Impacts of coffee fragmented landscapes on biodiversity and microclimate with emerging monitoring technologies.

Elise Damstra

A thesis submitted for the degree of Doctor of Philosophy

Department of Life Sciences,

Imperial College London

2022

Declaration of Originality

I declare that all work presented in this thesis is my own, unless otherwise stated.

Specific acknowledgements:

Chapter 2:

This work was done in collaboration with Cristina Banks-Leite (Imperial College London) as part of an opportunity from Oxford Bibliographies who commissioned a summary of the most important literature on the Ecology of the Atlantic Forest. I wrote most of the review with input from C.B.L.; C.B.L. led the section on conservation policy.

Chapter 3:

Andrea Larissa Boesing selected the study sites in Brazil and received the appropriate licenses. Field work was carried out by A.L.B. and myself. The study was designed with discussions from C.B.L. and Kate E. Jones I led the quantitative approach.

Chapter 4:

Oisin Mac Aodha designed and trained the CNN model and retrained it for the Brazilian dataset. The UK CNN model was trained with data from multiple sources including the Bat Conservation Trust and Martyn Cooke. Annotations for the UK training data were annotated by myself, O.M.A. and Liz Walsh. The Brazilian training data was put together and labelled by me. Identification of Brazilian bat calls was done with the assistance of Adriana Arias-Aguilar

Chapter 5:

The data for this chapter was provided by Climate Edge, a start-up founded by Paul Baranowski and Peter Baker. Analysis ideas and code snippets were provided by Rob Ewers (Imperial College London).

Copyright Declaration

Acknowledgements

A year ago I sent in a letter resigning from this PhD to both my supervisors. I had reached breaking point due to poor mental health, stress and the global pandemic. Instead of accepting my resignation, Cristina and Kate recommended I take an extended period of leave, which I did. A month into this leave I hit rock bottom: my (ex)husband surprised me with a divorce, I was forced to move home to my parents and I felt utterly lost. As such, I am the last person who thought I'd be handing this in. How things have changed. I have many people to thank for their support throughout this journey, first and foremost my supervisory team of Cristina Bank-Leite and Kate Jones who got me through some very dark times with kindness but also knowing when to push me for more. Thank you to the wonderful members of the CALE lab: Flavia, Jenna, Ben, Paula, Jack and Michelle who made the office (until the pandemic struck) a fun, collaborative working environment. I would also like to thank other Silwood Park PhD students and staff, special mentions go to Paul Beasley and Martin Selby for being so patient when I needed remote help, and to James Rosindell and Will Pearse for their support. Finally, I'd like to thank my family: to Richard, Helen, Marie-José, Cynthia, Tina, Kim, Cecile, and Theo for always being there for me.

Thesis Abstract

Habitat fragmentation and loss are causing biodiversity declines across the globe. As biodiversity is unevenly distributed, with many hotspots located in the tropics, conserving and protecting these areas is important to preserve as many species as possible. Chapter 2 presents an overview of the Ecology of the Atlantic Forest, a highly fragmented biodiversity hotspot. A major driver of habitat fragmentation is agriculture, and in the tropics coffee is major cash crop. Developing methods to monitor biodiversity effectively without labour intensive surveys can help us understand how communities are using fragmented landscapes and better inform management practices that promote biodiversity. Acoustic monitoring offers a promising set of tools to remotely monitor biodiversity. Developments in machine learning offer automatic species detection and classification in certain taxa. Chapters 3 and 4 use acoustic monitoring surveys conducted on fragmented landscapes in the Atlantic Forest to quantify bird and bat communities in forest and coffee matrix, respectively. Chapter 3 shows that acoustic composition can reflect local avian communities. Chapter 4 applies a convolutional neural network (CNN) optimised on UK bat calls to a Brazilian bat dataset to estimate bat diversity and show how bats preferentially use coffee habitats. In addition to monitoring biodiversity, monitoring microclimate forms a key part of climate smart agriculture for climate change mitigation. Coffee agriculture is limited to the tropics, overlapping with biodiverse regions, but is threatened by climate change. This presents a challenge to countries strongly reliant on coffee exports such as Brazil and Nicaragua. Chapter 5 uses data from microclimate weather stations in Nicaragua to demonstrate that sun-coffee management is vulnerable to supraoptimal temperature exposure regardless of local forest cover or elevation.

Contents

List of Figures and Tables

Chapter 1 – Introduction

1.1: Biodiversity crisis

Biodiversity is in crisis, currently in the midst of a 6th mass extinction event (Ceballos *et al.*, 2015), with not only loss of species but loss of biomass across all taxa (Goulson, 2019). Ecosystems operate within fragile equilibria that have evolved over millennia which though resilient to some change, are not able to respond to the scale and pace of threats they are currently faced with (Hoegh-Guldberg, 2012). Amphibians, mammals and birds are experiencing the highest rates of extinction (Johnson *et al.*, 2017). These dramatic declines are caused by anthropogenic influence through direct drivers such as land-use change leading to habitat degradation, pollution, invasive species and pathogens, overexploitation and climate change (Mace, 2011). Biodiversity provides essential services to humans such as pollination, pest control, nutrient cycling, seed dispersal, etc., which they are less able to provide when their previously stable ecosystems have altered functionality and lowered stability (Isbell *et al.*, 2017). Reversing current declines is a conservation priority that has been ratified by the United Nations in their Sustainable Development Goals (SDGs) (United Nations, 2015), which requires a detailed understanding of the current state of biodiversity and its vulnerability at a suitable scale.

Land use change is the main driver of local species loss globally, with a ~76.5% reduction in within-sample richness predicted by 2095 (Newbold *et al.*, 2015), much of which is driven by conversion to agriculture (Hurtt *et al.*, 2011). This is particularly pronounced in the tropics (Murphy and Romanuk, 2014), where forests are being lost at a rate of over 200 000 km^2 per year (Hansen *et al.*, 2013). Between 1980 and 2000, 55% of new agricultural land in the tropics came from intact forests and an additional 28% was derived from disturbed forests (Gibbs *et al.*, 2010), which is of particular concern due to the exceptional natural value of intact forests (Watson *et al.*, 2018). Much of this is driven by demand for crops that are physiologically limited to the tropics e.g. oil palm, rubber and coffee (Verbist, Putra and Budidarsono, 2005; Li *et al.*, 2007; Ambinakudige and Choi, 2009; Vijay *et al.*, 2016). However, agricultural intensification drives economic development, as it lowers food insecurity, reducing poverty rates (Galford *et al.*, 2013), but its benefits are unevenly distributed between and within countries (World Resources Institute, 2005). Agriculture is also a substantial contributor to greenhouse gas emissions, and thus climate change (Tubiello *et al.*, 2015). Though technological advances may improve crop yields to sufficiently meet demand, this may not be sufficient to discourage further tropical forest loss (Carrasco *et al.*, 2014).

Climate change is estimated to cause global temperature to rise by up to 2°C by 2050 (Yerlikaya, Ömezli and Aydoğan, 2020), presenting a major threat to both biodiversity and agriculture. Extensive and rapid shifts in suitable ranges have been predicted for cultivated crops (Arora, 2019) and biodiversity (Colwell *et al.*, 2008), leaving many species at risk of extinction. Predictions in range shifts of tropical crops imply that new regions of suitability often fall within areas that are currently forested (Magrach and Ghazoul, 2015). These range shifts are under-considered in most conservation policies (Pecl *et al.*, 2017). To ensure future food security and biodiversity, there has been much focus on nature-based solutions, which purportedly mitigate and adapt against climate change while simultaneously benefitting biodiversity (Naumann *et al.*, 2011), although there is an overreliance on afforestation strategies in policy (Seddon *et al.*, 2019). The interactions and feedbacks between climate change, agriculture and biodiversity remain poorly understood and are a key research priority (Ortiz *et al.*, 2021) if the effects of climate change are to be properly mitigated.

To address the problem of biodiversity decline and demonstrate efficacy of various conservation solutions, monitoring is essential. This is not limited to biodiversity, but also specific local variables that may influence the local community including microclimate, land use, habitat quality and anthropogenic activity. The scale of this monitoring (spatially and temporally) is critical, as some tropical biodiversity hotspots are topographically complex, creating fine-scale microclimatic variation (Trew and Maclean, 2021), and current widely used environmental data is at a coarser scale, which is a source of inaccuracy (Storlie *et al.*, 2014). To fill this gap, scalable monitoring solutions are required that can be implemented at local scales, ideally with the collaboration of citizen science and farmers.

1.2: Monitoring

Terrestrial biodiversity it not uniformly distributed across the globe, there are regions that are home to disproportionately high numbers of species, but have lost large extents of their original extent, known as "biodiversity hotspots" (Myers *et al.*, 2000; Kobayashi, Okada and Mori, 2019). This is a useful concept to identify regions for conservation and research priority, as 2.4% of land surface contains over half of all plant species and 43% of bird, mammal and herpetofauna species (Conservation International, 2021). Of the 36 formally identified hotspots, 24 are located within or on the tropics (CEPF, 2021). Hotspots have suffered high levels of biotic decline (Newbold *et al.*, 2016) to beyond the limit of normal ecological functioning. Despite being an urgent conservation priority, biodiversity data in the tropics is relatively lower than regions nearer the poles (Collen *et al.*, 2008) and even when areas are designated for protection, they are poorly monitored (Oliveira *et al.*, 2017), which gives rise to questions about the efficacy of such areas. Tropical regions therefore need scalable monitoring solutions to collect useful information about species and communities in order to assess species responses to change.

Technological advances have resulted in rapid advances in monitoring with regards to methodologies and the volume of data that can be collected. Techniques such as eDNA analysis, acoustic monitoring, camera traps and remote sensing are providing ways to identify communities of species present without constant presence in the field (Bohmann *et al.*, 2014; Rocchini *et al.*, 2018; Gibb *et al.*, 2019), though they are still mostly at a point where these methods should be paired with observer-based methods (Stephenson, 2020). These methods differ in the species they can optimally monitor, which should be considered within the context of designing monitoring protocols. These advances in technology allow for more citizen science participation in data collection which could result in accelerated scaling of monitoring (Chandler *et al.*, 2017; Pocock *et al.*, 2018).

1.2.1: Acoustic monitoring

Acoustic monitoring is a rapidly developing field using sound to detect patterns and trends in biodiversity as well as other sounds of interest (Browning *et al.*, 2017), with the number of studies implementing these tools increasing fifteenfold since 1992 (Sugai *et al.*, 2019). Recent technological advances have lowered the cost of recording equipment, e.g. Audiomoths and Raspberri Pi are affordable recorders that can be customised to different recording frequencies and schedules (Upton and Halfacree, 2014; Hill *et al.*, 2019). Data storage capacity is also inexpensive and no longer a limiting factor in many cases. However, the ease of data collection has now created a bottleneck in analysis methodologies.

1.2.2: Analysis of acoustic data

Once acoustic data has been collected, many analytical approaches are available that vary in complexity and robustness depending on the question being answered and factors such as study region and taxa of interest. Acoustic data are complex, so summarising sound features with acoustic indices that act as proxies for biodiversity metrics has proved popular (Gibb *et al.*, 2019). Soundscape analyses are used to characterise the acoustic characteristics of landscapes, including all present sound (Deichmann *et al.*, 2017; Furumo and Mitchell Aide, 2019; Sethi *et al.*, 2020). More sophisticated tools are also emerging using machine learning to automate event detection and classification to species level, particularly for bats and birds in North America and Europe (Szewczak, 2015; Mac Aodha *et al.*, 2018; Stowell *et al.*, 2018). Machine learning tools require a high level of input data to be robust, which currently poses a bottleneck, particularly in high diversity regions of the tropics.

Acoustic indices are some of the more commonly used tools for analysing acoustic data, likely due to their ease of access and implementation. Over 60 acoustic indices have been developed (Bradfer-Lawrence *et al.*, 2019), many of which can be easily calculated using accessible programs. Some indices, such as the Acoustic Complexity Index (ACI) have been shown to positively correlate to peaks in avian activity and avian species richness (Gage *et al.*, 2017; Hilje, Stack and Sánchez-Azofeifa, 2017). However, other studies found no correlations to species richness (Buxton *et al.*, 2018). Lack of agreement on how indices relate to biodiversity may be linked to a failure to consider the effect of local vegetation, (Darras *et al.*, 2016) or the urban environment (Fairbrass *et al.*, 2017). As a result of this mixed transferability between ecosystems, it is recommended to combine multiple indices as well as complementing them with other methods (such as soundscape analysis) when conducting biodiversity surveys (Fuller *et al.*, 2015; Bradfer-Lawrence *et al.*, 2019).

Machine learning methods have accelerated the potential for automated acoustic monitoring. Deep learning is a subset of machine learning and is still a relatively new field for bioacoustics, but significant work such as the publication of the AudioSet dataset and the VGGish neural network architecture (Hershey *et al.*, 2017) have made deep learning more accessible to biodiversity researchers (Stowell, 2022) and has successfully been used for anomaly detection (Sethi *et al.*, 2019). However, to robustly train convolutional neural networks (CNNs) to identify and classify sounds of interest (Mac Aodha *et al.*, 2018) these methods require large training datasets that have been expertly verified. This is a major hurdle to automated species detection, particularly in biodiverse regions. However, if a CNN is optimised to detect a particular taxa in one region, some aspects may be generalisable and can thus be applied to the same taxa in new regions without retraining.

Bats are a highly biodiverse taxonomic group, with 1470 described species, which represents approximately one fifth of all mammal species (Solari and Baker, 2007; Lei and Dong, 2016). Around 1000 of these species use echolocation to navigate and hunt for prey (Boonman *et al.*, 2013), most of which are ultrasonic, making them a good target for CNNs as few other sounds overlap at these frequencies. Call shape, frequency range and interval between pulses are strongly linked to body size, diet and skull morphology, so many species can be distinguished by their call (Bogdanowicz, Fenton and Daleszczyk, 1999; Arias-Aguilar *et al.*, 2018), though intraspecific variation can still be a confounding factor (Russo, Ancillotto and

Jones, 2017). As bats are present on all continents except Antarctica, convergent evolution has resulted in cases of different species occupying similar niches developing similar calls (Wowk *et al.*, 2001). The development of regional bat detection and classification tools is well underway, but as CNNs require extensive input data, better knowledge of tropical bat calls is necessary to develop robust classification tools.

Monitoring biodiversity and its abiotic surroundings is now more accessible than ever, with instruments that can be left in place to record data for days, weeks or indefinitely without the need for human interference. It must also be stated that despite these leaps in remote sensing tools, many species will be missed if using a single detection method e.g. acoustic monitoring will not detect quiet or non-vocal species. Therefore, to better capture an entire community, other detection methods should be used simultaneously. What these are will depend on the target species. With the rapid development of data collection tools, a current priority must be form standardized monitoring protocols and best practices for emerging tools. Once such protocols are in place, the potential of monitoring schemes to form part of smart future solutions is extensive.

1.2.3 Microclimatic monitoring

The scale of climatic data used to inform predictions of species distributions under climate change tends to be relatively coarse, with many studies relying on the popular WorldClim datasets (Hijmans *et al.*, 2005; Fick and Hijmans, 2017), with a combined number of citations over 23 400. This type of data is constructed by interpolation of across regions between weather stations, and the distance between these tends to be higher in the tropics. Large discrepancies have been uncovered between this data and locally collected data in highly heterogenous regions, such as tropical mountainous areas, as microclimatic effects in agricultural landscapes were not detected in the interpolated data (Faye *et al.*, 2014). Other remotely sensed datasets have performed better, but the scale of these data remains coarse (Deblauwe *et al.*, 2016). Considering this, predictions about how future climate change will affect the future of crops cultivated at high altitudes in topographically complex regions may be inaccurate and require re-examination with the inclusion of microclimatic data.

1.3 Coffee

The earliest evidence of coffee being consumed as a beverage comes from the 14th or 15th century in the Middle East (Haarer, 1956; Wrigley, 1988), and it is now the most consumed beverages in the world after water (Butt and Sultan, 2011). Demand has recently outpaced production, causing a reduction of stockpiles (Vegro and de Almeida, 2020). Its global popularity is likely to continue to grow, so ensuring a sustainable supply is not only important to consumers, but also the at least 125 million people who rely on it for their livelihood (Osorio, 2002). Many producers are smallholders, who are particularly vulnerable to market volatility and the threat of climate change (Verburg *et al.*, 2019). Due to the way coffee grows, automation is not possible to the extent it is in other agricultural commodities like wheat or maize, resulting in a high demand for labour and thus the number of people who rely on coffee production is high (Upendranadh and Subbaiah, 2016). Two varieties of coffee form the majority of the market: *Coffea arabica* and *Coffea canephora* var. *robusta*. The former is the more popular variety deemed better tasting and of higher quality, and is the focus of this work, so when referring to coffee, I am referring to *C. arabica*. To maintain production in the long-term, sustainable management practices need to be implemented and supported.

Coffee is native to Africa, originating from the mountains of what are now Ethiopia, Kenya and Sudan (Waller, Bigger and Hillocks, 2007). Coffee is adapted to grow in the tropics, and is cultivated in Latin and South America, Africa and Asia, with Brazil leading global production, producing ~30% of the export market (Mussatto *et al.*, 2011). Coffee is a shade adapted species, evolved to live in forest undergrowth (Waller, Bigger and Hillocks, 2007), and traditionally farmed in forest. It has a narrow optimal climatic niche of 18-21°C mean annual temperature, though some varieties can tolerate higher means up to 24°C (DaMatta, 2004), though temperatures above this result in heat stress and mean annual temperature below 17°C result in depressed growth (DaMatta and Ramalho, 2006). Coffee also requires a minimum of 1000 mm annual rainfall (though some varieties can tolerate as little as 750 mm), the seasonality of which is important (Fain *et al.*, 2017). This limits suitable growing regions to high altitudes of the tropics. However, with climate change, the minimum suitable altitude is increasing and the area considered suitable for coffee growth is predicted to decrease considerably: in Africa by 2080 up to 100% loss has been predicted (Davis *et al.*, 2012), 30% loss by 2050 in Mesoamerica (Ovalle-Rivera *et al.*, 2015a), and globally 50% of suitable area could be lost (Bunn, Läderach, Rivera, *et al.*, 2015). Current areas of natural forest at high altitudes will likely be lost to create more space for cultivation (Magrach and Ghazoul, 2015).

Coffee has traditionally been cultivated as a shade crop, grown in a polyculture with taller trees grown for fruit production and timber (Campanha *et al.*, 2004). This type of agriculture is similar to forest, providing habitat for higher levels of biodiversity (Philpott *et al.*, 2008).

However, demand and market forces have driven intensification, particularly in Latin America (Moguel and Toledo, 1999), where it is also thought that more exposure to light and a higher density of plants will help prevent fungal growth (Muschler, 1997). However, a growing concern over climate change and biodiversity loss as shaded management reduces water loss and lowers ambient temperatures (Lin, 2010; Rigal *et al.*, 2020). In addition, international certification schemes from organizations such as Fairtrade, Rainforest Alliance and Bird Friendly Coffee offering higher prices have encouraged better practices, which have shown to aid conservation (Hardt *et al.*, 2015), though their impact on farmer income has been debatable (Valkila, 2009; Beuchelt and Zeller, 2011). Sun coffee monocultures are particularly vulnerable to climate change, and as this management is most dominant globally, representing 41% of coffee cultivation area (traditional shade is only 24% (Jha *et al.*, 2014)), shifts in management must be enacted quickly to minimize losses.

Climate change does not simply present a direct threat to the physiology of coffee plants, it is also predicted to increase the range and intensity of common pests and diseases of coffee such as the coffee borer beetle *Hypothenemus hampei* and rust caused by the fungus *Hemileia vastatrix* (Avelino *et al.*, 2015; Jaramillo *et al.*, 2009). However, re-analysis with models constructed with high spatio-temporal resolution reject the finding that rust will increase with climate change (Bebber, Castillo and Gurr, 2016). *H. hampei* are predicted to double their number of generations per year in some coffee producing regions form 5 to 10, posing a serious threat to quality and quantity of harvests (Jaramillo *et al.*, 2011). Borer beetles are hard to control with pesticides (Silva and Beauvais, 2010), but high levels of biodiversity have been shown to reduce outbreak intensity. Shade coffee farms have been found to have higher rates of pest control by birds and ants than sun coffee (Milligan *et al.*, 2016). However, it has also been found that the structure of the surrounding habitat may have a stronger effect on biodiversity than the management type itself (Boesing, Nichols and Metzger, 2017). Bats have also been found to reduce arthropods by up to 84% on coffee plantations (Williams-Guillén, Perfecto and Vandermeer, 2008). Therefore, providing more habitat for biodiversity is mutually beneficial for the crop and enables the provision of ecosystem services.

Though the benefits of biodiversity and optimizing management to minimize climatic extremes caused by climate change are broadly understood, they tend to be poorly monitored. Monitoring species diversity on coffee farms consists of labour-intensive mist-netting and point counts, requiring good knowledge on local species. The inclusion of microclimatic data into broader models has demonstrated the importance of finer scale climate data (Storlie *et al.*, 2014) and could explain local variation in coffee yields, pest risk and biodiversity. As coffee is grown in topographically complex regions, understanding how management, microclimate and biodiversity interact will be important at a local scale to best inform the most sustainable practices in the future. This type of work is starting to be done (Karungi *et al.*, 2018), but to scale the pace of such work, more rapid monitoring protocols needs to be established.

1.4 Study region introduction

1.4.1 The Atlantic Forest domain

Brazil is the world's biggest producer and exporter of coffee, and Minas Gerais is the most productive coffee-producing region, growing primarily *Coffea arabica* (Volsi *et al.*, 2019). Though traditionally grown as a shade crop, economic pressure has driven intensification, leading to large, sun-exposed monoculture *C. arabica* plantations. Climate change since the 1970s in Minas Gerais has already caused a mean annual temperature increase of ~1°C and a

20% reduction in coffee yield (Koh *et al.*, 2020). Interspersed between these plantations are fragments of forest that range in size and level of environmental degradation. These remain largely due to the enactment of the Brazilian Forest Code, which has stipulations such as "areas on hilltops, mountains and ridges with a minimum height of 100 metres…must maintain native habitat" (Machado and Anderson, 2016).The presence of forest on farms has been shown to increase pest control services by birds and bats and thus increase coffee yield (Williams-Guillén, Perfecto and Vandermeer, 2008; Karp *et al.*, 2013; Jordani, Hasui and Silva, 2015). Forest fragments within coffee landscapes are an important area of study as they represent some of the only remaining habitat for biodiversity in the region (Rocha, Passamani and Louzada, 2011).

The Atlantic forest of southern Brazil is recognised as a biodiversity hotspot (Myers *et al.*, 2000), but is highly threatened, with some arguing that it is the hottest hotspot (Laurance, 2009). The remaining natural habitat exists as an archipelago of forest fragments surrounded by urban area or agriculture (Joly, Metzger and Tabarelli, 2014). This combination of high endemic biodiversity and high fragmentation signifies that it is a strategically important area for conservation efforts to protect remaining species, as the changes already made have driven biotic homogenisation (Lôbo *et al.*, 2011). Despite legal conservation measures including the Forest Code mandating zero deforestation of primary forest since 2006, problems with enforcement have led to non-compliance (Sparovek *et al.*, 2012) and the loss of over 180 000 hectares between 2008-2018, though rates of loss are decreasing (SOS Mata Atlantica and INPE, 2019). This region of Brazil is also its most populated: the states of São Paulo, Rio de Janeiro and Minas Gerais have a combined population over 85 million (IBGE, 2020), mostly living in urban areas. Conservation measures thus have to protect as many species as possible while not impeding economic development.

The Atlantic Forest is one of the most studied regions with regards to species responses to habitat fragmentation, representing 63% of all fragmentation based studies in Brazil, with most studies focusing on the impact of fragment size (Teixido *et al.*, 2020). The remaining native vegetation is highly fragmented, mostly consisting of fragments smaller than 50 ha in size (Ribeiro *et al.*, 2011), and an estimated 28% of its initial extent remaining, (Rezende *et al.*, 2018), much of which is on private land (Sparovek *et al.*, 2012). There has been debate around the term "fragmentation" and how the effects of habitat loss versus fragmentation per se (defined as different spatial configurations at a given level of habitat loss) may interact (Didham, Kapos and Ewers, 2012), as landscape scale species responses to fragmentation per se have been found to be non-significant (Fahrig *et al.*, 2019), though this has been strongly contested (Fletcher *et al.*, 2018). A tipping point of ~30% forest cover has been found in several studies, below which richness and community intactness decline across multiple taxa (Banks-Leite *et al.*, 2014; Lima and Mariano-Neto, 2014). Despite a rich literature on species responses to fragmentation (e.g. Banks-Leite, Ewers and Metzger, 2012; Uezu and Metzger, 2016; Delciellos *et al.*, 2018), knowledge gaps remain in behavioural responses and community composition responses (Teixido *et al.*, 2020).

Birds are one of the more studied taxonomic groups within the Atlantic Forest, as they are a conservation target. Out of at least 832 identified species (Hasui *et al.*, 2018), 223 have been identified as endemic to the domain (Vale *et al.*, 2018). Birds have been the main indicator used to examine the effects of fragmentation in the Atlantic forest and are the primary focus of many studies (Piratelli *et al.*, 2008; Uezu and Metzger, 2016). There is interest in how different functional groups of birds occupy and use fragmented landscape and how this relates to their vulnerability to extinction (Pardini *et al.*, 2009). As many Atlantic Forest birds are passerines, they are a good potential target for acoustic monitoring. Acoustic methods may detect differences in communities as well as responses to fragmentation and land use change.

1.4.2 Field work region introduction

Field work was carried out for Chapters 3 and 4 between January and March 2019 in the largest coffee producing area of Brazil (Volsi *et al.*, 2019), in a region on the border between the states of Minas Gerais and São Paulo. This region is within the Atlantic Forest domain and is regulated by the Brazilian Forest Code, which requires that landowners conserve a minimum of 20% of native forest (Brock *et al.*, 2021). Thus, the landscape is mostly heterogenous mosaics of coffee agriculture with forest patches. The type of coffee

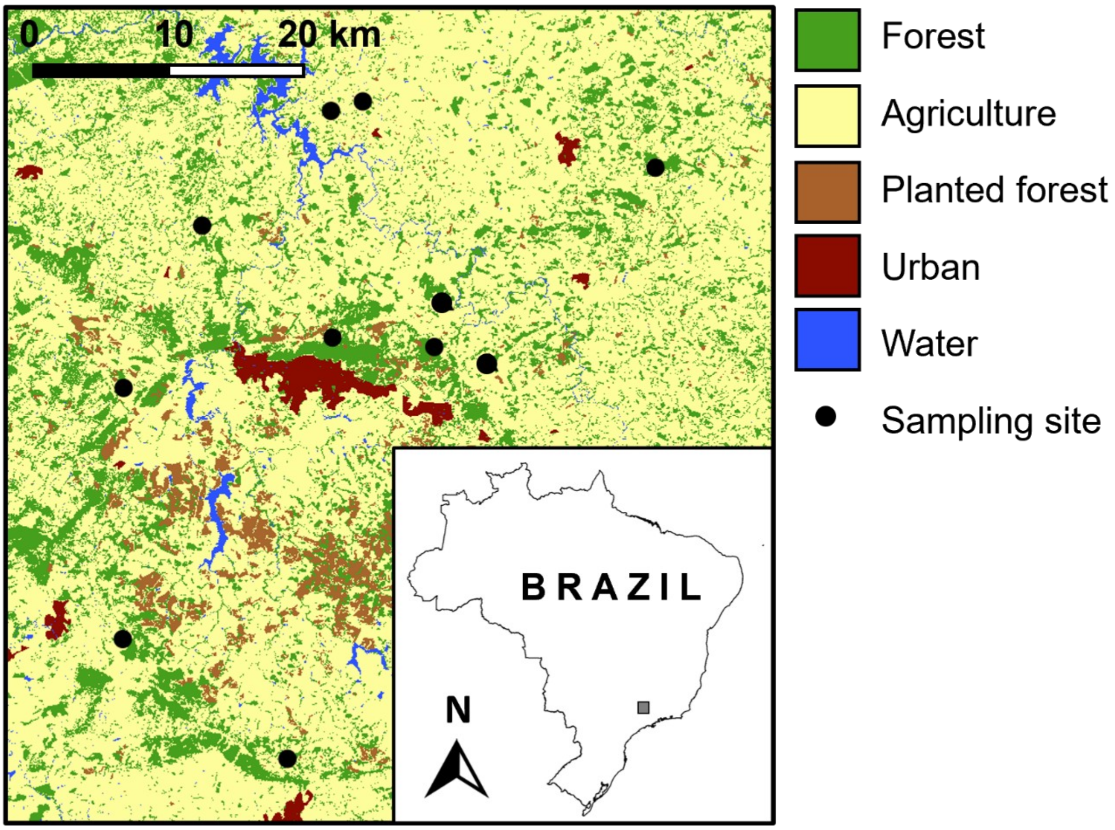


Figure 1.1: *Map of sampling locations. Inset map of Brazil shows grey square where of the location. Colours denote land use, key on the right. The main urban region in the centre is the town of Poços de Caldas*

management in this region is intensive, unshaded monoculture with varying plant densities and most farms apply herbicides, pesticides and fertilizers (Moguel and Toledo, 1999; Librán-Embid, De Coster and Metzger, 2017). This region has a subtropical highland climate with rainy summers and dry winters (Pompeu *et al.*, 2009).

The field work and sampling schedule was adapted to be performed alongside the existing field work and research permissions of a post-doctoral researcher from the University of São Paulo (Andrea Larissa Boesing). Her aim was to use mist-nets and a telemetry set-up to compare the activity of habitat specialist versus habitat generalist birds at coffee farms and adjacent forest. This field work was initially going to be performed twice, once during a season without pest outbreaks, and then later in the year (late April-June) when pest outbreaks were expected. Mist-nets were set up along the forest edge and in the coffee (more details in Chapter 3). Three days of mist-netting were performed at each sampling site, during which sufficient birds of interest were captured and tagged with telemetry devices. The sites were then unattended for 7-10 days, after which the telemetry and acoustic equipment were collected. This was repeated at 9 coffee farms across a gradient of local forest cover, as we were interested in the effect of local forest cover on the bird community and bird activity.

To record sound at the sample sites, Audiomoths (Hill *et al.*, 2019) were deployed. As both bats and birds were of interest, two recorders were deployed at each sampling point: one at a sampling rate of 48 kHz, which recorded from 04:00-22:00 and one at a sampling rate of 384 kHz, which recorded from 22:00-04:00. This allowed the memory cards on the Audiomoths to fill at similar rates. The Audiomoths were placed in plastic zip-lock bags at 1.5 m above the ground, attached to branches or coffee plants with cable ties. Most recorders performed well, though some did suffer minor water damage as field work was conducted during the rainy season. Unfortunately, due to equipment damage and logistical problems, the acoustic recording only took place during the first field work season, without the pest outbreaks.

1.4.3 Nicaraguan field work

Nicaragua is a small central American country, which relies heavily on coffee exports for income. Over 60% of producers are smallholders, farming on less than two hectares (ICO, 2020), making them highly vulnerable to small fluctuations in production. Central America is particularly vulnerable to the impacts of climate change, with elevated temperatures already causing problems with yield due to hydric stress in 2019 in Honduras (IHCAFE, 2020). A consequence of increasing temperatures is increased risk of disease outbreaks such as leaf rust, which previously did not occur above elevations of 1000 m (Bacon *et al.*, 2017). As Nicaragua has such high vulnerability to climate change as well as a high proportion of smallholders, providing microclimatic monitoring and making localised recommendations based on this monitoring may provide more income security for these farmers.

Coffee production is most intense in the high-altitude regions, such as Jinotega, which is where the field work was conducted for Chapter 5. Field work was conducted by Climate-Edge, who selected sites and deployed weather stations at selected coffee farms to measure microclimate.

1.5 Thesis aims and objectives

Chapter 2 is published in Oxford Bibliographies and provides an overview of the important literature on the ecology of the Atlantic Forest. This is aimed to help undergraduates navigate the literature on this region and direct them to useful data sources. It discusses the history of the domain, how various taxonomic groups are affected by the current state of fragmentation, and current conservation policy. The structure was limited by constraints such as maximum references per section. In the original publication, each individual piece of literature was summarised in a sentence or two beneath the section from which it was taken. Permission to reproduce and the publication in its original format is attached in Appendix A.

Chapter 3 investigates how acoustic composition changes across a forest cover gradient and whether this is driven by avian community composition.

Chapter 4 examines how far a UK trained bat classification algorithm can be applied to acoustic data recorded in Brazil, a region without a formal call library.

Chapter 5 investigates the relative strength of microclimates in sun-coffee farms in Nicaragua using data from portable weather stations.

Chapter 6 discusses the implications of findings from the chapters within the thesis and what they mean for future monitoring schemes and research.

1.6 Thesis collaborators and contributions

Chapter 2: The Ecology of the Atlantic Forest

This work was done in collaboration with Cristina Banks-Leite (Imperial College London) as part of an opportunity from Oxford Bibliographies who commissioned a summary of the most important literature on the Ecology of the Atlantic Forest. Oxford Bibliographies provide online research guides aimed at undergraduates to provide a thorough summary of a topic. Each article features an annotated bibliography, where each reference is summarised in terms of its highlights and overall contribution to the field. I wrote most of the review with input from C.B.L.; C.B.L. led the section on conservation policy.

Publication reference:

Damstra-Oddy, E. & Banks-Leite, C. (2021) Ecology of the Atlantic Forest. In: David Gibson (ed.). *Oxford Bibliographies in Ecology*. . New York, Oxford University Press. p. doi:10.1093/OBO/9780199830060-0233.

Chapter 3: Acoustic composition reflects avian community composition across local forest-cover gradients in a coffee growing region in Minas-Gerais, Brazil

I led this work in collaboration with Andrea Larissa Boesing, C.B.L. and Kate E. Jones (UCL/ Institute of Zoology). A.L.B. selected the study sites in Brazil and received the appropriate licenses. Field work was carried out by A.L.B. and myself. The study was designed with discussions from C.B.L. and K.E.J. I led the quantitative approach. The code I developed for this chapter was then used by a Masters student, Charis Declaudure, at Silwood Park (partly supervised by me) who compared my acoustic composition methodology to various acoustic indices with point count data. Her results showed acoustic composition outperformed all acoustic indices, and the manuscript to her upcoming publication can be found in Appendix B.2.

Chapter 4: Applying a deep learning bat detection model to a challenging Brazilian acoustic data set

I led this work in collaboration with A.L.B., C.B.L., K.E.J. and Oisin Mac-Aodha (University of Edinburgh). A.L.B. selected the study sites in Brazil and received the appropriate licenses. Data collection was carried out by A.L.B. and myself. The study was designed with discussions from C.B.L., K.E.J. and O.M.A. O.M.A. designed and trained the CNN model and retrained it for the Brazilian dataset. The UK CNN model was trained with data from multiple sources including the Bat Conservation Trust and Martyn Cooke. Annotations for the UK training data were annotated by myself, OMA and Liz Walsh. The Brazilian training data was put together and labelled by me.

This work contributed to an upcoming publication:

Mac Aodha, O., Balvanera, S.M., Damstra, E., Cooke, M., et al. (n.d.) *Deep Joint Detection and Classification of Bat Echolocation calls*.

The current manuscript can be found in Appendix C.2. My personal contribution to this work was the data I collected in Brazil as well as personally labelling over 3000 audio files of UK bat species to train the model. I also helped design the labelling protocol for others to follow, ensuring consistency.

Chapter 5: Sun coffee plantations provide weak microclimatic buffering effects but not against supraoptimal temperatures.

The data for this chapter was provided by Climate Edge, a start-up founded by Paul Baranowski and Peter Baker. The study was conceived with discussions from P.B., P.B., and C.B.L., with analysis ideas and contributions from Rob Ewers (Imperial College London). C.B.L., K.E.J., P.B., and P.B. contributed to the manuscript.

Chapter 2: The Ecology of the Atlantic Forest

2.1: Introduction

Extending along the southern coast of Brazil, into Argentina and Paraguay, the Atlantic Forest is a domain that once covered 150 Mha and includes many distinct forest subtypes and ecosystems. Its large latitudinal (29˚) and altitudinal (0–2,800 m above sea level) range, as well as complex topography in the region, has created microclimates within forest subtypes, which has led to biodiversity specifically adapted to narrow ecological ranges. The region is incredibly species-rich and is home to charismatic or economically important species such as the black and golden lion tamarin, the red-browned Amazon parrot, and the highly prized palm heart from *Euterpe edulis*. Through widespread human-driven change dating back to the arrival of European settlers in 1500, this realm has been extensively reduced, fragmented, and modified. Nowadays, this region is home to about 130 million people (60 percent of the Brazilian population) and is responsible for producing 70 percent of Brazil's GDP, putting a strain on natural resources and providing challenges to conservation. Due to its high levels of endemic species coupled with a high threat of habitat loss and fragmentation, the Atlantic Forest has been identified as a "biodiversity hotspot." Numerous studies have assessed the effects of habitat transformation on biodiversity and the consensus is that the majority of species are negatively affected. It was puzzling however that few species had actually gone extinct in the wild, even if some extinctions might have gone undetected. Extinctions do not immediately follow habitat change, there is often a time lag of many decades between habitat transformation and extinction. This may suggest that many species in the Atlantic Forest are "living deads," as despite their presence the available habitat no longer supports their requirements. It also suggests that there is a window of opportunity to restoring the domain to avert extinctions before they are realized. However, in this extinction lag may be ending as recent research, such as that by , showing that between 5-7 bird species have likely been driven to extinction, with a further 9 now critically endangered. Current research and policy actions are geared toward optimizing restoration and increasing the extent of native forest cover, bringing hope to the conservation of this unique domain.

2.2: General Overview of Forest Ecology and Forest Extent

Several good sources give detailed overviews of the forest ecology. Galindo-Leal & Gusmão Câmara (2003) is a good general introduction to the history, biodiversity, and human impacts in the region as well as conservation management strategies; however, certain aspects of this text pertaining to current trends and conservation management are likely outdated. Metzger & Sodhi (2009) a special issue in Biological Conservation, focuses on conservation issues in the Atlantic Forest. Joly, Metzger & Tabarelli (2014) provides a comprehensive review of the history of disturbance, the ecology, the ongoing effects of fragmentation, and how climate change is impacting and will impact the Atlantic Forest. Rates of land use change in the Atlantic Forest tend to be monitored on a country-by-country basis, with remote sensing efforts in Brazil, Argentina, and Paraguay revealing different levels of deforestation, as seen in Azevedo *et al.* (2018); Izquierdo, De Angelo & Aide (2008); and Huang *et al.* (2009), respectively. In Brazil, low levels of deforestation are mostly matched by reforestation, which means that the amount of forest cover has either been stable or slowly increased in the past decades. The amount of remaining forest cover in Brazil has been measured by several groups and has been repeatedly updated since the 2000s as high-resolution satellite imagery becomes available. Galindo-Leal and Câmara 2003 reported that about 7–8 percent of Atlantic Forest still remained. Ribeiro *et al.* (2009) calculated the existence of 11–16 percent of native vegetation, but when secondary forest fragments are excluded, the estimated remaining forest would stand at around 8 percent. More recently, using RapidEye imagery with 5 m of resolution, the authors of Rezende *et al.* (2018) have shown that there is actually 26 percent of native vegetation. Because each group used different methods to assess the extent of forest cover, these estimates cannot be compared, and do not indicate that the amount of forest has increased over time.

2.3: Historical Background

The earliest evidence of human activity within the Atlantic Forest is from at least 3220 years ago, and consists of indigenous settlements causing modest disturbance, leaving ceramic evidence, such as that presented by Scheel-Ybert, Beauclair and Buarque (2014). Since the arrival of Portuguese settlers in the 16th century, deforestation for urbanization and agriculture has been rampant. Two of the world's most populous cities (São Paulo and Rio de Janeiro) are located within the Atlantic Forest realm, and Brazil's economy also relies heavily on agriculture, including commodities such as coffee, cocoa, sugar, rice, soybean, and cotton, as explained by Martinelli *et al.* (2010). Much of Brazil's agriculture is exported and/or transported by trucks across Brazil, requiring an extensive road network, which, as shown in Freitas, Hawbaker & Metzger (2010), further exacerbates land use change and forest fragmentation. As for pre-Anthropocene history, there has been some research into the underlying evolutionary and paleoclimatic drivers responsible for the centers of endemism found within this region, such as the study Carnaval & Moritz (2008). Álvarez-Presas *et al.* (2011) used planarians as model organisms to understand patterns of biodiversity, whereas the authors of Carnaval *et al.* (2009) used frogs as indicators in their paleoclimatic models. Fjeldsa & Rahbek (2006) used climate and phylogenetic data to explain the higher diversity of tanagers in the Atlantic Forest around Rio de Janeiro. Batalha-Filho *et al.* (2013) used bird data to demonstrate how the Amazonian and Atlantic Forests were connected, linking the connection to geotectonic events.

2.4: Habitat Types: Present and Future

The Atlantic Forest is commonly divided either into forest types (a particular community of plant species that define a region can be referred to as a physiognomy) or into biogeographical subregions. The types of forest present are: dense and open ombrophilous forests, mixed ombrophilous forests, seasonal forests, and semideciduous as well as mangrove forests. The most studied forest physiognomy is the ombrophilous dense forest, which is believed to have higher levels of biodiversity and endemism than other physiognomies, although this could be due to differences in sampling effort. However, Scarano (2009) argues that these peripheral forest subtypes, including restingas and swamp forests, should be given more priority, making the case that they have a high conservation value due to their oligarchic diversity, with a few dominant species but many rare species at local scales. The western extent of the forest that extends into Argentina (an introduction to this forest can be found in Chebez & Hilgert (2003) and Paraguay (Cartes (2003) presents a background to this forest) forms part of the subtropical semideciduous forest. Silva & Casteleti (2003) defined eight biogeographical subregions: Araucaria, Bahia, Brejos Nordestinos, Diamantina, Interior, Pernambuco, Serra do Mar, and São Francisco. Five of these regions are considered to be centers of endemism, as shown by Tabarelli *et al.* (2010) (cited under Biodiversity of the Atlantic Forest). Assessments of the potential impacts of climate change on the Atlantic Forest have revealed the area to be extremely vulnerable: Lemes, Melo & Loyola (2014) found that as species ranges shift under climate change, protected areas will need to shift with them, particularly for vulnerable taxa like amphibians. Scarano & Ceotto (2015) reviewed the vulnerability of both biodiversity and society of the Atlantic Forest to climate change and discusses important adaptive practices.

2.5: Biodiversity of the Atlantic Forest

Myers *et al.*, (2000) showed that nearly half of all plant species and more than a third of mammals, birds, reptiles, and amphibians are found in only 1.4 percent of the Earth's land surface area, regions which are known as biodiversity hotspots. Hence, by focusing conservation policies and protecting these areas, a disproportionately high level of biodiversity could be protected. A follow-up book about biodiversity hotspots, Mittermeier *et al.* (2005), discusses that 40 percent of the 20,000 plant species, 16 percent of the 688 bird species, 27 percent of the 261 mammal species, 31 percent of 200 reptile species, and 60 percent of 280 amphibian species are endemic to the Atlantic Forest, which means they can only be found within this realm. Together, they represent over 8,650 species, 8,000 of which are tree species. The Atlantic Forest has shown some of the highest levels of biodiversity in the world. Martini *et al.* (2007) identified an area containing 144 species of trees (above diameter at breast height > 4.8 cm) within 0.1 ha in southern Bahia, which is the second highest concentration of tree species in the world. Among animal species, one notable example is the golden lion tamarin, an endangered primate species that Lapenta & Procópio-de-Oliveira (2008) found to have a role in the seed dispersal of ninety-seven species of plants. The Atlantic Forest has high levels of endemism and of habitat loss, making it one of the most endangered biodiversity hotspots. The species that are still present are often trapped within small fragments and unable to migrate, as discussed in Tabarelli *et al.* (2010).

Amphibians are a taxon of particular concern due to their high endemicity and the increasing threat posed by chytrid fungus, as shown in Carnaval *et al.* (2006).

2.6: Biodiversity Data

There have been large-scale efforts to gather data on Atlantic Forest biodiversity. There are several public data sets known as the Atlantic data papers, which represent a high proportion of diversity present in the region, including: 94 species of mammal (Souza *et al.*, 2019), 26 primate species (Culot *et al.*, 2019), 745 bird species (Hasui *et al.*, 2018), 2,095 epiphyte species (Ramos *et al.*, 2019), 279 butterfly species (Santos *et al.*, 2018), 528 amphibian species (Vancine *et al.*, 2018), and 98 bat species (Muylaert *et al.*, 2017). Several of these data sets monitor assemblages (e.g. Culot, *et al.* (2019) over time. As well as simple occurrence records, some of the data sets collate information on species traits and interactions, such as plant–frugivore interactions in Bello *et al.* (2017), and bird traits including body mass and wing length in Rodrigues *et al.* (2019). Going beyond simple species occurrence data allows researchers to investigate patterns over time such as demographics and make inferences about how anthropogenic stressors affect morphology. Interaction data is particularly valuable as it can uncover species mutualisms and the extent to which certain species depend on others. This contributes to the deeper understanding of how ecological communities work, which can be useful to inform better conservation policies. For the state of São Paulo, there is also information available on the SinBiota 2.1 platform, as described in Mira *et al.* (2011). This was created by the Biota/Fapesp program to integrate information generated by all researchers funded by this program.

2.7: Habitat Transformation

Habitat loss, fragmentation, and degradation are the major threats to biodiversity in the Atlantic Forest, thus there is a substantial body of literature on this subject. Lôbo *et al.* (2011) showed that habitat transformation has over time modified diverse communities within five physiognomic subtypes (including evergreen, semideciduous, and open forest) of the Atlantic Forest into a homogenized set of disturbance-specialist species. Indeed, studying in southern Bahia (world's second highest concentration of tree species), the authors of Benchimol *et al.* (2017) found that forest loss leads to nonrandom floristic shifts, such that shade-intolerant species (e.g., pioneers) become more common than shade-tolerant species below 30 percent of forest cover. The loss of species also drives evolutionary changes in seed size, as shown by Galetti *et al.* (2013). Within the evergreen and semideciduous forests, Santos *et al.* (2008) demonstrated that fragmentation and the resulting creation of more edges have severely reduced functional traits of tree assemblages. But habitat loss is not the only driver of species loss. Using a multi-taxa database with over 2,200 community-level estimates from 1,097 sites, Püttker *et al.* (2020) showed that forest-dependent species respond negatively to habitat loss and fragmentation, and that in areas with over 30 percent of forest cover, habitat fragmentation was as important as or more important than habitat loss in driving changes in species richness. This was not the first time that it has been shown that the effects of fragmentation on species are dependent on the amount of forest cover. Pardini *et al.* (2010) demonstrated that within the ombrophilous dense forests, the size of a forest fragment only positively affects biodiversity when the landscape level forest cover is intermediate (around 30 percent of forest cover). The impacts of edge effects have been shown for a variety of abiotic and biotic conditions. Magnago *et al.* (2015) showed that forest edges are drier and warmer, and these abiotic changes affect forest structure. Changes to habitat structure can then have knock-on effects on other taxa. For instance, Filgueiras, Iannuzzi & Leal (2011) found that dung beetle diversity was impacted by the impoverished flora of small patches. Banks-Leite, Ewers & Metzger (2010) found that edge effects likely drive the patch area effects on birds in the Atlantic Forest. This is because large patches experience a weaker influence of edge effects than small patches, which have higher edge-to-area ratio.

2.8: Vertebrates and Habitat Transformation

Birds are one of the most commonly studied taxa in the Atlantic Forest due to their diversity and sensitivity. Responses of birds to fragmentation have been well documented: Zurita & Bellocq (2010) found forest cover to be the main driver of differences in bird communities in Argentina, while Morante-Filho *et al.* (2015) showed the number of bird species in southern Bahia abruptly changed at a threshold of 50 percent forest cover. Banks-Leite, Ewers & Metzger (2012) showed that around São Paulo responses of bird communities to fragmentation do not conform to the classical ecological species–area relationship. Instead, the main drivers of changes in bird communities across a gradient of disturbance are purported to be individual species reaching their extinction threshold. Despite conservation efforts, many species remain at high risk of extinction (as demonstrated in Canale *et al.* (2012)), particularly large mammals such as the jaguar, which, as shown by Paviolo *et al.* (2016) have undergone high rates of extirpation. Umetsu & Pardini (2007) found that small mammals, particularly endemic species, have been found to be sensitive to land use change. The response of bats is less clear, as Gorresen & Willig (2004) found the highest levels of bat diversity in moderately fragmented landscapes. With regards to the response of amphibians to fragmentation, Becker *et al.* (2007) found habitat loss to be a key driver of amphibian declines, particularly for forest species; and amphibians are particularly vulnerable to the coupled effects of fragmentation and climate change, as shown by Loyola *et al.* (2014). The effects of fragmentation on reptiles are less studied, but Lion *et al.* (2016) demonstrated that reptiles can benefit from even small forest fragments.

2.9: Conservation and Policy

The Atlantic Forest is protected by the Forest Code, which is an environmental law created in Brazil in 1965, when most of the deforestation had already taken place. By law, landowners are required to set aside 20 percent of their land for native habitat, as well as protect riparian forests, hilltops, and other environmentally sensitive areas. The Forest Code was revised recently, weakening the protection of the Atlantic Forest, as discussed by Soares-Filho *et al.* (2014). This revision is particularly problematic given that Banks-Leite *et al.* (2014) has shown that at least 30 percent of native habitat is required to protect biodiversity within the Atlantic Forest. Due to the pressing need to preserve its unique yet endangered biota, a group of academics, NGOs, industry, and government formed the Atlantic Forest Restoration Pact, an initiative which aims to restore 15 Mha of habitat in the Atlantic Forest by 2050. This pledge comes as part of Brazil's commitment to the Bonn Challenge. Crouzeilles *et al.* (2019) shows that the Atlantic Forest Restoration Pact has already facilitated the restoration of roughly 700,000 ha, estimating that by 2020 there will be 1.5 Mha under restoration. Rezende *et al.* (2018) has estimated that if landowners comply with the new Forest Code to restore riparian forest (i.e., forest strip along rivers), by 2038 the vegetation cover in the Atlantic Forest will be close to 35 percent, bringing hope to the preservation of this charismatic and species-rich system.

Chapter 3: Acoustic composition reflects the avian community across local forest-cover gradients in a coffee growing region in Minas-Gerais, Brazil

3.1: Abstract

Understanding how groups of species respond to local land use and forest cover can contribute to more efficient conservation policy, but it is still challenging to obtain data on biodiversity. Acoustic analysis is emerging as a tool to rapidly assess biodiversity, but current methodologies vary in the extent to which they truly measure the underlying community. We use a holistic soundscape approach to examine how local avian community composition varies with local acoustic composition of Brazilian coffee farms and adjacent forest along a gradient of forest cover. This study built on recently developed methods to uncover drivers of acoustic differences between land uses and across a forest cover gradient. I found that 21% of variance between soundscapes could be explained by a forest cover gradient as well as different land uses, and that this variance correlated with bird community composition (Pearson's $r = 0.6$). Specific indicator frequencies driving the acoustic differences were identified and these were distributed throughout the day, highlighting the importance collecting acoustic data across longer time spans rather than concentrated around dawn. A majority of indicator frequencies were found to be from avian sources, though insects also contributed.

3.2: Introduction

Acoustic monitoring of biodiversity is a rapidly expanding field of ecology, as a lot of data can be collected without the need for labour intensive field work or detailed knowledge of local biodiversity in a non-invasive manner. It has the potential to contribute to rapid assessments of biodiversity which are necessary for achieving the UN 2030 Sustainable Development Goals (United Nations, 2015) . By building on emerging acoustic methods and thus creating techniques to assess communities from soundscape data is key to making rapid assessments a reliable form of monitoring.

The potential of acoustic data as a tool to monitor biodiversity in a less labour intensive manner has been boosted by low-cost passive acoustic monitoring devices such as AudioMoths (Hill *et al.*, 2019), customisable products such as Raspberry Pis (Upton and Halfacree, 2014), and the rapid development of machine learning technology (Stowell *et al.*, 2018). However, currently a gap exists in acoustic analytical tools between generalised acoustic indices and overly specific algorithms trained to detect a particular species or taxa, though this gap is starting to be filled (Buxton *et al.*, 2018). The drawback of producing specific algorithms is they require extensive training datasets, which take time and expert verification to compile, as well as computational expertise to design, train and test the algorithm (Gibb *et al.*, 2019). This is still a challenge for taxa like birds, as many other sounds are also made in the same frequency range, though progress is being made with neural network methodologies (Stowell *et al.*, 2018). However, collecting and verifying sufficient data in megadiverse regions such as the Neotropics make automation a challenge.

Over 60 acoustic indices have been developed to summarise various aspects of acoustic data that correlate with components of biodiversity, often species richness (Rajan et al., 2019; Sueur et al., 2014), however they have also been found to be inconsistent between biomes (Eldridge et al., 2018). Guidelines are improving, and more recent recommendations for the use of acoustic indices proposes using a combination of at least four indices (Wimmer, Williamson and Roe, 2014; Buxton *et al.*, 2018; Bradfer-Lawrence *et al.*, 2019). In work that has used acoustic indices to examine avian responses across land use gradients, the responses of indices have been mixed (Shamon *et al.*, 2021). So far links between indices and community composition of birds are uncertain, as specific drivers of acoustic indices are difficult to uncover, and often differences in acoustic indices between land uses is presented without the drivers of these indices being fully explored (Scarpelli, Ribeiro and Teixeira, 2021). In addition, recent analysis of combinations of seven acoustic indices in the Atlantic Forest have found that they are more correlated to the number of vocalisations rather than richness or diversity (Gaspar, 2021). Though acoustic indices are sometimes used well, and protocols are improving, they may not necessarily reflect underlying community composition at local scales.

Soundscapes are summaries of the sound in a particular place and are a useful tool to visualise patterns of acoustic activity in time (Pijanowski *et al.*, 2011). An approach that explores broad patterns in soundscapes has been developed (Campos-Cerqueira and Mitchell Aide, 2017; Furumo and Mitchell Aide, 2019), based on the idea that spectrograms of sound across various locations can be treated in the same way as species composition data can be compared across sites. In this methodology, each time-frequency pixel within an averaged spectrogram is treated as if it were a species, and each respective amplitude value indicates the presence or absence of that sound. This is based on “acoustic niche hypothesis” which states that birds living in the same community separate their calls by either time or frequency to avoid overlaps with other species (Farina *et al.*, 2011). Therefore, if a site has an occupied acoustic niche that is not occupied in another site, this can be detected as an indicator time-frequency, which can lead to the identification of taxa contributing to the acoustic composition of certain locations or differences in temporal acoustic composition. If conducted in parallel with in situ monitoring of biodiversity, the validity of these methods as indicators of local biodiversity trends can be demonstrated.

The impact of fragmentation on avian communities is well studied in the Atlantic Forest of Brazil, but thus far most studies rely on intensive field work to monitor communities (Morante-Filho *et al.*, 2015; Boesing, Nichols and Metzger, 2018). Here, we set out to answer how acoustic community composition varies across a land use and degradation gradient and whether acoustic community composition reflects the composition of avian communities. By monitoring acoustics at different local land uses across a gradient of surrounding forest cover, biotic indicators of each landscape type can be recorded. I predict that soundscapes will vary most between land use types as well as across a forest cover gradient. It is expected that the edge of the forest will sound more similar to the forest than the agricultural land use. I hypothesise that there will be a correlation between the bird community and the acoustic communities across sites and as a result most indicator frequencies be driven by bird taxa. In addition, I predict that the times at which the indicator frequencies occur will be widely dispersed across the day, as this method extracts the main sources of difference between sites is likely to occur when one site contains sound when others so not. We present a potential suite of generalised acoustic analysis methods that monitor and capture acoustic indicators of environmental degradation using a field case study done in Brazil's Atlantic Forest.

3.3: Methods

Study area

Fieldwork was conducted between January and March 2019 on the boundary of Minas Gerais and São Paulo states of Brazil, within 50km of the town of Poços de Caldas. The altitude in this region is 750-1200 m a.s.l., with a mean annual temperature of 17.7ºC and average annual rainfall of 1695 mm (Pompeu *et al.*, 2009). Prior to agricultural intensification, the original forest physiognomies in this region were a mixture of ecotonal, semi-deciduous and ombrophilous dense (Joly, Metzger and Tabarelli, 2014). Sampling was conducted at 11 sites within a 250 000 hectare highly fragmented region (21º47'S, 46º33'W), which has 13% native forest cover and 80% agriculture (dominated by *Coffea arabica* monoculture plantations). The remaining 7% landcover consists of urbanised areas, water bodies and forest plantations. In total, 11 sites (Figure 3.1A) were surveyed using a paired forest-coffee plantations design (i.e. sampling sites were composed by a farm adjacent to a forest fragment). Sites were selected to reflect a wide range of landscape level forest cover, which ranged from 9% to 60% within a 1000 m radius. The scale at which responses to

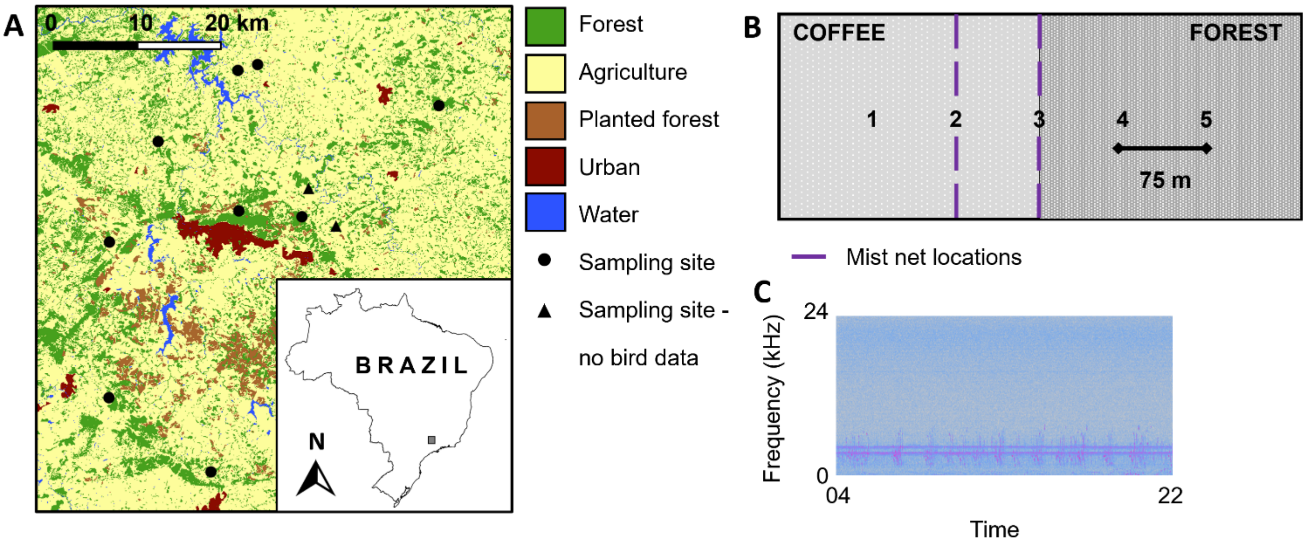


Figure 3.2: ***A:*** *Map of study area land use and sampling sites.* ***B:*** *Diagram of AudioMoth placement at each sampling site as well as placements of mist nets at relevant sites.* ***C:*** *Diagram of acoustic sampling rate schedule, at each sampling point.*

fragmentation are best measured varies between species, but 1000 m has been found to be significant and is commonly used across multiple studies (Boscolo and Metzger, 2009; Pereira, Oliveira and Torezan, 2013; Regolin *et al.*, 2017). This field work and bird sampling permissions were granted under the research permit number SISBIO 64745/1.

Acoustic sampling

Acoustic sampling was conducted using AudioMoth recorders (Hill *et al.*, 2018). These are small, low-cost acoustic recorders that can record customised recording schedules at sampling rates up 384kHz. At each sampling site, five recorders were deployed every 75 m along a transect (see Figure 3.1B). The recorders were deployed at 1m from the ground, in plastic bags to protect from rain. The recorders were set to record 1 minute every 5 minutes from 04:00 to 22:00 at a sampling rate of 48 kHz. This schedule was selected due to previous work in the Atlantic Forest showing that recording every fifth minute retains the most information across sites (Pieretti et al., 2015). The recorders were deployed for 8-14 consecutive days. Recorders were firstly placed at the edge, then 75m and 150m into the forest and coffee plantation in such a way that they were at least 75 m and 150 m from any edges respectively. Due to this constraint, in a few of the smaller forest fragments, the recorder at transect point 5 was not deployed at three sites. Though distances between recorders of 100m would have been ideal (Dixon, Baker and Ellis, 2020), this was not possible due to the size of small forest fragments.

The temporal resolution of acoustic data collection is also important and must consider the patterns of activity of the target taxa. However, this has resulted in acoustic monitoring surveys for bird activity focusing heavily (sometimes exclusively) at dawn (Venier *et al.*, 2012; Farina *et al.*, 2015), with limited recording during the middle of the day (Scarpelli, Ribeiro and Teixeira, 2021). This type of temporal sampling misses species that are acoustically active at other times of day (Metcalf, Barlow, Devenish, *et al.*, 2021) and also may miss differences in behaviours of the same species at different sites. Higher temporal resolution in acoustic sampling can contribute to fill knowledge gaps in how communities of species and their behaviour respond to their local environment (Teixido *et al.*, 2020). Therefore, to better assess as much of the community as possible, many shorter recordings large sections of the day should be done.

Bird sampling

At nine of the eleven sites (Figure 3.1A), mist-netting was performed for three consecutive days (this matched the first three days of recording). Six 12 × 3 m with 29 mm mesh Ecotone mist nets were placed along the edge of the forest as well as 100m into the coffee plantation, parallel to the edge (Figure 3.1B). Mist nets were operated between sunrise until 6 hours later, to capture the higher bird activity time, totalling 216 mist net hours per site, where one mist-net hour represents one mist net open for one hour (Roos, 2010). All captured birds were identified to species level in the field and marked (by painting neon pink nail polish onto a talon) before release, to avoid double counting individuals.

To determine the composition of the bird community, a Principle Coordinates Analysis (PCoA) was done, with a Bray-Curtis similarity and the first axis (PCo1) was selected for further investigation. This was done using the vegan package (Oksanen *et al.*, 2019) in R version 4.0.2 (R Core Team, 2020).

Forest cover data

Forest cover was calculated within a 1000m radius of each site, using QGIS (QGIS Development Team, 2021). A 5 m resolution land-cover map was provided by A.L.B. for this work.

Acoustic analysis

All acoustic and statistical analyses were performed in R version 4.0.2 (R Core Team, 2020). A general soundscape approach, as described in (Furumo and Mitchell Aide, 2019) was selected as it compares the each time/frequency point explicitly to reflect temporal acoustic composition. These methods create an average soundscape for each site, then treats each soundscape as an acoustic community which can be compared to other acoustic communities. This was done by creating spectrograms of all recordings taken at the same time of day for each recorder, then a mean spectrogram for that minute is calculated. The frequency bin width of the spectrograms was set to 0.2kHz in width and all files were resampled to a frequency of 24 kHz, which corresponds to a Nyquist frequency of 12 kHz, to make sure all avian activity was included, as most birds call up to 8kHz (Kasten *et al.*, 2012). This was done by adapting the spectrogram binning function described in Kasten *et al.* (2012). The lower 1 kHz was removed, as these bandwidths contain the most geophony (noise created by wind, rain etc.) (Furumo and Mitchell Aide, 2019). The mean amplitude of each frequency bin for the spectrograms was obtained so a single vector for each recorded minute of the day could be calculated. These vectors were collated into a soundscape matrix.

Once the matrices were obtained for each recorder, they were flattened into vectors and arranged into a data structure analogous to a community matrix where each column represented a single time frequency bin and each row represented a different site. By approaching soundscapes as a community, each time-frequency bin can be thought of as a species.

The next step was to determine whether there were clear differences in acoustic composition across both the land-use transects and the forest cover gradient. To uncover whether differences in soundscapes could be correlated to bird communities, the soundscapes were trimmed to only include the sound from 04:30 to 19:30. This corresponded to half an hour before sunrise and half an hour after sunset during the fieldwork season to capture a maximum of bird activity.

PCoA were performed on these trimmed soundscapes using Bray-Curtis similarity, and the first axis (PCo1) was selected for further investigation. This was done using the vegan package (Oksanen *et al.*, 2019) in R.

Statistical analysis

A linear mixed effect model was constructed to investigate the relationship between the acoustic composition PCoA 1 score, forest cover and land use. To determine the optimal model structure of the forest cover and land use variables, the beyond optimal model selection protocol from Zuur et al. (2009) was used. This protocol determines the order in which fixed and random effects are selected, and then the best model is selected using AIC. Models were compared using the Akaike Information Criterion (AIC) (Burnham and Anderson, 2016) and the selected model had ΔAIC>2 than the next best model.

Sampling site was a random factor within the model. Model assumptions were verified by plotting the residuals against fitted values (see Appendix B.1 Figure 1). All model verification and analysis was done using the nlme package in R (Pinheiro *et al.*, 2020).

Then to examine the relationship between the acoustic community and the bird community, a simple linear model was run between the PCoA of the acoustic data and the bird data.

Indicator analysis

To determine the acoustic niches driving the differences between different sites, indicator analysis was performed on the soundscape community matrix using the labdsv package in R (Roberts, David and Roberts, 2019). This indicator frequency analysis is based on indicator species analysis, where the value is highest when all individuals of a species are found within a single group of sites (Dufrêne and Legendre, 1997). Specific time-frequency points were considered in the same way as species would be. Amplitude values that were less than 0.1 were set to zero to ignore less important frequencies. The diurnal soundscapes were firstly grouped by land-use, then categorised according to forest cover in 10% bins. The indicator analysis then extracted particular time-frequencies which were strongly associated with certain forest cover categories. This type of analysis removes sounds common across multiple categories, and highlights those that are strongly associated to a single category. Specific time-frequencies were considered indicators if they a p-value below 0.05, based on 10.000 iterations.

To determine the source of clusters of the indicator frequencies, one minute in the centre of the cluster was selected, and a random subset of ten files recorded at that minute for the corresponding land use and forest cover level were systematically listened to. For each file, the taxa producing the sounds made at the frequency of interest were noted. For broader clusters, multiple minutes within the cluster were listened to. Details of files listened to are in Table 4 of the Supplementary Information.

3.4: Results

Soundscape and bird community analysis

Around 77, 000 60-second files were recorded and analysed for this work, representing over 53 full days of sound. For spectrograms of average soundscapes, see supplementary material.

A total of 773 birds were captured, 471 in the coffee and 302 at the edge. A total of 77 species were captured: between 9 and 29 species of birds were captured in the coffee (mean 17, s.d. 6.48), and between 10 and 26 species were captured at the edge (mean 17, s.d. 4.74). The most common species were *Zonotrichia capensis* (122 individuals) and *Sporophila caerulescens* (121 individuals) totalling 43% of all captures.

The mixed effects linear model showed that across the forest cover gradient there was a directional shift in composition of both acoustic and bird communities which was consistent between land use categories (Figure 3.2). Forest and edge acoustic communities were more similar to one another than the coffee acoustic community within each site.

Table 3.1: *Model selection results. A is acoustic composition (PCoA 1 value), F is forest cover within 1000m, L is land use category, and S is site.*

Model structure	**AIC**
A~F+L+(1|S)	**-62.78473**
A~F*L+(1|S)	-43.34625
A~F*L	-44.17307
A~F*L+(1|S +1|F)	-42.39815

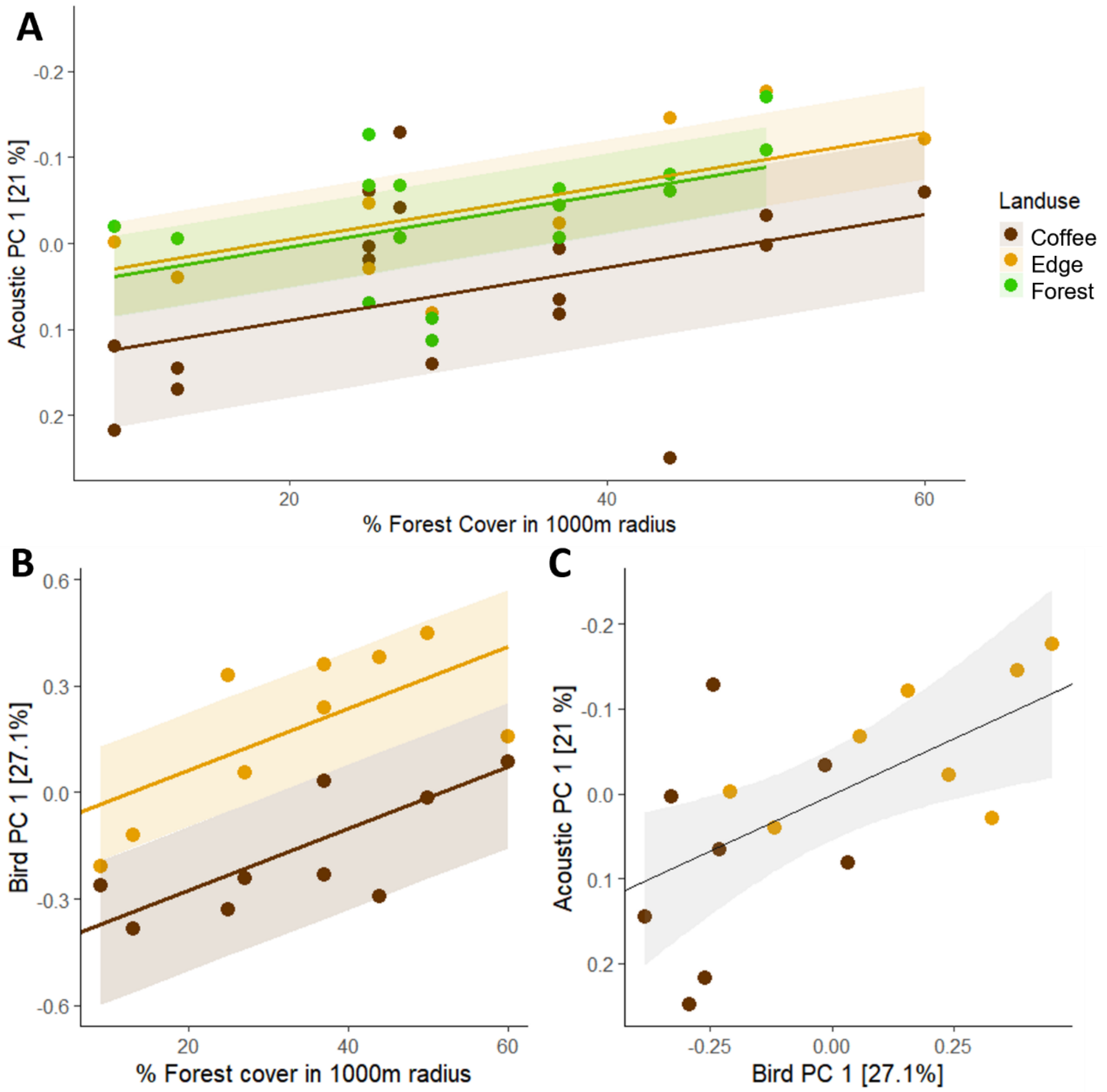


Figure 3.2: *A: Graph of soundscape PC 1 against forest cover within 1000 m radius. Land uses are denoted by colour (brown = coffee, orange = edge, green = forest). Lines represent the best fit model, showing the different intercepts between land uses B: Graph of bird PC 1 against forest cover within 1000 m radius, with lines denoting the best model fit. C: Graph showing PC 1 of soundscape against PC 1 of bird data, the line denotes a linear model fit with an adjusted* R^2 *of 0.36 , p=0.0086.*

The best model was a linear mixed effect model with forest cover and landuse as additive fixed effects see Table 3.1 for outputs. This model structure was the best fit for both the soundscape and bird community PC data.

Table 3.2: *Output parameters of selected model with lowest AICs for both the soundscape model and the bird model. The intercept values denote the baseline of coffee land use. Models are both linear mixed effects model with the formula: PCoA output ~ Forestation + Landuse with site as a categorical random effect. Model validation plots and full model outputs are in the Supplementary Materials.*

	Soundscape model				**Bird model**			
Predictors	*Estimates*	*df*	*C.I.*	*p-value*	*Estimates*	*df*	*C.I.*	*p-value*
(Intercept – coffee land use)	*0.152*	*32*	*0.063 – 0.241*	***0.0015*****	*-0.473*	*8*	*-0.063 – -0.315*	***<0.0001 ******
Forest cover	*-0.003*	*9*	*-0.006 – -0.0004*	***0.0311****	*-0.0087*	*7*	*0.005 – 0.0125*	***0.0001******
Land use (Edge)	*-0.094*	*32*	*-0.149 – -0.040*	***0.0012*****	*0.3644*	*8*	*0.218 – 0.511*	***0.0002******
Land use (Forest)	*-0.086*	*32*	*-0.133 – -0.039*	***0.0008*****				
Conditional R^2	*0.571*				*0.719*			

Indicator frequency analysis

The indicator frequency analysis showed a contrast in both number and temporal distribution of indicator frequencies between land uses (Figure 3.3). Within the coffee land use, surrounding forest cover of 30-39% had the most significant time- frequency indicators, whereas in the forest, more significant indicators were found for lower surrounding forest cover of 10-19%. There was a distinct difference in the temporal distribution of the indicator frequencies, as in the coffee land use, the number of indicators peaked in the morning, but in the forest the peak number of indicators occurred in the early afternoon. The sites with forest cover above 50% had very few indicator frequencies (none in the coffee sites and only 3 in the forest/edge sites).

By selecting clusters of indicator frequencies, and listening to the raw data, the underlying biotic driver could be determined (see Table 4 of Appendix B.1 for additional detail). Sometimes both birds and insects were present within a certain frequency band, so a single biotic driver could not be isolated.

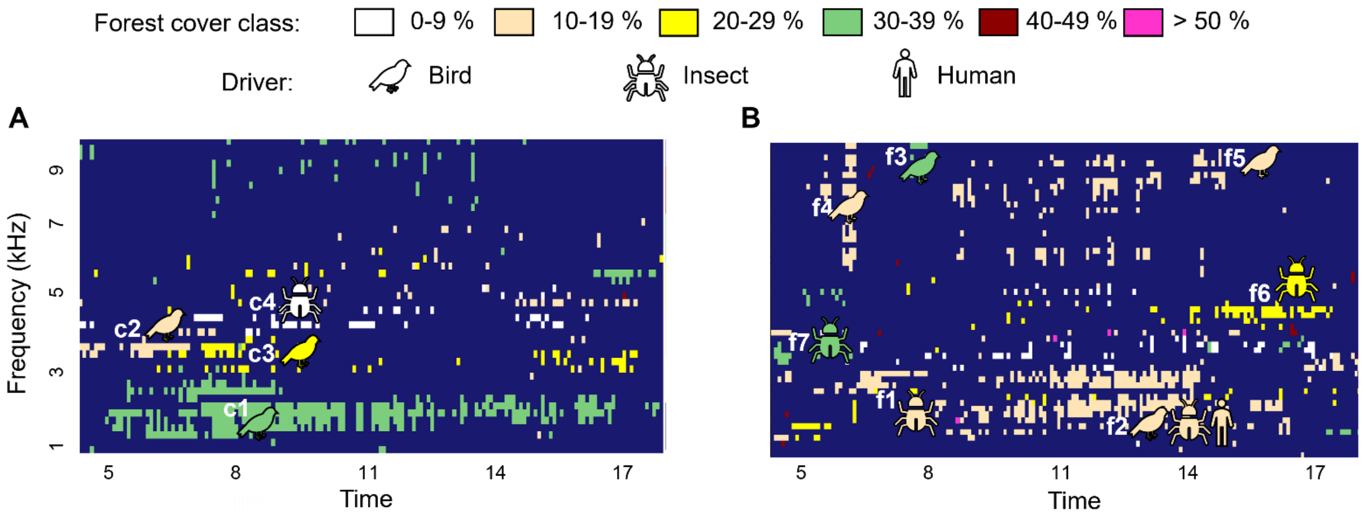


Figure 3.3: *Spectrograms showing times and frequencies of indicator frequencies at different levels of local forest cover. The taxa driving the certain clusters of indicator frequencies are marked with either a bird, insect or human icon.* ***A:*** *Indicator frequencies between coffee samples.* ***B:*** *Indicator frequencies between forest samples*

3.5: Discussion

This study shows that acoustic composition can reflect avian community composition across a landscape gradient of change and habitat types. It can also identify the main drivers of difference between sites as indicator frequencies were driven by birds, and a few by insects. This allows for direct identification of drivers of acoustic trends, without the need to use proxy indicators such as acoustic diversity index or acoustic complexity index. This is an important finding as it shows acoustic composition is a straightforward, intuitive method to quantify soundscapes in biodiverse regions. Acoustic composition has the potential to make biodiversity monitoring more accessible, reliable and scalable.

The main difference between soundscapes was between coffee, edge and forest land-uses, though edge and forest were similar, which was expected because of the small size and high levels of degradation of forest patches. Though a significant relationship was detected between the PCo1 of acoustic composition and the local forest cover, the lower limit of the 95% confidence interval was close to zero (0.0004), so care should be taken when making conclusions from this data. However, the mist-netting data showed a more significant relationship between community composition and forest cover as well as a significant effect of land use. As the soundscapes used in the this analysis were averaged over several days, the sounds that are present are highly characteristic of the recording site, which means they are biased towards loud insect sounds and territorial bird species that call at a particular time of day (Barbosa *et al.*, 2017). This will make their call frequencies more pronounced when averaging across several days than a call from a species that is more sporadic in its acoustic activity.

The wide temporal distribution of indicator frequencies confirms that acoustic differences between sites are not only driven by biotic sounds around dawn, and highlights the importance of high temporal resolution in acoustic sampling (Metcalf, Barlow, Devenish, *et* *al.*, 2021). The difference between the temporal distribution of indicator frequencies between coffee and forested land uses is unexpected. These results imply that the acoustic composition around dawn contributes to much of the differences across the forest cover gradient between coffee land use, and in contrast that between forested land uses, the drivers of difference are more evenly distributed throughout the day. One drawback of exploring indicator frequencies as opposed to manual species detection through systematic listening is that if multiple species call within the same frequency range at the same time it may obscure true diversity. This is likely to occur in the Atlantic Forest as bird calls have been shown remain relatively clustered in bandwidth and time (Kleyn, Kaizer and Passos, 2021). Therefore, despite detectable differences between land uses and across a fragmentation gradient, acoustic monitoring that is verified by experienced ornithologists may be more useful until such a time when robust species classification algorithms are built for the region.

I found that birds drove most of the difference between coffee land use across a forest cover gradient, but a mixture of birds and insects drove differences between forested sites. Other studies using similar methods found that insects were the primary driver of acoustic space use, insects though much of this contribution occurred nocturnally (Aide *et al.*, 2017), so though the differences in acoustic composition may correlate with differences in the bird community, the role of insects must be considered. This is consistent with an index-based approach study done in Africa, where insect acoustic richness peaked at night, while avian acoustic richness was higher in the day (Lopes, Roberto and De Brito, 2018). Since doing this work, a masters student (Charis Declaudure) used the acoustic composition methods at Silwood Park in the UK paired with avian point counts and found acoustic composition to be a more useful metric than any commonly used acoustic indices (see Appendix B.2). As species were not identified in this work, links to specific drivers from the mist-netting was not done, which highlights the continued demand for expert ornithologists who can link calls to species.

All monitoring methods have gaps and biases, which is why using a combination of techniques is optimal. Though acoustic monitoring is a useful tool to assess overall patterns in avian communities, non-vocal birds are missed through this type of analysis (Vold, Handel and McNew, 2017). Using mist-nets to validate the findings of acoustic analysis has limitations as the types of birds typically captured in mist nets are small, understorey birds. Ground mist nets deployed in forested areas will primarily capture understorey birds, and not the full community of birds present within a forest (Robinson, Lees and Blake, 2018). However, understorey birds are more sensitive to habitat fragmentation and forest loss, which is better detected by mist netting surveys (Uezu and Metzger, 2011). Overall, conducting point counts or have an expert ornithologist listen to recorded audio could provide better validation for acoustic data.

This holistic methodology can be useful in monitoring communities over time in changing landscapes as well as for cross-site comparisons, as it could detect temporal shifts in biotic sounds across seasons. Having a framework to monitor such changes and identify acoustic and hence biotic drivers of these changes could add to the current toolbox of acoustic methods. This method was able to detect broad differences in composition without the need for extensive machine learning tools or high-performance computing, lending it useful to form part of a low-cost toolkit. However, recent unsupervised machine learning approaches have been able to detect temporal trends in soundscapes as well as differences in avian communities (Sethi *et al.*, 2020). If a good quality bird call detection and classification tool were to be developed, the specific sources of acoustic difference could be linked to particular species.

Chapter 4: Applying a deep learning bat detection model to a challenging acoustic data set

4.1: Abstract

Monitoring the current state of biodiversity and its responses to change over time is necessary to put forward effective conservation solutions. Tropical regions are highly biodiverse but lack data relative to the high numbers of species present. Acoustic monitoring of biodiversity is a rapidly expanding field, but as data accumulation has become cheaper, there remains a need for accurate analyses tools. Bats are a highly diverse group that are well-suited to acoustic monitoring but current acoustic analysis tools have primarily been designed for bats in the northern hemisphere.

Here, we show that by using an existing convolutional neural network (CNN) optimised with data from the UK, and retraining it with data from new location in Brazil, the model can successfully detect bat calls in the new region, with a high accuracy (0.9645). Two families of bats were detected (Molossidae and Vespertilionidae), for which the model had an average precision of 0.9523 and 0.9767 respectively. We also extracted spectral features of the detected calls and identified 17 bat call sonotypes (6 and 11 sonotypes within Molossidae and Vespertilionidae respectively). With further manual examination, calls could be linked to 5 bat genera. Then, we explored how these sonotypes responded to local land use type and forest cover. Though there was no effect of local forest cover, both families of bats were significantly more abundant in coffee land use than forest land use overall, as were 4 of the identified sonotypes. This shows how applying a CNN specifically designed for bat calls can contribute to detection and identification of bat communities in biodiverse regions, which can then be used to monitor ecological trends.

4.2: Introduction

Accurate biodiversity monitoring tools are critical to assess ongoing effects of anthropogenic change (Schmeller *et al.*, 2017). Halting declines in biodiversity forms part of the 15th goal of the United Nations Sustainable Development Goals (United Nations, 2015), hence reliable data on biodiversity trends is essential to review progress towards this. Traditional in situ monitoring is labour intensive and expensive, hence the recent rise in the field of bioacoustic monitoring. With the development of cheap Passive Acoustic Monitoring (PAM) sensors, such as AudioMoths and Raspberry Pi recorders (Upton and Halfacree, 2014; Hill *et al.*, 2019), acoustic monitoring has become a popular option with its potential to collect vast amounts of data with ease and extract biologically relevant information (Browning *et al.*, 2017). Considering ongoing advances in remote sensing and machine learning, there is a strong potential for PAM to provide high quality monitoring in areas that are otherwise difficult to access.

There has been a fifteen fold rise in the publication of PAM-based studies between 1992 and 2018 (Sugai *et al.*, 2019), and have been implemented in both terrestrial and aquatic environments, particularly where visual detection is difficult such as in dense forests, underwater or nocturnal monitoring (Desjonquères, Gifford and Linke, 2020). Audiomoths (Hill *et al.*, 2018) are a popular low-cost recording device that are beginning to be widely used for acoustic data collection as they perform well when compared to more expensive acoustic recording devices (Toenies and Rich, 2021). This increase in use of PAM has resulted in the rapid accumulation of data and a bottleneck forming with regards to analytical tools (Gibb *et al.*, 2019; Stowell, 2022). More sophisticated machine learning based tools such as convolutional neural networks (CNNs) are emerging, which can identify the presence of particular species or taxa (Gibb *et al.*, 2019; Stowell, 2022). To develop a robust species identification CNN, it must be trained on spectrogram data with an associated species label, which still requires qualified human labellers, though this dataset can be augmented (Stowell, 2022). This process is a hurdle in regions lacking in acoustic knowledge on target species, which is a challenge in megadiverse regions (Arias-Aguilar *et al.*, 2018).

Bats are a highly diverse order of mammals, with 1470 identified species, they represent ~20 % of all known mammal species (Solari and Baker, 2007; Lei and Dong, 2016), and are sensitive to land-use change and habitat fragmentation (Meyer, Struebig and Willig, 2015). Around 1000 of these species use echolocation (Boonman *et al.*, 2013), which makes them a good target for acoustic monitoring, particularly as they use ultrasound which does not overlap in frequency with most other taxa. Bat call detection tools (which locate bat calls within sound files, but do not identify the type of bat calling) are more generalisable than classifiers (classifiers sort detected calls into distinct groups such as species) as bat calls are quite distinctive from other sounds (Mac Aodha *et al.*, 2018). However, once calls have been detected, classifying them into genera or species requires more detailed expertise of local call structures. In the early 2000s, bat identification tools were primarily based on multivariate statistical methods (Russo and Jones, 2002; MacSwiney G., Clarke and Racey, 2008), which were then overtaken by machine learning methods (Britzke *et al.*, 2011; Walters *et al.*, 2012; Zamora-Gutierrez *et al.*, 2016; Roemer, Julien and Bas, 2021). Though machine learning methods have produced encouraging results, certain groups of bats remain difficult to discern from one another (Zamora-Gutierrez *et al.*, 2016). More recently, deep learning approaches have shown to have high detection and classification accuracy, but require extensive verified reference calls for training, so call libraries must be compiled for regions of interest (Mac Aodha *et al.*, 2018; Chen *et al.*, 2020). Available commercial tools to detect and classify bat calls from acoustic data are heavily skewed towards Europe and North America (Sugai *et al.*, 2019) and vary in their agreement (Lemen *et al.*, 2015; Nocera *et al.*, 2019). Megadiverse regions such as Brazil lack such libraries, but by applying and adapting existing deep learning models with transfer learning, bat call libraries for new regions can start to be collated. BatDetect was the first CNN model developed for bat call detection within audio files (Mac Aodha *et al.*, 2018; Zamora-Gutierrez *et al.*, 2021). It has been updated and the most recent version aims to both detect calls and classify them into pre-determined classes, which are determined by the user (Mac Aodha *et al.*, no date). The Bat Detect model takes in information about the location of calls within a spectrogram (maximum and minimum frequency, start and end time of call) and their associated class, then uses this to predict the presence, location and class of calls within other audio data. This is in contrast to existing deep learning based solutions that only identify the start time of calls (Mac Aodha *et al.*, 2018). BatDetect is free and open source (this includes all code and training data) in contrast to commercial bat detection software and can also be run without GPU. The BatDetect model can be applied to new regions by retraining and evaluating it with a new set of training and test data, without editing the analytical pipeline. New data from regions lacking libraries can thus apply BatDetect not only to detect calls, but also to analyse and classify calls into classes, which can contribute to developing species ID tools for these regions.

Brazil is a megadiverse country, with over 180 known species of bat present (Arias-Aguilar *et al.*, 2018), of which 121 species occur in the Atlantic Forest domain that extends along the south east coast (Delgado-Jaramillo *et al.*, 2020). The Atlantic Forest domain is threatened by fragmentation and land use change, having lost almost three quarters of its initial extent (Rezende *et al.*, 2018), and bat communities are negatively affected by more deforested, fragmented landscapes (Faria, 2006). Currently no formal call libraries exist, though call information has been compiled for 65 species with varying levels of detail (Arias-Aguilar *et al.*, 2018). This information ranges from descriptions of the acoustic features of a few calls, to more detailed call libraries based on recorded passes in the field, many of which are from recordings in the north of Brazil and not from the Atlantic Forest domain. When bat calls cannot be conclusively linked to a particular species due to recording quality or lack of information, detected calls can be divided into call types, or sonotypes. Sonotypes can be used as proxies for species until more robust tools exist to measure the impact ecological stressors such as habitat fragmentation and land use on local bat communities. As this region is a biodiversity hotspot, creating a species ID tool for monitoring bats is important. Understanding the responses of Neotropical bats to land use change is a priority as they fulfil important ecological functions such as pest control, pollination and seed dispersal (Kunz *et al.*, 2011). Currently Neotropical bat research is biased towards a single family, Phyllostomidae (this particular bias can rapidly be reversed using acoustic monitoring rather than mist netting), and major knowledge gaps exist regarding the extent to which anthropogenic disturbance affects bat communities and hence their ability to provide ecosystem services (Meyer, Struebig and Willig, 2015). Bat richness, functional richness and abundance have been shown to be positively related with local forest cover (García-Morales *et al.*, 2016). However, an extensive meta-analysis of Neotropical bat publications found that overall more bat species were present within human-use land uses than in well preserved forests, though responses varied by factors such as family and feeding habit (García-Morales, Badano and Moreno, 2013). Monitoring bats using acoustic methods could help reveal how they use different land uses and further assess their response to habitat fragmentation and land use change.

Here, I use BatDetect, an adaptable, open-source acoustic bat call detection and classification tool that is trained and tested on data from the UK with a graphical user interface (GUI) for quick and efficient visualisation and labelling of audio events to answer three questions: First, whether a CNN pipeline optimised using UK bat data can successfully be applied to data from a new region (Brazil's Atlantic Forest), how many types of call (or sonotype) can be identified from the model outputs, and finally how local land use affects the composition of the bat community. These methods could drive acoustic bat monitoring towards more detailed global datasets on bat biodiversity and ecosystem health.

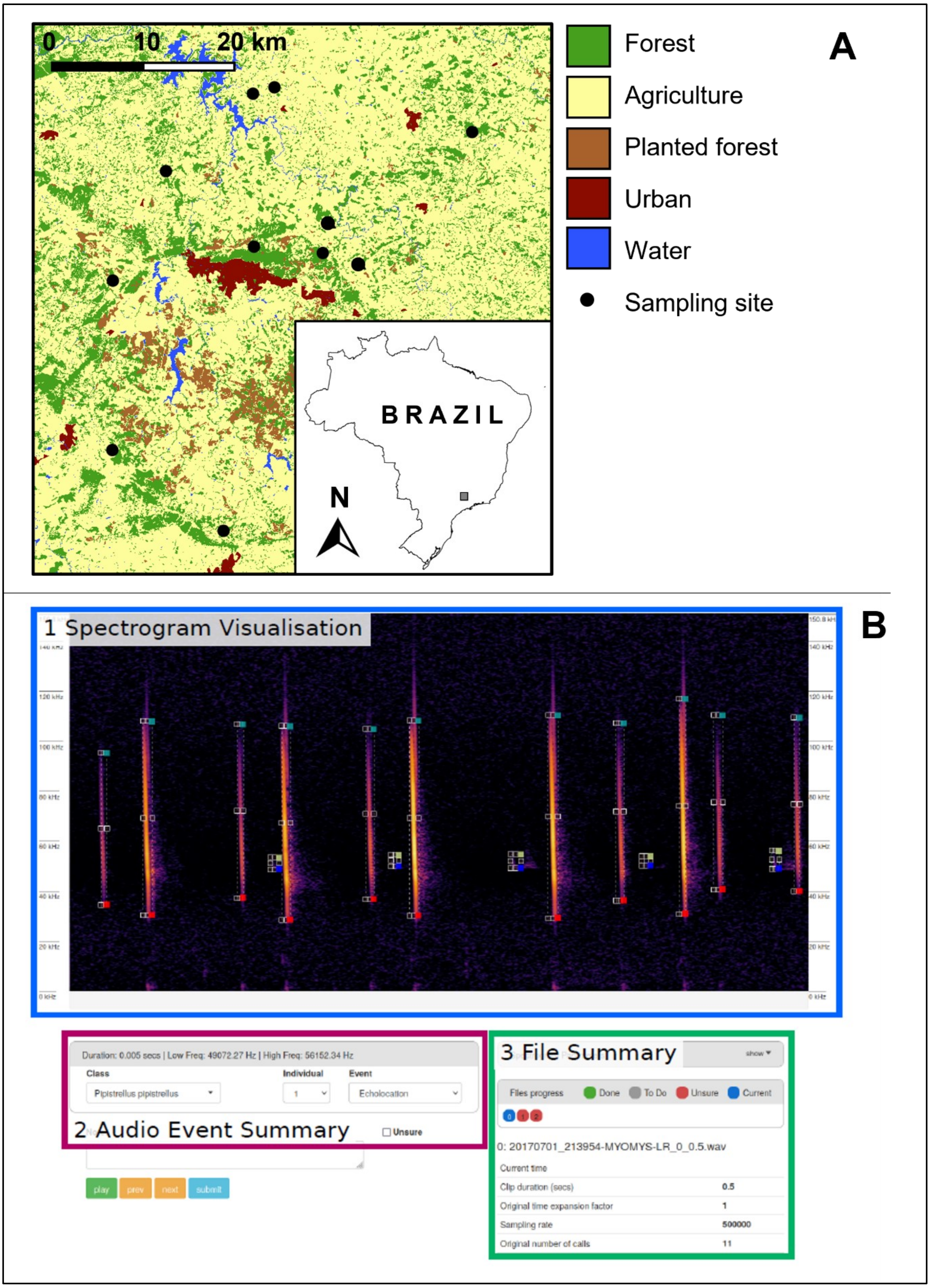
4.3: Methods

4.3.1: Data collection and annotation

I collected acoustic data between January and March 2019 in south-eastern Brazil (Figure 4.1), where Audiomoth (Hill *et al.*, 2019) recorders were set to record a sampling rate of 395 kHz for one minute every five minutes between 22:00 and 04:00. The recorders were deployed on 11 coffee farms, in forest patches adjacent to the coffee, and on the edge between the two habitats. Five recorders were placed at each farm along a transect (two in the coffee: 75 m and 150 m from the edge, one at the edge, and two in the forest: 75 m and 150 m from the edge).

Using the GUI developed as part of the BatDetect suite, a subsample of files (320 10-second files) was fully annotated, where every bat call was labelled to family level. This training set contained recordings from all sites and recorders as well as range of files containing previously identified false positives. This allowed the model to learn to exclude the range of background sounds present at all sites. Based on the identified calls, the labels used were: "Vespertillionidae_unknown", "Molossidae_unknown". This dataset had an additional challenge that within most audio files, a recurring click was present in the ultrasonic range which resembled bat echolocation calls, which had the potential to result in thousands of false positives.

Call labelling was done using a GUI which forms part of the updated Bat Detect suite. For consistency, the boxes were drawn from the highest frequency of the main harmonic to the lowest frequency, excluding any echo. Most calls were part of a pass, which is a series of echolocation calls made by a bat moving past the recorder that start quieter, then get louder as the bat approaches the recorder, and gets quieter as the bat moves further away. Only the main harmonic of each call was annotated. Calls were labelled as part of a pass, and if the best quality calls within the pass (usually in the middle of the pass) could be identified to family level, this was extrapolated to all calls within the pass. In total, 2413 calls were labelled as unknown Molossidae and 5320 calls were labelled as unknown Vespertilionidae in the 320 files.



Duration: 0.005 secs | Low Freq: 49072.27 Hz | High Freq: 56152.34 Hz

Class: Pipistrellus pipistrellus

Individual: 1

Event: Echolocation

☐ Unsure

0: 20170701_213954-MYOMYS-LR_0_0.5.wav

Current time	
Clip duration (secs)	0.5
Original time expansion factor	1
Sampling rate	500000
Original number of calls	11

Figure 4.3: *A: Map of sampling site locations. Audiomoth recorders were deployed at each site in both forest and agricultural land use.* ***B:*** *BatDetect GUI (adapted from Mac Aodha et al. (in prep)). The spectrograms are visualised (1) and the annotator selects the appropriate*

class and event (2) when drawing boxes. An overview of files that have been labelled is displayed in the bottom right (3).

4.3.2: Brazil model training and testing

To train the BatDetect model for this data, the labelled Brazilian dataset was split into a training and test set. The training set consisted of 272 files and the test set had 48 files (15 %). All files were fully annotated. In the training set, there were 2413 calls labelled as unknown Molossidae, and 5320 calls labelled as unknown Vespertilionidae. The number of labelled calls in the training and test sets was 8630 and 1368 respectively. All files were 10 seconds in duration.

Once the model was trained, a detection threshold had to be set before running on all the data. All detected calls have an associated probability score, the higher the score, the more certain the model is that it is a bat call. To select an appropriate threshold, twelve files were manually labelled (containing 425 true positives). The aim was to find the lowest threshold that resulted in 0 false positive detections (100% precision), which was 0.32. At this level, there were 0 false positives, and 91.3% of true positives were detected (see Supplementary Information for detail). The trained model was run with this threshold, so only calls with a probability score above 0.32 were included in the final outputs.

4.3.3: Identifying sonotypes

To identify the number of sonotypes, selected spectral features of all detected calls were clustered using k-means. Within known species, the minimum frequency of bat calls is often the parameter with least variation (Murray, Britzke and Robbins, 2001). This makes it a good variable to estimate the number of sonotypes. The number of clusters should give a good indication of the types of calls within the dataset, which can then be associated to genera or species.

The spectral features selected for clustering were minimum frequency, call duration and frequency of maximum power. Within all detections, unknown Molossidae and unknown Vespertilionidae, five subsamples of 25 000 random detected calls were analysed to find the optimum number of clusters for the k-means clustering. The optimal number of clusters was determined using the Gap statistic. The k-means analysis was run in R using the packages stats and cluster (R Core Team, 2020; Maechler *et al.*, 2022).

4.3.4: Calculating abundance

Using sonotypes as a proxy for species (assuming a different species produces each sonotype detected), abundance for each sonotype at each sampling point was calculated. Abundance was calculated as number of calls of a sonotype detected by a deployed recorder divided by the number of sampling minutes (Lintott *et al.*, 2013). This gave the number of calls per minute at each sampling point for each sonotype.

4.3.5: Measuring the impact of land use

Sonotype abundance was compared between coffee and forest land uses, but only for sites where recordings were captured at 75 m into the coffee and 75 m into the forest (eight out of eleven sites). To determine if land use type affected the abundance of bats, a paired Wilcoxon rank test was done between coffee and forest land uses for each individual sonotype as well as sonotypes grouped by bat family (Molossidae or Vespertilionidae).

The impact of local forest cover, richness and diversity (measured using Shannon's H) were measured against forest cover within 1 km. Bat species have been shown to respond to forest cover at this scale in Mexico (García-Morales *et al.*, 2016), though the most appropriate scale varies between species (Boughey *et al.*, 2011). An ANCOVA was run for Shannon's H against forest cover, grouped by land use. The diversity data met the relevant assumptions (normality of residuals and homogeneity of variance) which were confirmed by non-significant Shapiro-Wilk and Levene's tests. The species richness data did not meet the required assumptions for an ANCOVA, but showed no clear relationship with land use or forest cover. Forest cover calculations were done with QGIS (QGIS Development Team, 2021) using a land use map at a 5 m scale provided by Andrea Larissa Boesing.

4.4 Results

4.4.1: Model performance

The trained model had an average precision of 0.9259 (average precision is the proportion of detections that were true positives) and a per class precision of 0.952 for unknown Molossidae and 0.977 for unknown Vespertilionidae (Figure 4.2).

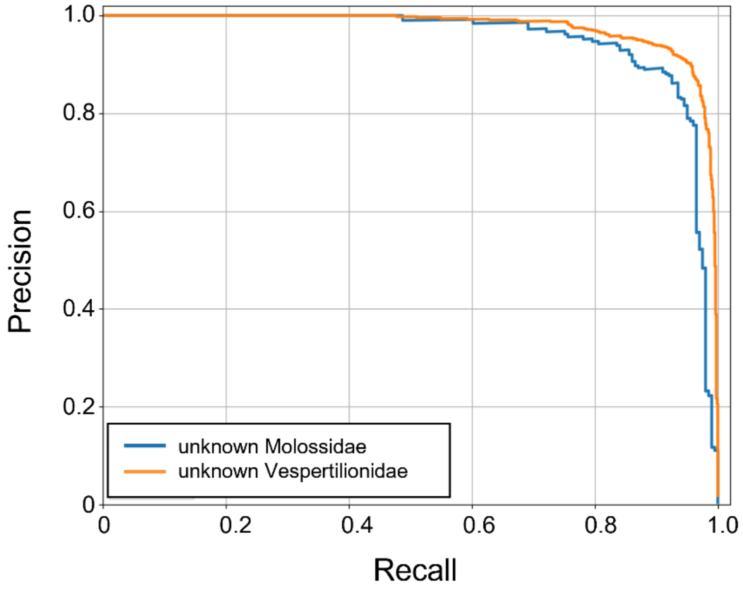


Figure 4.4: *Precision recall curve for Brazil adapted model. The precision is the proportion of correctly classed detections per class and recall is the proportion of calls per class that were detected.*

4.4.2: Brazil model output

In total, 817 377 calls were detected above the selected threshold (0.32) in 18 920 one-minute recordings. 305 353 calls were classified as unknown Vespertilionidae, and 512 024 were classified as unknown Molossidae.

4.4.3: Identifying number of sonotypes

40 recorders successfully captured sound when deployed (19 in coffee, 6 at edges, and 15 in the forest). The number of sonotypes within all detected calls, as well as within unknown Vespertilionidae and unknown Molossidae was calculated using k-means clustering. Within all data, unknown Vespertilionidae, and unknown Molossidae, the k-means clustering showed an optimum of 17, 11 and 6 clusters respectively. As the number of clusters for all data was equal to the sum of the number of clusters in unknown Molossidae and unknown Vespertilionidae, only these were further examined.

Based on exploration of the data and personal communication with Arias-Aguilar, calls could be associated with five genera with the two identified families: *Eumops*, *Myotis*, *Histoitus, Lasiurus* and *Rhogeesa.* Arias-Aguilar is the author of the an extensive database of Brazilian bat species call and call feature summaries (Arias-Aguilar *et al.*, 2018). A wide subset of good-quality data was shared with Arias-Aguilar and they identified calls as far as possible based on their extensive experience and regional knowledge. Based on the duration and frequency of the Molossidae calls, these are likely calls by *Eumops* bat. The shape and duration of V1 may be those of *Rhogeesa,* V6 with *Myotis* and V9 with *Lasiurus*.

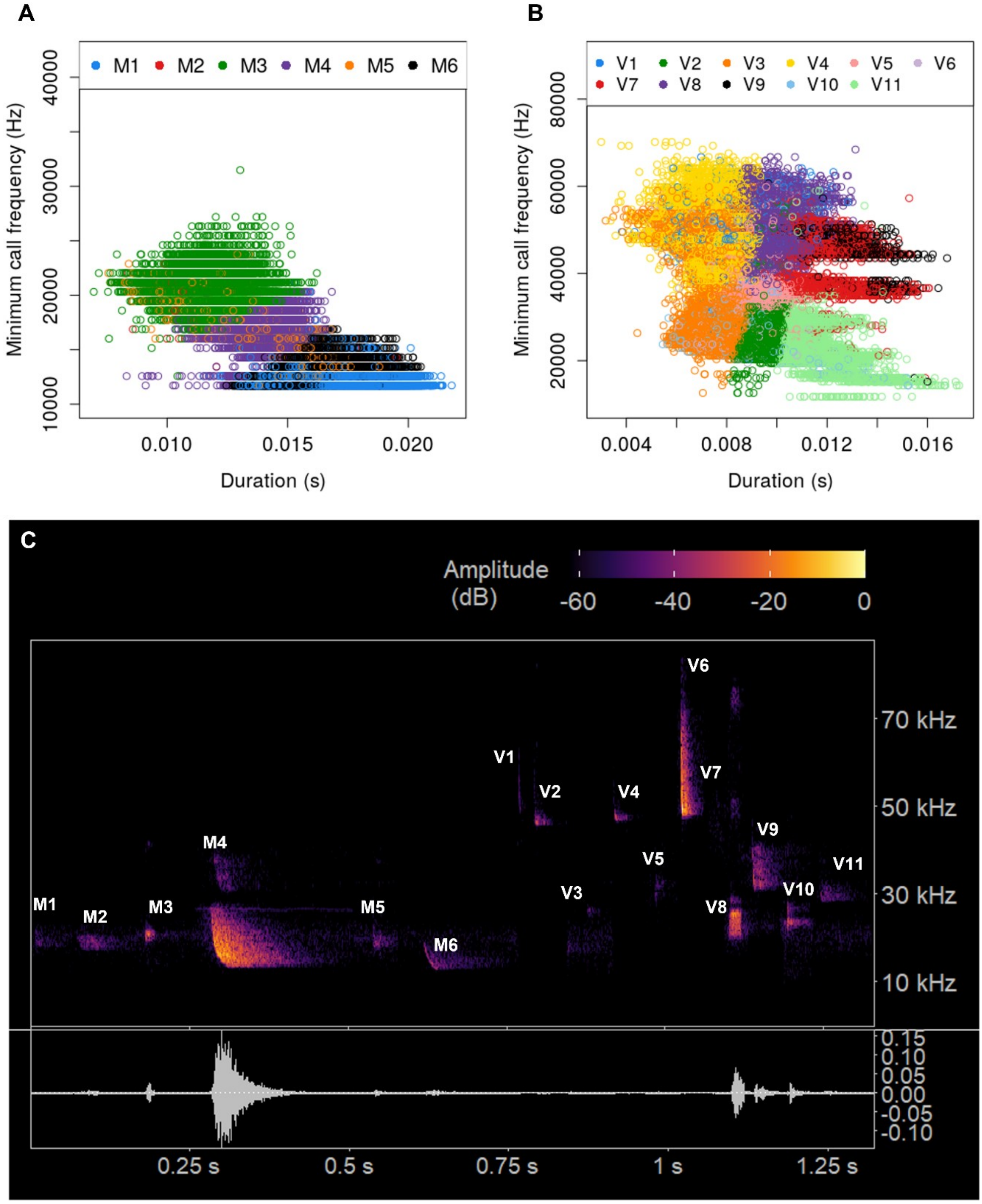


Figure 4.5: A: *Scatterplot showing duration and minimum call frequency of bat calls detected by bat detect and classified as unknown Molossidae. Colour denotes the cluster assigned by the k-means clustering and each represents a distinct sonotype.* ***B:*** *Scatterplot showing duration and minimum call frequency of bat calls detected by bat detect and classified as unknown Vespertilionidae. Colour denotes the cluster assigned by the k-means*

clustering and each represents a distinct sonotype. ***C:*** *Spectrogram plot showing an example of each detected sonotype, of a typical call from within each sonotypes cluster. Certain clusters all had faint calls (e.g. V3, V5 and V7).*

4.4.4: Sonotype response to local forest cover and land use

Neither sonotype richness nor Shannon's diversity showed a relationship with local forest cover (Figure 4.4).

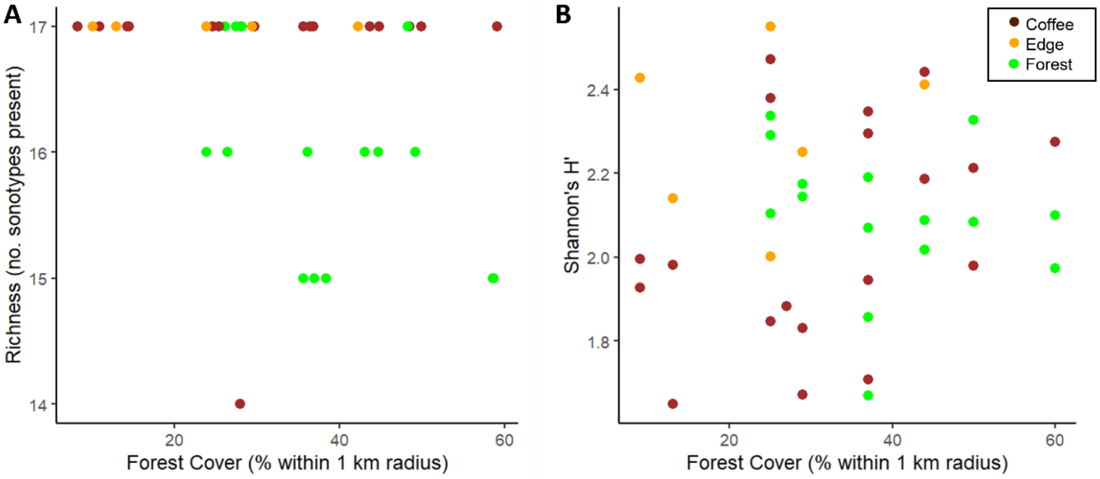


Figure 4.6: *The relationship between local forest cover and detected bat community* ***A:*** *Sonotype richness (number of sonotypes present) across a local forest cover gradient. Points have been slightly jittered along the y-axis for visibility. Colour denotes land use at the recorder (brown = coffee, orange = edge, green = forest).* ***B:*** *Sonotype diversity across local forest cover gradient. No significant trends were detected within land use types or across the whole dataset.*

There was a significant effect of land use on the number of calls per minute for both bat families overall when applying a paired Wilcoxon rank sum test. Across the eight sites with coffee and forest pairs, the Molossidae sonotypes had significantly more calls per minute in the coffee land use than the forest, with an effect size (r) of 0.529, p = 0.0126 (Figure 4.5A). Vespertilionidae sonotypes also had significantly more calls per minute in the coffee land use than the forest, with an effect size (r) of 0.271, p = 0.0109 (Figure 4.5B). When comparing the effect of land use for each sonotype, four sonotypes (M1, M4, M6 and V2) had significantly higher call abundance in the coffee than the forest (Figure 4.5C).

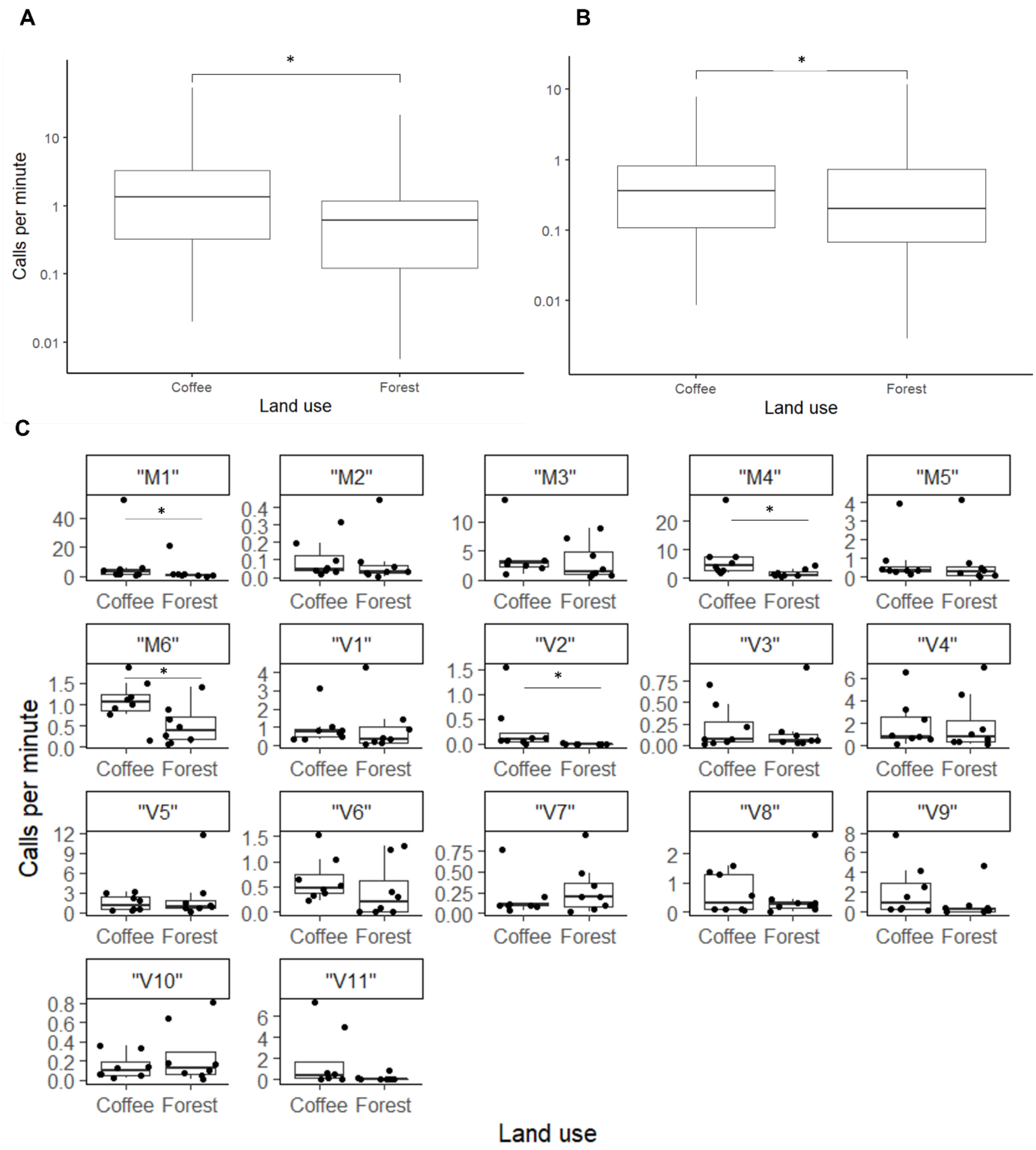


Figure 4.7*: The response of bat sonotypes to land use, irrespective of local forest cover.* ***A:*** *The difference in calls per minute for Molossidae sonotypes in coffee and forest (r= 0.529, p=0.0126*, paired Wilcoxon rank sum test).* ***B:*** *The difference in calls per minute for Vespertilionidae sonotypes in coffee and forest (r=0.271, p=0.0109**, paired Wilcoxon rank sum test).* ***C:*** *The difference in calls per minute for each individual sonotype. Four sonotypes had significantly more calls per minute in the coffee than the forest: M1, M4, M6 and V2 with*

paired Wilcoxon rank sum effect sizes of 0.74, 0.74, 0.69 and 0.55, and p-values of 0.039, 0.039, 0.05 and 0.015 respectively.

4.5: Discussion

The results show that the retrained model can successfully be applied to a new geographical region, with an overall precision of (0.9259) and recall across all classes (class AP 0.9645). The results also show that 17 distinct sonotypes can be separated and used to calculated relative abundance. Bat call abundance was higher overall in coffee land use than forest land use. However, sonotype diversity was not significantly related to local forest cover across land uses or overall. This demonstrates that this approach can be scaled to create a call library for new regions, with the potential to develop a species ID tool.

The model's high average precision of 0.9259 highlights its ability to correctly detect most calls present within this dataset. This result is particularly good given the challenging dataset that contained clicks that had the potential to be detected as bat calls. Compared to other available bat detection tools, the BatDetect model outperforms other available tools such as Tadarida (Mac Aodha *et al.*, no date; Bas, Bas and Julien, 2017), though it is difficult to make accurate comparisons as currently other tools tend to output bat detections, and not classified detections based on a labelled subset. The high model precision can be partly attributed to thorough and extensive labelling of the training dataset. Labelled training data is the main barrier to developing species identification tools and thus collecting bat call data alongside species verification in the field, automated species identification tools can be kickstarted. The number of calls detected shows the potential for rapid data collection once a model has been trained. However, the species verification work still represents a significant hurdle.

The k-means clustering analysis showed 17 sonotypes present. A single species could not be attributed to each sonotype due to lack of libraries for the region, though expert verification strongly suggested the presence of five genera: *Eumops*, *Myotis*, *Histoitus, Lasiurus* and *Rhogeesa*. This is a problem for south and east Brazil as even recent acoustic bat surveys use tools that have not been specifically trained for the target region (Pereira, Falcão and Bernard, 2022), maintaining a reliance on sonotypes and acoustic indices (de Aguiar Silva *et al.*, 2022). This highlights the gap in bat call knowledge for many Brazilian species. For North America and Europe, researchers have a choice of commercial species classifiers available (Nocera *et al.*, 2019), which in not the case for South America. Though the k-means approach managed to broadly detect sonotypes, this remains an estimate due to the number of similar calls as the data set is biased towards quieter calls as bats approach the recorder. It may also be the case that the sonotype groupings separate different phases of an approach from the same species as the shape and volume of the call change throughout the pass. However, clustering broadly allowed for local richness and abundance to be calculated as well as an analysis of how the sonotypes present responded to local ecological factors.

Land use had a significant effect on the number of calls per minute for both Molossidae and Vespertilionidae groups as well as for 4 of the 17 sonotypes, with higher abundance in the coffee compared to the forest. This is consistent with observations from other studies, which showed that insectivorous bats prefer open habitats and may not enter forested areas with dense understories as vegetational complexity may restrict flight (Suarez-Rubio, Ille and Bruckner, 2018; Laurindo *et al.*, 2020). As most sonotypes were detected at each sampling site, richness was not affected by land use or forest cover variables. This lack of difference in richness implies that the bat communities in this study region are relatively homogenous due to historical habitat loss and extensive fragmentation. Homogenisation of communities has been observed across other taxa in the Atlantic Forest including epiphytes and birds, showing the importance of preserving large areas of unfragmented forest (Hatfield *et al.*, 2020; Parra-Sanchez and Banks-Leite, 2020).

There were a few limitations to this work that can be improved upon for future studies. Much of the information on Brazilian bat acoustics is from calls recorded in the north of the country, or even other South American countries (Arias-Aguilar *et al.*, 2018), where typical species calls are described based on a single pass. Bat species are known to alter their calls depending on the habitat they are traversing, which can be source of misidentification (Pedro and Simonetti, 2014). Intraspecific acoustic variation is also present in different geographic communities (Farrell, Corben and Gannon, 2000), but not enough is known about how this affects Brazilian bat species, and would require rigorous acoustic surveys coupled to capture studies. Some bat calls had a prominent harmonic above the main call. The labelling protocol I helped develop did not label these as part of the call, but some were still detected by the model. Testing the impact of labelling entire calls with the harmonic on model performance should be done. The bat detection model retained little information on call shape or direction, so calls within the same frequency range and of similar duration but from different species were not distinguishable.

This study shows that adapting a CNN with a relatively small additional dataset can detect bat calls to a high degree of precision. Even without a call library, estimates of diversity were made and used to measure responses to ecological stressors. However, the detected sonotypes appeared to be distributed across all sampling sites, implying that the bat community in this region is homogenous which may be a result of decades of habitat fragmentation. Automated bat species monitoring with acoustic tools can be achieved in more biodiverse regions such as the Atlantic Forest with the creation of a high-quality call library.

Chapter 5: Sun coffee plantations provide weak microclimatic buffering effects but not against supraoptimal temperatures.

5.1: Abstract

Coffee (*Coffea arabica*), one of the most consumed global commodities, is threatened by climate change, as it has a narrow optimal thermal niche, and is particularly vulnerable to high temperature exposure which causes physiological stress. Though traditionally farmed as a shade crop, economic pressure has driven many producers to intensify their management to entirely sun-exposed coffee. However, it is possible that higher levels of surrounding forest cover could provide sufficient microclimatic buffering (defined here as the extent to which local temperature differs from regional temperature) to protect crops from sun-exposure. Here, I show, using low-cost sensors collecting hourly soil and air temperature data from ten Nicaraguan sun-coffee farms, that microclimate buffering causes lower mean and minimum temperatures, but none against maximum temperatures. The degree of microclimate buffering is independent of altitude and the level of local forest cover (defined as % forest within 1000 m). The lack of a relationship between the degree of buffering and the levels of forest cover or altitude strongly suggests that mitigation strategies are necessary to protect coffee from rising temperatures. The outlook for Nicaraguan sun-coffee is uncertain, and that adaptation strategies are essential to maintain production in this region.

5.2: Introduction

Coffee is one of the most consumed global commodities, with over 3 billion cups being drunk per day (USDA, 2020). Demand continues to grow, having recently outpaced production, causing a decline in global coffee stock (Vegro and de Almeida, 2020) . *Coffea arabica,* the more popular variety of coffee due to its higher quality, represents the majority of the global market share, and is vulnerable to climate change as it is primarily grown in the high elevation areas of the tropics (Bunn, Läderach, Rivera, *et al.*, 2015), with Brazil, Colombia, Ethiopia and other countries of central and south America being the most productive (seven of the top ten producers are in Central and South America (USDA, 2020)). Though traditionally farmed as part of an agroforest system, high demand has led to intensification and many farmers now grow sun-exposed coffee (Jha *et al.*, 2014). To ensure coffee production remains sustainable, predicting the effect of climate change on different farming strategies in relation to their local geography and biodiversity is essential (Coltri *et al.*, 2019; Ennis and Philpott, 2019).

C. arabica has a narrow optimal climatic niche, requiring mean annual temperatures of around 20°C but no lower than 18°C and no higher than 24°C, and around 1200mm of annual rainfall (DaMatta and Ramalho, 2006). Extended periods of over 30°C and frost exposure can be particularly damaging to coffee crops (DaMatta, 2004). High altitude areas of the tropics meet these criteria, with Central American countries being particularly suitable, having cultivated coffee for over 200 years (Staver *et al.*, 2001). However, global temperatures have already risen by approximately 1°C relative to pre-industrial times (Haustein *et al.*, 2017), with an increase of at least 1.5 degrees by 2050 predicted and higher increases possible (IPCC, 2018), lowering the suitability of this area for coffee production (Läderach *et al.*, 2017). Increased variations in rainfall and temperature in the 21st century, particularly by the El Niño-La Niña cycle, have reflected in a high fluctuation in coffee yields, (Baker and Haggar, 2007). Predicting the impacts of climate change on coffee agriculture is difficult as climate change also appears to act synergistically with other threats to coffee production, such as increased pest risk and outbreak intensity (Jaramillo *et al.*, 2009), higher disease risk (Avelino *et al.*, 2015) and lower pollinator diversity (Imbach *et al.*, 2017). To lessen the severity of climate change on coffee agriculture, better understanding the effect of different management and local environmental variables on microclimatic buffers is necessary.

Predictions about the future of coffee using species distribution modelling have mainly found that most countries currently producing coffee will lose suitable areas (Laderach *et al.*, 2011; Bunn, Läderach, Rivera, *et al.*, 2015; Ovalle-Rivera *et al.*, 2015a). However, commonly used climatic data sets used in species distribution modelling have major discrepancies with satellite data or data collected in situ within tropical agricultural landscapes (Faye *et al.*, 2014; Deblauwe *et al.*, 2016), as local conditions can cause localised microclimatic buffering (Frey *et al.*, 2016). Microclimatic buffering can be defined as the difference in temperature between two sites measured by the slope of the regression between a pair of sites (Ewers and Banks-Leite, 2013). For instance, if the temperature in one location increases by 1°C and temperature at the second location only increases by 0.1°C, this represents a 0.9°C buffer. Microclimatic buffering is regularly seen in forests, which are usually cooler than temperatures in open land use given the protective shade cover from canopies (Ewers and Banks-Leite, 2013), both above ground and in the soil. Quantifying microclimate is becoming much cheaper with low-cost data loggers and weather stations (Climate Edge, 2021). Within agroforest coffee systems, farmers can create microclimatic buffers by planting coffee in the shade of other trees to mitigate temperature extremes (Lin, 2007), but the extent to which sun-coffee monocultures are buffered is not known.

Forests can only act as thermal buffers if they are intact (De Frenne *et al.*, 2019) and as forests become more fragmented, their buffering ability is reduced (Ewers and Banks-Leite, 2013; Lin *et al.*, 2020; Mendes and Prevedello, 2020). Croplands adjacent to forests experience higher temperatures than the forest interior, as they lack shade protection (Pinto *et al.*, 2010; Senior *et al.*, 2017). Trees may not only reduce maximum temperatures as they can also protect against high windspeeds, which may benefit forest-adjacent agriculture (Pezzopane *et al.*, 2011). Sun-coffee planted in an area of high local forest cover may therefore benefit from lower wind speeds and cooler adjacent regions. However, these studies do not specifically examine how far buffering may extend over forest edges to different land uses and how different levels of surrounding forest cover could affect this relationship.

Elevation plays an important role in climate as air pressure decreases at higher elevations (Jacob, 1999). In the context of climate change, higher elevation regions have been warming at higher rates in subtropical regions (Dong *et al.*, 2014). Predictions of future habitat suitability for growing coffee often show that the minimum lower elevation boundary will shift upwards (Ovalle-Rivera *et al.*, 2015a). For instance, in Nicaragua the lowest suitable elevation for coffee is predicted to increase from 800 m to 1200 m by 2050 (Laderach *et al.*, 2011). Elevation influences humidity and light intensity (Karungi *et al.*, 2018), which also contribute to the local microclimate. In addition, the elevational topography of a region is important, as more complex topography results in wider heterogeneity of microclimate (Opedal, Armbruster and Graae, 2014). Examining the impact of elevation on microclimatic buffering has not been fully investigated and may help to identify elevations that mitigate more against temperature extremes.

Here, I investigate two questions: firstly, whether there is evidence for microclimatic buffering on sun-coffee farms, and secondly how the extent of this buffering varies across a local forest cover between 28% and 78% and an elevational gradient covering a range of 600 m. I expect weak thermal buffering against daily maximum temperatures on sun-coffee farms given the lack of shade trees. However, I also expect that higher elevations and higher levels of forest cover should confer more protection from higher temperatures. For this case study, weather stations were deployed on ten sun-grown coffee farms in the region of Jinotega, Nicaragua, Central America. Nicaragua relies heavily on coffee exports, with coffee representing approximately 7.2% of total GDP (Flores *et al.*, 2002). It is at particular risk of loss of coffee suitability with climate change (Gourdji *et al.*, 2015), as it is affected by the El Nino-southern oscillation (Almeida Silva *et al.*, 2020) and Nicaragua is ranked 4th on the Global Climate Risk Index (Kreft *et al.*, 2013). The northern Nicaragua region of Jinotega is currently one of the country's most productive coffee producing areas (UCA Soppexcca, 2011), thus this is a highly relevant area to study the impact of microclimatic buffering.

5.3: Methods

Data Collection

I used hourly air and soil temperature data collected from Climate Edge® (Climate Edge, 2021) weather stations deployed on ten coffee farms ranging in elevation from 504 to 1101m above sea level in Jinotega, Nicaragua (Figure 5.1) between May and September 2016 (Table S1). Climate Edge is a start-up that set up weather stations at farms with multiple sensors to provide customised management recommendations (Climate Edge, 2021). Temperatures were recorded for a mean of 125 days per farm during the rainy season (Figure 5.3). These weather stations consisted of two DS18B20 (Maxim Integrated Products, 2019) sensors (one for soil temperature and one for air temperature, measured 1 m above ground), accurate to ±0.5ºC. They were calibrated against Lascar temperature data loggers (Lascar Electronics, Salisbury, UK) to ensure consistency between stations. These sensors are relatively cheap and can therefore be deployed widely. Information concerning precise locations of the weather stations on the coffee farms relative to the crops was not available and location data was based on farm addresses and estimations based on conversations with members of Climate Edge.

Microclimatic buffering is a relative measure, for example, forests buffer against high temperatures *relative* to adjacent unforested areas (Ewers and Banks-Leite, 2013). Therefore, a different dataset from the same region, but not from a coffee farm was required. I obtained daily mean, maximum and minimum air temperature data from a regional weather station (sometimes the data were collected every two days), station ID number 787340 (National Climatic Data Center *et al.*, 2017), located on the outer boundary of the city of Jinotega at 985 m altitude. The daily temperature was given to the nearest 0.1 ºC, but maximum and minimum were only given to the nearest integer. This weather station is referred to as the regional station to distinguish it from the *in situ* weather stations at the coffee farms.

I calculated the percentage of forest cover within a 1000 m radius of the coffee farm weather stations using 30 m resolution tree cover maps (Hansen *et al.*, 2013), recalculated to 2015 levels of forest cover by including subsequent maps of gain and loss onto the initial forest cover map from 2000. Metadata on the forest quality was not available. A 1000 m buffer was chosen as previous studies examining microclimatic variation have used landscape variables at this scale (Fridley, 2009). The elevation data was obtained from 1-arc second resolution

SRTM data from USGS (USGS, 2019). Mapping analysis was done using QGIS (QGIS Development Team, 2021).

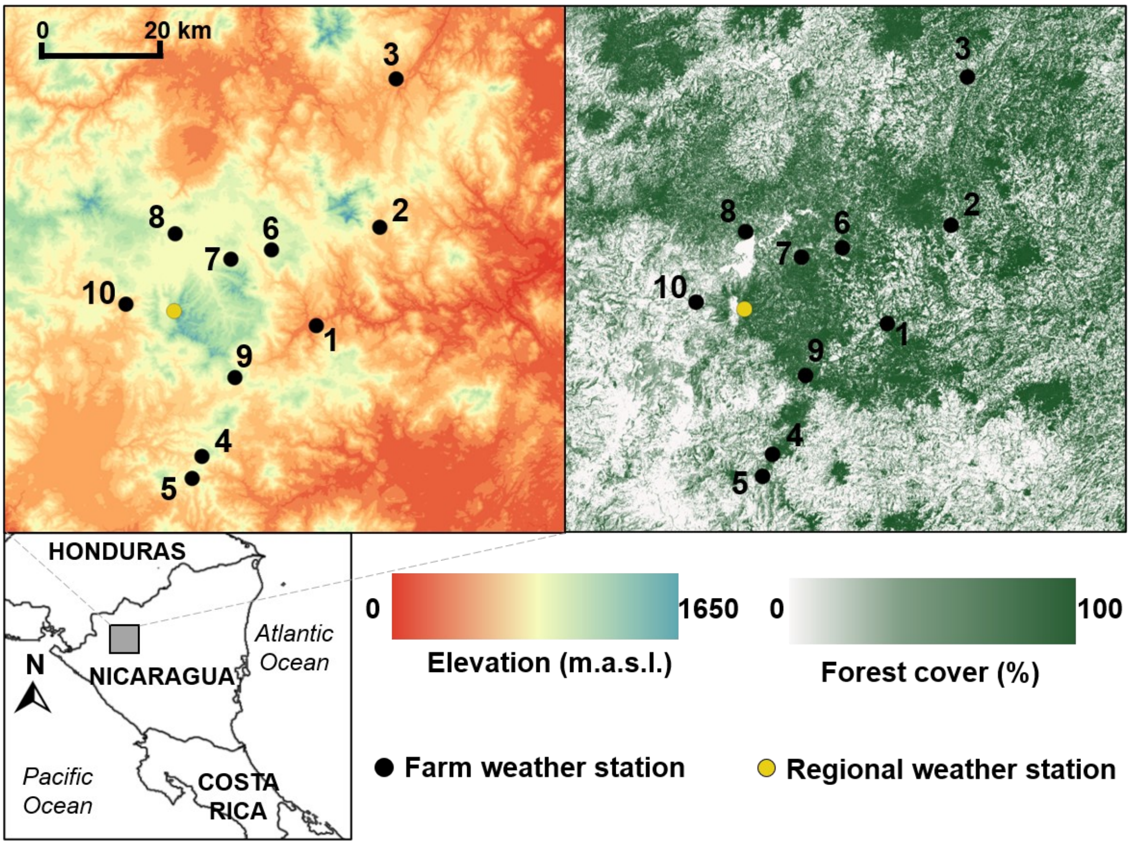


Figure 5.8*: Distribution of coffee plantation sampling locations in Jinotega, Nicaragua and the two environmental variables examined (elevation and forest cover)*. *Black dots denote the farm weather stations, and the yellow dot denotes the location of the regional weather station.* ***A:*** *Elevation map (USGS, 2019)* ***B:*** *Local forest cover* (Hansen *et al.*, 2013).

Statistical analyses

Overall microclimatic buffering

To assess whether there was any overall buffering across farms, I first calculated the daily mean, maximum and minimum temperatures for each farm for the entire period of recording for both the above ground and soil temperatures. This could then be directly compared to the daily mean, maximum and minimum temperature obtained from the regional weather station. For each group of data (mean, maximum and minimum for above ground and soil temperature respectively), temperature data from the farm weather stations were regressed against the regional weather station data, and the slope of the linear regression was taken to represent the buffering effect. In each model, the data collected at the farm was the dependent variable and the regional weather station data was the independent variable. The data met the required assumptions for a simple linear regression. A buffering effect of 1 means there was no difference between the local and regional data, whereas a buffering effect below one shows a cooling effect at the farm, and a buffering effect above 1 represents that farm temperatures were higher than the regional weather station temperatures. Figure 5.2 shows a graphical description of the methodology. To calculate the buffering in soil, the same process was repeated, but as soil temperature data was not available from the public station, the buffering in soil was calculated relative to air temperatures. The methodology from Ewers & Banks-Leite (2013) was followed. All data analysis was performed in R (R Core Team, 2020), with the application of the dplyr package (Wickham *et al.*, 2020).

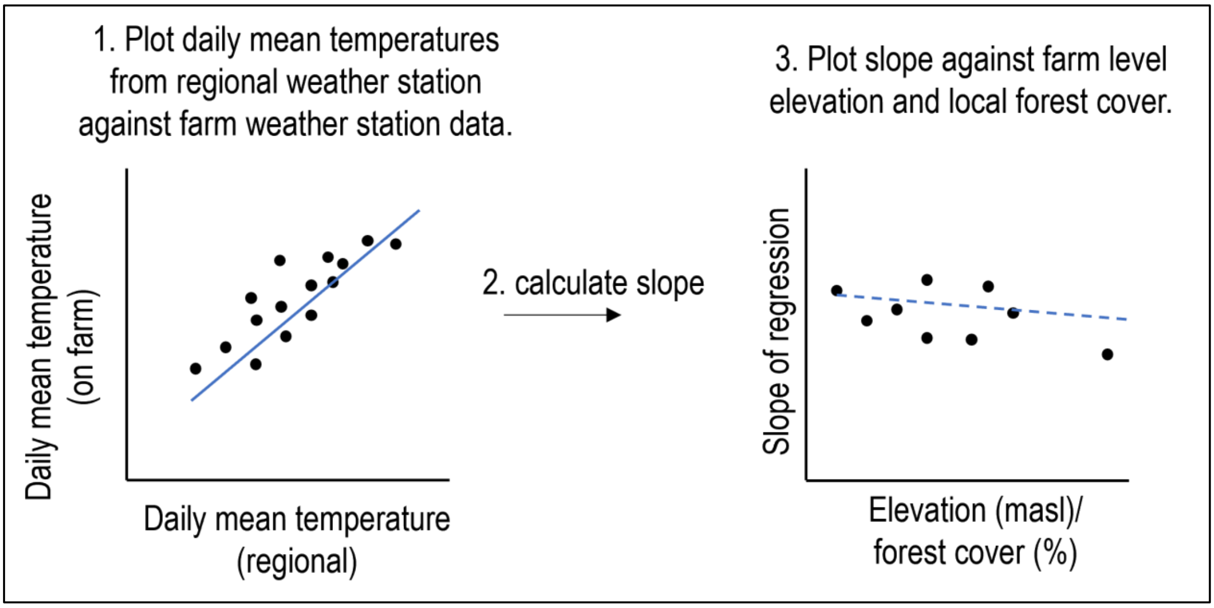


Figure 5.9: *Diagram showing methods process to obtain the buffering data.*

The effect of elevation and forest cover on microclimatic buffering

To assess whether the extent of buffering varied with either local forest cover or the elevation of the farm, the previously calculated coefficients of regression at each farm were regressed using a multiple linear regression against both elevation (in m) and local forest cover (as a percentage of forest cover within a 1000 m radius) using the “lm” function in base R (R Core Team, 2020). This was also calculated independently for above ground and soil mean, maximum, and minimum temperatures according to the formula: buffer ~ elevation + forest cover.

To format the dates to create Figure 5.3A, the package lubridate was used (Grolemund and Wickham, 2011). The package pracma was used to create the boxplots in Figures 5.2B and 5.2C (Borchers, 2021). All code is available at: https://github.com/emdo1/Microclimate.

5.4: Results

Overall microclimatic buffering

The mean daily temperature across all farms for the duration of the study was 22.06 °C above ground, 23.41°C in the soil, and 23.86 °C at the regional weather station (Figure 5.3 A and B). Daily maximum air temperatures reached in excess of 30°C for an average of 27.6 days per farm (range from 0 to 86 days). The mean of regional temperature data daily maximum temperatures was 26.88 °C (s.d. 1.79 °C) and the mean of the regional daily minimum temperatures was 20.08 °C (s.d. 0.89 °C).

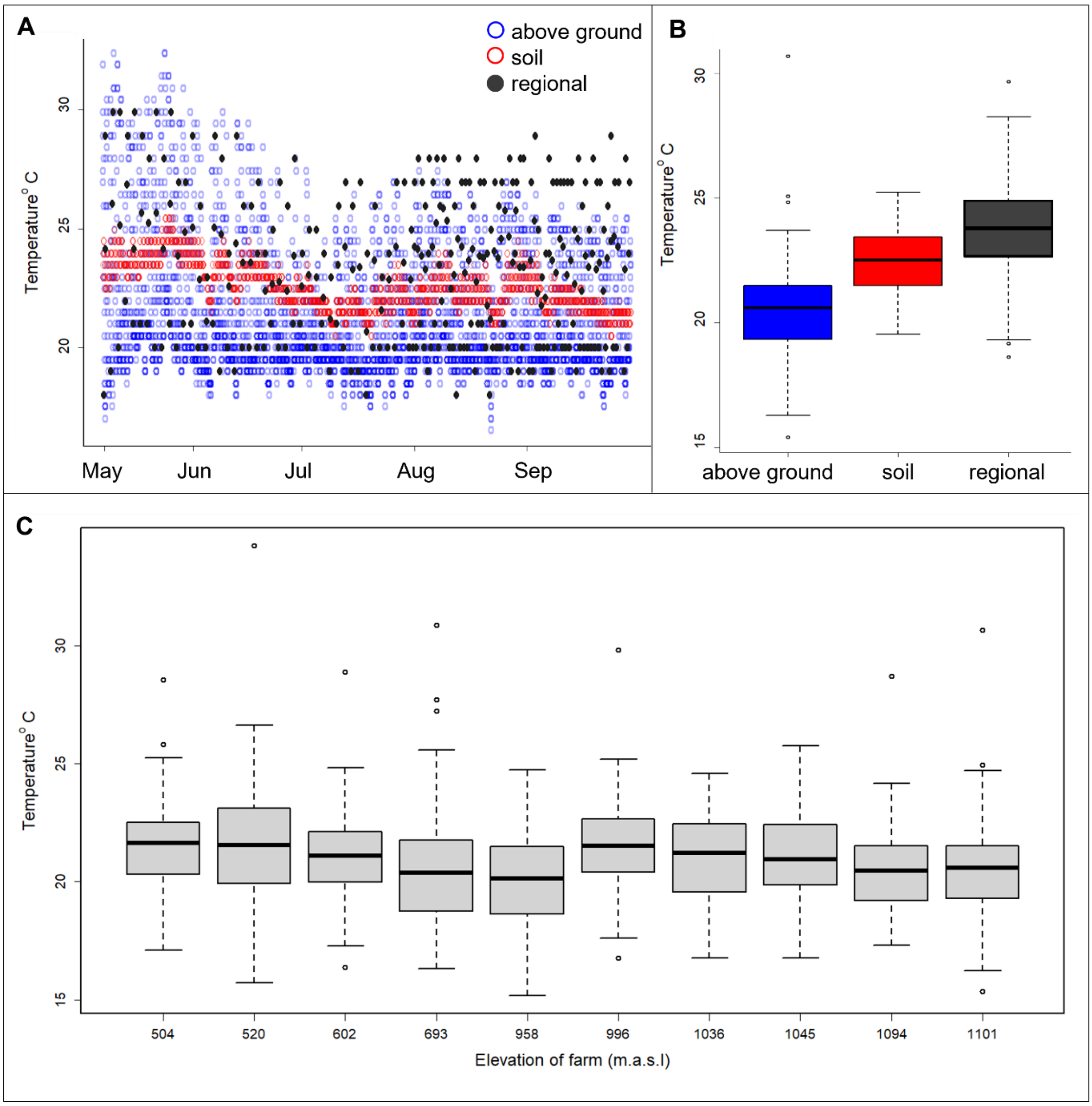


Figure 5.10: *Summaries of temperature data collected across the study period, with a detailed example from a single farm as well as boxplots of above ground temperatures from all farms.* ***A:*** *Temperature data collected at farm 4 (elevation 1101 m), where the scatterplot shows hourly above ground temperature (blue), hourly soil temperature (red) and daily mean, max and minimum data from regional weather (grey).* ***B:*** *Boxplots of above ground, soil and regional temperatures from farm 4, where the bold line represents the median temperature over the study period, the box represents the interquartile range (IQR) and the whisker plots represent are 1.5 times the IQR and the dots represent outliers.* ***C:*** *Boxplots of above ground temperature collected at each farm, which are ranked and denoted by their elevations the bold line is the median buffering effect, the grey box is the IQR, the whiskers are 1.5 times the IQR and the circles are outliers.*

An overall buffering effect of the farms compared to the regional weather station was observed for both mean and minimum above ground temperatures (Figure 5.4A), but there was no significant buffering against maximum temperatures. If no buffering were observed, the regression coefficients would be close to 1. The closer the coefficient is to zero, the stronger the buffering effect. At several farms, there was even negative buffering, where the slope of maximum local temperature was more than 1 (this means if the regional maximum

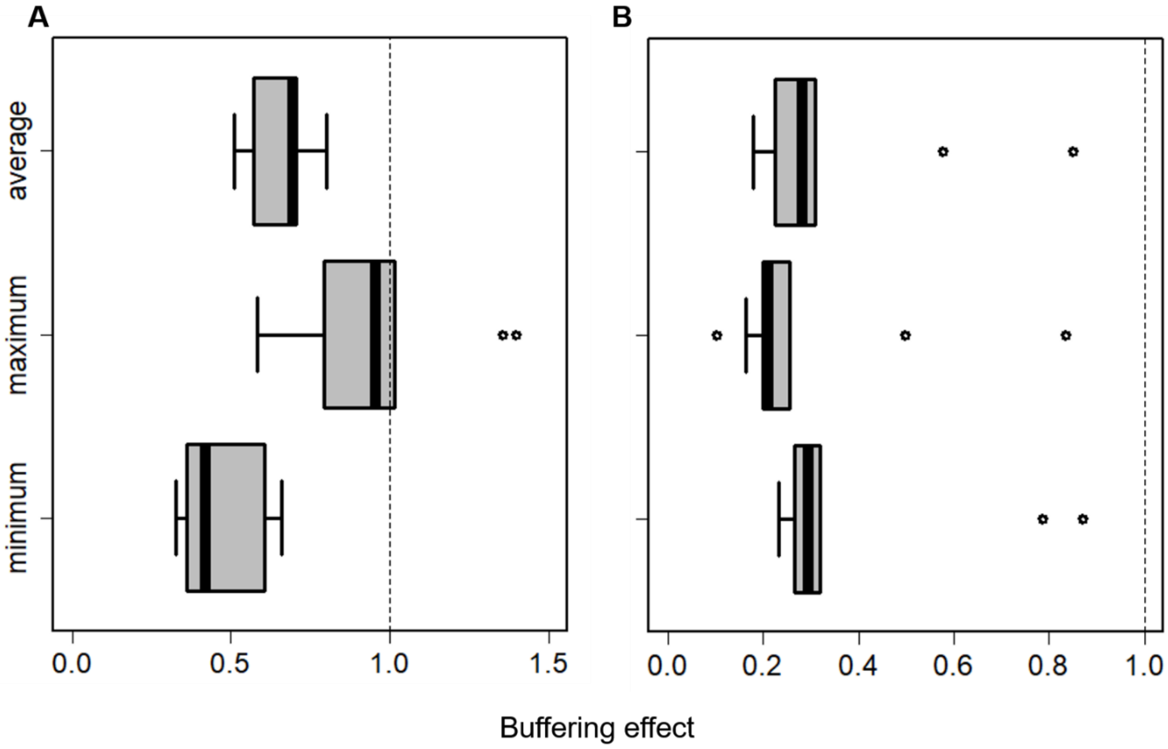


Figure 5.4: *Boxplots of the microclimatic buffers calculated from regressions of farm temperatures against the regional weather station temperatures. The dotted lines at 1.0 represent a regression coefficient of 1, which represents zero buffering effect. The closer the buffer is to zero, the stronger the buffering effect.* ***A:*** *Buffers calculated for daily mean, maximum and minimum at 1 m above ground temperature.* ***B*** *Buffers calculated for daily mean, maximum and minimum for soil temperature. The black line is the median buffering effect, the grey box is the interquartile range (IQR), the whiskers are 1.5 times the IQR and the circles are outliers*

increased by 1°C, the local temperature increased by up to 1.4°C. All soil temperatures showed clear buffering effects (Figure 5.4B) against ambient air temperatures recorded from the regional weather station.

The effect of elevation and forest cover on microclimatic buffering

I found a significant relationship between the extent of buffering of minimum air temperatures and elevation (Figure 5.5B, p = 0.032), but found no other statistically significant relationships between buffering of the air or soil temperatures on farms compared

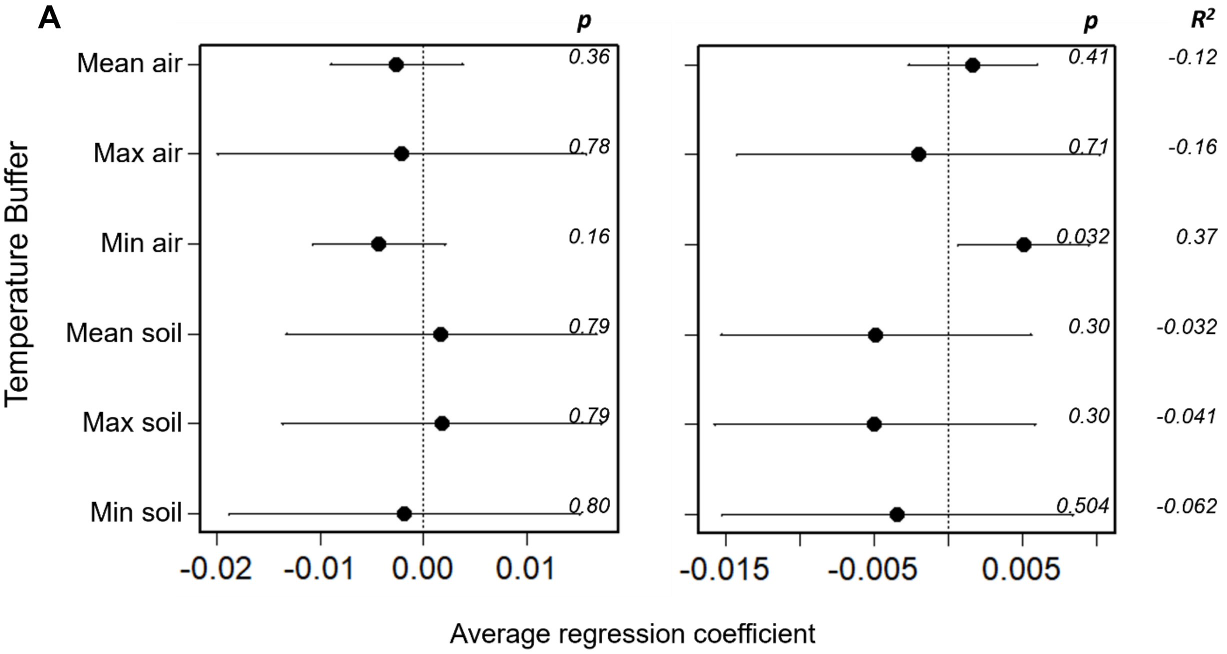


Figure 5.5: *Summaries of multiple linear regressions (buffer~ Forest cover+ Elevation) investigating the impact of forest cover and elevation on the extent of buffering. The dots denote the regression coefficient of the multiple linear regression and the lines denote the 95% confidence interval; the corresponding p-values and adjusted* R^2 *for each model is displayed on the right-hand side.* ***A:*** *Regression coefficients and associated confidence intervals for forest cover and buffering. No significant trends were observed.* ***B:*** *Regression coefficients and associated confidence intervals for elevation and buffering. A significant trend was observed between the minimum air temperature buffer and elevation (p=0.032).*

to the regional weather station temperatures and the elevation or the percentage of forest cover at the farms (Figure 5.5).

5.5: Discussion

I found that the current conditions in sun-coffee do offer some level of microclimatic buffering for both mean and minimum temperatures but provides no mitigation against high peaks in temperature during the day. The extent of local forest cover and elevation did not have any effect on the level of buffering compared to the regional weather station, either above ground or in soil. A significant trend was observed between minimum temperature and elevation, where lower elevation sites had a stronger buffering effect for minimum temperatures. Nonetheless, the overall pattern of results suggests that the higher elevation sun-coffee farms and those surrounded by higher levels of forest cover are just as vulnerable to exposure to temperature extremes as lower elevation sun-coffee farms in deforested areas.

Overall microclimatic buffering

The most important finding is the lack of any thermal protection against supraoptimal temperatures; on two of the farms, there was even a negative buffering effect (Figure 5.4A). High temperatures pose one of the greatest threats to coffee, as long-term exposure results in depressed growth and development (Haggar and Schepp, 2012). Without changes in management, it is unlikely that these regions will be as productive for coffee in the long-term, as relative to other coffee growing regions, Nicaragua is considered relatively hot and dry (Bunn, Läderach, Jimenez, *et al.*, 2015). It has been predicted that approximately 50% shade cover will be necessary at altitudes around 1000 m.a.s.l. (which is considered “low elevation” for coffee) to maintain the suitability of these areas for this crop (Rahn *et al.*, 2018).

The effect of elevation and forest cover

The results show that elevation had a significant effect on the extent of buffering for minimum temperatures (*p=0.032*), where lower elevation sites had a stronger buffering effect for minimum temperatures. This suggests that lower elevations may buffer against increases in minimum temperatures. Increases in minimum temperatures have been linked to wider susceptibility to coffee rust in Central America (Avelino *et al.*, 2015), resulting in outbreaks at higher altitude, which this result corroborates if higher altitudes are less able to protect against higher minimums. Though no previous research has examined the effect of elevation on the extent of buffering, Karungi et al. (2018) found an inverse relationship between light intensity and humidity across all types of management, which varied across elevations: mid-level elevations (1500-1679 m.a.s.l.) had low humidity and high light intensity, whereas low and high elevations had the reverse. However, the lowest elevation farms in their study exceed the altitude of the highest elevation farm in our study, so the relationships between humidity, light-intensity, buffering and elevation may vary considerably and warrant further investigation. For coffee growing regions with a climate like that of Nicaragua, the median elevation for *C. arabica* is predicted to increase from ~800 m to ~1200 m by 2050 (Bunn, Läderach, Jimenez, *et al.*, 2015), and as elevation has not been shown to affect the extent of microclimatic buffering on mean or maximum temperatures, changes in management are essential to ensure sustained production.

The lack of relationship against local forest cover shows that though forests are able to create their own microclimate, this has no effect on adjacent non-forested land uses. Even if the farm is surrounded by over 80% forest, there is no additional microclimatic buffering than at a farm without any surrounding forest. This demonstrates that buffering within forest interiors is largely due to shade and wind protection. Coffee with high levels of shade cover has been found to be cooler with higher humidity both above ground and in soil (Lin, 2010).

This is because the presence of protective shade trees reduces light intensity, reduces wind speeds (Pezzopane *et al.*, 2011), and limits water loss in the soil (de Carvalho *et al.*, 2021), which is important for soil macrofauna (Prayogo *et al.*, 2019). These microclimate studies have also found other variables such as light intensity, humidity and topography, to be impactful. Having more data on light intensity, humidity and hill shading may have revealed trends between these factors and local elevation, as the topographic layout of a farm and how many hours of sunlight the crop is exposed to has an impact on microclimate.

Soil temperature buffering

The mean soil temperature across all farms was higher than the above ground air temperature, which is unexpected. No relationship between soil buffering and either elevation or forest cover was detected. However, the lack of an equivalent below ground temperature data set for direct comparison means drawing strong conclusions from this part of the study is premature, particularly as other studies found below-ground buffering effects to be strongest (Ewers and Banks-Leite, 2013).

Limitations of this study:

Metadata on the location of the weather stations at each farm was lacking. For instance, the coordinates of the exact location of each weather station were not recorded. This is important as it affects the reliability of the elevation data, as coffee farms are often located on hillsides. It was assumed that the stations were located adjacent to the coffee growing the closest to the farmhouse for accessibility reasons. Additionally, a lack of data concerning topography and light intensity means that the role of topographic shading, which plays an important role in microclimatic buffering (Ashcroft, 2010), could not be explored. However, current versions of the Climate Edge weather stations have sensors to measure light intensity, humidity, dew point and windspeed, which will allow a more complete analysis of these variables to be performed in future. There were gaps in the regional weather station data, as it sometimes only recorded temperature every second day. In addition, having higher temporal resolution from the regional data would have allowed daily patterns of temperature to be more fully explored, as for example, the minimum temperature at night has been found to be a significant related to yield (Craparo *et al.*, 2015). The fact that minimum and maximum temperature data from the regional weather station were only available as integers implies that it is only collected as integers, which is a source of uncertainty. Though a simpler analysis between only Climate Edge data sets could have been performed, we decided it would be preferable to have a baseline dataset from a non-agricultural area to quantify buffering. It would have been ideal to compare the farm soil temperature data to soil temperature data from the same location as the regional weather station.

Conclusion

This study highlights the importance of quantifying microclimatic buffers, as despite temperatures differing between farms, and evidence for some buffering against mean temperatures, sun-exposed coffee is vulnerable to supraoptimal temperatures regardless of local environmental variables. Future studies should examine more environmental variables and also collect microclimate data from a range of management types, to explore to relative protection provided by different managements. This may also fill knowledge gaps on how exposure to different temperatures affect various physiological processes such as flowering (Rahn *et al.*, 2018). Microclimates could also be explored at a higher spatial resolution, as within a single coffee plant, leaf temperatures have been shown to experience a difference of up to 5°C depending on which direction they are facing (Bis *et al.*, 2020). This would allow for better informed management, particularly on larger farms. The future of Nicaraguan coffee remains uncertain, but with demand showing no signs of subsiding, providing better localised projections and mitigation solutions is essential (Bro, 2020) as the region remains highly vulnerable to climatic extremes (Hannah *et al.*, 2013).

Chapter 6: General Discussion

Loss of habitat and fragmentation present major threats to global biodiversity, particularly in areas of the tropics, which are home to a disproportionately high number of species (Myers *et al.*, 2000). Monitoring these species using low-cost methods such as passive acoustic monitoring (PAM) is key to understanding their responses to stressors. The tropics are also economically important for coffee agriculture, which only grows in high altitudes of the tropics. Ensuring continued production under less certain climatic conditions is economically important to countries like Brazil and Nicaragua which rely heavily on coffee exports. Monitoring how coffee, biodiversity and local microclimate interact can inform future climate smart management of agriculture. Developing methods that can accurately reflect biodiversity trends is necessary to accurately inform conservation policy as well as sustainable agriculture policy.

6.1: Acoustic monitoring: best uses and implementations

Acoustic monitoring is a rapidly developing field particularly in the field of biodiversity monitoring. The advantages of acoustic monitoring are the ease of data collection and the wide range of environments which can be monitored using acoustics: marine, freshwater and terrestrial (Browning *et al.*, 2017). The range of analytical tools available has rapidly expanded, from more general soundscape and index approaches to more modern machine learning models (Gibb *et al.*, 2019; Stowell, 2022) that detect and classify specific sounds of interest. Acoustic indices remain popular due to their ease of application though they have mixed effects in their ability to accurately quantify biodiversity and require careful selection (Buxton *et al.*, 2018). Not many tools exist between indices and more complex machine learning methods, which require more computational knowledge to apply. However, the acoustic composition approach presented in Chapter 3 both reflects biodiversity and does not require extensive machine learning knowledge. As the number of available tools expand, more consensus is needed in acoustic survey protocols that are region specific.

Birds have been a key target for acoustic monitoring surveys. As in other fields, automated species detection models are more advanced in temperate regions than in the tropics, with BirdNET being able to identify 984 North American and European species (Kahl *et al.*, 2021). However, in the tropics, researchers often have to resort to acoustic indices, manual listening, spectrogram analysis and expert ornithologist verification (Metcalf, Barlow, Marsden, *et al.*, 2021). In some cases where automation was attempted, comparison with manual inspection revealed low detection rates (Szymański *et al.*, 2021). Expansion of tools such as BirdNET into tropical regions will allow for easier monitoring of communities. Once species identification tools are developed, the next hurdle is developing accurate methods to estimate abundance, which is starting to be explored (Pérez-Granados *et al.*, 2019).

Acoustic composition methods, as applied in Chapter 3, show potential for monitoring of avian and insect biodiversity. The advantage of this approach is its transferability to other regions and the ease with which drivers can be detected and identified. The methodology in Chapter 3 was applied in a master's project by C. Declaudure (Appendix B.2), with point count data and commonly used acoustic indices as well. Acoustic composition was shown to best reflect the avian community above any individual acoustic index or combination of indices. However, with the advance of machine learning methodologies, acoustic composition may fill a gap between simplistic acoustic indices and automated species detection but is likely to be overtaken by species detection methods when these become available. Other similar approaches using machine learning and soundscape analysis suggest the field could move away from acoustic indices (Sethi *et al.*, 2020), although existing large-scale monitoring projects are implementing index based approaches (Towsey *et al.*, 2018).

In the field of species-level acoustic monitoring, bats are one of the most advanced taxa, as in some countries all species can be identified by echolocation calls alone. As is the case for most biodiversity monitoring, tools available in North America and Europe are far ahead of those in the tropics. Live bat detection systems in London have already been successfully deployed with smart monitors that detect bat approaches, and identify the species based on a pass (Gallacher *et al.*, 2021). The challenge for UK bats is smaller as there are only 18 species present (Bat Conservation Trust, 2022), whereas Brazil is known to have over 180 (Arias-Aguilar *et al.*, 2018). However, the work presented in Chapter 4 shows that a CNN designed to detect bat calls in the UK can successfully be applied to data from Brazil. Across all bats, the number of echolocation call shapes is limited (estimated to seven distinct types) (Jones and Teeling, 2006), and although there is variation between species, a sufficiently trained CNN should detect the presence of echolocating bats. This work is at the forefront of species-specific acoustic monitoring in biodiversity hotspots, providing a sound analytical pipeline that can be applied and expanded to other regions. The next challenge is creating robust classifiers for the tropics, which requires the creation of reliable call libraries. Beyond monitoring bat species, machine learning has the potential to distinguish between call types such as social calls and feeding buzzes, which can give insights into bat activity and behaviour.

Acoustic monitoring has a promising future and the potential to fill many knowledge gaps. With machine learning and more specifically, deep learning methods, species-level detections are being achieved more an increasing number of taxa. Recording devices with a power supply (e.g. solar power) and ability to transfer data wirelessly can remain in remote locations for extensive periods of time without interference (Sethi *et al.*, 2018). More recently, biodegradable sensors have been developed which can be deployed without being collected (Sethi *et al.*, 2022). In Australia, sensor networks have been deployed for the automatic detection of cane toads, an invasive species, to provide early warning systems (Roe *et al.*, 2018). Models for non-biotic sounds such as gunshots are also improving (Sigmund and Hrabina, 2021), which can detect levels of human impacts such as hunting pressures. If most acoustically active species could be detected and classified, rolling out networks of acoustic monitors could remotely monitor communities, anomalies, population trends and responses to ecological stressors. Considering the pace of technological advances, some of the approaches used in this thesis may already be redundant, particularly the generalised soundscape approaches. Using current technology, an unsupervised machine learning approach paired with expert ornithologist listening would be a better alternative to what I did in chapter 3. Expert labelled clips could then contribute to a training library for Brazilian biodiversity and an open data bank. Acoustic monitoring is going to be a key part of biodiversity monitoring across all ecosystems and will be able to provide detailed information on acoustically active biodiversity at a global scale.

6.2: Microclimate monitoring for coffee

Coffee crops are sensitive to extremes in temperature and are at risk of physiological damage with climate change particularly due to supraoptimal temperature exposure (Ovalle-Rivera *et al.*, 2015b). Coffee is grown mainly in mountainous regions with topographic complexity, therefore that farms within close geographic proximity may have very different microclimates to one another. Elevation of coffee farms has a considerable effect on microclimate as well as taste of the bean (Smith, 2018; Getachew *et al.*, 2022). Understanding the impacts of microclimate on coffee growth can inform better management and mitigate the impacts of climate change. Microclimate also impact soil fauna (Karungi *et al.*, 2018) and the likelihood of pest and disease outbreaks (Liebig *et al.*, 2019), a phenomenon that is set to increase with climate change (Jaramillo *et al.*, 2009). Monitoring microclimate at farm-level will likely form part of future agricultural management so farmers can customise management to their local needs.

Deployment of microclimate weather stations can form part of climate smart agriculture as farm-level problems can be solved by better understanding of local weather (Tenzin *et al.*, 2017). This is only possible due to the availability of low-cost sensors that can be customised to measure desired variables of import and deployed. To better understand the impact of local microclimate on coffee production, and build on the trends observed in Chapter 5, more data on coffee yield, farm-level management as well as more detail on farm topography is needed. If some farms do transition to more agroforest management, then performing a pairwise study over the course of several years would be informative to compare the performance of coffee under different conditions, as in other locations, shade coffee has shown more resilience against extremes in temperature (Vaast *et al.*, 2016). The options for farmers to mitigate against temperature extremes is limited. Apart from changing management, new technologies such as kaolin spray have been successful in reducing leaf surface temperature and doubled the yield compared to non-treated sun coffee (Steiman, Bittenbender and Idol, 2007), though this is expensive. The future of coffee in Nicaragua is currently uncertain, as lower elevations are no longer suitable, and there is little space at higher altitudes for farms to expand into.

As the cost of sensors and data storage is low, there is global interest in exploring and maximising opportunities they offer. Integrating information from multiple sources forms part of agricultural modernisation within the Internet of Things paradigm (Sinha and Dhanalakshmi, 2022). Under this paradigm, a range of sensors can contribute to broad automation within the agricultural sector with minimal human intervention. Though the work conducted in chapter 5 used data from such sensors, the full potential of these weather stations was not realised.

6.3: The future of coffee, implications for stakeholders

Coffee is at high risk of yield and quality loss as a result of climate change (Bunn, Läderach, Rivera, *et al.*, 2015; Ovalle-Rivera *et al.*, 2015b; Kath *et al.*, 2020). This will likely lead to volatility in production, and thus price. The majority of coffee producers are smallholders (an estimated 20 million globally people rely on coffee agriculture for income (Toledo and Moguel, 2012)), who cannot afford losses in yield, so require strategies to provide a stable income, which may be through climate smart agriculture, diversification or management changes (Rahn *et al.*, 2018; Djufry, Wulandari and Villano, 2022). Though some recent literature has argued that increased atmospheric CO_2 may mitigate the physiological stress caused by supraoptimal temperatures on coffee plants limits (Verhage, Anten and Sentelhas, 2017; Rakocevic *et al.*, 2018; DaMatta *et al.*, 2019), this is not sufficient to prevent all damage (Marques *et al.*, 2021). Unlike other cash crops, developing new varieties of *C. arabica* is difficult, as it is an allotetraploid with low levels of variation (Scalabrin *et al.*, 2020). Certification schemes such as Fairtrade can provide higher income per quantity of coffee, which has been a popular option for many smallholders, but results have been mixed (Bray and Neilson, 2017; Jena, Stellmacher and Grote, 2017). Though mostly grown under direct sunlight, shade is considered advantageous for coffee production (Biruk, 2018) and transitioning to mixed agroforest management is gaining popularity (Rahn *et al.*, 2018).

Farmers may be hesitant to transition from sun-coffee to agroforest due to uncertainty, potential yield loss from coffee, higher labour costs and insufficient models and knowledge in the region (Sagastuy and Krause, 2019). Best practices for agroforest management rely on local knowledge and research on which species perform best with coffee for intercropping (Santos *et al.*, 2011), a field that requires more development for Brazil. A range of certification schemes exist that encourage shade adaptation such as Rainforest Alliance and Bird Friendly Coffee, which have demonstrably improved economic outcomes for farmers whilst simultaneously promoting conservation and biodiversity (Philpott *et al.*, 2007). Schemes such as these provide synergistic solutions that benefit both coffee farmers and biodiversity in the longer term. However, these schemes are often run by private companies and becoming certified can be a difficult process for farmers. There is space to bring in policy solutions and incentives that have the same results without private certification companies.

6.4: Implications for the Atlantic Forest, conservation and policy

The Atlantic Forest remains a key conservation target for biodiversity, but with an estimated 26% of its original extent remaining, much of which is made up of smaller fragments and only 30% of this forest within protected areas (Rezende *et al.*, 2018), there is room for more policy intervention. To preserve the maximum amount of biodiversity, at least 30% native vegetation cover is required (Banks-Leite *et al.*, 2014). Currently the Brazil Forest Code requires native vegetation to be maintained in areas including hilltops and steep slopes (Areas of Permanent Protection) (Brancalion *et al.*, 2016), and evidence suggests that the presence of native forest fragments positively impacts coffee production in Brazil's most productive coffee regions (Latini *et al.*, 2020). Agricultural subsidies could be made available for farmers keen to transition to agroforest schemes to offset initial losses. Even small changes like tree planting around plantation edges should be promoted for intensive smallholders. Policies intended to promote agroforest have had mixed impacts so should be made carefully with strong inputs from smallholders involved (Vaast *et al.*, 2016) to benefit both future coffee production and local biodiversity.

Given the low cost and ease of deployment of PAM, it can play a significant role in ongoing and future monitoring schemes. However, acoustic data, particularly ultrasonic data results in a lot of data being collected. More efficient systems with on-board algorithms that can extract species identity without storing the data exist for British bats (Gallacher *et al.*, 2021). If a similar system could be developed for Brazil, monitoring Atlantic Forest bat populations in real-time could become a reality. Acoustic monitoring have already demonstrated that poaching pressure remains high in protected areas of the Atlantic Forest, highlighting the need for urgent intervention (Pardo *et al.*, 2022). Implementing an acoustic-composition based monitoring scheme could provide useful insights into avian community trends while better species detection tools are developed, as acoustic-indices have highly variable outputs (Scarpelli, Ribeiro and Teixeira, 2021).

6.5: Limitations of fieldwork design

The use of PAM for data collection with Audiomoths (Hill *et al.*, 2018) allowed for hundreds of hours of sound to be collected at desired sampling rates with relative ease. However, in the years since this fieldwork was carried out, Audiomoths and other devices have been significantly improved. For instance, those used in this study had to be pre-ordered months in advance and had no protective covering, so they were deployed in resealable plastic bags containing silica packets to absorb any excess moisture. This was insufficient protection for some devices, as they malfunctioned due to water damage, though more recently, Audiomoths are available with protective cases that are suitable for all weather and underwater deployment (Open Acoustic Devices, 2022). The recorders were also limited by the capacity of the 32 GB SD cards. Due to time constraints, sound detection ranges of the devices in different land uses were not explored. Low-cost PAM devices are fragile, but can now be better protected and other studies have adapted recorders to send data remotely via mobile networks (Sethi *et al.*, 2018), removing the need for on-board storage and creating the potential for very long-term monitoring, which is ideal for remote locations.

The field work study design was pre-organised by my collaborator, Andrea Larissa Boesing, who had previously selected the coffee farms and obtained the relevant permits. Though initially the plan was to deploy the recorders for two seasons (one without, and one during an outbreak of coffee borer beetle), acoustic recording was only done for the first season. This changed the questions being explored as seasonal comparisons and impact of pest could no longer be explored. The field work schedule only permitted for 3 days of mist netting at each site, which may not have fully captured the avian communities at all sites, particularly if the weather was cold or rainy. Given more time and if more visits could have been made to each farm, I would have extended the recording schedule for recorders at both sampling rates, recording 1 minute every 5 minutes for 24 hours. This would then include more of the dusk bat activity and any diurnal ultrasonics. Ideally more data on the vegetation density in the forest fragments would have been collected as this may explain the variance in bat presence within the forest.

6.6: Main conclusions

Acoustic monitoring of tropical biodiversity can give useful insights into species responses to ecological stressors, though tools available for tropical regions lag behind those available in more temperate regions. Acoustic composition methods show that bird communities strongly correlate with local forest cover in coffee-fragmented regions of the Atlantic Forest. This reflects other work from the Atlantic Forest highlighting the sensitivity of bird communities to fragmentation (Martensen *et al.*, 2012). The potential for highly localised predictions is more tangible than ever and I foresee a future where microclimate, biodiversity and soil health data can be integrated for individual farms (as part of IoT), so climate smart agriculture models can make recommendations that promote both crop and ecosystem health.

Bat calls in the Atlantic Forest can successfully be detected using a CNN designed for the UK, though a lack of call libraries limits current species classification approaches. The richness of the bat communities was relatively homogenous across the sample sites, though abundance was significantly higher within coffee land use than in adjacent forest patches.

Microclimatic buffering on Nicaraguan sun-coffee farms showed some buffering against mean and minimum temperatures, but the lack of buffering against high temperatures means that the crop remains vulnerable to the ongoing effects of climate change. As local forest cover and elevation did not impact the buffering, changes in management towards climate-smart agriculture is necessary.

Tropical regions such as the Atlantic Forest can provide habitat for their valuable biodiversity as well as continue producing high quality coffee, though not by keeping current practices in place.

References

de Aguiar Silva, C. *et al.* (2022) ‘Listening in the dark: acoustics indices reveal bat species diversity in a tropical savannah’, *Bioacoustics* [Preprint]. doi:10.1080/09524622.2022.2053741/SUPPL_FILE/TBIO_A_2053741_SM4257.PDF.

Aide, T.M. *et al.* (2017) ‘Species richness (of insects) drives the use of acoustic space in the tropics’, *Remote Sensing*, 9(11). doi:10.3390/rs9111096.

Almeida Silva, K. *et al.* (2020) ‘Influence of El Niño and La Niña on coffee yield in the main coffee-producing regions of Brazil’, *Theoretical and Applied Climatology*, 139(3–4), pp. 1019–1029. doi:10.1007/s00704-019-03039-9.

Álvarez-Presas, M. *et al.* (2011) ‘Land planarians (Platyhelminthes) as a model organism for fine-scale phylogeographic studies: understanding patterns of biodiversity in the Brazilian Atlantic Forest hotspot’, *Journal of Evolutionary Biology*, 24(4), pp. 887–896. doi:10.1111/j.1420-9101.2010.02220.x.

Ambinakudige, S. and Choi, J. (2009) ‘Global coffee market influence on land-use and land-cover change in the Western Ghats of India’, *Land Degradation and Development*, 20(3), pp. 327–335. doi:10.1002/ldr.921.

Mac Aodha, O. *et al.* (2018) ‘Bat detective - Deep learning tools for bat acoustic signal detection’, *PLoS computational biology*, 14(3), p. e1005995.

Mac Aodha, O. *et al.* (no date) *Deep Joint Detection and Classification of Bat Echolocation calls*.

Arias-Aguilar, A. *et al.* (2018) ‘Who’s calling? Acoustic identification of Brazilian bats’, *Mammal Research*. Springer Berlin Heidelberg, pp. 231–253. doi:10.1007/s13364-018-0367-z.

Arora, N.K. (2019) ‘Impact of climate change on agriculture production and its sustainable solutions', *Environmental Sustainability*, 2(2), pp. 95–96. doi:10.1007/s42398-019-00078-w.

Ashcroft, M.B. (2010) 'Identifying refugia from climate change', *Journal of Biogeography*, 37(8), pp. 1407–1413. doi:10.1111/J.1365-2699.2010.02300.X.

Avelino, J. *et al.* (2015) 'The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions', *Food Security*, 7(2), pp. 303–321. doi:10.1007/s12571-015-0446-9.

Azevedo, T. *et al.* (2018) *MapBiomas initiative: Mapping annual land cover and land use changes in Brazil from 1985 to 2017.* Available at: http://adsabs.harvard.edu/abs/2018AGUFM.B22A..04A (Accessed: 30 January 2020).

Bacon, C.M. *et al.* (2017) 'Vulnerability to Cumulative Hazards: Coping with the Coffee Leaf Rust Outbreak, Drought, and Food Insecurity in Nicaragua', *World Development*, 93, pp. 136–152. doi:10.1016/J.WORLDDEV.2016.12.025.

Baker, P.S. and Haggar, J. (2007) 'Global Warming: the impact on global coffee', in. Long Beach, pp. 1–14.

Banks-Leite, C. *et al.* (2014) 'Using ecological thresholds to evaluate the costs and benefits of set-asides in a biodiversity hotspot', *Science*, 345(6200), pp. 1041–1045. doi:10.1126/SCIENCE.1255768.

Banks-Leite, C., Ewers, R.M. and Metzger, J.-P. (2010) 'Edge effects as the principal cause of area effects on birds in fragmented secondary forest', *Oikos*, 119(6), pp. 918–926. doi:10.1111/j.1600-0706.2009.18061.x.

Banks-Leite, C., Ewers, R.M. and Metzger, J.P. (2012) 'Unraveling the drivers of community dissimilarity and species extinction in fragmented landscapes', *Ecology*, 93(12), pp. 2560–2569. doi:10.1890/11-2054.1.

Barbosa, K.V. de C. *et al.* (2017) 'Use of small Atlantic Forest fragments by birds in

Southeast Brazil', *Perspectives in Ecology and Conservation*, 15(1), pp. 42–46. doi:10.1016/j.pecon.2016.11.001.

Bas, Y., Bas, D. and Julien, J.-F. (2017) 'Tadarida: A Toolbox for Animal Detection on Acoustic Recordings', *Journal of Open Research Software*, 5(1). doi:10.5334/jors.154.

Bat Conservation Trust (2022) *UK Bats - Types of bats - Bat Conservation Trust*. Available at: https://www.bats.org.uk/about-bats/what-are-bats/uk-bats (Accessed: 5 October 2022).

Batalha-Filho, Henrique *et al.* (2013) 'Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events', *J Ornithol*, 154, pp. 41–50. doi:10.1007/s10336-012-0866-7.

Bebber, D.P., Castillo, Á.D. and Gurr, S.J. (2016) 'Modelling coffee leaf rust risk in Colombia with climate reanalysis data', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1709). doi:10.1098/RSTB.2015.0458.

Becker, C.G. *et al.* (2007) 'Habitat split and the global decline of amphibians', *Science*, 318(5857), pp. 1775–1777. doi:10.1126/science.1149374.

Bello, C. *et al.* (2017) 'Atlantic frugivory: a plant–frugivore interaction data set for the Atlantic Forest', *Ecology*. Ecological Society of America, p. 1729. doi:10.1002/ecy.1818.

Benchimol, M. *et al.* (2017) 'Translating plant community responses to habitat loss into conservation practices: Forest cover matters', *Biological Conservation*, 209, pp. 499–507. doi:10.1016/J.BIOCON.2017.03.024.

Beuchelt, T.D. and Zeller, M. (2011) 'Profits and poverty: Certification's troubled link for Nicaragua's organic and fairtrade coffee producers', *Ecological Economics*, 70(7), pp. 1316–1324.

Biruk, A. (2018) 'Impact of shade on morpho-physiological characteristics of coffee plants, their pests and diseases: A review’, *African Journal of Agricultural Research*, 13(39), pp. 2016–2024. doi:10.5897/AJAR2018.13408.

Bis, A.P.P. *et al.* (2020) ‘Microclimatic characterization of a conilon coffee plantation grown in an east-west orientation’, *Australian Journal of Crop Science*, 14(3), pp. 431–438.

Boesing, A.L., Nichols, E. and Metzger, J.P. (2017) ‘Effects of landscape structure on avian-mediated insect pest control services: a review’, *Landscape ecology*, 32(5), pp. 931–944.

Boesing, A.L., Nichols, E. and Metzger, J.P. (2018) ‘Biodiversity extinction thresholds are modulated by matrix type’, *Ecography*, 41(9), pp. 1520–1533. doi:10.1111/ECOG.03365.

Bogdanowicz, W., Fenton, M.B. and Daleszczyk, K. (1999) ‘The relationships between echolocation calls, morphology and diet in insectivorous bats’, *Journal of Zoology*, 247(3), pp. 381–393. doi:10.1111/j.1469-7998.1999.tb01001.x.

Bohmann, K. *et al.* (2014) ‘Environmental DNA for wildlife biology and biodiversity monitoring’, *Trends in Ecology and Evolution*. Elsevier Ltd, pp. 358–367. doi:10.1016/j.tree.2014.04.003.

Boonman, A. *et al.* (2013) ‘It’s not black or white-on the range of vision and echolocation in echolocating bats’, *Frontiers in Physiology*, 4, p. 248. doi:10.3389/FPHYS.2013.00248/BIBTEX.

Borchers, H.W. (2021) ‘Package “pracma” ’.

Boscolo, D. and Metzger, J.P. (2009) ‘Is bird incidence in Atlantic forest fragments influenced by landscape patterns at multiple scales?’, *Landscape Ecology 2009 24:7*, 24(7), pp. 907–918. doi:10.1007/S10980-009-9370-8.

Boughey, K.L. *et al.* (2011) ‘Effects of landscape-scale broadleaved woodland configuration and extent on roost location for six bat species across the UK’, *Biological*

Conservation, 144(9), pp. 2300–2310. doi:10.1016/J.BIOCON.2011.06.008.

Bradfer-Lawrence, T. *et al.* (2019) ‘Guidelines for the use of acoustic indices in environmental research’, *Methods in Ecology and Evolution*. Edited by V. Zamora-Gutierrez, 10(10), pp. 1796–1807. doi:10.1111/2041-210X.13254.

Brancalion, P.H.S. *et al.* (2016) ‘A critical analysis of the Native Vegetation Protection Law of Brazil (2012): Updates and ongoing initiatives’, *Natureza e Conservacao*, 14, pp. 1–15. doi:10.1016/J.NCON.2016.03.003.

Bray, J.G. and Neilson, J. (2017) ‘Reviewing the impacts of coffee certification programmes on smallholder livelihoods’, *https://doi.org/10.1080/21513732.2017.1316520*, 13(1), pp. 216–232. doi:10.1080/21513732.2017.1316520.

Britzke, E.R. *et al.* (2011) ‘Acoustic identification of bats in the eastern United States: A comparison of parametric and nonparametric methods’, *The Journal of Wildlife Management*, 75(3), pp. 660–667. doi:10.1002/JWMG.68.

Bro, A.S. (2020) ‘Climate Change Adaptation, Food Security, and Attitudes toward Risk among Smallholder Coffee Farmers in Nicaragua’, *Sustainability 2020, Vol. 12, Page 6946*, 12(17), p. 6946. doi:10.3390/SU12176946.

Brock, R.C. *et al.* (2021) ‘Implementing Brazil’s Forest Code: a vital contribution to securing forests and conserving biodiversity’, *Biodiversity and Conservation*, 30(6), pp. 1621–1635. doi:10.1007/S10531-021-02159-X/FIGURES/3.

Browning, E. *et al.* (2017) *Acoustic Monitoring*, *WWF Conservation Technology Series*. Available at: https://www.wwf.org.uk/conservationtechnology/documents/Acousticmonitoring-WWF-guidelines.pdf (Accessed: 15 November 2018).

Bunn, C., Läderach, P., Rivera, O.O., *et al.* (2015) ‘A bitter cup: climate change profile of global production of Arabica and Robusta coffee’, *Climatic Change*, 129(1–2), pp. 89–

101.

Bunn, C., Läderach, P., Jimenez, J.G.P., *et al.* (2015) 'Multiclass Classification of Agro-Ecological Zones for Arabica Coffee: An Improved Understanding of the Impacts of Climate Change', *PLOS ONE*, 10(10), p. e0140490. doi:10.1371/JOURNAL.PONE.0140490.

Burnham, K.P. and Anderson, D.R. (2016) 'Multimodel Inference: Understanding AIC and BIC in Model Selection', *http://dx.doi.org/10.1177/0049124104268644*, 33(2), pp. 261–304. doi:10.1177/0049124104268644.

Butt, M.S. and Sultan, M.T. (2011) 'Coffee and its Consumption: Benefits and Risks', *Critical Reviews in Food Science and Nutrition*, 51(4), pp. 363–373. doi:10.1080/10408390903586412.

Buxton, R.T. *et al.* (2018) 'Efficacy of extracting indices from large-scale acoustic recordings to monitor biodiversity', *Conservation Biology*, 32(5), pp. 1174–1184. doi:10.1111/cobi.13119.

Campanha, M.M. *et al.* (2004) *Growth and yield of coffee plants in agroforestry and monoculture systems in Minas Gerais, Brazil*. Available at: https://link.springer.com/content/pdf/10.1023%2FB%3AAGFO.0000049435.22512.2d.pdf (Accessed: 7 December 2018).

Campos-Cerqueira, M. and Mitchell Aide, T. (2017) 'Changes in the acoustic structure and composition along a tropical elevational gradient'. doi:10.22261/JEA.PNCO7I.

Canale, G.R. *et al.* (2012) 'Pervasive Defaunation of Forest Remnants in a Tropical Biodiversity Hotspot', *PLoS ONE*. Edited by J. Chave, 7(8), p. e41671. doi:10.1371/journal.pone.0041671.

Carnaval, A.C. *et al.* (2009) 'Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot', *Science*, 323(5915), pp. 785–789. doi:10.1126/science.1166955.

Carnaval, A.C. and Moritz, C. (2008) ‘Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest’, *Journal of Biogeography*, 35(7), pp. 1187–1201. doi:10.1111/j.1365-2699.2007.01870.x.

Carnaval, A.C.O.D.Q. *et al.* (2006) ‘Amphibian chytrid fungus broadly distributed in the Brazilian Atlantic rain forest’, *EcoHealth*, 3(1), pp. 41–48. doi:10.1007/s10393-005-0008-2.

Carrasco, L.R. *et al.* (2014) ‘A double-edged sword for tropical forests’, *Science*. American Association for the Advancement of Science, pp. 38–40. doi:10.1126/science.1256685.

Cartes, J.L. (2003) ‘Brief history of conservation in the Interior Atlantic Forest’, in Leal, C.G. and Câmara, I.G. (eds) *The Atlantic Forest of South America: Biodiversity Status, Threats, and Outlook*. Island Press, pp. 269–287.

de Carvalho, A.F. *et al.* (2021) ‘Microclimate and soil and water loss in shaded and unshaded agroforestry coffee systems’, *Agroforestry Systems*, 95(1), pp. 119–134. doi:10.1007/s10457-020-00567-6.

Ceballos, G. *et al.* (2015) ‘Accelerated modern human-induced species losses: Entering the sixth mass extinction’, *Science Advances*, 1(5), p. e1400253. doi:10.1126/sciadv.1400253.

CEPF (2021) *Biodiversity hotspots defined | CEPF*. Available at: https://www.cepf.net/our-work/biodiversity-hotspots/hotspots-defined (Accessed: 29 June 2021).

Chandler, M. *et al.* (2017) ‘Contribution of citizen science towards international biodiversity monitoring’, *Biological Conservation*, 213, pp. 280–294. doi:10.1016/j.biocon.2016.09.004.

Chebez, J. and Hilgert, N. (2003) ‘Brief history of conservation in the Paraná Forest’, in Leal, C.G. and Câmara, I.G. (eds) *The Atlantic Forest of South America: Biodiversity Status, Threats, and Outlook*. Washington: Island Press, pp. 141–159. Available at:

nd&pg=PA141&dq=parana+conservation+forest+argentina+chebez&ots=XPtsmzywKP&sig=Wl-iHBOU-9Ah5P0abC10fWjI3V4 (Accessed: 30 January 2020).

Chen, X. *et al.* (2020) 'Automatic standardized processing and identification of tropical bat calls using deep learning approaches', *Biological Conservation*, 241, p. 108269. doi:10.1016/j.biocon.2019.108269.

Climate Edge (2021) *Climate Edge, Service platform*. Available at: https://www.climate-edge.com/service-platform (Accessed: 1 June 2021).

Collen, B. *et al.* (2008) 'The Tropical Biodiversity Data Gap: Addressing Disparity in Global Monitoring', *Tropical Conservation Science*, 1(2), pp. 75–88. doi:10.1177/194008290800100202.

Coltri, P.P. *et al.* (2019) 'Low levels of shade and climate change adaptation of Arabica coffee in southeastern Brazil', *Heliyon*, 5(2), p. e01263. doi:10.1016/j.heliyon.2019.e01263.

Colwell, R.K. *et al.* (2008) 'Global warming, elevational range shifts, and lowland biotic attrition in the wet tropics', *Science*, 322(5899), pp. 258–261. doi:10.1126/science.1162547.

Conservation International (2021) *Biodiversity Hotspots*. Available at: https://www.conservation.org/priorities/biodiversity-hotspots (Accessed: 29 June 2021).

Craparo, A.C.W. *et al.* (2015) 'Coffea arabica yields decline in Tanzania due to climate change: Global implications', *Agricultural and Forest Meteorology*, 207, pp. 1–10. doi:10.1016/J.AGRFORMET.2015.03.005.

Crouzeilles, R. *et al.* (2019) 'There is hope for achieving ambitious Atlantic Forest restoration commitments', *Perspectives in Ecology and Conservation*, 17(2), pp. 80–83. doi:10.1016/j.pecon.2019.04.003.

Culot, L. *et al.* (2019) ‘ATLANTIC-PRIMATES: a dataset of communities and occurrences of primates in the Atlantic Forests of South America’, *Ecology*, 100(1). doi:10.1002/ecy.2525.

DaMatta, F.M. (2004) ‘Ecophysiological constraints on the production of shaded and unshaded coffee: a review’, *Field Crops Research*, 86(2–3), pp. 99–114.

DaMatta, F.M. *et al.* (2019) ‘Why could the coffee crop endure climate change and global warming to a greater extent than previously estimated?’, *Climatic Change*, 152(1), pp. 167–178. doi:10.1007/s10584-018-2346-4.

DaMatta, F.M. and Ramalho, J.D.C. (2006) ‘Impacts of drought and temperature stress on coffee physiology and production: a review’, *Brazilian Journal of Plant Physiology*, 18(1), pp. 55–81. doi:10.1590/S1677-04202006000100006.

Darras, K. *et al.* (2016) ‘Measuring sound detection spaces for acoustic animal sampling and monitoring’, *Biological Conservation*, 201, pp. 29–37. doi:10.1016/J.BIOCON.2016.06.021.

Davis, A.P. *et al.* (2012) ‘The Impact of Climate Change on Indigenous Arabica Coffee (Coffea arabica): Predicting Future Trends and Identifying Priorities’, *PLoS ONE*. Edited by B. Fenton, 7(11), p. e47981. doi:10.1371/journal.pone.0047981.

Deblauwe, V. *et al.* (2016) ‘Remotely sensed temperature and precipitation data improve species distribution modelling in the tropics’, *Global Ecology and Biogeography*, 25(4), pp. 443–454. doi:10.1111/geb.12426.

Deichmann, J.L. *et al.* (2017) ‘Soundscape analysis and acoustic monitoring document impacts of natural gas exploration on biodiversity in a tropical forest-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)’, *Ecological Indicators*, 74, pp. 39–48. doi:10.1016/j.ecolind.2016.11.002.

Delciellos, A.C. *et al.* (2018) ‘Habitat fragmentation affects individual condition: evidence from small mammals of the Brazilian Atlantic Forest', *Journal of Mammalogy*, 99(4), pp. 936–945. doi:10.1093/JMAMMAL/GYY078.

Delgado-Jaramillo, M. *et al.* (2020) 'Assessing the distribution of a species-rich group in a continental-sized megadiverse country: Bats in Brazil', *Diversity and Distributions*, 26(5), pp. 632–643. doi:10.1111/DDI.13043.

Desjonquères, C., Gifford, T. and Linke, S. (2020) 'Passive acoustic monitoring as a potential tool to survey animal and ecosystem processes in freshwater environments', *Freshwater Biology*. Blackwell Publishing Ltd, pp. 7–19. doi:10.1111/fwb.13356.

Didham, R.K., Kapos, V. and Ewers, R.M. (2012) 'Rethinking the conceptual foundations of habitat fragmentation research', *Oikos*, 121(2), pp. 161–170. doi:10.1111/J.1600-0706.2011.20273.X.

Dixon, A.P., Baker, M.E. and Ellis, E.C. (2020) 'Agricultural Landscape Composition Linked with Acoustic Measures of Avian Diversity', *Land 2020, Vol. 9, Page 145*, 9(5), p. 145. doi:10.3390/LAND9050145.

Djufry, F., Wulandari, S. and Villano, R. (2022) 'Climate Smart Agriculture Implementation on Coffee Smallholders in Indonesia and Strategy to Accelerate', *Land 2022, Vol. 11, Page 1112*, 11(7), p. 1112. doi:10.3390/LAND11071112.

Dong, D. *et al.* (2014) 'Temperature trend–altitude relationship in China during 1963–2012', *Theoretical and Applied Climatology 2014 122:1*, 122(1), pp. 285–294. doi:10.1007/S00704-014-1286-9.

Dufrêne, M. and Legendre, P. (1997) 'Species assemblages and indicator species: the need for a flexible asymmetrical approach', *Ecological Monographs*, 67(3), pp. 345–366. doi:10.1890/0012-9615(1997)067[0345:saaist]2.0.co;2.

Ennis, K.K. and Philpott, S.M. (2019) 'Seasonal and microhabitat differences alter ant predation of a globally disruptive coffee pest', *Agriculture, Ecosystems & Environment*, 284, p. 106597. doi:10.1016/J.AGEE.2019.106597.

Ewers, R.M. and Banks-Leite, C. (2013) ‘Fragmentation Impairs the Microclimate Buffering Effect of Tropical Forests’, *PLoS ONE*. Edited by G. Bohrer, 8(3), p. e58093. doi:10.1371/journal.pone.0058093.

Fahrig, L. *et al.* (2019) ‘Is habitat fragmentation bad for biodiversity?’, *Biological Conservation*, 230, pp. 179–186. doi:10.1016/J.BIOCON.2018.12.026.

Fain, S.J. *et al.* (2017) ‘Climate change and coffee: assessing vulnerability by modeling future climate suitability in the Caribbean island of Puerto Rico’, *Climatic Change 2017 146:1*, 146(1), pp. 175–186. doi:10.1007/S10584-017-1949-5.

Fairbrass, A.J. *et al.* (2017) ‘Biases of acoustic indices measuring biodiversity in urban areas’, *Ecological Indicators*, 83, pp. 169–177. doi:10.1016/J.ECOLIND.2017.07.064.

Faria, D. (2006) ‘Phyllostomid bats of a fragmented landscape in the north-eastern Atlantic forest, Brazil’, *Journal of Tropical Ecology*, 22(5), pp. 531–542. doi:10.1017/S0266467406003385.

Farina, A. *et al.* (2011) ‘Avian soundscapes and cognitive landscapes: theory, application and ecological perspectives’, *Landscape Ecology 2011 26:9*, 26(9), pp. 1257–1267. doi:10.1007/S10980-011-9617-Z.

Farina, A. *et al.* (2015) ‘Spatial and temporal variation of bird dawn chorus and successive acoustic morning activity in a Mediterranean landscape’, *http://dx.doi.org/10.1080/09524622.2015.1070282*, 24(3), pp. 269–288. doi:10.1080/09524622.2015.1070282.

Farrell, M.J.O., Corben, C. and Gannon, W.L. (2000) ‘Geographic variation in the echolocation calls of the hoary bat (Lasiurus cinereus)’, *Acta Chiropterologica*, 02(2), pp. 185--195.

Faye, E. *et al.* (2014) ‘Strong Discrepancies between Local Temperature Mapping and

Interpolated Climatic Grids in Tropical Mountainous Agricultural Landscapes', *PLOS ONE*, 9(8), p. e105541. doi:10.1371/JOURNAL.PONE.0105541.

Fick, S.E. and Hijmans, R.J. (2017) 'WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas', *International Journal of Climatology*, 37(12), pp. 4302–4315. doi:10.1002/joc.5086.

Filgueiras, B.K.C., Iannuzzi, L. and Leal, I.R. (2011) 'Habitat fragmentation alters the structure of dung beetle communities in the Atlantic Forest', *Biological Conservation*, 144(1), pp. 362–369. doi:10.1016/j.biocon.2010.09.013.

Fjeldså, J. and Rahbek, C. (2006) 'Diversification of tanagers, a species rich bird group, from lowlands to montane regions of South America', *Integrative and Comparative Biology*, 46(1), pp. 72–81. doi:10.1093/icb/icj009.

Fletcher, R.J. *et al.* (2018) 'Is habitat fragmentation good for biodiversity?', *Biological Conservation*, 226, pp. 9–15. doi:10.1016/J.BIOCON.2018.07.022.

Flores, M. *et al.* (2002) *Centroamérica: El impacto de la caída de los precios del café*.

Freitas, S.R., Hawbaker, T.J. and Metzger, J.P. (2010) 'Effects of roads, topography, and land use on forest cover dynamics in the Brazilian Atlantic Forest', *Forest Ecology and Management*, 259(3), pp. 410–417. doi:10.1016/j.foreco.2009.10.036.

De Frenne, P. *et al.* (2019) 'Global buffering of temperatures under forest canopies', *Nature Ecology and Evolution*, 3(5), pp. 744–749. doi:10.1038/s41559-019-0842-1.

Frey, S.J.K. *et al.* (2016) 'Spatial models reveal the microclimatic buffering capacity of old-growth forests', *Science Advances*, 2(4), p. e1501392. doi:10.1126/SCIADV.1501392.

Fridley, J.D. (2009) 'Downscaling Climate over Complex Terrain: High Finescale (<1000 m) Spatial Variation of Near-Ground Temperatures in a Montane Forested Landscape (Great Smoky Mountains)', *Journal of Applied Meteorology and Climatology*, 48(5), pp. 1033–1049. doi:10.1175/2008JAMC2084.1.

Fuller, S. *et al.* (2015) ‘Connecting soundscape to landscape: Which acoustic index best describes landscape configuration?’, *Ecological Indicators*, 58, pp. 207–215. doi:10.1016/J.ECOLIND.2015.05.057.

Furumo, P.R. and Mitchell Aide, T. (2019) ‘Using soundscapes to assess biodiversity in Neotropical oil palm landscapes’, *Landscape Ecology*, 34(4), pp. 911–923. doi:10.1007/s10980-019-00815-w.

Gage, S.H. *et al.* (2017) ‘Acoustic patterns at the Samford Ecological Research Facility in South East Queensland, Australia: The Peri-Urban SuperSite of the Terrestrial Ecosystem Research Network’, *Ecological Informatics*, 38, pp. 62–75. doi:10.1016/J.ECOINF.2017.01.002.

Galetti, M. *et al.* (2013) ‘Functional extinction of birds drives rapid evolutionary changes in seed size’, *Science*, 340(6136), pp. 1086–1090. doi:10.1126/science.1233774.

Galford, G.L. *et al.* (2013) ‘Intensification of tropical agriculture as seen by satellite’, *AGUFM*, 2013, pp. B53E-01. Available at: https://ui.adsabs.harvard.edu/abs/2013AGUFM.B53E..01G/abstract (Accessed: 29 June 2021).

Galindo-Leal, C. and Gusmão Câmara, I. de (2003) ‘The Atlantic forest of South America: biodiversity status, threats and outlook.’, *The Atlantic forest of South America: biodiversity status, threats and outlook.* [Preprint].

Gallacher, S. *et al.* (2021) ‘Shazam for bats: Internet of Things for continuous real-time biodiversity monitoring’. doi:10.1049/smc2.12016.

García-Morales, R. *et al.* (2016) ‘Deforestation Impacts on Bat Functional Diversity in Tropical Landscapes’, *PLOS ONE*, 11(12), p. e0166765. doi:10.1371/JOURNAL.PONE.0166765.

García-Morales, R., Badano, E.I. and Moreno, C.E. (2013) ‘Response of Neotropical Bat

Assemblages to Human Land Use', *Conservation Biology*, 27(5), pp. 1096–1106. doi:10.1111/COBI.12099.

Gaspar, L.P. (2021) *Predicting bird diversity using acoustic indices within the Atlantic Forest biodiversity hotspot.* Instituto de Biociências, Rio Claro. Available at: https://repositorio.unesp.br/handle/11449/204710.

Getachew, M. *et al.* (2022) 'Shade tree canopy cover affects coffee plant traits across elevations in coffee farms in southwest Ethiopia', *Nordic Journal of Botany*, 2022(2), p. e03383. doi:10.1111/NJB.03383.

Gibb, R. *et al.* (2019) 'Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring', *Methods in Ecology and Evolution*. British Ecological Society, pp. 169–185. doi:10.1111/2041-210X.13101.

Gibbs, H.K. *et al.* (2010) 'Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s', *Proceedings of the National Academy of Sciences of the United States of America*, 107(38), pp. 16732–16737. doi:10.1073/pnas.0910275107.

Gorresen, P.M. and Willig, M.R. (2004) 'Landscape Responses of Bats to Habitat Fragmentation in Atlantic Forest of Paraguay', *Journal of Mammalogy*, 85(4), pp. 688–697. doi:10.1644/BWG-125.

Goulson, D. (2019) 'The insect apocalypse, and why it matters', *Current Biology*. Cell Press, pp. R967–R971. doi:10.1016/j.cub.2019.06.069.

Gourdji, S. *et al.* (2015) 'Historical climate trends, deforestation, and maize and bean yields in Nicaragua', *Agricultural and Forest Meteorology*, 200, pp. 270–281. doi:10.1016/j.agrformet.2014.10.002.

Grolemund, G. and Wickham, H. (2011) 'Dates and Times Made Easy with {lubridate}', *Journal of Statistical Software*, 40(3), pp. 1–25. Available at: https://www.jstatsoft.org/v40/i03/.

Haarer, A.E. (1956) ‘Modern coffee production.’, *Modern coffee production.* [Preprint].

Haggar, J. and Schepp, K. (2012) ‘Coffee and Climate Change Impacts and options for adaption in Brazil, Guatemala, Tanzania and Vietnam Coffee and Climate Change Impacts and options for adaption in Brazil, Guatemala, Tanzania and Vietnam’, *National Resources Institute*, 4. Available at: www.nri.org (Accessed: 13 July 2021).

Hannah, L. *et al.* (2013) ‘Global Climate Change Adaptation Priorities for Biodiversity and Food Security’, *PLOS ONE*, 8(8), p. e72590. doi:10.1371/JOURNAL.PONE.0072590.

Hansen, M.C. *et al.* (2013) ‘High-Resolution Global Maps of 21st-Century Forest Cover Change’, *Science*, 342(6160), pp. 850–853. doi:10.1126/SCIENCE.1244693.

Hardt, E. *et al.* (2015) ‘Does certification improve biodiversity conservation in Brazilian coffee farms?’, *Forest Ecology and Management*, 357, pp. 181–194.

Hasui, É. *et al.* (2018) ‘ATLANTIC BIRDS: a data set of bird species from the Brazilian Atlantic Forest’, *Ecology*, 99(2), pp. 497–497. doi:10.1002/ecy.2119.

Hatfield, J.H. *et al.* (2020) ‘Mediation of area and edge effects in forest fragments by adjacent land use’, *Conservation Biology*, 34(2), pp. 395–404. doi:10.1111/COBI.13390.

Haustein, K. *et al.* (2017) ‘A real-time Global Warming Index’, *Scientific Reports 2017 7:1*, 7(1), pp. 1–6. doi:10.1038/s41598-017-14828-5.

Hershey, S. *et al.* (2017) ‘CNN architectures for large-scale audio classification’, *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 131–135. doi:10.1109/ICASSP.2017.7952132.

Hijmans, R.J. *et al.* (2005) ‘Very high resolution interpolated climate surfaces for global land areas’, *International journal of climatology*, 25(15), pp. 1965–1978.

Hilje, B., Stack, S. and Sánchez-Azofeifa, A. (2017) ‘Lianas Abundance is Positively Related with the Avian Acoustic Community in Tropical Dry Forests’, *Forests 2017, Vol. 8,* *Page 311*, 8(9), p. 311. doi:10.3390/F8090311.

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*. Edited by N. Isaac, 9(5), pp. 1199–1211. doi:10.1111/2041-210X.12955.

Hill, A.P. *et al.* (2019) 'AudioMoth: A low-cost acoustic device for monitoring biodiversity and the environment', *HardwareX*, 6, p. e00073. doi:10.1016/j.ohx.2019.e00073.

Hoegh-Guldberg, O. (2012) 'The adaptation of coral reefs to climate change: Is the Red Queen being outpaced?', *Scientia Marina*, 76(2), pp. 403–408. doi:10.3989/SCIMAR.03660.29A.

Huang, C. *et al.* (2009) 'Assessment of Paraguay's forest cover change using Landsat observations', *Global and Planetary Change*, 67(1–2), pp. 1–12. doi:10.1016/j.gloplacha.2008.12.009.

Hurtt, G.C. *et al.* (2011) 'Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands', *Climatic Change 2011 109:1*, 109(1), pp. 117–161. doi:10.1007/S10584-011-0153-2.

IBGE (2020) *2020 Population Census* . Available at: https://www.ibge.gov.br/en/statistics/social/population/22836-2020-census-censo4.html?=&t=series-historicas (Accessed: 7 July 2021).

ICO (2020) 'Daily Coffee Prices', *ICO Statistics* [Preprint]. Available at: https://www.ico.org/coffee_prices.asp.

IHCAFE (2020) *Los Retos de la Caficultura Hondureña ante el COVID-19.* Available at: https://promecafe.net/wp-content/uploads/2020/05/Analisis-de-la-caficultura-hondureña-ante-el-COVID-19-01052020.pdf.

Imbach, P. *et al.* (2017) 'Coupling of pollination services and coffee suitability under climate change', *Proceedings of the National Academy of Sciences*, 114(39), pp. 10438–10442. Available at: http://www.pnas.org/content/pnas/114/39/10438.full.pdf (Accessed: 16 February 2018).

IPCC (2018) *Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* . Edited by V. Masson-Delmotte et al. Available at: https://archive.ipcc.ch/report/sr15/pdf/sr15_citation.pdf (Accessed: 12 July 2021).

Isbell, F. *et al.* (2017) 'Linking the influence and dependence of people on biodiversity across scales', *Nature*. Nature Publishing Group, pp. 65–72. doi:10.1038/nature22899.

Izquierdo, A.E., De Angelo, C.D. and Aide, T.M. (2008) 'Thirty years of human demography and land-use change in the Atlantic Forest of Misiones, Argentina: An evaluation of the forest transition model', *Ecology and Society*, 13(2). doi:10.5751/es-02377-130203.

Jacob, D.J. (1999) *Introduction to Atmospheric Chemistry*, *Introduction to Atmospheric Chemistry*. Princeton University Press. doi:10.1515/9781400841547/HTML.

Jaramillo, J. *et al.* (2009) 'Thermal Tolerance of the Coffee Berry Borer Hypothenemus hampei: Predictions of Climate Change Impact on a Tropical Insect Pest', *PLOS ONE*, 4(8), p. e6487. doi:10.1371/JOURNAL.PONE.0006487.

Jaramillo, J. *et al.* (2011) 'Some like it hot: the influence and implications of climate change on coffee berry borer (Hypothenemus hampei) and coffee production in East Africa', *PLoS One*, 6(9), p. e24528.

Jena, P.R., Stellmacher, T. and Grote, U. (2017) 'Can coffee certification schemes increase incomes of smallholder farmers? Evidence from Jinotega, Nicaragua', *Environment, Development and Sustainability*, 19(1), pp. 45–66. doi:10.1007/S10668-015-9732-

0/TABLES/10.

Jha, S. *et al.* (2014) ‘Shade Coffee: Update on a Disappearing Refuge for Biodiversity’, *BioScience*, 64(5), pp. 416–428. doi:10.1093/BIOSCI/BIU038.

Johnson, C.N. *et al.* (2017) ‘Biodiversity losses and conservation responses in the Anthropocene’, *Science*. American Association for the Advancement of Science, pp. 270–275. doi:10.1126/science.aam9317.

Joly, C.A., Metzger, J.P. and Tabarelli, M. (2014) ‘Experiences from the Brazilian Atlantic Forest: ecological findings and conservation initiatives’, *New Phytologist*, 204(3), pp. 459–473. doi:10.1111/nph.12989.

Jones, G. and Teeling, E.C. (2006) ‘The evolution of echolocation in bats’, *Trends in Ecology & Evolution*, 21(3), pp. 149–156. doi:10.1016/J.TREE.2006.01.001.

Jordani, M.X., Hasui, É. and Silva, V.X. da (2015) ‘Natural enemies depend on remnant habitat size in agricultural landscapes’, *Journal of Forestry Research 2015 26:2*, 26(2), pp. 469–477. doi:10.1007/S11676-015-0043-Y.

Kahl, S. *et al.* (2021) ‘BirdNET: A deep learning solution for avian diversity monitoring’, *Ecological Informatics*, 61. doi:10.1016/j.ecoinf.2021.101236.

Karp, D.S. *et al.* (2013) ‘Forest bolsters bird abundance, pest control and coffee yield’, *Ecology Letters*, 16(11), pp. 1339–1347. doi:10.1111/ele.12173.

Karungi, J. *et al.* (2018) ‘Elevation and cropping system as drivers of microclimate and abundance of soil macrofauna in coffee farmlands in mountainous ecologies’, *Applied Soil Ecology*, 132, pp. 126–134. doi:10.1016/j.apsoil.2018.08.003.

Kasten, E.P. *et al.* (2012) ‘The remote environmental assessment laboratory’s acoustic library: An archive for studying soundscape ecology’, *Ecological Informatics*, 12, pp. 50–67. doi:10.1016/j.ecoinf.2012.08.001.

Kath, J. *et al.* (2020) ‘Not so robust: Robusta coffee production is highly sensitive to temperature’, *Global Change Biology*, 26(6), pp. 3677–3688. doi:10.1111/gcb.15097.

Kleyn, T., Kaizer, M. da C. and Passos, L.F. (2021) ‘Sharing sound: Avian acoustic niches in the Brazilian Atlantic Forest’, *Biotropica*, 53(2), pp. 658–670. doi:10.1111/BTP.12907.

Kobayashi, Y., Okada, K. ichi and Mori, A.S. (2019) ‘Reconsidering biodiversity hotspots based on the rate of historical land-use change’, *Biological Conservation*, 233, pp. 268–275. doi:10.1016/j.biocon.2019.02.032.

Koh, I. *et al.* (2020) ‘Climate risks to Brazilian coffee production’, *Environmental Research Letters*, 15(10), p. 104015. doi:10.1088/1748-9326/ABA471.

Kreft, S. *et al.* (2013) ‘Who Suffers Most From Extreme Weather Events? Weather-related Loss Events in 2013 and 1994 to 2013’. Available at: www.germanwatch.org (Accessed: 8 July 2021).

Kunz, T.H. *et al.* (2011) ‘Ecosystem services provided by bats’, *Annals of the New York Academy of Sciences*, 1223(1), pp. 1–38. doi:10.1111/J.1749-6632.2011.06004.X.

Laderach, P. *et al.* (2011) ‘Predicted impact of climate change on coffee supply chains’, in *The economic, social and political elements of climate change*. Springer, pp. 703–723.

Läderach, P. *et al.* (2017) ‘Climate change adaptation of coffee production in space and time’, *Climatic Change*, 141(1), pp. 47–62. doi:10.1007/s10584-016-1788-9.

Lapenta, M.J. and Procópio-de-Oliveira, P. (2008) ‘Some Aspects of Seed Dispersal Effectiveness of Golden Lion Tamarins (Leontopithecus Rosalia) in a Brazilian Atlantic Forest’:, *http://dx.doi.org/10.1177/194008290800100205*, 1(2), pp. 122–139. doi:10.1177/194008290800100205.

Latini, A.O. *et al.* (2020) ‘Reconciling coffee productivity and natural vegetation conservation in an agroecosystem landscape in Brazil’, *Journal for Nature Conservation*, 57, p. 125902. doi:10.1016/J.JNC.2020.125902.

Laurance, W.F. (2009) ‘Conserving the hottest of the hotspots’, *Biological Conservation*, 142(6), p. 1137. doi:10.1016/J.BIOCON.2008.10.011.

Laurindo, R.S. *et al.* (2020) ‘Feeding habits define habitat use by bats in an agricultural landscape of the Atlantic Forest’, *Revista mexicana de biodiversidad*, 91. doi:10.22201/IB.20078706E.2020.91.3223.

Lei, M. and Dong, D. (2016) ‘Phylogenomic analyses of bat subordinal relationships based on transcriptome data’, *Scientific Reports*, 6. doi:10.1038/SREP27726.

Lemen, C. *et al.* (2015) ‘The Problem of Low Agreement among Automated Identification Programs for Acoustical Surveys of Bats’, *Western North American Naturalist*, 75(2), pp. 218–225. doi:10.3398/064.075.0210.

Lemes, P., Melo, A.S. and Loyola, R.D. (2014) ‘Climate change threatens protected areas of the Atlantic Forest’, *Biodiversity and Conservation*, 23(2), pp. 357–368. doi:10.1007/s10531-013-0605-2.

Li, H. *et al.* (2007) ‘Demand for rubber is causing the loss of high diversity rain forest in SW China’, in *Plant Conservation and Biodiversity*. Springer Netherlands, pp. 157–171. doi:10.1007/978-1-4020-6444-9_11.

Librán-Embid, F., De Coster, G. and Metzger, J.P. (2017) ‘Effects of bird and bat exclusion on coffee pest control at multiple spatial scales’, *Landscape Ecology*, 32(9), pp. 1907–1920. doi:10.1007/S10980-017-0555-2/FIGURES/6.

Liebig, T. *et al.* (2019) ‘Interactive effects of altitude, microclimate and shading system on coffee leaf rust’, *Journal of Plant Interactions*, 14(1), pp. 407–415. doi:10.1080/17429145.2019.1643934/SUPPL_FILE/TJPI_A_1643934_SM1486.DOCX.

Lima, M.M. and Mariano-Neto, E. (2014) ‘Extinction thresholds for Sapotaceae due to forest cover in Atlantic Forest landscapes’, *Forest Ecology and Management*, 312, pp. 260–270. doi:10.1016/J.FORECO.2013.09.003.

Lin, B.B. (2007) ‘Agroforestry management as an adaptive strategy against potential microclimate extremes in coffee agriculture’, *Agricultural and Forest Meteorology*, 144(1–2), pp. 85–94. Available at: https://ac.els-cdn.com/S0168192307000548/1-s2.0-S0168192307000548-main.pdf?_tid=f0193730-b8a1-11e7-99e1-00000aab0f27&acdnat=1508839313_822a7bec6f96d501ddc7342fb8a150f0 (Accessed: 24 October 2017).

Lin, B.B. (2010) ‘The role of agroforestry in reducing water loss through soil evaporation and crop transpiration in coffee agroecosystems’, *Agricultural and Forest Meteorology*, 150(4), pp. 510–518.

Lin, H. *et al.* (2020) ‘Forests buffer thermal fluctuation better than non-forests’, *Agricultural and Forest Meteorology*, 288–289, p. 107994. doi:10.1016/j.agrformet.2020.107994.

Lintott, P.R. *et al.* (2013) ‘Testing the effectiveness of surveying techniques in determining bat community composition within woodland’, *Wildlife Research*, 40(8), pp. 675–684. doi:10.1071/WR13153.

Lion, M.B. *et al.* (2016) ‘The Conservation Value of Small Fragments For Atlantic Forest Reptiles’, *Biotropica*, 48(2), pp. 265–275. doi:10.1111/btp.12277.

Lôbo, D. *et al.* (2011) ‘Forest fragmentation drives Atlantic forest of northeastern Brazil to biotic homogenization’, *Diversity and Distributions*, 17(2), pp. 287–296. doi:10.1111/j.1472-4642.2010.00739.x.

Lopes, L.C., Roberto, M. and De Brito, M. (2018) ‘What do insects, anurans, birds, and mammals have to say about soundscape indices in a tropical savanna’, *Article in Journal of Ecoacoustics* [Preprint]. doi:10.22261/JEA.PVH6YZ.

Loyola, R.D. *et al.* (2014) ‘Clade-specific consequences of climate change to amphibians in Atlantic Forest protected areas’, *Ecography*, 37(1), pp. 65–72. doi:10.1111/j.1600-0587.2013.00396.x.

Mace, G.M. (2011) ‘Drivers of Biodiversity Change’, in Leader-Williams, N., Adams, W.M., and Smith, R.J. (eds) *Trade-offs in Conservation: Deciding What to Save* . John Wiley & Sons, pp. 349–364. Available at: https://books.google.no/books?hl=en&lr=&id=ju30VGWm1pwC&oi=fnd&pg=PA349&dq=drivers+of+biodiversity+loss&ots=dTlwX_jgAo&sig=zNz8KcaozKJkgN5zxSzED0Rc0Wc&redir_esc=y (Accessed: 29 June 2021).

Machado, F. and Anderson, K. (2016) *Brazil’s new Forest Code: A guide for decision-makers in supply chains and governments*. Available at: www.panda.org/amazon (Accessed: 9 September 2020).

MacSwiney G., M.C., Clarke, F.M. and Racey, P.A. (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(5), pp. 1364–1371. doi:10.1111/j.1365-2664.2008.01531.x.

Maechler, M. *et al.* (2022) ‘cluster: Cluster Analysis Basics and Extensions’. Available at: https://orcid.org/0000-0001-9143-4880 (Accessed: 20 April 2022).

Magnago, L.F.S. *et al.* (2015) ‘Microclimatic conditions at forest edges have significant impacts on vegetation structure in large Atlantic forest fragments’, *Biodiversity and Conservation*, 24(9), pp. 2305–2318. doi:10.1007/s10531-015-0961-1.

Magrach, A. and Ghazoul, J. (2015) ‘Climate and Pest-Driven Geographic Shifts in Global Coffee Production: Implications for Forest Cover, Biodiversity and Carbon Storage’, *PLOS ONE*. Edited by F. Moreira, 10(7), p. e0133071. doi:10.1371/journal.pone.0133071.

Marques, I. *et al.* (2021) ‘A Transcriptomic Approach to Understanding the Combined Impacts of Supra-Optimal Temperatures and CO2 Revealed Different Responses in the Polyploid Coffea arabica and Its Diploid Progenitor C. canephora’, *International*

Journal of Molecular Sciences 2021, Vol. 22, Page 3125, 22(6), p. 3125. doi:10.3390/IJMS22063125.

Martensen, A.C. *et al.* (2012) 'Associations of Forest Cover, Fragment Area, and Connectivity with Neotropical Understory Bird Species Richness and Abundance', *Conservation Biology*, 26(6), pp. 1100–1111. doi:10.1111/j.1523-1739.2012.01940.x.

Martinelli, L.A. *et al.* (2010) 'Agriculture in Brazil: Impacts, costs, and opportunities for a sustainable future', *Current Opinion in Environmental Sustainability*, pp. 431–438. doi:10.1016/j.cosust.2010.09.008.

Martini, A.M.Z. *et al.* (2007) 'A hot-point within a hot-spot: A high diversity site in Brazil's Atlantic Forest', *Biodiversity and Conservation*, 16(11), pp. 3111–3128. doi:10.1007/s10531-007-9166-6.

Maxim Integrated Products, I. (2019) *DS18B20*. Available at: www.maximintegrated.com (Accessed: 8 July 2021).

Mendes, C.B. and Prevedello, J.A. (2020) 'Does habitat fragmentation affect landscape-level temperatures? A global analysis', *Landscape Ecology*, 35(8), pp. 1743–1756. doi:10.1007/s10980-020-01041-5.

Metcalf, O.C., Barlow, J., Devenish, C., *et al.* (2021) 'Acoustic indices perform better when applied at ecologically meaningful time and frequency scales', *Methods in Ecology and Evolution*, 12(3), pp. 421–431. doi:10.1111/2041-210X.13521.

Metcalf, O.C., Barlow, J., Marsden, S., *et al.* (2021) 'Optimizing tropical forest bird surveys using passive acoustic monitoring and high temporal resolution sampling', *Remote Sensing in Ecology and Conservation*. Edited by N. Pettorelli and C. Astaras, p. rse2.227. doi:10.1002/RSE2.227.

Metzger, J.P. (2009) 'Conservation issues in the Brazilian Atlantic forest', *Biological Conservation*, 142(6), pp. 1138–1140. doi:10.1016/j.biocon.2008.10.012.

Meyer, C.F.J., Struebig, M.J. and Willig, M.R. (2015) 'Responses of Tropical Bats to Habitat Fragmentation, Logging, and Deforestation', *Bats in the Anthropocene: Conservation of Bats in a Changing World*, p. 63. doi:10.1007/978-3-319-25220-9_4.

Milligan, M.C. *et al.* (2016) 'Quantifying pest control services by birds and ants in {K}enyan coffee farms', *Biological conservation*, 194, pp. 58–65.

Mira, C. *et al.* (2011) 'The SinBiota 2.0 biodiversity information system', *Proceedings - 2011 7th IEEE International Conference on eScience, eScience 2011*, pp. 142–149. doi:10.1109/ESCIENCE.2011.28.

Mittermeier, R.A. *et al.* (2005) 'A Brief History of Biodiversity Conservation in Brazil', *Conservation Biology*, 19(3), pp. 601–607. doi:10.1111/j.1523-1739.2005.00709.x.

Moguel, P. and Toledo, V.M. (1999) 'Biodiversity conservation in traditional coffee systems of Mexico', *Conservation biology*, 13(1), pp. 11–21.

Morante-Filho, J.C. *et al.* (2015) 'Birds in anthropogenic landscapes: The responses of ecological groups to forest loss in the Brazilian Atlantic forest', *PLoS ONE*, 10(6). doi:10.1371/journal.pone.0128923.

Murphy, G.E.P. and Romanuk, T.N. (2014) 'A meta-analysis of declines in local species richness from human disturbances', *Ecology and Evolution*, 4(1), pp. 91–103. doi:10.1002/ece3.909.

Murray, K.L., Britzke, E.R. and Robbins, L.W. (2001) 'Variation in Search-Phase Calls of Bats ', *Journal of Mammalogy*, 82(3), pp. 728–737. Available at: https://academic.oup.com/jmammal/article/82/3/728/2372701 (Accessed: 23 August 2021).

Muschler, R. (1997) 'Shade or sun for ecologically sustainable coffee production: a summary of environmental key factors', *III Semana Cient{\'\i}fica del Centro Agronómico Tropical de Investigación y Enseñanza (CATIE). Turrialba*, pp. 109–112.

Mussatto, S.I. *et al.* (2011) ‘Production, composition, and application of coffee and its industrial residues’, *Food and Bioprocess Technology*, 4(5), p. 661.

Muylaert, R. d. L. *et al.* (2017) ‘ATLANTIC BATS: a data set of bat communities from the Atlantic Forests of South America’, *Ecology*, 98(12), pp. 3227–3227. doi:10.1002/ecy.2007.

Myers, N. *et al.* (2000) ‘Biodiversity hotspots for conservation priorities’, *Nature*, 403(6772), pp. 853–858. doi:10.1038/35002501.

National Climatic Data Center *et al.* (2017) *Global Surface Summary of the Day -GSOD*. Available at: https://www7.ncdc.noaa.gov/CDO/cdosubqueryrouter.cmd (Accessed: 24 October 2017).

Naumann, S. *et al.* (2011) *Assessment of the potential of ecosystem-based approaches to climate change adaptation and mitigation in Europe: Final report to the European Commission*.

Newbold, T. *et al.* (2015) ‘Global effects of land use on local terrestrial biodiversity’, *Nature*, 520(7545), pp. 45–50. doi:10.1038/nature14324.

Newbold, T. *et al.* (2016) ‘Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment’, *Science*, 353(6296), pp. 291–288. doi:10.1126/science.aaf2201.

Nocera, T.S. *et al.* (2019) ‘Let’s Agree to Disagree: Comparing Auto-Acoustic Identification Programs for Northeastern Bats’, *Journal of Fish and Wildlife Management*, 10(2), pp. 346–361. doi:10.3996/102018-JFWM-090.

Oksanen, J. *et al.* (2019) *Package ‘vegan’ Title Community Ecology Package Version 2.5-6*.

Oliveira, U. *et al.* (2017) ‘Biodiversity conservation gaps in the Brazilian protected areas’, *Scientific Reports*, 7(1). doi:10.1038/s41598-017-08707-2.

Opedal, Ø.H., Armbruster, W.S. and Graae, B.J. (2014) ‘Linking small-scale topography with microclimate, plant species diversity and intra-specific trait variation in an alpine landscape', *http://dx.doi.org/10.1080/17550874.2014.987330*, 8(3), pp. 305–315. doi:10.1080/17550874.2014.987330.

Open Acoustic Devices (2022) *AudioMoth | Open Acoustic Devices*. Available at: https://www.openacousticdevices.info/audiomoth (Accessed: 3 October 2022).

Ortiz, A.M.D. *et al.* (2021) 'A review of the interactions between biodiversity, agriculture, climate change, and international trade: research and policy priorities', *One Earth*, 4(1), pp. 88–101. doi:10.1016/J.ONEEAR.2020.12.008.

Osorio, N. (2002) *Global coffee crisis: A threat to sustainable development*. Available at: www.ico.org (Accessed: 7 July 2021).

Ovalle-Rivera, O. *et al.* (2015a) 'Projected shifts in Coffea arabica suitability among major global producing regions due to climate change', *PloS one*, 10(4), p. e0124155.

Ovalle-Rivera, O. *et al.* (2015b) 'Projected Shifts in Coffea arabica Suitability among Major Global Producing Regions Due to Climate Change', *PLOS ONE*. Edited by R.D. Loyola, 10(4), p. e0124155. doi:10.1371/journal.pone.0124155.

Pardini, R. *et al.* (2009) 'The challenge of maintaining Atlantic forest biodiversity: A multi-taxa conservation assessment of specialist and generalist species in an agro-forestry mosaic in southern Bahia', *Biological Conservation*, 142(6), pp. 1178–1190. doi:10.1016/J.BIOCON.2009.02.010.

Pardini, R. *et al.* (2010) 'Beyond the Fragmentation Threshold Hypothesis: Regime Shifts in Biodiversity Across Fragmented Landscapes', *PLOS ONE*, 5(10), p. e13666. doi:10.1371/JOURNAL.PONE.0013666.

Pardo, J.M. *et al.* (2022) 'Predicting poaching hotspots in the largest remnant of the Atlantic Forest by combining passive acoustic monitoring and occupancy models', *Biological Conservation*, 272, p. 109600. doi:10.1016/J.BIOCON.2022.109600.

Parra-Sanchez, E. and Banks-Leite, C. (2020) ‘The magnitude and extent of edge effects on vascular epiphytes across the Brazilian Atlantic Forest’, *Scientific Reports 2020 10:1*, 10(1), pp. 1–11. doi:10.1038/s41598-020-75970-1.

Paviolo, A. *et al.* (2016) ‘A biodiversity hotspot losing its top predator: The challenge of jaguar conservation in the Atlantic Forest of South America’, *Scientific Reports*, 6. doi:10.1038/srep37147.

Pecl, G.T. *et al.* (2017) ‘Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being’, *Science*. American Association for the Advancement of Science. doi:10.1126/science.aai9214.

Pedro, A.R.S. and Simonetti, J.A. (2014) ‘Variation in search-phase calls of Lasiurus varius (Chiroptera: Vespertilionidae) in response to different foraging habitats’, *Journal of Mammalogy*, 95(5), pp. 1004–1010. doi:10.1644/13-MAMM-A-327.

Pereira, C.G., Falcão, F. and Bernard, E. (2022) ‘One size doesn’t fit all: Singularities in bat species richness and activity patterns in wind-energy complexes in Brazil and implications for environmental assessment’, *Zoologia (Curitiba)*, 39. doi:10.1590/S1984-4689.V39.E21041.

Pereira, L.C.D.S.M., Oliveira, C.C.C. and Torezan, J.M.D. (2013) ‘Woody Species Regeneration in Atlantic Forest Restoration Sites Depends on Surrounding Landscape’, *Natureza & Conservação*, 11(2), pp. 138–144. doi:10.4322/natcon.2013.022.

Pérez-Granados, C. *et al.* (2019) ‘Vocal activity rate index: a useful method to infer terrestrial bird abundance with acoustic monitoring’, *Ibis*, 161(4), pp. 901–907. doi:10.1111/IBI.12728.

Pezzopane, J.R.M. *et al.* (2011) ‘Microclimate in coffee plantation grown under grevillea trees shading’, *Acta Scientiarum - Agronomy*, 33(2), pp. 201–206. doi:10.4025/actasciagron.v33i2.7065.

Philpott, S.M. *et al.* (2007) 'Field-testing ecological and economic benefits of coffee certification programs', *Conservation Biology*, 21(4), pp. 975–985.

Philpott, S.M. *et al.* (2008) 'Biodiversity loss in Latin American coffee landscapes: review of the evidence on ants, birds, and trees', *Conservation Biology*, 22(5), pp. 1093–1105.

Pijanowski, B.C. *et al.* (2011) 'Soundscape Ecology: The Science of Sound in the Landscape', *BioScience*, 61(3), pp. 203–216. doi:10.1525/bio.2011.61.3.6.

Pinheiro, J. *et al.* (2020) *Package 'nlme' Title Linear and Nonlinear Mixed Effects Models*. Available at: https://bugs.r-project.org (Accessed: 23 September 2020).

Pinto, S.R.R. *et al.* (2010) 'Landscape attributes drive complex spatial microclimate configuration of Brazilian Atlantic forest fragments', *Tropical Conservation Science*, 3(4), pp. 389–402. doi:10.1177/194008291000300404.

Piratelli, A. *et al.* (2008) 'Searching for bioindicators of forest fragmentation: passerine birds in the Atlantic forest of southeastern Brazil', *Brazilian Journal of Biology*, 68(2), pp. 259–268. doi:10.1590/S1519-69842008000200006.

Pocock, M.J.O. *et al.* (2018) 'A Vision for Global Biodiversity Monitoring With Citizen Science', in *Advances in Ecological Research*. Academic Press Inc., pp. 169–223. doi:10.1016/bs.aecr.2018.06.003.

Pompeu, P. *et al.* (2009) 'Classificação e apresentação dos parâmetros climáticos de Poços de Caldas-MG.', in *Ill Congresso Latino Americano de Ecologia, São Lourenço-MG.*

Prayogo, C. *et al.* (2019) 'Soil macrofauna diversity and structure under different management of pine-coffee agroforestry system', *Journal of Degraded and Mining Lands Management*, 6(3), pp. 1727–1736. doi:10.15243/jdmlm.2019.063.1727.

Püttker, T. *et al.* (2020) 'Indirect effects of habitat loss via habitat fragmentation: A cross-taxa analysis of forest-dependent species', *Biological Conservation*, 241(December 2019), p. 108368. doi:10.1016/j.biocon.2019.108368.

QGIS Development Team (2021) ‘QGIS Geographic Information System’. Available at: http://www.qgis.org.

R Core Team (2020) ‘R: A Language and Environment for Statistical Computing’. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.r-project.org/.

Rahn, E. *et al.* (2018) ‘Exploring adaptation strategies of coffee production to climate change using a process-based model’, *Ecological Modelling*, 371, pp. 76–89. doi:10.1016/j.ecolmodel.2018.01.009.

Rakocevic, M. *et al.* (2018) ‘Structural and functional changes in coffee trees after 4 years under free air CO2 enrichment’, *Annals of Botany*, 121(5), pp. 1065–1078. doi:10.1093/aob/mcy011.

Ramos, F.N. *et al.* (2019) ‘ATLANTIC EPIPHYTES: a data set of vascular and non-vascular epiphyte plants and lichens from the Atlantic Forest’, *Ecology*, 100(2). doi:10.1002/ecy.2541.

Regolin, A.L. *et al.* (2017) ‘Forest cover influences occurrence of mammalian carnivores within Brazilian Atlantic Forest’, *Journal of Mammalogy*, 98(6), pp. 1721–1731. doi:10.1093/JMAMMAL/GYX103.

Rezende, C.L. *et al.* (2018) ‘From hotspot to hopespot: An opportunity for the Brazilian Atlantic Forest’, *Perspectives in Ecology and Conservation*, 16(4), pp. 208–214. doi:10.1016/J.PECON.2018.10.002.

Ribeiro, M.C. *et al.* (2009) ‘The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation’, *Biological Conservation*, 142(6), pp. 1141–1153. doi:10.1016/J.BIOCON.2009.02.021.

Ribeiro, M.C. *et al.* (2011) ‘The Brazilian Atlantic Forest: A Shrinking Biodiversity Hotspot’, in *Biodiversity Hotspots*. Springer Berlin Heidelberg, pp. 405–434.

doi:10.1007/978-3-642-20992-5_21.

Rigal, C. *et al.* (2020) ‘Coffee production during the transition period from monoculture to agroforestry systems in near optimal growing conditions, in Yunnan Province’, *Agricultural Systems*, 177, p. 102696. doi:10.1016/J.AGSY.2019.102696.

Roberts, D.W., David, M. and Roberts, W. (2019) *Package ‘labdsv’ Title Ordination and Multivariate Analysis for Ecology*.

Robinson, W.D., Lees, A.C. and Blake, J.G. (2018) ‘Surveying tropical birds is much harder than you think: a primer of best practices’, *Biotropica*, 50(6), pp. 846–849. doi:10.1111/BTP.12608.

Rocchini, D. *et al.* (2018) ‘Measuring β-diversity by remote sensing: A challenge for biodiversity monitoring’, *Methods in Ecology and Evolution*, 9(8), pp. 1787–1798. doi:10.1111/2041-210X.12941.

Rocha, M.F., Passamani, M. and Louzada, J. (2011) ‘A Small Mammal Community in a Forest Fragment, Vegetation Corridor and Coffee Matrix System in the Brazilian Atlantic Forest’, *PLOS ONE*, 6(8), p. e23312. doi:10.1371/JOURNAL.PONE.0023312.

Rodrigues, R.C. *et al.* (2019) ‘ATLANTIC BIRD TRAITS : a data set of bird morphological traits from the Atlantic forests of South America’, *Ecology*, 100(6), p. e02647. doi:10.1002/ecy.2647.

Roe, P. *et al.* (2018) ‘Catching toad calls in the cloud: Commodity edge computing for flexible analysis of big sound data’, *Proceedings - IEEE 14th International Conference on eScience, e-Science 2018*, pp. 67–74. doi:10.1109/ESCIENCE.2018.00022.

Roemer, C., Julien, J.F. and Bas, Y. (2021) ‘An automatic classifier of bat sonotypes around the world’, *Methods in Ecology and Evolution*, 12(12), pp. 2432–2444. doi:10.1111/2041-210X.13721.

Roos, A.L. (2010) ‘Capturando Aves.’, *Ornitologia e conservação: ciência aplicada,* *técnicas de pesquisa e conservação.*, (January), pp. 77–104.

Russo, D., Ancillotto, L. and Jones, G. (2017) ‘Bats are still not birds in the digital era: echolocation call variation and why it matters for bat species identification1’, *https://doi.org/10.1139/cjz-2017-0089*, 96(2), pp. 63–78. doi:10.1139/CJZ-2017-0089.

Russo, D. and Jones, G. (2002) ‘Identification of twenty-two bat species (Mammalia: Chiroptera) from Italy by analysis of time-expanded recordings of echolocation calls’, *Journal of Zoology*, 258(1), p. S0952836902001231. doi:10.1017/S0952836902001231.

Sagastuy, M. and Krause, T. (2019) ‘Agroforestry as a Biodiversity Conservation Tool in the Atlantic Forest? Motivations and Limitations for Small-Scale Farmers to Implement Agroforestry Systems in North-Eastern Brazil’, *Sustainability 2019, Vol. 11, Page 6932*, 11(24), p. 6932. doi:10.3390/SU11246932.

Santos, B.A. *et al.* (2008) ‘Drastic erosion in functional attributes of tree assemblages in Atlantic forest fragments of northeastern Brazil’, *Biological Conservation*, 141(1), pp. 249–260. doi:10.1016/j.biocon.2007.09.018.

Santos, J.P. dos *et al.* (2018) ‘Atlantic butterflies: a data set of fruit-feeding butterfly communities from the Atlantic forests’, *Ecology*. Ecological Society of America, p. 2875. doi:10.1002/ecy.2507.

Santos, R. *et al.* (2011) ‘Coffee Yield and Microenvironmental Factors in a Native Tree Agroforestry System in Southeast Minas Gerais, Brazil’, *Journal of Sustainable Agriculture*, 36(1), pp. 54–68. doi:10.1080/10440046.2011.608468.

Scalabrin, S. *et al.* (2020) ‘A single polyploidization event at the origin of the tetraploid genome of Coffea arabica is responsible for the extremely low genetic variation in wild and cultivated germplasm’, *Scientific Reports*, 10(1), pp. 1–13. doi:10.1038/s41598-020-61216-7.

Scarano, F.R. (2009) ‘Plant communities at the periphery of the Atlantic rain forest: Rare-species bias and its risks for conservation’, *Biological Conservation*, 142(6), pp. 1201–1208. doi:10.1016/j.biocon.2009.02.027.

Scarano, F.R. and Ceotto, P. (2015) ‘Brazilian Atlantic forest: impact, vulnerability, and adaptation to climate change’, *Biodiversity and Conservation*, 24(9), pp. 2319–2331. doi:10.1007/s10531-015-0972-y.

Scarpelli, M.D.A., Ribeiro, M.C. and Teixeira, C.P. (2021) ‘What does Atlantic Forest soundscapes can tell us about landscape?’, *Ecological Indicators*, 121, p. 107050. doi:10.1016/J.ECOLIND.2020.107050.

Scheel-Ybert, R., Beauclair, M. and Buarque, A. (2014) ‘The forest people: Landscape and firewood use in the Araruama region, southeastern Brazil, during the late Holocene’, *Vegetation History and Archaeobotany*, 23(2), pp. 97–111. doi:10.1007/S00334-013-0397-Z/METRICS.

Schmeller, D.S. *et al.* (2017) ‘Building capacity in biodiversity monitoring at the global scale’, *Biodiversity and Conservation*. Springer Netherlands, pp. 2765–2790. doi:10.1007/s10531-017-1388-7.

Seddon, N. *et al.* (2019) ‘Grounding nature-based climate solutions in sound biodiversity science’, *Nature Climate Change*. Nature Publishing Group, pp. 84–87. doi:10.1038/s41558-019-0405-0.

Senior, R.A. *et al.* (2017) ‘A pantropical analysis of the impacts of forest degradation and conversion on local temperature’, *Ecology and Evolution*, 7(19), pp. 7897–7908. doi:10.1002/ece3.3262.

Sethi, S.S. *et al.* (2018) ‘Robust, real-time and autonomous monitoring of ecosystems with an open, low-cost, networked device’, *Methods in Ecology and Evolution*, 9(12), pp. 2383–2387. doi:10.1111/2041-210X.13089.

Sethi, S.S. *et al.* (2019) ‘Combining machine learning and a universal acoustic feature-set yields efficient automated monitoring of ecosystems’, *bioRxiv*, p. 865980. doi:10.1101/865980.

Sethi, S.S. *et al.* (2020) ‘Characterizing soundscapes across diverse ecosystems using a universal acoustic feature set’, *Proceedings of the National Academy of Sciences of the United States of America*, 117(29), pp. 17049–17055. doi:10.1073/pnas.2004702117.

Sethi, S.S. *et al.* (2022) ‘Biodegradable sensors are ready to transform autonomous ecological monitoring’, *Nature Ecology & Evolution 2022 6:9*, 6(9), pp. 1245–1247. doi:10.1038/s41559-022-01824-w.

Shamon, H. *et al.* (2021) ‘Using ecoacoustics metrices to track grassland bird richness across landscape gradients’, *Ecological Indicators*, 120, p. 106928. doi:10.1016/j.ecolind.2020.106928.

Sigmund, M. and Hrabina, M. (2021) ‘Efficient Feature Set Developed for Acoustic Gunshot Detection in Open Space’, *Elektronika ir Elektrotechnika*, 27(4), pp. 62–68. doi:10.5755/J02.EIE.28877.

da Silva, J.M.C. and Casteleti, C.H.M. (2003) ‘Status of the biodiversity of the Atlantic Forest of Brazil’, in Leal, C.G. and Câmara, I.G. (eds) *The Atlantic Forest of South America: Biodiversity Status, Threats, and Outlook*. Washington: Island Press, pp. 43–59.

Silva, M.H. and Beauvais, S.L. (2010) ‘Human health risk assessment of endosulfan. I: Toxicology and hazard identification’, *Regulatory Toxicology and Pharmacology*, 56(1), pp. 4–17.

Sinha, B.B. and Dhanalakshmi, R. (2022) ‘Recent advancements and challenges of Internet of Things in smart agriculture: A survey’, *Future Generation Computer Systems*, 126, pp. 169–184. doi:10.1016/J.FUTURE.2021.08.006.

Smith, J. (2018) ‘Coffee Landscapes: Specialty Coffee, Terroir, and Traceability in Costa Rica’, *Culture, Agriculture, Food and Environment*, 40(1), pp. 36–44. doi:10.1111/CUAG.12103.

Soares-Filho, B. *et al.* (2014) ‘Cracking Brazil’s Forest Code’, *Science*, 344(6182), pp. 363–364. doi:10.1126/science.1246663.

Solari, S. and Baker, R.J. (2007) *Mammal Species of the World: A Taxonomic and Geographic Reference by D. E. Wilson; D. M. Reeder*, *Journal of Mammalogy*. Oxford Academic. doi:10.1644/06-MAMM-R-422.1.

SOS Mata Atlantica and INPE (2019) *Atlas dos Remanescentes Florestais da Mata Atlântica*. Available at: http://mapas.sosma.org.br (Accessed: 21 July 2021).

Souza, Y. *et al.* (2019) ‘ATLANTIC MAMMALS : a data set of assemblages of medium- and large-sized mammals of the Atlantic Forest of South America’, *Ecology*, 100(10). doi:10.1002/ecy.2785.

Sparovek, G. *et al.* (2012) ‘The revision of the Brazilian Forest Act: increased deforestation or a historic step towards balancing agricultural development and nature conservation?’, *Environmental Science & Policy*, 16, pp. 65–72. doi:10.1016/J.ENVSCI.2011.10.008.

Staver, C. *et al.* (2001) ‘Designing pest-suppressive multistrata perennial crop systems : shade-grown coffee in Central America’, *Agroforestry Systems*, 53(2), pp. 151–170.

Steiman, S., Bittenbender, H.C. and Idol, T.W. (2007) ‘Analysis of Kaolin Particle Film Use and Its Application on Coffee’, *HortScience*, 42(7), pp. 1605–1608. Available at: https://journals.ashs.org/hortsci/view/journals/hortsci/42/7/article-p1605.xml (Accessed: 6 April 2021).

Stephenson, P.J. (2020) ‘Technological advances in biodiversity monitoring: applicability, opportunities and challenges’, *Current Opinion in Environmental Sustainability*, 45, pp.

36–41. doi:10.1016/j.cosust.2020.08.005.

Storlie, C. *et al.* (2014) ‘Stepping inside the niche: Microclimate data are critical for accurate assessment of species’ vulnerability to climate change’, *Biology Letters*, 10(9). doi:10.1098/rsbl.2014.0576.

Stowell, D. *et al.* (2018) *Automatic acoustic detection of birds through deep learning: the first Bird Audio Detection challenge*. Available at: https://arxiv.org/pdf/1807.05812.pdf (Accessed: 20 June 2019).

Stowell, D. (2022) ‘Computational bioacoustics with deep learning: a review and roadmap’, *PeerJ*, 10, p. e13152. doi:10.7717/PEERJ.13152/SUPP-1.

Suarez-Rubio, M., Ille, C. and Bruckner, A. (2018) ‘Insectivorous bats respond to vegetation complexity in urban green spaces’, *Ecology and Evolution*, 8(6), pp. 3240–3253. doi:10.1002/ECE3.3897.

Sugai, L.S.M. *et al.* (2019) ‘Terrestrial Passive Acoustic Monitoring: Review and Perspectives’, *BioScience*, 69(1), pp. 15–25. doi:10.1093/biosci/biy147.

Szewczak, J. (2015) ‘SonoBat^{TM} Software for Bat Call Analysis’. Available at: www.sonobat.com.

Szymański, P. *et al.* (2021) ‘Passive acoustic monitoring gives new insight into year-round duetting behaviour of a tropical songbird’, *Ecological Indicators*, 122, p. 107271. doi:10.1016/J.ECOLIND.2020.107271.

Tabarelli, M. *et al.* (2010) ‘Prospects for biodiversity conservation in the Atlantic Forest: Lessons from aging human-modified landscapes’, *Biological Conservation*, 143(10), pp. 2328–2340. doi:10.1016/j.biocon.2010.02.005.

Teixido, A.L. *et al.* (2020) ‘Major biases and knowledge gaps on fragmentation research in Brazil: Implications for conservation’, *Biological Conservation*, 251, p. 108749. doi:10.1016/J.BIOCON.2020.108749.

Tenzin, S. *et al.* (2017) 'Low cost weather station for climate-smart agriculture', *2017 9th International Conference on Knowledge and Smart Technology: Crunching Information of Everything, KST 2017*, pp. 172–177. doi:10.1109/KST.2017.7886085.

Toenies, M. and Rich, L.N. (2021) 'Advancing bird survey efforts through novel recorder technology and automated species identification', *California Fish and Wildlife*, 107(2), pp. 56–70. doi:10.51492/cfwj.107.5.

Toledo, V.M. and Moguel, P. (2012) 'Coffee and Sustainability: The Multiple Values of Traditional Shaded Coffee', *Journal of Sustainable Agriculture*, 36(3), pp. 353–377. doi:10.1080/10440046.2011.583719.

Towsey, M. *et al.* (2018) 'Long-duration, false-colour spectrograms for detecting species in large audio data-sets Peer review: Citation'. doi:10.22261/JEA.IUSWUI.

Trew, B.T. and Maclean, I.M.D. (2021) 'Vulnerability of global biodiversity hotspots to climate change', *Global Ecology and Biogeography*. Blackwell Publishing Ltd, pp. 768–783. doi:10.1111/geb.13272.

Tubiello, F.N. *et al.* (2015) 'The Contribution of Agriculture, Forestry and other Land Use activities to Global Warming, 1990-2012', *Global Change Biology*, 21(7), pp. 2655–2660. doi:10.1111/gcb.12865.

UCA Soppexcca (2011) *Jinotega, the capital of coffee in Nicaragua*. Available at: http://www.soppexcca.org/en/jinotega.html (Accessed: 8 July 2021).

Uezu, A. and Metzger, J.P. (2011) 'Vanishing bird species in the Atlantic Forest: Relative importance of landscape configuration, forest structure and species characteristics', *Biodiversity and Conservation*, 20(14), pp. 3627–3643. doi:10.1007/S10531-011-0154-5/TABLES/3.

Uezu, A. and Metzger, J.P. (2016) 'Time-Lag in Responses of Birds to Atlantic Forest Fragmentation: Restoration Opportunity and Urgency', *PLOS ONE*, 11(1), p.

e0147909. doi:10.1371/JOURNAL.PONE.0147909.

Umetsu, F. and Pardini, R. (2007) ‘Small mammals in a mosaic of forest remnants and anthropogenic habitats - Evaluating matrix quality in an Atlantic forest landscape’, *Landscape Ecology*, 22(4), pp. 517–530. doi:10.1007/s10980-006-9041-y.

United Nations (2015) *Transforming our world: the 2030 Agenda for Sustainable Development*. Available at: https://sustainabledevelopment.un.org/post2015/transformingourworld (Accessed: 7 September 2020).

Upendranadh, C. and Subbaiah, C.A. (2016) ‘Labour shortage in coffee plantation areas : Coping strategies of small growers in Kodagu district’, *Globalisation, Development and Plantation Labour in India*, pp. 167–195. doi:10.4324/9781315620510-7.

Upton, E. and Halfacree, J. (2014) *Raspberry Pi User Guide*. John Wiley & Sons, Ltd. Available at: https://books.google.es/books?hl=en&lr=&id=g-XhBQAAQBAJ&oi=fnd&pg=PA123&dq=raspberry+pi&ots=EmnqVtBUXs&sig=bBT3qV7KhVBBgSr2fkEDjSn_aBA&redir_esc=y#v=onepage&q=raspberry pi&f=false (Accessed: 3 May 2021).

USDA (2020) *Coffee: World Markets and Trade*. Available at: https://apps.fas.usda.gov/psdonline/circulars/coffee.pdf (Accessed: 31 March 2021).

USGS (2019) *USGS EROS Archive - Digital Elevation - Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global*. Available at: https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm-1-arc?qt-science_center_objects=0#qt-science_center_objects (Accessed: 3 August 2019).

Vaast, Philippe *et al.* (2016) ‘Coffee and Cocoa Production in Agroforestry—A Climate-Smart Agriculture Model’, *Climate Change and Agriculture Worldwide*, pp. 209–224.

doi:10.1007/978-94-017-7462-8_16.

Vale, M.M. *et al.* (2018) ‘Endemic birds of the Atlantic Forest: traits, conservation status, and patterns of biodiversity’, *Journal of Field Ornithology*, 89(3), pp. 193–206. doi:10.1111/JOFO.12256.

Valkila, J. (2009) ‘Fair Trade organic coffee production in {N}icaragua — sustainable development or a poverty trap?’, *Ecological Economics*, 68(12), pp. 3018–3025.

Vancine, M.H. *et al.* (2018) ‘ATLANTIC AMPHIBIANS: a data set of amphibian communities from the Atlantic Forests of South America’, *Ecology*, 99(7), pp. 1692–1692. doi:10.1002/ecy.2392.

Vegro, C.L.R. and de Almeida, L.F. (2020) ‘Global coffee market: Socio-economic and cultural dynamics’, in *Coffee Consumption and Industry Strategies in Brazil: A*. Woodhead Publishing, pp. 3–19. doi:10.1016/B978-0-12-814721-4.00001-9.

Venier, L.A. *et al.* (2012) ‘Evaluation of an automated recording device for monitoring forest birds’, *Wildlife Society Bulletin*, 36(1), pp. 30–39. doi:10.1002/WSB.88.

Verbist, B., Putra, A.E.D. and Budidarsono, S. (2005) ‘Factors driving land use change: Effects on watershed functions in a coffee agroforestry system in Lampung, Sumatra’, in *Agricultural Systems*. Elsevier BV, pp. 254–270. doi:10.1016/j.agsy.2005.06.010.

Verburg, R. *et al.* (2019) ‘An innovation perspective to climate change adaptation in coffee systems’, *Environmental Science & Policy*, 97, pp. 16–24. doi:10.1016/J.ENVSCI.2019.03.017.

Verhage, F.Y.F., Anten, N.P.R. and Sentelhas, P.C. (2017) ‘Carbon dioxide fertilization offsets negative impacts of climate change on Arabica coffee yield in Brazil’, *Climatic Change*, 144(4), pp. 671–685. doi:10.1007/s10584-017-2068-z.

Vijay, V. *et al.* (2016) ‘The impacts of oil palm on recent deforestation and biodiversity loss’, *PLoS ONE*, 11(7), p. e0159668. doi:10.1371/journal.pone.0159668.

Vold, S.T., Handel, C.M. and McNew, L.B. (2017) ‘Comparison of acoustic recorders and field observers for monitoring tundra bird communities’, *Wildlife Society Bulletin*, 41(3), pp. 566–576. doi:10.1002/WSB.785.

Volsi, B. *et al.* (2019) ‘The dynamics of coffee production in Brazil’, *PLoS ONE*, 14(7). doi:10.1371/journal.pone.0219742.

Waller, J.M., Bigger, M. and Hillocks, R.J. (2007) *Coffee pests, diseases and their management*. CABI.

Walters, C.L. *et al.* (2012) ‘A continental-scale tool for acoustic identification of European bats’, *Journal of Applied Ecology*. Edited by J. Minderman, 49(5), pp. 1064–1074. doi:10.1111/j.1365-2664.2012.02182.x.

Watson, J.E.M. *et al.* (2018) ‘The exceptional value of intact forest ecosystems’, *Nature Ecology and Evolution*. Nature Publishing Group, pp. 599–610. doi:10.1038/s41559-018-0490-x.

Wickham, H. *et al.* (2020) ‘dplyr: A Grammar of Data Manipulation’. Available at: https://cran.r-project.org/package=dplyr.

Williams-Guillén, K., Perfecto, I. and Vandermeer, J. (2008) ‘Bats limit insects in a neotropical agroforestry system.’, *Science*, 320(5872), p. 70. doi:10.1126/science.1152944.

Wimmer, J., Williamson, I. and Roe, P. (2014) ‘The use of acoustic indices to determine avian species richness in audio-recordings of the environment’, *Ecological Informatics*, 21, pp. 110–119. doi:10.1016/J.ECOINF.2013.11.007.

World Resources Institute (2005) ‘Millennium Ecosystem Assessment’, *Ecosystems and human well-being: Biodiversity synthesis* [Preprint].

Wowk, B. *et al.* (2001) ‘Echolocation behavior and signal plasticity in the Neotropical bat Myotis nigricans (Schinz, 1821) (Vespertilionidae): a convergent case with European species of Pipistrellus?', *Behavioral Ecology and Sociobiology 2001 50:4*, 50(4), pp. 317–328. doi:10.1007/S002650100379.

Wrigley, G. (1988) 'Coffee/Coffea/spp.(Rubiaceae)..', *Evolution of crop plants. 2. ed.*. [Preprint].

Yerlikaya, B.A., Ömezli, S. and Aydoğan, N. (2020) 'Climate Change Forecasting and Modeling for the Year of 2050', in *Environment, Climate, Plant and Vegetation Growth*. Springer International Publishing, pp. 109–122. doi:10.1007/978-3-030-49732-3_5.

Zamora-Gutierrez, V. *et al.* (2016) 'Acoustic identification of Mexican bats based on taxonomic and ecological constraints on call design', *Methods in Ecology and Evolution*. Edited by R. Freckleton, 7(9), pp. 1082–1091. doi:10.1111/2041-210X.12556.

Zamora-Gutierrez, V. *et al.* (2021) 'The Evolution of Acoustic Methods for the Study of Bats', pp. 43–59. doi:10.1007/978-3-030-54727-1_3.

Zurita, G.A. and Bellocq, M.I. (2010) 'Spatial patterns of bird community similarity: Bird responses to landscape composition and configuration in the Atlantic forest', *Landscape Ecology*, 25(1), pp. 147–158. doi:10.1007/s10980-009-9410-4.

Zuur, A.F. *et al.* (2009) 'Mixed Effects Modelling for Nested Data', in. Springer, New York, NY, pp. 101–142. doi:10.1007/978-0-387-87458-6_5.

Appendix A:

A.1 Permission to reproduce Ecology of the Atlantic Forest review within thesis

15th March 2022

Dear Oxford Bibliographies,

I am completing my PhD thesis at Imperial College London.

I seek your permission to reprint, in my thesis an extract from: [Damstra-Oddy & Banks-Leite “Ecology of the Atlantic Forest.” *Oxford Bibliographies* in "Ecology". Ed. David Gibson. New York: Oxford University Press, 2021. https://www.oxfordbibliographies.com/view/document/obo-9780199830060/obo-9780199830060-0233.xml?rskey=uoXvCO&result=71 . The extract to be reproduced is: main text.

I would like to include the extract in my thesis which will be added to Spiral, Imperial's institutional repository http://spiral.imperial.ac.uk/ and made available to the public under a Creative Commons Attribution-NonCommercial 4.0 International Licence.

If you are happy to grant me all the permissions requested, please return a signed copy of this letter. If you wish to grant only some of the permissions requested, please list these and then sign.

Yours sincerely,

Elise Damstra

EDamstra

Permission granted for the use requested above:

I confirm that I am the copyright holder of the extract above and hereby give permission to include it in your thesis which will be made available, via the internet, for non-commercial purposes under the terms of the user licence.

[please edit the text above if you wish to grant more specific permission]

Signed:

Name: Daniel Rogers

Organisation: Oxford University Press USA

Job title: Editorial Manager

A.2 Original publication

Oxford Bibliographies
Your Best Research Starts Here

Ecology of the Atlantic Forest

Elise Damstra-Oddy, Cristina Banks-Leite

LAST MODIFIED: 12 JANUARY 2021
DOI: 10.1093/OBO/9780199830060-0233

Introduction

Extending along the southern coast of Brazil, into Argentina and Paraguay, the Atlantic Forest is a domain that once covered 150 Mha and includes many distinct forest subtypes and ecosystems. Its large latitudinal (29˚) and altitudinal (0–2,800 m above sea level) range, as well as complex topography in the region, has created microclimates within forest subtypes, which has led to biodiversity specifically adapted to narrow ecological ranges. The region is incredibly species-rich and is home to charismatic or economically important species such as the black and golden lion tamarin, the red-browned Amazon parrot, and the highly prized palm heart from *Euterpe edulis*. Through widespread human-driven change dating back to the arrival of European settlers in 1500, this realm has been extensively reduced, fragmented, and modified. Nowadays, this region is home to about 130 million people (60 percent of the Brazilian population) and is responsible for producing 70 percent of Brazil's GDP, putting a strain on natural resources and providing challenges to conservation. Due to its high levels of endemic species coupled with a high threat of habitat loss and fragmentation, the Atlantic Forest has been identified as a "biodiversity hotspot." Numerous studies have assessed the effects of habitat transformation on biodiversity and the consensus is that the majority of species are negatively affected. It is puzzling however that few species have actually gone extinct in the wild, even if some extinctions might have gone undetected. Extinctions do not immediately follow habitat change, there is often a time lag of many decades between habitat transformation and extinction. This may suggest that many species in the Atlantic Forest are "living deads," as despite their presence the available habitat no longer supports their requirements. It also suggests that there is a window of opportunity to restoring the domain to avert extinctions before they are realized. Current research and policy actions are geared toward optimizing restoration and increasing the extent of native forest cover, bringing hope to the conservation of this unique domain.

General Overview of Forest Ecology and Forest Extent

Several good sources give detailed overviews of the forest ecology. Galindo-Leal and Câmara 2003 is a good general introduction to the history, biodiversity, and human impacts in the region as well as conservation management strategies; however, certain aspects of this text pertaining to current trends and conservation management are likely outdated. Metzger and Sodhi 2009, a special issue in *Biological Conservation*, focuses on conservation issues in the Atlantic Forest. Joly, et al. 2014 provides a comprehensive review of the history of disturbance, the ecology, the ongoing effects of fragmentation, and how climate change is impacting and will impact the Atlantic Forest. Rates of land use change in the Atlantic Forest tend to be monitored on a country-by-country basis, with remote sensing efforts in Brazil, Argentina, and Paraguay revealing different levels of deforestation, as seen in Azevedo, et al. 2018; Izquierdo, et al. 2008; and Huang, et al. 2009, respectively. In Brazil, low levels of deforestation are mostly matched by reforestation, which means that the amount of forest cover has either been stable or slowly increased in the past decades. The amount of remaining forest cover in Brazil has been measured by several groups and has been repeatedly updated since the 2000s as high-resolution satellite imagery becomes available. Galindo-Leal and Câmara 2003 reported that about 7–8 percent of Atlantic Forest still remained. Ribeiro, et al. 2009 calculated the existence of 11–16 percent of native vegetation, but when secondary forest fragments are excluded, the estimated remaining forest would stand at around 8 percent. More recently, using RapidEye imagery with 5 m of resolution, the authors of Rezende, et al. 2018 have shown that there is actually 26 percent of native vegetation. Because each group used different methods to assess the extent of forest cover, these estimates cannot be compared, and do not indicate that the amount of forest has increased over time.

Azevedo, T., C. M. Souza, J. Shimbo, and A. Alencar. 2018. *MapBiomas initiative: Mapping annual land cover and land use*

2018, Washington D.C.

MapBiomas is an excellent interactive tool to visualize high-resolution land use change in Brazil as well as a reliable source of free geographical data to use in research.

Galindo-Leal, C. G., and I. G. Câmara, eds. 2003. *The Atlantic Forest of South America: Biodiversity status, threats, and outlook*. Vol. 1. Washington, DC: Island Press.

This book provides an excellent initial overview of the history and threats to the Atlantic Forest, divided into sections by country.

Huang, C., S. Kim, K. Song, et al. 2009. Assessment of Paraguay’s forest cover change using Landsat observations. *Global and Planetary Change* 67.1–2: 1–12.

This analysis of Paraguay’s forest cover shows that most forests outside of protected areas were lost by the 2000s, demonstrating the importance of protected areas.

Izquierdo, A. E., C. D. De Angelo, and T. M. Aide. 2008. Thirty years of human demography and land-use change in the Atlantic Forest of Misiones, Argentina: An evaluation of the forest transition model. *Ecology and Society* 13.2: 3.

This paper examines changes in forest cover in Misiones, showing that there has been increased planting of pine and eucalyptus monocultures associated with a loss in natural forest.

Joly, C. A., J. P. Metzger, and M. Tabarelli. 2014. Experiences from the Brazilian Atlantic Forest: Ecological findings and conservation initiatives. *New Phytologist* 204.3: 459–473.

This article provides a good discussion, with the use of conceptual models, of how large-scale landscape ecological processes can help maintain biota, as well as providing a research agenda that would conserve biodiversity of tropical forests.

Metzger, J. P., and N. Sodhi, eds. 2009. *Special issue: Conservation issues in the Atlantic Forest. Biological Conservation* 142.6: 1137–1252.

This special issue of *Biological Conservation* contains some key articles on the Atlantic Forest including the highly impactful paper Ribeiro, et al. 2009 on how much of the Atlantic Forest remains and the distribution of the remnants.

Rezende, C. L., F. R. Scarano, E. D. Assad, et al. 2018. From hotspot to hopespot: An opportunity for the Brazilian Atlantic Forest. *Perspectives in Ecology and Conservation* 16.4: 208–214.

This study presents high-resolution (5-m) remote sensing data of the Brazilian Atlantic Forest to reveal that there are in fact 32 Mha of forest within the domain, which corresponds to 28 percent of the original extent (2 percent of the area is planted forest).

Ribeiro, M. C., J. P. Metzger, A. C. Martensen, F. J. Ponzoni, and M. M. Hirota. 2009. The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. *Biological Conservation* 142.6: 1141–1153.

This article provides a highly detailed analysis of how much forest cover remains in the Atlantic Forest by taking into account even quite small patches of forest. It discusses how the majority of the remaining patches are small, isolated, and composed of secondary vegetation, and it proposes four main strategies for protecting and restoring the Atlantic Forest.

Historical Background

The earliest evidence of human activity within the Atlantic Forest is from approximately 500 CE, and consists of indigenous settlements causing modest disturbance. Since the arrival of Portuguese settlers in the 16th century, deforestation for urbanization and agriculture has been rampant. Two of the world's most populous cities (São Paulo and Rio de Janeiro) are located within the Atlantic Forest realm, and Brazil's economy also relies heavily on agriculture, including commodities such as coffee, cocoa, sugar, rice, soybean, and cotton, as explained by Martinelli, et al. 2010. Much of Brazil's agriculture is exported and/or transported by trucks across Brazil, requiring an extensive road network, which, as shown in Freitas, et al. 2010, further exacerbates land use change and forest fragmentation. As for pre-Anthropocene history, there has been some research into the underlying evolutionary and paleoclimatic drivers responsible for the centers of endemism found within this region, such as the study Carnaval and Moritz 2008. Álvarez-Presas, et al. 2011 used planarians as model organisms to understand patterns of biodiversity, whereas the authors of Carnaval, et al. 2009 used frogs as indicators in their paleoclimatic models. Fjeldså and Rahbek 2006 used climate and phylogenetic data to explain the higher diversity of tanagers in the Atlantic Forest around Rio de Janeiro. Batalha-Filho, et al. 2013 used bird data to demonstrate how the Amazonian and Atlantic Forests were connected, linking the connection to geotectonic events.

Álvarez-Presas, M., F. Carbayo, J. Rozas, and M. Riutort. 2011. Land planarians (Platyhelminthes) as a model organism for fine-scale phylogeographic studies: Understanding patterns of biodiversity in the Brazilian Atlantic Forest hotspot. *Journal of Evolutionary Biology* 24.4: 887–896.

Paleoclimatic models have previously had variable success in predicting the observed habitat stability in southern Atlantic Forest. This research uses two land planarians as model organisms because they have low dispersal capability, whereas previous work focused on species with high dispersal ability. Results suggest that there were no recent colonizations or population expansions, indicating a long-term stability scenario.

Batalha-Filho, H., J. Fjeldså, P. H. Fabre, and C. Y. Miyaki. 2013. Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. *Journal of Ornithology* 154.1: 41–50.

This study uses phylogenetic and distribution data of birds to unravel the spatiotemporal dynamics of how the Atlantic and Amazon forests, which used to be connected, may have diverged.

Carnaval, A. C., M. J. Hickerson, C. F. Haddad, M. T. Rodrigues, and C. Moritz. 2009. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science* 323.5915: 785–789.

This article uses frogs as indicators to examine different potential scenarios of community responses to climate change in the late-Quaternary period. The results show that the Brazilian Atlantic Forest had a relatively unstable climate, with the exception of three refugia that provided a suitable habitat for many neotropical species in the late Pleistocene. The authors suggest conservation priorities based on their results.

Carnaval, A. C., and C. Moritz. 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography* 35.7: 1187–1201.

This research models the range of the Atlantic Forest under current and past climatic scenarios to investigate whether these can predict current patterns of biodiversity distribution. The past range of the forest was verified using fossil pollen data. The authors find evidence for two refugia (Bahia and Pernambuco) and suggest that southern forests may have been more unstable.

Fjeldså, J., and C. Rahbek. 2006. Diversification of tanagers, a species rich bird group, from lowlands to montane regions of South America. *Integrative and Comparative Biology* 46.1: 72–81.

This paper links rates of speciation within tanagers to geological events, and links the elevated levels of speciation in the Rio de Janeiro area and the Andean forelands to mountains being uplifted in the late Pleistocene.

Freitas, S. R., T. J. Hawbaker, and J. P. Metzger. 2010. Effects of roads, topography, and land use on forest cover dynamics in the Brazilian Atlantic Forest. *Forest Ecology and Management* 259.3: 410–417.

This article examines the effect of road density, land use, and topography on forest fragmentation, deforestation, and regrowth. Roads were found to have the strongest links to deforestation and forest fragmentation, as they facilitate both processes by creating increased accessibility.

Martinelli, L. A., R. Naylor, P. M. Vitousek, and P. Moutinho. 2010. Agriculture in Brazil: Impacts, costs, and opportunities for a sustainable future. *Current Opinion in Environmental Sustainability* 2.5–6: 431–438.

This is a thorough article that sets out the current state of Brazilian agriculture, the potential for sustainable development, and the major hurdles to achieving socioeconomic development without further environmental damage. Though it is about the whole of Brazil, it highlights key areas of conflict between economic and environmental priorities.

Habitat Types: Present and Future

The Atlantic Forest is commonly divided either into forest types (a particular community of plant species that define a region can be referred to as a physiognomy) or into biogeographical subregions. The types of forest present are: dense and open ombrophilous forests, mixed ombrophilous forests, seasonal forests, and semideciduous as well as mangrove forests. The most studied forest physiognomy is the ombrophilous dense forest, which is believed to have higher levels of biodiversity and endemism than other physiognomies, although this could be due to differences in sampling effort. However, Scarano 2009 argues that these peripheral forest subtypes, including restingas and swamp forests, should be given more priority, making the case that they have a high conservation value due to their oligarchic diversity, with a few dominant species but many rare species at local scales. The western extent of the forest that extends into Argentina (an introduction to this forest can be found in Chebez and Hilgert 2003) and Paraguay (Cartes 2003 presents a background to this forest) forms part of the subtropical semideciduous forest. Silva and Casteleti 2003 defined eight biogeographical subregions: Araucaria, Bahia, Brejos Nordestinos, Diamantina, Interior, Pernambuco, Serra do Mar, and São Francisco. Five of these regions are considered to be centers of endemism, as shown by Tabarelli, et al. 2010 (cited under Biodiversity of the Atlantic Forest). Assessments of the potential impacts of climate change on the Atlantic Forest have revealed the area to be extremely vulnerable: Lemes, et al. 2014 found that as species ranges shift under climate change, protected areas will need to shift with them, particularly for vulnerable taxa like amphibians. Scarano and Ceotto 2015 reviewed the vulnerability of both biodiversity and society of the Atlantic Forest to climate change and discusses important adaptive practices.

Cartes, J. L. 2003. Brief history of conservation in the Interior Atlantic Forest. In *The Atlantic Forest of South America: Biodiversity status, threats, and outlook*. Edited by C. G. Leal and I. G. Câmara, 269–287. Washington, DC: Island Press.

This book chapter presents a thorough introduction to a lesser-researched area of the Atlantic Forest, the Paraguayan Atlantic Forest. The author details the key differences to Brazilian and Argentinian forests and discusses its history.

Chebez, J., and N. Hilgert. 2003. Brief history of conservation in the Paraná Forest. In *The Atlantic Forest of South America: Biodiversity status, threats, and outlook*. Edited by C. G. Leal and I. G. Câmara, 141–159. Washington, DC: Island Press.

A descriptive history of the Argentinian Atlantic Forest from its early history to current land use in the region as well as an introduction to the local economic concerns and conservation efforts.

Colombo, A. F., and C. A. Joly. 2010. Brazilian Atlantic Forest lato sensu: The most ancient Brazilian forest, and a biodiversity hotspot, is highly threatened by climate change. *Brazilian Journal of Biology* 70.3: 697–708.

This study examines how the distribution of thirty-eight different species of tree that are typical to the Brazilian Atlantic Forest would be affected by different potential future climate scenarios. They use species distribution models to show that at best, they would lose about 25 percent of their distribution, but under the worst scenarios, 50 percent could be lost, demonstrating the sensitivity of the region to climate change.

Lemes, P., A. S. Melo, and R. D. Loyola. 2014. Climate change threatens protected areas of the Atlantic Forest. *Biodiversity and Conservation* 23.2: 357–368.

This article makes the excellent point that as species ranges change with climate change, protected areas need to expand to contain them, particularly in highland areas. By using amphibians as an example, the paper shows that without changes to protected areas, the number of species within them will decline.

Neves, D. M., K. G. Dexter, R. T. Pennington, et al. 2017. Dissecting a biodiversity hotspot: The importance of environmentally marginal habitats in the Atlantic Forest Domain of South America. *Diversity and Distributions* 23.8: 898–909.

This article shows that marginal habitats with poor environmental protection are very important to maintain high species richness in the Atlantic Forest, as 45 percent of Atlantic Forest endemic species only occur within these habitats.

Scarano, F. R. 2009. Plant communities at the periphery of the Atlantic rain forest: Rare-species bias and its risks for conservation. *Biological Conservation* 142.6: 1201–1208.

Scarano makes the case for a wider range of Atlantic Forest subtypes on the periphery of the core forest to be made a higher conservation priority despite lower diversity and endemism.

Scarano, F. R., and P. Ceotto. 2015. Brazilian Atlantic forest: Impact, vulnerability, and adaptation to climate change. *Biodiversity and Conservation* 24.9: 2319–2331.

This review identifies vulnerabilities of biodiversity and society in the Atlantic Forest, as over 60 percent of Brazil's population live within it. It discusses the role of ecosystem base adaptation strategies and highlights examples of good adaptive practice.

Silva, J. M. C., and C. H. M. Casteleti. 2003. Status of the biodiversity of the Atlantic Forest of Brazil. In *The Atlantic Forest of South America: Biodiversity status, threats, and outlook*. Edited by C. G. Leal and I. G. Câmara, 43–59. Washington, DC: Island Press.

This chapter summarizes the threat posed by deforestation to the Brazilian Atlantic Forest and how much of the biodiversity is now threatened due to habitat loss. It also outlines potential solutions and conservation strategies.

Biodiversity of the Atlantic Forest

Myers, et al. 2000 showed that nearly half of all plant species and more than a third of mammals, birds, reptiles, and amphibians are found in only 1.4 percent of the Earth's land surface area, regions which are known as biodiversity hotspots. Hence, by focusing conservation policies and protecting these areas, a disproportionately high level of biodiversity could be protected. A follow-up book about biodiversity hotspots, Mittermeier, et al. 2005, discusses that 40 percent of the 20,000 plant species, 16 percent of the 688 bird species, 27 percent of the 261 mammal species, 31 percent of 200 reptile species, and 60 percent of 280 amphibian species are endemic to the Atlantic Forest, which means they can only be found within this realm. Together, they represent over 8,650 species, 8,000 of which are tree species. The Atlantic Forest has shown some of the highest levels of biodiversity in the world. Martini, et al. 2007 identified an area containing 144 species of trees (above diameter at breast height > 4.8 cm) within 0.1 ha in southern Bahia, which is the second highest concentration of tree species in the world. Among animal species, one notable example is the golden lion tamarin, an endangered primate species that Lapenta and Procópio-de-Oliveira 2008 found to have a role in the seed dispersal of ninety-seven species of plants. The Atlantic Forest has high levels of endemism and of habitat loss, making it one of the most endangered biodiversity hotspots. The species that are still present are often trapped within small fragments and unable to migrate, as discussed in Tabarelli, et al. 2010. Amphibians are a taxon of particular concern due to their high endemicity and the increasing threat posed by chytrid fungus, as shown in Carnaval, et al. 2006.

Carnaval, A. C. O. Q., R. Puschendorf, O. L. Peixoto, V. K. Verdade, and M. T. Rodrigues. 2006. Amphibian chytrid fungus broadly distributed in the Brazilian Atlantic Rain Forest. *EcoHealth* 3.1: 41–48.

This research reports the results of histological screenings on nearly one hundred anurans from the Atlantic rain forest, showing a wide spread of chytrid fungus, with an infection record coinciding with the first observed declines in amphibians.

Lapenta, M. J., and P. Procópio-de-Oliveira. 2008. Some aspects of seed dispersal effectiveness of golden lion tamarins (*Leontopithecus rosalia*) in a Brazilian Atlantic forest. *Tropical Conservation Science* 1.2: 122–139.

This article examines the role of golden lion tamarins in the dispersal of seeds by analyzing the seeds found in fecal deposits. This evidence shows that the golden lion tamarin may be a keystone species, as it provides ecosystem services for so many other species.

Martini, A. M. Z., P. Fiaschi, A. M. Amorim, and J. L. da Paixão. 2007. A hot-point within a hot-spot: A high diversity site in Brazil's Atlantic Forest. *Biodiversity and Conservation* 16.11: 3111–3128.

This paper describes an area of exceptionally high biodiversity within the Atlantic Forest, located within southern Bahia, revealing it to be the second most diverse area in the world.

Mittermeier, R. A., P. R. Gill, M. Hoffmann, et al. 2005. *Hotspots revisited: Earth's biologically richest and most endangered terrestrial ecoregions*. Washington, DC: CEMEX.

This book takes a detailed look at the world's biodiversity hotspots, looking into levels of endemism for many taxonomic groups.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. Da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403.6772: 853.

This is a key paper on the subject of biodiversity hotspots. It identifies areas of the world that are most concentrated in biodiversity and puts them forward as a "silver bullet" conservation strategy to conserve the highest level of biodiversity within the smallest area.

Tabarelli, M., A. V. Aguiar, M. C. Ribeiro, J. P. Metzger, and C. A. Peres. 2010. Prospects for biodiversity conservation in the Atlantic Forest: Lessons from aging human-modified landscapes. *Biological Conservation* 143.10: 2328–2340.

This review provides a good introduction to the history of Atlantic Forest disturbance and discusses in detail various potential conservation options, highlighting the value of protected areas and old-growth forests. It also contains a useful table of studies that have examined the effects of habitat loss on different taxa and the conservation insights they provide.

Biodiversity Data

There have been large-scale efforts to gather data on Atlantic Forest biodiversity. There are several public data sets known as the Atlantic data papers, which represent a high proportion of diversity present in the region, including: 94 species of mammal (see Souza, et al. 2019), 26 primate species (see Culot, et al. 2019), 745 bird species (see Hasui, et al. 2018), 2,095 epiphyte species (see Ramos, et al. 2019), 279 butterfly species (see Santos, et al. 2018), 528 amphibian species (see Vancine, et al. 2018), and 98 bat species (see Muylaert, et al. 2017). Several of these data sets monitor assemblages (e.g. Culot, et al. 2019) over time. As well as simple occurrence records, some of the data sets collate information on species traits and interactions, such as plant–frugivore interactions in Bello, et al. 2017, and bird traits including body mass and wing length in Rodrigues, et al. 2019. Going beyond simple species occurrence data allows researchers to investigate patterns over time such as demographics and make inferences about how anthropogenic stressors affect morphology. Interaction data is particularly valuable as it can uncover species mutualisms and the extent to which certain species depend on others. This contributes to the deeper understanding of how ecological communities work, which can be useful to inform better conservation policies. For the state of São Paulo, there is also information available on the SinBiota 2.1 platform, as described in Mira, et al. 2011. This was created by the Biota/Fapesp program to integrate information generated by all researchers funded by this program.

Bello, C., M. Galetti, D. Montan, et al. 2017. Atlantic frugivory: A plant–frugivore interaction data set for the Atlantic Forest. *Ecology* 98.6: 1729.

This data set contains pairwise observations between 331 vertebrate species and 788 plant species from the Atlantic Forest.

Culot, L., L. A. Pereira, I. Agostini, et al. 2019. ATLANTIC-PRIMATES: A dataset of communities and occurrences of primates in the Atlantic Forests of South America. *Ecology* 100.1: e02525.

This detailed data set contains georeferenced locations of all twenty-six primate species found in the Atlantic Forest as well as one introduced species. It describes 700 primate communities made up of over 8,000 individual single-species occurrences and covers the entire Atlantic Forest range including Brazil, Argentina, and Paraguay.

Hasui, É., J. P. Metzger, R. G. Pimentel, et al. 2018. ATLANTIC BIRDS: A data set of bird species from the Brazilian Atlantic Forest. *Ecology* 99.2: 497.

This is a large data set comprising over 33,000 individual birds of 832 species from the Atlantic Forest in Brazil. It includes location, date, sampling mode, altitude, and type of habitat.

Mira, C., P. Feijao, T. Duque-Estrada, J. Meidanis, and C. A. Joly. 2011. The SinBiota 2.0 biodiversity information system. In *Proceedings 2011 Seventh IEEE International Conference on e-Science. 5–8 December 2011*. pp. 142–149. Los Alamitos, CA: IEEE.

This conference paper discusses the history of the SinBiota platform and future plans to include new technologies.

Muylaert, R. d. L., R. D. Stevens, C. E. L. Esbérard, et al. 2017. ATLANTIC BATS: A data set of bat communities from the Atlantic Forests of South America. *Ecology* 98.12: 3227.

This data set comprises information on over 90,000 individual bat captures totaling ninety-eight species. The data reported were collected in 205 sites from 135 studies.

Ramos, F. N., S. R. Mortara, N. Monalisa-Francisco, et al. 2019. ATLANTIC EPIPHYTES: A data set of vascular and non-vascular epiphyte plants and lichens from the Atlantic Forest. *Ecology* 100.2: e02541.

An extensive data set containing 2,095 species of epiphytes from the Atlantic Forest including 89,270 individual records. The data set includes records from 1824 to early 2018.

Rodrigues, R. C., É. Hasui, J. C. Assis, et al. 2019. ATLANTIC BIRD TRAITS: A data set of bird morphological traits from the Atlantic forests of South America. *Ecology* 100.6: e02647.

A large data set of bird traits from 67,197 individuals of 711 species. The traits recorded include sex, body mass, body length, and reproductive stage.

Santos, J. P. dos, A. V. L. Freitas, K. S. Brown, et al. 2018. Atlantic butterflies: A data set of fruit-feeding butterfly communities from the Atlantic forests. *Ecology* 99.12: 2875.

A large data set of butterfly occurrence records from the Atlantic Forest including 7,062 presence records for 279 species dating from 1949 to 2018.

Souza, Y., F. Gonçalves, L. Lautenschlager, et al. 2019. ATLANTIC MAMMALS: A dataset of assemblages of medium and large-sized mammals of the Atlantic Forest of South America. *Ecology* 100.10: e02785.

This data set compiles 129 studies resulting in occurrence data of ninety-four species of mammal in the Atlantic Forest across 244 assemblages.

Vancine, M. H., K. da Silva Duarte, Y. S. de Souza, et al. 2018. ATLANTIC AMPHIBIANS: A data set of amphibian communities from the Atlantic Forests of South America. *Ecology* 99.7: 1692.

This data set of Atlantic Forest amphibians includes 17,619 records of 528 species. The records date between 1940 and 2017.

Habitat Transformation

Habitat loss, fragmentation, and degradation are the major threats to biodiversity in the Atlantic Forest, thus there is a substantial body of literature on this subject. Lôbo, et al. 2011 showed that habitat transformation has over time modified diverse communities within five physiognomic subtypes (including evergreen, semideciduous, and open forest) of the Atlantic Forest into a homogenized set of disturbance-specialist species. Indeed, studying in southern Bahia (world's second highest concentration of tree species), the authors of Benchimol, et al. 2017 found that forest loss leads to nonrandom floristic shifts, such that shade-intolerant species (e.g., pioneers) become more common than shade-tolerant species below 30 percent of forest cover. The loss of species also drives evolutionary changes in seed size, as shown by Galetti, et al. 2013. Within the evergreen and semideciduous forests, Santos, et al. 2008 demonstrated that fragmentation and the resulting creation of more edges have severely reduced functional traits of tree assemblages. But habitat loss is not the only driver of species loss. Using a multi-taxa database with over 2,200 community-level estimates from 1,097 sites, Püttker, et al. 2020 showed that forest-dependent species respond negatively to habitat loss and fragmentation, and that in areas with over 30 percent of forest cover, habitat fragmentation was as important as or more important than habitat loss in driving changes in species richness. This was not the first time that it has been shown that the effects of fragmentation on species are dependent on the amount of forest cover. Pardini, et al. 2010 demonstrated that within the ombrophilous dense forests, the size of a forest fragment only positively affects biodiversity when the landscape level forest cover is intermediate (around 30 percent of forest cover). The impacts of edge effects have been shown for a variety of abiotic and biotic conditions. Magnago, et al. 2015 showed that forest edges are drier and warmer, and these abiotic changes affect forest structure. Changes to habitat structure can then have knock-on effects on other taxa. For instance, Filgueiras, et al. 2011 found that dung beetle diversity was impacted by the impoverished flora of small patches. Banks-Leite, et al. 2010 found that edge effects likely drive the patch area effects on birds in the Atlantic Forest. This is because large patches experience a weaker influence of edge effects than small patches, which have higher edge-to-area ratio.

Banks-Leite, C., R. M. Ewers, and J. P. Metzger. 2010. Edge effects as the principal cause of area effects on birds in fragmented secondary forest. *Oikos* 119.6: 918–926.

This article demonstrates that edge and area effects are intrinsically confounded in fragmented landscapes, but that the magnitude of edge-to-interior differences increases in larger patches. When controlling for edge effects, the authors show that patch size does not affect birds.

Benchimol, M., E. Mariano-Neto, D. Faria, et al. 2017. Translating plant community responses to habitat loss into conservation practices: Forest cover matters. *Biological Conservation* 209:499–507.

This study shows that species richness of plants declines drastically below 30 percent of forest cover, and that in deforested areas the assemblages of seedlings and saplings is very different from that of mature trees. These results suggest that the incredibly speciose forest of southern Bahia will progressively lose species as mature trees are replaced by younger individuals.

Filgueiras, B. K., L. Iannuzzi, and I. R. Leal. 2011. Habitat fragmentation alters the structure of dung beetle communities in the Atlantic Forest. *Biological Conservation* 144.1: 362–369.

This research is based on field work performed in nineteen forest fragments of varying size in the Atlantic Forest in northeastern Brazil. It demonstrates not only that fragment area directly affects dung beetle diversity, but that low tree species diversity and lower levels of shade-tolerant plants also lowered dung beetle diversity.

Galetti, M., R. Guevara, M. C. Côrtes, et al. 2013. Functional extinction of birds drives rapid evolutionary changes in seed size. *Science* 340.6136: 1086–1090.

This article makes the link between the reduction in seed size of a keystone palm species and the functional extinction of large-gape seed dispersers in the Brazilian Atlantic Forest. This shows a short-term adaptation to loss of large birds.

Lôbo, D., T. Leão, F. P. Melo, A. M. Santos, and M. Tabarelli. 2011. Forest fragmentation drives Atlantic forest of northeastern Brazil to biotic homogenization. *Diversity and Distributions* 17.2: 287–296.

This article makes the important observation that land cover change is altering tree flora communities and that since 1980 similarity in species composition has increased by 20–40 percent.

Magnago, L. F. S., M. F. Rocha, L. Meyer, S. V. Martins, and J. A. A. Meira-Neto. 2015. Microclimatic conditions at forest edges have significant impacts on vegetation structure in large Atlantic forest fragments. *Biodiversity and Conservation* 24.9: 2305–2318.

This study provides a good working example of how edge effects impact microclimatic conditions and how these changes affect vegetation structure at edges.

Pardini, R., A. A. Bueno, T. A. Gardner, P. I. Prado, and J. P. Metzger. 2010. Beyond the fragmentation threshold hypothesis: Regime shifts in biodiversity across fragmented landscapes. *PloS One* 5.10: e13666.

This article presents a model that describes mechanisms and consequences of changes in biodiversity within fragmented landscapes, taking factors such as local extinction risk and immigration rates. They show that patch size only positively affects biodiversity in landscapes with intermediate amounts of cover (30 percent). In highly forested (50 percent) or deforested (10 percent) landscapes, species richness is not affected by patch size.

Püttker, T., R. Crouzeilles, M. Almeida-Gomes, et al. 2020. Indirect effects of habitat loss via habitat fragmentation: A cross-taxa analysis of forest-dependent species. *Biological Conservation* 241:108368.

This article analyzes data from the Synthesis in Atlantic Forest Ecology and Sustainability Group. This group has collated data from various projects, putting together a database on amphibians, reptiles, birds, mammals, spiders, harvestmen, beetles, butterflies, termites, bees, ants, and other insects, as well as bryophytes, pteridophytes, and higher plants. Here they report their first results on the relative effects of habitat loss and fragmentation on forest specialists.

Santos, B. A., C. A. Peres, M. A. Oliveira, A. Grillo, C. P. Alves-Costa, and M. Tabarelli. 2008. Drastic erosion in functional attributes of tree assemblages in Atlantic forest fragments of northeastern Brazil. *Biological Conservation* 141.1: 249–260.

This article examines data of tree assemblages in a hyperfragmented part of the Atlantic Forest in northeastern Brazil, showing that smaller fragments and higher edge length within the evergreen and semideciduous physiognomies have drastic effects on functional traits from seed size to proportion of pioneer and emergent species.

Vertebrates and Habitat Transformation

Birds are one of the most commonly studied taxa in the Atlantic Forest due to their diversity and sensitivity. Responses of birds to fragmentation have been well documented: Zurita and Bellocq 2010 found forest cover to be the main driver of differences in bird communities in Argentina, while Morante-Filho, et al. 2015 showed the number of bird species in southern Bahia abruptly changed at a threshold of 50 percent forest cover. Banks-Leite, et al. 2012 showed that around São Paulo responses of bird communities to fragmentation do not to conform to the classical ecological species–area relationship. Instead, the main drivers of changes in bird communities across a gradient of disturbance are purported to be individual species reaching their extinction threshold. Despite conservation efforts, many species remain at high risk of extinction (as demonstrated in Canale, et al. 2012), particularly large mammals such as the jaguar, which, as shown by Paviolo, et al. 2016, have undergone high rates of extirpation. Umetsu and Pardini 2007 found that small mammals, particularly endemic species, have been found to be sensitive to land use change. The response of bats is less clear, as Gorresen and Willig 2004 found the highest levels of bat diversity in moderately fragmented landscapes. With regards to the response of amphibians to fragmentation, Becker, et al. 2007 found habitat loss to be a key driver of amphibian declines, particularly for forest species; and amphibians are particularly vulnerable to the coupled effects of fragmentation and climate change, as shown by Loyola, et al. 2014. The effects of fragmentation on reptiles are less studied, but Lion, et al. 2016 demonstrated that reptiles can benefit from even small forest fragments.

Banks-Leite, C., R. M. Ewers, and J. P. Metzger. 2012. Unraveling the drivers of community dissimilarity and species extinction in fragmented landscapes. *Ecology* 93.12: 2560–2569.

This article demonstrates that using species–area relationship in fragmentation studies does not hold for bird communities in the Atlantic Forest and that changes in community composition are primarily driven by species-level extinction thresholds.

Becker, C. G., C. R. Fonseca, C. F. B. Haddad, R. F. Batista, and P. I. Prado. 2007. Habitat split and the global decline of amphibians. *Science* 318.5857: 1775–1777.

This article shows that as well as habitat loss and the rapid spread of chytrid fungus driving the losses of amphibians globally, fragmentation of habitat is also a key driver, particularly in species that migrate to forests from an aquatic larval stage.

Canale, G. R., C. A. Peres, C. E. Guidorizzi, C. A. F. Gatto, and M. C. M. Kierulff. 2012. Pervasive defaunation of forest remnants in a tropical biodiversity hotspot. *PloS One* 7.8.

This paper examines local extinctions of 18 mammal species in 196 patches of the Brazilian Atlantic Forest, and finds that of a possible 3,528 populations that could have persisted, only 767 have, and each forest patch only retains an average of 3.9 species.

Gorresen, P. M., and M. R. Willig. 2004. Landscape responses of bats to habitat fragmentation in Atlantic forest of Paraguay. *Journal of Mammalogy* 85.4: 688–697.

This article reports the outcome of an extensive survey effort on different landscape within the Atlantic Forest of Paraguay, finding the highest species richness in partly forested fragments.

Lion, M. B., A. A. Garda, D. J. Santana, and C. R. Fonseca. 2016. The conservation value of small fragments for Atlantic Forest reptiles. *Biotropica* 48.2: 265–275.

This fragmentation study focuses on reptile communities in the Brazilian Atlantic Forest; though the main predictor of reptile species richness and abundance is fragment size, both the matrix quality and shape of the fragment contribute to the communities present.

Loyola, R. D., P. Lemes, F. T. Brum, D. B. Provete, and L. D. Duarte. 2014. Clade-specific consequences of climate change to amphibians in Atlantic Forest protected areas. *Ecography* 37.1: 65–72.

This paper shows that the suitable range for most amphibian species would contract under climate change and that the responses are clade-specific. It starts to piece together how understanding the changes in the phylogenetic pool may lead to a more comprehensive idea of the effect of climate change on assembly-related processes.

Morante-Filho, J. C., D. Faria, E. Mariano-Neto, and J. Rhodes. 2015. Birds in anthropogenic landscapes: The responses of ecological groups to forest loss in the Brazilian Atlantic Forest. *PLoS One* 10.6.

The authors examined bird species grouped by whether they were forest specialists or habitat generalists and ran models to investigate the effect of fragmentation. The results find that all ecological groups show a similar forest cover threshold value of 50 percent, where species numbers abruptly change.

Paviolo, A., C. De Angelo, K. M. Ferraz, et al. 2016. A biodiversity hotspot losing its top predator: The challenge of jaguar conservation in the Atlantic Forest of South America. *Scientific Reports* 6.1: 1–16.

A thorough review of the status of the jaguar's habitat in the Atlantic Forest, combining information from fourteen research groups. It shows that jaguars only persist in 2.8 percent of the Atlantic Forest in low densities.

Umetsu, F., and R. Pardini. 2007. Small mammals in a mosaic of forest remnants and anthropogenic habitats—evaluating matrix quality in an Atlantic forest landscape. *Landscape Ecology* 22.4: 517–530.

This paper investigates matrix use by small mammals of the Atlantic Forest, showing that assemblages of small mammals are highly dissimilar between native vegetation and anthropogenic land use. Endemic small mammals typically occupy native vegetation.

Zurita, G. A., and M. I. Bellocq. 2010. Spatial patterns of bird community similarity: Bird responses to landscape composition and configuration in the Atlantic forest. *Landscape Ecology* 25.1: 147–158.

This paper examines how continuous and fragmented forest landscapes affect the similarity of bird communities, finding that forest cover explains most of the variation. It also finds that native bird communities are more resilient to forest loss in landscapes dominated by planted trees.

Conservation and Policy

The Atlantic Forest is protected by the Forest Code, which is an environmental law created in Brazil in 1965, when most of the deforestation had already taken place. By law, landowners are required to set aside 20 percent of their land for native habitat, as well as protect riparian forests, hilltops, and other environmentally sensitive areas. The Forest Code was revised recently, weakening the protection of the Atlantic Forest, as discussed by Soares-Filho, et al. 2014. This revision is particularly problematic given that Banks-Leite, et al. 2014 has shown that at least 30 percent of native habitat is required to protect biodiversity within the Atlantic Forest. Due to the pressing need to preserve its unique yet endangered biota, a group of academics, NGOs, industry, and government formed the Atlantic Forest Restoration Pact, an initiative which aims to restore 15 Mha of habitat in the Atlantic Forest by 2050. This pledge comes as part of Brazil's commitment to the Bonn Challenge. Crouzeilles, et al. 2019 shows that the Atlantic Forest Restoration Pact has already facilitated the restoration of roughly 700,000 ha, estimating that by 2020 there will be 1.5 Mha under restoration. Rezende, et al. 2018 has estimated that if landowners comply with the new Forest Code to restore riparian forest (i.e., forest strip along rivers), by 2038 the vegetation cover in the Atlantic Forest will be close to 35 percent, bringing hope to the preservation of this charismatic and species-rich system.

Banks-Leite, C., R. Pardini, L. R. Tambosi, et al. 2014. Using ecological thresholds to evaluate the costs and benefits of set-asides in a biodiversity hotspot. *Science* 345.6200: 1041–1045.

This study analyzes the responses of mammals, birds, and amphibians to habitat loss to show that the minimum amount of area required to maintain biodiversity is 30 percent of forest cover. The authors used these results to plan a domain-wide restoration strategy and demonstrate that with 6.5 percent of Brazil's annual expenditure on agricultural subsidies, it would be possible to restore priority areas back to 30 percent of cover.

Crouzeilles, R., E. Santiami, M. Rosa, et al. 2019. There is hope for achieving ambitious Atlantic Forest restoration commitments. *Perspectives in Ecology and Conservation* 17.2: 80–83.

The authors document the amount of forest that has been restored between 2011 and 2015 and how much restoration they expect to be under way by 2020. They discuss how this progress is due to the activities promoted by the Atlantic Forest Restoration Pact. They discuss how these activities could help restoration elsewhere.

Rezende, C. L., F. R. Scarano, E. D. Assad, et al. 2018. From hotspot to hopespot: An opportunity for the Brazilian Atlantic Forest. *Perspectives in Ecology and Conservation* 16:208–214.

This study presents the most updated estimate of extent of forest cover within the Atlantic Forest. It also discusses that landowners are legally required to restore 5.2 Mha of currently degraded riparian area. Through adequate enforcement, restoring these areas could cause native vegetation cover to increase to 35 percent of forest cover.

Soares-Filho, B., R. Rajão, M. Macedo, et al. 2014. Cracking Brazil's Forest Code. *Science* 344.6182: 363–364.

This policy forum discusses the origins of Brazil's Forest Code, the changes that were implemented recently, and their impacts on forest cover and carbon storage.

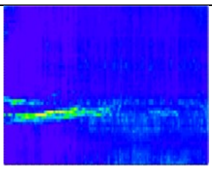
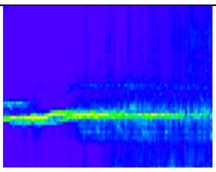
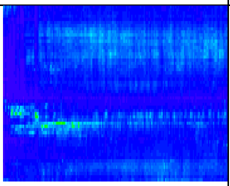
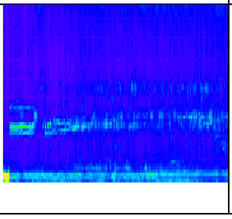
Appendix B:

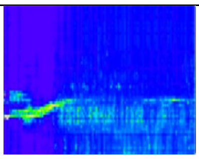
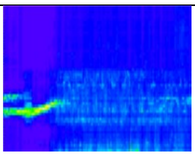
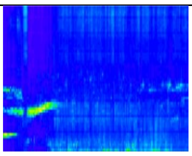
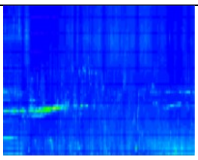
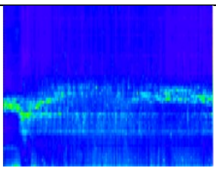
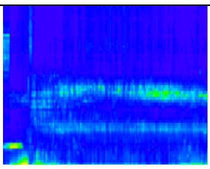
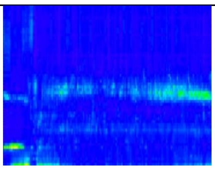
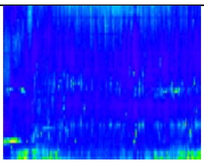
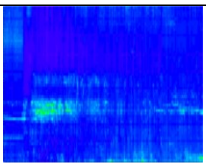
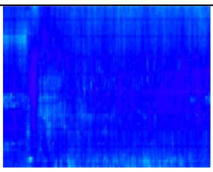
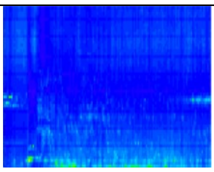
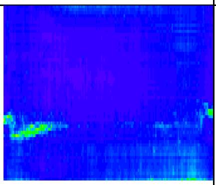
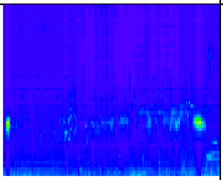
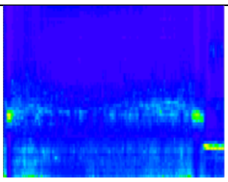
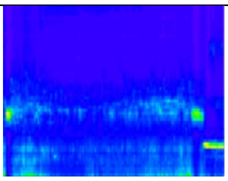
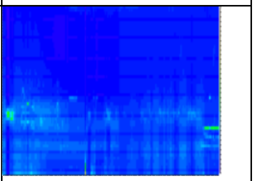
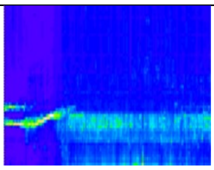
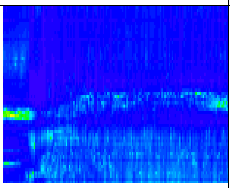
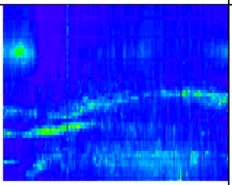
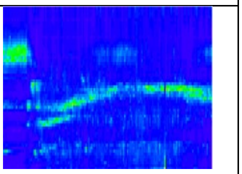
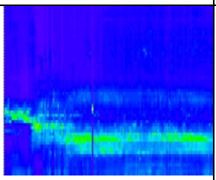
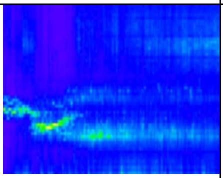
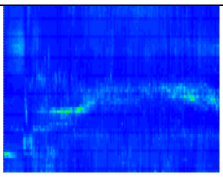
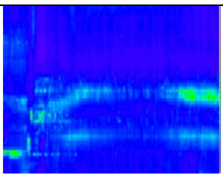
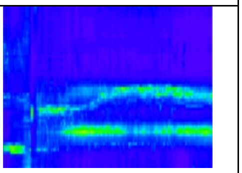
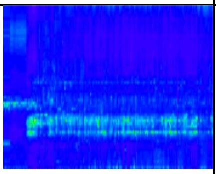
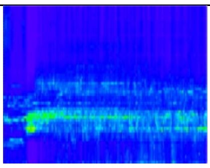
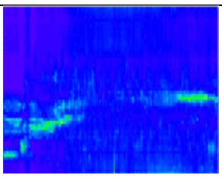
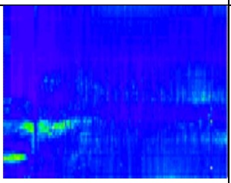
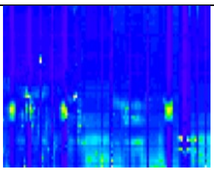
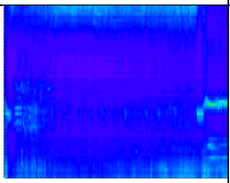
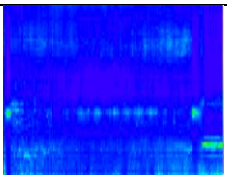
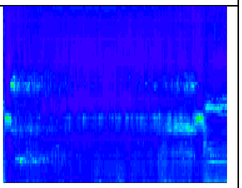
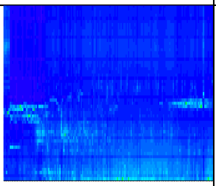
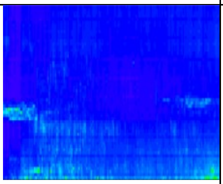
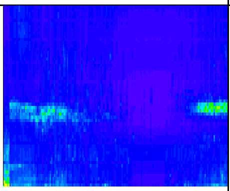
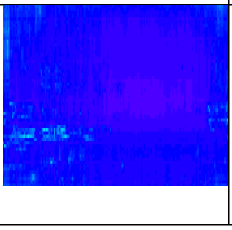
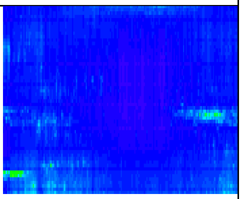
B.1: Supplementary Information for Chapter 3

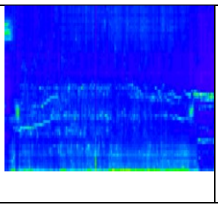
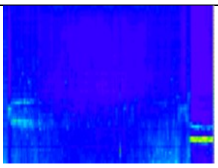
Table B.2: Farm location and forest cover metadata

Farm ID	Forest cover (% within 1000m radius)	Recorders deployed	Recorders collected	Latitude	Longitude	Mist netting
1	9	09/02/2019	18/02/2019	-21.6231	-46.5461	Y
2	13	04/01/2019	18/01/2019	-21.6171	-46.526	Y
3	25	19/01/2019	31/02/2019	-21.7822	-46.4457	N
4	25	09/01/2019	21/01/2019	-21.6959	-46.6279	Y
5	27	25/02/2019	09/03/2019	-22.0338	-46.5738	Y
6	29	07/01/2019	18/01/2019	-21.7439	-46.4738	N
7	37	27/01/2019	06/02/2019	-21.7987	-46.6777	Y
8	37	22/01/2019	01/02/2019	-21.9579	-46.6782	Y
9	44	20/02/2019	05/03/2019	-21.6591	-46.3403	Y
10	50	05/02/2019	17/02/2019	-21.7728	-46.4806	Y
11	60	09/03/2019	20/03/2019	-21.7668	-46.5452	Y

Table B.3: Spectrograms of mean soundscapes from the transect points at each sample site. The Farm IDs correspond to those in Table 1. The x-axis corresponds to 4:30 am to 19:30 pm. The y-axis represents frequency from 1000 Hz to 12,000 Hz.

Farm ID	Transect				
	1	2	3	4	5
1					-

2					-
3	-				
4	-				-
5					
6		-			
7					
8					-
9	-				
10					

11		-		-	-

CSV files of the bird communities and soundscape matrices as well as metadata are available.

Code to run the PCoA analysis, modelling, and plotting are also available

Table B.4: Model selection results. A is acoustic composition (PCoA 1 value), F is forest cover within 1000m, L is land use category, and S is site.

Model structure	**AIC**
A~F+L+(1|S)	**-62.78473**
A~F*L+(1|S)	-43.34625
A~F*L	-44.17307
A~F*L+(1|S +1|F)	-42.39815

Figure B.11: Model validation plots. Residuals vs fitted for (A) soundscape model and (B) bird model

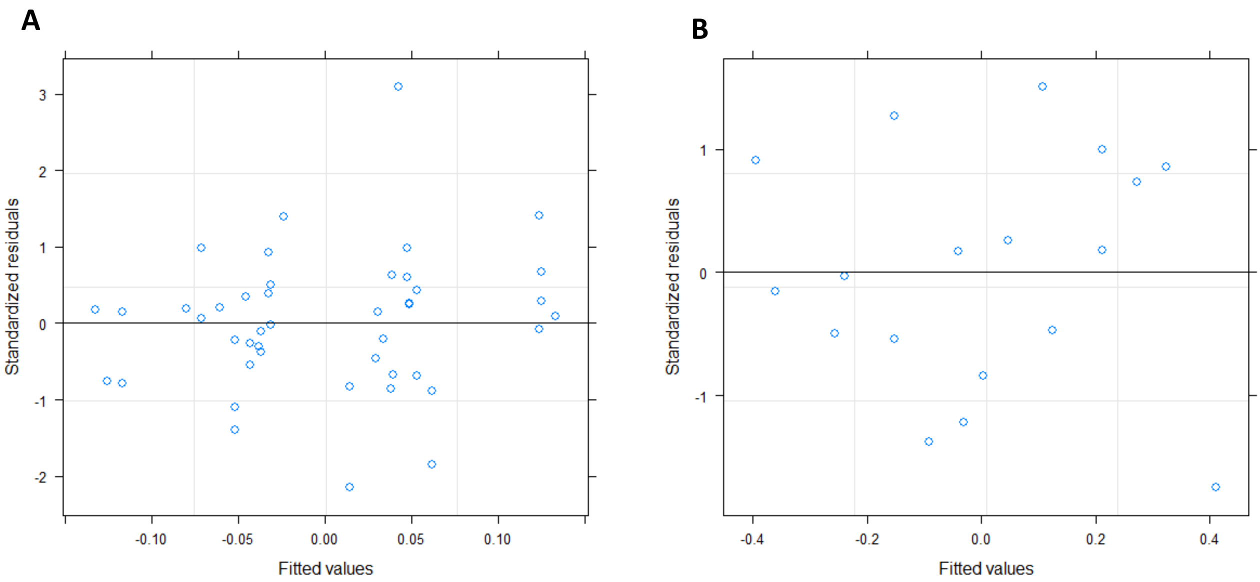


Table B.5: Output parameters of selected model with lowest AICs for both the soundscape model and the bird model. The intercept values denote the baseline of coffee land use. Models are both linear mixed effects model with the formula: PCoA output ~ Forestation + Landuse with site as a categorical random effect.

	Soundscape model				***Bird model***			
Predictors	*Estimates*	*df*	*C.I.*	*p-value*	*Estimates*	*df*	*C.I.*	*p-value*
(Intercept – coffee land use)	0.152	32	0.063 – 0.241	**0.0015****	-0.473	8	-0.063 – -0.315	**<0.0001 *****
Forest cover	-0.003	9	-0.006 – -0.0004	**0.0311***	-0.0087	7	0.005 – 0.0125	**0.0001*****
Land use (Edge)	-0.094	32	-0.149 – -0.040	**0.0012****	0.3644	8	0.218 – 0.511	**0.0002*****

Land use (Forest)	-0.086	32	-0.133 – -0.039	**0.0008****	
Random Effects:					
σ^2	0.004				0.00002
ICC	0.315				<0.00001
N_{Site}	11				9
No. obs	45				18
Marginal R^2	0.374				0.719
Conditional R^2	0.571				0.719

Table B.6: Table of files listened to, to determine source of indicator frequencies. Only files with drivers present are included.

Cluster	Forest cover	Time	Frequency	Sound	Date
c1	37	07:56	3000-7000	bird	24/01/2019
c1	37	07:56	4200-4500	Insect	24/01/2019
c1	37	08:06	2000-2100	bird	30/01/2019
c1	37	08:06	6500-9000	bird	30/01/2019
c1	37	08:36	3000-6700	bird	28/01/2019
c1	37	08:36	6200-9000	bird	28/01/2019
c1	37	06:31	2600	birds	28/01/2019
c1	37	06:31	2600	birds	28/01/2019
c1	37	06:31	2600	birds	30/01/2019
c1	37	06:31	2600	birds	27/01/2019

c1	37	06:31	2600	birds	28/01/2019
c1	37	06:31	2600	birds	31/01/2019
c1	37	06:31	2600	birds	27/01/2019
c1	37	08:51	1600-2200	birds	05/02/2019
c1	37	08:51	2000-3000	birds	20/01/2019
c1	37	08:51	2800	bird	26/01/2019
c4	9	09:16	4400	insect	12/02/2019
c4	9	09:16	4400	insect	17/02/2019
c4	9	09:16	4400	insect	11/02/2019
c4	9	09:16	4400	insect	16/02/2019
c2	13	06:21	3600-4000	bird and insect	13/01/2019
c2	13	06:21	3600-4000	bird	16/01/2019
c2	13	06:21	3600-4000	bird	06/01/2019
c2	13	06:21	3600-4000	bird	17/01/2019
c2	13	06:21	3600-4000	bird	08/01/2019
c3	27	08:21	3200-5400	geophony	26/02/2019
c3	27	08:21	3200-5400	bird and insect	01/03/2019
c3	27	08:21	3200-5400	bird and insect	26/02/2019
c3	27	08:21	3200-5400	bird and insect	28/02/2019
c3	29	08:21	3200-5400	bird and insect	14/01/2019
c3	29	08:21	3200-5400	insect	12/01/2019
c3	25	08:21	3200-5400	bird and insect	14/01/2019
c3	25	08:21	3200-5400	bird	19/01/2019
f1	13	06:51	3600-4200	insect	12/01/2019
f1	13	06:51	3600-4200	insect	13/01/2019
f1	13	06:51	3600-4200	insect	06/01/2019
f1	13	06:51	3600-4200	insect	05/01/2019
f1	13	06:51	3600-4200	insect	05/01/2019
f1	13	06:41	3800-4200	insect, some bird	06/01/2019

f4	13	06:26	8000-12000	bird	05/01/2019
f4	13	06:26	8000-12000	bird	10/01/2019
f3	37	07:51	11000-12000	bird	29/01/2019
f3	37	07:51	11000-12000	bird	24/01/2019
f3	37	07:51	11000-12000	bird	25/01/2019
f3	37	07:51	11000-12000	bird	27/01/2019
f3	37	07:51	11000-12000	bird	31/01/2019
f3	37	07:51	11000-12000	insect	02/02/2019
f3	37	07:51	11000-12000	bird	01/02/2019
f6	25	17:21	6200-6600	bird and insect	19/01/2019
f6	25	17:21	6200-6600	bird	10/01/2019
f6	25	17:21	6200-6600	insect	14/01/2019
f6	29	17:21	6200-6600	insect	08/01/2019
f6	29	17:21	6200-6600	insect	16/01/2019
f6	29	17:21	6200-6600	insect	16/01/2019
f6	27	17:21	6200-6600	insect	25/02/2019
f6	27	17:21	6200-6600	insect	04/03/2019
f6	25	17:21	6200-6600	insect	19/01/2019
f5	13	15:56	10800-11800	bird	07/01/2019
f5	13	15:56	10800-11800	bird	06/01/2019

f5	13	15:56	10800- 11800	bird	12/01/2019
f2	13	12:26	2800-4400	bird	04/01/2019
f2	13	12:26	2800-4400	bird	05/01/2019
f2	13	12:26	2800-4400	anthrophony	07/01/2019
f2	13	12:26	2800-4400	bird	08/01/2019
f2	13	12:26	2800-4400	bird and insect	13/01/2019
f2	13	11:06	2200-4200	bird	13/01/2019
f2	13	11:06	2200-4200	bird	06/01/2019
f2	13	11:06	2200-4200	bird	07/01/2019
f7	37	04:41	4600-5200	insect	23/01/2019
f7	37	04:41	4600-5200	insect	27/01/2019
f7	37	04:41	4600-5200	insect	25/01/2019
f7	37	04:41	4600-5200	insect	29/01/2019
f7	37	04:41	4600-5200	insect	30/01/2019
f7	37	04:41	4600-5200	insect	28/01/2019

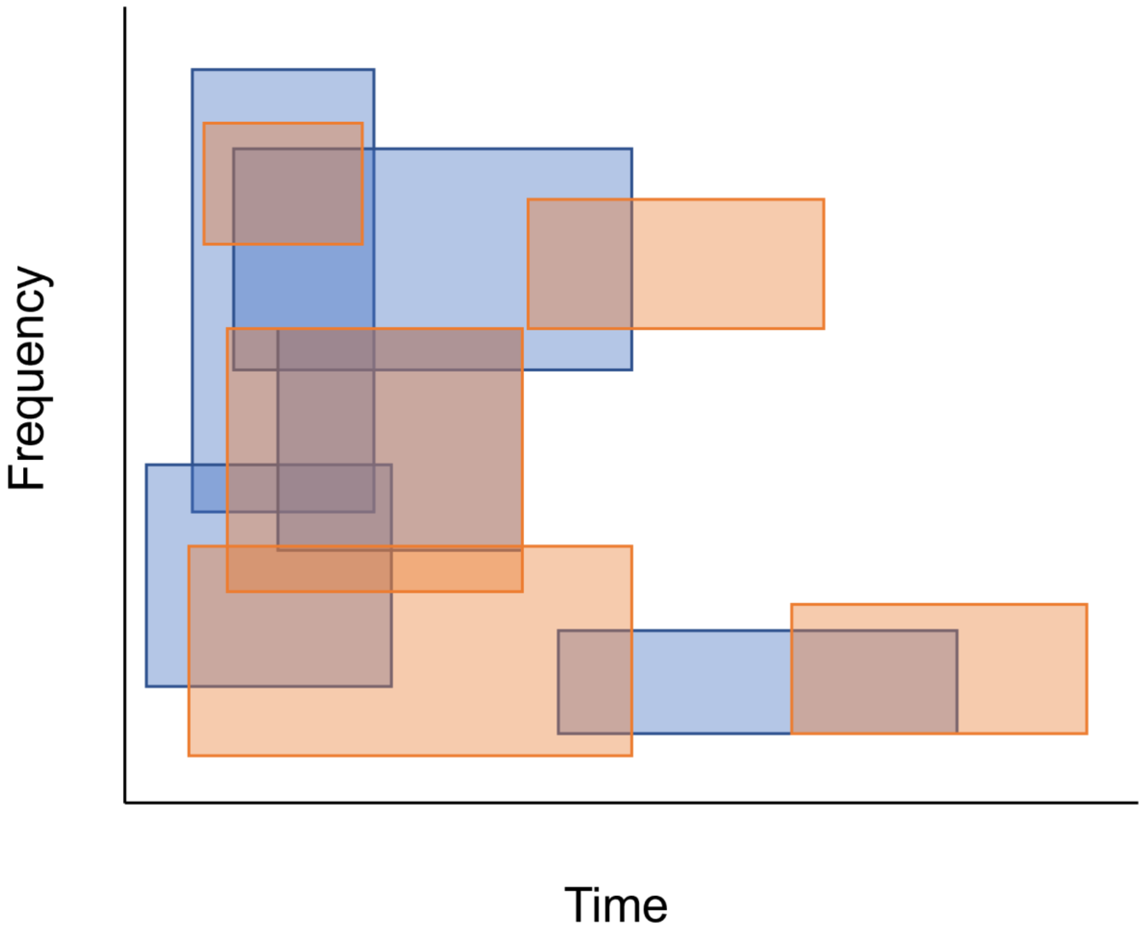


Figure B.12: *Conceptual diagram to explain acoustic niche hypothesis: Each box represents the typical calling frequency and time of a species, the box colour represents the land use. When running the indicator frequency analysis, only times and frequencies that do not overlap any others will remain. Therefore, some species will be lost, which could be the source of difference in community.*

B.2: Manuscript of Charis Declaudure's upcoming publication

Comparing soundscape methods for bird diversity monitoring

Declaudure, Charis*[1]; Damstra, Elise[1]; Lawson, Jenna[1]; Xiuhan Zhang[1]; Pearse, Will[1]; Banks-Leite, Cristina[1]

* Corresponding author: charis.declaudure@gmail.com

1 – Georgina Mace Centre, Silwood Park Campus, Imperial College London

Abstract

Biodiversity monitoring is key to understand anthropogenic pressures on ecosystems. The use of audio recording units (ARUs) in hand with acoustic indices within soundscape ecology has greatly improved efficiency of biodiversity monitoring. However, the effectiveness of acoustic indices is debated, and a novel soundscape method that measures acoustic composition could potentially be more biologically meaningful. In this study, we used 38 ARUs across a gradient of habitat types in rural England to assess the effectiveness of a combination of nine acoustic indices and acoustic composition for monitoring the bird community. Our results show that most acoustic indices performed poorly. Although the combination of NDSI and LFC were significantly correlated to bird species richness and composition, acoustic indices did not reveal the same ecological trends across habitats as observed with the bird community. On the other hand, acoustic composition was significantly correlated to both species richness and composition, and revealed similar ecological trends across habitats. These results suggest that acoustic composition is a more reliable method for monitoring biodiversity. We recommend its effectiveness as a biodiversity monitoring method be tested further in different global climatic regions and at various global spatial and temporal scales.

Keywords: *acoustic composition, acoustic indices, biodiversity monitoring, community composition, soundscapes.*

Introduction

Biodiversity monitoring is essential to understand the magnitude of biodiversity change induced by anthropogenic pressures (Proença et al., 2017). Gaps and biases undermine biodiversity monitoring through inconsistencies in the type of data collected and the sampling scale, as well as taxonomic biases. These all need to be addressed by new, global, long-term, and effective monitoring schemes (Proença et al., 2017).

The most common method for collecting data on biological communities involve manual surveys (Doser et al., 2021; Farnsworth et al., 2005). Birds are one of the easiest taxa to survey and point counts are the most common surveillance method. A trained observer reports all birds seen and heard; at a site, during a set amount of time, and in a set radius (Matsuoka et al., 2014). However, this type of manual survey is labour-intensive and may be biased, particularly in tropical regions where bird diversity is very high and requires extensive observer training (Buxton et al., 2016). Technological innovations have allowed for the use of ARUs (Autonomous Recording Units) in large scale studies, conducted over longer periods of time, passively, and with no observer effect (Deichmann et al., 2018; Hill et al., 2018). Furthermore, the vocal nature of birds renders ARUs particularly pertinent tools for monitoring them (Sekercioglu, Wenny & Whelan, 2016). For instance, Celis-Murillo, Deppe & Allen, (2009), found that species detection probability and relative abundance measures were better captured through ARUs than through point counts. They reason that ARUs avoid bias and are deployed for longer periods allowing to detect more cryptic species, which rarely vocalise and are harder to detect on site (Celis-Murillo, Deppe & Allen, 2009).

Within ecology, ARUs can be used for bioacoustics or soundscape analyses. Bioacoustics are the monitoring of one or more focus species (Deichmann et al., 2018).

However, the vast amounts of data ARUs generate make bioacoustics demanding, as complex machine learning algorithms are required to identify presence or absence of species of interest within audio files (Doser et al., 2021; Obrist et al., 2010.). As birds, especially songbirds (oscine passerines), can adopt wide ranges of vocalisations, creating automated systems of individual species recognition seems arduous (Obrist et al., 2010). By contrast, soundscape analyses are aimed at monitoring the effect of geophysical and anthropogenic factors on entire acoustic communities, therefore averting the need to create automated systems for individual species recognition (Deichmann et al., 2017; Sueur et al., 2014). Soundscape methods are considered the sound-focused equivalent of spatial remote sensing, due to their non-invasiveness and speed (Scarpelli, Ribeiro & Teixeira, 2021). As birds occupy a large part of the terrestrial soundscapes (from 4kHz to 8kHz and above), acoustic diversity measures could inform on avian diversity and help improve general biodiversity assessments by validating both theoretical frameworks and methodologies (Gasc et al., 2017).

Using soundscapes, acoustic indices can be extracted through processing large amounts of acoustic data (Hill et al., 2018). Over 60 indices have been proposed, but there is still some discussion as to whether these indices are biologically meaningful (Sueur et al., 2014). Just as is the case for biodiversity metrics (Mason et al., 2005; Vandewalle et al., 2010), no one acoustic index can capture soundscapes as a whole (Sueur et al., 2014). Towsey et al., (2014) recommended the use of a combination of indices to best characterise soundscapes, as efficiency of the representation can be increased up to 87%. However, acoustic indices should not be applied to all communities identically as they can be affected by geophony, dominant species, vegetation structure, and scale (Deichmann et al., 2018; Scarpelli, Ribeiro & Teixeira, 2021). There have indeed been discrepancies in the results of acoustic-focused avian studies (Bradfer-Lawrence et al., 2020). Some studies have found that only a few acoustic indices were only weakly correlated with avian diversity derived from manual survey data (Bradfer-Lawrence et al., 2020; Dröge et al., 2021; Mammides et al., 2017). Another study found a weak correlation of only a few indices with avian diversity in experimentally controlled environments (Zhao et al., 2019). This highlights the need to continue assessing the reliability of acoustic indices as indicators of avian diversity, and potentially propose new, more effective methods.

While some acoustic indices have been shown to correlate with avian species richness, a single aspect of bird diversity, little is known on the relationship between soundscapes and avian community composition (Depraetere et al., 2012; Dröge et al., 2021; Smith et al., 2020; Sueur et al., 2014). A relatively new methodology, acoustic composition, was developed to encompass community composition in soundscape analyses (Furumo & Mitchell Aide, 2019). Assuming sound recordings from specific sites are themselves communities, it is possible to treat single pixels (time-frequency bins) of spectrograms as species. Levels of amplitude indicate the abundance of these pseudo-species, which are then arranged in a "site x pseudo-species" matrix treated similarly to a traditional "site x species" matrix. A way of determining how well novel methods capture biodiversity is to compare its trends with trends of diversity metrics from manual surveys in habitat type gradients (Pijanowski et al., 2011).

In this study, we assessed the effectiveness of acoustic indices and acoustic composition in monitoring the bird community in rural England. Specifically, we asked: 1) How well does a combination of acoustic indices and acoustic composition capture variation in avian species richness and community composition? 2) Do acoustic indices and acoustic composition reveal similar ecological trends as those detected by traditional methods of sampling bird community?

Methods

Study area

The study was conducted in the Silwood Park, an Imperial College campus, located in South-East England, 25 miles west of London in Ascot, Berkshire (51-24’23.162N, 0-38’55.33”W). The 100 hectares area is composed of cultivated fields, parklands, and “Habitats of Principal Importance”, such as acid grasslands, an orchard, and wet woodlands surrounding a lake. South-East England is a low elevation (<100m above sea level on average) area with a mean annual rainfall of 697mm and mean annual temperature of 10℃.

Data collection: Fractal design

Sites were selected using a fractal design as part of “The Ecological Fractal Network” (Pearse, 2021). Fractals provide an excellent framework for capturing environmental gradients at several scales, and thus provide a good compromise between accuracy of community composition estimates and sampling effort (Simpson & Pearse, 2021). Their nested structures allow for an optimal number of sampling points in contrast to a grid, for instance, in which the number of sampling points increases with scale (Marsh & Ewers, 2013). The fractal network created for Silwood Park is composed of nine sampling points on the vertices of an equilateral triangle with 900m-long sides (to fit the size of Silwood Park, 1st order, Figure 1). These large triangles were subset into three 300m-long sided equilateral triangles (2nd order), themselves subset into three 100m-long sided triangles (3rd order). In the centre, the pattern was completed with a hexagon and a final central sampling point. All sample points were placed at least 100m from one another. Two additional 100m sided equilateral triangles were created to place sampling points in habitat types poorly represented by the fractal design. This amounted to 38 sites, which were representative of the area, as they varied in habitat, vegetation structure, and distances to roads, buildings, and the lake. Site coordinates were obtained with a GPS (GPSmap

62s) accurate to approximately 10-15meters (Garmin Ltd, 1996-2021). Sampling points were located in four habitat types: (1) Grassland, (2) Edge, (3) Woodland, (4) Wet woodland (Table IIIs in Appendix).

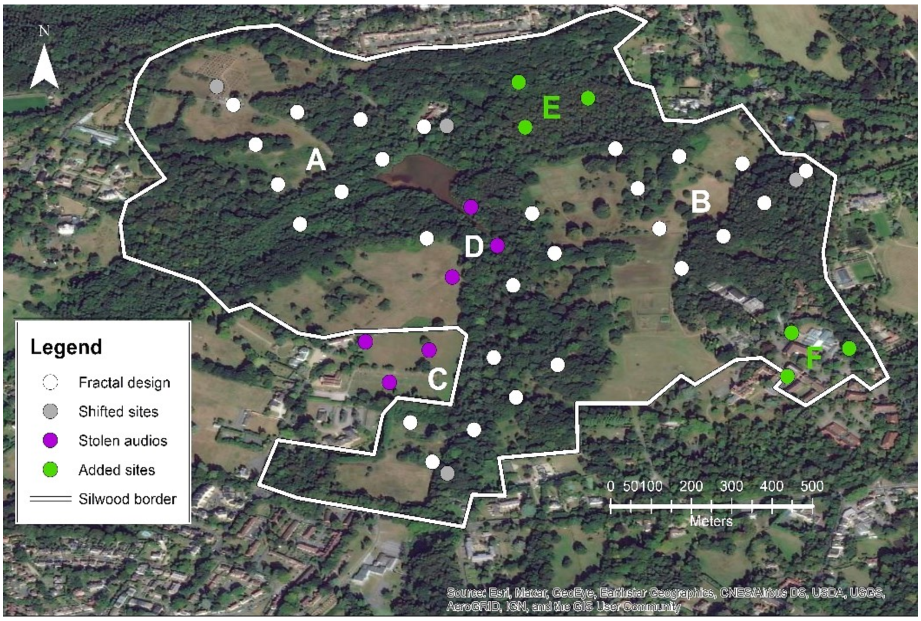


Figure 1: Map of sample sites within Silwood Park Campus. All sites within the fractal design (white points) were sampled, however some acoustic monitors were stolen (purple points). For this reason, we added some sites (green points). We also shifted some points slightly for practical purposes, such as lack of presence of trees to place audios or proximity to homes (grey points).

Acoustic monitoring

Data were collected every two weeks from 12th March to 1st May 2021, thereby capturing the beginning of Spring, the breeding season of most European birds and when bird activity is at its peak (Lack, 2008). AudioMoth (Hill et al., 2018) devices were placed at around 2 meters high and slightly tilted (~30°) towards the sky on the nearest trees to the GPS locations to best capture bird vocalisation. The peak activity periods of birds are early mornings and late evenings, which correspond to dawn and dusk choruses (Morgan et al., 1983). Using the AudioMoth application (Hill et al., 2018), the devices were set to record continuously during these dawn and dusk choruses for seven days every two weeks (Table IIs in Appendix). This resulted in 2,940 1-minute, 5.49MB *.wav* audio files per week. The sample rate was set to 48kHz as a trade-off between optimisation of birdsong resolution (most birds sing between 4-8kHz) and SD card space preservation. Whilst file names corresponded to date and time of recording, some were in hexadecimal format, so a python script was run to modify the names using Spyder (Rossum & Boer, 1991). The ARUs for sites 9D6 and 8C1 were faulty, and most of the audio files contained static, so these sites were removed from the analyses (see Figure 4s in Appendix).

Point-counts

Point count surveys were conducted by CD once a month per site during the acoustic recording periods (Table IVs). These consisted of a 5-minute buffer time and a 10-minute counting period during which the identity and abundance (i.e. number of individuals) of all bird species detected within a 50-meter radius were reported. Point counts were not conducted in rain, or when wind speed was higher than 12mph.

Acoustic analyses

Acoustic composition

For the acoustic composition analysis, we used the packages *seewave* (Sueur, Aubin & Simonis, 2008) *bioacoustics* (Marchal, Fabianek & Scott, 2021), and *tuneR* (Ligges *et al.*, 2018) in R (R Core Team, 2020). The code creates a spectrogram for each minute file, summarises the amplitude in frequency bins of 1kHz, then averages the amplitudes within the time-frequency bins of consecutive days. This produces matrices with minutes as columns and frequency bins as rows for each site. This code was run over 4 weeks-worth of audio files ranging from 6-9am and 5-7pm. Each site matrix was then gathered into one vector and all site vectors were compiled to produce one final "site x pseudo-species" matrix, termed an "acoustic composition matrix".

Acoustic indices

We used AnalysisPrograms.exe in a PowerShell script developed by the Ecoacoustics Research Group (ERG) of Queensland University of Technology to generate both spectral and summary indices for each 1-minute audio file (Towsey et al., 2018). Summary indices are single statistics representing various aspects of acoustic energy distributions within time periods across a broad frequency range, whereas spectral indices are vectors which summarise aspects of energy distribution in time periods over specific frequency bands (Oliveira *et al.*, 2021). In total, we calculated 9 acoustic indices: events per second (EPS), high frequency cover (HFC), medium frequency cover (MFC), low frequency cover (LFC), acoustic complexity index (ACI), temporal entropy (TE), cluster count (CC), three gramm count (TGC), normalised difference soundscape index (NDSI) (see Table IIIs in Appendix for a description of each).

Files of the same day were concatenated using another PowerShell script from the ERG. The program generates spectral images from the spectral indices (False Colour Spectrograms, FCS) that allowed identification of periods where geophony (wind, rain or rushing water) covered all frequencies of the spectrum, potentially masking biophony. These obstructive periods were removed from the output files of indices, as they could have biased analyses by inflating or reducing the values of some indices (Scarpelli, Ribeiro & Teixeira, 2021). FCS depict frequency (y axis) over time (x axis) (Figure 2). Colours (red, green, and blue) are inferred from three orthogonal Acoustic index. AnalysisPrograms.exe generates, by default, two spectrograms based on these groups of spectral indices: (a) ACI-ENT-EVN, and (b) BGN-

PMN-EVN. To remove obstructive periods, their characteristics were identified on spectrograms by listening to corresponding audio files for rain or wind. For instance, rain appears as a red curtain covering the spectrogram (a) or a blue curtain in the spectrogram (b). Similarly, wind is yellow on spectrogram (a), and green on (b).

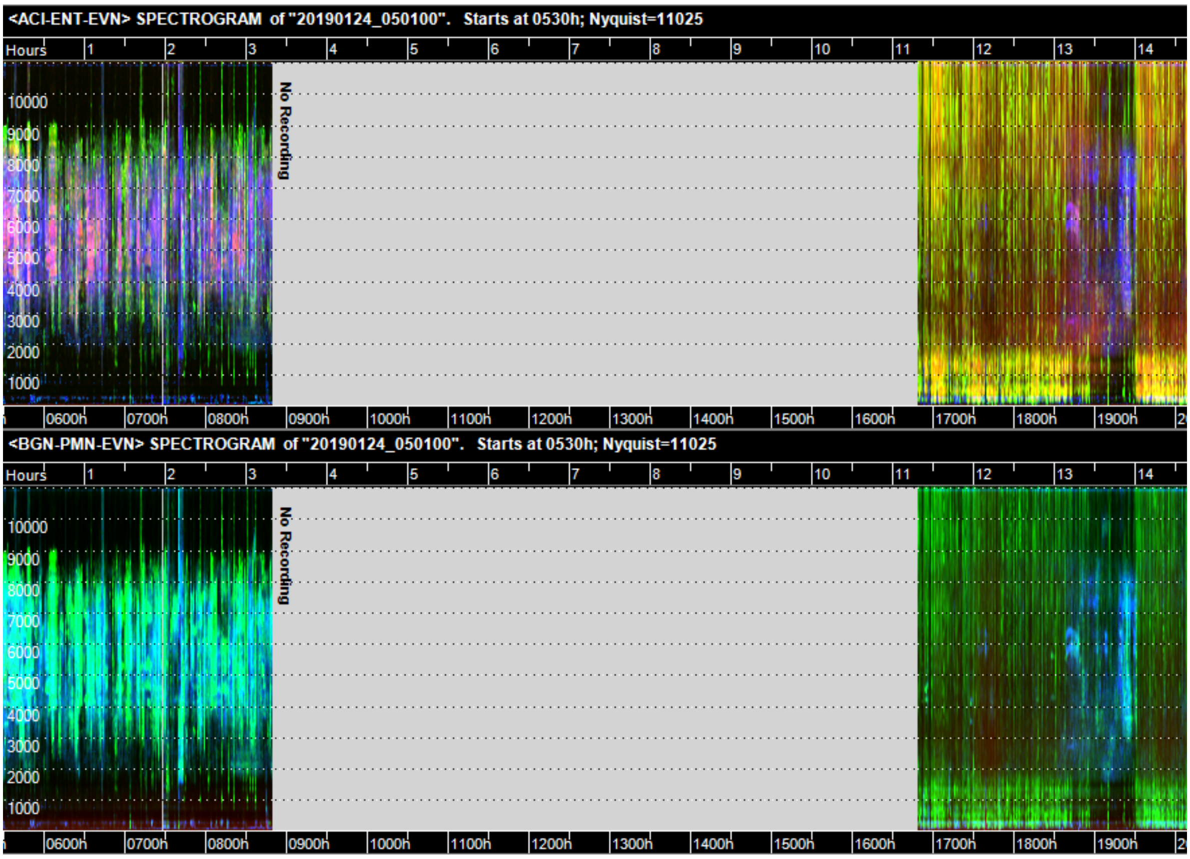


Figure 2: Example of false colour spectrograms (FCS) generated by AnalysisPrograms.exe used to identify periods of rain. They depict 24hour soundscapes, with frequency (x axis) over time (y axis) and colours inferred from the combination of spectral indices. There is a clear difference in the morning, where the soundscape is dominated by biophony, in pink and blue, and the evening where geophony dominates (represented by yellow, red and green). Times when then spectrogram was dominated by geophany, such as here, was deleted from the summary indices to avoid inflation or reduction of acoustic indices' values.

Statistical analyses

All analyses were performed in R version 4.0.3. (R Core Team, 2020).

Computing model variables

We calculated species richness per site from the point count data using R packages *reshape* (Wickam, 2007) and *vegan* (Oksanen *et al*., 2020). To calculate community composition from the point count data, we calculated the Bray-Curtis dissimilarity index in *vegan* and then performed a Principal Coordinates Analyses (PCoA) using "cmdscale". We applied the same approach to calculate acoustic composition from the matrix containing the pseudo-species amplitudes. This allowed us to reduce the data and use the scores from each PCoA as proxies for community composition and acoustic composition.

Preliminary analysis of the data revealed that the two values of goodness of fit for both the PCoAs were above 0.8 and the first two axes of these PCoAs represented together 38.56% (acoustic composition), and 54.54% (community composition) of variation. However, the results obtained with the second axes were essentially the same as those obtained for the first axes, therefore we here present only the results from the first axis (see Appendix for full results).

To calculate acoustic diversity, we first z-scaled the Acoustic indices as they varied in range. Most indices were moderately to highly correlated ($r^2 > 0.4$), so we conducted Principal Component Analysis to reduce the dimensionality of the data. The scores of the first two Principal Component (PC) axes were used in the analyses.

Preliminary analysis of the data revealed that the acoustic diversity PC1 accounted for 54.3% of the variation in the data, while acoustic diversity PC2 accounted for 21.9% of the variation (Figure 3s in Appendix). Together, the acoustic diversity PC1 and PC2 represented 76.2% of the variation, and thus were both kept.

Modelling

Species richness, a count variable, was log-transformed to correct for its positive skew, and all variables were scaled. We created six linear models to explore the relationship between each soundscape measurements and the point count measurements. For clarity, community composition axes computed from PCoA on the point count matrix will be referred to as “species composition”. Species richness calculated based on the point count data will simply be referred to as “species richness”. As for soundscape metrics, we will refer to “acoustic diversity PC1” and “acoustic diversity PC2”, and “acoustic composition”.

To understand how well a combination of acoustic diversity and acoustic composition capture variations in bird species richness and composition, we ran a series of single regressions, where we had as response variables: acoustic diversity (PC1 or PC2) and acoustic composition, and as explanatory variables: species richness and species composition.

To answer the question “Do we detect the same ecological patterns in bird diversity with data collected from point counts and soundscape analyses?” we explored how these variables changed across the different Silwood habitat types. We thus created a set of linear regressions with each diversity/composition metric as a response, and habitat type as an explanatory variable. Species richness was a count variable following a Poisson distribution.

Spatial autocorrelation (SAC) violates the assumption of error independence in linear models (Moran, 1950) . SAC was verified in model residuals using *spdep* (Bivand & Wong, 2018). First, we created a matrix of nearest neighbours using latitude and longitude values for each site with “knearneigh” (setting k=2, 3rd order of the fractal design). It was converted to a list using “knn2nb”, and then a neighbourhood matrix was created in list format with “nb2listw”. We used the “moran.test” function to calculate Moran’s I values. It calculates a correlation value, whilst considering the neighbourhood matrix, and estimates similarities among outliers of a variable (here the residuals) that are spatially aggregated. It then compares the estimate to an expected variable in which value similarity is not influenced by position. If the observed and expected estimates are non-significantly different, the H0 is accepted, there is no spatial autocorrelation, and the model’s assumptions are not violated. The observed statistic was tested using “moran.mc”, with nsim=600 permutations of the residuals, which also takes the spatial weights list into account.

For the models where H0 was rejected, and there was spatial autocorrelation in model residuals, we constructed simultaneous autoregressive models (SAR models), which use a neighbourhood matrix to account for the relationship values of a variable of interest (here model residuals) and those of neighbours (F. Dormann et al., 2007). These models allowed to account for SAC when necessary, using *spatialreg* (Bivand, Hauke, & Kossowski, 2013). The possible models were “spatial error models” (SAR_{err}), which assume that the autoregressive process is found in the error term, “spatial lag models” (SAR_{lag}), which assume that it takes place in the response variable, and “spatial mixed models” (SAR_{mix}), which assume it takes place in both the response and explanatory variables (Kissling & Carl, 2007). To identify the correct model structure, we conducted Lagrange multiplier diagnostics for spatial dependence (Anselin et al., 1996).

Moran’s I estimate was significantly different from the expected value in three models (Tables Vs and VIs in Appendix). Using the Lagrange multiplier diagnostics for spatial dependence, we selected SAR_{lag} for the first two models, and SAR_{err} for the third. The SAR models effectively eliminated SAC in the residuals and the summary statistics are in Tables VIIs and VIIIs of the Appendix.

Results

In all, 52 bird species were identified (Table IVs in Appendix), with a mean species richness of 9.3 and a mean abundance of 22.9 during point counts. The most common species detected were *Corvus monedula* (Jackdaw), *Cyanistes caeruleus* (Blue tit), and *Parus major* (Great tit). Although we should have collected a total of 196 hours of audio files per site 6 audio devices were stolen during week 4 (30/04/2021) (Figure 1) and other malfunctions occurred, such as SD or batteries being removed from their slots or rain and wind knocking the audios over. Therefore, a total of 6,958 hours of recording was returned.

No significant relationship nor trend was found between acoustic diversity PC1 and species richness or species composition (Figure 3A). Acoustic diversity PC2 was negatively correlated with species richness (F=9.169, p=0.005, Figure 3B), and positively correlated with species composition (F=6.498, p=0.004, Figure 3B).

Acoustic composition had a significant positive relationship with species richness (F=8.312, p=0.007, Figure 3C), and negatively correlated with species composition (F=8.732, p=0.001, Figure 3C).

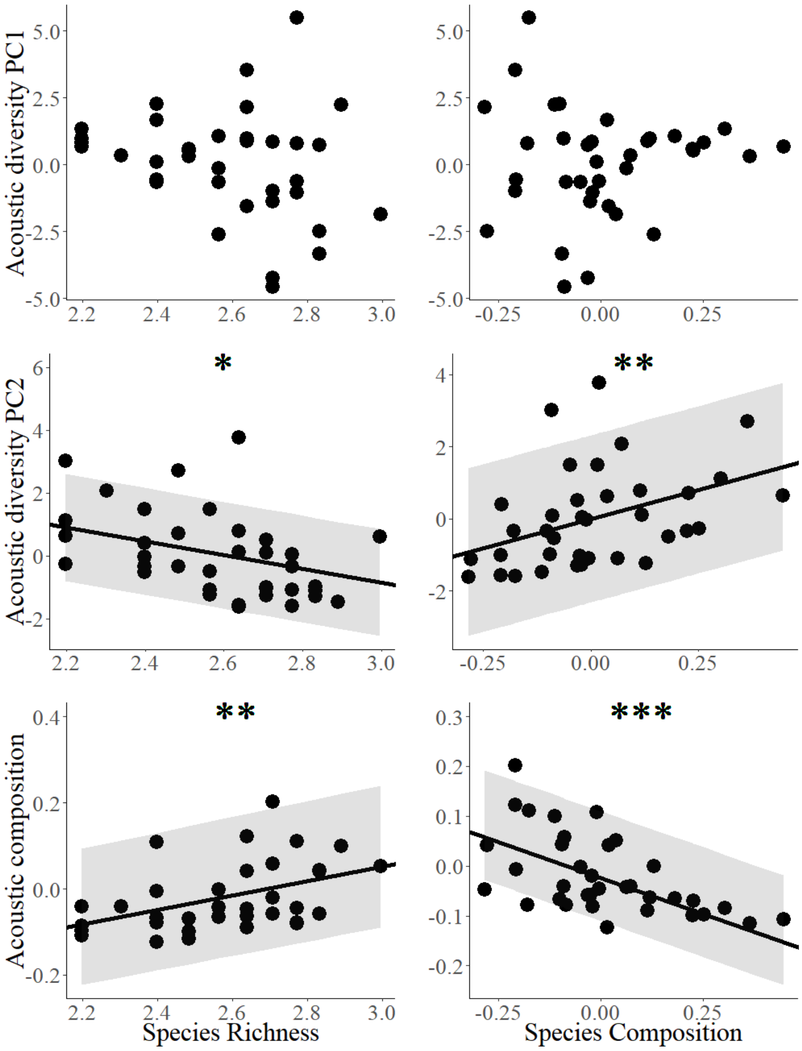


Figure 3: Regression plots depicting relationships between soundscape metrics and point count metrics. The significance level is symbolised through stars: * <0.05, ** <0.01, *** <0.001. The estimated slope of significant relationships is represented with a black line, and its respective 95% confidence interval with a grey ribbon.

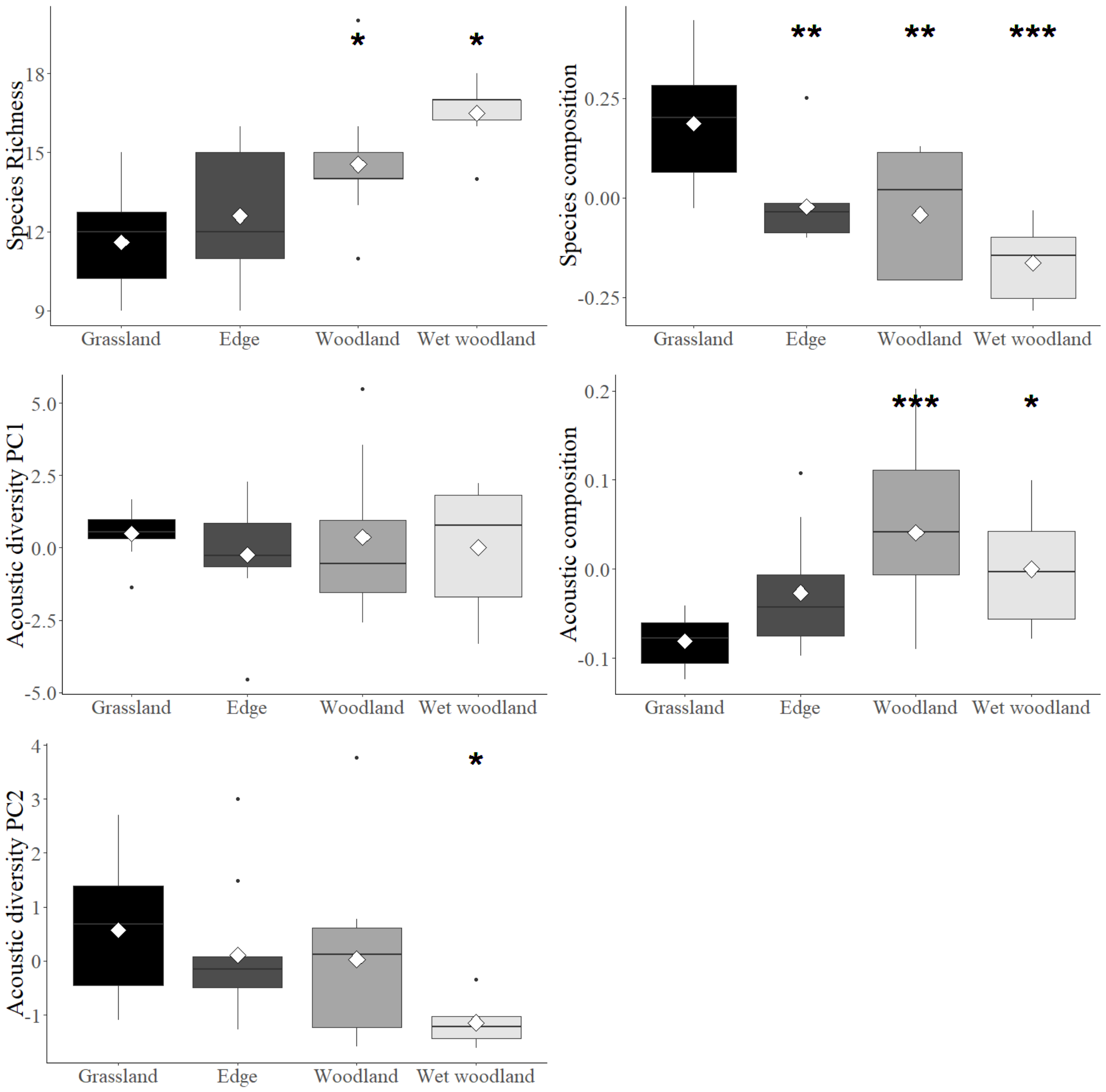


Figure 4: Trends of diversity metrics according to habitat types. Boxplots are completed with white diamonds to represent means. The linear models took grassland habitats as a base level for mean comparison. The significance level is symbolised with stars (<0.05, **<0.01, ***<0.001).*

Species richness was higher in woodlands and wet woodlands, and species composition in the grassland was different to all other habitats (Figure 4). Acoustic composition followed the same trend as species richness and was significantly different in woodland and wet woodland compared to other habitats (Figure 4). Acoustic diversity PC1 did not vary significantly according to habitat type, but acoustic diversity PC2 was lower in wet woodlands compared to other habitats (Figure 4).

Discussion

Our results show that most acoustic diversity indices, as represented by the first PCA axis, were a poor measure of either bird species richness or species composition. However, the second axis, which is highly influenced by NDSI and LFC, was correlated to changes in species richness and composition, but did not reflect the same ecological trends across habitats. Acoustic composition, on the other hand, was strongly correlated to species richness and composition, and it also reflected similar ecological trends across habitats. These findings suggest that acoustic composition is a simpler, more effective, and intuitive method for measuring soundscapes; and, it has the potential to make biodiversity monitoring easier and more reliable.

Acoustic composition was correlated to species richness whilst the combination of acoustic indices was not. Acoustic indices are single values per minute over large frequency bands, usually averaged over long periods of time (Oliveira et al., 2021). This could lead to signal masking, where dominant species or noises mask target species (Metcalf et al., 2021). Metcalf et al., (2021) calculated two acoustic indices (ACI, Bioacoustic Index) at narrower time-frequency bins and determined that the efficacy of the acoustic indices was significantly increased. The acoustic composition method averages amplitudes (the pseudo-abundances) per minute and per frequency, then keeps all measures in site matrices (Damstra, unpublished; Furumo & Mitchell Aide, 2019). This implies that the time-frequency bins are much narrower and birds that rarely sing or sing at a particular time of day will be given more weight. However, there is a possibility that birds of different species vocalising at the same frequency and time could be perceived as the same pseudo-species in the acoustic composition matrix. For example, Goldcrest (*Regulus regulus*) calls are similar to Fircrest (*Regulus ignicapillus*) songs, which are both produced at around 9-10 kHz. As these two species are often seen foraging together at Silwood, they may have been attributed the same pseudo-species. In that sense, acoustic composition could underestimate species richness. However, the time-frequency bins remain narrower using acoustic composition rather than the acoustic indices, and thus reflect the diversity of vocal species more accurately.

The acoustic composition method gave a representative measurement of the avian species composition in Silwood Park, whilst the combination of acoustic indices did not. There was a clear difference between the grassland acoustic community and the wetland acoustic community, confirmed by manual surveys. This shows that the acoustic composition method reflects real change in avian communities across habitats. This conforms with a recent study on landscape structure in grasslands, which found that a measure of acoustic composition represented species composition of birds better than single indices (Müller et al., 2022). Furumo & Mitchell Aide (2019), who were the first to use acoustic composition as a way of determining differences in species composition between habitat types, discovered specialist bird species exclusive to forests and absent in plantations, highlighting the value of the acoustic composition method for conservation efforts.

The combination of acoustic indices was ineffective in capturing trends of bird diversity metrics calculated from point counts. These results were unexpected because British biophony is mostly dominated by birds (Gasc et al., 2017), and the acoustic indices associated with this PCA axis have been shown in other studies to be correlated to biodiversity (Buxton et al., 2016). It is possible that these acoustic indices are reflecting the number of bird calls rather than the variety of calls (Alcocer et al., 2022, Bradfer-Lawrence et al., 2020; Lawson, unpublished). In other words, highly vocal species present in some habitats, such as jackdaws (*Corvus monedula*) or ring-necked parakeets (*Psittacula krameri*) in grasslands, could inflate acoustic diversity values so that no change can be detected across habitats. Previous studies have also raised red flags over the use of acoustic indices. A study testing acoustic indices in urban environments found that aircraft noise, which is quite frequent over our study site, could bias results so that they are not correlated with variety of bird calls (Devos, 2016). Another found that most of the acoustic indices tested were biased by anthrophony or geophony (Fairbrass et al., 2017). This implies that a study using a metric of combined acoustic indices might overlook significant biodiversity change in habitats and infer that none of the habitats harbour a greater number of species, or a more sensitive community.

It is noteworthy that acoustic diversity PC2 was correlated with the species composition axes. Acoustic diversity PC2 represents acoustic activity in the lower frequency bands (1-1000Hz) and the ratio of activity between the lower (1000-2000Hz) and mid-frequency bands (1000-8000Hz). Previous studies have suggested that NDSI captures anthrophony (Fuller et al., 2015; Kasten et al., 2012; Machado, Aguiar & Jones, 2017), though these results are difficult to generalise across regions. Avian communication is indeed affected by human noise, or noise pollution, with evidence emerging to suggest that bird populations alter the frequency of their calls in urban soundscapes (Schindler, Gerber & Quinn, 2020; Phillips et al., 2020). While we cannot be certain what environmental changes acoustic diversity PC2 is reflecting (i.e., bird community or anthrophony), it is clear that this metric is not able to capture the same ecological trends across habitats as acoustic composition. Future research should further investigate what sounds are captured by NDSI and LFC and the extent to which anthrophony affects bird communication and behaviour, as this has profound consequences for population monitoring.

As this study demonstrated, ARUs are sensitive to malfunction and data can be lost at many stages of the fieldwork process. Furthermore, pre-processing of the large amounts of audio data is extremely long, but less so using the novel method. Pre-processing for one month's worth of audio data per site took approximately 24h50min using the acoustic indices, while it took 4h45min using the acoustic composition method. Future research could aim to increase efficiency of pre-programming and lower the cost of ARUs that allow real-time transfer of audio data. In the meantime, the advantages of ARUs outweigh the limitations and allow informative soundscape methods, such as the novel acoustic composition method, to replace laborious manual surveys.

In this study, the acoustic composition method captured bird diversity measurements better than a combination of biotic acoustic indices. This novel method shows great promise for enabling effective, and reliable ways of conducting fieldwork and processing large amounts of audio data. It could help quantify various components of biodiversity, thus improving methods of testing and approving action plans to preserve ecosystem health. "Remote sensing" data collection such as the acoustic composition method may contribute to creating a global and unbiased database of biodiversity change. We recommend this method be tested at different spatial and temporal scales to test if it is unaffected by scale-dependency, and in various climatic regions where dominant acoustic taxa are not birds, to test its robustness as a measure of biodiversity.

Acknowledgements

I wish to thank Yewshen Lim for helping regularly with data collection, and other less regular helpers.

Data and code availability

The code I wrote and needed to run the code for the fulfilment of this thesis can be found at: Sounds_Silwood_Code

The code written by Elise Damstra for the acoustic composition method can be found at: https://github.com/emdo1/Acoustic_compostition

The Analysesprograms.exe code for pre-processing and calculation of the acoustic indices can be found at: https://zenodo.org/record/1188744#.YSY4IkuSlPY

References

Journal articles and books

Alcocer I, Lima H, Sugai LSM, Llusia D. (2022) Acoustic indices as proxies for biodiversity: a meta-analysis. Biol Rev Camb Philos Soc. doi: 10.1111/brv.12890. Epub ahead of print. PMID: 35978471

Anselin, L., Bera, A.K., Florax, R. & Yoon, M.J. (1996) Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics. 26 (1), 77–104. doi:10.1016/0166-0462(95)02111-6.

Bradfer-Lawrence, T., Bunnefeld, N., Gardner, N., Willis, S.G. & Dent, D.H. (2020) Rapid assessment of avian species richness and abundance using acoustic indices. Ecological Indicators. 115, 106400. doi:10.1016/j.ecolind.2020.106400.

Buxton, R.T., Brown, E., Sharman, L., Gabriele, C.M. & McKenna, M.F. (2016) Using bioacoustics to examine shifts in songbird phenology. Ecology and Evolution. 6 (14), 4697–4710. doi:10.1002/ece3.2242.

Celis-Murillo, A., Deppe, J.L. & Allen, M.F. (2009) Using soundscape recordings to estimate bird species abundance, richness, and composition. Journal of Field Ornithology. 80 (1), 64–78. doi:10.1111/j.1557-9263.2009.00206.x.

Deichmann, J.L., Acevedo-Charry, O., Barclay, L., Burivalova, Z., Campos-Cerqueira, M., d'Horta, F., Game, E.T., Gottesman, B.L., Hart, P.J., Kalan, A.K., Linke, S., Nascimento, L.D., Pijanowski, B., Staaterman, E. & Mitchell Aide, T. (2018) It's time to listen: there is much to be learned from the sounds of tropical ecosystems. Biotropica. 50 (5), 713–718. doi:10.1111/btp.12593.

Deichmann, J.L., Hernández-Serna, A., Delgado C., J.A., Campos-Cerqueira, M. & Aide, T.M. (2017) Soundscape analysis and acoustic monitoring document impacts of natural gas exploration on biodiversity in a tropical forest. Ecological Indicators. 74, 39–48. doi:10.1016/j.ecolind.2016.11.002.

Depraetere, M., Pavoine, S., Jiguet, F., Gasc, A., Duvail, S. & Sueur, J. (2012) Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland. Ecological Indicators. 13 (1), 46–54. doi:10.1016/j.ecolind.2011.05.006.

Devos, P. (2016), August. Soundecology indicators applied to urban soundscapes. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings 253 (5), 3631-3638. Institute of Noise Control Engineering.

Doser, J.W., Finley, A.O., Weed, A.S. & Zipkin, E.F. (2021) Integrating automated acoustic vocalization data and point count surveys for estimation of bird abundance. Methods in Ecology and Evolution. 12 (6), 1040–1049. doi:10.1111/2041-210X.13578.

Dröge, S., Martin, D.A., Andriafanomezantsoa, R., Burivalova, Z., Fulgence, T.R., Osen, K., Rakotomalala, E., Schwab, D., Wurz, A., Richter, T. & Kreft, H. (2021) Listening to a changing landscape: Acoustic indices reflect bird species richness and plot-scale vegetation structure across different land-use types in north-eastern Madagascar. Ecological Indicators. 120, 106929. doi:10.1016/j.ecolind.2020.106929.

F. Dormann, C., M. McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., G. Davies, R., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., R. Peres-Neto, P., Reineking, B., Schröder, B., M. Schurr, F. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography. 30 (5), 609–628. doi:10.1111/j.2007.0906-7590.05171.x.

Fairbrass, A.J., Rennert, P., Williams, C., Titheridge, H. & Jones, K.E. (2017) Biases of acoustic indices measuring biodiversity in urban areas. Ecological Indicators. 83, 169–177. doi:10.1016/j.ecolind.2017.07.064.

Farnsworth, G.L., Nichols, J.D., Sauer, J.R., Fancy, S.G., Pollock, H., Shriner, S.A. & Simons, T.R. (2005) Statistical Approaches to the Analysis of Point Count Data: A Little Extra Information Can Go a Long Way. 9.

Fuller, S., Axel, A.C., Tucker, D. & Gage, S.H. (2015) Connecting soundscape to landscape: Which acoustic index best describes landscape configuration? Ecological Indicators. 58, 207–215. doi:10.1016/j.ecolind.2015.05.057.

Furumo, P.R. & Mitchell Aide, T. (2019) Using soundscapes to assess biodiversity in Neotropical oil palm landscapes. Landscape Ecology. 34 (4), 911–923. doi:10.1007/s10980-019-00815-w.

Gasc, A., Francomano, D., Dunning, J.B. & Pijanowski, B.C. (2017) Future directions for soundscape ecology: The importance of ornithological contributions. The Auk. 134 (1), 215–228. doi:10.1642/AUK-16-124.1.

Hill, A.P., Prince, P., Piña Covarrubias, E., Doncaster, C.P., Snaddon, J.L. & Rogers, A. (2018) AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment N. Isaac (ed.). Methods in Ecology and Evolution. 9 (5), 1199–1211. doi:10.1111/2041-210X.12955.

Kasten, E.P., Gage, S.H., Fox, J. & Joo, W. (2012) The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology. Ecological Informatics. 12, 50–67. doi:10.1016/j.ecoinf.2012.08.001.

Kissling, W.D. & Carl, G. (2007) Spatial autocorrelation and the selection of simultaneous autoregressive models. Global Ecology and Biogeography. 0 (0), 070618060123007-??? doi:10.1111/j.1466-8238.2007.00334.x.

Lack, D. (1950) The breeding season of European birds. Ibis. 92 (2), 288–316. doi:10.1111/j.1474-919X.1950.tb01753.x.

Machado, R.B., Aguiar, L. & Jones, G. (2017) Do acoustic indices reflect the characteristics of bird communities in the savannas of Central Brazil? Landscape and Urban Planning. 162, 36–43. doi:10.1016/j.landurbplan.2017.01.014.

Mammides, C., Goodale, E., Dayananda, S.K., Kang, L. & Chen, J. (2017) Do acoustic indices correlate with bird diversity? Insights from two biodiverse regions in Yunnan Province, south China. Ecological Indicators. 82, 470–477. doi:10.1016/j.ecolind.2017.07.017.

Marsh, C.J. & Ewers, R.M. (2013) A fractal-based sampling design for ecological surveys quantifying β-diversity D. Murrell (ed.). Methods in Ecology and Evolution. 4 (1), 63–72. doi:10.1111/j.2041-210x.2012.00256.x.

Mason, N.W.H., Mouillot, D., Lee, W.G. & Wilson, J.B. (2005) Functional richness, functional evenness and functional divergence: the primary components of functional diversity. Oikos. 111 (1), 112–118. doi:10.1111/j.0030-1299.2005.13886.x.

Matsuoka, S.M., Mahon, C.L., Handel, C.M., Sólymos, P., Bayne, E.M., Fontaine, P.C. & Ralph, C.J. (2014) Reviving common standards in point-count surveys for broad inference across studies. The Condor. 116 (4), 599–608. doi:10.1650/CONDOR-14-108.1.

Metcalf, O.C., Barlow, J., Devenish, C., Marsden, S., Berenguer, E. & Lees, A.C. (2021) Acoustic indices perform better when applied at ecologically meaningful time and frequency scales R. Freckleton (ed.). Methods in Ecology and Evolution. 12 (3), 421–431. doi:10.1111/2041-210X.13521.

Moran, P.A. (1950) Notes on continuous stochastic phenomena. Biometrika, 37(1/2), pp.17-23.

Morgan, B.J.T., North, P.M., Ralph, C.J. & Scott, J.M. (1983) Estimating Numbers of Terrestrial Birds. Biometrics. 39 (4), 1123. doi:10.2307/2531357.

Müller, S., Gossner, M.M., Penone, C., Jung, K., Renner, S.C., et al. (2022) Land-use intensity and landscape structure drive the acoustic composition of grasslands. Agriculture, Ecosystems & Environment. 328, 107845. doi:10.1016/j.agee.2021.107845.

Obrist, M.K., Pavan, G., Sueur, J., Riede, K., Llusia, D. & Márquez, R. (2010) Chapter 5 Bioacoustics approaches in biodiversity inventories. Abc Taxa. 8, 68-99.

Oliveira, E.G., Ribeiro, M.C., Roe, P. & Sousa-Lima, R.S. (2021) The Caatinga Orchestra: Acoustic indices track temporal changes in a seasonally dry tropical forest. Ecological Indicators. 129, 107897. doi:10.1016/j.ecolind.2021.107897.

Phillips, J.N., Rochefort, C., Lipshutz, S., Derryberry, G.E., Luther, D. & Derryberry, E.P. (2020) Increased attenuation and reverberation are associated with lower maximum frequencies and narrow bandwidth of bird songs in cities. Journal of Ornithology. 161 (2), 593–608. doi:10.1007/s10336-020-01751-2.

Pijanowski, B.C., Villanueva-Rivera, L.J., Dumyahn, S.L., Farina, A., Krause, B.L., Napoletano, B.M., Gage, S.H. & Pieretti, N. (2011) Soundscape Ecology: The Science of Sound in the Landscape. BioScience. 61 (3), 203–216. doi:10.1525/bio.2011.61.3.6.

Proença, V., Martin, L.J., Pereira, H.M., Fernandez, M., McRae, L., Belnap, J., Böhm, M., Brummitt, N., García-Moreno, J., Gregory, R.D., Honrado, J.P., Jürgens, N., Opige, M., Schmeller, D.S., Tiago, P. & van Swaay, C.A.M. (2017) Global biodiversity monitoring: From data sources to Essential Biodiversity Variables. Biological Conservation. 213, 256–263. doi:10.1016/j.biocon.2016.07.014.

Scarpelli, M.D.A., Ribeiro, M.C. & Teixeira, C.P. (2021) What does Atlantic Forest soundscapes can tell us about landscape? Ecological Indicators. 121, 107050. doi:10.1016/j.ecolind.2020.107050.

Schindler, A.R., Gerber, J.E. & Quinn, J.E. (2020) An ecoacoustic approach to understand the effects of human sound on soundscapes and avian communication. Biodiversity. 21 (1), 15–27. doi:10.1080/14888386.2020.1713889.

Sekercioglu, Ç.H., Wenny, D.G. & Whelan, C.J. (2016) Why Birds Matter: Avian Ecological Function and Ecosystem Services. University of Chicago Press. doi:10.7208/chicago/9780226382777.001.0001.

Simpson, E.G. & Pearse, W.D. (2021) Fractal triads efficiently sample ecological diversity and processes across spatial scales. Oikos, 130: 2136-2147. https://doi.org/10.1111/oik.08272

Smith, D.G., Truskinger, A., Roe, P. & Watson, D.M. (2020) Do acoustically detectable species reflect overall diversity? A case study from Australia's arid zone N. Pettorelli & V. Lecours (eds.). Remote Sensing in Ecology and Conservation. 6 (3), 286–300. doi:10.1002/rse2.173.

Sueur, J., Farina, A., Gasc, A., Pieretti, N. & Pavoine, S. (2014) Acoustic Indices for Biodiversity Assessment and Landscape Investigation. Acta Acustica united with Acustica. 100 (4), 772–781. doi:10.3813/AAA.918757.

Towsey, M., Wimmer, J., Williamson, I. & Roe, P. (2014) The use of acoustic indices to determine avian species richness in audio-recordings of the environment. Ecological Informatics. 21, 110–119. doi:10.1016/j.ecoinf.2013.11.007.

Vandewalle, M., de Bello, F., Berg, M.P., Bolger, T., Dolédec, S., et al. (2010) Functional traits as indicators of biodiversity response to land use changes across ecosystems and organisms. Biodiversity and Conservation. 19 (10), 2921–2947. doi:10.1007/s10531-010-9798-9.

Zhao, Z., Xu, Z., Bellisario, K., Zeng, R., Li, N., Zhou, W. & Pijanowski, B.C. (2019) How well do acoustic indices measure biodiversity? Computational experiments to determine effect of sound unit shape, vocalization intensity, and frequency of vocalization occurrence on performance of acoustic indices. Ecological Indicators. 107, 105588. doi:10.1016/j.ecolind.2019.105588.

Software and R packages

Marchal, J., Fabianek, F., & Scott C. (2021) bioacoustics: Analyse Audio Recordings and Automatically Extract Animal Vocalizations. R package version 0.2.5. CRAN.R-project.org/package=bioacoustics

Bivand, R. S., Hauke, J., and Kossowski, T. (2013). Computing the Jacobian in Gaussian spatial autoregressive models: An illustrated comparison of available methods. Geographical Analysis, 45(2), 150-179. https://doi.org/10.1111/gean.12008

Bivand, R. S, & Wong, D. W. S. (2018) Comparing implementations of global and local indicators of spatia lassociation TEST, 27(3), 716-748. https://doi.org/10.1007/s11749-018-0599-x

Oksanen, J.F., Blanchet, G., Friendly, M., Kindt, R., Legendre, P., McGlin, D., Minchin, PR., O'Hara, Simpson, G.L., Solymos, P., Stevens, M.H., Szoecs, E., and Wagner, H. (2020). vegan: Community Ecology Package. Rpackage version 2.5-7. CRAN.R-project.org/package=vegan.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/.

Rossum, G., Boer, J. (1991) Interactively Testing Remote Servers Using the Python Programming Language. CWI Quarterly. 4(4)283–303.

Sueur J., Aubin T., & Simonis C. (2008) “Seewave: a free modular tool for sound analysis and synthesis.” Bioacoustics. 18, 213-226. doi/abs/10.1080/09524622.2008.9753600.

Towsey, M., Truskinger, A., Cottman-Fields, M., Roe. P. (2018). Ecoacoustics Audio Analysis Software v18.03.0.41 (Version v18.03.0.41). Zenodo. doi.org/10.5281/zenodo.1188744

Ligges, U., Krey, S., Mersmann, O., and Schnackenberg, S. (2018). tuneR: Analysis of Music and Speech. https://CRAN.R-project.org/package=tuneR

Webpages

Pearse, W. (2021) EcoFracNet. Site made with Jekyll, from Allan Lab template. ecofracnetwork.github.io/ [Accessed 11th August 2021]

Garmin Ltd. (1996-2021) Garmin.com. www.garmin.com/en-GB/ [Accessed 9th March 2021]

Appendix

Methods

Habitat types:

Using ArcMap 10.6.1., I selected sample sites from Prof. Crawley's vegetation surveys that were closest to my own sampling sites and classified them accordingly (Table Is). I chose to keep the four main habitat types (grassland, edge, woodland, wet woodland), and changed the category of sites in less frequent habitats to the closest main habitats based on similarity. I changed habitat category of sites to "Edge" when these were at the border of habitats differing in openness.

Table Is: Classification of sites into habitat and dates and times of point count surveys. Each site was surveyed once a month.

Fractal Site	Longitude	Latitude	Habitat Type	March point count dates	Point count time 1	April point count dates	Point count time 2
0A1	-0.652	51.414	Edge	13.03	07:05	13.04	18:12
0A2	-0.652	51.415	Grassland	13.03	07:23	13.04	18:30
0A3	-0.651	51.415	Edge	13.03	07:36	13.04	18:45
1A1	-0.649	51.414	Wet woodland	1.04	17:25	29.04	06:46
1A2	-0.650	51.414	Woodland	1.04	17:10	29.04	07:03
1A3	-0.648	51.414	Wet woodland	29.03	06:56	29.04	06:30
2A1	-0.651	51.412	Wet woodland	29.03	07:27	1.05	08:25
2A2	-0.651	51.413	Grassland	29.03	07:42	1.05	08:40
2A3	-0.650	51.413	Wet woodland	29.03	07:12	1.05	08:05
3B1	-0.643	51.413	Edge	2.04	07:45	26.04	17:36
3B2	-0.644	51.414	Edge	2.04	07:30	26.04	17:20
3B3	-0.642	51.414	Woodland	2.04	07:15	26.04	17:00
4B1	-0.641	51.413	Woodland	1.04	07:10	30.04	18:45
4B2	-0.641	51.413	Grassland	1.04	07:25	30.04	19:20
4B3	-0.640	51.413	Woodland	1.04	07:40	30.04	19:00
5B1	-0.642	51.411	Edge	16.03	07:20	14.04	18:15

5B2	-0.643	51.412	Grassland	16.03	07:04	14.04	18:30
5B3	-0.641	51.412	Woodland	16.03	06:50	14.04	18:45
6C1	-0.649	51.409	Grassland	3.04	07:28	30.04	08:05
6C2	-0.649	51.410	Grassland	3.04	07:55	30.04	07:45
6C3	-0.648	51.409	Grassland	3.04	07:40	30.04	08:15
7C1	-0.646	51.409	Grassland	13.03	16:00	16.04	06:32
7C2	-0.647	51.409	Edge	13.03	16:20	16.04	06:18
7C3	-0.645	51.409	Woodland	13.03	15:45	16.04	06:07
8C1	-0.648	51.407	Woodland	16.03	16:13	13.04	06:11
8C2	-0.648	51.408	Wet woodland	16.03	16:30	13.04	06:30
8C3	-0.647	51.408	Wet woodland	16.03	16:53	13.04	06:45
9D1	-0.648	51.411	Grassland	3.04	08:12	30.04	07:30
9D2	-0.648	51.412	Edge	15.03	07:40	16.04	17:53
9D3	-0.647	51.413	Woodland	15.03	07:00	16.04	17:25
9D4	-0.647	51.412	Wet woodland	15.03	07:20	16.04	17:38
9D5	-0.646	51.411	Edge	29.03	17:24	14.04	07:40
9D6	-0.646	51.412	Grassland	29.03	16:52	14.04	07:10
9D7	-0.645	51.412	Edge	29.03	17:10	14.04	07:25
E1	-0.646	51.414	Wet woodland	3.04	06:58	27.04	18:25
E2	-0.646	51.415	Woodland	3.04	06:43	27.04	18:55
E3	-0.645	51.415	Woodland	3.04	07:14	27.04	18:40
F1	-0.640	51.409	Grassland	17.03	07:02	27.04	08:00
F2	-0.640	51.410	Edge	17.03	06:48	27.04	07:40
F3	-0.639	51.410	Edge	17.03	07:15	27.04	10:13

Sampling dates and times:

Every two weeks, audios were placed at sampling sites to record 0h30 minutes before sunrise and 1h30 after sunrise, and 1.30 before sunset and 0.30h after sunset (Table IIs). During recording periods at dawn and dusk, I also conducted point counts, one a month per site (Table Is).

Table IIs: Timetable of recording periods, capturing the start of Spring season, the peak activity period of birds.

Week	**Start date**	**End date**	**Dawn times**	**Dusk times**
1	12/03	18/03	6-9.30am	3-6.30pm
2	27/03	03/04	6-9.30am	3.30-7pm
3	10/04	17/04	5.30-9am	5-8.30pm
4	24/04	01/05	5.30-9am	5-8.30pm

Acoustic Indices:

Table IIIs: Description of each acoustic index mentioned in this thesis, and reference to the first paper in which they were developed.

Index	**Description of measure**	**Reference**
Background Noise (BGN)	Remaining noise (dB) in the waveform after removal of acoustic activity. This index is used in calculations of several other indices but was not included in our analyses.	Towsey (2014)
Events per Second (EPS)	The number of acoustic events (= the signal envelope crosses a BGN + 3dB threshold) in a second.	Towsey (2014)
High Frequency Cover (HFC)	The number of noise-reduced spectrogram cells exceeding 3dB above BGN in the high frequency band (8000-11025Hz).	Towsey (2014)
Mid Frequency Cover (MFC)	The number of noise-reduced spectrogram cells exceeding 3dB above BGN in the mid frequency band (1000-8000Hz).	Towsey (2014)
Low Frequency Cover (LFC)	The number of noise-reduced spectrogram cells exceeding 3dB above BGN in the low frequency band (1-1000Hz). Is often linked to human-induced noise.	Towsey (2014)
Acoustic Complexity Index (ACI)	Short-time averaged changes in energy across frequency bins in consecutive spectrums (main value is the average over 1 minute). It is meant to exclude constant low-frequency sounds and noises (geophony or anthrophony) and include short signals, reflecting complexity when many differences occur.	Pieretti, Farina, et Morri (2011)

Temporal Entropy (H[t], or TE)	Concentration of mean energy, or Shannon entropy of probability mass function, of each frequency bin in the amplitude envelope.	(Sueur et al., 2008)
Cluster Count (CC)	Number of spectral clusters (= spectral diversity) in the mid frequency band (1000-8000Hz).	Towsey (2014)
Three Gramm Count (TGC)	Derived measure from CLS and repeated sequences	Towsey (2014)
Normalised Difference Soundscape Index (NDSI)	NDSI= (biophony – anthrophony) / (biophony + anthrophony). Gives a value between 0 and 1, where 1 reflects pure biotic sounds, and 0 reflects pure anthrophony.	(Kasten *et al.*, 2012)

Preliminary results:

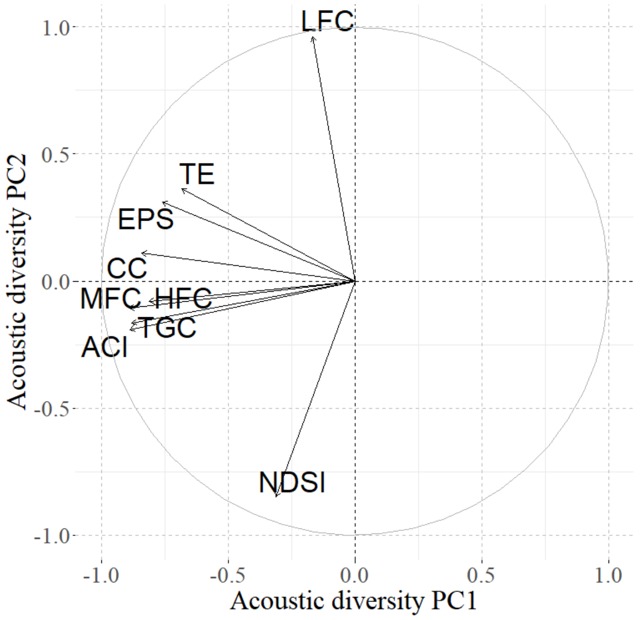


Figure 1s: Variable correlation plot depicting correlations between each index and the first two PCs. Most acoustic indices contribute to PC1, whereas NDSI and LFC are loaded onto PC2.

Results

Table IVs: The 52 bird species identified during the point counts.

Barn owl, *Tyto alba*	Common kingfisher, *Alcedo atthis*
Blackbird, *Turdus merula*	Long-tailed tit, *Aegithalos caudatus*
Blackcap, *Sylvia atricapilla*	Eurasian magpie, *Pica pica*
Blue tit, *Cyanistes caeruleus*	Mallard duck, *Anas platyrhynchos*
Buzzard, *Buteo buteo*	Mandarin duck, *Aix galericulata*
Canada goose, *Branta canadensis*	Mistle thrush, *Turdus viscivorus*
Carrion crow, *Corvus corone*	Common Moorhen, *Gallinula chloropus*
Chaffinch, *Fringilla coelebs*	Eurasian nuthatch, *Sitta europaea*
Chiffchaff, *Phylloscopus collybita*	Ring-necked pheasant, *Phasianus colchicus*
Coal tit, *Periparus ater*	Pied wagtail, *Motacilla alba*
Dunnock, *Prunella modularis*	Raven, *Corvus corax*
Egyptian goose, *Alopochen aegyptiaca*	Red kite, *Milvus milvus*
Eurasian jay, *Garrulus glandarius*	Redwing, *Turdus iliacus*
Feral pigeon, *Columba livia*	Ring-necked parakeet, *Psittacula krameri*
Firecrest, *Regulus ignicapilla*	European robin, *Erithacus rubecula*
Goldcrest, *Regulus regulus*	Rook, *Corvus frugilegus*
Goldfinch, *Carduelis carduelis*	Siskin, *Spinus spinus*
Great spotted woodpecker, *Dendrocopos major*	Song thrush, *Turdus philomelos*
Great tit, *Parus major*	Common starling, *Sturnus vulgaris*
Green woodpecker, *Picus viridis*	Stock dove, *Columba oenas*
Greenfinch, *Chloris chloris*	Tawny owl, *Strix aluco*
Grey heron, *Ardea cinerea*	Teal duck, *Anas crecca*
Grey wagtail, *Motacilla cinerea*	Eurasian treecreeper, *Certhia familiaris*
Greylag goose, *Anser anser*	Willow warbler, *Phylloscopus trochilus*

Jackdaw, Corvus monedula	Woodpigeon, *Columba palumbus*
Common kestrel, *Falco tinnunculus*	Eurasian wren, *Troglodytes troglodytes*

Model results:

Moran's I:

Moran's I estimates of linear model residuals were both calculated with "moran.test" and tested with "moran.mc" in *spdep* (Bivand & Wong, 2013). After identification of significant p-values, we used the Lagrange multiplier diagnostic estimates and respective p-values to choose from the three Simultaneous Autoregressive Models (SAR) models in *spatialreg* (Table Vs; Table VIs, Bivand, Hauke, & Kossowski, 2013). We concluded that SAR_{lag} model was the most appropriate except for one formula, which required a SAR_{err} model. Calculating Moran's I on residuals of SAR models confirmed they had removed the spatial autocorrelation appropriately. The nature of the relationships was unaffected by spatial autocorrelation, only the estimate values (which were largely uninformative as most variables were PCA axes) and the levels of significance were slightly modified from linear models to SAR models.

Table Vs: Moran's I estimate of lm residuals for the first set of models (Soundscape metric ~ point count metric). Significant p.values are shaded. Lagrange multiplier diagnostic estimates and respective p.values are given for the models where SAC was detected. The Moran’s I estimates recalculated after fitting the appropriate SAR are presented in the last section of the table.

	Moran's I from residuals of lms			**Lagrange multiplier diagnostics for spatial dependence**				**Moran's I from residuals of SARs**		
Model formula	**Obs**	**Exp**	**p**	**LMerr**	**p**	**LMlag**	**p**	**Obs**	**Exp**	**p**
Acoustic diversity PC1 ~ species richness	0.109	-0.029	0.178							
Acoustic diversity PC1 ~ species composition	0.153	-0.029	0.113							
Acoustic diversity PC2 ~ species richness	0.303	-0.029	0.013	3.675	0.055	2.907	0.015	-0.068	-0.029	0.604
Acoustic diversity PC2 ~ species composition	0.19	-0.029	0.074							
Acoustic composition ~ species richness	-0.08	-0.029	0.629							
Acoustic composition ~ species composition	-0.326	-0.029	0.973	4.127	0.042	2.368	0.124	-0.001	-0.029	0.43

Table VIs: Moran's I estimate of lm residuals for the second set of models (Diversity metric ~ Habitat type). Significant p.values are shaded. Lagrange multiplier diagnostic estimates and respective p.values are given for the models where SAC was detected. The Moran’s I estimates recalculated after fitting the appropriate SAR are presented in the last section of the table.

	Moran's I from residuals of lms			**Lagrange multiplier diagnostics for spatial dependence**				**Moran's I from residuals of SARs**		
Model formula	Obs	Exp	p	LMerr	p	LMlag	p	Obs	Exp	p
Species richness	0.059	-0.029	0.285							
Species composition	0.032	-0.029	0.346							
Acoustic composition	-0.136	-0.029	0.755							
Acoustic diversity PC1	0.102	-0.029	0.195							
Acoustic diversity PC2	0.314	-0.029	0.012	3.82	0.051	6.46	0.011	-0.084	-0.029	0.642

The ARUs at sites 9D6 and 8C1 were faulty and influenced model results drastically, thus were deleted from the dataset.

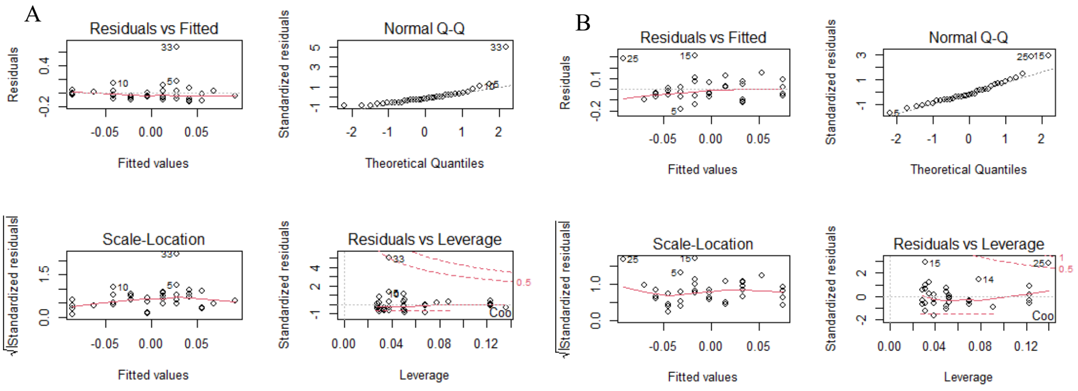


Figure 4s: diagnostic plots of two models showing outliers that were removed from appropriate models: A] Acoustic composition PC1, site 9D6 was an outlier (row 33) so was removed from dataset for the models involving this response variable. B] Acoustic composition PC2, site 8C1 was an outlier (row25) so was removed from dataset for the models involving this response variable.

Table VIIs: Model summary statistics of the first set of models (Soundscape metric ~ point count metric)

Model formula	**Linear models**				**SAR models**		
Model formula	**Fstat**	**Rsquare**	**AdjRsquare**	**p-value**	**z**	**Wald**	**p-value**
Acoustic diversity PC1 ~ species richness	2.098	0.058	0.03	0.158			
Acoustic diversity PC1 ~ species composition	0.198	0.012	-0.048	0.824			
Acoustic diversity PC2 ~ species richness	9.169	0.212	0.189	0.005	2.557	6.54	0.011
Acoustic diversity PC2 ~ species composition	6.498	0.2825	0.239	0.004			
Acoustic composition ~ species richness	8.312	0.2	0.177	0.007			
Acoustic composition ~ species composition	8.732	0.353	0.312	0.001	-2.202	4.8486	0.028

Table VIIIs: Model summary statistics of the second set of models (Diversity metric ~ Habitat type)

	Linear models				**SAR models**		
Model formula	**Fstat**	**Rsquare**	**AdjRsquare**	**p-value**	**z**	**Wald**	**p-value**
Species richness (Poisson distribution)		0.395 (pseudo)					
Species composition	9.789	0.487	0.437	0.000			
Acoustic composition	5.309	0.339	0.276	0.005			
Acoustic diversity PC1	0.286	0.027	-0.067	0.836			
Acoustic diversity PC2	2.176	0.174	0.094	0.111	2.822	7.962	0.005

References for Appendix

Bivand, R. S., Hauke, J., and Kossowski, T. (2013) Computing the Jacobian in Gaussian spatial autoregressivemodels: An illustrated comparison of available methods. *Geographical Analysis*, 45(2), 150-179. https://doi.org/10.1111/gean.12008

Bivand, R. S, & Wong, D. W. S. (2018) Comparing implementations of global and local indicators of spatia lassociation *TEST*, 27(3), 716-748. https://doi.org/10.1007/s11749-018-0599-x

Kasten, E.P., Gage, S.H., Fox, J. & Joo, W. (2012) The remote environmental assessment laboratory’s acoustic library: An archive for studying soundscape ecology. *Ecological Informatics*. 12, 50–67. doi:10.1016/j.ecoinf.2012.08.001.

Pieretti, N., Farina, A., Morri, D. (2011) A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI). *Ecological Indicators*. 11(3), 868-873. https://doi.org/10.1016/j.ecolind.2010.11.005.

Towsey, M., Wimmer, J., Williamson, I. & Roe, P. (2014) The use of acoustic indices to determine avian species richness in audio-recordings of the environment. *Ecological Informatics*. 21, 110–119. doi:10.1016/j.ecoinf.2013.11.007.

Appendix C:

C.1: Supplementary Information for Chapter 4

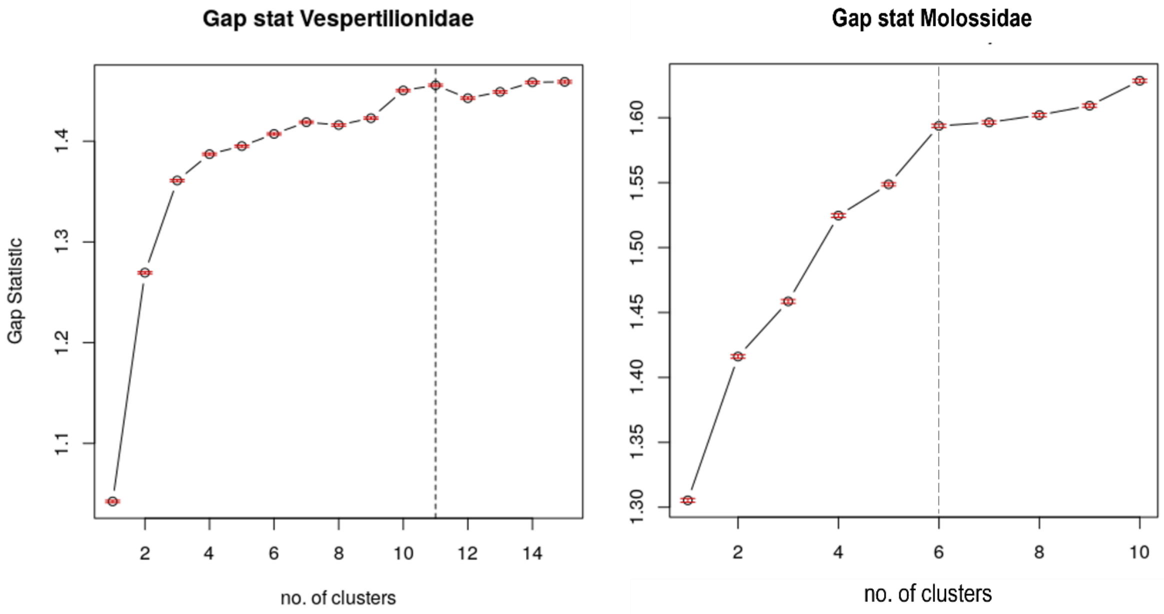


Figure 13: *Examples of gap statistic plots for unknown Vespertilionidae and Molossidae clustering analysis. The gap statistic was plotted for five random subsets of 25 000 calls from each class.*

Table 7: *Mean spectral features for each sonotype. Detection probability is the mean model detection probability. Maximum and minimum frequency are the highest and lowest frequencies of the call. Duration describes the length of the call. Bandwidth is the difference between the maximum and minimum frequencies. Frequency of maximum power is the frequency within the call that is loudest.*

Sonotype	**Detection probability**	**Maximum frequency (Hz)**	**Minimum frequency (Hz)**	**Duration (s)**	**Bandwidth (Hz)**	**Frequency of Maximum power (Hz)**
M1	0.494115	21019.19	13703.24	0.016625	7315.432	15634.31

M2	0.40232	23191.26	17167.92	0.013132	6022.936	18653.57
M3	0.45369	25551.65	19912.4	0.011975	5638.784	21821.17
M4	0.489878	22767.32	16070.55	0.014193	6696.422	17695.85
M5	0.420383	29398.07	12731.47	0.017085	16666.15	15018.39
M6	0.444194	22628.4	16710.96	0.014046	5917.057	18328.53
V1	0.467486	29303.59	21917.03	0.01108	7386.033	24216.84
V2	0.563413	68373.25	44185.95	0.011791	24186.9	47027.31
V3	0.499497	59933.85	48423.39	0.008802	11509.99	50811.88
V4	0.427428	32269.16	24922.18	0.009327	7346.536	27178.98
V5	0.415714	37778.43	29666.39	0.007921	8111.573	31964.6
V6	0.566488	44756.15	31320.28	0.009116	13435.37	34001.56
V7	0.406482	35683.69	27935.95	0.008466	7747.262	29915.07
V8	0.459514	61426.75	51458.76	0.007764	9967.535	53815.77
V9	0.515853	58446.03	48418.52	0.009476	10027.08	50210.74
V10	0.406699	31732.68	24853.34	0.00911	6878.873	27100.93
V11	0.587272	64360.52	47746.73	0.009857	16613.38	50214.05

C.2: Manuscript of upcoming publication

RESEARCH ARTICLE

Deep Joint Detection and Classification of Bat Echolocation Calls

Oisin Mac Aodha[1,2] | Santiago Martínez Balvanera[3] | Elise Damstra[4] | Martyn Cooke[5] | Philip Eichinski[6] | Michel Barataud[7] | Katherine Boughey[8] | Giada Giacomini[8] | M. Cristina MacSwiney G.[9] | Martin K. Obrist[10] | Stuart Parsons[11] | Thomas Sattler[12] | Kate E. Jones[3]

[1]School of Informatics, University of Edinburgh, UK

[2]The Alan Turing Institute, UK

[3]Centre for Biodiversity and Environmental Research, University College London, UK

[4]Department of Life Sciences, Imperial College London, UK

[5]...

[6]School of Electrical Engineering and Computer Science, Queensland University of Technology, Australia

[7]...

[8]Bat Conservation Trust, UK

[9]Centro de Investigaciones Tropicales. Universidad Veracruzana, Xalapa, México

[10]Biodiversity and Conservation Biology, Swiss Federal Research Institute WSL, Switzerland

[11]University of the Sunshine Coast, Australia

[12]Swiss Ornithological Institute, Switzerland

Correspondence
Oisin Mac Aodha
Email: oisin.macaodha@ed.ac.uk

Funding information
...

Abstract

1. Acoustic monitoring is an effective and scalable way to assess the health of important bioindicators like bats in the wild. However, the large amounts of resulting noisy data collected requires accurate tools for automatically determining the presence of different species of interest. Machine learning-based solutions offer the potential to reliably perform this task, but can require expertise in order to train and deploy.

2. We propose, BatDetect, a novel deep learning-based pipeline for jointly detecting and classifying bat species from acoustic data. Distinct from existing deep learning-based acoustic methods, BatDetect's outputs are interpretable as they indicate where in time and frequency a predicted echolocation call occurs. BatDetect also makes use of surrounding temporal information in order to improve its predictions, while still remaining efficient at deployment time.

3. We present experiments on four challenging datasets, from three distinct geographical regions. BatDetect results in a mean average precision of 0.88 for a dataset containing 17 bat species from the UK. This is significantly better than the 0.71 obtained by a traditional call parameter extraction-based baseline.

4. We show that the same pipeline, without any modifications, can be applied to acoustic data from multiple different regions and species groups. The data annotation, model training, and evaluation tools proposed will enable practitioners to easily develop and deploy their own models. BatDetect lowers the barrier to entry currently preventing researchers from availing of effective deep learning powered bat acoustic classifiers.

KEYWORDS
Bioacoustics, Bats, Passive Acoustic Monitoring, Deep Learning, Acoustic Event Detection

1 | INTRODUCTION

Reliable biodicators are necessary to enable us to better measure the impact of climate change and accelerating habitat loss. Bats have previously been identified as one promising candidate for this role due to their global distribution, taxonomic diversity, and sensitivity to environmental and habitat change (Jones et al., 2009). However, despite making up approximately one fifth of all mammalian diversity, we know comparatively less about them in relation to other well studied taxonomic groups (Frick et al., 2020). In order for them to fulfill this potential, there is a growing need for robust and reliable tools for monitoring their populations (Russo et al., 2021).

Recent advances in hardware and software have resulted in low-cost solutions for automated bioacoustic monitoring. This enables us to unobtrusively monitor wild populations at unprecedented spatial and temporal scales via audio (Gibb et al., 2019). In the context of bats, there is a rich history of using acoustic methods for monitoring purposes (Zamora-Gutierrez et al., 2021) by leveraging the fact that bats use sound to navigate and communicate (Jones and Siemers, 2011; Prat et al., 2016). Machine learning-based approaches have been extensively used by extracting acoustic features from audio recordings and then classifying which species are present in the input audio (Parsons and Jones, 2000; Walters et al., 2012; Zamora-Gutierrez et al., 2016; Bas et al., 2017; Roemer et al., 2021). In this line of work, the extracted features are typically manually crafted so that they encode discriminative information related to the temporal and frequency-based characteristics of bat echolocation calls.

However, bat calls are complex and varied. They can exhibit regional, habitat, and species-specific variation which makes them challenging to precisely characterise using hand-designed rules (Walters et al., 2013; Russo et al., 2018). This is in addition to other complicating factors such as background noise and other vocalising species (e.g. small mammals and insects) that can be present in ultrasonic audio recordings. Deep learning-based approaches attempt to address these challenges by learning discriminative representations directly from the raw input data. They have been shown to be highly successful across a wide variety of applications in ecological monitoring (Christin et al., 2019), in addition to bioacoustics (Stowell, 2021).

The first deep learning-based methods applied to bat acoustic monitoring focused on determining the presence of bats versus background noise (Mac Aodha et al., 2018) or the species present (Chen et al., 2020; Kobayashi et al., 2021) from short audio clips, i.e. typically shorter than 50 milliseconds. The disadvantage of these approaches is that they cannot capture longer temporal information such as the interval between individual pulses which can sometimes be an important discriminative signal. To address this issue, other work has used longer input recordings in order to capture multiple individual calls in a sequence (Paumen et al., 2021; Zualkernan et al., 2020; Tabak et al., 2021). Unfortunately the higher dimensionality of the data, due to the longer input audio recording, can necessitate larger models and thus requires more supervised data at training time. Compact and efficient models are necessary in the context of low powered deployments on edge-based monitoring devices (Gallacher et al., 2021; Zualkernan et al., 2021). In addition, there is also an increased chance that more than one species could be present in the longer input recording (Dierckx et al., 2022). This last point is especially problematic as it violates the 'one species per input' assumption of conventional classification approaches.

Despite this recent progress in deep learning-based solutions for bat monitoring, there is still a gap between the latest research advances and the open-source tools available to practitioners. In this work, we attempt to address this gap by proposing a novel pipeline for bat echolocation call detection and species classification from acoustic data. Our approach, called BatDetect, combines the strengths of the short temporal window-based methods with the benefits of the longer-range temporal reasoning of the call sequence based methods. Our main contributions are: (i) An efficient model for joint detection and classification of bat echolocation calls. (ii) This model provides interpretable predictions that illustrate where in the input spectrogram, in terms of frequency and time, the model has detected a call. (iii) We evaluate the effectiveness of our proposed approach on four challenging datasets, collected from three distinct geographical regions, and show that it is superior to existing call parameter-based methods. (iv) We provide open-source tools for our full pipeline in order to enable practitioners to annotate data, train, and deploy models on their own datasets.

2 | MATERIALS AND METHODS

2.1 | Acoustic event detection

Distinct acoustic vocalisation events created by a species of interest (e.g. a bat echolocation call or a bird song) can be characterised by the start time of the event, the duration of the event, and the minimum and maximum frequency bands that the event spans. Our goal is develop a model $g()$ that takes a ultrasonic audio recording as input, represented as a spectrogram $\mathbf{x}$, and outputs a set of predictions related to the events of interest in the input audio file, $O = g(\mathbf{x})$. In our case these events will be bat echolocation calls. Each prediction from the model, $\mathbf{o} \in O$, represents a distinct event and contains information characterising the time and frequency components of the event. Specifically, each predicted event, $\mathbf{o} = [t_{\text{start}}, t_{\text{end}}, f_{\text{min}}, f_{\text{max}}, \mathbf{p}_{\text{species}}]$, represents the start time, end time, minimum frequency, maximum frequency of the event, along with the predicted probability indicating which species the model thinks is present. Here, $\mathbf{p}_{\text{species}}$ is a $C + 1$ dimensional vector that sums to one, and represents the probability of the species the model thinks emitted the call, for each one of C different species plus one additional background class (i.e. 'Not bat'). Note, that this representation is distinct from conventional acoustic classification models that only attempt to determine the species present in a short duration input spectrogram, i.e. $y = g(\mathbf{x})$, where $y \in \{1, ..., C + 1\}$ is an integer denoting the predicted species label.

2.2 | Detection and classification model

We implement our joint classification and detection model $g()$ as a deep neural network. Our model is inspired by computationally efficient one-stage object detection methods from computer vision (Zhou et al., 2019). Unlike two-stage methods that first propose a set of regions of interest and then assign each region to a class (i.e. a species), one-stage approaches directly predict the location and size of each class in the input.

Our model makes use of a U-Net-style architecture (Ronneberger et al., 2015), with an encoder that extracts features from the input spectrogram, followed by a decoder that generates the predicted size and location of each echolocation call along with the corresponding species id. The model also uses skip connections which facilitate the sharing of higher resolution feature information (in terms of frequency and time) from the encoder to the decoder. The output of the decoder is a distribution for over time and frequency indicating where the model thinks a set of calls are present and also the sizes (in terms of frequency range and duration) of the calls. As a final step, we pass this output to a non-maximal suppression layer, implemented via max pooling, in order to extract the local peak detections (Zhou et al., 2019). This step prevents the model from predicting multiple calls very close to each other (i.e. within a few milliseconds). A high-level depiction of the model is illustrated in Figure 1.

A common issue with many current deep learning-based bat call detection and classification models, e.g. (Mac Aodha et al., 2018; Chen et al., 2020; Kobayashi et al., 2021), is that they typically only utilise very short temporal input windows (e.g. less than 50 milliseconds) to determine if a species is present. This prevents these models from reasoning about inter-pulse temporal information that can exist between individual calls and can span hundreds or thousands of milliseconds. This issue could be partially addressed by using more computationally expensive backbone encoder models that have a larger temporal receptive field size, e.g. (Simonyan and Zisserman, 2015; He et al., 2016). However, the downside of such models is that they are much larger, and thus have more parameters that need to be trained. This larger size necessitates larger supervised training datasets and results in a reduction in speed at inference time. To overcome this problem, without having to increase the size and capacity of the encoder, we introduce a self-attention layer into the middle of our network. Transformer-based self-attention architectures (Vaswani et al., 2017) are among the current most performant models in natural language processing owning to their ability to capture long-range dependencies that occur in the input data. The introduction of this layer allows our model to ‘attend’ to information from different points in time in the input audio file in order to increase or decrease its estimated likelihood that a given species is present at the current time step. Note that this self-attention layer only operates along the temporal dimension and is thus very computationally efficient.

Our entire model is trained end-to-end using a three component loss function which includes a detection loss, a classification loss, and an event size loss. The first two losses are implemented using a focal loss (Lin et al., 2017), and the final one uses an L1 penalty. The model and associated training and evaluation code are implemented using the `PyTorch` deep learning framework (Paszke et al., 2019). A detailed description of the audio pre-processing steps,

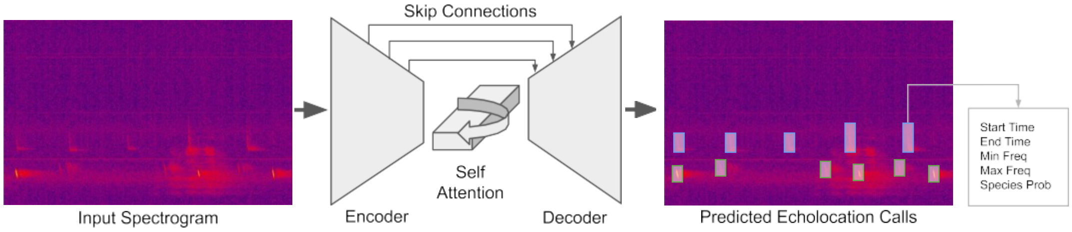


FIGURE 1 Overview of BatDetect, our echolocation call classification and detection model. In contrast to most existing deep learning-based bat call classifiers, our model directly predicts the time in file of each event of interest, along with the duration of the event, the frequency range, and the species. It utilises a self-attention layer in the middle of the model so that it can reason over a longer temporal scale.

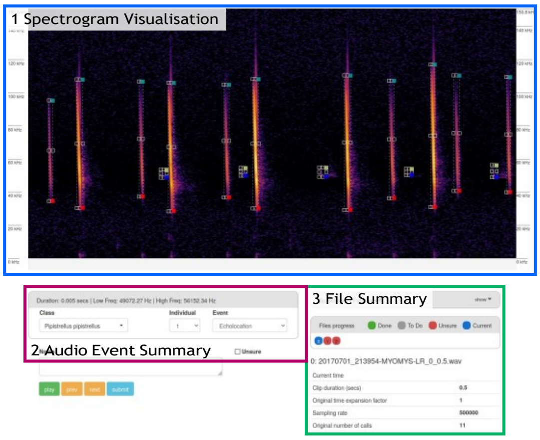
model architecture, training losses, and training settings are provided in the supporting information.

2.3 | Audio annotation interface

Our model requires supervision in the form of bounding boxes encompassing each individual echolocation call present in an audio file. In order to obtain this, we developed an audio annotation interface to enable human annotators to efficiently draw boxes and to assign a species class label to every audible echolocation call in a given input file. The interface is implement using the `Flask` web framework (Flask, 2021) and is is depicted in Figure 2. As a result, we can deploy the interface on the web to allow annotators to annotate remotely or we can also deploy it locally on an annotator's own device.

The interface has been optimised to speed up the annotation process. For example, we pre-cache the spectrogram generation step for the next file to be annotated so that the annotator does not have to wait when switching between files. In addition, it is possible to change the spectrogram visualisation settings in order to trade-off frequency resolution for temporal resolution, or vice versa. The annotations are stored together in a separate JSON file for each audio file using a format similar to the one used for the COCO dataset (Lin et al., 2014). Annotators can playback the audio file using a time expansion factor of ten to ensure that the ultrasonic signals of interest are audible.

Unless otherwise specified, the audio files that we annotated had information at the file-level related to which species were present in the recording. Annotators were instructed to draw boxes around each individual echolocation call, irrespective of how faint the call was. They then assigned the recording-level species class label to an annotation unless it differed from a prototypical echolocation call for that species. Harmonics were not annotated as part of the main call. In cases where it was not possible to assign the correct class label, or when multiple species were present in a file, annotators marked unknown calls as being from the generic 'Bat' class.



Duration: 0.005 secs | Low Freq: 49072.27 Hz | High Freq: 56152.34 Hz

Class: Pipistrellus pipistrellus

Individual: 1

Event: Echolocation

☐ Unsure

Files progress: ● Done ● To Do ● Unsure ● Current

0: 20170701_213954-MYOMYS-LR_0_0.5.wav

Current time	
Clip duration (secs)	0.5
Original time expansion factor	1
Sampling rate	500000
Original number of calls	11

FIGURE 2 Our audio annotation interface has three main components: (i) spectrogram visualisation and playback, (ii) editing of echolocation call annotations, and (iii) file-level metadata display.

2.4 | Audio datasets

We train and evaluate our model on three different full spectrum ultrasonic acoustic datasets. Additional details for each, including visual examples and per-species counts, are available in the supporting information.

2.4.1 | UK data

This dataset contains audio data from 17 bat species that breed in the UK, and has been collated from six different sources. In total there are 2,809 distinct audio files, with an average duration of 1.04 seconds, and the dataset contains a total of 34,635 annotated echolocation calls. To increase our robustness to background noise, we supplement this data with 4,225 additional, 0.384 second duration, files from the iBats Program (Jones et al., 2013). This adds an additional 6,842 annotated bat calls that do not have an associated species id. Finally, we also add 345, one second duration, empty files (i.e. no bats present) from London, UK, collected using the recording devices from Gallacher et al. (2021).

We split the UK data into two sets, UK_{same} and UK_{diff}. For UK_{same} we randomly assign files to the test set by ensuring a maximum of four files per species, per data source. The remaining files are kept for the training set. This results in 7,010 training files and 369 test files, containing 36,955 and 4,522 calls respectively. UK_{diff} is a more challenging split. Here we hold-out the largest single data source for testing. This leaves 5,991 training and 1,468 test files, containing 24,315 and 17,162 echolocation calls. This second split represents a more challenging test-case where the data is guaranteed to be very different from the training set. This also results in a reduction in the overall amount of training data, both in terms of sheer quantity but also diversity. Both variants of the dataset retain the 4,570 files without species labels as part of their respective training sets.

2.4.2 | Yucatan data

The second dataset consists of 1,193 one second audio clips extracted from 285 passive acoustic recordings from the Yucatan peninsula in Mexico. The data was collected as part of a study by MacSwiney G et al. (2008). It is smaller in size than the UK dataset, but is representative of the type of data that would be feasible to collect and annotate as part of a smaller-scale monitoring project. The annotations from the original study were used and then expanded to ensure that all audible echolocation events were annotated. The final annotated dataset contains 9,981 echolocation calls from 17 different species. We divided the data into 911 training and 282 test clips, making sure to separate at the original recording-level, and not the clip-level, to ensure that clips from the same recording were not in both sets.

2.4.3 | Brazil data

Our final dataset represents an orthogonal challenge to the first two. It contains 320, ten second duration, recordings collected between January and March 2019 in south-eastern Brazil using AudioMoth (Hill et al., 2018) recorders. Here we have access to the recordings but do not have any species id metadata. As a result, instead of annotating the calls with species labels, we instead created 'meta-categories' based on the dominant frequency component exhibited by each call. This resulted in three distinct call groups in the final annotated dataset. Like the other datasets, this annotation was performed manually, where the protocol again stipulated that all echolocation call instances in each recording should be annotated. We split the data into 256 train files and 64 test files, which resulted in 7,989 and 2,010 calls respectively.

2.5 | Baseline model

In order to evaluate the effectiveness of our model, we compare it to a traditional bat call parameter/feature extraction pipeline. To do this, we use the Tadarida-D model from Bas et al. (2017) which consists of two main components: (i) a bat echolocation call detector and (ii) a echolocation call feature extractor. The extracted call features are a set of numerical values that encode information about the shape and frequency content of each individual detected bat call. In the case of Tadarida-D, this amounts to 268 features for each detected event. For additional details about the specific set of call features in Tadarida-D, please consult the original paper (Bas et al., 2017).

For each of our datasets, we first run Tadarida-D to detect the calls and extract the call features. Then for each detected event in the training set we compute the overlap between the event (using the reported time in file, duration, and frequency range from Tadarida-D) and our ground truth annotations. We select the detection that overlaps most in time and frequency with a given ground truth annotation and then assign the species label from the ground truth to that event. If a detected event does not match to a ground truth annotation it is assigned to the 'Not bat' class. Each ground truth annotation can only be assigned to one predicted detection. Finally, we train a Random Forest (Breiman, 2001) classifier on the extracted calls using the implementation from `scikit-learn` (Pedregosa et al., 2011), using the default parameters. It is worth emphasising that while we are using Tadarida-D, our baseline is not directly equivalent to the full Tadarida method as we do not make use of their pre-trained models, labeling interface, or classification code. However, this baseline allows us to control for the impact of the training data as we can ensure that we are using the same audio and ground truth annotations at training and test time for both methods. This baseline also does not make use of the additional echolocation events that only have the generic 'Bat' class label. However, this is only relevant for the UK datasets.

2.6 | Evaluation metrics

We use four different evaluation metrics to quantify the performance of our model. The first, detection average precision ('AP Det'), evaluates the ability of the model to correctly identify all valid echolocation calls in the test data. This metric calculates the precision and recall resulting from varying a threshold on the model output predictions for the 'Bat' versus 'Not bat' task. We then average over these different thresholds to quantify the area under the precision-recall curve, using the interpolation method from Everingham et al. (2010). A prediction is counted as a true positive if its estimated start time overlaps with a ground truth echolocation call by at most ten milliseconds. This is the same evaluation criteria used in Mac Aodha et al. (2018).

'AP Det' does not evaluate the ability of the model to accurately assign the correct species label to a prediction. To address this, we also report the mean average precision across the classes ('mAP Class'). This involves taking the per-class average precision and then averaging this value over each class. This also has the added effect of weighting each class equally, irrespective of the number of calls for each class in the test set. Here, we exclude calls for which there are no ground truth species labels available.

'mAP Class' suffers from one major limitation. As the classes are evaluated independently, it does not highlight cases where the underlying model may be poorly calibrated and thus require different output thresholds for each class. Calibration issues like this can result from class-level data imbalances in the training data. To overcome this limitation, we also report a third precision based metric which we refer to as 'Top Class'. Here we simply take the top predicted class label, along with its corresponding probability, for each detected call and then evaluate the average precision as above. Unlike 'mAP Class', this metric can be biased if there is a large imbalance in the classes in the test set.

The final metric, 'File Acc', evaluates the file-level classification accuracy. For this metric only, we exclude test files that have been manually annotated as containing more than one species. In order to convert the multiple possible individual call predictions for a given file into a single file-level class label, we threshold each of the individual

TABLE 1 Performance of our BatDetect model compared to the Random Forest baseline with uses traditional bat echolocation call features. We evaluate both models using the same four test datasets, and BatDetect performs best in all cases. For each of the metrics, higher numbers are better, and the results are averaged over three runs.

	BatDetect (Ours)				Random Forest Baseline			
Dataset	AP Det	mAP Class	Top Class	File Acc	AP Det	mAP Class	Top Class	File Acc
UK_{same}	0.971	0.884	0.843	0.866	0.890	0.706	0.638	0.800
UK_{diff}	0.964	0.810	0.690	0.780	0.903	0.587	0.47	0.687
Yucatan	0.927	0.769	0.809	0.805	0.651	0.409	0.454	0.603
Brazil	0.926	0.962	0.940	1.000	0.883	0.912	0.910	1.000

detections and remove any detection below the threshold. We then sum the per-class probabilities of the remaining detections and choose the class with the highest sum as the file-level prediction. Finally, we report the file-level accuracy corresponding to the single best threshold across all files. The best possible score for each of these four metrics is 1.0, and the worst is 0.0.

3 | RESULTS

3.1 | Detection and classification performance

In Table 1 we present the main results comparing the performance of our model, BatDetect, to the Random Forest baseline that uses Tadarida-D call features. The results represent the average of three different models, each trained with different random initialisation. We observe that across all datasets, and the four evaluation metrics, BatDetect performs best. The Random Forest baseline also performs well on the comparatively easy Brazil dataset, but struggles on the other three. The difference in performance is between 0.05 and 0.36 mean average precision (‘mAP Class’), across the datasets.

We can see that BatDetect’s detection performance, reported via ‘AP Det’, is strong. This indicates that the model is capable of correctly detecting the vast majority of calls. However the lower performance for the two call-level classification metrics (‘mAP Class’ and ‘Top Class’) indicates that it can have difficulty identifying the correct species for a given call in some situations. Table 1 also highlights the challenge posed by the more difficult UK_{diff} dataset in contrast to the performance on UK_{same}. In the supporting information, we illustrate the impact that the amount of training data per-class has on test performance, and broadly observe that more data increases performance.

In Figure 3 we display the per-class precision-recall curves for BatDetect. We also show precision-recall curves at the genus-level. For these genus results, we do not retrain the models, but instead sum the predictions for each species belonging to a given genus, convert the ground truth class label to the genus label, and then evaluate in the same way as the species-level curves. We also display the file-level confusion matrix. Like the 'File Acc' metric, we only report results for files that have one reported species in them. By comparing the genus-level results in the second column for the challenging *Myotis* calls to the corresponding species-level ones in the first column for both UK datasets, we see that the model is capable of resolving the classification task to the genus-level for these calls, but has difficulty for some at the species level. This difficulty is most apparent when looking at the confusion matrix for UK_{diff} in the second row of Figure 3. Here we see that our model confuses some *Myotis* species at the file-level.

We visualise the model's predictions for a subset of files in Figure 4. We observe that BatDetect is capable of detecting faint calls, and also handles situations where multiple species are present in a recording. The model is also robust to background noise. This is most apparent in the example from the Brazil dataset recorded using an AudioMoth (Hill et al., 2018) on the bottom row of the figure. In this example we can see a repetitive high frequency signal, most prominent at ∼60kHz that repeats every 50 milliseconds. Despite this structured noise, our model does not produce any false positives in this example.

It takes BatDetect just under four minutes to process and save the results for 424, ten second duration, 384kHz AudioMoth recordings using a GPU, i.e. 70.6 minutes of ultrasonic data in total. Tadarida-D takes 2.5 minutes for detection and feature extraction for the same data. Note that this processing time does not include the evaluation of the Random Forest and also Tadarida-D does not utilise a GPU. This benchmarking was performed on a desktop computer which contained an Intel i7-6850K CPU and an Nvidia TITAN Xp GPU.

3.2 | Impact of self-attention

In Table 2 we present results a variant of the model on the $\mathrm{UK}_{\mathrm{diff}}$ dataset. Here we report the results for when we remove the self-attention layer, i.e. 'No Self-Attn', again averaged over three different runs. We observe a large drop in performance when compared to the full model. Notably, the detector results illustrated by 'AP Det' are not impacted, but two of the classification metrics, 'mAP Class' and 'Top Class', show a large decrease when removing this layer. This points to the value of longer temporal range reasoning when resolving species classification that is provided by the self-attention layer. In the supporting information, we provide a visualisation of how the self-attention layer makes use of information from different points in time in order to improve its species-level predictions.

TABLE 2 Performance of two different variants BatDetect on the $\mathrm{UK}_{\mathrm{diff}}$ test set. Note, the results in the first row are the same as the BatDetect results in the second row of Table 1. 'No Self-Attn' is the same as the full BatDetect model but the self-attention layer has been removed at training and test time.

Dataset	AP Det	mAP Class	Top Class	File Acc
Full model	0.964	0.810	0.690	0.780
No Self-Attn	0.962	0.725	0.614	0.790

4 | DISCUSSION

4.1 | Model performance

BatDetect performs significantly better than the traditional call feature-based baseline tested. For the vast majority of species in the $\mathrm{UK}_{\mathrm{same}}$ dataset, BatDetect results in high precision at high recall rates (see Figure 3). This is important as it enables practitioners to trade-off recall for precision to ensure that they obtain reliable, high confidence, predictions from the model. The file-level accuracy is 78% and 86.6% for the $\mathrm{UK}_{\mathrm{diff}}$ and $\mathrm{UK}_{\mathrm{same}}$ datasets, where a large percentage of the mistakes can be attributed to known challenging species, i.e. the *Myotis* species. While we observe a performance drop for $\mathrm{UK}_{\mathrm{diff}}$, the $\mathrm{UK}_{\mathrm{same}}$ results indicate that training on larger quantities of more representative data results in a more effective model.

Unlike existing deep learning-based classifiers, our model produces interpretable predictions in the form of time and frequency boxes around the detected calls (see Figure 4). This is valuable as it will enable practitioners to inspect the model predictions to better understand any failure cases they may observe for their datasets. BatDetect can efficiently use information from longer input time scales via the self-attention layer without significantly increasing the amount of computation performed at test time. This results in a model that can perform inference ∼17 times faster than real time using a GPU, i.e. 17 minutes of recorded ultrasonic audio takes one minute to fully process.

Perhaps most importantly, we showed that the same pipeline, without any modifications, can be applied to audio data from three distinct regions. This is valuable as it will allow practitioners to focus on collecting and annotating datasets for their species of interest. Our annotation interface assists this process and will enable researchers to make annotations available to others in a standardised and open format.

4.2 | Limitations

BatDetect performs well across the four datasets tested, however it still suffers from some limitations. We rely on the availability of diverse, and exhaustively annotated, training data. Collecting such data can be challenging, in addition to being time consuming to annotate. This limitation is common to any supervised learning-based method. While methods for semi-supervised and self-supervised training offer the potential to learn effective models with limited to no training supervision, diverse labelled data is still needed to evaluate the performance of the developed models. Bat calls can exhibit plasticity depending on the population sampled (Montauban et al., 2021). As a result, care needs to be taken to ensure that the collected training datasets are representative of the downstream deployment situation. Finally, our training datasets currently only contain annotated echolocation calls, and thus the model cannot make predictions for other types of calls, e.g. social calls or feeding buzzes. With appropriate training data, this could be addressed.

For a given input recording, BatDetect returns a list of detections along with their corresponding time in file and species id. It is left up to the user to decide how to best merge the individual detections into a set of 'bat passes', where a pass constitutes a sequence of individual calls. This summary step can be important, as practitioners often derive to use a grouping-based heuristic based on the time between detected calls as in Mac Aodha et al. (2018). The high recall rates of BatDetect means that this type approach is less likely to separate individual bat passes into multiple different ones. In contrast, methods that produce high numbers of false negatives run the risk of overcounting the number of passes as they can miss faint calls in a sequence, and thus incorrectly break them up into a number of shorter passes.

5 | CONCLUSIONS

We presented BatDetect, a general-purpose model for detecting and classifying bat echolocation calls in challenging high-frequency audio data. We showed that the same model, without modifications, can be trained and evaluated successfully on data from different geographical regions. In addition to pretrained models, we also make data and code for our models and annotation interface available to stimulate future research.

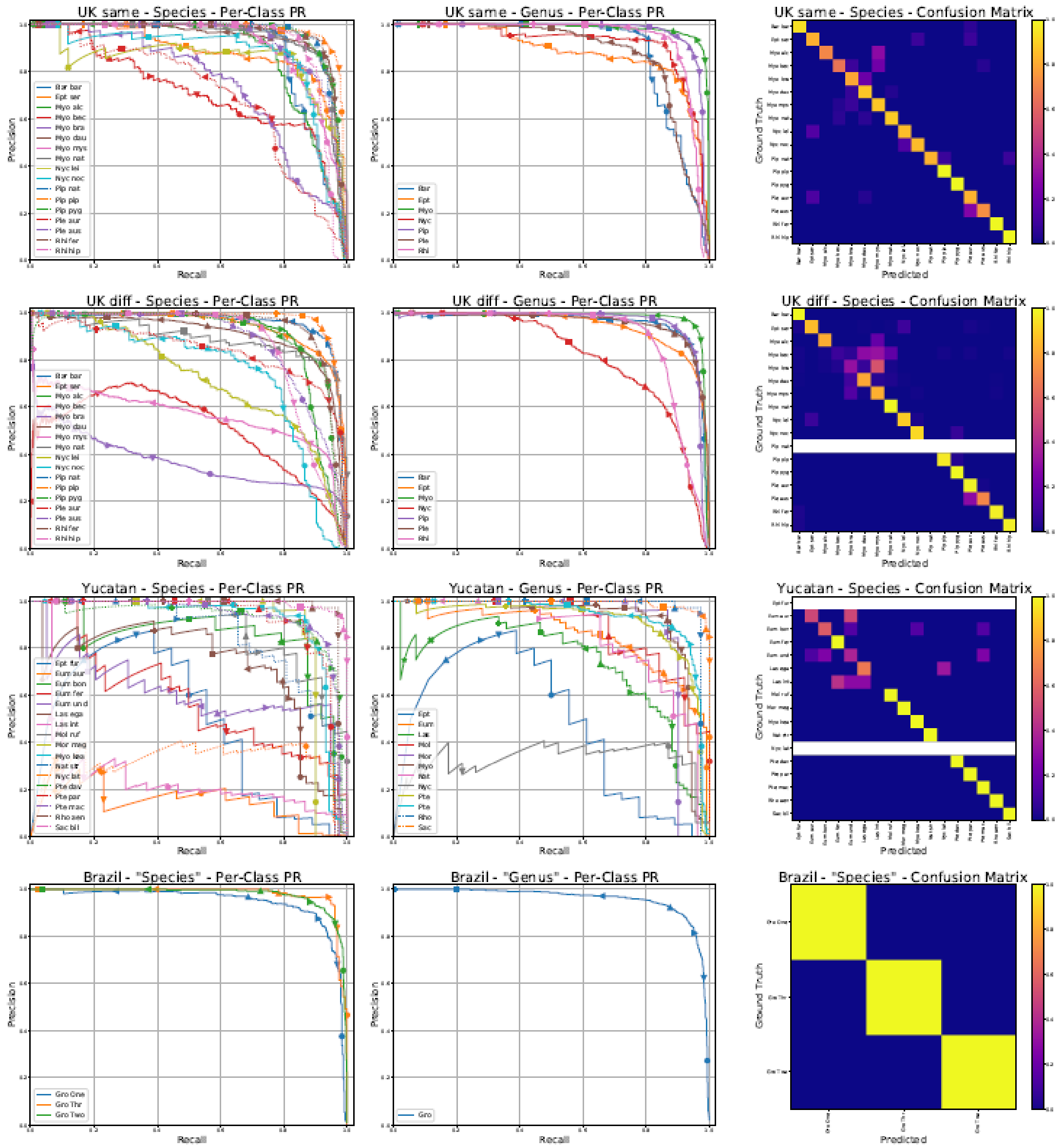


FIGURE 3 Precision-recall (PR) and confusion matrices for our BatDetect model for the four different test sets. The first column depicts the per-species precision-recall curves and the second column is the per-genus equivalent. The third column illustrates the file-level confusion matrix, where white rows indicate that there were no species of that type in the filtered test set. Each row depicts a different dataset.

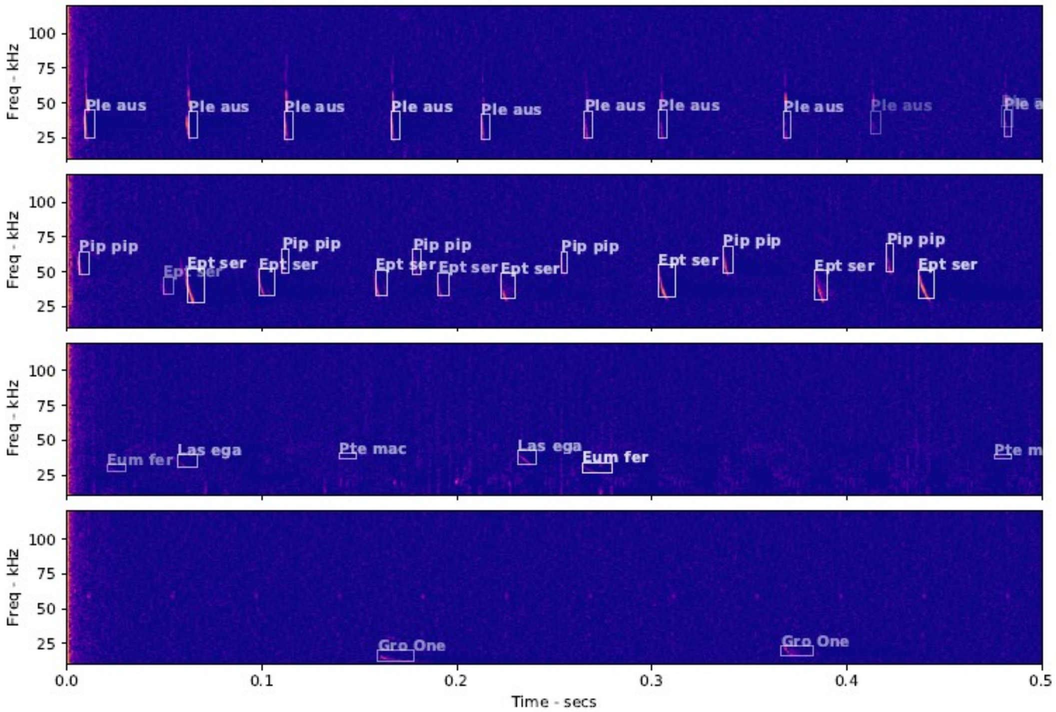


FIGURE 4 Predictions from our BatDetect model. Each row represents a different audio file selected from the test sets of the UK_{same}, UK_{diff}, Yucatan, and Brazil datasets, ordered from top to bottom. The intensity of an individual predicted bounding box indicates the model's confidence, with a brighter white value indicating more confident. The text above each box corresponds to the highest probability class label.

References

Bas, Y., Bas, D. and Julien, J.-F. (2017) Tadarida: a toolbox for animal detection on acoustic recordings. *Journal of open research software*.

Breiman, L. (2001) Random forests. *Machine learning*, **45**, 5-32.

Chen, X., Zhao, J., Chen, Y.-h., Zhou, W. and Hughes, A. C. (2020) Automatic standardized processing and identification of tropical bat calls using deep learning approaches. *Biological Conservation*, **241**, 108269.

Christin, S., Hervet, É. and Lecomte, N. (2019) Applications for deep learning in ecology. *Methods in Ecology and Evolution*, **10**, 1632-1644.

Dierckx, L., Beauvois, M. and Nijssen, S. (2022) Detection and multi-label classification of bats. In *International Symposium on Intelligent Data Analysis*, 53-65.

Everingham, M., Van Gool, L., Williams, C. K., Winn, J. and Zisserman, A. (2010) The pascal visual object classes (voc) challenge. *International journal of computer vision*.

Flask (2021) Python web framework. https://flask.palletsprojects.com.

Frick, W. F., Kingston, T. and Flanders, J. (2020) A review of the major threats and challenges to global bat conservation. *Annals of the New York Academy of Sciences*, **1469**, 5-25.

Gallacher, S., Wilson, D., Fairbrass, A., Turmukhambetov, D., Mac Aodha, O., Kreitmayer, S., Firman, M., Brostow, G. and Jones, K. (2021) Shazam for bats: Internet of things for continuous real-time biodiversity monitoring. *IET Smart Cities*.

Gibb, R., Browning, E., Glover-Kapfer, P. and Jones, K. E. (2019) Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in Ecology and Evolution*, **10**, 169-185.

He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep residual learning for image recognition. In *Conference on computer vision and pattern recognition*.

Hill, A. P., Prince, P., Piña Covarrubias, E., Doncaster, C. P., Snaddon, J. L. and Rogers, A. (2018) Audiomoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. *Methods in Ecology and Evolution*, **9**, 1199-1211.

Jones, G., Jacobs, D. S., Kunz, T. H., Willig, M. R. and Racey, P. A. (2009) Carpe noctem: the importance of bats as bioindicators. *Endangered species research*, **8**, 93-115.

Jones, G. and Siemers, B. M. (2011) The communicative potential of bat echolocation pulses. *Journal of Comparative Physiology A*, **197**, 447-457.

Jones, K. E., Russ, J. A., Bashta, A.-T., Bilhari, Z., Catto, C., Csősz, I., Gorbachev, A., Győrfi, P., Hughes, A., Ivashkiv, I., Koryagina, N., Kurali, A., Langton, S., Collen, A., Margiean, G., Pandourski, I., Parsons, S., Prokofev, I., Szodoray-Paradi, A., Szodoray-Paradi, F., Tilova, E., Walters, C. L., Weatherill, A. and Zavarzin, O. (2013) *Indicator Bats Program: A System for the Global Acoustic Monitoring of Bats*, chap. 10, 211-247.

Kobayashi, K., Masuda, K., Haga, C., Matsui, T., Fukui, D. and Machimura, T. (2021) Development of a species identification system of japanese bats from echolocation calls using convolutional neural networks. *Ecological Informatics*, **62**, 101253.

Lin, T.-Y., Goyal, P., Girshick, R., He, K. and Dollár, P. (2017) Focal loss for dense object detection. In *International conference on computer vision*.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P. and Zitnick, C. L. (2014) Microsoft coco: Common objects in context. In *European conference on computer vision*.

Mac Aodha, O., Gibb, R., Barlow, K. E., Browning, E., Firman, M., Freeman, R., Harder, B., Kinsey, L., Mead, G. R., Newson, S. E., Pandourski, I., Parsons, S., Russ, J., Szodoray-Paradi, A., Szodoray-Paradi, F., Tilova, E., Girolami, M., Brostow, G. and Jones, K. E. (2018) Bat detective–deep learning tools for bat acoustic signal detection. *PLOS Computational Biology*, **14**, 1-19.

MacSwiney G, M. C., Clarke, F. M. and Racey, P. A. (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.

Montauban, C., Mas, M., Tuneu-Corral, C., Wangensteen, O. S., Budinski, I., Martí-Carreras, J., Flaquer, C., Puig-Montserrat, X. and López-Baucells, A. (2021) Bat echolocation plasticity in allopatry: a call for caution in acoustic identification of pipistrellus sp. *Behavioral Ecology and Sociobiology*, **75**, 1-15.

Parsons, S. and Jones, G. (2000) Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks. *Journal of experimental biology*, **203**, 2641-2656.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L. et al. (2019) Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, **32**, 8026-8037.

Paumen, Y., Mälzer, M., Alipek, S., Moll, J., Lüdtke, B. and Schauer-Weisshahn, H. (2021) Development and test of a bat calls detection and classification method based on convolutional neural networks. *Bioacoustics*, 1-12.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E. (2011) Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825-2830.

Prat, Y., Taub, M. and Yovel, Y. (2016) Everyday bat vocalizations contain information about emitter, addressee, context, and behavior. *Scientific Reports*, **6**, 1-10.

Roemer, C., Julien, J.-F. and Bas, Y. (2021) An automatic classifier of bat sonotypes around the world. *Methods in Ecology and Evolution*, **12**, 2432-2444.

Ronneberger, O., Fischer, P. and Brox, T. (2015) U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, 234-241.

Russo, D., Ancillotto, L. and Jones, G. (2018) Bats are still not birds in the digital era: echolocation call variation and why it matters for bat species identification. *Canadian Journal of Zoology*, **96**, 63-78.

Russo, D., Salinas-Ramos, V. B., Cistrone, L., Smeraldo, S., Bosso, L. and Ancillotto, L. (2021) Do we need to use bats as bioindicators? *Biology*, **10**, 693.

Simonyan, K. and Zisserman, A. (2015) Very deep convolutional networks for large-scale image recognition. In *International Conference on Learning Representations*.

Stowell, D. (2021) Computational bioacoustics with deep learning: a review and roadmap. *arXiv preprint arXiv:2112.06725*.

Tabak, M. A., Murray, K. L., Reed, A. M., Lombardi, J. A. and Bay, K. J. (2021) Automated classification of bat echolocation call recordings with artificial intelligence. *Ecological Informatics*, 101526.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. and Polosukhin, I. (2017) Attention is all you need. In *Advances in neural information processing systems*.

Walters, C. L., Collen, A., Lucas, T., Mroz, K., Sayer, C. A. and Jones, K. E. (2013) Challenges of using bioacoustics to globally monitor bats. In *Bat evolution, ecology, and conservation*, 479-499.

Walters, C. L., Freeman, R., Collen, A., Dietz, C., Brock Fenton, M., Jones, G., Obrist, M. K., Puechmaille, S. J., Sattler, T., Siemers, B. M. et al. (2012) A continental-scale tool for acoustic identification of european bats. *Journal of Applied Ecology*, **49**, 1064-1074.

Zamora-Gutierrez, V., Balvanera, S. M., Esquivelzeta, E. R. et al. (2021) The evolution of acoustic methods for the study of bats. In *50 Years of Bat Research*, 43-59.

Zamora-Gutierrez, V., Lopez-Gonzalez, C., MacSwiney Gonzalez, M. C., Fenton, B., Jones, G., Kalko, E. K., Puechmaille, S. J., Stathopoulos, V. and Jones, K. E. (2016) Acoustic identification of mexican bats based on taxonomic and ecological constraints on call design. *Methods in Ecology and Evolution*, **7**, 1082-1091.

Zhou, X., Wang, D. and Krähenbühl, P. (2019) Objects as points. *arXiv preprint arXiv:1904.07850*.

Zualkernan, I., Judas, J., Mahbub, T., Bhagwagar, A. and Chand, P. (2020) A tiny cnn architecture for identifying bat species from echolocation calls. In *International Conference on Artificial Intelligence for Good*, 81-86.

– (2021) An aiot system for bat species classification. In *International Conference on Internet of Things and Intelligence System*, 155-160.

Appendix D:

D.1 Supplementary Information for Chapter 5

Table D.1: Table summarising weather station altitude and location from Climate Edge and NOAA

Data provider	Source	Latitude	Longitude
Climate Edge	Farm altitude 504m	13.49032	-85.593
Climate Edge	Farm altitude 520m	13.05751	-85.7331
Climate Edge	Farm altitude 602m	13.23013	-85.6217
Climate Edge	Farm altitude 693m	13.0952	-86.0675
Climate Edge	Farm altitude 958m	12.96643	-85.8761
Climate Edge	Farm altitude 996m	12.78955	-85.9513
Climate Edge	Farm altitude 1036m	13.17411	-85.8832
Climate Edge	Farm altitude 1045m	13.21876	-85.9809
Climate Edge	Farm altitude 1094m	13.19047	-85.8119
Climate Edge	Farm altitude 1101m	12.82815	-85.9341
NOAA	Jinotega altitude 985m	13.08	-85.98