
5 CONCLUSIONS 28
6 ACKNOWLEDGEMENTS 30
7 REFERENCES 34

CHAPTERS

I: PROTAX-Sound: A probabilistic framework for automated animal sound identification . 45
II: Animal Sound Identifier (ASI): software for automated identification of vocal animals . . . 73
III: Spatio-temporal scaling of biodiversity in acoustic tropical bird communities 121

The thesis constitutes of the following articles, which are referred to in the text by their Roman numerals:

I **de Camargo, U.M.,** Somervuo, P. & Ovaskainen, O. (2017) PROTAX-Sound: A probabilistic framework for automated animal sound identification. PLOS ONE, 12, e0184048.

II Ovaskainen, O., **de Camargo, U.M.** & Somervuo, P. (2018) Animal Sound Identifier (ASI): software for automated identification of vocal animals. Ecology Letters, 21, 1244-1254.

III **de Camargo, U.M.,** Roslin, T. & Ovaskainen, O. (2019) Spatio-temporal scaling of biodiversity in acoustic tropical bird communities. Ecography 42: 1–12.

Table of contributions

	I	II	III
Initial idea	UC, OO	OO, UC	UC, TR, OO
Study design	UC	UC	UC
Empirical Data	UC	UC	UC
Analysis	UC, PS, OO	UC, OO, PS	OO, UC
Manuscript preparation	UC, OO	OO, UC	UC, TR, OO

OO: Otso Ovaskainen
PS: Panu Somervuo
TR: Tomas Roslin
UC: Ulisses Camargo

ABSTRACT

Modern technologies for the automated acoustic monitoring of animal communities enable species surveys that yield data in unprecedented volumes. Interpretation of these data bring new challenges related to the need of automated species identification.

Coupling automated audio recording with automated species identification has enormous potential for biodiversity assessment studies, but it has posed many challenges to the effective use of techniques in real-world situations.

This thesis develops new methods in the field of bioacoustics applied to automated monitoring of vocal species in terrestrial environments. Specifically, I developed automated methods to classify acoustic ecological data generated under the two most common contexts used in ecology: identification of vocalization data stored in acoustic libraries of sounds and identification of vocalizations in audio data collected from the field, through e.g., acoustic monitoring programs.

The methods bring key developments across the entire pipeline for automated acoustical identification, connecting techniques from the data acquisition in the field to the ecological modelling of data identified utilizing automated classification methods.

I show the performance of methods over huge datasets, compare them with alternative cutting-edge techniques and provide an ample study case of Amazonian bird communities to show the tools in practice. The methods in this thesis are available as open source and ready-to-use software capable to work directly on field data collected from acoustic monitoring efforts.

TIIVISTELMÄ

Nykyaikaiset/modernit teknologiat/tekniikat eläinyhteiskuntien automattiseen akustiseen monitorointiin mahdollistavat lajitutkimuksen, joka tuottaa ennennäkemättömän määrän tutkimusaineistoa. Tällaisen tutkimusaineiston tulkinta aiheuttaa uusia haasteita (kuten) tarpeen automatisoituun lajitunnistukseen.

Automatisoitu audiotallennus yhdistettynä automaattiseen lajitunnistukseen luo uusia mahdollisuuksia biodiversiteetin inventointiin/ luontotyyppien seurantatutkimukseen, mutta ne ovat myös aiheuttaneet monia haasteita menetelmän käyttämiseen todellisissa tilanteissa.

Tämä tutkimus kehittää uusia menetelmiä maalla elävien ääntelevien lajien automaattiseen seurantaan bioakustiikan tutkimuksen kentälle. Kehitin ennenkaikkea automatisoituja menetelmiä kahden tyypillisimmän akustisessa ekologiassa käytetyn aineiston; lajiäänitteiden tietokantojen sekä lajiäänitteiden maastoaineiston tallenteiden, luokitteluun.

Nämä menetelmät kehittävät olennaisesti koko automatisoidun akustisen tunnistuksen kenttää yhdistäen maastoaineiston automatisoidun keruun automaattisten luokitusmenetelmien avulla tunnistettujen tietojen ekologiseen mallintamiseen.

Osoitan menetelmien toimivuuden (käytännössä) erittäin suurten aineistojen avulla vertaillen niitä tämänhetkisiin huipputekniikoihin sekä tarjoan laajan Amazonin lintuyhdyskuntia koskevan tapaustutkimuksen/tutkimusesimerkin osoittaen näin välineiden/menetelmien toimivuuden käytännössä. Tutkimuksessa tuotetut menetelmät ovat saatavilla avoimen lähdekoodin sekä käyttöönotettavan/toimivan ohjelmiston muodossa maastoaineiston käsittelyä varten.

SUMMARY

Ulisses Moliterno de Camargo

Helsinki Lab of Ornithology, Finnish Museum of Natural History, University of Helsinki

1 INTRODUCTION

A key challenge for ecological research is to understand the interactions between biotic and abiotic factors affecting the spatio-temporal dynamics of individuals, populations, species and communities (Ovaskainen et al. 2017). Understanding of fundamental ecology, together with robust and cost-effective methods of large-scale and long-term biodiversity monitoring forms the basis for more applied research, such as the evaluation of the consequences of environmental change (Laurance et al. 2011). However, acquiring adequately replicated large-scale and long-term data remains a major challenge both in ecological research and biodiversity monitoring (Ferraz et al. 2008), especially for remote areas and species-rich communities, and for taxa that require experts for species identification (Cohn-Haft et al. 1997).

For vocal taxa such as birds (Potamitis 2015, Campos-Cerqueira and Aide 2016, Frommolt 2017), bats (MacSwiney G et al. 2008, Armitage and Ober 2010, Newson et al. 2017) and frogs (Measey et al. 2017, Dutilleux and Curé 2018), automated audio recording offers a powerful tool for acoustic monitoring schemes, capable of capturing information at adequate ecological scales (Aide et al. 2013). In recent years, the ability to track and monitor wildlife populations has greatly increased and numerous types of sensors are currently available for ecological studies (Porter et al. 2009, Rundel et al. 2009). These technical developments have enabled a significant increase in the amount (Campos-Cerqueira and Aide 2016) and accuracy of data (Trifa et al. 2008, Collier et al. 2010), but simultaneously they pose new challenges in the storage, processing, analysis, archiving and interpretation of big data (Ovaskainen et al. 2018, Stowell et al. 2018).

With the advent of appropriate recording units (waterproof, wireless, easily transportable and with large memory capacities; Mennill et al. 2012, Merchant et al. 2015, Whytock and Christie 2017, Hill et al. 2018), the major challenges for large-scale biological monitoring are no longer related to data collection in the field but to the technical aspects of automatic data processing and, most importantly, the interpretation of the data we generate. From the technical side, there is need to create analytical tools capable to automatically manage, process and analyze big data. In relation to data interpretation challenges, there is still need to understand the observational process for passive acoustic recorders. Very little is known about how the acoustically detected and identified species relate to the total species community.

While several methods have been proposed to automatically identify species from audio recordings (e.g. work reported in LifeCLEF classification challenges; Goëau et al. 2015, Goëau et al. 2016, Goëau et al. 2017), reliable automated identification algorithms capable to operate in large scale and that would reach even close to the same level of species identification as obtained by manual identification by experts are still lacking (Stowell et al. 2016). Additionally, most of the available methods lack the architecture capable to integrate the many steps of an automated pipeline into a ready-to-use software useful for biological monitoring purposes.

Coupling automated audio recording and automated species identification has enormous potential for biodiversity assessment studies but has posed many challenges related to effectively use and interpretation of the techniques in real-world situations and by non-experts. This thesis brings analytical tools to address such challenges. In the following sections, the main technical and ecological concepts and challenges involved in such methods are introduced. The remaining text is organized to provide a summary of the research objectives, material and methods, results, discussion and conclusions of the Chapters I–III of this thesis.

1.1 Automated acoustic monitoring

A pipeline for automated acoustic monitoring usually contains the following basic elements (Figure 1): I) a field protocol for automated sound recording; II) a framework to process audio files and identify species automatically; and III) statistical tools capable to extract relevant biological information from the audio recordings.

1.2 Automated sound recording

Any bird sampling method have its advantages and disadvantages. Depending on the objectives of a particular study, some are more suitable than others, and there is not a single method that will always be the optimal approach to every biological question. Historically, studies on natural bird populations have been based on methods that capture individuals, such as mark-recapture techniques, or more recently, tagging birds with geolocators or GPS devices (Bibby et al. 2000). Although marking techniques can allow the unequivocal identification of individuals, they are often costly and time-consuming, especially in areas which are difficult to access. Further, the use of invasive methods can have both short-term (due to the capture and handling process; Caro 1998) and long-term effects on individuals, including avoidance of the capture area (Marques et al. 2013), stress-related susceptibility to disease (Menu et al. 2000, Schmutz and Morse 2000), increased susceptibility to predation and poorer reproductive success (Moorhouse and Macdonald 2005). Non-invasive methods such as visual and aural counts (e.g., distance sampling methods), or the use of aerial photos, videos and audio recording techniques, on the other hand, can offer an alternative with fewer adverse welfare implications and avoid the problems associated with biases from animal handling (Terry et al. 2005).

Regarding acoustical methods, virtually all vocal species have unique acoustic patterns that differ significantly among species and in many cases also among individuals (Kroodsma and Miller 1996), potentially yielding a natural tag that allow for population monitoring without a need for direct contact with the study objects (Petrusková et al. 2016). At least in theory, vocalization signals can be used to obtain both life-history information (species, sex, identity or behavior) and ecological information (habitat use, survival, recruitment, immigration and emigration), and useful for many animal studies (Payne et al. 2003, MacSwiney G et al. 2008, Scott Brandes 2008, Enari et al. 2017, Suter et al. 2017, Wrege et al. 2017).

Equipment is constantly evolving to lower costs and better quality. In the last ten years, several alternatives for acoustic recorders became available for the scientific community (e.g., Audio Moth, Sole, Wildlife acoustics SM4; Mennill et al. 2012, Whytock and Christie 2017, Hill et al. 2018). Researchers are exploring now alternatives to expensive commercially available options and adopting field devices designed and built in partnership with engineers (see Hill et al. 2018 for a review on recent equipment).

Although very promising, one must emphasize that acoustical methods are still in its infancy as a bird sampling method and the majority of published literature is still experimental and of small-scale nature. There are advantages of applying automated acoustic recording to the monitoring of vocal species (Box 1). These advantages, however, are matched by the high complexity of acoustic landscapes (Box 2) and the many limitations of techniques available to analyze and interpret such data. There is no single method suitable to answer all ecological questions, and any researcher willing to apply acoustical methods needs to account for the pros and cons of using such techniques to answer the biological questions of interest.

1.3 Automated species identification

The set of techniques used in automated species identification largely vary depending on the goals of the study and the type of output data to be generated. The sections below aim to provide an overview of the techniques commonly utilized in automated species identification. Automated frameworks usually have the following general steps: I) the pre-processing of the audio files, generally aimed to prepare the data to be classified by filtering out unwanted noise and segmenting the candidate sounds to be classified; II) the extraction of relevant audio features and statistical summaries useful for audio classification; and III) the training of the statistical models, audio classification and the extraction of relevant biological information that will be used in downstream statistical analysis.

Pre-processing audio: signal enhancement and segmentation. The pre-processing of audio recordings aims to increase overall signal-to-noise ratio and segment the audio into regions which are informative for the classification algorithms. The processing usually starts with the cleaning of audio for background noise, followed by the segmentation of audio into regions of interest that contains the target signals to be classified.

The removal of background noise is a very important step as low recognition accuracy is often attributed to the common case of noises overlapping with signals (Baker and Logue 2007). However, it is difficult to clearly define what is noise and often impossible to eliminate noise without degrading the signal of interest itself (Priyadarshani et al. 2018). The Wiener filter technique is often used to eliminate noise in speech recognition and animal sound identification (Jingdong et al. 2006). Alternatively, one can treat the spectrograms as images and apply image processing methods to reduce noise (Potamitis 2014, de Camargo et al. 2017).

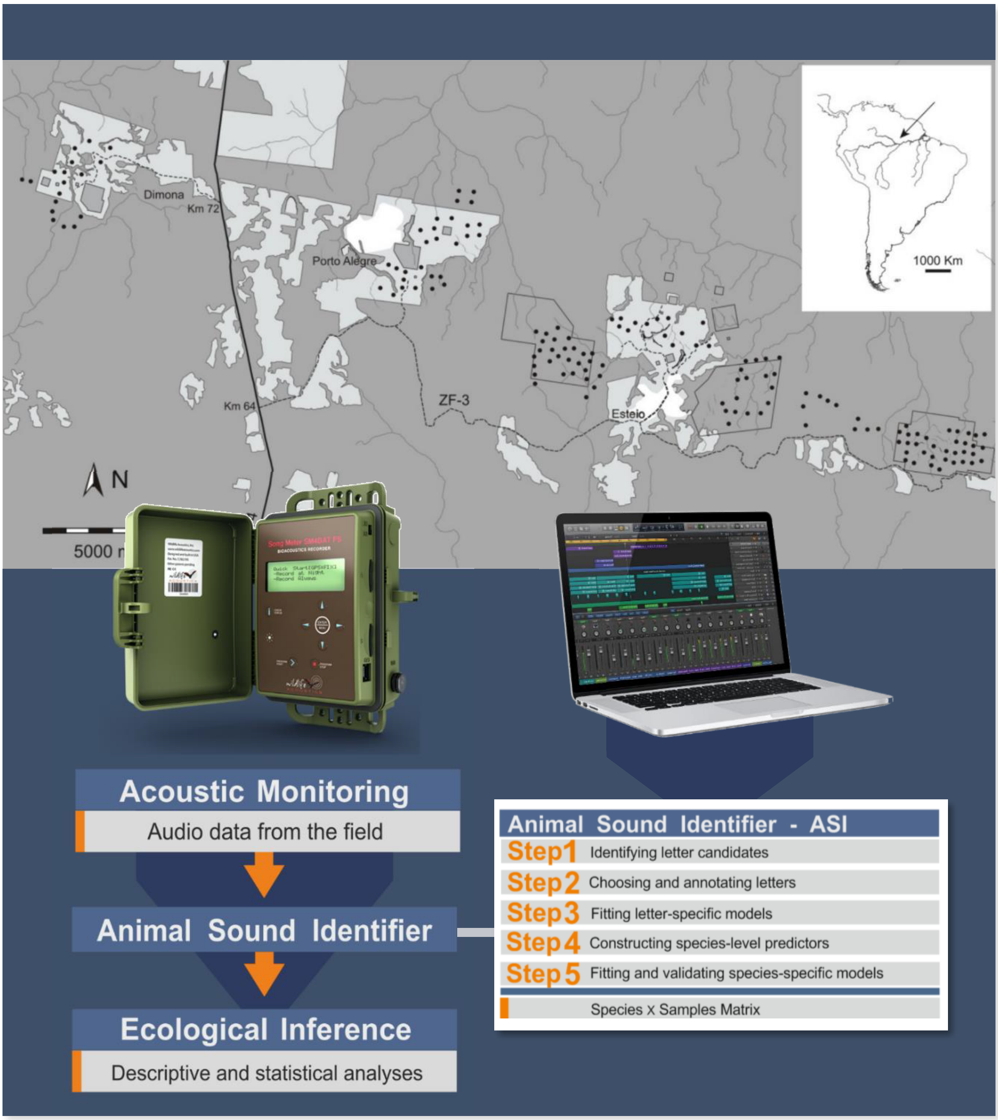

Fig. 1. Acoustic monitoring pipeline. Example of automated pipeline consisting of three phases: a) A sampling protocol employing autonomous recorders to collect and store data from multiple sites in the field; b) The use of software for automated species identification (e.g., Animal Sound Identifier, Chapter II); and c) Ecological inference with statistical analysis using the biological data automatically extracted from the field audio recordings (e.g., Chapter III).

Box 1. Advantages of automated acoustic recording in biological monitoring

1) Unavailability of better alternative methods: Acoustic methods can provide ways to detect individuals where other approaches for data collection would not be applicable or where human observers would not be able to operate efficiently. Nocturnal species, for instance, may be readily detected acoustically (Bardeli et al. 2010, Goyette et al. 2011, Digby et al. 2013, Oppel et al. 2014). Species that are difficult to detect visually because of their small size or cryptic behaviour may still be detectable acoustically (Hüppop et al. 2006, Hilje and Aide 2012, Digby et al. 2013). Similarly, many species can be difficult to identify by sight but are easily distinguishable by vocalizations (Kerosky et al. 2012).

2) Increase in the amount and scale of data: Many ecological phenomena take place over large spatial or temporal scales and their study requires equally extensive datasets. Automated acoustic recording allows the collection of large amounts of data unobtainable using other, more labour-intensive method (Ribeiro et al. 2017). Species that are rare or vocalize infrequently might not be detected in short-term studies and obtaining enough temporal replication may need extensive data collection periods which is not feasible when using human observers (Bardeli et al. 2010, Goyette et al. 2011).

3) Less requirements for human participation: Automated acoustic methods cut down on the need for experts in the field and make it easier to employ staff without previous experience or training on species identification. The analysis of field recordings in a computer, with the ability to consult reference material and revisit unclear sound segments can be a less daunting task to non-experts than e.g. performing point counts in the field (Goyette et al. 2011).

4) Better accuracy and quality of data and analysis: With human observers, differences in observer expertise (observer bias) and fluctuating attention levels (observer fatigue) can cause variation in the sensitivity and accuracy of observations. Passive acoustics offer a way to standardize data collection over different times of day or seasons of the year, simultaneously at various locations (e.g., Larkin et al. 2002, Carstensen et al. 2006, Wrege et al. 2010). Passive acoustic methods create a permanent record of the monitoring period, and thus allows error checking at any time, and future use of more powerful identification methods, as well as the availability of raw data to be used to explore new research questions.

BOX2. Soundscape diversity and its measurement: signal and noise.

Within the living sources of sounds, insects are the most frequent ones. Crickets and cicadas produce sounds from 3-4 kHz and 6-8 kHz (Aide et al. 2017). Amphibians are also very common and use their vocalizations to attract pairs (Gerhardt 1994). Frequencies of frog choruses range from 2-5 kHz. Almost all birds use sound to communicate, sometimes exhibiting very complex patterns. Sounds are used to attract mates, defend territories and send danger alarms (Kroodsma 2015). Most bird songs occur in the 2-6 kHz range. Many terrestrial mammals produce sounds and bats produce ultrasounds to locate prey (Kloepper et al. 2017).

Communication patterns can be complex. Birds are well-known for their circadian singing patterns, with vocal activity peaking early in the morning and in the early evening (Henwood and Fabrick 1979). When multiple animals sing at the same time, the spatio-temporal dynamics of the soundscape rapidly becomes very complex. In order not to overlap, animals vocalize in different frequencies, varying the duration and moments singing events happen (Young 1981).

Background noise can be as diverse as signals, both the stationary noise (e.g. constant noise of rain, wind, rivers, etc.) and non-stationary noise (i.e. varying in frequency and time), such as sounds from other animals, human-generated (airplanes, cars, gunshots, etc.) or environmental sounds like tree-falls, branches cracking, etc. In many situations, sound sources happen independently from each other and often unpredictably. Other sounds, however, are correlated either in positive (e.g. close to rivers frogs sing more actively) or negative way (e.g. when a predator vocalize the prey stay quiet). All this diversity sums up to constitute complex soundscapes that sound monitoring schemes and equipment need to account for (Fig. 2).

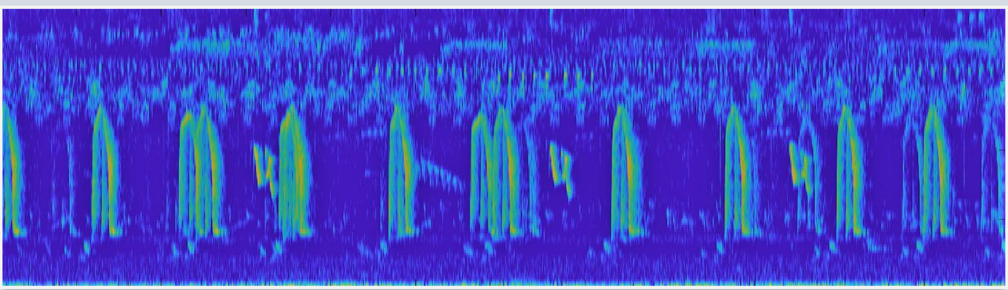

Fig 2. Soundscape complexity and acoustic recording. Spectrogram image of 1-minute audio segment from Amazon rainforest containing insects and multiple bird species vocalizing. Background noise consists of wind and a nearby stream.

No matter the technique, however, structured noises that truly resemble animal sounds cannot be fully eliminated and are further processed to be eliminated only at the classification stage (i.e., by the classifier assorting a very small probability of it belonging to target species).

The way audio is segmented is determinant for the quality of final species classification. A good segmentation algorithm can provide useful data regardless of the choice for the subsequent algorithm used for classification. By just segmenting audio to regions of interest, i.e. containing the query vocalizations to be classified, and even without a classifier, human analysts can benefit from semi-automated approaches and concentrate identification effort only to regions containing sounds of interest (Andreassen et al. 2014).

It is desirable to divide a song or a series of calls into smaller units such as syllables but the definition of where to segment is not a simple task. Merging syllables that are very close to each other is a common practice (Fagerlund 2007) or segment it more carefully by hand (Anderson et al. 1996, Chen and Maher 2006, Somervuo et al. 2006, Fox et al. 2008).

Many segmentation methods work on the assumption that the sections where the birds sing carry more energy than the other parts of the recording (Harma and Somervuo 2004, Somervuo et al. 2006, Juang and Chen 2007, Towsey et al. 2012, Jinnai et al. 2018). However, this assumption is valid for recordings with low background noise levels, being hardly the case for most data recorded in field conditions of complex environments (Box 2).

Another approach is to segment the spectrogram image of audio by applying image processing techniques (e.g. median clipping techniques; Potamitis 2014, Lasseck 2015a, Lasseck 2015b). Other techniques such as shape morphology (Potamitis 2015) and morphological opening (erosion and dilation; de Oliveira et al. 2015) are commonly used to further improve image-based segmentation methods.

Feature extraction. In the context of acoustic species recognition, a feature is an individual measurable property of an audio recording which provides useful information for an algorithm to perform species classification (Sandsten et al. 2016). Feature vectors can be viewed as equivalent to the vectors of explanatory variables used in statistical models such as linear regression. As when choosing explanatory variables, choosing informative, discriminative and independent features is crucial for building effective classification algorithms (Aggarwal 2014).

A set of numerical features is usually described by a vector which can be used directly to feed classification algorithms or further processed (e.g. multiplied by weighted vectors, logarithmic transformed, etc.) to construct new features which may have better resolution for classification. However, the many techniques used to transform and explore features can generate new features that are redundant or too large to be managed in practice. Therefore, a preliminary step in many applications consists of selecting a subset of features to facilitate model learning, to improve model generalization and its interpretability (Tang et al. 2014).

In order to reduce the dimensionality of the feature space, several dimensionality reduction techniques can be used (e.g., PCA-related techniques; Sorzano et al. 2014). Choosing and selecting the best set of features is a field of science of its own, related to feature engineering and much explored in models with supervised learning (Tang et al. 2014). It requires the experimentation of multiple possibilities and some level of domain-specific knowledge. The process of automating feature engineering is called feature learning, where an algorithm not only uses features for learning patterns but learns the best features itself (Stowell and Plumbley 2014).

There are many toolboxes readily available to easily extract acoustical features (see Moffat et al. 2015 for a detailed evaluation of the major feature extraction tools). Common approaches are to use the short-time Fourier transformed data in a scale that matches how humans hear sounds (i.e., the perceptual features, such as the Bark and the Mel scale), to use Linear Predictive Coding (LPC) and its extension, the Linear Prediction Cepstrum Coefficients (Zbancioc and Costin 2003); and Mel Frequency Cepstral Coefficients (MFCC). MFCCs has been widely used in human speech recognition (Makhoul and Schwartz 1995) and extended to animal vocalizations (Kogan and Margoliash 1998, Clemins et al. 2005, Briggs et al. 2009, Chen and Li 2013). An alternative path of research is to use Discrete Wavelet Transformation features (Bastas et al. 2012) with some successful applications to birdsongs (Selin et al. 2006).

Model training and audio classification. Machine learning methods generally take feature vectors as inputs and compute their representations in order to cluster similar inputs together in the feature space. The goal is to find features that makes the examples of one vocalization from one species similar to each other, but dissimilar to other vocalizations from any other species, i.e., easily separable by an algorithm.

Once a feature representation has been chosen, feature vectors extracted from each segment of the sound file can be fed into a standard statistical model (e.g., machine learning algorithms), which will group those representations that are similar either in an unsupervised (i.e., without human labelling of the vocalizations used for training) or supervised fashion (i.e., species labels are provided by human experts and used in model training). Alternatively, selected examples of a vocalization can be treated as templates of a particular type of vocalization, and the distances between vectors representing each reference vocalization and a query vocalization can be computed, with the closest pair being directly declared a match to the query vocalization. There are optimized distance measures derived exclusively for particular applications, such as the geometric distance (Jinnai et al. 2018) used by SoundID software.

There are multiple machine learning algorithms, and the focus of this thesis is on techniques utilized in the classification of animal sounds. A variety of different types of neural networks have been used for birdsong recognition (e.g., Cai et al. 2007, Sprengel et al. 2016). Support Vector Machines are another approach frequently used (e.g., Fagerlund 2007). There has also been research applying decision trees (Lasseck 2015a) and random forest and bagging techniques (Campos-Cerqueira et al. 2017, de Camargo et al. 2017). These techniques are relatively simple to implement and have shown good results for birdsong recognition (Digby et al. 2013, Potamitis 2014, Stowell and Plumbley 2014, Lasseck 2015b), however, they perform classification by dividing the audio files into frames and do not necessarily consider how the dynamics of sounds evolve in time. Hidden Markov Models are capable to incorporate these dynamics by creating a time-dependent probability distribution showing how likely certain syllables are to follow from others in a sequence (Kogan and Margoliash 1998, Kwan et al. 2004, Katahira et al. 2011, Kirschel et al. 2011).

Finally, spectrograms cross-correlation can be used to feed a simple classifier by taking a segment of the spectrogram and computing the cross-correlation with a set of reference template calls (e.g., Lasseck 2015a) . This method is simple yet proven to be successful when scanning for a specific species with limited call variations (Goyette et al. 2011, Frommolt and Tauchert 2014, Ulloa et al. 2016).

1.4 Ecological inference

The use of automated species identification methods in ecological and biodiversity research is limited both by the technical challenges pertaining to the methodologies themselves and the ability to meaningfully interpret results from acoustic data. Using acoustic data to characterize communities is naturally nothing new in itself (Cohn-Haft et al. 1997) and modeling frameworks based on detection/non-detection data generated by aural detections (e.g., site-occupancy models) are generally suitable to analyze data generated by automated methods.

One of the main uses of acoustic data is to characterize species' use of habitats across space and time utilizing e.g. occurrence probabilities. The simplicity of input data used in such models (detection/non-detection data) makes occurrence approaches ideal to model the data obtained using automated classification methods. Occurrence is widely used to understand habitat relationships and patterns of species co-occurrence, population dynamics and species phenology (MacKenzie et al. 2003, Chambert et al. 2015). Another benefit is that scientific literature is rich in modelling frameworks that account for observational errors in detection/non-detection data used in occupancy models (McClintock et al. 2010, Miller et al. 2012). Chambert et al. (2018) developed methods specifically to account for both false negatives and false positives in data generated by autonomous identification methods.

The challenges related to modeling and interpretation of data in resolution finer than detection/non-detection data are much greater (e.g., trying to extract abundance information from acoustic data). These include the difficulty to separate between the number of vocalizing individuals and variation in vocal activity, i.e. is the observed variation due to changes in the number of individuals or to changes in individual vocalization activity? Vocal activity levels have potential to provide more detailed information on the species' activity and behavior but requires clear choices on metrics used to quantify and interpret the variation in activity and behavior (Figueira et al. 2015, Pérez-Granados et al. 2019). Sound production in birds can vary both seasonally (depending e.g. on hormonal cycles) and in space (due to habitat, species composition, e.g., presence of predators potentially change vocal behavior). This natural variation, together with the lack of robust methods to separate and interpret the effects of the observational process make it much more complex to use vocal activity as a reliable proxy for animal abundance.

Another way to interpret numbers of individuals is through territory mapping, as demonstrated by (Bardeli et al. 2010) with Savi's warbler (Locustella luscinoides). However, territory mapping is a very intensive method to estimate abundance and is not used often for this reason. Distance sampling is now a much more common and efficient technique used to study birds and their abundance patterns (Buckland et al. 2001). Localization of calls can also be possible using arrays of microphones

(Blumstein et al. 2011, Frommolt and Tauchert 2014). Lastly, the number of individuals can be approached directly by attempting to identify separate individuals instead of species (e.g., Kirschel et al. 2011).

Automated acoustic species identification is still in its infancy as a methodology to study real-life biological questions in terrestrial environments. New modelling methods and insights on how to interpret such data still of experimental nature and are blooming together with the new field (Barré et al. 2019). We will see many more developments of the field in the upcoming years (Servick 2014, Burivalova et al. 2019).

1.5 Challenges for the automated acoustic monitoring of birds

The field of automated acoustic monitoring has seen much research in optimizing recording equipment and identification algorithms but relatively less emphasis was given to the development of complete architectures (i.e., capable to support from the data acquisition in the field, storage and backup, data management, pre-processing, analysis, post-processing and output results). With acoustic data collections growing in quantity, cloud-based systems need to be connected to library of sounds and provide users with services for e.g. the search and identification of query audio. Systems need to be broad enough in scope and access so both scientific users and enthusiasts can benefit of such data. This is no simple task.

The LifeCLEF classification challenges (Goëau et al. 2015, Goëau et al. 2016, Goëau et al. 2017) have yield a huge progress in proof-of-concept research related to automated bird species recognition. On each edition algorithms perform better when compared to previous years. However, only recently they started exploring real-world data continuously collected from the field. More importantly, however, the results would bring a more practical benefit if the focus of the challenges would move from a competition between techniques to a scheme aimed to combine successful models. It is common case that one algorithm is good at detecting e.g. high-pitched birds, while others are good at detecting e.g. low-pitched birds. One can combine their outputs by using different techniques (e.g. weighted average, majority vote, etc.) in order to improve classification. That is the idea behind bagging of models, for example (Prasad et al. 2006). When combining techniques, independent models could be explicitly designed to excel in different tasks, or alternatively, they could simply be multiple attempts to solve the same problem (e.g., by using different combinations of feature sets and classifiers). At the end of the pipeline results should be combined to produce a final probabilistic output that is reliable and meaningful for the users (de Camargo et al. 2017).

Lastly, most approaches are not implemented as ready-to-use software, limiting its usage by users lacking the programming or statistical skills needed to implement and optimize the algorithms by themselves. Additionally, these cutting-edge techniques have proven successful but at the cost of significant computational resources (i.e., need to run in super clusters capable to process terabytes of data) making it impossible for home users to process audio files in their home computers in feasible time.

2 THESIS OUTLINE

In this thesis, I present three articles that address different angles of applying automated acoustic methods to the biological monitoring of species. Emphasis is given to the quantification of uncertainty and the reliability of automated identifications, as well as the direct application of the methods to continuous data collected from the field. By using tropical bird communities as a case study, I develop methods that increase the practical toolbox of techniques used in Bioacoustics. Specifically, I aim at answering the following questions:

How to reliably identify, search and retrieve query sounds against libraries of references?

The goal of Chapter I is to adapt approaches for assessing identification uncertainty of DNA barcoding data to the context of acoustic species identification, thus enabling a robust quantification of acoustic identification uncertainty. The method is aimed to operate e.g. as part of a framework to classify and organize data uploaded by users of online audio libraries. Online acoustic libraries such as the Macaulay Library of the Cornell Lab of Ornithology, Xeno-canto Library and the Internet Bird Collection usually bring audio segments with vocalizations representing a single species aimed to be used as references for its vocal repertoire. It is common that for many species there are only one or a few recordings representing the vocalizations and often the vocal repertoire is not complete in the database. Users should be capable to automatically compare their own audio data against these references and get the species identifications together with an assessment of the classification uncertainty, so they can be confident of their identifications.

How to reliably identify species vocalizing in continuous data from the field?

The goal of Chapter II is to provide a reliable and ready-to-use software to the automated identification of vocal species from continuous field recordings. Acoustic monitoring studies usually have focus on the acoustic community and recordings cover entire soundscapes, frequently and over the long-term (by collecting many samples across weeks, months and years). This approach usually generates data containing multiple target species (e.g., insects, frogs, birds and mammals) in the same audio recording.What makes real-world data extremely challenging is that audio from field conditions are complex, with vocalizations of the target species overlapping with each other and with many types of background noise. Chapter II aims to develop software to deal with the complexity of soundscape data, ideal to be used in monitoring schemes of target species across huge quantity of audio segments recorded in real field conditions.

How to extract ecological information from extensive acoustic datasets?

The goal of Chapter III is to illustrate the potential of the methods in this thesis to process big ecological data. The overall aim of this case study is to examine whether the impacts of habitat fragmentation that happened 30 years ago in Amazon can still be heard in the soundscape of the modern-day landscapes. In particular, I ask what is the relative contribution of the spatial, temporal and habitat dimensions to variation in bird acoustic communities in a previously fragmented tropical rainforest? And does the functional diversity of birds scale similarly with space and time as does species diversity, when both are recorded by bioacoustics means?

While all chapters in this thesis bring specific study cases, the methods and results are general and ready to be implemented with different data, focusing either on specific taxa or on entire acoustic communities. Altogether, Chapters I-III develop and apply automated methods capable to perform the acoustical biological monitoring of vocal species.

3 MATERIAL AND METHODS

3.1 Study area and sampling design

The empirical bird data utilized in this thesis (in Chapters II and III) comes from sites from the Biological Dynamics of Forest Fragments Project (BDFFP), 60 km north of Manaus, Amazonas, Brazil. The area was completely covered by old growth forests until 1980, when newly established cattle ranches located East-West across the BR-174 highway started clear-cutting forest. Approximately 15% of the area was deforested in the early 1980's but gradual abandonment of pasturelands starting in the middle of that decade resulted in the current mosaic of forest fragments, secondary forest, old-growth and pastures (Fig. 2; Laurance et al. 2011). Presently, most of the once deforested area is covered by secondary forest. Forty-four of our sampling sites were in secondary forest and 107 in old growth, distributed along more than 40km across the study area. Secondary forest sites range in age of regrowth from 18 to 29 years (Mesquita et al. 2001). To attain independence between points, I kept them as isolated as possible from each other, the minimum distance between a point and its closest neighbour was an average 462 m, ranging from 306 to 906 m; I sampled sites for five consecutive years (2010-2014) during the dry season, between June and October, using 15 to 25 autonomous Song Meter SM2 Digital Field Recorders (Wildlife Acoustics). Each site was sampled with one recorder tied to a tree at approximately 1.5 m above ground, operating from four to six consecutive days. I had to sample sites in blocks of 15-25, but I alternated longitudinal positions and forest types of the blocks to avoid any correlation between sampling time and longitude or forest type. One recording day consists of three hours of continuous recording starting 40 minutes before sunrise. In total, the five-year audio dataset consists of more than 11,000 hours of field recordings.

3.2 Audio databases and study species

Chapter I utilizes tropical bird vocalizations extracted from the Xeno-Canto collaborative sound library (https://www.xeno-canto.org). This is part of the same dataset used in the Bird task of LifeCLEF classification challenges (Goëau et al. 2015, Goëau et al. 2016, Goëau et al. 2017), enabling to compare results from this thesis to classification made by other methods. The dataset comprises the 200 tropical bird species most numerously represented in Xeno-Canto, gathered from field sites in Brazil, Colombia, Venezuela, Guyana, Suriname and French Guiana. Audio files are stereo and recorded at sampling rate of 44,100Hz, with generally good quality but with variation in the level of noise due to e.g. weather conditions and the amount of background species, as common when building reference databases from heterogeneous sound sources.

Chapters II and III utilized the acoustic data collected from field stations at BDFFP sites as described in previous section. The five-year audio dataset consists of more than 11,000 hours of field recordings divided in 661110 1-minute segments for analysis. Chapter II focuses on the nocturnal subset of Amazon species and Chapter III focuses on the diurnal component of Amazonian bird assemblages. The total number of bird species in the study area is almost 400 (Cohn-Haft et al. 1997), as registered by multiple methods (e.g., point counts, mist netting, etc.). The manual aural processing of 300 hours of audio has documented a richness of about 250 species. The species detected by the automated methods (ca. 77) exhibits a wide variety of phylogenetic traits, habitat preferences and foraging strategies, offering a good representation of the local community (Cohn-Haft et al. 1997). Species inhabit primary and secondary forests, pastures, open water and campinaranas. They also differ in their microhabitat use, foraging at the understory, midstory, or at the canopy level, as well as at forest edges, tree-fall gaps and small streams. Finally, there are several differences in the degree of sociality and foraging strategies including mono- and mixed-species flocks, solitary species, species that join mixed-species assemblages at fruiting or flowering trees, army-ants followers and lekking species (Stouffer and Bierregaard 1995, Cohn-Haft et al. 1997).

3.3 Data pre-processing for automated acoustic identification

Methods used in chapters I-III use the same pre-processing techniques for the data. Because of dealing with short and long audio segments, in Chapter I the entire audio clip was used as a single audio segment to be classified, while in Chapter II-III the continuous audio clips were cropped to thousands of 1-minute segments. In all cases stereo channels were mixed together into a mono audio file. A fast Fourier transformation algorithm was used to get data in the frequency domain and the Weiner filter was applied to remove background noise (Jingdong et al. 2006, Ovaskainen et al. 2018).

The de-noised audio files were then used to extract acoustic features useful for subsequent sound classification. Regardless of how the different pipelines handle data, methods in Chapters I-III use the normalized cross-correlation calculated over spectrograms images of the sound segments as the basic acoustic feature. The normalized cross-correlation ranges between 0-1 and it can be used as a similarity measure between two images, with two identical images having a cross-correlation value of 1 (Lewis 1995). Sliding one spectrogram image over the other will generate a vector of similarities over each pixel of the images and the maximum peak is the point that scores the maximum similarity between the two images. This maximum is stored and used to calculate the different predictors used by models from Chapter I and Chapter II-III as detailed in the methods sections of each chapter.

3.4 Audio classification

Models in Chapters I-III utilize supervised learning to perform sound classification. This means that labelled data (i.e., either from a reference database or labelled by an expert as on Chapters I and Chapters II-III, respectively) is used to estimate the parameters of the models and these parameters are then used to apply the models to the query data to be classified. In Chapter I this happens with a clearly defined training phase for the parameter estimation followed by a classification phase to estimate the probabilities of belonging to the study species for each query segment. Chapter II-III utilizes the same idea but the training and classification phases happen at the same time: After each audio segment used for training is classified by the expert, models are re-fit and all audio segments are classified for the probability of species being detected.

In Chapter I a multinomial logistic regression model is used to convert the predictors calculated from cross-correlation features (or outputs by other classifiers) into a prediction of which species the sample represents. The method performs multispecies classification, thus the outcome of the classifier is a vector containing the classification probabilities for all possible species in the reference database, plus the possibility of query audio belonging to an "unknown" class. Each audio segment that is classified has one vector of probabilities of length number of species plus the unknown species class and these probabilities sum up to 1.

Chapters II-III utilize simple logistic regression to map different predictors into probabilities at two levels: in the letter level the model estimates the probability of the target letter (a letter is defined as a sound unit that is selected by an expert and relevant for classification) being present within the highlighted region of the audio track, whereas the species level estimates the probability of the target species vocalizes throughout the 1-minute audio segment. Chapter II-III utilize species-specific models for classification and outputs a matrix of probabilities with each 1-minute audio segment represented by a row and each species represented by a column in the matrix. In this case the probabilities of a given row does not sum up to 1 as an audio segment can have identifications of multiple species singing within the same minute.

3.5 Data post-processing and ecological modelling

Data outputted by PROTAX-Sound and ASI methods both can be used to generate detection data for downstream statistical analysis. In Chapter I, data can be used to generate identifications of query sound segments by taking the species with the maximum probability value. In Chapters II-III the probability matrix can be transformed into detection/non-detection data simply by thresholding the matrix by some chosen value (e.g. 90% probability). Values above the threshold level will get value 1 and 0 otherwise.

Of course, the best way of propagating uncertainty from the classification methods presented in this thesis to the downstream analysis is to take advantage of the full probability matrices, instead of taking the maximum value or using an arbitrary threshold. When building e.g. Bayesian models of bird community dynamics, or joint-species distribution models, the collection of such detection matrices can be considered as a prior for the true occurrence matrix. Then one can sample the posterior distribution of the true occurrence matrix, thus enabling to propagate species identification uncertainty through the community modelling analyses. Other option is to use the data outputted by our methods as input of further modelling frameworks and add other levels of error quantification and corrections against bias (Chambert et al. 2018, Barré et al. 2019).

4 RESULTS AND DISCUSSION

Here I present the most relevant findings of this thesis and discuss how these findings relate to thesis objectives and to broader applications of automated acoustic techniques. In Chapter I, I developed methods capable to reliably search and retrieve query searches against libraries of sounds. In Chapter II, I developed methods to perform the reliable identification of vocalizations in data obtained directly from the field, without the need of using references from databases. In Chapter III, I presented the spatio-temporal scaling of biodiversity in acoustic tropical bird communities, a study case illustrating how essential ecological patterns can be successfully extracted from a massive amount of audio data by using the methods developed in this thesis.

4.1 A reliable method to search and retrieve query sounds against libraries of references. In Chapter I, I have utilized recent developments in probabilistic taxonomic classification methods for DNA sequences to develop PROTAX-Sound, a statistical framework for probabilistic species identification of audio samples. I have demonstrated that PROTAX-Sound is able to convert similarities of audio features extracted from candidate sounds into classification probabilities for each of the target species, making the assessment of species identification uncertainty reliable and ready to be propagated to downstream analyses. The performance of PROTAX-Sound method is shown, as well as its ability to combine different acoustic features and classifiers into an optimized framework for audio classification. The framework is very flexible, allowing any combination of audio similarity measures and classifiers to be used as predictors of PROTAX-Sound model. The main feature of the approach is to provide a probability of placement to each taxon existent in the reference database, as well as a probability of the sound not belonging to any of the species in the reference set. This is a much-needed feature for search engines that provide taxonomical search based on comparison of query audio against a library of sounds.

The distribution of the probabilities outputted by PROTAX-Sound reflected the range of difficulties in acoustic species identification that is faced also by an ornithologist conducting similar identifications manually, making the interpretation of uncertainty very intuitive for the users. For example, in cases with much uncertainty, PROTAX-Sound has the ability to assign the highest probability to 'unknown species', indicating that the similarity between the query sample and the best matching reference sample is no better than the matches between reference samples belonging to different species. Another feature of the method is the ability to provide reliable identification at higher taxonomical levels (e.g., family level), even if the uncertainty at lower levels is very high (e.g., species level). Therefore, sometimes PROTAX-Sound, equally as a human expert, cannot make a confident identification, or is capable to reliably identify a segment only at the level of a group of species, not at the level of an individual species.

Statistical methods with great potential for automated identification are continuously appearing in the scientific literature, and PROTAX-Sound provides a statistically rigorous method to combine the strengths of different techniques. While I have illustrated the use of PROTAX-Sound specifically for identifying bird sounds, it provides a general framework to classify the sounds of any vocal animals. The framework provides a robust starting point for probabilistic identification of animal sounds, making it possible to propagate the unavoidable uncertainty in species identifications to biological inference derived from audio data, or to identification and search engines used in audio libraries.

4.2 The reliable identification of vocalizations in continuous data from the field. In chapter II, I developed Animal Sound Identifier (ASI) with a focus on improving classification of continuous data collected from the field. The methods and pipeline are illustrated by classifying thousands of sampling units (1-min segments of the data) for the occurrences of the vocalisations of 14 Amazonian crepuscular and nocturnal bird species. The key novelty of the method compared to other approaches are that ASI does not require any a priori references of the target vocalisations, but it finds them directly from field recordings. The direct use of field recordings differs from using reference audio files from online libraries in many ways, including less variation on technical recording quality, the type of background noise, and the geographic region from which the vocalisations originate, all factors that reduce classification accuracy. Importantly, ASI generates training data adaptively, thus asking the user to classify only such training data for which classification by the present model would be uncertain, which data are thus especially valuable for improving classification accuracy. ASI was designed not only to provide accurate classifications, but also to make efficient use of human time.

While ASI provides a major step forward on automated classification of animal vocalisations, it clearly involves several limitations that I hope future research efforts will improve on. First, as the classification models are based on training data provided by the user, an upper limit for the performance of ASI is clearly set by the level of expertise of the user. Most obviously, if the user is not able to identify the species behind a certain vocalisation type, ASI will not be able to classify those vocalisation types either. Additionally, the predicted classifications will always involve some uncertainty. Whether or not removing such uncertainty by post-classification validation is possible or necessary depends on the type of the data and the purpose of the study. As the key benefit of ASI is that it is able to classify massive amounts of data rather than a small sample of it, the disadvantage of having some level of classification error is likely to be more than compensated by the ample supply of data, as long as the recall and precision rates are sufficiently high for the signal to dominate the noise.

4.3 Ecological inference from extensive acoustic datasets: The spatio-temporal scaling of biodiversity in acoustic communities. In chapter III, I combined the classic fields of Species–Area and Species–Time Relations with the novel and rapidly emerging field of bioacoustics to derive new insights in community ecology. Using automated analysis of 11,000 hours of audio recordings, I developed Species–Area and Species–Time Relations using acoustic data to examine whether the impact of habitat fragmentation 30 years ago can still be heard in the soundscape of the modern-day landscapes. I found that both species-level and functional diversity accumulate faster in primary forest than in secondary forest, and that soundscape turnover in relatively small space (some hundreds of meters) was much higher than turnover over relatively long time (years). Overall, these findings suggest that habitat modification can be heard as a long-lasting imprint on the soundscape of regenerating habitats even after 30 years of abandon and identify Soundscape–Area and Soundscape –Time Relations based on the automated analysis of acoustic communities as promising tools for biodiversity research, applied biomonitoring and restoration ecology.

Traditionally, biodiversity is assessed using a variety of methods that are generally costly, limited in space and time, and most importantly, they rarely include a permanent database record. Furthermore, most fauna monitoring protocols require the presence of experts in the field in order to document species presence. This study shows how automated techniques can be successfully applied to classify massive amounts of acoustic data, thereby quantifying ecological relations between time, space, and habitats.

On the technical side, the automated identification results showed very good performance over real-world datasets. Out of the 63 species in the study, for 3 species the classification precision of ASI was poor (i.e. lower than 0.5). Out of the remaining species, for 3 species precision was moderate (i.e. between 0.5 and 0.7), for 19 species it was good (between 0.7 and 0.9), and for the remaining 38 species it was very good (i.e., higher that 0.9). The main reason for such a good performance is that ASI method directly utilizes field recordings to build training data. Consequently, variation in training data reflects the variation in the data to be classified, e.g. the type of background noise, the geographic region from which the vocalisations originate, etc.

The ecological results show that past habitat modification can still be heard as a long-lasting imprint on the acoustic communities of regenerating habitats at BDFFP. This finding was evidenced by higher species diversity and much higher functional diversity in primary forest than in the rapidly-regenerating secondary forest, and in larger differences between communities located in these different habitats than between same-habitat communities. We currently lack a good understanding of the rate with which biodiversity can be expected to revert to their original state after disturbance (Gardner et al. 2007, Laurance 2007) and of the right indicator to measure this. At the BDFFP sites, secondary vegetation is often high, and some species groups appear to have reverted to the pre-fragmentation state (Dunn 2004, Quintero and Roslin 2005, Stouffer et al. 2006). Our findings suggest that the soundscape provides a sensitive indicator of enduring effects of disturbance, and that automated bioacoustics provide accurate tools for recording them (Deichmann et al. 2018, Burivalova et al. 2019).

5 CONCLUSIONS

While rapidly developing and receiving increasing interest among biologists and conservationists, automated acoustic monitoring is still as an early stage methodology to be adopted by non-experts and widely used in large-scale studies to address complex biological monitoring questions. The majority of published literature is still methodology-oriented or, if applied to real-life biological situations, of an experimental and small-scale nature. My goal in this thesis was to explore both of this challenges and purpose new methods in the form of accessible tools for non-experts.

Automated acoustic monitoring methods improved by human input may be the key to optimize and enable classification results in practice (Chapter II). Human input may be needed to e.g., control for errors made by the detectors, to optimize methods to work with real data, to validate models, etc. Even simple methods that performs only automated event detection instead of species classification, i.e. all detected events still must be classified manually (e.g., Farnsworth & Russell 2007), is already a significant reduction in the amount of data that needs to be handled manually. In chapter II, I developed ASI method based on this idea: human input is still necessary but should be optimized in order to maximize the gain in information and minimize the effort in manual work. In this thesis, human input is used to provide better training data and improve the classification power of the models. In fact, the combination of information from experts and enthusiasts, together with cutting-edge technical developments is a fruitful path not only in automated species recognition, but in many other applications in Machine Learning. The power or collaborative networks and crowdsourced training data is becoming apparent and more problems are being solved by such approaches (Abhigna et al. 2018). Finally, methods need to be made even less computer intensive and designed as user-friendly software.

In this thesis, I successfully classified data from thousands of hours across hundreds of sites. The new feasibility of taking on such tasks opens new avenues for community ecology, and for biodiversity research (Rajan et al. 2018, Burivalova et al. 2019). The use of automated recorders at a higher number of sites at the same time and with the same technical conditions relieves the concerns related to sampling different sites by different people at different times (Ribeiro et al. 2017). In dealing with the resulting data, we are no longer limited by availability to expert listeners who can identify the species from their sounds in the field, or by the impossible task of manually listening to all audio recordings in a given study (Ferraz et al. 2008, Ribeiro et al. 2017). Instead, experts can make optimal use of their knowledge by annotating candidate sounds, validating and verifying automated classification results (Chapter II). This allows them to free up their expertise from endless hours of routine tasks and instead offer targeted insights were truly needed. The routine tasks can then be performed automatically.

Understanding the dynamics and trends in animal wildlife is an important component in the assessment of environmental change. Studies on complex and megadiverse ecosystems such as the Amazon are highly relevant because they offer an unique opportunity to understand how the spatio-temporal dynamics of species are modulated by the joint effects of intrinsic factors (e.g. influence of hydrographic networks, biogeochemical processes, vegetation dynamics and species interactions), land use change (e.g. decrease in forest area, fragmentation and edge effects) and climate change (e.g. changes in mean and variability of temperature and precipitation). However, such complex interactions among multiple factors pose major challenges for research. A successful study requires a multidisciplinary approach with a strong knowledge of natural history, a well-designed sampling scheme, and efficient analytical skills to extract the biologically relevant information from the data, including the quantification of various types of uncertainties associated with the observed responses. This thesis brings a new set of tools to face such challenges, taking advantage of recent technological and computational advances from multidisciplinary fields to provide an automated framework for ecologists. Continuous research in automation will not allow the classification performance of methods to improve until perfection but will finally make the bridge between successful proof-of-concept methods and the large-scale real applications that are truly useful for biologists and conservationists.

6 ACKNOWLEDGMENTS

Wabi-Sabi is a Japanese philosophy based on accepting the imperfect and transient nature of life. It is originated from ancient tea ceremonies in which the prized utensils were handmade, irregular and imperfect. There is no direct western translation for Wabi-Sabi, but essentially it is the art of accepting and finding beauty in the imperfect, impermanent and incomplete nature of things.

Learning to appreciate the outcomes of this doctoral study was my own "tea ceremony". The following paragraphs are meant to acknowledge the many people that influenced me on this work. I would like to express my deepest gratitude for their personal and professional support. As the Japanese utensils, this section is imperfect, impermanent and incomplete.

Otso, I sincerely thank you for all the opportunities, knowledge, experiences, and enthusiasm to the development of this project during almost seven years. I have learned a great deal from you, from optimizing work time during eating-meeting lunches to the great skiing skills into Russian wilderness. I admire and try to mirror your approach to work in my own professional life. I thank you for the guidance and patience, for the straightforward communication and practical advice on where to move. No matter how busy you were, I always felt your commitment towards the finishing of this work and to my supervision. Thank you for all that and especially for the final push towards the completion of this PhD.

Panu, you have been an informal supervisor of this thesis, but most importantly, the person always available for the technical complains about signal processing, machine learning and general coding of any sort. You are a natural problem-solver and I admire how rapid you can come up with solutions and get work done. Without your valuable input and positive attitude (plus your pessimist mindset), this thesis would not be feasible. Thank you!

The third scientific advisor throughout my studies is Gonçalo Ferraz. Gonçalo, you have been a great friend and inspiration to me. More than the supervision during my Masters in Amazon, you always helped me in so many ways that I lack words to express my gratitude. I took your project of recording Amazon birds as my own flag, turned the task of automated processing acoustic data into my own PhD project and now we are getting the results of such a big effort. I am happy that this thesis provides us tools to finish the project we started in 2010. More important than the achievement itself, I value your friendship and unconditional support for all these years. Our conversations and opinions about politics, arts and life aspects that are external to academy are the greatest memories from your supervision period in Manaus I will carry for life. Your commitment to your beliefs and your contribution to the scientific training of students in Brazil motivated me to push my career as a scientist, and made me believe individuals can truly make a difference.

Next I would like to thank Professor David Miller for examining my thesis and acting as my opponent in the defense. I would also like to thank Professor Karl-L. Schuchmann and Dr. Alison Johnston who took the time and effort to pre-examine an early version of the thesis. The comments I received from you all were very constructive and improved this thesis for sure. I would like also to express my gratitude for Tomas Roslin and Patrik Byholm. Tomas, thank you for your contribution to this thesis. Your addition to the team of collaborators and positive attitude towards the end of the thesis made all the difference. Your final push made it possible for me to accomplish the last PhD tasks and I am forever grateful for that. Patrik, thank you for the support and insights during all the advisory board meetings and for believing in such complex project since the start. Your excitement when seeing the achievements of this thesis was a big boost of motivation for me to make it to the end.

This PhD started way before moving to Finland. The five years of acoustic data collection that motivated all my work in automated species identification started back in 2009 when I moved to Amazon for my Master's under Gonçalo's supervision. This was the happiest period of my professional life, much because of the natural innocence of an early-career student amazed by doing fieldwork in the Amazon forest, but mostly due to the very especial place we were based in Manaus, surrounded by very special people. For the friendship and all sorts of support, a huge thank you to the brave ones living everyday life closer to mine: Murilo Dias and Elisa Xavier, Diego Brandão, Camila dos Anjos, Marco Aurélio and Eduardo Braga. Gabriela Pinho, Laynara Lugli and Camila Duarte, big thanks to you as well. You are all very special friends that I have spent an amazing good time and shared so much in life. Gonçalo's population ecology Lab is the other half of what made that period in Amazon so rich and happy, with a permanent warm place in my heart. For all the Lab time and amazing field seasons we did together, I thank Sandra Freitas, Marconi Cerqueira, Letícia Soares, Thiago Belisario, and especially Gabriel McCrate, Juliana Bonanomi, Francisco Diniz, Luiza Figueira and Humberto Mohr, who shared most of the forest stories and adventures I have to tell nowadays. Being at work was so pleasant because of you, and the field excursions we had are surely unforgettable. Thanks also to the people from BDFFP and INPA in general. Since 2012 I have been living in Finland. One can only imagine the contrast between the +40º C Manaus in Amazon and the -30º C in Helsinki. I was lucky enough to arrive at the Metapopulation Research Group, and its rich diversity of nationalities provided the warmth I needed to survive the period of adaptation, as well as the support for the many years of hard work to come. As that is the moment I am sure to forget some people, I would like to just thank all the MRG members for the support and good times together. I have never been in such a great academic environment to discuss scientific ideas, full of brilliant, humble and very helpful people.

From the MRG gang, I would like to acknowledge in especial Henna Fabritius, Luisa Woestmann, Anna Norberg, Henjo de Knegt, and Silvija Budaviciute for sharing office and the countless hours of bad jokes, procrastination-deep-existential discussions and general peer support. Henna and Luisa in particular, thanks a lot for the extra support and trust to share many conversations and thoughts in important times. Tanjona Ramiadantsoa, Victoria Veach, Johana Eklund, Joona Lehtomäki, Jussi Juosimo, Toby Fountain, Anne Duplouy, Marjo Saastamoinen, Anna-Liisa Laine, Anniina Mattila and Elena Rosa, you were also great companies that kept the mood up for countless times. I would like to finish MRG section by thanking very special people, the main pieces

of the engine that powered MRG for many years: Sami Ojanen and Viia Forsblom (and Jenni Villa and Kaisla Torppa). Your contribution to this work is beyond measure. Thank you for always being available to solve issues in a very smooth and well-humored way, no matter how silly (or repetitive) the question I had. You made the handling of my everyday-life in Finland much simpler! Finally, another important component for the PhD machine to move smoothly: Anni Tonteri. Thank you for your dedication to LUOVA students. You all made all the difference and are part of this big achievement being possible.

Moving a bit to the personal side, big thanks to the many people that provided extensive support and the true friendship that allowed life in Finland to be good, as well as lots of fun to counter balance the most difficult times. I would like to thank Maria Delgado, Matti Pellinen and Nerea Abrego for forming the very special crew joining me together with Otso to the amazing and adventurous times of skiing in Russian wilderness.

The other big source of friends I have made in Finland all happened because of Capoeira. Capoeira is such a beautiful art, the type makes one proud to be Brazilian. It is an activity difficult to explain for non-practitioners, that some think as a dance, or as a fighting style. Capoeira is all of that and much more, a game played with mind and the entire body, and more importantly, formed by a community of friends. It symbolizes the good aspects of Brazilian culture and its African ancestry. All that passionate speech said, I would like to thank all my friends from Cordão de Ouro Capoeira group, especially Mestre Pium, a Master in the art and good friend; also thanks to Diana, Erika, Marjaana, Natalia and Nadja for the friendship. Not from my own group but that became part of my life because of capoeira, big thanks to Roosa, Laura, Ismo, Sabugosa, Wilhelmiina, Liisa, Pirata, Noora, Nora, Niina, Johanna, Eveliina, Nicholas and Viola. Some from capoeira became even closer, as the great friends Ilkka and Sini. Thank you for all the good times together, dinners and weekends in the mökki, full of good wine and food. Apart from capoeira, I would like to mention Léo Custódio and Dan Santos for the friendship, advices and good times together. You all are very important for keeping me sane and to my decision to stay and live my life in this country. Thank you!

Even deeper into personal side, I would like to thank for all the support and love from my family. The type of difficulties related to studying in a foreign country, as different as Brazil and Finland, can only be understood by living it. I was lucky enough to have family’s full support to my life choices, and even the luxury of their visit for two consecutive summers now at the end of my studies. I was

double lucky to have a Finnish family that adopted me fully, which I feel truly welcomed and loved in every time we visit. I would like to thank my Finnish family members Kari and Riitta, Anni, Jari, Eetu, Sirkka and Emmi's many cousins and their beautiful families.

I want to thank in especial for the love and support from my Brazilian family: Mom, you are the best mom, so loving and supportive about my life choices, always putting my education first. Thank you very much for all the support and for being here visiting me and Emmi. It was very important to us! I would like to be more skilled to translate into words your immaterial contribution to my PhD, career and life. I cannot say thank you enough for my big sister Eurídice. Tata, you are a model for me, of a strong woman, a great mother and an excellent person. Throughout my life you have always been a safe harbor whenever I needed advice and courage to pursue my dreams. Thank you for being my sister, my second mother and always present friend! I always enjoy our conversations about all aspects of life, and it is a pleasure to share my thoughts, values and world view with you. And to my dad Antonio Carlos. Dad, I wish so much I just have had more time to share my adult life with you, especially now after so many years of learning and adventures from Amazon and Finland. So many things happened! All the support you gave me since my early education has brought me here and I wanted to share this entire path with you. I love you all from the deepest of my heart.

Finally, and most importantly, I would like to thank Emmi. Thank you for sharing your life with me and all the support you gave throughout the development of this thesis and on every aspect of our lives. Only the closest ones know how heavy the PhD studies can become, how big part of my life is this work, and the every-day effects it had on ourselves. You were here all the time giving me your full support. I admire and try to mirror your approach and kindness to the world, love for diversity and strength towards what you believe. You are my source of motivation to build something beautiful in life and the partner I choose to share it with. Thank you so much for all. I love you.

7 REFERENCES

Abhigna, B. S., et al. 2018. Crowdsourcing – A step towards advanced machine learning. - Procedia Computer Science 132: 632-642.

Aggarwal, C. C. 2014. Data classification: Algorithms and applications. - Taylor & Francis.

Aide, T., et al. 2017. Species richness (of insects) drives the use of acoustic space in the tropics. - Remote Sensing 9: 1096.

Aide, T. M., et al. 2013. Real-time bioacoustics monitoring and automated species identification. - PeerJ 1: e103.

Anderson, S. E., et al. 1996. Template-based automatic recognition of birdsong syllables from continuous recordings. - The Journal of the Acoustical Society of America 100: 1209-1219.

Andreassen, T., et al. 2014. Semi-automatic long-term acoustic surveying: A case study with bats. - Ecological Informatics 21: 13-24.

Armitage, D. W. and Ober, H. K. 2010. A comparison of supervised learning techniques in the classification of bat echolocation calls. - Ecological Informatics 5: 465-473.

Baker, M. C. and Logue, D. M. 2007. A comparison of three noise reduction procedures applied to bird vocal signals. - Journal of Field Ornithology 78: 240-253.

Bardeli, R., et al. 2010. Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring. - Pattern Recognition Letters 31: 1524-1534.

Barré, K., et al. 2019. Accounting for automated identification errors in acoustic surveys. - Methods in Ecology and Evolution 00: 1– 18

Bastas, S., et al. 2012. A novel feature extraction algorithm for classification of bird flight calls. - In: 2012 IEEE International Symposium on Circuits and Systems. pp. 1676-1679.

Bibby, C. J., et al. 2000. Bird census techniques. - Academic Press.

Blumstein, D. T., et al. 2011. Acoustic monitoring in terrestrial environments using microphone arrays: Applications, technological considerations and prospectus. - Journal of Applied Ecology 48: 758-767.

Briggs, F., et al. 2009. Audio classification of bird species: A statistical manifold approach. - In: 2009 Ninth IEEE International Conference on Data Mining. pp. 51-60.

Buckland, S.T., et al. 2001. Introduction to Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press, Oxford, UK.

Burivalova, Z., et al. 2019. The sound of a tropical forest. - Science 363: 28-29.

Cai, J., et al. 2007. Sensor network for the monitoring of ecosystem: Bird species recognition. - In: 2007 3rd International Conference on Intelligent Sensors, Sensor Networks and Information. pp. 293-298.

Campos-Cerqueira, M. and Aide, T. M. 2016. Improving distribution data of threatened species by combining acoustic monitoring and occupancy modelling. - Methods in Ecology and Evolution 7: 1340-1348.

Campos-Cerqueira, M., et al. 2017. Have bird distributions shifted along an elevational gradient on a tropical mountain? - Ecology and Evolution 7: 9914-9924.

Caro, T. 1998. Behavioral Ecology and Conservation Biology. - Oxford University Press.

Carstensen, J., et al. 2006. Impacts of offshore wind farm construction on harbour porpoises: acoustic monitoring of echolocation activity using porpoise detectors (T-PODs). - Marine Ecology Progress Series 321: 295-308.

Chambert, T., et al. 2015. Testing hypotheses on distribution shifts and changes in phenology of imperfectly detectable species. - Methods in Ecology and Evolution 6: 638-647.

Chambert, T., et al. 2018. A new framework for analysing automated acoustic species detection data: Occupancy estimation and optimization of recordings post-processing. - Methods in Ecology and Evolution 9: 560-570.

Chen, S. and Li, Y. 2013. Automatic recognition of bird songs using time-frequency texture. - In: 2013 5th International Conference and Computational Intelligence and Communication Networks. pp. 262-266.

Chen, Z. and Maher, R. C. 2006. Semi-automatic classification of bird vocalizations using spectral peak tracks. - The Journal of the Acoustical Society of America 120: 2974-2984.

Clemins, P. J., et al. 2005. Automatic classification and speaker identification of African elephant (Loxodonta africana) vocalizations. - The Journal of the Acoustical Society of America 117: 956-963.

Cohn-Haft, M., et al. 1997. A new look at the "species-poor" central Amazon: The avifauna north of Manaus, Brazil. - Ornithological Monographs 205-235.

Collier, T. C., et al. 2010. Acoustic localization of antbirds in a Mexican rainforest using a wireless sensor network. - The Journal of the Acoustical Society of America 128: 182-189.

de Camargo, U. M., et al. 2017. PROTAX-Sound: A probabilistic framework for automated animal sound identification. - PLOS ONE 12: e0184048.

de Oliveira, A. G., et al. 2015. Bird acoustic activity detection based on morphological filtering of the spectrogram. - Applied Acoustics 98: 34-42.

Deichmann, J. L., et al. 2018. It's time to listen: there is much to be learned from the sounds of tropical ecosystems. - Biotropica 50: 713-718.

Digby, A., et al. 2013. A practical comparison of manual and autonomous methods for acoustic monitoring. - Methods in Ecology and Evolution 4: 675-683.

Dunn, R. R. 2004. Recovery of faunal communities during tropical forest regeneration. - Conservation Biology 18: 302-309.

Dutilleux, G. and Curé, C. 2018. Automated acoustic monitoring of endangered common spadefoot toad populations reveals patterns of vocal activity. - Freshwater Biology 0:

Enari, H., et al. 2017. Feasibility assessment of active and passive acoustic monitoring of sika deer populations. - Ecological Indicators 79: 155-162.

Fagerlund, S. 2007. Bird species recognition using support vector machines. - EURASIP Journal on Advances in Signal Processing 2007: 038637.

Ferraz, G., et al. 2008. Biological monitoring in the Amazon: Recent progress and future needs. - Biotropica 40: 7-10.

Figueira, L., et al. 2015. Autonomous sound monitoring shows higher use of Amazon old growth than secondary forest by parrots. - Biological Conservation 184: 27-35.

Fox, E. J. S., et al. 2008. Call-independent individual identification in birds. - Bioacoustics 18: 51-67.

Frommolt, K.-H. and Tauchert, K.-H. 2014. Applying bioacoustic methods for long-term monitoring of a nocturnal wetland bird. - Ecological Informatics 21: 4-12.

Frommolt, K.-H. 2017. Information obtained from long-term acoustic recordings: applying bioacoustic techniques for monitoring wetland birds during breeding season. - Journal of Ornithology 158: 659-668.

Gardner, T. A., et al. 2007. Predicting the uncertain future of tropical forest species in a data vacuum. - Biotropica 39: 25-30.

Gerhardt, H. C. 1994. The evolution of vocalization in frogs and toads. - Annual Review of Ecology and Systematics 25: 293-324.

Goëau, H., et al. 2015. Lifeclef bird identification task 2015. - In: CLEF working notes 2015.

Goëau, H., et al. 2016. LifeCLEF Bird Identication Task 2016: The arrival of Deep learning. - In: CLEF working notes 2016.

Goëau, H., et al. 2017. LifeCLEF Bird Identication Task 2017. - In: CLEF working notes 2017.

Goyette, J. L., et al. 2011. Detecting tropical nocturnal birds using automated audio recordings. - Journal of Field Ornithology 82: 279-287.

Harma, A. and Somervuo, P. 2004. Classification of the harmonic structure in bird vocalization. - In: 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing. pp. V-701-4 vol.5.

Henwood, K. and Fabrick, A. 1979. A quantitative analysis of the dawn chorus: Temporal selection for communicatory optimization. - The American Naturalist 114: 260-274.

Hilje, B. and Aide, T. M. 2012. Calling activity of the common Tink frog (Diasporus Diastema) (Eleutherodactylidae) in secondary forests of the Caribbean of Costa Rica. - Tropical Conservation Science 5: 25-37.

Hill, A. P., et al. 2018. AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. - Methods in Ecology and Evolution 9: 1199-1211.

Hüppop, O., et al. 2006. Bird migration studies and potential collision risk with offshore wind turbines. - Ibis 148: 90-109.

Jingdong, C., et al. 2006. New insights into the noise reduction Wiener filter. - IEEE Transactions on Audio, Speech, and Language Processing 14: 1218-1234.

Jinnai, M., et al. 2018. Design considerations in an automatic classification system for bird vocalisations using the Two-dimensional Geometric Distance and cluster analysis.

Juang, C.-F. and Chen, T.-M. 2007. Birdsong recognition using prediction-based recurrent neural fuzzy networks. - Neurocomputing 71: 121-130.

Katahira, K., et al. 2011. Complex sequencing rules of birdsong can be explained by simple Hidden Markov Processes. - PLOS ONE 6: e24516.

Kerosky, S. M., et al. 2012. Bryde's whale seasonal range expansion and increasing presence in the Southern California Bight from 2000 to 2010. - Deep Sea Research Part I: Oceanographic Research Papers 65: 125-132.

Kirschel, A. N. G., et al. 2011. Territorial dynamics of Mexican Ant-thrushes Formicarius moniliger revealed by individual recognition of their songs. - Ibis 153: 255-268.

Kloepper, L., et al. 2017. The soundscape of bat swarms. - The Journal of the Acoustical Society of America 142: 2504-2504.

Kogan, J. A. and Margoliash, D. 1998. Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: a comparative study. - J Acoust Soc Am 103: 2185-96.

Kroodsma, D. 2015. The singing life of birds: the art and science of listening to birdsong. - Houghton Mifflin Harcourt.

Kroodsma, D. E. and Miller, E. H. 1996. Ecology and evolution of acoustic communication in birds. - Comstock Pub.

Kwan, C., et al. 2004. Bird classification algorithms: theory and experimental results. - In: 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing. pp. V-289.

Larkin, R. P., et al. 2002. Nocturnal Flight Calls of Dickcissels and Doppler Radar Echoes over South Texas in Spring (Llamadas durante vuelos nocturnos sobre Texas de Spiza americana durante la primavera y el eco en un radar Doppler). - Journal of Field Ornithology 73: 2-8.

Lasseck, M. 2015a. Towards automatic large-scale identification of birds in audio recordings. - In: J. Mothe, et al. (eds), Experimental IR Meets Multilinguality, Multimodality, and Interaction: 6th International Conference of the CLEF Association, CLEF'15, Toulouse, France, September 8-11, 2015, Proceedings. Springer International Publishing, pp. 364-375.

Lasseck, M. 2015b. Improved automatic bird identification through decision tree based feature selection and bagging. - In: L. Cappellato, et al. (eds), The Cross Language Image Retrieval Track (CLEF) Conference. CEUR.

Laurance, W. F. 2007. Have we overstated the tropical biodiversity crisis? - Trends in Ecology & Evolution 22: 65-70.

Laurance, W. F., et al. 2011. The fate of Amazonian forest fragments: A 32-year investigation. - Biological Conservation 144: 56-67.

MacKenzie, D. I., et al. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. - Ecology 84: 2200-2207.

MacSwiney G, M. C., et al. 2008. What you see is not what you get: the role of ultrasonic detectors in increasing inventory completeness in Neotropical bat assemblages. - Journal of Applied Ecology 45: 1364-1371.

Makhoul, J. and Schwartz, R. 1995. State of the art in continuous speech recognition. - Proceedings of the National Academy of Sciences of the United States of America 92: 9956-9963.

Marques, J. T., et al. 2013. Optimizing sampling design to deal with mist-net avoidance in Amazonian birds and bats. - PLOS ONE 8: e74505.

McClintock, B. T., et al. 2010. Unmodeled observation error induces bias when inferring patterns and dynamics of species occurrence via aural detections. - Ecology 91: 2446-2454.

Measey, G. J., et al. 2017. Counting chirps: acoustic monitoring of cryptic frogs. - Journal of Applied Ecology 54: 894-902.

Mennill, D. J., et al. 2012. Field test of an affordable, portable, wireless microphone array for spatial monitoring of animal ecology and behaviour. - Methods in Ecology and Evolution 3: 704-712.

Menu, S., et al. 2000. Effects of neck bands on survival of Greater Snow Geese. - The Journal of Wildlife Management 64: 544-552.

Merchant, N. D., et al. 2015. Measuring acoustic habitats. - Methods in Ecology and Evolution 6: 257-265.

Mesquita, R. C. G., et al. 2001. Alternative successional pathways in the Amazon Basin. - Journal of Ecology 89: 528-537.

Miller, D. A. W., et al. 2012. Experimental investigation of false positive errors in auditory species occurrence surveys. - Ecological Applications 22: 1665-1674.

Moffat, D., et al. 2015. An evaluation of audio feature extraction toolboxes.

Moorhouse, T. P. and Macdonald, D. W. 2005. Indirect negative impacts of radio-collaring: sex ratio variation in water voles. - Journal of Applied Ecology 42: 91-98.

Newson, S. E., et al. 2017. Potential for coupling the monitoring of bush-crickets with established large-scale acoustic monitoring of bats. - Methods in Ecology and Evolution 8: 1051-1062.

Oppel, S., et al. 2014. Estimating population size of a nocturnal burrow-nesting seabird using acoustic monitoring and habitat mapping. - Nature Conservation 7:

Ovaskainen, O., et al. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. - Ecology Letters 20: 561-576.

Ovaskainen, O., et al. 2018. Animal Sound Identifier (ASI): software for automated identification of vocal animals. - Ecology Letters 21: 1244-1254.

Payne, K. B., et al. 2003. Elephant calling patterns as indicators of group size and composition: the basis for an acoustic monitoring system. - African Journal of Ecology 41: 99-107.

Pérez-Granados, C., et al. 2019. Vocal activity rate index: a useful method to infer terrestrial bird abundance with acoustic monitoring. - Ibis 0:

Petrusková, T., et al. 2016. Repertoire-based individual acoustic monitoring of a migratory passerine bird with complex song as an efficient tool for tracking territorial dynamics and annual return rates. - Methods in Ecology and Evolution 7: 274-284.

Porter, J. H., et al. 2009. New eyes on the world: Advanced sensors for Ecology. - BioScience 59: 385-397.

Potamitis, I. 2014. Automatic classification of a taxon-rich community recorded in the wild. - PLoS One 9: e96936.

Potamitis, I. 2015. Unsupervised dictionary extraction of bird vocalisations and new tools on assessing and visualising bird activity. - Ecological Informatics 26, Part 3: 6-17.

Prasad, A. M., et al. 2006. Newer classification and regression tree techniques: Bagging and random forests for Ecological prediction. - Ecosystems 9: 181-199.

Priyadarshani, N., et al. 2018. Automated birdsong recognition in complex acoustic environments: a review. - Journal of Avian Biology 49: jav-01447.

Quintero, I. and Roslin, T. 2005. Rapid recovery of dung beetle communities following habitat fragmentation in central Amazonia. - Ecology 86: 3303-3311.

Rajan, S. C., et al. 2018. Rapid assessment of biodiversity using acoustic indices. - Biodiversity and Conservation

Ribeiro, J. W., et al. 2017. Passive acoustic monitoring as a complementary strategy to assess biodiversity in the Brazilian Amazonia. - Biodiversity and Conservation 26: 2999-3002.

Rundel, P. W., et al. 2009. Environmental sensor networks in ecological research. - New Phytol 182: 589-607.

Sandsten, M., et al. 2016. Robust feature representation for classification of bird song syllables. - EURASIP Journal on Advances in Signal Processing 2016: 68.

Schmutz, J. A. and Morse, J. A. 2000. Effects of neck collars and radio transmitters on survival and reproduction of Emperor Geese. - The Journal of Wildlife Management 64: 231-237.

Scott Brandes, T. 2008. Automated sound recording and analysis techniques for bird surveys and conservation. - Bird Conservation International 18: S163-S173.

Selin, A., et al. 2006. Wavelets in recognition of bird sounds. - EURASIP Journal on Advances in Signal Processing 2007: 051806.

Servick, K. 2014. Eavesdropping on ecosystems. - Science 343: 834-837.

Somervuo, P., et al. 2006. Parametric representations of bird sounds for automatic species recognition. - IEEE Transactions on Audio, Speech, and Language Processing 14: 2252-2263.

Sorzano, C., et al. 2014. A survey of dimensionality reduction techniques.

Sprengel, E., et al. 2016. Audio based bird species identification using deep learning techniques. - In: Conference and Labs of the Evaluation Forum (CLEF) pp. 547-559-547-559.

Stouffer, P. C. and Bierregaard, R. O. 1995. Use of Amazonian forest fragments by understory insectivorous birds. - Ecology 76: 2429-2445.

Stouffer, P. C., et al. 2006. Long-term landscape change and bird abundance in Amazonian rainforest fragments. - Conserv Biol 20: 1212-23.

Stowell, D. and Plumbley, M. D. 2014. Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. - PeerJ 2: e488.

Stowell, D., et al. 2016. Bird detection in audio: a survey and a challenge. - In: 2016 IEEE International Workshop on Machine Learning for Signal Processing.

Stowell, D., et al. 2018. Automatic acoustic detection of birds through deep learning: The first Bird Audio Detection challenge. - Methods in Ecology and Evolution 0:

Suter, S. M., et al. 2017. Non-invasive acoustic detection of wolves. - Bioacoustics 26: 237-248.

Tang, J., et al. 2014. Feature selection for classification: A review. - In: C. C. Aggarwal (ed) Data Classification: Algorithms and Applications. CRC Press, p. 37.

Terry, A. M. R., et al. 2005. The role of vocal individuality in conservation. - Frontiers in Zoology 2: 10.

Towsey, M., et al. 2012. A toolbox for animal call recognition. - Bioacoustics 21: 107-125.

Trifa, V. M., et al. 2008. Automated species recognition of antbirds in a Mexican rainforest using hidden Markov models. - The Journal of the Acoustical Society of America 123: 2424-2431.

Ulloa, J. S., et al. 2016. Screening large audio datasets to determine the time and space distribution of Screaming Piha birds in a tropical forest. - Ecological Informatics 31: 91-99.

Whytock, R. C. and Christie, J. 2017. Solo: an open source, customizable and inexpensive audio recorder for bioacoustic research. - Methods in Ecology and Evolution 8: 308-312.

Wrege, P. H., et al. 2010. Use of acoustic tools to reveal otherwise cryptic responses of forest elephants to oil exploration. - Conservation Biology 24: 1578-1585.

Wrege, P. H., et al. 2017. Acoustic monitoring for conservation in tropical forests: examples from forest elephants. - Methods in Ecology and Evolution.

Young, A. M. 1981. Temporal selection for communicatory optimization: The dawn-dusk chorus as an adaptation in tropical cicadas. - The American Naturalist 117: 826-829.

Zbancioc, M. and Costin, M. 2003. Using neural networks and LPCC to improve speech recognition. - In: Signals, Circuits and Systems, 2003. SCS 2003. International Symposium on. pp. 445-448 vol.2