Kent Academic Repository

Yoh, Natalie (2022) *The use of acoustics for monitoring tropical bats in Southeast Asia.* Doctor of Philosophy (PhD) thesis, University of Kent,.

The use of acoustics for monitoring tropical bats in Southeast Asia

Natalie Yoh

Durrell Institute of Conservation and Ecology
School of Anthropology and Conservation
University of Kent, Canterbury

Thesis submitted for the degree of Doctor of Philosophy in Biodiversity Management
April 2022
Word count: 39,865

Rhinolophus trifoliatus

The use of acoustics for monitoring tropical bats in Southeast Asia

Natalie Yoh

Supervised by:

Dr Matthew J. Struebig

Dr Jake E. Bicknell

Covid-19 Impact Statement

N. Yoh kindly requests her examiners to note how the COVID-19 pandemic impacted this thesis.

During the lockdown periods (March–June and October–December 2020) N. Yoh was confined to work from home without a working environment conducive to focused work. The COVID-19 pandemic also disrupted travel plans commencing after March 2020. This included fieldwork, training events, and conference attendance. Originally the PhD included two field seasons in Borneo, with the second taking place in Kalimantan in June/July 2020. This was not possible as travel restrictions remained in place until 2022. There had been additional training opportunities overseas planned, including potential internships with external collaborators in USA and Spain. These trips were also not possible due to travel restrictions. The biggest impacts were on **Chapter 3** which required reference call data and advanced statistical approaches. Data were subsequently collected through a network of online collaborations and these skills were self-taught through online resources.

She would like to take this opportunity to thank her supervisors who made working through this disruptive period much more manageable and who took extra steps to maintain the lab community online.

Author signature: Natalie Yoh

Date: 12th April 2022

Main supervisor signature:

Date: 19th April 2022

Acknowledgements

I am immensely grateful and indebted to the endless number of people who have supported me throughout my PhD without whom this thesis would not have been possible. First and foremost, I would like to thank my supervisory team Matthew Struebig, Jake Bicknell, and (unofficially) Tigga Kingston for their patience, support, and encouragement. I would also like to thank the Kent administrative staff for all the work they have done behind the scenes to support my PhD.

I would like to thank everyone on our field expedition to Crocker. To the other members of the bat team, Katie Fitzgerald, Alex Troutman, Su, and Kuja, for having the opportunity to learn about Borneo bats with them and for the endless hours spent restringing harp traps, running from night wasps, and sharing mountains of PJ sandwiches. To Theo Sumnicht for spending his nights off driving me back and forth from my survey sites and to the rest of the ant team, Robin Verble, Lilly Germeroth, and Siti, who gave me a whole new appreciation for these species.

I would also like to thank all my co-authors who shared their data and experience with me, often from afar during the pandemic. All of them contributed in one way or another, both to the work in this thesis and to my development as a researcher. Without them, I would know a great deal less about Southeast Asian bats.

So much of my academic journey has been supported by the endless patience and guidance of Adrià López-Baucells and Ricardo Rocha. Without their technical support, professional advice, the opportunities they've given me, and of course their friendship, I have no doubt I would not be in this position without them and I will forever be grateful. I would also like to thank the bat teams at the Natural History Museum of Granollers and the Norwegian University of Life Sciences for always making me feel welcome in their research communities.

There are far too many people that deserve thanks from the DICE staff and student community. However, I'd particularly like to acknowledge Dave Seaman, Anna Jemmett, and Sophus Zu Ermgassen for their constant support, humour, and the never-ending coffee. I'm so glad I got to share and navigate the last few years with them. Although I cannot name everyone, I'd additionally like to give a special mention to Dee Rundle, Jess Fisher, Addy Lowe, Claire Stewart, Keira Pratt-Boyden, Helen Pheasey, Maria Voigt, and Michaela Lo, as well as the Women in Canterbury Conservation Network and our Defaunation lab group.

Outside academia, I’d like to thank my close friends and family who may not always understand what I am doing but are there for me regardless. To my two brothers, Connor and Toby, and to Suzy and Alex, and to my parents for laughter and commiserations throughout all the ups and downs. In particular, I’d like to thank Anna Jenkins for her unwavering support no matter what.

To April, for the endless hours chewing over R scripts, sonograms, mist-nets, and long-distance vegetable gardens. For letting me waffle on about naming conventions and algorithms. For my first Northern bat. I will be forever grateful we got to do this part of the PhD journey together.

Lastly, I’d like to formally thank the people and services who encouraged and supported me when it was easy to not. To those at CAMHS Burnley, Early Break, Early Invention, and the Colne and West Craven Police, and finally, Becky Wyatt to whom I am eternally grateful.

Author declaration

N. Yoh wrote all of the chapters, with editorial suggestions made by PhD supervisors J.E. Bicknell and M.J Struebig. Chapters 2-5 include collaborations with researchers outside of the supervisory team, as outlined below.

Chapter 2 originated from discussions between T. Kingston, M. Struebig, A. Mahyudin, I. Azhar, K.V. Fitzgerald, and N. Yoh. The methodology was developed by all the co-authors listed above. N. Yoh led the design for the acoustic component of the project and data collection. Data analysis was performed by N. Yoh and R. Yu. N. Yoh, T. Kingston, K.V. Fitzgerald, and R. Yu wrote the chapter for publication with N. Yoh as the first author.

Chapter 3 originated from discussions between M. Struebig and H. Bernard who conceived and designed the study. Fieldwork was completed by K. Mullin, S. Mitchell, and M. Struebig. The bat acoustic data was originally manually processed by K. Mullin and S. Basrur. All analysis of these data was subsequently performed by N. Yoh with assistance from D.J.I. Seaman. N. Yoh wrote the chapter for publication, with assistance from K. Mullin and M. Struebig. N. Yoh and K. Mullin share the first authorship.

Chapter 4 originated from discussions between N. Yoh, M. Struebig, J.E. Bicknell, and T. Kingston. Data was collected by N. Yoh, T. Kingston, M. Struebig, E. McArthur, O.E. Aylen, J.C.C. Huang, S.L. Mitchell, E.R. Jinggong, B.P.Y.H. Lee, and F.A.A. Khan. The classifier pipeline was developed by N. Yoh with support from M. Struebig and T. Kingston. All data analysis was conducted by N. Yoh. N. Yoh wrote the pipeline as an open-access script hosted from GitHub. N. Yoh wrote the chapter for publication as the first author, with collaborative input from J.E. Bicknell, T. Kingston, and M. Struebig.

Chapter 5 originated from discussions between M. Struebig and H. Bernard who conceived and designed the study. M. Struebig led the fieldwork. The bat acoustic data was manually processed by N. Yoh. All analysis of these data was subsequently performed by N. Yoh with assistance from D.J.I. Seaman and N.J. Deere. N. Yoh wrote the chapter for publication as the first author, with collaborative input from J.E. Bicknell, and M. Struebig.

Abstract

Halting the loss of biodiversity is one of our century's greatest challenges. In Southeast Asia, the biodiversity crisis is driven by unprecedented rates of forest clearance and degradation. Here, conservation efforts face the inherent complexity of trying to protect species in landscapes that also support human livelihoods. Most of our understanding of how bats are affected by land-use change in the tropics is limited to those species which can be monitored using live-trapping techniques. However, acoustic monitoring is an important survey method for monitoring the whole bat community in other regions of the world. In this thesis, advanced technological and statistical approaches are used to investigate how tropical bats respond to land-use change in Borneo. Specifically, the thesis explores the application of acoustic monitoring for bats in this region.

In the first chapter, I outline how conservation zoning can be effective at protecting tropical bat diversity. To do so, I use a combination of live-trapping and acoustic monitoring to assess bat diversity and activity within The Crocker Range Biosphere Reserve compared to the surrounding agriculture. This includes demonstrating how acoustic data can be used to analyse differences in bat activity between habitats, even if it is not possible to classify calls to sonotype/species. Namely, this research highlights the importance of strict protective zones for conserving forest specialist species.

Next, I refine the bat call classification process by using acoustic data manually classified to sonotype/species to assess the value of conservation set-asides. These data were combined with forest structural metrics measured using airborne LiDAR to assess the effectiveness of riparian reserves (areas of native forest along waterways) for conserving bats in oil-palm landscapes. Using Bayesian occupancy modelling, I demonstrate how - unlike other components of tropical biodiversity - forest quality is more important than riparian reserve width for maintaining suitable habitat for bats. This provides important evidence for designing effective conservation policies for tropical mammals.

The lack of reference call libraries and automated classification tools in Southeast Asia impedes ongoing acoustic monitoring for bats in this region. Acoustic monitoring generates large datasets - often approaching the scale of big data - which requires time and expertise to manually process. These costs limit the application of bioacoustics as a feasible method for monitoring bats in tropical regions, not least in Southeast Asia - thus, limiting our understanding of how bat communities persist in human-modified landscapes. Therefore, for the third data chapter, I focus on developing a semi-automated framework for classifying bat calls in Southeast Asia when reference libraries may be limited. As proof of concept, this framework is then used to develop a classifier for the bats of Borneo using reference calls for 52% of all 81 known echolocating species on the island. This classifier was developed using free software to ensure the same framework could be applied to other regions of Southeast Asia.

Last, I apply the newly-developed classifier to examine patterns of bat activity in response to habitat disturbance at the Stability of Altered Forest Ecosystems Project in Sabah. The data used in this assessment represented a seven-fold increase in the number of recording hours compared to the combined effort of the previous studies. This acoustic dataset is combined with forest structural metrics measured using airborne LiDAR to assess responses across a gradient of habitat disturbance - comprising old-growth forest, logged forest, and monoculture tree plantations. In this chapter, I demonstrate how logged forests can provide important habitat for bats in human-modified landscapes, maintaining acoustic diversity and activity of many common species, particularly when compared to tree plantations. However, I also demonstrate how old-growth forest remains important for conserving forest specialist species.

Collectively, in this thesis, I document how bats can benefit from conservation initiatives that protect landscape features within human-modified landscapes. Crucially however, I also demonstrate that as disturbance intensity increases, less resilient species are lost from human-modified landscapes. Whilst I provide important empirical contributions for understanding bat responses to disturbance, there remain substantial questions about how best to conserve these species and promote sustainable land use. The pipeline and classifier provided in this thesis will help improve accessibility for future acoustic studies to address these questions in Borneo and other areas of Southeast Asia. Therefore, the findings and resources presented here are an important step towards evidence-based land management for conserving Southeast Asian bats in human-modified landscapes.

Keywords: Chiroptera; Acoustic Monitoring; Land-use change; Machine Learning; Disturbance gradients; Borneo; Oil palm; Selective Logging; Tropical forests.

Table of contents

Acknowledgements .... **iv**

Author declaration .... **vi**

Abstract .... **vii**

Table of contents .... **ix**

List of tables .... **xii**

List of supplementary tables .... **xiii**

List of figures .... **xiv**

List of supplementary figures .... **xv**

Chapter 1 Introduction **1**

1.1 General background 1

1.2 Forest loss and disturbance 1

1.2.1 Deforestation related threats to tropical ecosystems 1
1.2.2 Drivers of tropical forest loss 2

1.3 Mechanisms for biodiversity loss & change 3

1.3.1 Mechanisms for change 3
1.3.2 Habitat quantity 3
1.3.3 Habitat suitability 3
1.3.4 Connectivity 4
1.3.5 The feedback loop of forest loss 5

1.4 Management & monitoring 5

1.4.1 The need for monitoring 5
1.4.2 The potential for acoustic monitoring 6

1.5 Potential applications for acoustics & tropical bats 7

1.6 Bat responses to land-use change 11

1.7 Bats & their role as acoustic bioindicators 14

1.8 Thesis structure 17

Chapter 2 Bat Ensembles Differ in Response to Use Zones in a Tropical Biosphere Reserve **22**

2.1 Abstract 23

2.2 Introduction 24

2.3 Materials and methods 27

2.3.1 Study Sites 27

2.3.2 Species Capture and Identification 29
2.4 Results 31
2.5 Discussion 45
2.6 Conclusions 50
2.7 Acknowledgements 50
2.8 Supplementary information 51

Chapter 3 Riparian Reserves Promote Insectivorous Bat Activity in Oil Palm Dominated Landscapes 60
3.1 Abstract 61
3.2 Introduction 62
3.3 Materials and methods 64
3.4 Results 73
3.4.1 Activity 73
3.4.1.1 Total activity and foraging activity between habitat types 73
3.4.1.2 Total activity in relation to habitat covariates 75
3.4.2 Site occupancy 78
3.5 Discussion 79
3.5.1 Bat Activity in Riparian Reserves 79
3.5.2 Foraging Activity in Riparian Reserves 80
3.5.3 Importance of Habitat Structure 80
3.5.4 Detection 81
3.6 Conclusions 82
3.7 Acknowledgements 82
3.8 Supplementary information 83

Chapter 4 A machine learning framework to classify Southeast Asia echolocating bats 94
4.1 Abstract 95
4.2 Introduction 96
4.3 Materials and methods 102
4.3.1 Input data 102
4.3.1.1 File Processing 103
4.3.1.2 Call parameters 108
4.3.1.3 Call classification via machine learning 110
4.3.1.4 Subset for training and testing data 110
4.3.1.5 Constructing the models 110
4.3.1.6 k-means clustering 111
4.3.2 Performance testing 111
4.3.2.1 Testing success rate on bootstrapped data 111
4.3.2.2 Accuracy thresholds for manual verification 111
4.4 Results 113
4.4.1 Minimum training dataset size 113
4.4.2 Call parameter importance 114
4.4.3 Success in performing classifications 117
4.4.4 Manual verification 120

4.5 Discussion ... 123
4.6 Conclusions ... 126
4.7 Availability ... 126
4.8 Acknowledgements ... 127
4.9 Supplementary information ... 128

Chapter 5 Benign effects of logging on aerial insectivorous bats in the palaeotropics revealed by remote sensing technologies ... 155
5.1 Abstract ... 156
5.2 Graphical abstract ... 158
5.3 Introduction ... 159
5.4 Materials and methods ... 161
5.4.1 Study area and sampling design ... 161
5.4.2 Defining the forest disturbance gradient ... 162
5.4.3 Acoustic sampling, processing, and bat identification ... 163
5.4.4 Statistical analysis ... 164
5.4.4.1 Acoustic diversity along a disturbance gradient ... 164
5.4.4.2 Bat community composition and habitat association ... 165
5.5 Results ... 165
5.5.1 Acoustic diversity ... 168
5.5.2 Bat community composition and association with habitats ... 169
5.6 Discussion ... 174
5.7 Conclusions ... 176
5.8 Acknowledgements ... 176
5.9 Supplementary information ... 177

Chapter 6 Discussion ... 191
6.1 Contributions to the research field ... 191
6.1.1.1 Impacts of land-use change ... 191
6.1.1.2 Survey design and tool development ... 194
6.1.1.3 Borneo as a case study ... 195
6.2 Implications for policy and practice ... 196
6.3 Future research ... 198
6.3.1.1 Activity versus abundance ... 198
6.3.1.2 Other landscape factors ... 198
6.3.1.3 Technological advancements ... 199
6.4 Conclusion ... 200

References ... 201

Appendix I Co-authored publications ... 232

Appendix II Borneo Bat Classifier user script ... 237

List of tables

Table 1.1 Naming conventions for each of the Borneo bat call groups defined in the thesis 20

Table 2.1 Number of individual captures of bat species across the four zones of CRBR and the surrounding agriculture 33

Table 2.2 Morphological data for 30 bat species from CRBR 36

Table 3.1 Bat activity and foraging activity alongside rivers sampled at the SAFE project Sabah 66

Table 3.2 Call parameters of constant frequency bats 69

Table 4.1 Number of bat passes used to train and test the Borneo Bat Classifier 105

Table 4.2 Performance (%) of random forest models in each stage of the Borneo Bat Classifier 118

Table 4.3 Confusion matrix for correct and incorrect classifications to call type 120

Table 5.1 Total bat passes per sonotype/species in each habitat along the disturbance gradient 166

List of supplementary tables

Table S2.1 Detailed site description for sampling localities within CRBR 51

Table S2.2 Relationship between bat activity and survey time 55

Table S3.1 Habitat characteristics of riparian sites 83

Table S3.2 Roost preference and foraging guild for insectivorous bat species present in Borneo 84

Table S3.3 Equations used for the final Generalised linear mixed effect models 88

Table S3.4 GLMM model comparisons for each group 89

Table S3.5 Likelihood ratio results for comparisons between top GLMM models for each group 91

Table S3.6 Measures for the predictive quality of each hierarchical Bayesian multispecies occupancy model 92

Table S3.7 Estimated regression parameters, standard errors, z-values and P-values for the Poisson GLMM 93

Table S4.1 Survey information relating to reference call data 128

Table S4.2 Call parameters used in Borneo Bat Classifier 134

Table S4.3 Sonotype descriptions 135

Table S4.4 Comparison of computational power for each random forest model 137

Table S4.5 Confusion matrix for correct and incorrect identifications for the CF classifier model 138

Table S4.6 Confusion matrix for correct and incorrect identifications for the FMqCF classifier model 139

Table S4.7 Overall classifier performance 140

Table S5.1 Site descriptions for each of the acoustic survey points in Sabah, Borneo 177

Table S5.2 Candidate models used to characterise bat activity at the 100m scale 185

Table S5.3 Biodiversity estimates of acoustic diversity (sonotype) richness summarised at for each habitat. 188

Table S5.4 Results of the post-hoc Dunn test comparing bat activity between the four habitat types 189

List of figures

Figure 1.1 The different pipelines for supervised and unsupervised machine learning........... 9

Figure 1.2 Map of study locations included in this thesis... 18

Figure 2.1 Map of study area in CRBR, Sabah, Malaysia 28

Figure 2.2 Photographic portraits of the 30 species captured in CRBR 32

Figure 2.3 Rarefaction curves for species captured in harp traps and mist nets 41

Figure 2.4 Violin plots demonstrating the variation in acoustic bat activity recorded within the four zones of the CRBR, Malaysia 43

Figure 2.5 Differences in call structure between the acoustic bat activity recorded across the four zones of the CRBR, Borneo, Malaysia... 44

Figure 3.1 Location of the 15 rivers sampled for bats via acoustic monitoring in the Stability of Altered Forest Ecosystems (SAFE) landscape of Sabah, Malaysian Borneo 65

Figure 3.2 Bat activity and foraging activity between rivers and habitats 74

Figure 3.3 The relationship between bat activity and significant habitat covariates in oil palm 76

Figure 3.4 The relationship between bat activity and significant habitat covariates within both forest and oil palm 77

Figure 3.5 Caterpillar plots of the output from the hierarchical Bayesian multi-species occupancy model 78

Figure 4.1 The four call types used in the Borneo bat classifier... 100

Figure 4.2 Sonograms for the four call types used in the Borneo bat classifier................... 101

Figure 4.3 Map for the reference calls used to train the Borneo bat classifier..................... 103

Figure 4.4 Pipeline for designing the semi-automated classifier 110

Figure 4.5 Diagram for the user application of the classifier framework 112

Figure 4.6 Comparison of different training data sizes for each classification algorithm ... 114

Figure 4.7 Comparison of variable importance in each classifier stage............................... 116

Figure 4.8 Percentage of echolocation calls requiring manual verification using the Borneo Bat Classifier... 122

Figure 5.1 Map of the Stability of Altered Forest Ecosystem Project (SAFE) and sampling sites in Maliau Basin, Malaysian Borneo.. 162

Figure 5.2 Species/sonotype diversity along the disturbance gradient 168

Figure 5.3 Nonmetric multi-dimensional scaling showing the (dis)similarities between bat communities across the four habitat types. 169

Figure 5.4 A comparison of overall nightly bat activity across each of the four habitat types. 171

Figure 5.5 Caterpillar plots from the General Additive Models for Location, Shape, and Space (including 95% confidence interval).. 172

Figure 5.6 Caterpillar plots from the General Additive Models for Location, Shape, and Space (including 95% confidence interval) for the dominant sonotypes/species 173

Figure 6.1 Visual summary of the main thesis findings.. 192

List of supplementary figures

Figure S2.1 Results of the pairwise comparison of mean bat activity 56

Figure S2.2 Scatterplot of temporal variation in bat activity 57

Figure S2.3 The relative importance of acoustic variables 58

Figure S2.4 Sonogram of high frequency *Hipposideros* species 59

Figure S4.1 Comparison of the performance of six supervised machine learning algorithms 148

Figure S4.2 Sonograms for “constant frequency” species 149

Figure S4.3 Sonograms for “frequency modulated quasi constant frequency” and “quasi constant frequency” sonotypes/species 150

Figure S4.4 Sonograms for “frequency modulated” call type 151

Figure S5.1 Example sonograms for each of the four call types used in the study 190

Chapter 1 Introduction

1.1 General background

Humans have modified three-quarters of the world's land area (Luyssaert *et al.* 2014; Arneth *et al.* 2019). As well as providing food and shelter for growing populations, adapting the environment has been important for developing social security and cultural identities. Currently, agriculture, urbanisation, forestry, and wildfires are the main drivers of land-use change (Curtis *et al.* 2018). However, land-use change has had profound implications for biodiversity (Bradshaw, Sodhi and Brook 2009; Cowie, Bouchet and Fontaine 2022), ecosystem functioning (Cardinale *et al.* 2006), and the global climate (Coe *et al.* 2013; Schulte to Bühne *et al.* 2021). The associated habitat loss and disturbance have fuelled the current biodiversity crisis which many describe as the start of the Sixth Mass Extinction (Ceballos, Ehrlich and Dirzo 2017; Giam 2017; Cowie, Bouchet and Fontaine 2022). The number of species threatened with extinction is forecast to continue rising over the next 50 years, particularly in the most biodiverse but often economically poor regions (Newbold *et al.* 2015; Ceballos, Ehrlich and Dirzo 2017; Tilman *et al.* 2017). Therefore, we need to act urgently to address the underlying causes of anthropogenic disturbance and reinforce current conservation policies to safeguard global biodiversity (Tilman *et al.* 2017).

1.2 Forest loss and disturbance

1.2.1 Deforestation related threats to tropical ecosystems

Forests cover around 31% of the world's land area (over four billion hectares; FAO & UNEP, 2020). They are important reservoirs of biodiversity (Mora *et al.* 2011; Giam *et al.* 2012), regulate global climate (Coe *et al.* 2013), water quality (Mello *et al.* 2018), and support the livelihoods of a quarter of the people on the planet (Newton *et al.* 2020). However, in the last 30 years alone, there has been a global net loss of forest the size of Libya (178 million hectares; FAO & UNEP, 2020). Agricultural land now constitutes a greater area of land than forests (FAO 2021). Deforestation has a cascade of environmental impacts, including a reduction in global carbon sequestration and extensive declines in habitat availability for tens of thousands of species (Raven 1988; Malhi *et al.* 2008; Coe *et al.* 2013).

However, forest clearance is not the only way the value of forests for biodiversity and climate regulation is diminished. Forest disturbance or degradation impacts a vastly greater area than forest clearance. Grantham *et al.* (2020) found that only 25.9% of remaining global forests were free of any anthropogenic modification and almost 60% had suffered significant modification. Forest degradation results in a loss of forest ecological integrity and produces more carbon dioxide than highway vehicles annually (2.1 billion tons; Pearson *et al.*, 2017; Grantham *et al.*, 2020). As degradation impacts habitat suitability, it leads to changes in the composition of ecological communities and causes localised species extinctions (Laurance *et al.* 2018; Grantham *et al.* 2020). Ninety-five percent of annual deforestation occurs in the tropics (Ritchie 2021). Tropical forests represent 10% of the world's land - a third of all forests - but support two-thirds of global terrestrial diversity (Raven 1988; Giam *et al.* 2012; Giam 2017). Current deforestation rates are expected to exceed Wright's (2010) predictions of 36% loss of remaining tropical forest by 2050 (Hansen *et al.* 2013) and intrinsically this will have far-reaching consequences, not only for tropical species but biodiversity worldwide. It is estimated over 80% of eukaryotic species remain undiscovered (7.4-10 million) and the vast majority of these occur in tropical forests (Mora *et al.* 2011; Giam *et al.* 2012). Therefore, many species reliant on tropical forests are likely to go extinct before ever being documented (Giam *et al.* 2012; Cowie, Bouchet and Fontaine 2022).

1.2.2 Drivers of tropical forest loss

Conversion of forest to agriculture and pasture remains the dominant driver of land-use change in the tropics (Song *et al.* 2018). However, the rates of forest clearance and their drivers vary between continents. The highest mean rates of forest clearance occur in Latin America (Hansen *et al.* 2013; Malhi *et al.* 2014; Rosa *et al.* 2016). In 2020, the Brazilian Amazon, the world's largest tropical forest, suffered the highest rates of deforestation for a decade (Silva Junior *et al.* 2021). Approximately 70% of this deforestation can be attributed to cattle ranching (Malhi *et al.* 2008; Aleixandre-Benavent *et al.* 2018). Increasing global demand for soy is further fueling forest loss in regions such as Mato Grosso, Brazil (Silva and Lima 2018). Whilst forest is not generally cleared for soybean plantations, displacing cattle ranches indirectly instigates further deforestation elsewhere (Barona *et al.* 2010).

Unlike Mesoamerica and South Asia, much of the Southeast Asian tropical forest estate was largely intact until the 1970s (Ghazoul and Sheil 2010). However, as international demand rapidly accelerated, Malaysia, Indonesia, and to a smaller extent Thailand became the epicentre of global oil palm production. Asia now has the highest relative rates of forest clearance predominantly in Malaysia and Indonesia (Hansen *et al.* 2013; Malhi *et al.* 2014; Rosa *et al.* 2016). Between 2005-2015, this equated to a forest loss rate of approximately 8 million ha yr^{-1} (Estoque *et al.* 2019). Over the last two decades alone, oil palm production has increased by 260 % (FAO 2018). There is also growing concern about the environmental costs of additional large-scale monocultures, including coconut (*Cocos nucifera*) which is predominantly produced in the Philippines and Indonesia (Magrach and Sanz 2020; Meijaard *et al.* 2020). Africa, meanwhile, has deforestation rates that are less than half those observed in Asia and are primarily the result of smallholder clearance (Malhi *et al.* 2014; Leblois, Damette and

Wolfersberger 2017). Hence drivers of deforestation differ between regions and this should be an important consideration when establishing conservation policies.

1.3 Mechanisms for biodiversity loss & change

1.3.1 Mechanisms for change

There are multiple mechanisms through which species are affected by land-use change such as forest loss and disturbance. These principally affect habitat quantity, suitability, and/or connectivity for various species. The impact of land-use change on biodiversity may be immediate (e.g., localised extinctions occurring quickly once habitat is lost) or delayed (e.g., where remaining habitat cannot support populations long-term; Hylander and Ehrlén, 2013). Numerous, multifaceted mechanisms contribute to the overall impact on species, including fragmentation (reduced patch size, isolation, edge effects), resource reduction, biotic homogenization, increased risk of invasion, predation, hunting, parasites/disease, and genetic drift (Holl, Luong and Brancalion 2022; Fahrig 2017; Ewers *et al.* 2011). For example, habitat configuration has been shown to influence dung beetle β-diversity as communities in core forest differ are distinct from communities in edge-affected forest (Filgueiras *et al.* 2016).

1.3.2 Habitat quantity

Forest loss inevitably results in less habitat available for species, as many are unable to adapt to modified ecological conditions (Laurance *et al.* 2018). However, deforestation is rarely uniform across a landscape and often leads to a patchwork of forest patches isolated in agricultural landscapes (Curran, 2004). Patch size is an important determinant in the ability of the remaining habitat to support biodiversity (Ewers and Didham 2006), particularly for species that have large home ranges, which are highly specialised, or occupy higher trophic levels. The average patch size for tropical rainforest is 4.33 km^2 (FAO & UNEP 2020). For context, the home ranges of sambar deer (*Rusa unicolor*) and their predator the Malayan tiger (*Panthera tigris jacksoni*) are estimated between 2.4–11.8 km^2 and over 70 km^2 respectively (Chatterjee *et al.* 2014; Lazarus *et al.* 2021). As well as the initial loss of resources caused by loss of habitat, fragmentation also results in crowding effects whereby more individuals occupy a smaller space and are reliant on the resources it contains (Ewers and Didham 2006). If a patch does not have adequate resource availability, then it cannot support viable populations long-term, leading to the defaunation of these remaining forests.

1.3.3 Habitat suitability

The quality of the habitat in remaining forest patches is not uniform. The margins of these artificially created fragments are subjected to biotic and abiotic changes which exceed the natural intrinsic variation in conditions (Laurance *et al.* 2018). These *edge effects* deteriorate a habitat's suitability for many forest-dependent species. They include a myriad of effects from changes to soil moisture content and increased exposure to wind, to elevated tree mortality and increased disease risk (Laurance *et al.* 2018). Edge effects are thought to impact approximately 85% of global vertebrate species (Haddad *et al.* 2015) and half of tropical forest is predicted to become forest edge by the end of this century (Fischer *et al.* 2021).

Logging also impacts the quality of forest, both at its edge and interior (Asner *et al.* 2009; Costantini, Edwards and Simons 2016). Selective logging impacts a quarter of the world's tropical forests which alters forest structure, localised microclimatic conditions, and can disrupt ecosystem functioning (Asner *et al.* 2009; Blaser *et al.* 2011; Costantini, Edwards and Simons 2016). In Borneo, species declines in response to logging have been reported for multiple taxa including mammals, birds, and butterflies (Cleary *et al.* 2009; Hamer *et al.* 2014; Chapman *et al.* 2018). However, these impacts are less severe than other forms of anthropogenic disturbance to tropical forests (Bicknell *et al.* 2014; Edwards, Tobias, *et al.* 2014). Therefore, many argue logging concessions provide a financial incentive for preserving large areas of forest that may otherwise be at risk from clearing (Meijaard and Sheil 2007; Putz *et al.* 2012).

Species' responses to logging - as with disturbance more broadly - are strongly influenced by life-history traits and niche breadths (Cleary *et al.* 2009; Hamer *et al.* 2014). Species dependent on old-growth tree species for foraging or roosting (e.g., cavity roosting bats; Struebig *et al.*, 2013) experience the greatest population declines in response to logging as these trees are preferential for harvesting (Edwards, Tobias, *et al.* 2014; Costantini, Edwards and Simons 2016). Species with narrow ecological niches or large-bodied/long-lived species with low fecundity are also more vulnerable (Edwards, Tobias, *et al.* 2014; Costantini, Edwards and Simons 2016). Different feeding guilds have varying sensitivities to disturbances such as logging. Insectivorous and animalivorous species are thought to be particularly sensitive to disturbance (Edwards, Tobias, *et al.* 2014; Brändel *et al.* 2020). However, there is often interspecific variation in response within broad feeding guilds (Cleary *et al.* 2007; Brändel *et al.* 2020). For example, in Central Kalimantan, understory insectivorous birds increased in abundance within logging concessions at the expense of terrestrial insectivores and hornbills (Bucerotidae spp.; Cleary *et al.*, 2007). Therefore, pooling species into broad feeding guilds may mask species-specific responses (Hamer *et al.* 2014). Importantly, hornbills are key seed dispersers but are highly sensitive to logging disturbance. These shifts in community assemblage disrupt ecosystem functioning within and neighbouring logging concessions (Edwards, Tobias, *et al.* 2014). Therefore, this recognition of conservation value should not be interpreted as impunity for all remaining primary forest to be made available for timber harvesting (Meijaard and Sheil 2007).

1.3.4 Connectivity

Ensuring connectivity between remaining habitat patches can alleviate some of these impacts of habitat loss and disturbance by facilitating dispersal and greater access to resources (Kupfer, Malanson and

Franklin 2006). Matrix composition is important in determining the isolation of a forest fragment by acting as a filter for species movements. Low contrast matrices (such as mature secondary forest) facilitate higher connectivity than high contrast matrices (such as plantations) by reducing the gradient of change (Ricketts 2001). However, connectivity is not only related to the physical distance between patches (structural connectivity) but also a species' ecological tolerance and dispersal ability within the surrounding matrix (the functional connectivity; Crooks and Sanjayan, 2006). A prime example of this are understory passerines. Despite their capacity for flight, the increased light associated with agriculture acts as a barrier between forest patches for many species (Sodhi *et al.* 2011; Hamer *et al.* 2014). Tropical forest specialists typically lack the ecological flexibility necessary to cross high-contrast matrices, even for short distances. In contrast, generalist species have a much greater capacity which can lead to a higher prevalence of invasive species in fragmented forests (Laurance *et al.* 2018).

1.3.5 The feedback loop of forest loss

The effects of land-use change are not isolated to a specific location at a specific time. Rather, they contribute to a feedback loop that threatens the ecological integrity of global forests and the global climate (Runyan, D'Odorico and Lawrence 2012; Coe *et al.* 2013; Schulte to Bühne *et al.* 2021). Increased access (e.g., road infrastructure), increased risk of fire, and more extreme weather conditions all resulting from the initial disturbance lead to a legacy of future forest loss. Therefore, protecting the integrity of tropical forests is of urgent, global socio-ecological importance.

1.4 Management & monitoring

1.4.1 The need for monitoring

There is growing resistance against traditional conservation strategies, such as strictly protected areas or fortress conservation, which evict people entirely from large areas of high conservation value (Rai *et al.* 2021). These methods in isolation cannot halt biodiversity loss (Mora and Sale 2011; Balmford 2021) and they have severe consequences for those reliant on resources in these areas (Rai *et al.* 2021). Around 70% of protected areas in the tropics are currently occupied by human communities (Terbough, 2022). Therefore, there is an ever-growing acknowledgement of the need to balance sustainable resource use and conservation to achieve greater conservation success and better societal outcomes. Environmental certification schemes aim to minimise the environmental impacts of crops, such as oil palm (Murphy, Goggin and Paterson 2021). Whilst these initiatives demonstrate a positive commitment to addressing conservation issues, it is unclear how effective these schemes are for protecting biodiversity (Koh and Wilcove 2008; Edwards *et al.* 2010). As such, it is vital to monitor the implementation and effectiveness of conservation initiatives (Salafsky *et al.* 2019). However, there is often a mismatch between the optimum timelines of research and the time available for practitioners to act on imminent and escalated threats to biodiversity (Grantham *et al.* 2009; Cardinale *et al.* 2012;

Jarvis *et al.* 2020). This means there is often incomplete baseline knowledge about species, community dynamics, or the knock-on effects of implementing specific conservation actions (Grantham *et al.* 2009). This can result in ineffective conservation strategies. Conservation funding is limited and failed conservation actions have real consequences for the wildlife they aim to protect. Therefore, dynamic, responsive research and revisions of conservation actions are necessary to meet conservation goals.

1.4.2 The potential for acoustic monitoring

Despite the importance of evidence-based management, it is not possible to monitor every species. As well as limited time and funding, there are many additional constraints to carrying out ecological monitoring (Gardner *et al.*, 2008). Where a species occurs is the most fundamental data used to aid conservation but accurate species distribution data is still missing for most taxa (Fisher-Phelps *et al.* 2017). Of the 141,601 extant species included on the IUCN Red List of Threatened Species, over 14% are classified as data deficient and this only represents a fraction of understudied species (IUCN 2022). A lack of existing ecological knowledge, along with differences in societal preferences (Troudet *et al.* 2017), accessibility between regions (Reddy and Davalos 2003; Fisher-Phelps *et al.* 2017), and the relative difficulty or cost of sampling methods (Pawar 2003) results in spatial and taxonomic bias across monitoring efforts. This highlights the need for alternative, cost-effective, and scalable methods for ecological monitoring (Cardinale *et al.* 2012; Gibb *et al.* 2019).

Recent technological advances have dramatically increased the potential applications for bioacoustic monitoring. By monitoring environmental and biotic sound, bioacoustic monitoring provides a non-invasive technique for monitoring individuals, populations, and ecosystem health (Browning *et al.* 2017; Gibb *et al.* 2019; Chhaya *et al.* 2021). Bioacoustic monitoring is currently used to achieve many conservation goals such as monitoring how species respond to land-use change, mapping biodiversity, monitoring poaching pressure, and can be utilised for a broad range of taxa (Wrege *et al.* 2017; Gibb *et al.* 2019; Sugai *et al.* 2019). It is also used as a complementary tool alongside camera trap and live-trapping survey methods (Buxton *et al.* 2018; Appel *et al.* 2021). For example, it can be used to monitor how hunting pressure (acoustic detections of gunshots) affects animal behaviour (e.g., avoidance behaviour detected using camera traps; Buxton *et al.*, 2018).

Bioacoustic data is generally matched to known reference sound data to determine which species made a vocalization. But where this is not possible, acoustic indices also provide a way to characterise ecological communities and soundscapes (the acoustic environment). This approach can describe a range of characteristics for acoustic communities or soundscapes, such as the amplitude, evenness, richness, or heterogeneity, which in turn can be used to monitor responses to environmental change (Sueur *et al.* 2014). It can reliably be used as a proxy for monitoring non-soniferous species (such as Nicrophorine burying beetles; Dodgin, Hall and Howard, 2020) and to rapidly assess how habitats are affected by land-use change or the impact of conservation actions. An intermediary approach is the targeted monitoring of acoustic bioindicators. This approach relies on knowing how a specific species or component of the soundscape responds to a set of environmental conditions. A bioindicator can therefore be used to assess the condition of an ecosystem without monitoring each component independently (Holt and Miller 2010). For example, four species of bush bird are effective bioindicators of mining disturbance in Australia (Read, Parkhurst and Delean 2015). This provides a cost-effective, passive method to easily monitor mining disturbance over large spatial scales without recording all species' responses. However only specific species/acoustic signals can be used as bioindicators, and these will vary between spatial context and the conservation question (Carignan and Villard 2002; Rice and Rochet 2005; Russo and Jones 2015).

1.5 Potential applications for acoustics & tropical bats

Automated or semi-automated classifiers are increasingly used to alleviate the time requirements of manually classifying animal vocalisations (Valletta *et al.* 2017; Kwok 2019; Tabak *et al.* 2021). As such, open-source and commercial tools are being developed to expand the capacity to monitor bats more extensively (Bas, Bas and Julien 2017; *SonoBat-4* 2017). This reduces the demand for experts and increases local capacity to conduct acoustic surveys. These data produced can be used to estimate species occupancy and monitor spatial and temporal changes to community composition and activity (Browning *et al.* 2017; Gibb *et al.* 2019; Sugai *et al.* 2019), which in turn can be used as a proxy for other metrics of biodiversity (Browning *et al.* 2017; Doohan *et al.* 2019). Classifying calls in this way is useful to assess the impact of specific threats or conservation actions or forecast how bats are likely to be affected by climate change. Therefore, it can be used to inform conservation policy and practice.

Most modern classifiers are developed using machine learning techniques. Machine learning is a branch of artificial intelligence that develops problem-solving tools without the need to explicitly program the solution (Bianco *et al.* 2019). Several different models can be used to achieve this (Figure 1.1). In supervised machine learning, the model uses a labelled dataset to learn which set of parameters best distinguishes between species. For identifying bat calls, this labelled dataset can either be D-dimensional vectors of call parameters or, increasingly, graphical representations of calls (e.g. spectrographs; Frick, 2013; Bas, Bas and Julien, 2017; Kobayashi *et al.*, 2021). Supervised machine learning includes techniques such as neural networks, random forests, and support vector analysis, and is the dominant model used for call recognition in bioacoustics (Bianco *et al.* 2019). In many machine learning applications, this is not appropriate either due to a lack of data (e.g., lack of reference calls for rare species) or where there are no prior hypotheses regarding the underlying patterns in these data (e.g., delineating the taxonomy between species; Derkarabetian *et al.*, 2019). Unsupervised machine learning provides an alternative approach by identifying structure within unlabeled datasets (e.g., a set of bat calls where the species identification is unknown) and organizing these data based on inherent similarity. Unsupervised machine learning techniques (e.g., *k*-means clustering) have been used to classify animal vocalizations for gibbons and dolphins but have not been widely applied to other areas of acoustic monitoring (Frasier *et al.* 2017; Clink and Klinck 2021). The application of other machine learning models for acoustic monitoring, such as semi-supervised machine learning (which uses a combination of labelled and unlabelled training data) and reinforcement learning, has rarely been tested. As such, almost all bioacoustic classifiers are built using supervised machine learning models (Bianco *et al.* 2019).

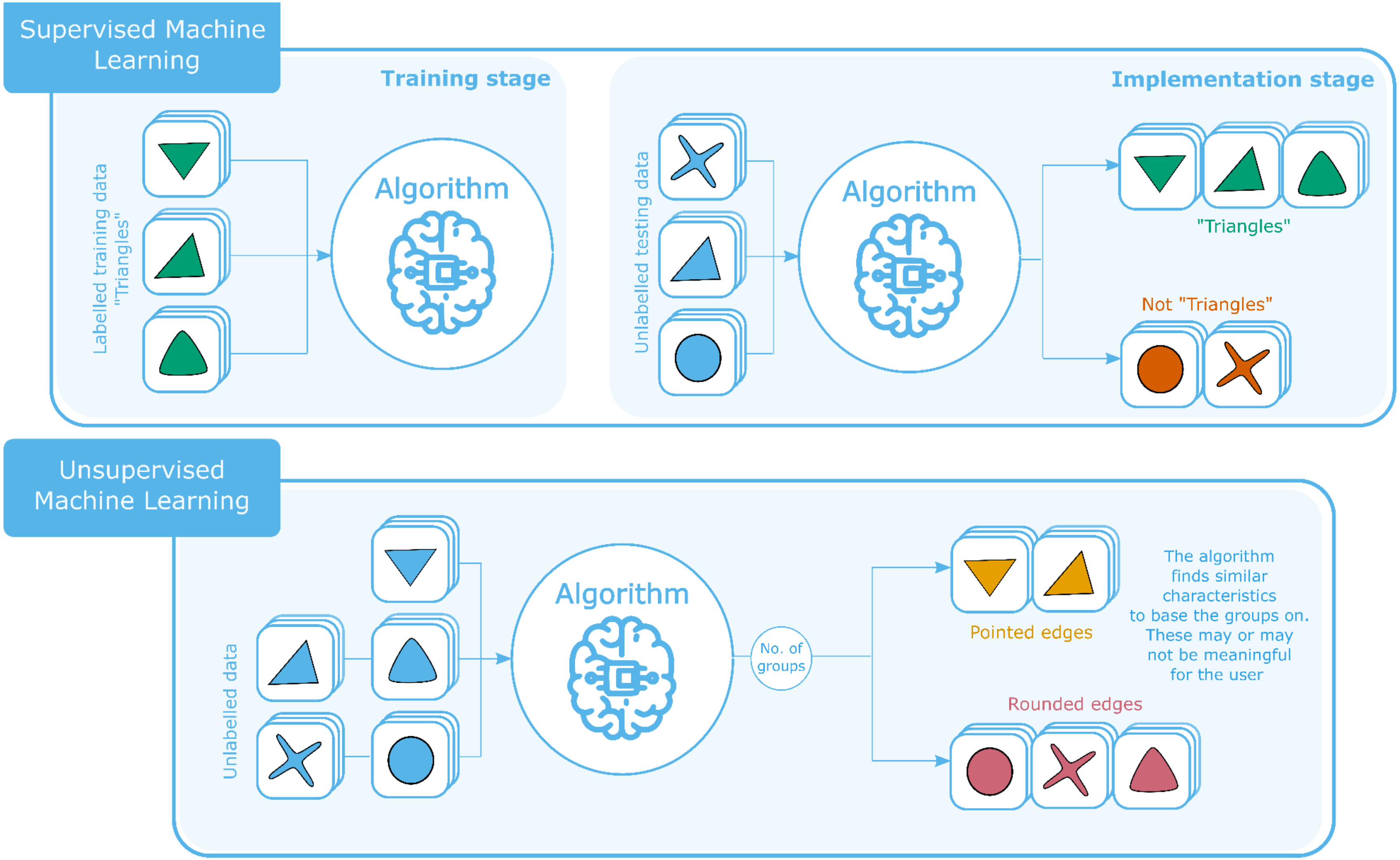


Figure 1.1 The different pipelines for supervised and unsupervised machine learning.

Whilst automated classifiers provide a quick and repeatable method for analysing acoustic data, they are not without criticism. The accuracy of many classifiers is not effectively tested, particularly for closed-source algorithms in commercial tools (Gibb *et al.* 2019), and poor or inconsistent accuracy is a significant limitation of automated classifiers. The training data used will ultimately dictate the quality of any classifier. If a classifier is trained on a very narrow pool of calls (e.g., from the same individual, in the same location, or in an unnatural setting) then the tool will not encompass all the natural variation in call characteristics or recording conditions, making species identification unreliable. This has the potential to misinform conservation actions and policies.

One proposed solution is the use of semi-automated classifiers. This approach automatically classifies calls to species, but the operator then manually verifies a subset of these calls based on a predefined set of conditions (e.g., a species where the classifier is known to perform poorly or for species of high conservation concern; see López-Baucells *et al.*, 2019 for an example). This reduces the volume of calls to be processed but integrates quality control into the identification process. As outlined by Kwok (2019), so long as an operator is aware of a classifier's accuracy, they can take steps to mitigate bias and interpret the results appropriately. These "human-in-the-loop" systems have been shown to help mitigate the inherent bias in machine learning outputs across many computing applications (Wu *et al.* 2022). That being said, by reintroducing a manual verification step, the potential for operator bias is reintroduced (Jennings, Parsons and Pocock 2008). The expertise of the operator is a consistent factor affecting species identification across all acoustic studies (Jennings, Parsons and Pocock 2008). However, Lundberg *et al.* (2021) demonstrated how pairing automated tools with manual confirmation as part of a citizen science project can provide high-quality results comparable to manual validation by experts. An alternative option, depending on the research question, is to trade specificity for better accuracy. The hurdle for many bat classifiers is interspecific overlap in call characteristics which can result in inflated detections (e.g., false positives) or missed detections (e.g., false negatives) for species' (Russo, Ancillotto and Jones 2018). However, this overlap in call similarity can be used as an advantage to train classifiers that recognise different call groups. Recently, Roemer, Julien and Bas (2021) used this approach to develop a universal tool for classifying bat calls globally into one of eight sonotypes - indicative of different ecological niches. Whilst it is not appropriate for species-specific monitoring, this tool can be used to quickly monitor changes to acoustic bat communities over time. It is therefore important to identify what is required of automated classifiers versus their idealised application.

Combining all these advantages - low-cost detectors, citizen science, semi-automated classifiers, and scall indicator groups - dramatically expedites the large-scale, long-term potential for acoustic monitoring so long as the limitations of these techniques are acknowledged. Therefore, classifiers should not be perceived as perfect tools which determine species identification with absolute certainty, but rather as optimization tools that require careful assessment and interpretation.

1.6 Bat responses to land-use change

Most research studying how tropical bats respond to land-use change is concentrated in the neotropics. Here, a large proportion of the bat community can be sampled using mist-netting and thus does not rely on acoustic monitoring (Meyer, Struebig and Willig 2016). As such, phyllostomids dominate the literature in terms of what is known about how tropical bats respond to forest disturbance. However, the assemblage structures between the neo- and palaeotropics are very different. Palaeotropical bat faunas have much higher numbers of insectivorous species and fewer frugivorous/nectarivorous species compared to those found in neotropical forests (Francis 1990). In Borneo, insectivorous species represent over 80% of the bat community and most species are poorly sampled using live-trapping methods (Kingston 2013; Phillipps and Phillipps 2016). Therefore, studies focused on phyllostomid species are unlikely to be indicative of bat responses in Southeast Asia or Africa. Even when excluding this geographic bias, few generalizations can be concluded about how tropical bats respond to disturbance. What we do know is that bat responses are highly species and ensemble specific (Meyer, Struebig and Willig 2016). Functional traits, such as diet, wing morphology, home range size, roost preference, dispersal ability, and stratum preference, all affect a species' sensitivity to patch size, isolation, and edge effects (Struebig *et al.* 2013; Farneda *et al.* 2015; Meyer, Struebig and Willig 2016; Núñez *et al.* 2019).

1.6.1 Why landscape context matters for bats

Bats exhibit a broad range of dispersal abilities. Consequently, landscape composition can be more important for bats compared to many other taxa (Meyer, Struebig and Willig 2016). Dispersal ability is governed by mobility (e.g., wing morphology and echolocation type) but also by external pressures, such as light sensitivity and vulnerability to predators (Kingston, Francis, *et al.* 2003; Azam *et al.* 2016). A combination of these factors means that even though bats have the ability to fly, forest interior specialists typically have a low dispersal ability in human-modified habitats. Hence, the contrast of the surrounding matrix compared to forest plays a significant role in governing the relative importance of different components of the landscape configuration, e.g., isolation versus fragment area. Where there is a high-contrast matrix (e.g., intensive agriculture; Meyer and Kalko, 2008), isolation is the most important determinant for species richness and composition. However, in landscapes with low matrix contrast (e.g., between forest and rubber plantations; Struebig *et al.*, 2008), patch size may be more significant. Therefore, more studies are recognizing the need for matrix-inclusive approaches when evaluating bat responses (Kupfer, Malanson and Franklin 2006; Meyer, Struebig and Willig 2016). This includes assessing responses along a gradient of varying habitat quality (e.g., from old-growth forest, logged forest, regenerating forest, to agriculture). Some studies have found that measures of forest quality (e.g., canopy height) are better predictors of assemblage level responses (e.g., changes to community composition) compared to measures of landscape configuration (Meyer and Kalko 2008;

Henry, Cosson and Pons 2010). However, this is not consistent across all landscapes (see Meyer, Struebig and Willig, 2016) and will vary between different spatial scales for different species (Pinto and Keitt 2008; Henry, Cosson and Pons 2010; Klingbeil and Willig 2010) and in relation to other landscape features, such as the proximity to water bodies (Pinto and Keitt 2008; Torrent *et al.* 2018). There can also be spatiotemporal variation in the relative importance of different landscape characteristics. These can relate to differences in food availability, dietary flexibility, and different energy requirements across the year (e.g. for reproduction; Klingbeil and Willig, 2010; Ferreira *et al.*, 2017).

1.6.2 Bat responses to logging in the tropics

There is considerable bias concerning which aspects of land-use change are assessed in the relevant bat literature. Meyer, Struebig and Willig (2016) found there were comparatively few studies assessing how tropical bats respond to logging or agroforestry compared to other aspects of habitat modification. To date, these studies reveal idiosyncratic responses to different logging regimes for different bat ensembles (Clarke, Rostant and Racey 2005; Presley *et al.* 2008; Castro-Arellano *et al.* 2009; Bicknell, Struebig and Davies 2015; Castro *et al.* 2021). For example, two logging-effect studies using mist-net captures in South America (Trinidad and Brazil) showed that gleaning insectivores and animalivores were negatively impacted by logging disturbance, whereas frugivorous and nectarivorous bats benefited from the increased foraging opportunities in new forest gaps (Clarke, Rostant and Racey 2005; Presley *et al.* 2008). Therefore, the impacts of logging are better captured by monitoring changes to community composition rather than changes in species richness. In Guyana, Bicknell, Struebig and Davies (2015) showed that phyllostomid bats benefit from low-intensity harvest techniques, such as reduced impact logging. However, even under reduced impact guidelines, logging has been shown to affect the behaviour, physiology, and resource availability of other bat species in both the neotropics (Brazil; Castro *et al.*, 2021) and palaeotropics (Borneo; Seltmann *et al.*, 2017; Hemprich-Bennett *et al.*, 2021). Therefore, whilst these guidelines may protect generalist bat species, logging can lead to local extinctions of rare, forest specialists (Presley *et al.* 2008). Again, how logging sites are embedded in the wider landscape context will ultimately moderate their ability to support bat populations (Presley *et al.* 2008).

The most comprehensive study investigating how logging impacts palaeotropical insectivorous bat communities was conducted in Sabah, Borneo. Struebig *et al.* (2013) investigated how different logging regimes impacted forest bats using harp-trapping, a technique that more reliably samples these taxa in Southeast Asia than the mist nets commonly applied in other regions (Francis 1989; Kingston 2013). The results of this study mirror those in the neotropics with no consistent association between species richness and logging disturbance, but distinct alterations in the species composition between old-growth and logged forest. Although more reliable than mist-netting, harp trapping is still limited by a low detection probability for many aerial-insectivorous species (Kingston 2013). The inclusion of species that have previously been under-sampled in both the neotropics and palaeotropics may fundamentally affect our understanding of the relative impacts of different logging regimes.

1.6.3 The value of tree plantations for bats

Forest clearance for large-scale tree plantations, such as oil palm or rubber, has a much greater impact on tropical diversity than logging (Gibson *et al.* 2011). Despite being one of the main drivers of forest clearance in Southeast Asia, there are few studies assessing how oil palm or other types of tree plantation impact bats. Those that have been conducted have again been constrained to studying bats that can be well sampled using mist-nets. In the case of Southeast Asia, this is predominantly frugivorous pteropodids. Palm height and crop density are negatively correlated with pteropodid species richness in Peninsular Malaysia (Azhar *et al.* 2015). Overall bat abundance has been shown to increase in certain oil palm plantations (Danielsen and Heegaard 1995; Shafie *et al.* 2011). However, it is heavily dominated by common species, such as *Cynopterus* spp., which are typically associated with disturbance (Danielsen and Heegaard 1995; Shafie *et al.* 2011). In Borneo, Fukuda *et al.* (2009) found that oil palm was a poor habitat substitute for most frugivorous and insectivorous bats. Therefore, it is vital to assess differences in community composition between habitats, not just overall abundance. Pteropodids differ from Yangchiroptera in almost all the functional traits known to influence a species' sensitivity to disturbance (e.g., dispersal ability, diet, roost preference). Therefore, their responses are very unlikely to reflect those of the complete bat assemblage.

Most studies assessing the value of agroforestry for bats have focused on shade-grown coffee (*Coffea arabica, Coffea canephora*) and cacao (*Theobroma cacao*) plantations which are common in many tropical countries (Meyer, Struebig and Willig 2016). These plantations are thought to provide a low-contrast matrix between forest fragments due to their structural similarity with natural forest. At low intensities, coffee has been shown to support phyllostomid species richness and forest insectivore activity (Pineda *et al.* 2005; Williams-Guillén and Perfecto 2011). Similarly, structurally-complex shade cacao plantations have been shown to support high levels of bat diversity, including gleaning animalivores (Pardini *et al.* 2009; Meyer, Struebig and Willig 2016). However, there are shifts in community composition and the availability of food resources across all guilds as management intensity increases (Pineda *et al.* 2005; Williams-Guillén and Perfecto 2011). Landscape context remains an important factor influencing bat species richness in coffee or cacao plantations (Numa, Verdú and Sánchez-Palomino 2005; Faria and Baumgarten 2007). In Brazil, cacao plantations that were isolated (> 1 km) from remanent forest were dominated by a limited number of generalist species (Faria and Baumgarten 2007). This reinforces the importance of considering connectivity and source-sink dynamics within human-modified landscapes even in low-contrast matrices.

1.7 Bats & their role as acoustic bioindicators

1.7.1 Potential bioindicators for forest disturbance

Alongside birds, bats are recognised as the best acoustic bioindicators for monitoring land management practices (Doohan *et al.* 2019). Bats have historically and continue to be the most popular candidates for acoustic monitoring as they are difficult to survey via traditional techniques due to their nocturnal behaviour but they have ‘acoustically active’ lifestyles (Browning *et al.* 2017; Sugai *et al.* 2019). Acoustic monitoring of bats has been used effectively to monitor river quality, forest disturbance, urbanisation, bioaccumulation, and climate change (Russo *et al.* 2021).

Bats constitute at least 1,446 species - approximately a quarter of mammal diversity and represent a broad range of physiological and ecological diversity (Burgin *et al.* 2018; Simmons and Cirranello 2021a). They provide ecosystem services to forest health and regeneration, such as seed dispersal, pollination, and nutrient cycling, as well as to agriculture and human health through pest predation (Kunz *et al.* 2011; Ramírez-Fráncel *et al.* 2022). A global review by Ramírez-Fráncel *et al.* (2022) found bats consumed over 752 insect species and contributed to the reproduction of at least 549 plant species, including many commercial species (e.g., agave, durian, mango, and shea trees; Ghanem and Voigt, 2012). For example, in Thailand, the wrinkle-lipped free-tailed bat (*Chaerephon plicatus*) consumes 20,000 metric tonnes of insects per year including many rice pests such as planthoppers (*Sogatella* spp.; Leelapaibul, Bumrungsri and Pattanawiboon, 2005). As such, bats were historically revered in many cultures as symbols of good fortune, predominantly across the Asia-Pacific region (Rocha, López-Baucells and Fernández-Llamazares 2021).

Despite their importance, 80% of bat populations are declining globally and bats are often excluded from studies on tropical forest mammals (Welch and Beaulieu 2018). According to the IUCN (2022), almost 1,000 bat species need conservation action or research attention, and much less is known about the population status of bat species compared to other mammals and birds (Frick, Kingston and Flanders 2020). Southeast Asia is a global hotspot for bat diversity with at least 388 species (Simmons & Cirranello, 2021). However, almost a quarter of Southeast Asia’s bats are predicted to be extirpated by the end of the century due to deforestation (Lane, Kingston and Lee 2006). Land-use change and the associated habitat loss and disturbance are frequently cited as the primary threat to bats in Southeast Asia (Kingston 2013; IUCN 2022).

1.7.2 Southeast Asian bat community structure

Several eco-morphological factors influence a bat species’ tolerance to disturbance, including diet, morphology, foraging style, and roosting ecology (Farneda *et al.* 2015; Sagot and Chaverri 2015; Meyer, Struebig and Willig 2016; Núñez *et al.* 2019). Traits are not independent as they evolve concurrently to maximise a species' suitability to a specific ecological niche. Insectivory is the dominant diet for bats worldwide (Neuweiler 2000) and Kingston (2013) described three main guild types for insectivorous bat species in Southeast Asia depending on where they forage: forest interior insectivores, edge and gap insectivores, and open-space insectivores. Bats that forage stationary/fluttering insects from the forest interior (forest interior insectivores) typically have broad, short wings (low-aspect-ratio and low wing loading; body mass (kg)/wing area (m^2); Norberg, Rayner and Lighthill, 1987; Neuweiler, 2000). This grants them high manoeuvrability in cluttered forest environments. They typically roost in trees or other foliage in small groups or harems (although not exclusively; Kunz, 1982). Within this guild, there are several echolocation call strategies to detect prey that may otherwise be camouflaged by the echoes from background vegetation (Arlettaz, Jones and Racey 2001), but most use echolocation calls characterised as low-intensity, high frequency, with high repetition rates. In Southeast Asia, this includes members of the Rhinolophidae, Hipposideridae, Murininae, and Kerivoulinae (Liu *et al.* 2008; Hughes *et al.* 2011; McArthur and Khan 2021).

Open-space insectivores hawk airborne insects during flight. Their prey is more widely dispersed than insects found within cluttered forest environments, and therefore, they need an increased range of prey detection and energy-efficient flight. Consequently, they have narrow, long wings (high-aspect-ratio and high wing-loading; Norberg, Rayner and Lighthill, 1987; Neuweiler, 2000) and use narrowband, low-frequency, high-intensity calls (Denzinger and Schnitzler 2013). In Southeast Asia, this includes the families Molossidae and Emballonuridae (Kingston, Jones, *et al.* 2003; Hughes *et al.* 2011). Bats that forage along the forest edge or in forest gaps (edge and gap insectivores) need to balance the requirements of foraging in cluttered environments (e.g., the ability to decipher background echoes from prey) against the requirements of navigating more open spaces. To do this, many species have high call plasticity which enables them to adjust their call to meet different environmental requirements. This may involve adjusting parameters such as call bandwidth, peak frequency, or call shape (Kalko and Schnitzler 1993; Schnitzler, Moss and Denzinger 2003). They can also take advantage of anthropogenic resources. For example, using buildings for roosting (e.g., the Black-bearded tomb bat; *Taphozous melanopogon*; Phillipps and Philipps 2016).

However, there are exceptions to these broad feeding guilds. In Southeast Asia, the Malayan slit-faced bat (*Nycteris tragata*) and Lesser false vampire bat (*Megaderma spasma*) are both forest interior insectivores, but they often rely on hearing to detect their prey (Denzinger and Schnitzler 2013). As such, they have enhanced auditory adaptions (e.g., elongated ears) and use echolocation calls more commonly resembling the phyllostomid bats in the American tropics (Hughes *et al.* 2011; Yoh, Syme, *et al.* 2020). These two species have been known to hunt animal prey, such as fish, amphibians, and small mammals (Phillipps and Phillipps 2016). As such, they closely resemble neotropical animalivorous bat species such as the big-eared woolly bat (*Chrotopterus auritus*). Nectarivorous and frugivorous are less reliant on echolocation to locate food resources and instead often use smell and vision (Raghuram *et al.* 2011; Denzinger and Schnitzler 2013). These include members of the families Pteropodidae (exclusively in the palaeotropics) and Phyllostomidae (exclusively in the neotropics).

1.7.3 Limitations of acoustic monitoring in the tropics

In temperate regions, acoustic monitoring is applied extensively in bat research. However, there are substantial difficulties when trying to implement similar initiatives in many tropical countries; both in the field and when processing these data (Meyer 2015; Meyer, Struebig and Willig 2016; Fisher-Phelps *et al.* 2017). Arguably the greatest limitation to using acoustic monitoring is the lack of call libraries and the taxonomic uncertainty for bat species in many tropical regions (Russo and Voigt 2016; Gibb *et al.* 2019). Our understanding of bat echolocation is not geographically homogenous and reference calls for many tropical bat species' calls are not well documented, particularly across Asia and Central Africa (Walters *et al.* 2013).

The initial costs of establishing a local reference call library are high (Kershenbaum *et al.* 2016). It requires intensive surveying efforts across numerous sites using multiple complementary methods (including live capture to confirm species identification in hand). Inherently areas of high species diversity also require more survey effort to effectively sample the whole community than areas with lower diversity. This is compounded by the general practical difficulties of conducting fieldwork in remote areas which may be hard to access (e.g., no transport infrastructure), politically challenging, or hazardous. Like much ecological research, bat studies in Southeast Asia are spatially biased toward protected areas that often have greater accessibility for research (Fisher-Phelps *et al.* 2017). Therefore, data demonstrating the distribution and habitat use of species more broadly is limited. Even when reference calls and spatial data have been collected, there are considerable differences between the in-country and regional infrastructure for disseminating such data for others to use (Fisher-Phelps *et al.* 2017; Stephenson *et al.* 2017).

There are two additional shortfalls when using acoustic monitoring to monitor bats: overlap in call characteristics and the time required to process large volumes of acoustic data (Russo, Ancillotto and Jones 2018; Gibb *et al.* 2019; Russo *et al.* 2021). Storing and processing the datasets produced by acoustic monitoring has formidable logistical and analytical challenges (Gibb *et al.* 2019). Acoustic datasets often approach the scale of big data (multi-gigabyte to petabyte) and recording in ultrasound (full spectrum with sampling rates often >200 kHz) creates additional storage requirements (Gibb *et al.* 2019; Sugai *et al.* 2019). Therefore, the time and expertise required to manually process such large datasets can be a significant barrier for those wanting to use acoustic monitoring (Browning *et al.* 2017). Additionally, identifying bat calls can be more difficult compared to other taxa. The calls used to monitor most taxonomic groups are calls used for communication. These can be acoustically simple (e.g., amphibians) or complex (e.g., bird song), but their structure must be stable as they provide signals to conspecifics (e.g., territorial calls, mating calls). Therefore, such calls should exhibit limited variability between individuals. In contrast, bat echolocation calls are primarily used to interpret their environment as a form of "self-communication" (Simmons et al., 1977). Therefore, echolocating bats adjust their call structure in response to different habitat and behavioural requirements (Kalko and Schnitzler 1993). Moreover, as calls are not primarily for intraspecific communication, overlap in call structure in species-rich assemblages is common which can limit our capacity to differentiate between species. Interspecific variability is also subject to phylogenetic constraints, and relationships with morphology (e.g., body mass is negatively correlated with call frequency; Jones, 1999), and species can exhibit geographical variability in call frequency across their range, between sexes, or age classes (Pham *et al.* 2021; Russo, Ancillotto and Jones 2018; Walters *et al.* 2013). Therefore, bats exhibit individual and species-specific call plasticity whilst interspecific variation remains constrained in complex systems. This can limit our ability to distinguish between species calls, even between common species in well-studied regions if there is substantial call overlap (Montauban et al., 2021). As such, these limitations have hampered the application of acoustic monitoring for bats in the tropics and their use as bioindicators in these regions (Russo *et al.* 2021).

1.8 Thesis structure

Our understanding of how bats respond to land-use change has been limited to species that we can capture in the forest understory using harp traps or mist nets. There remains very limited information regarding how the vast majority of bats in Southeast Asia are responding to land-cover change. Therefore, this thesis used acoustic monitoring as a complementary method to assess how the bat community is responding to land-use change in Southeast Asia. Specifically, the research assessed the effectiveness of different landscape management strategies for protecting bats in anthropogenic landscapes. This includes the effectiveness of protective zoning (**Chapter 2**), retaining riparian buffers in plantations (**Chapter 3**), and the impacts of logging on bat communities **(Chapter 5**). To do this, the thesis focused on case studies from two sites in Sabah Borneo (Figure 1.2): Crocker Range Biosphere Reserve (**Chapter 2**) and the Stability of Altered Forest Ecosystems Project (**Chapters 3** and **5**), as well as incorporating acoustic data from additional sites across Borneo (**Chapter 4**).

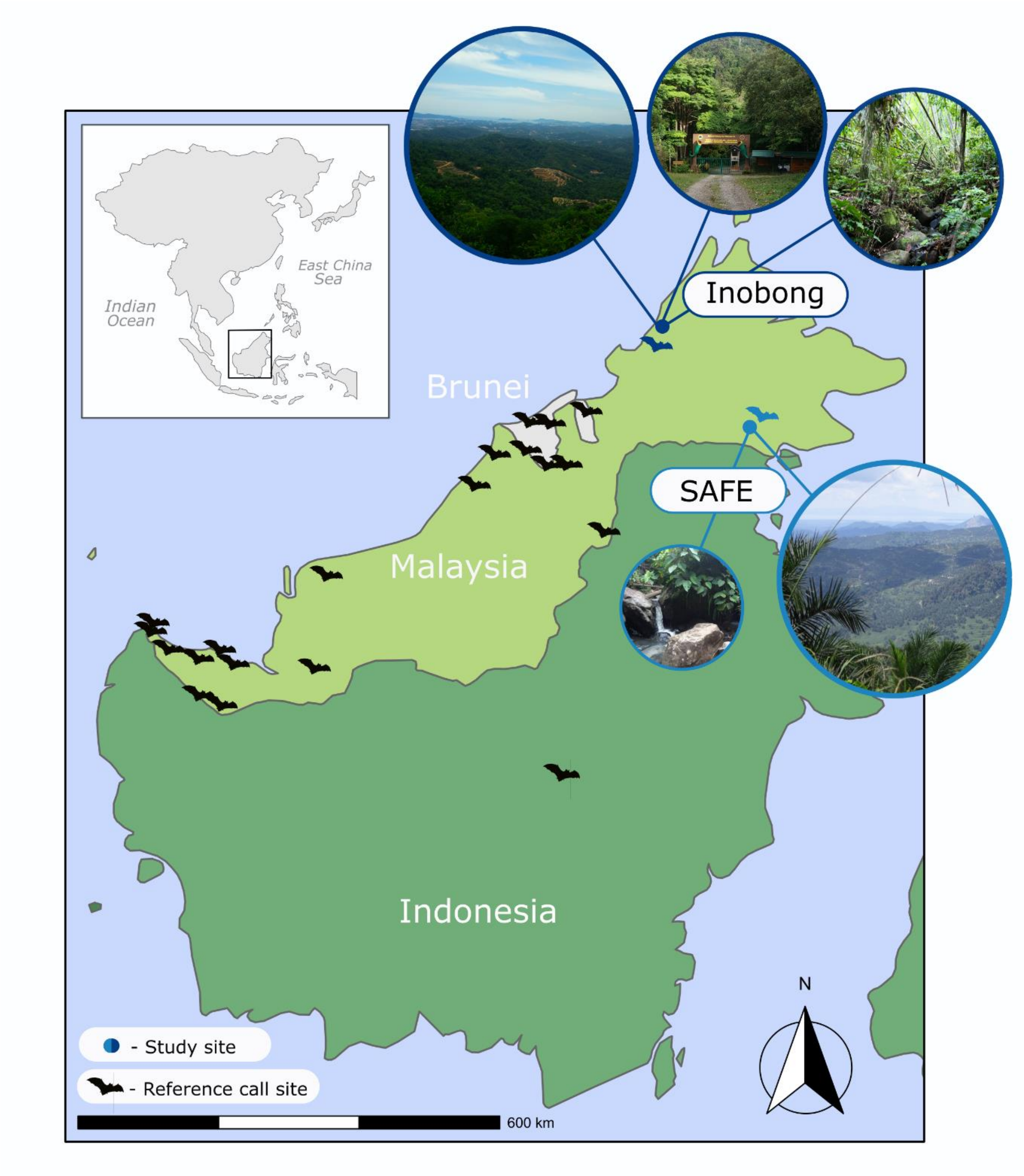


Figure 1.2 Map of study locations included in this thesis

The location of the two main study sites in Sabah in relation to other states - Crocker Range Biosphere Reserve (Chapter 2) and the Stability of Altered Forest Ecosystems Project (SAFE; Chapters 3 and 5) - and the sampling locations for the reference calls used in Chapter 4.

The thesis is structured into three parts. The first investigates how bats respond to land-use change using acoustic monitoring techniques, but without access to automated identification tools (**Chapters 2** and **3**).

Chapter 2 assesses the effectiveness of land-use management strategies for protecting bat activity and diversity in the Crocker Range Biosphere Reserve, Sabah, using both live-capture and acoustic methods. In this chapter, principal component analysis is used to assess differences in acoustic composition between several conservation zones in the reserve, rather than identifying calls into biologically meaningful groups. This chapter has been published as the following:

> Yoh, N., *et al.* (2020). Bat Ensembles Differ in Response to Use Zones in A Tropical Biosphere Reserve. *Diversity, 12*(2). doi: 10.3390/d12020060

Chapter 3 uses an acoustic monitoring dataset from a landscape in the Tawau district, Sabah, to help determine the value of protected riparian buffers for bats in agricultural areas. In this chapter, bat calls are manually identified to call type or species enabling a more refined analysis of bat activity between protected areas and plantations. This chapter has been published as the following:

> Yoh*, N., Mullin*, K., *et al.* (2020). Riparian reserves promote insectivorous bat activity in oil palm dominated landscapes. *Frontiers in Forests and Global Change, 3*. doi: 10.3389/ffgc.2020.00073 *contributed equally

The second part (**Chapter 4**), outlines a machine learning pipeline to help analyse large acoustic datasets in Southeast Asia, such as those provided in **Chapters 2** and **3**. The application of this pipeline is then demonstrated by developing a semi-automated tool for classifying bat calls in Borneo (**Chapter 4**). This chapter has been published as the following:

> Yoh, N., *et al*. (2022). A machine learning framework to classify Southeast Asian echolocating bats. Ecological Indicators, 136. doi: 10.1016/j.ecolind.2022.108696

Finally, the third section of the thesis (**Chapter 5**) reexamines how bats respond to land-use change with the additional aid of the semi-automated identification tool. This research investigates how echolocating bats are affected by habitat disturbance over a land-use gradient from old-growth forest through to logged forests and tree plantations. The chapter is based on another acoustic monitoring campaign at the Stability of Altered Forest Ecosystems Project in Tawau district, Sabah. This chapter is currently under peer review with the Journal of Applied Ecology.

Each of the chapters was written as independent studies and **Chapters 2-4** have been subject to peer review prior to publication. The requirements of reviewers in this process led to differences in the naming conventions between different call groups. These are outlined in Table 1.1.

Table 1.1 Naming conventions for each of the Borneo bat call groups defined in the thesis

Naming conventions for each of the Borneo bat call groups defined in the thesis, their corresponding foraging guild, and their assigned group as described by the universal classifier presented in Roemer, Julien and Bas (2021).

Family	**Genus/species**	**Foraging Guild**	**Chap. 3**	**Chap. 4**	**Chap. 5**	**Roemer *et al.* (2021)**
Emballonuridae						
	Taphozous spp.	Open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Saccolaimus saccolaimus	Open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Emballonura spp.	Forest edges	FMqCF	qCF	*Emballonura alecto/monticola*	QCF-FMd
Nycteridae						
	Nycteris tragata	Forest interior	MHFM	FM	Whispering bats	FMd
Megadermatidae						
	Megaderma spasma	Forest interior	MHFM	FM	Whispering bats	FMd
Rhinolophidae						
	Rhinolophus spp.	Forest interior/edge	CF	CF	Rhinolophoidea bats	FMu-CF-FMd
Hipposideridae						
	Hipposideros spp.	Forest interior	CF	CF	Rhinolophoidea bats	CF-FMd
	Coelops spp.	Forest interior	CF	CF	Rhinolophoidea bats	CF-FMd
Miniopteridae						
	Miniopterus spp.	Open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
Vespertilionidae						
Kerivoulinae						
	Kerivoula spp.	Forest interior	FMb	FM	Whispering bats	FMd
	Phoniscus spp.	Forest interior	FMb	FM	Whispering bats	FMd
Murininae						
	Murina spp.	Forest interior	FMb	FM	Whispering bats	FMd
	Harpiocephalus harpia	Forest interior	FMb	FM	Whispering bats	FMd

Vespertilioninae						
	Pipistrellus spp.	Edge/open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Falsistrellus spp.	Edge/open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Arielulus spp.	Edge/open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Hypsugo kitcheneri	Edge/open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Tylonycteris spp.	Forest edges	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Glischropus tylopus	Forest edges	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Philetor brachypterus	Edge/open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Scotophilus collinus	Edge/open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Hesperoptenus spp.	Edge/open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Myotis spp.	Forest edges	FMb	FMqCF	Edge/open space hawking bats	FMd-QCF
Molossidae						
	Mops mops	Open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF
	Chaerephon spp.	Open space	FMqCF	FMqCF	Edge/open space hawking bats	FMd-QCF

Acronyms: constant frequency (CF), frequency modulating quasi-constant frequency (FMqCF), quasi-constant frequency (qCF), broad-band frequency modulated (FMb), frequency modulated (FM), multi-harmonic frequency modulated (MHFM), frequency modulated upwards sweeping (FMu), and frequency modulated downwards sweeping (FMd).

Chapter 2 Bat Ensembles Differ in Response to Use Zones in a Tropical Biosphere Reserve

Natalie Yoh[1], Isham Azhar[2,3,4,†], Katheryn V. Fitzgerald[2,†], Rieka Yu[5,†], Tenaja Smith-Butler [2], Azniza Mahyudin [6,7] and Tigga Kingston [2,3,*]

1. Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, Canterbury, Kent, CT2 7NR, UK
2. Department of Biological Sciences, Texas Tech University, Lubbock, TX 79409, USA
3. Southeast Asian Bat Conservation Research Unit, Lubbock, TX 79409, USA
4. Institute of Borneo Studies, University College Sabah Foundation, Sembulan, 88100 Kota Kinabalu, Malaysia
5. Department of Biology, University of Missouri, St. Louis, MO 63121, USA
6. Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Malaysia
7. Biotechnology Research Institute, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Malaysia

*Correspondence: Tigga.Kingston@ttu.edu; Tel.: +1-806-834-2594

[†] These authors contributed equally to the work.

This chapter is based on the following publication:

Yoh, N., *et al.* (2020). Bat Ensembles Differ in Response to Use Zones in A Tropical Biosphere Reserve. *Diversity, 12*(2). doi: 10.3390/d12020060

Author Contributions: Conceptualization, T.K., N.Y., K.V.F., I.A., and A.M.; Methodology, T.K., N.Y., A.M., and I.A.; Software, N.Y.; Validation, T.K., N.Y., and R.Y.; Formal analysis, N.Y. and R.Y.; Investigation, T.K., N.Y., K.V.F., I.A., R.Y., A.M., and T.S.-B.; Resources, T.K. and A.M.; Data Curation N.Y., K.V.F., T.K., R.Y., and T.S.-B.; Writing—Original Draft Preparation, T.K., N.Y., R.Y., K.V.F., and I.A.; Writing—Review and Editing, T.K.; Visualization, N.Y., T.S.-B., R.Y., and K.V.F.; Supervision, T.K.; Project Administration—T.K., A.M., and I.A.; Funding Acquisition, T.K. and A.M.

2.1 Abstract

Biosphere reserves, designated under The United Nations Education, Scientific and Cultural Organization’s (UNESCO) Man and Biosphere Programme, aim to sustainably integrate protected areas into the biological and economic landscape around them by buffering strictly protected habitats with zones of limited use. However, the effectiveness of biosphere reserves and the contribution of the different zones of use to protection is poorly known. We assessed the diversity and activity of bats in the Crocker Range Biosphere Reserve (CRBR) in Sabah, Malaysia, using harp traps, mist nets and acoustic surveys in each zone—core, buffer, transition and in agricultural plots outside of the reserve. We captured 30 species, bringing the known bat fauna of CRBR to 50 species, half of Borneo’s bat species. The reserve is therefore an area of high bat diversity compared to other areas within the Heart of Borneo initiative. Species composition and acoustic activity varied among zones and by foraging ensemble, with the core and buffer showing particular importance for conserving forest-dependent insectivorous bats. In contrast, forest-interior species declined within the transition and extralimital agriculture indicating a negative response to land-use change for this ensemble. Frugivorous bats were found in all zones but were the most abundant and most species-rich ensemble within agricultural sites. Although sampling was limited, bat diversity and activity was low in the transition zone compared to other zones, indicating potential for management practices that increase food availability and enhance biodiversity value. We conclude that the inner zones of the CRBR are important for supporting bat diversity, but the value of the transition zone can be improved. Evaluations such as this are important to ensure protected area schemes are functioning effectively. We conclude that collectively UNESCO Biosphere zones can provide valuable habitat for conserving bat diversity in the Southeast Asian tropics.

Keywords: Man and Biosphere; Crocker Range; Chiroptera; bat diversity; land-use change; foraging ensembles

2.2 Introduction

Tropical forests represent ~10% of total land cover yet support approximately two-thirds of global terrestrial diversity (Giam 2017; Raven 1988). Tragically, this diversity is under threat from the conversion of forests to agriculture and pasture, which remains the dominant driver of land-use change in the tropics today (Song *et al.* 2018). This has instigated an increase in the extent and quantity of protected areas in the tropics to conserve remaining biodiversity (Jenkins and Joppa 2009; DeFries *et al.* 2007). However, human activity at the periphery of a protected area can affect its functionality for conserving biodiversity (DeFries *et al.* 2007). This is a growing concern for tropical forest protected areas, which are becoming increasingly isolated within agricultural landscapes (Curran *et al.* 2004). Moreover, protected areas themselves are not void of anthropogenic activity as ~70% of protected areas in the tropics are occupied by human communities (Terborgh and Peres 2002). For these and additional reasons (Hill and Curran 2003), the success of a protected area is considered dependent on its integration into the wider landscape (Laurance *et al.* 2012).

Biosphere reserves promote the integration and management of protected areas in a landscape context by buffering strictly protected habitats with zones of limited use. They are designated by The United Nations Education, Scientific and Cultural Organization (UNESCO), as part of the Man and Biosphere (MAB) Programme, launched in 1971 (Ishwaran, Persic and Tri 2008). The programme currently lists 669 reserves, including 701 sites, across 124 countries (Reed 2019; UNESCO 2017). Biosphere reserves are comprised of three functional zones: the core area(s), the buffer zone, and the transition area, ranging in their accessibility for human land-use (Reed 2019; UNESCO 2014). The core area is strictly protected, limiting human disturbance to an ecosystem. Its primary function is to conserve biodiversity and protect ecosystem services. Limited low-impact human activities, such as non-destructive research and environmental education, are permitted where they complement conservation goals. The buffer zone surrounds or adjoins the core area, providing an intermediary buffer between the core and transition zone. The buffer permits greater human use, including activities such as environmental education, recreation, ecotourism, and applied research, that are compatible with sound ecological practices. At the landscape level, buffers help maintain connectivity for biodiversity components residing in core areas. The primary function of the transition zone is for sustainable development. Development may include agriculture, settlements, and other developments that improve an area’s resources. Stakeholders in such development should include local communities, non-governmental organisations, and management agencies. Activities in this zone must be socioculturally and ecologically sustainable. The functions of these zones are designed to be complementary with goals to maintain anthropogenic, biological and cultural diversity (Reed 2019; UNESCO 2017). Currently, the Southeast Asia Biosphere Reserve Network (SeaBRnet) consists of 35 reserves governing > 30 mn hectares (UNESCO 2019).

With over 1400 species distributed globally, bats make up the second largest mammalian order and provide key ecosystem services as seed dispersers, pollinators, and as agents of pest suppression (Kunz *et al.* 2011). Many economically valued crops, such as durian in Southeast Asia (Sheherazade, Ober and Tsang 2019) and agave in the Americas (Fleming and Holland 2018), are primarily pollinated by bats. Additionally, they can serve as bioindicators of ecosystem health due to their sensitivity to changes in the environment, climate change, water loss, and noise pollution (Jones *et al.* 2009). Unfortunately, bats are proving highly vulnerable to human activities (Voigt and Kingston 2016; Frick, Kingston and Flanders 2020), with 21% of species assessed as threatened (EX, CR, EN, VU) or nearly so (NT) by the IUCN (IUCN 2019) and, globally, four out of five bat populations are declining (Welch and Beaulieu 2018). Bat diversity peaks in tropical regions, particularly tropical rainforests (Kingston, Francis, *et al.* 2003), but these habitats are rapidly being degraded and converted to other land-uses, with negative consequences for bat populations worldwide (Frick, Kingston and Flanders 2020; García-Morales, Badano and Moreno 2013; Meyer, Struebig and Willig 2016). Southeast Asian tropical forests suffered the highest rates of deforestation amongst all tropical regions during the 1990s (Miettinen, Shi and Liew 2011), and only half of Borneo's forest cover remains as a result of agriculture, intensive logging, and palm and pulp plantations (Fitzherbert *et al.* 2008; Struebig *et al.* 2012; Gaveau *et al.* 2013). Bats make up 40% of the island's terrestrial mammal diversity (Payne 2000), and approximately one-third of the 98 species are dependent on large tracts of unmodified forests (Phillipps and Phillipps 2016). Forest conversion to agriculture in Singapore resulted in bat extinction rates estimated to range from 33–72%, depending on the group, and it is estimated that if current deforestation rates persist across Southeast Asia, as many as 40% of species may be lost from the region by the end of the century (Lane, Kingston and Lee 2006).

Bats do not respond uniformly to land-use change and their response varies between foraging ensembles and species. Old-World bats can be divided into four broad foraging ensembles. Plant-visiting bats, members of the family Pteropodidae, are primarily frugivorous, with a few species specialised for nectarivory. All of the other eight families found on Borneo are insectivorous and fall into one of three foraging ensembles, based on where they forage relative to background vegetation: (1) "open space bats" forage in the open spaces above forests or cleared/agricultural land; (2) "edge or gap" bats hunt insects along the edge of vegetation stands or in small clearings within forest or over water; and (3) "forest interior bats" are able to detect and pursue prey within the vegetatively cluttered habitats of the forest (Kingston 2013). Each ensemble is associated with a combination of wing and echolocation traits that maximise performance in their preferred foraging habitat, but greatly constrain performance in other habitats (Kingston 2013; Norberg, Rayner and Lighthill 1987; Schnitzler and Denzinger 2011; Schmieder *et al.* 2012; Senawi and Kingston 2019), therefore resulting in differences amongst responses to land-use change.

Forest interior bats are proving particularly susceptible to forest loss and degradation (Kingston 2013; Norberg, Rayner and Lighthill 1987; Schnitzler and Denzinger 2011; Schmieder *et al.* 2012; Senawi and Kingston 2019; Struebig *et al.* 2008). Foraging in the vegetatively complex forest interior requires slow, manoeuvrable flight (Senawi and Kingston 2019) and clutter-tolerant echolocation, but these adaptations compromise performance in the more open habitats that result from land-use change. Slow flight can be energetically costly, and clutter-tolerant echolocation often has a limited range. Moreover, many species rely on roosts that are only available in numbers able to support populations in relatively unmodified forest (e.g., hollows in large standing or fallen trees). Reduction in native forest cover has particularly impacted bats found in the families Hipposideridae, Rhinolophidae, Megadermatidae, Nycteridae, as well as the Kerivoulinae and Murininae subfamilies of the Vespertilionidae (Kingston, Francis, *et al.* 2003; Kingston 2013; Huang *et al.* 2014).

Although the MAB Programme has been in operation for nearly 50 years, the effectiveness of biosphere reserves in protecting bat diversity is poorly documented, particularly in the wet tropics where bat diversity is greatest. Studies in tropical dry forests and grasslands of South Africa and Mexico suggest that, although core zones protect more species, buffers support a significant proportion of core diversity (Taylor *et al.* 2013; Linden *et al.* 2014; López-González *et al.* 2015; Weier *et al.* 2017), but forests in these habitats are less structurally complex and support fewer species than tropical rainforest. Crocker Range Biosphere Reserve (CRBR) was established as a UNESCO site under the MAB Programme in Sabah, Malaysia Borneo in 2014. It consists mainly of hill and lower montane tropical forests, and, as home to approximately 101 mammal, 259 bird, 47 reptile, 63 amphibian, and over 300 plant species, the reserve and its surroundings are considered a refuge for much of Borneo's biodiversity (UNESCO 2014; Yasuma *et al.* 2003; Das 2006), including bats. The last inventory of bat diversity in Crocker Range was conducted in 1999 (Yasuma *et al.* 2003; Tuen, Rahman and Salleh 2002), prior to the Range's entry into the MAB Programme. Forty-one species were reported from the survey and the prior literature (Tuen, Rahman and Salleh 2002), representing at least 40% of Borneo's total bat diversity and indicating the importance of CRBR to regional bat conservation.

The goal of this study was to determine the effectiveness of CRBR in protecting bat diversity. Specifically, we evaluated how diversity and activity vary among the biosphere zones and foraging ensembles, and contrast with neighbouring extralimital agriculture. Using a mix of harp traps, mist nets and acoustic surveys, we surveyed bats in each zone and the surrounding extralimital agriculture. We added eight confirmed species to the known bat fauna for CRBR and conclude that CRBR is an effective reserve, protecting half of Borneo’s bat diversity. Species composition and activity varied among zones, with the core and buffer particularly important for conserving forest-interior insectivores.

2.3 Materials and methods

2.3.1 Study Sites

Bat surveys were conducted in the CRBR, also known as Crocker Range Park, from 19–22 July 2017, 24 June–26 July 2018 and 14 June–18 July 2019. CRBR comprises more than 350,000 hectares of forested area, ranging from lowland to montane forest up to 2350m a.s.l. (Kitayama 1992). The reserve is characterised by three different protection zones: the core zone, the buffer zone, and the transition zone (Reed 2019) (Figure 2.1). The buffer and transition zone of CRBR is occupied by approximately 400 communities surrounded by natural hills with lower montane forest, secondary forest and mixed agriculture such as oil palm, rubber, paddy field, and commercial crops (UNESCO 2017). Live trapping and acoustic surveys were conducted at 17 sites covering different protection zones within the reserve, and eight sites within the extralimital zone on the northeast edge of the reserve.

Survey sites in the northeast corner of CRBR were accessed from Mahua Substation and Malungung Control Post in 2017 and 2018, and through Inobong Substation in the northwest edge of the reserve in 2019. Most of the natural vegetation where bats were sampled from the three distinct protection zones is hill dipterocarp forest, occurring between 500m a.s.l. to about 1000m a.s.l. The average temperature is between 24–27◦C, and the average annual precipitation is between 1880–2621mm (Weatherbase 2019). The extralimital zones are mostly agriculture plots, with associated commercial crops including banana (*Musa* spp.), rambutan (*Nephelium lappaceum*), soursop (*Annona* spp.), mango (*Mangifera* spp.), rubber trees (*Hevea brasiliensis*), oil palm (*Elaeis guineensis*), and durian (*Durio* spp.). Full details of each site are given in Supplementary Table S2.1.

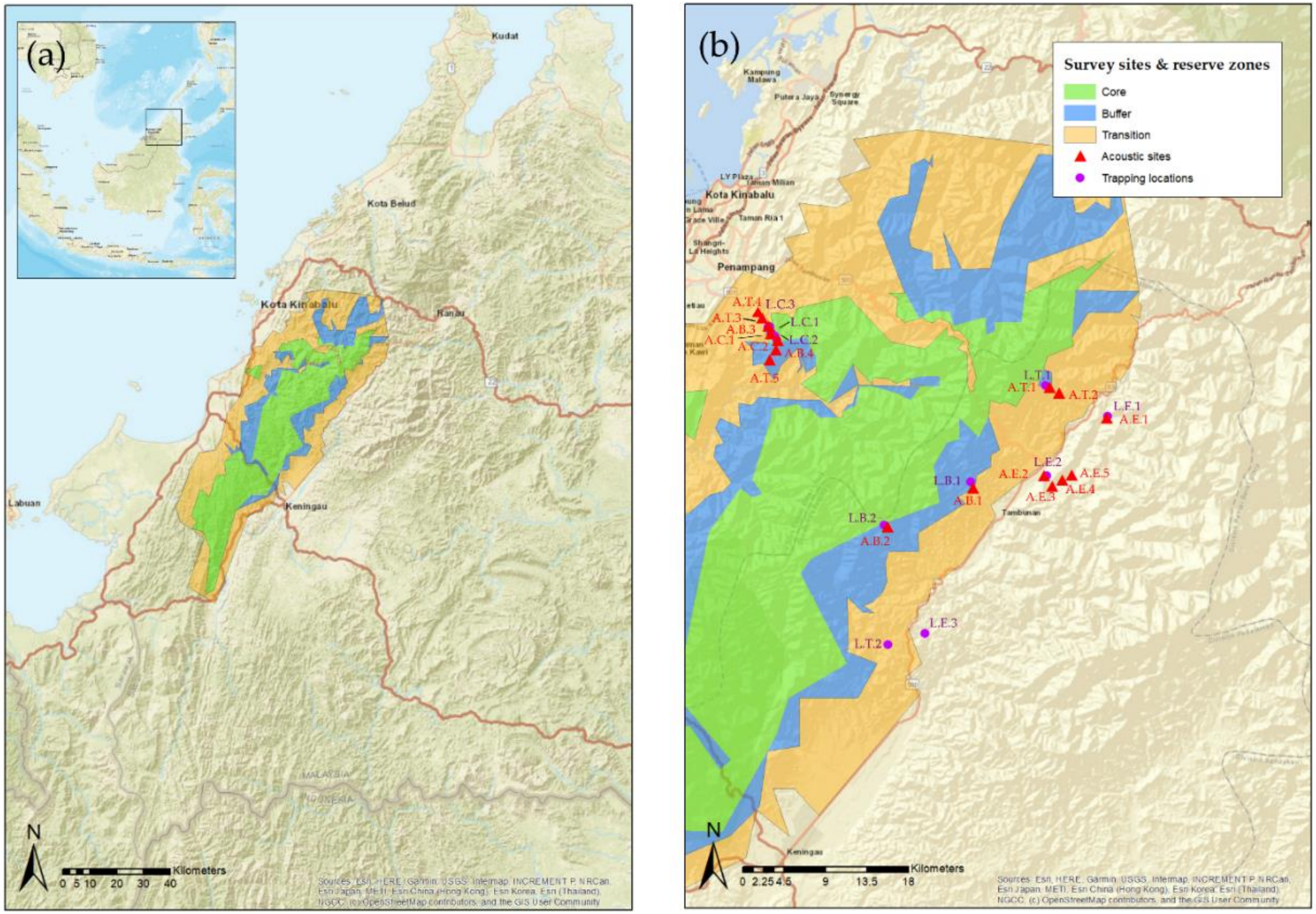


Figure 2.1 Map of study area in CRBR, Sabah, Malaysia

a) Location of the CRBR showing the distribution of the MAB zones green-core, blue-buffer, yellow-transition. Inset illustrates the location of Sabah on Borneo. b) Area of sampling sites, containing trapping and acoustic transect locations L – live trapping sites, A – acoustic sites, C – Core, B- Buffer, T- transition, E- extralimital agriculture.

2.3.2 Species Capture and Identification

Bats were captured at each site using four-bank harp traps (Francis 1989) positioned across trails or small streams. Ground polyester mist nets (9m; 36mm mesh size; four shelves) and double stacked nets (eight shelves) were positioned across established trails, at forest edges, and across small ponds and streams. Mist nets were checked every 15-30 minutes from 18:00 to 23:00 hours, or until bat activity declined. Meanwhile, harp traps were checked several times in the evening, and once after times of peak activity in the morning. Bat sampling was avoided during periods of heavy rain. Hand nets were used occasionally ad-hoc to capture low flying bats in open areas and bats at roosts.

Captured bats were identified following Phillipps and Phillipps (2016), Kingston, Lim, and Akbar (2006), and Payne, Francis, and Phillipps (1985). Forearm length (mm) and body mass (g) were recorded for each individual. Juveniles were distinguished from adults by examining the epiphyseal fusion of phalanges (Anthony 1988). The reproductive status of females was determined by examining the mammary glands and the surrounding area, and was categorised as non-reproductive, pregnant, lactating and post-lactating (Anthony 1988). Each captured individual was released at the point of capture within six hours. Occasional voucher specimens were collected to confirm the identification of uncertain records and were deposited at the Institute of Tropical Biology and Conservation, Universiti Malaysia Sabah.

The procedures followed animal care and use guidelines of Texas Tech University's Institutional Animal Care and Use Committee and was approved as protocol 17026-07 category C. This project was also approved Sabah Biodiversity Council and conducted under access license JKM/MBS .1000-2/2 JLD.6 (53).

2.3.3 Acoustic Sampling and Analysis

Anabat Walkabout Bat Detectors v1.2 (Broken-Brow and Thompson 2017) with model-specific microphone adapters were used to record bat activity along 2-km transects. The adapters improve directionality and the quality of recordings in tropical environments where there are high volumes of insect noise and where bat species produce echolocation calls at high frequencies (< 150kHz). Transects consisted of six sample points situated 200m apart to ensure independence. Where possible, acoustic transects intersected live-capture sites however this was not possible for all locations (Figure 2.1). The detectors were programmed to record continuously for ten minutes at each point; each of which was sampled twice per survey. Detectors recorded activity in real-time with a full spectrum resolution of 16-bit and a sampling rate of 500kHz. Files were recorded in 15-second sections. Where possible surveys were conducted in the evening, commencing at 18:30 and concluding before 21:00. However, if this was not possible due to weather restrictions dawn surveys were conducted in the extralimital agriculture (43% surveys), commencing at 05:00.

For sites A.E.1:5, A.T.1:2, and A.B.1:2, two surveyors simultaneously started at sample point one and sample point six and recorded along the transect in alternate directions such that each point was surveyed twice per night. These sites were surveyed once each. Surveys at sites A.T.3:5, A.B.3:4, A.C.1:2 commenced from one starting point (either sample point one or sample point six) and surveyors retraced their steps for the second sample of the night. Consequently, the recording period for these sites was longer. Transects A.T.3:5, A.B.3, and A.C.1 were repeated three times, and A.C.2 and A.B.4 twice for replication. This equates to five surveys in extralimital agriculture, fourteen in the transition zone, seven in the buffer zone, and five in the core zone. As there was a significant relationship between bat activity and time of surveying (Supplementary Figure S2.1, Table S2.2), only recordings from the first sample of each point were used to compare activity between the different zones to minimise temporal bias.

The number of pulses (e.g., individual vocalisations) was used as a sample unit for calculating bat activity. Fifteen-second recordings were filtered using Kaleidoscope V5.1 (Wildlife Acoustics Inc, USA) for those with two or more search-phase pulses. Once filtered, Anabat Insight (Broken-Brow 2018) was used to automatically extract call characteristics and calculate total activity per sample point, between 15kHz–500kHz with a smoothness five. Call characteristics included: characteristic frequency (Fc), characteristic slope (Sc), duration (Dur), maximum frequency (Fmax), minimum frequency (Fmin), mean frequency (Fmean), time between calls (TBC), frequency of knee (Fk), time from the start of the call to the knee (Tk), average call smoothness (Quality), the slope of the first five points in a pulse (S1), and time from start to the characteristic component (Tc). These data were then cleaned of any pulses with a peak frequency (Fmean) < 20kHz to remove false positives/noise, which removed 44,657 records. Species identification was not conducted for this study and therefore activity represents total bat activity for all species. As bats produce a series of pulses in quick succession when approaching their prey, feeding activity may artificially inflate activity.

2.3.4 Statistical Analysis

To compare acoustic activity, activity per transect was calculated as the sum of activity per corresponding sample point per night divided by sampling effort (minutes of recording). A Kruskal–Wallis H test was then conducted to determine if there was a difference in bat activity between the different zones and the extralimital agriculture. Pairwise comparisons were also conducted between the mean bat activity of each zone and extralimital agriculture using nonparametric, bootstrapped 95% confidence intervals. Principal component analysis (PCA) in R (packages "stats v3.6.1" and "factoextra v1.0.5") (R Core Team 2017) was also employed to investigate the similarity between the call structure of the acoustic activity observed between the different zones. In order to assess survey completeness and compare species richness of sites, rarefaction analyses for each site were done based on sampling effort in iNEXT v2.0.19 (Chao *et al.* 2014; Hsieh, Ma and Chao 2016). We performed separate analyses for individuals captured in harp traps and mist nets. Species richness was rarefied by sample units and then rescaled to individuals. For harp traps, sample units were one harp trap night and for mist nets, sample units were one mist net meter/hour.

2.4 Results

2.4.1 Bat Assemblage

A total of 260 individuals of 30 species were captured belonging to 5 families (Tables 2.1 and 2.2, Figure 2.2). An additional species, *Hipposideros doriae*/*Coelops robinsoni*, was identified acoustically, bringing the total to 31 species. Crocker Range was last surveyed for bats in 1999 (Tuen, Rahman and Salleh 2002). That preliminary study used mist-nets, harp traps, and museum inventories to document 41 bat species from the area. Comparing species inventories, our study recorded nine new species in CRBR (*Cynopterus minutus, Rousettus amplexicaudatus, Hipposideros ridleyi, Hypsugo* cf *kitcheneri, Kerivoula intermedia, Phoniscus atrox, Tylonycteris pachypus, Nycteris tragata and H. doriae/C. robinsoni),* bringing the new total to 50 species for CRBR.

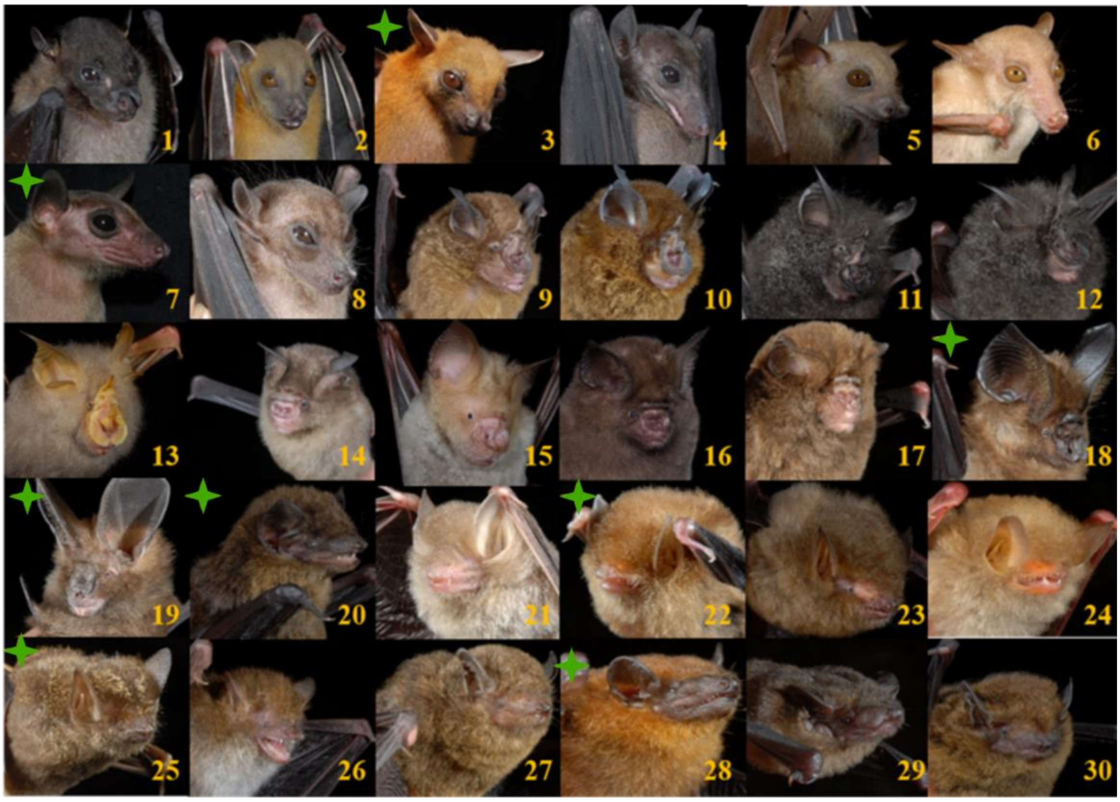


Figure 2.2 Photographic portraits of the 30 species captured in CRBR

New locality records in CRBR are indicated by green star. (1) *Balionycteris maculate*, (2) *Cynopterus brachyotis*, (3) *Cynopterus minutus*, (4) *Eonycteris major*, (5) *Megaerops ecuadatus*, (6) *Macroglossus minimus*, (7) *Rousettus amplexicaudatus*, (8) *Penthetor lucasi*, (9) *Rhinolophus acuminatus*, (10) *Rhinolophus borneensis*, (11) *Rhinolophus luctus*, (12) *Rhinolophus sedulus*, (13) *Rhinolophus trifoliatus*, (14) *Hipposideros cervinus*, (15) *Hipposideros diadema*, (16) *Hipposideros dyacorum*, (17) *Hipposideros galeritus*, (18) *Hipposideros ridleyi*, (19) *Nycteris tragata*, (20) *Hypsugo* cf *kitcheneri*, (21) *Kerivoula hardwickii*, (22) *Kerivoula intermedia*, (23) *Kerivoula papiliosa*, (24) *Kerivoula pellucida*, (25) *Phoniscus atrox*, (26) *Murina suilla*, (27) *Glyschropus tylopus*, (28) *Tylonycterus pachypus*, (29) *Tylonycterus robustula*, (30) *Pipistrellus javanicus.*

Table 2.1 Number of individual captures of bat species across the four zones of CRBR and the surrounding agriculture

Number of individual captures (from harp traps; HT, mist nets; MN and stacked nets; SN) of 30 bat species grouped into four zones and surrounding agriculture (see Figure 2.1). Effort is determined by the number of hours actively trapping. Mist nets and stack net hours are calculated by multiplying the net meter length with active netting hours (m x hr). * Indicates new species record in CRBR. ** Additional captures made with hand nets, X – identified from acoustic transect. Foraging Ensemble (FE) is coded as P – plant visiting, O – open-space insectivore, E – edge or gap insectivore, F – forest interior insectivore.

	FE	**Core**			**Buffer**			**Transition**			**Extralimital**			**Grand Total**
		HT	MN/SN	Total captures	HT	MN/SN	Total captures	HT	MN/SN	Total captures	HT	MN/SN	Total captures	Captures
Effort (Hours)		699	1,569		248	1,307		100	258		332	1,291		
Individual trap sites		56	26		23	19		8	10		20	27		
Pteropodidae														
Balionycteris maculata	P	1	3	**4**		2	**2**			**0**		3	**3**	**9**
Cynopterus brachyotis	P		10	**10**		8	**8**		7	**7**		3	**3**	**28**
**Cynopterus minutus*	P		7	**7**			**0**			**0**		5	**5**	**12**
Eonycteris major	P			**0**		1	**1**			**0**		2	**2**	**3**
Megaerops ecuadatus	P	2	2	**4**			**0**			**0**			**0**	**4**
Macroglossus minimus	P	1	1	**2**		8	**8**			**0**		29	**29**	**39**
Penthetor lucasi	P			**0**		1	**1**			**0**		1	**1**	**2**
**Rousettus*	P			**0**			**0**			**0**		1	**1**	**1**

amplexicaudatus														
Hipposideridae														
Coelops robinsoni	F			**X**			**0**			**0**				
Hipposideros cervinus	F	9		**9**			**0**			**0**			**0**	**9**
Hipposideros diadema	F/E			**0**		1	**1**						**0**	**1**
Hipposideros dyacorum	F	1		**1**			**0**			**0**			**0**	**1**
Hipposideros galeritus	F			**0**	1		**1**			**0**			**0**	**1**
**Hipposideros ridleyi*	F			**0**	2		**2**			**0**			**0**	**2**
Rhinolophidae														
Rhinolophus acuminatus	F			**0**			**0**	1		**1**			**0**	**1**
Rhinolophus borneensis	F			**0**			**0**	4		**4**			**0**	**4**
Rhinolophus luctus	F/E			**0**			**0**			**0**		1	**1**	**1**
Rhinolophus sedulus	F	1		**1**	1		**1**			**0**		1	**1**	**3**
Rhinolophus trifoliatus	F	9	1	**10**	1		**1**			**0**		1	**1**	**12**
Vespertilionidae														
Glischropus tylopus	E	6	3	**11****		24	**24**		3	**3**		4	**4**	**42**
**Hypsugo cf kitcheneri*	E			**0**			**0**			**1****			**0**	**1**
**Kerivoula intermedia*	F			**0**	4		**4**			**0**			**0**	**0**
Kerivoula papillosa	F			**0**	1		**1**	3		**3**			**0**	**4**
Kerivoula pellucida	F			**0**			**0**			**0**	1		**1**	**1**
Kerivoula hardwickii	F	15		**15**	1		**1**			**0**	2		**2**	**18**

Murina suilla	F			**0**	1		**1**			**0**			**0**	**1**
Pipistrellus javanicus	E			**0**		1	**1**			**0**			**0**	**1**
**Phoniscus atrox*	F	1		**1**			**0**			**0**			**0**	**1**
**Tylonycteris pachypus*	E		6	**6**		2	**2**			**0**			**0**	**8**
Tylonycteris robustula	E	10	10	**20**	2	23	**25**			**0**			**0**	**45**
Nycteridae														
**Nycteris tragata*	F	1		**1**			**0**			**0**			**0**	**1**
Total Captures		59	43	**102**	14	67	**85**	9	3	**19**	3	51	**54**	**260**
Species Richness		12	9	**16**	9	9	**18**	3	2	**6**	2	11	**13**	**31**

Table 2.2 Morphological data for 30 bat species from CRBR

Morphological data for 30 bat species from CRBR. (Mean ± SD); Smallest value in data set, Min.; Largest value in data set, Max.; M, male; F, female, n = number of individuals.

	Measurements							
	Forearm length (mm)				Body mass (g)			
Taxa	**M**		**F**		**M**		**F**	
	Mean ± SD	No.	Mean ± SD	No.	Mean ± SD	No.	Mean ± SD	No.
Pteropodidae								
Balionycteris maculata	-	0	41.59 ± 2.28 (36.9 – 44.6)	9	-	0	14.89 ± 3.16 (9.25 – 19)	9
Cynopterus brachyotis	58.79 ± 3.14 (54.3 – 66.1)	12	59.07 ± 2.49 (53.6 – 64.7)	15	28.93 ± 4.36 (20 – 38.1)	13	31.67 ± 5.92 (25 – 44)	15
Cynopterus minutus	57. 35 ± 0.56 (56.7 – 58.1)	6	56.73 ± 1.34 (54. 1 – 57.7)	6	24.33 ± 2.07 (22 – 28)	6	29.33 ± 6.35 (19 – 37)	6
Eonycteris major	84.8 ± 3.25 (82.5 – 87.1)	2	76.5	1	120	1	77	1
Megaerops ecuadatus	53.4	1	52.23 ± 1.05 (51.2 – 53.3)	3	23	1	25.33 ± 4.16 (22 – 30)	3
Macroglossus minimus	40.51 ± 1.63 (37.9 – 44.3)	15	40.56 ± 1.36 (38.1 – 43.7)	20	16.99 ± 2.99 (10 –20.5)	15	17.49 ± 3.31 (12.25 – 24)	19

Penthetor lucasi	60.2	1	60.5	1	36	1	42	1
Rousettus amplexicaudatus	-	0	68.4	1	-	0	44	1
Hipposideridae								
Hipposideros cervinus	50.1 ± 1.05 (49.1 – 51.2)	3	51.93 ± 1.19 (50 – 53.7)	6	10.5 ± 0.87 (10 – 11.5)	3	11.16 ± 1.21 (9.5 – 12.5)	6
Hipposideros diadema	-	0	88.7	1	-	0	51	1
Hipposideros dyacorum	42.3	1	-	0	7	1	-	0
Hipposideros galeritus	49.7	1	-	0	9.5	1	49.7	1
Hipposideros ridleyi	49.2	1	48	1	8.25	1	9.25	1
Rhinolophidae								
Rhinolophus acuminatus	-	0	49.4	1	-	0	14.25	1
Rhinolophus borneensis	40.25 ± 0.07 (40.2 – 40.3)	2	43.95 ± 0.5 (43.6 – 44.3)	2	9.5	2	8.88 ± 0.53 (8.5 – 9.25)	2
Rhinolophus luctus	64	1	-	0	29	2	-	0
Rhinolophus sedulus	-	0	46.97 ± 6.7 (42.1 – 54.6)	3	-	0	14.33 ± 5.77 (11 – 21)	3

Rhinolophus trifoliatus	52.61 ± 2.3 (48.7 – 54.9)	7	53.16 ± 1.87 (49.9 – 54.6)	5	14.61 ± 1.98 (12.5– 7.25)	5	16.65 ± 1.90 (14.25 – 19)	7
Nycteridae								
Nycteris tragata	-	0	51.6	1	-	0	17.5	1
Vespertilionidae								
Glischropus tylopus	30.44 ± 0.84 (28.7 – 31.9)	20	30.70 ± 2.2 (22.2 – 33.3)	22	4.44 ± 0.43 (3 – 5)	20	4.46 ± 0.72 (3.4 – 6.5)	22
Hypsugo cf. *kitcheneri*	35	1	-	0	5.75	1	35	1
Kerivoula intermedia	28.75 ± 0.31 (28.4 – 29.1)	4	-	0	2.88 ± 0.14 (2.75 – 3)	4	28.75 ± 0.31 (28.4 – 29.1)	4
Kerivoula papillosa	38.4 ± 8.1 (29.1 – 43.9)	3	37.1	1	8.75 ± 3.46 (4.75–10.75)	3	6	1
Kerivoula pellucida	-	0	31.1	1	5.5	1	-	0
Kerivoula hardwickii	31.95 ± 1.09 (30.1 – 33.5)	13	33.12 ± 1.18 (31.5 – 34.8)	5	3.56 ± 0.41 (2.5 – 4)	12	4.65 ± 0.34 (4.25 – 5)	5
Murina suilla	-	0	31.4	1	-	0	4.75	1
Pipistrellus javanicus	-	0	34.3	1	-	0	6	1
Phoniscus atrox	33.9	1	-	0	4.5	1	-	0

Tylonycteris pachypus	25.36 ± 0.98 (24.3 – 26.4)	5	25.63 ± 0.97 (24.8 – 26.7)	3	3.35 ± 0.42 (2.75 – 3.7)	5	3.33 ± 0.58 (3 – 4)	3
Tylonycteris robustula	28.96 ± 2.41 (25.4 – 38.6)	22	28.19 ± 3.15 (18.3 – 38.3)	23	7.39 ± 1.24 (4.25 – 10)	22	6.80 ± 1.0 (4 – 9.5)	23

More species were recorded in the forested zones of core (16 species) and buffer (18 species) than in the transition (6 species) and extralimital (13 species) zones (Table 2.1). Estimated species richness based on harp traps was similarly greater in the forested zones (Figure 2.3), but rarefaction curves from mist net captures suggest greatest richness in extralimital agriculture, although there were no significant differences among core, buffer or extralimital agriculture (Figure 2.3b). Capture effort was low in the transition zone, but both richness curves appeared to level off at low values (< 4 species).

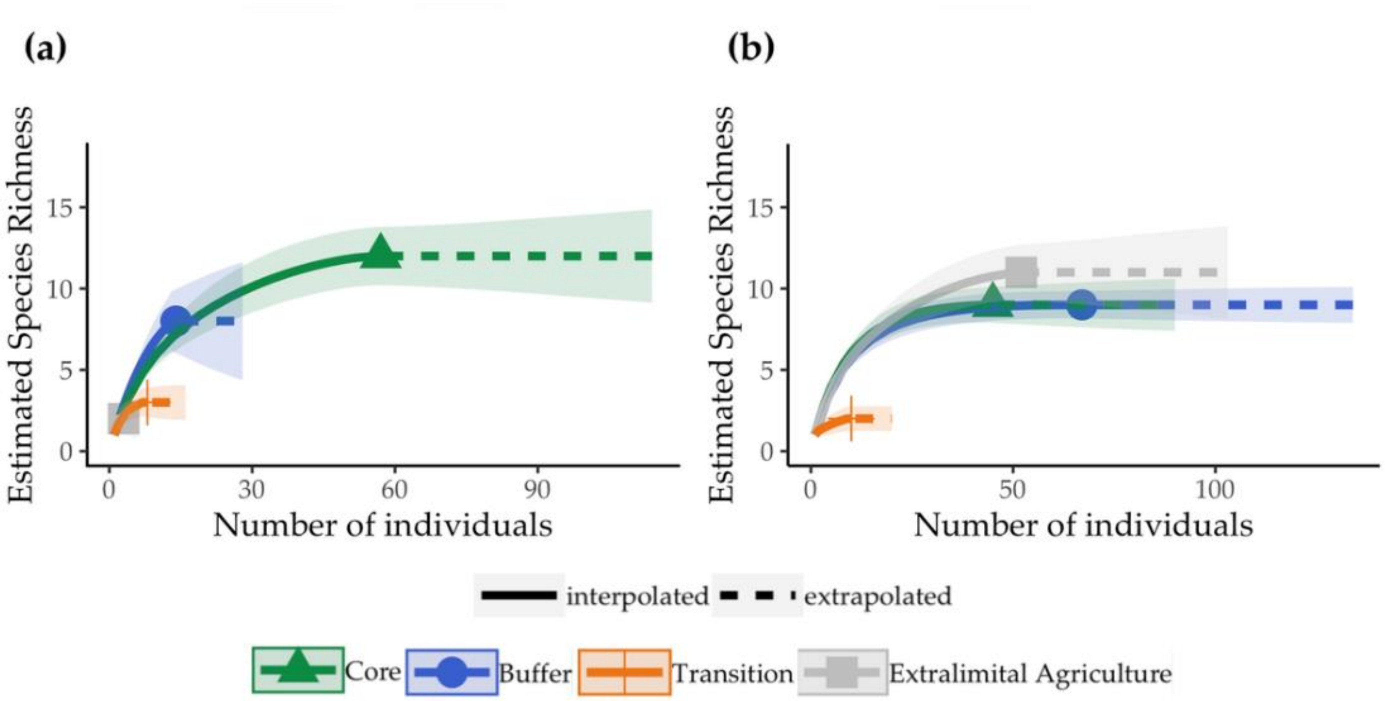


Figure 2.3 Rarefaction curves for species captured in harp traps and mist nets

Species richness was rarefied by (a) harp trap nights or (b) mist net meter/hour and rescaled to individuals for the four zones of the CRBR, Borneo, Malaysia. Colours and symbols indicate different MAB zones. Interpolated lines are the rarefaction curves and extrapolated lines are estimates for double the reference sample size. Shaded areas are the 95% confidence intervals for each curve.

Thirteen of the 23 species captured in core and buffer zones were exclusively captured in those zones. All 5 species of *Hipposideros* were caught exclusively in core and buffer zones. *Cynopterus brachyotis* and *Glischropus tylopus* were found in all four zones. *Rhinolophus trifoliatus, Macroglossus minimus, Rhinolophus sedulous,* and *Kerivoula hardwickii* were found in every zone except for the transition zone. *Rousettus amplexicaudatus, Kerivoula pellucida* and *Rhinolophus luctus* were captured only in the transition zone.

T. robustula and *G. tylopus* were the most abundant species, representing a third of overall captures (n = 87). In the buffer zone, they represented 58% of captures, largely because of the success of a single mist net placed over a small shallow pond (site L.B.2). In the core (L.C.1) mist nets placed near bamboo stands captured large numbers of individuals leaving roosts. *T. robustula* was not captured in transition and extralimital agriculture, and only three individuals of *G. tylopus* were captured in the transitional zone and only four in extralimital agriculture.

A single individual of *Hypsugo kitcheneri* (Thomas, 1915), one of the Borneo endemic bat species, was captured from the hostel at Mahua Substation. This species has been recorded from several localities in Borneo. The holotype of *H. kitcheneri* was collected from Buntok, Barito River in South Kalimantan (Francis and Hill 1986). Subsequently, one specimen was collected from Laham, East Kalimantan, and two specimens were collected from Sandakan, Sabah. However, the specimens from Sandakan have not been found since collection (Francis and Hill 1986; Medway and Harrisson 1963; Wilson and Mittermeier 2013). The specimen from this survey was deposited at the Institute of Tropical Biology and Conservation, Universiti Malaysia Sabah.

2.4.2 Acoustic Surveys

In total, 339,422 pulses were extracted from the recordings between 2018–2019. Of these 208,523 were recorded during the first sample of each point. Peak frequencies of pulses ranged 20–217kHz. The minimum acoustic threshold post-filtering was 20kHz. Acoustic activity was significantly different across the four habitat types (H = 15.093, df = 3, p-value = 0.002; Figure 2.4). Acoustic activity was lowest in the core zone but comparable across the buffer zone, transition zone, and extralimital agriculture. This is supported by the pairwise comparisons (Supplementary Figure S2.1).

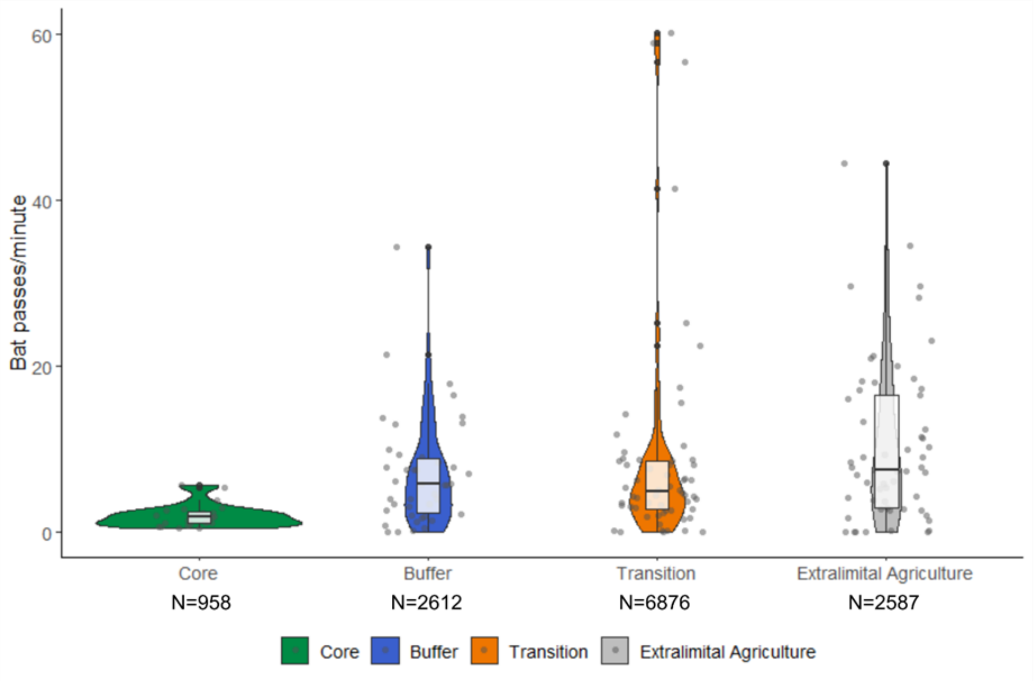


Figure 2.4 Violin plots demonstrating the variation in acoustic bat activity recorded within the four zones of the CRBR, Malaysia

Colours and symbols indicate different MAB zones. N = total number of pulses recorded per zone.

2.4.2.1 Call structure

Although there was overlap in the structure of the acoustic activity recorded across zones, there were identifiable differences in call duration, start and maximum frequency, and curvature of calls between zones (Figure 2.5). Longer calls were observed in the extralimital agriculture as identified by the PCA, scoring high values on PCA2. PCA loading contributions for PCA2 were dominated by curvature and call duration (ms) whereas PC1 was dominated by Fmax (kHz) and Fstart (kHz) (Supplementary Figure S2.3). The range of call curvature was greatest in the extralimital zone and lowest in the core zone. Calls in the core zone were typically short and high frequency. There was substantial variability in call structure in the buffer zone, where the greatest range of call duration and peak frequencies were recorded. However, whereas there were three distinct peaks in peak frequency recorded in extralimital agriculture (~20kHz, ~40kHz, and ~55kHz), a singular peak concentrated at ~55kHz was recorded in the buffer zone. Call structure in the transition zone reflects that of the buffer zone but demonstrates lower variability in call duration and higher variability in peak frequency. It also demonstrates a peak in peak frequency at ~55kHz, as well as another at ~20kHz. Using corresponding peaks across zones, the acoustic surveys suggest certain species are more active in the transition than in forest zones.

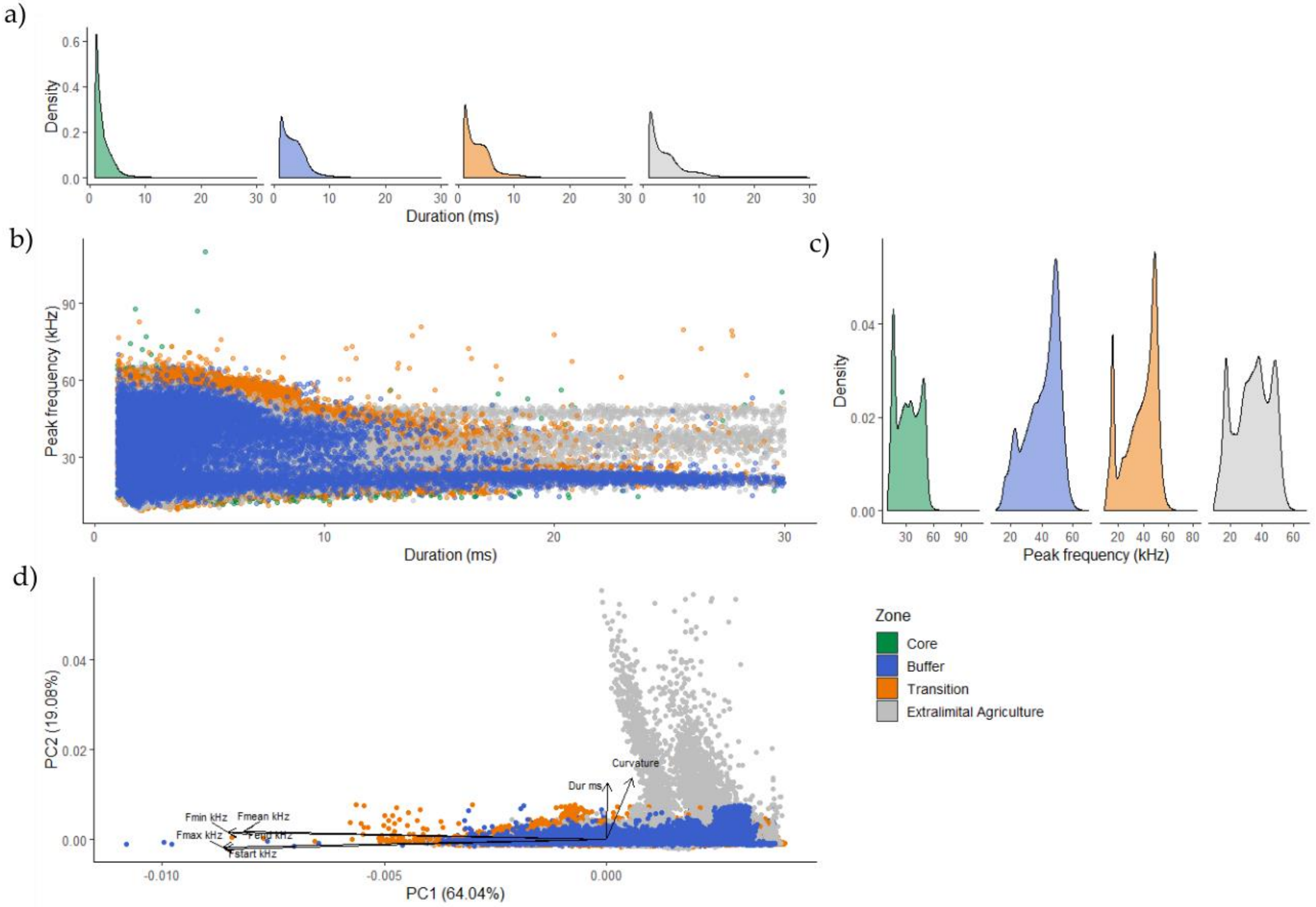


Figure 2.5 Differences in call structure between the acoustic bat activity recorded across the four zones of the CRBR, Borneo, Malaysia

a) Density plots showing the variability of call duration (ms) within each zone; b) scatterplot showing the relationship between two call variables within each zone; c) density plots showing variability in peak frequency (kHz) within each zone; and d) principal component analysis of call structure observed in each zone using six call characteristics, where "Fstart kHz", "Fmax kHz" and "Curvature" were the most heavily weighted components considering both dimensions.

2.5 Discussion

CRBR helps support bat diversity, with more than half of Borneo’s bat species now reported for the reserve. Notably, 16 of the 28 species captured in the CRBR in this study were not recorded in neighbouring extralimital agriculture. On a wider scale, the CRBR represents an area of relatively high value for bats compared to other protected areas within the Heart of Borneo (HoB) initiative (Struebig *et al.* 2010).

Species composition and activity varied among MAB zones with the core and buffer particularly important for conserving forest-interior insectivores. Thirteen species were captured exclusively in the core and buffer zone, including all species of Hipposideridae. The highest species richness for both edge/gap insectivores and forest interior insectivores was recorded in the buffer zone, however six forest interior species were only recorded in the core zone. Similarly, acoustic activity in the core zone was dominated by short (< 3ms), high-frequency calls typical of forest interior species of Vespertilionidae (subfamilies Kerivoulinae and Murininae) and Hipposideridae (Schmieder *et al.* 2012; Kingston *et al.* 1999; Kingston *et al.* 2000). In contrast, the peak of activity observed at ~55kHz within the buffer, transition, and agriculture is indicative of several edge-foraging species of Vespertilionidae, including *Myotis muricola and T. pachypus* (Hughes *et al.* 2011; Collen 2012). This is likely to be *T. pachypus,* an edge/gap insectivore with intermediate length calls (e.g., 3-10ms), which was captured abundantly across our study and was observed roosting in the core zone. Acoustic activity in the transition zone and agriculture was dominated by long calls with a low peak frequency of 20kHz, characteristic of open-space aerial foragers. This is likely to represent molossid species, such as *Mops mops* (peak frequency 18.5-23.3kHz; Kingston *et al.* 2003). However, in the agricultural sites, calls exhibited increased curvature and a greater range of low and mid-range peak frequencies, suggesting that diversity of open-space aerial foragers was greater in agriculture than the transition. The peak in activity of long calls at 30-40kHz is likely to represent molossids (such as *Chaerephon plicatus*), or emballonurids (such as *Saccolaimus saccolaimus*; Hughes *et al.* 2011; Kusuminda and Yapa 2017).

The decrease in forest-interior species within the transition and extralimital agriculture reflects previous studies that indicate this ensemble responds negatively to land-use change (Kingston 2013; Meyer, Struebig and Willig 2016). This includes species such as the Gilded Tube-nosed bat (*Murina rozendaali*) which is listed as Vulnerable by the IUCN and their populations continue to decline due to habitat loss and disturbance (IUCN 2022). Vegetation simplification is a consequence of agricultural intensification (Huang *et al.* 2019; Wordley *et al.* 2017) and ecomorphological traits characteristic of forest-dependent species (Kingston *et al.* 2003) are a limitation to forest ensembles persisting in these simplified, open spaces. Mechanisms for this decline in bats include reduced roost availability for plant-roosting species (Huang *et al.* 2019; Cortés-Delgado and Sosa 2014; Struebig *et al.* 2013), degraded roost quality (Phommexay *et al.* 2011), and reduced foraging success. If a species cannot persist in a matrix, connectivity between forest fragments is also greatly reduced (Meyer 2015). More research to better understand how we can conserve these forest-dependant species has been one of the four main research priorities for the Southeast Asian Bat Conservation Research Unit over the last decade (Kingston 2010a). Promisingly, our research shows that the core and buffer biosphere zones can help provide habitat refugees for these species.

Bats of the plant-visiting family Pteropodidae were captured in all four zones but dominated captures in the extralimital agriculture. Unlike forest-dependent insectivores, the plant-visiting pteropodids can travel larger distances (Start 1975; Wordley *et al.* 2018) as well as exploit the abundance of plant food resources available in the transition zone and extralimital agriculture (Furey, Mackie and Racey 2010; Acharya *et al.* 2015). This may explain the higher diversity and abundance of these species observed outside of the reserve in our study. Cave-roosting insectivores are also considered more tolerant of anthropogenic disturbance than forest-dependent species (Kingston *et al.* 2003; Kingston 2013; Struebig et al. 2013). Their low abundance across our study reflects the lack of available cave systems nearby rather than portraying a true response to disturbance.

Our findings support previous studies that the core and buffer zones are the most important zones for conserving bats in biosphere reserves (Taylor *et al.* 2013; López-González *et al.* 2015). However, our study differed from previous studies as species richness was comparable in the buffer and core (Taylor *et al.* 2013; López-González *et al.* 2015). In both Vhembe Biosphere Reserve in South Africa and La Michilía Biosphere Reserve in Mexico, more species were found in the core than buffer (Linden *et al.* 2014; López-González *et al.* 2015; Weier *et al.* 2017). The core zone has also been shown to provide important resources for species persisting across multiple zones. In Calakmul Biosphere reserve, Mexico, *Sturnira lilium* is a common species in the buffer and core but is reliant on primary forest in the core zone for roosting (Evelyn and Stiles 2003). We observed bamboo bats (*Tylonycteris* spp.) roosting within the core but foraging across multiple zones. Therefore, the core zone may provide roosting opportunities that help support greater populations of plant-roosting species in neighbouring MAB zones than may otherwise be possible. There may also be greater atmospheric attenuation in the core zone as a result of thicker vegetation. This reduces the distance at which calls can be detected. Therefore, the effective sampling size is likely to be lower in the core than the buffer resulting in the potential for artificially lower rates of activity.

Zoning provided by biosphere reserves has proved relevant to the conservation of other mammals, although the value of the core zone compared to the buffer zone differs across species. Similar to bats, forest-dependent species, including Sun bear (*Helarctos malayanus*), Clouded leopard (*Neofelis diardi*), and Marbled cat (*Pardofelis marmorata*) were only found in the core zone of Giam Siak Kecil-Bukit Batu Biosphere Reserve, Indonesia, whereas the generalist Palm civet (*Paradoxurus hermaphroditus*) was only found in the buffer zone (Motoko *et al.* 2012). Elsewhere, there appears to be no effect of the management zones on mammal richness and distribution in the reserve (e.g., Licona et al. 2011; Woldegeorgis and Wube 2012), or even greater species richness outside the reserve (Tabeni, Spirito and Ojeda 2013), but studies providing robust comparison across zones and extralimital habitats remain few.

Land-use change is the primary threat to bats in Southeast Asia (Kingston 2013), but our results support previous studies that show that ensembles are not equally affected by disturbance. Ecological characteristics such as roost type, foraging ensemble, and body size differentially confer vulnerability or resilience (Meyer, Struebig and Willig 2016). For example, different roost ensembles are affected by different drivers of disturbance. Foliage roosting species are affected by a reduction in roosting opportunities whereas cave-dwelling species are affected by both disturbance at the roost and greater commuter costs as foraging habitats are fragmented (Struebig *et al.* 2008; Nurul-Ain, Rosli and Kingston 2017). Body mass is positively correlated with resilience to disturbance whereas smaller species that emit higher peak frequencies responded negatively (Huang *et al.* 2019; Phelps *et al.* 2018). Mean body mass of the species recorded in our study was twice as high in the extralimital agriculture (26.44g) than in the core zone (13.11g; Table 2.2) however this does not consider the relative abundance of each species.

Insectivores are also known to exhibit a higher sensitivity to disturbance than herbivorous species (Boyles and Storm 2007). Our results highlighted a compositional shift from a dominant insectivorous bat assemblage in the core to one dominated by frugivorous species in extralimital agriculture. This may be the result of reduced prey availability, as land-use change alters insect communities (Ewers *et al.* 2015) and reduces insect biomass with consequences for bat diversity and activity (e.g., Phommexay *et al.* 2011). However, this compositional shift in the bat fauna was not observed within the transition zone. The transition zone sites were floristically less diverse than those in extralimital agricultural (see site descriptions) and predominantly comprised rubber and oil palm, therefore we suspect they did not provide abundant fruit resources necessary to support more frugivorous species. Along a disturbance gradient from primary forest to oil palm in lowland Borneo, capture rates were lowest in oil palm (Fukuda *et al.* 2009). This suggests the potential to improve the biodiversity value of transition habitats of CRBR, perhaps by implementing agroforestry crop systems that increase fruit and nectar resources. More diverse agroforestry, as well as an increase in patches of native vegetation, would also increase vegetative complexity and hence insect diversity and abundance important for edge and open space insectivorous bats.

Our results should be considered in the light of some limitations, particularly that sampling was not even across the reserve or the wider landscape. Differences in sampling effort relate to differing trap efforts between the different zones, access, and sampling efficiency between habitats. Whilst we were not able to ensure equal trap effort across zones, the species accumulation curves suggest inadequate sampling was limited to the transition zone for mist-netting and harp-trapping and the buffer zone for harp trapping. Therefore, it is possible that our results underestimate the overall diversity and abundance of bats utilizing the transition zone, and the diversity and abundance of forest-interior insectivores in the buffer zone. Our sampling was also restricted due to difficult terrain, and greater access was available at higher elevations. Elevation is known to negatively correlate with bat species richness (Cisneros *et al.* 2014), therefore it may have contributed to non-uniform capture rates across sites at differing elevations, rather than the effect being solely driven by land-use change. As our sampling was restricted to the wet season, seasonality may also have had a negative effect on capture rates (Meyer 2015). Future survey efforts should aim to isolate these surveying co-variates as random effects. Within the MAB zones, there is geographic variability in land-use practices and compliance with UNESCO legislation. Whilst we aimed to include a diversity of these practices in our study (e.g., oil palm, rubber, agriculture in transition zone), we did not investigate how specific practices, or how landscape-scale characteristics more broadly differed in their importance for protecting bat species within the reserve. Finally, the buffer zone is not continuous in CRBR, and in some areas the core abuts the transition zone. The buffer zone is designed as a low-contrast intermediary zone between the core and transition zones, thereby limiting edge effects related to the boundary between these two zones (Ishwaran, Persic and Tri 2008; Reed 2019). We therefore predict that the species richness and abundance of bat species would be lower where the core zone neighbours the transition zone directly, due to greater disturbance. However, in our study we did not compare areas of the MAB reserve with and without this buffer, thus future research is needed to validate whether the zone is fulfilling this role.

Trap efficacy and call detectability are also dependent on ensemble and habitat structure (Kingston 2013; Meyer 2015). As expected within our study, higher species richness was observed using harp traps compared to mist-nets within the core and buffer zones (e.g., forested zones) and higher species richness using mist-nets in extralimital agriculture (e.g., open spaces). Live trapping is also considered to be less effective in open spaces such as transition and agriculture, where bats are less likely to encounter a trap. High-frequency calls are more easily attenuated by the atmosphere than low-frequency calls, and the quantity of environmental clutter negatively correlates with the range of call detectability (Meyer 2015). For both these reasons, forest-interior insectivores such as *K. hardwickii* are more difficult to detect acoustically than edge/open space insectivores, and this is likely to explain why the acoustic activity was higher in all zones compared to core forest. By using live trapping and acoustic sampling as complementary techniques, we aimed to limit these biases. It is difficult to extrapolate robust conclusions from the acoustic data as it is not possible to calculate abundance or differentiate between species. Manual identification of calls within this study was not possible, due to time constraints and the lack of a call library. Further development of automated or semi-automated classifiers for the region will improve the capacity of acoustic sampling for monitoring bat populations within Southeast Asia. Despite this, the acoustic data aided in identifying areas where activity was inadequately represented by live-captures.

There were several unexpected findings within our study. Three species, *Kervioula papillosa*, *Rhinolophus borneensis*, and *Rhinolophus acuminatus*, were all recorded in transition despite being characteried as forest-specialists. However, all these records were from only one site (5◦47’53.20”, 116◦24017.93”), which is uncharacteristic of the transition zone as it is densely forested and neighbours the core forest directly. Therefore, we do not believe these records are representative of wider transition zone. Similarly, *Rousettus amplexicaudatus*, a cave-roosting species (Phillipps and Phillipps 2016), was captured within extralimital agriculture. However, as we only caught one individual and the species is capable of commuting 50 km nightly to find fruit resources (IUCN 2019), it is likely to be a transient individual rather than being locally common, as it is elsewhere in Sabah (Phillipps and Phillipps 2016).

Although we added eight species to the list for CRBR, we did not catch 20 species previously reported (Tuen, Rahman and Salleh 2002). This is most likely a consequence of our limited sampling effort and the survey methods used. Many of the species absent from our list are difficult to capture in nets or harp traps unless they are positioned close to a roost or established flyway (e.g., *Pipistrellus* spp., *Taphozous longimanus*, *Saccolaimus saccolaimus*). These species might ultimately be identified in the acoustic recordings, subject to further development of call classifiers and library development in the region. However, it is worth noting that several species that we did not catch commonly roost in large caves (e.g., *Miniopterus* spp., *S. saccolaimus*, *T. longimanus*, *Eonycteris spelaea*) or large tree hollows (e.g., *Dyacopterus spadiceus*, *Cheiromeles torquatus*), suggesting that disturbance at caves and land-use change may play a role. In a similar vein, some species are known to be declining (*C. torquatus, P. vampyrus*) (IUCN 2019) and may have been lost from the landscape.

2.6 Conclusions

In summary, the results of this study demonstrate that MAB programmes can be effective at promoting diversity in cooperation with human activity. The buffer zone provided adequate habitat for several forest-dependent and threatened species but not all. Therefore, it remains pivotal to ensure areas of primary forest are preserved. Nonetheless, buffer zones do help alleviate the effects of land-use change by reducing the contrast between protected areas and agriculture for many species. Our study did not demonstrate any additional conservation value of the transition zone compared to neighbouring agriculture, suggesting potential for agroforestry interventions that enhance vegetative complexity and retain patches of natural forest to improve the biodiversity value of the transition zone. Evaluations such as this are important to ensure protected area schemes are functioning effectively. We conclude that collectively UNESCO Biosphere zones can provide valuable habitat for conserving bat diversity in the Southeast Asian tropics.

2.7 Acknowledgements

We are grateful to the Sabah Biodiversity Council for granting the access license (permit JKM/MBS.1000-2/2 JLD.6 (53)), and Sabah Parks for permission to work in the CRBR. Specifically, Madam Rimi Repin, Deputy Director (Research and Education) of Sabah Parks, Thomas Yussop, Park Manager of Crocker Range Park, Yassin Miki, Park Manager of Kinabalu Park and Simun Limbawang, Head of Inobong Research Station (Kinabalu Park). We thank the District of Tambunan for cooperation in surveying outside the CRBR in the Tambunan Valley and are grateful to Anastasius Peter Modi, Angelo Asis, Helen Gotonis, and Amry Sorop for access to their land for surveys. We thank James Jupikely, Universiti Malaysia Sabah (UMS) Faculty of Sustainable Agriculture, Monica Suleiman, Director of Institute for Tropical Biology and Conservation, UMS, Jaya Seelan, UMS, Kalsum Yusah, UMS Michael Farmer, Texas Tech University, Sarah Fritts of Texas State University for logistical and participatory support of the project. Additional students participated in the fieldwork and we thank them for their contribution: Amirah Amat, Ummu Safiyyah Daud, Nurhani Adila Abdul Halim, Su Suzaine Sukiman and Nurkujayati Martin, Alex Troutman and Tyler Test. We are grateful to Joe Chun-Chia Huang for reviewing the high-frequency CF call and sharing his opinion on species identity and to two anonymous reviewers for the constructive comments. One reviewer very kindly shared some R code.

2.8 Supplementary information

Table S2.1 Detailed site description for sampling localities within CRBR

Site description, GPS coordinates, the abbreviations used to identify them in figure 2.1, their zones and sampling duration.

Zone	Localities	Latitude	Longitude	Description	Trapping Duration
Core	L.C.1	05°50.293	116°08.834	Mostly covered by lower montane and mixed dipterocarp forest, occurring at about 500 m asl. It is the start/end point of the Salt Trail, connecting the substation with Malungung Control Post in the north-east edge of CRBR.	19 nights
	L.C.2	05°50.856	116°08.527		
	L.C.3	05°51.426	116°08.404		
	A.C.1	05°51.313	116°08.347	The transect is along a ridge system on the Bansandon trail covered with mature trees with closed canopy and open understorey. There are clumps of bamboos along the transect.	3.5 night
	A.C.2	05°51.116	116°08.230	The transect is along a ridge system on the Salt trail covered with mature trees with closed canopy and open understorey. There are clumps of bamboos along the transect.	3.5 nights

Buffer	L.B.1	05°42'7.14"	116°20'7.05"	The trapping site consists of contrasting vegetation on the lower and upper portions of the hill. The lower portion of the hill is mostly covered by oil palm (*Elaeis guineensis*), durian (*Durio* spp.) and rambutan (Nephelium *lappaceum*). Meanwhile, the upper portion of the hill is an old growth forest. Trapping site occurs below 950m asl.	3 nights
	L.B.2	05°39'36.99"	116°15'2.80"	The trapping site is the start/end point of the Salt Trail that extends for 34km, connecting the control post with Inobong Substation in the north-west edge of CRBR. Covered by hill dipterocarp forest occurring below 900 m asl.	10 nights
	A.B.1	05°47'53.20"	116°24'17.93"	Transect is along a gravel road surrounded by oil palm plantation (*Elaeis guineensis*), mixed fruit orchard, secondary forest and an old growth forest at the top of the hill.	1 night
	A.B.2	05°39'36.99"	116°15'2.80"	Transect is on a terraced hill along a paved road. It is surrounded by forest edges with patches of agricultural plots.	1 night
	A.B.3	05°51'18.95"	116° 8'11.68"	Transect is on a paved road with patches of open space, but mostly covered with dense vegetation on both sides. There are street lightings along the transect.	3.5 night
	A.B.4	05°49'55.78"	116° 8'39.83"	Transect is along a main road surrounded by sparse secondary vegetation and plantation, mostly open space. This transect leads to the main entrance of the park.	2 night
Transition	L.T.1	05°47.841’	116°24.439’	Trapping site is covered by hill dipterocarp forest occurring below 1000m a.s.l.	4 nights

	L.T.2	05°32'36.88"	116°15'14.61"	Trapping site is a clear-cut area with small forest fragment at the bottom of the hill.	1 night
	A.T.1	05°47.841’	116°24.439’	Transect is around Mahua substation facilities surrounded by forest edges and a flowing river.	1 night
	A.T.2	05°32'36.88"	116°15'14.61"	Transect is along forest edges and a flowing river on one side. Land-use change presence around the park boundary, mostly for agriculture.	1 night
	A.T.3	05°51'47.89" N	116° 7'50.43"	Transect begins from the entrance to the park along wide road which is paved and unpaved along the route. There are several residential properties, areas of young forest with a low canopy, and areas of clearing.	3 nights
	A.T.4	05°52'26.52"	116° 7'42.66"	Transect is along a wide, dirt road surrounded by secondary forest with a low canopy, residential properties, agriculture, and oil palm small holdings.	3 nights
	A.T.5	05°49'19.55"	116° 8'16.05"	Transect is along a dirt road surrounded by secondary forest with sparse vegetation and open canopy, which leads to a rubber plantation. There is a human settlement nearby.	3 nights
Extralimital	L.E.1	05°45'57.45	116°28'11.40"	Trapping site is covered with abandoned rubber plantation, with some banana plants with thick undergrowth. This site occurs around 900 m a.s.l.	3 nights
	L.E.2	05°42'34.29"	116°24'28.98"	Trapping site is a secondary forest with patches of abandoned rubber plantation (*Hevea brasiliensis*) and mixed fruit plantation such as banana (*Musa* spp.), rambutan (*Nephelium*	5 nights

				lappaceum), soursop (*Annona* spp.) and mango (*Mangifera* spp.). This site occurs below 700m asl.	
	L.E.3	05°33'5.41"	116°17'25.27"	Trapping area was coved by monoculture rubber plantation occurring around 700 m asl.	1 night
	A.E.1	05°45'57.45	116°28'11.40"	Transect is along a gravel road surrounded by human settlement, abandoned rubber plantation (*Hevea brasiliensis*) and patches of secondary forest. Presence of street lighting along the transect.	1 night
	A.E.2	05°45'57.45	116°28'11.40"	Transect is along a gravel road surrounded by secondary forest on both sides.	1 night
	A.E.3	05°42'34.29"	116°24'28.98"	Transect is along a paved road surrounded by human settlement and agriculture plots such as rubber plantation (*Hevea brasiliensis*), banana (*Musa* spp.) orchard, mango (*Mangifera* spp.) orchard and patches of rich fields.	1 night
	A.E.4	05°42'34.29"	116°24'28.98"		1 night
	A.E.5	05°42'34.29"	116°24'28.98"		1 night

Table S2.2 Relationship between bat activity and survey time

The corresponding R² values for linear regression models for the relationship between bat activity and time of surveying. Models were conducted per CRBR zone and extralimital agriculture, as well as for all activity recorded.

	Adjusted R-squared	df	*p* value	
Core	0.019	193	0.030	.
Buffer	0.014	267	0.031	.
Transition	0.007	197	0.127	
Extralimital agriculture	-0.006	84	0.462	
Total activity	0.374	745	5.93E-08	***

Significance values where '.' <0.05 '' <0.01 '**' <0.001 '***' <0.0005

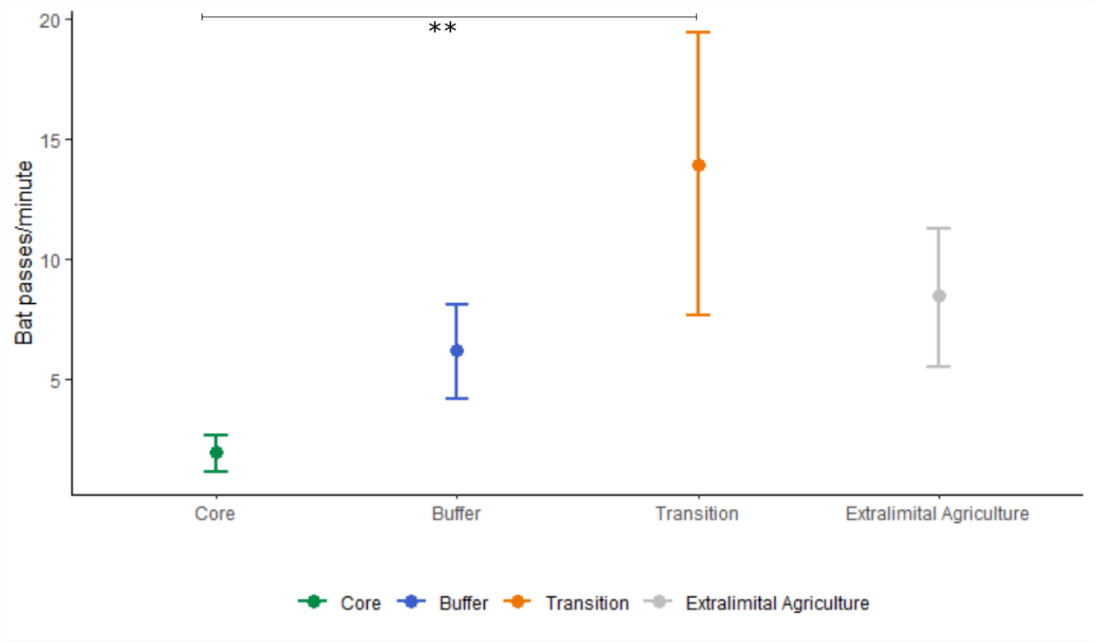


Figure S2.1 Results of the pairwise comparison of mean bat activity

Results of the pairwise comparison of mean bat activity recorded within the four zones of the CRBR, Borneo, Malaysia, with bootstrapped (nonparametric) 95% confidence intervals. ** < 0.01

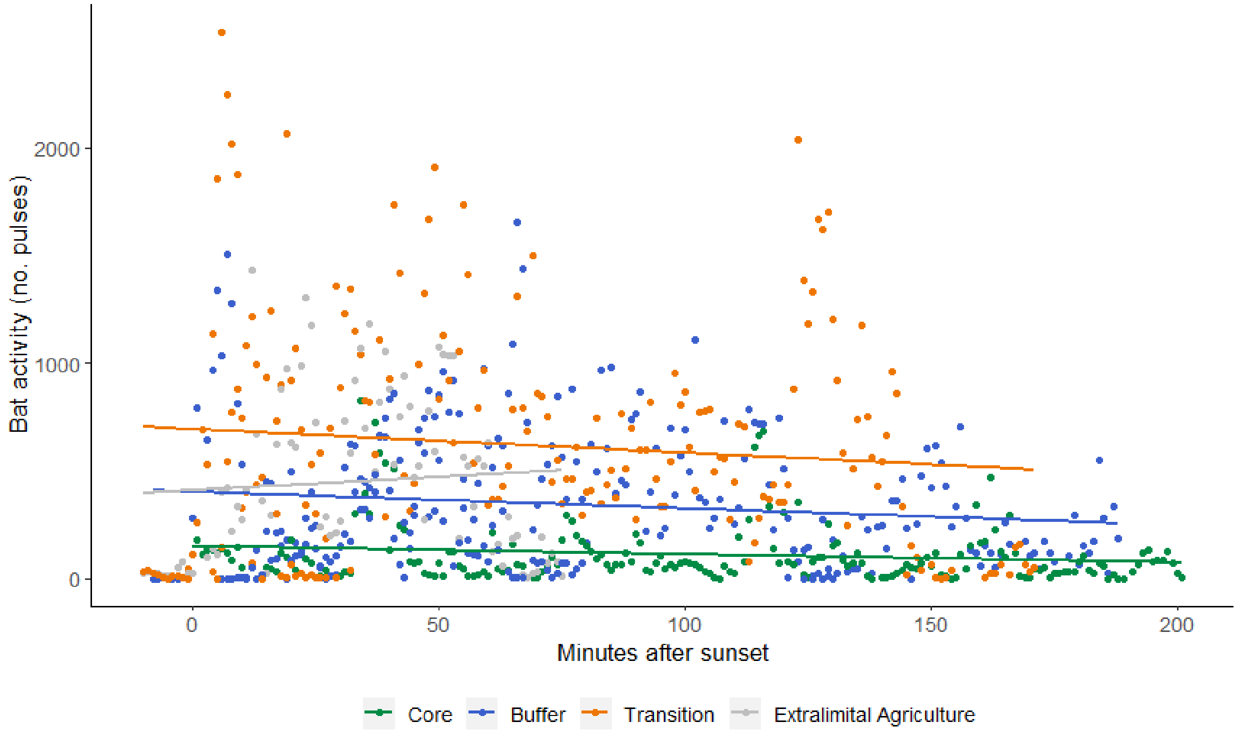


Figure S2.2 Scatterplot of temporal variation in bat activity

Acoustic bat activity recorded within the four zones of the CRBR, Borneo, Malaysia. Colors indicate different MAB zones.

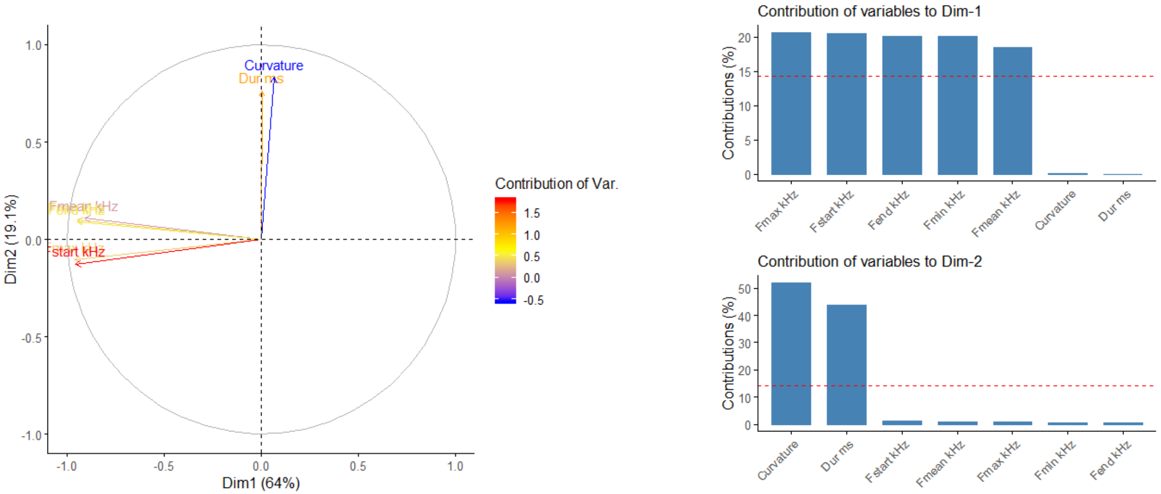
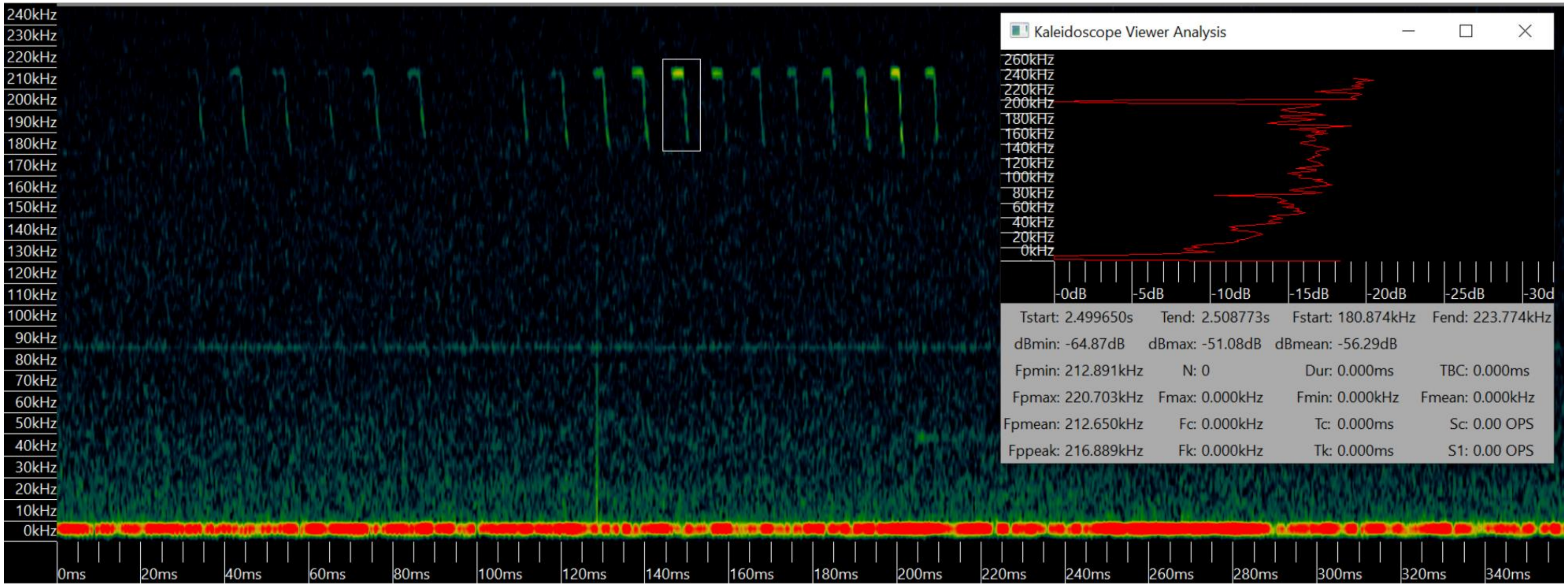


Figure S2.3 The relative importance of acoustic variables

Plot highlighting the relative importance of acoustic variables used for dividing calls along PCA 1 (Dim-1) and PCA 2 (Dim-2). The greater the contribution, the greater the importance.



Tstart: 2.499650s Tend: 2.508773s Fstart: 180.874kHz Fend: 223.774kHz
dBmin: -64.87dB dBmax: -51.08dB dBmean: -56.29dB
Fpmin: 212.891kHz N: 0 Dur: 0.000ms TBC: 0.000ms
Fpmax: 220.703kHz Fmax: 0.000kHz Fmin: 0.000kHz Fmean: 0.000kHz
Fpmean: 212.650kHz Fc: 0.000kHz Tc: 0.000ms Sc: 0.00 OPS
Fppeak: 216.889kHz Fk: 0.000kHz Tk: 0.000ms S1: 0.00 OPS

Figure S2.4 Sonogram of high frequency *Hipposideros* species

Sonogram of high frequency *Hipposideros* species, suspected to be *Hipposideros doriae*, with corresponding power spectrum. Call was recorded within the core zone of CRBR, Borneo, Malaysia, June 2019.

Chapter 3 Riparian Reserves Promote Insectivorous Bat Activity in Oil Palm Dominated Landscapes

Katherine E. Mullin[1†], **Natalie Yoh[1†],** Simon L. Mitchell[1], Saloni Basrur[1], Dave J. I. Seaman[1], Henry Bernard[2] and Matthew J. Struebig[1*]

1. Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, Canterbury, Kent, CT2 7NR, UK
2. Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah, Kota Kinabalu, Malaysia

*Correspondence: m.j.struebig@kent.ac.uk

† These authors contributed equally to the work.

This chapter is based on the following publication:

Yoh*, N., Mullin*, K., *et al.* (2020). Riparian Reserves Promote Insectivorous Bat Activity in Oil Palm Dominated Landscapes. *Frontiers in Forests and Global Change, 3*. doi: 10.3389/ffgc.2020.00073 *Authors contributed equally

Author Contributions: MS and HB conceived and designed the study. KM, SM, and MS conducted the fieldwork. KM and SB processed the bat calls. NY and DS performed the statistical analysis. NY, KM, and MS wrote the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

3.1 Abstract

The expansion of oil palm agriculture has contributed to biodiversity loss in Southeast Asia and elsewhere in the tropics. Riparian reserves (areas of native forest along waterways) have the potential to maintain forest biodiversity and associated ecological processes within these agricultural landscapes. Using acoustic sampling, we investigated the value of riparian reserves for insectivorous bats in oil palm plantations in Sabah, Malaysian Borneo. We compared general bat activity, foraging activity, and species occupancy between riparian areas in forest and riparian reserves in oil palm plantations. Overall bat activity varied little between riparian reserves in oil palm and riparian forest. Rather, activity was greatest in areas with a high forest canopy, irrespective of how much forest was available within or outside the riparian reserve. Bat foraging activity, as well as the occupancy of two species, was greatest in the forest sites, and while bats were detected in the oil palm riparian reserves, both foraging and occupancy were more associated with topographic ruggedness than forest amount or height. Our results indicate that habitat structure within riparian reserves may be more important than reserve size for supporting insectivorous bat diversity within oil palm landscapes. These findings provide important insights into the extent of the ecological benefits provided by conservation set-asides in forest-agricultural landscapes in the tropics.

Keywords: Chiroptera; Southeast Asia; palm oil; riparian buffer; acoustic monitoring; habitat fragmentation, land-use change, biodiversity

3.2 Introduction

The intensification and expansion of agriculture are major causes of tropical biodiversity declines (Phalan *et al.* 2013). Since the tropics support large numbers of species and some of the world's last pristine habitats, the ongoing biodiversity crisis occurs disproportionately in these regions (Rosa *et al.* 2016; Sodhi, Brook and Bradshaw 2013). Here industrial commodity crops, including but not limited to oil palm (*Elaeis guineensis)*, are becoming widespread. In the last decade palm oil production has more than doubled and demand is likely to rise, due to both its status as the world's primary and cheapest vegetable oil and its use as a biofuel feedstock (Meijaard *et al.* 2018). Southeast Asia dominates as the epicentre of oil palm production, with Indonesia and Malaysia being the leading producers (Meijaard *et al.* 2018). Seventeen percent of Malaysian plantations are thought to have needed forest clearing (Pirker *et al.* 2016). With oil palm plantations being a poor biodiversity substitute for native forest habitat (Fitzherbert *et al.* 2008), this raises concerns about the environmental sustainability of the crop, and the associated negative impacts agricultural expansion is having on tropical biodiversity.

One way to mitigate biodiversity losses in oil palm is to maximise natural habitat within estates by retaining habitat remnants as conservation set-asides (Meijaard *et al.* 2018). Protected forest patches are promoted within the Roundtable on Sustainable Palm Oil certification standard (RSPO 2018), and while these improve prospects for terrestrial mammals, patches are rarely large enough to sustain viable populations individually (N J Deere *et al.* 2020). Therefore, improving habitat area and connectivity between these patches is paramount, and will be particularly important for species with limited dispersal capabilities that are unable to traverse large expanses of open, agricultural land (Carroll *et al.* 2004).

Protecting riparian reserves (areas of native vegetation alongside rivers) is one potential way to help join up forest patches in tropical agriculture, but to date, they have been poorly studied in the context of oil palm (Luke *et al.* 2019). Riparian reserves (also known as buffers) are typically protected to maintain hydrology (Tabacchi *et al.* 2000) and nutrient regulation (Naiman, Decamps and McClain 2010), which can lead to potential co-benefits for biodiversity (Mitchell *et al.* 2018). Many tropical countries therefore protect riparian areas in some form, but there is a weak evidence-base supporting the size, or width, of the protection zone (Luke *et al.* 2019). Further evidence is considered a high priority within the oil palm research community (Padfield *et al.* 2019).

To date, bats have been poorly studied in the context of oil palm in part because they are difficult to capture and study in open habitats. However, bats make ideal indicator taxa since they are globally distributed, taxonomically stable, and can be monitored regularly to determine population trends if acoustic data are used (Jones et al. 2009). In addition, as primary invertebrate predators, bats are potentially important indicators of invertebrate abundance and their responses to environmental change. Bats comprise around a third of Southeast Asia's mammal diversity (Kingston 2010), and the conservation status of 17% of these species is currently unknown (IUCN 2019). The main threats to Asia's bats mirror those listed for other taxa; namely, deforestation, habitat fragmentation, and disturbance, and 22% are assessed as threatened (IUCN 2019). If these threats continue, Southeast Asia is projected to lose 40% of its bat species by the end of the century (Lane, Kingston and Lee 2006).

In palaeotropical forests, high bat diversity is associated with structurally complex, highly heterogeneous and lesser disturbed forest habitats (Struebig *et al.* 2013). Diversity and population viability are also greatest in larger forest patches (Struebig *et al.* 2008; Struebig *et al.* 2011). However, habitat disturbance and fragmentation affect bat species in different ways, with those in the families Hipposideridae and Rhinolophidae, and vespertilionid subfamilies Kerivoulinae and Murininae, being most affected (Struebig *et al.* 2008). The high sensitivity of these taxa is thought to be due to adaptations in wing morphology and echolocation call design that restrict their movements and foraging capabilities in more open habitats (Kingston, Jones, et al. 2003). Nevertheless, recapture distances and radiotracking studies confirm that some hipposiderid and rhinolophid species can commute several kilometres in a single night (Struebig *et al.* 2008), demonstrating that there is also substantial variation in species responses within these groups. In contrast, recapture and radiotracking studies of *Kerivoula* species rarely exceed 1km, and declines in abundance tend to be associated with the loss of tree cavities in disturbed forests since many of these species use tree features for roosting (Struebig *et al.* 2013). Other sub-components of insectivorous bat assemblages, such as those that forage in forest edges or open spaces outside or above the forest, may be able to adapt to disturbed habitats, and are frequently detected in farmland (Kingston, Francis, *et al*. 2003).

The importance of riparian reserves for bats has not been explicitly studied in the Southeast Asian tropics. However, in other parts of the world bats are known to benefit from protected waterways in agricultural areas. In the neotropics, for example, bats use riparian reserves as stepping stones between larger patches of vegetation (Peña-Cuéllar *et al.* 2015), and in Britain, reserves are known to be used for foraging due to the large number of insects supported by rivers (Vaughan, Jones and Harris 1996). In Mexico, riparian reserves were found to host a higher abundance and diversity of bats compared to other landscape features (Galindo-

González and Sosa 2003). Additionally, a study in Swaziland found that as well as high abundance and diversity there was also higher activity and species richness in riparian reserves compared to the surrounding savannah landscape (Monadjem and Reside 2008).

Here we investigate the use of riparian areas by palaeotropical bats in order to determine the optimum habitat features to support their conservation in Southeast Asian agricultural landscapes. We base our study in a region with widespread oil palm agriculture (Sabah, Malaysian Borneo), and use acoustic survey techniques as an alternative to traditional capture methods, which are known to poorly sample many insectivorous species, particularly in more open habitats. Specifically, we investigate how (i) total bat activity, and (ii) foraging activity varies between the riparian areas in forest and riparian reserves in oil palm, for all species and between call-types. We then identify the structural characteristics important for (iii) predicting bat activity and (iv) species-specific occurrence for a subset of taxa.

3.3 Materials and methods

3.3.1 Study Area

Research was conducted in and around the Stability of Altered Forest Ecosystems Project[1] within the Kalabakan Forest Reserve of Sabah, Borneo. The SAFE Project is a landscape experiment tracking changes to biodiversity and ecosystem functioning as logged mixed-dipterocarp forest is converted to plantation. Forest across the landscape has been logged multiple times since the 1970s, and during the fieldwork the remaining core areas of forest were being salvage-logged in preparation for additional plantation.

---

[1] (www.safeproject.net)

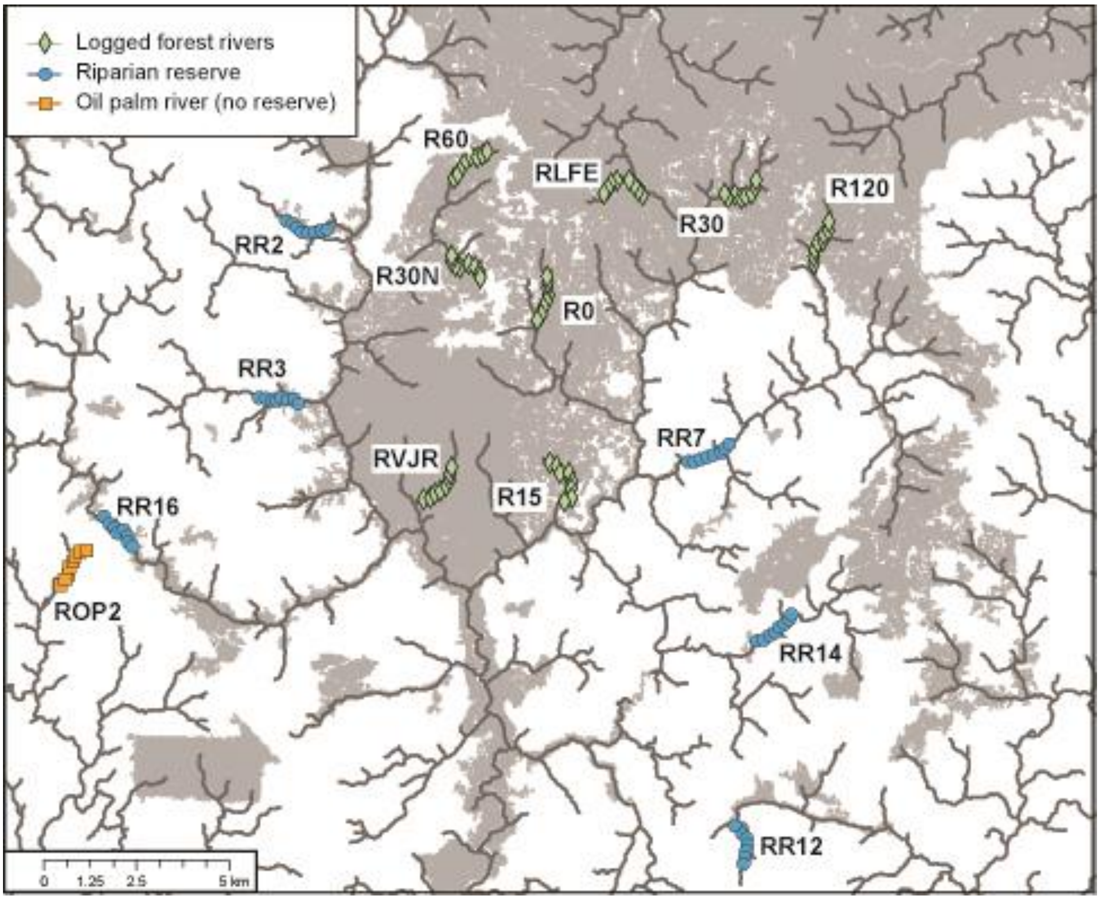


Figure 3.1 Location of the 15 rivers sampled for bats via acoustic monitoring in the Stability of Altered Forest Ecosystems (SAFE) landscape of Sabah, Malaysian Borneo
Each site comprised eight recording points, which were sampled multiple times. Forested area in grey (derived from Hansen *et al*., 2013 data for 2014); oil palm estates in white. Lines denote the river courses.

We sampled 15 riparian areas, comprising eight forested rivers in the SAFE project area, and seven riparian reserves in the surrounding oil palm estates, which were planted between 2000 and 2015 (Figure 3.1, Table S3.1). Forest sites were protected from further logging, but many remained highly disturbed from previous logging events. In Malaysia, riparian protection policies vary by state, with the prescribed reserve width in farmland depending on the size of the river and landscape context. In Sabah, companies are required to protect between 5 and 100m, with many reserves in oil palm being at least 20m wide from each riverbank (Luke *et al*. 2019). Therefore, the width of the riparian reserves in the oil palm varied substantially between and also within rivers, ranging from 0m (i.e., completely devoid of forest vegetation) in river ROP2, to >200m in parts of R14 and R16 - considering both riverbanks. All rivers were sampled at least twice (Table 3.1).

Table 3.1 Bat activity and foraging activity alongside rivers sampled at the SAFE project Sabah

Bat activity and foraging activity alongside rivers sampled at the SAFE project Sabah. Activity is partitioned into call-type and visit (1-4).

River	Total bat activity				Total foraging activity				CF activity				FMqCF activity				FMb activity			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Forested rivers																				
R0	56	65	11	-	6	8	1	-	14	4	1	-	42	60	7	-	0	1	3	-
R30	44	52	120	25	2	19	7	1	25	6	76	2	16	44	40	19	2	1	3	4
R15	153	155	46	216	15	38	0	23	38	49	2	166	114	98	41	50	0	8	3	0
R30N	198	115	111	141	15	10	6	1	53	63	43	28	145	52	66	87	0	0	2	23
R60	70	118	124	63	5	23	4	0	-	-	-	-	63	102	73	57	0	11	7	2
R120	65	157	291	192	4	19	17	9	-	-	-	-	29	77	171	144	0	5	16	6
RVJR	78	74	-	-	4	11	-	-	4	11	-	-	28	62	-	-	1	2	-	-
RLFE	29	73	52	-	0	6	2	-	0	4	0	-	23	65	48	-	6	4	3	-
Oil palm riparian reserves																				
RR2	55	44	-	-	10	2	-	-	19	17	-	-	31	17	-	-	5	10	-	-
RR3	207	94	-	-	31	3	-	-	14	13	-	-	191	77	-	-	2	4	-	-
RR7	125	120	182	-	10	4	7	14	0	1	16	22	99	98	135	244	23	20	25	20
RR12	123	90	237	-	21	16	9	-	21	16	9	-	108	79	150	-	0	3	0	-
RR14	107	73	-	-	27	1	-	-	27	1	-	-	92	69	-	-	1	0	-	-
RR16	44	49	269	-	4	8	19	-	4	8	19	-	16	24	199	-	2	4	6	-
ROP2	95	85	-	-	3	9	-	-	3	9	-	-	83	41	-	-	1	2	-	-

3.3.2 Acoustic Sampling

We implemented acoustic surveys along each river and the associated riparian forest/reserve using a walking-point-transect design. At each site, we recorded bats at eight sampling points positioned approximately 200m apart along a 2km stretch of river. Surveys were undertaken in the early evening during peak bat activity, commencing just before sunset (18:01-18:23, according to standardised times in Kota Kinabalu[2]). As heavy rainfall has implications for bat activity as well as the viability of recording equipment, we restricted surveys to rain-free nights, the number of nights abandoned due to rain was more than double that of the sample size achieved. At each point, we recorded bat activity for 10 minutes (80 minutes of recording per night), using a full-spectrum EM3+ ultrasonic recorder (Wildlife Acoustics Inc.). The detector was set to record intermittently, with sounds >18 db and >18 kHz triggering a recording. The sampling rate was set at 384 kHz, thus allowing frequencies up to 192 kHz to be recorded. The maximum duration of recordings was set at 10 seconds. We considered one trigger to be a bat pass of the microphone and used this as an index of bat activity. Triggers and associated data were stored on SDHC memory cards as raw WAV files.

The EM3+ microphone is omnidirectional, with optimal recordings ±30° vertically and 60° horizontally to the unit. As with all acoustic devices, the detection zone of the microphone varied for different species, depending on sound attenuation and vegetation clutter around the river. For this reason, we broke up the 10 minutes of recording into three components at each point: four minutes in the centre of the river, and 3 on each riverbank approximately 5m from the river edge. This minimised bias to species utilising merely the river, whilst also maximising detection of those bats in vegetation nearby. All rivers were sampled on at least two occasions, visiting sites sequentially from one end of the river to the other. Five rivers had four visits, five had three visits and five had two visits, with surveys spanning across 2014 and 2015. All visits used the same sampling points, and all recordings were processed following the same method. When logistically possible we sampled the rivers starting at the opposite end of the transect on the second sample in order to control for differing levels of bat activity over the period of the evening. For example, we expected higher bat activity earlier in the evening, around and just after dusk, with declining activity as the evening progressed.

---

[2] https://www.timeanddate.com/astronomy/malaysia/kota-kinabalu

3.3.3 Processing Acoustic Data

Data from each night were processed using the SonoBat SM2 Batch Attributer (Wildlife Acoustics PLC.) and then scrubbed using the SonoBat Batch Scrubber V5.2 to remove noise files. The remaining 16,530 files were processed using SonoBat 3.2p Batch Processor, and each sound file was manually inspected, then accepted or rejected as a genuine bat trigger. An independent trigger file was accepted as a genuine bat call if there was a minimum of three pulses (e.g. three individual vocalisations). To quantify bat activity at each river we used the number of genuine bat microphone triggers, noting that this could represent a single bat responsible for multiple triggers. During processing, we identified feeding buzzes through visual interpretation of the call and from playback. Feeding buzzes are a rapid increase in pulse rate as a bat closes on insect prey.

We used SonoBat call analysis software to automatically extract the following call parameters: pulse duration, and the highest frequency and the lowest frequency of the dominant harmonic. All calls were manually inspected alongside the call parameters to classify them into four call-types: constant frequency (CF; e.g. Rhinolophidae and Hipposideridae), frequency modulating quasi-constant frequency (FMqCF; e.g. Molossidae and several Vespertilioninae genre such as the bamboo bats, *Tylonycteris*), broad-band frequency modulating (FMb; e.g., Kerivoulinae and Murininae), and multi-harmonic frequency modulating (MHFM; e.g. *Megaderma spasma* and *Nycteris tragata*). Species utilizing CF, FMb, or MHFM calls are known to forage in cluttered narrow-spaces; CF species use flutter detection to find prey, whereas species using FMb and MHFM are active and passive gleaners respectively (Denzinger and Schnitzler 2013; Table S3.2). These three call types are therefore characteristic of species widely regarded as forest interior insectivores. On the other hand, species using FMqCF calls are characterised by aerial hawking to capture prey and are more commonly associated with forest edge and open spaces (Denzinger and Schnitzler 2013).

It was not possible to identify most calls to the species-level due to insufficient local knowledge of the echolocation calls for many species. However, horseshoe and leaf-nosed bats (Rhinolophidae and Hipposideridae respectively, and all CF species) could be confidently identified using the species-specific peak parameters of their constant frequency tone (Table 3.2). These call parameters were taken from a library of recordings from bats recorded in the landscape or elsewhere in northern Borneo. Analysis was therefore restricted to the level of call-type for most bats, with the addition of species-based analysis for CF bats.

Table 3.2 Call parameters of constant frequency bats

Call parameters of constant frequency (CF) bats identified in this study along with their activity in the riparian reserves.

Species	Low freq. (kHz)[a]	High freq. (kHz)[b]	Forest activity[c]	Oil palm activity[c]
Rhinolophus acuminatus	85	89	2	6
Rhinolophus borneensis	79	84	14	1
Rhinolophus sedulus	59	62	64	3
Rhinolophus trifoliatus [d]	47	52	777	291
Hipposideros cervinus	115	126	4	9
Hipposideros galeritus	108	114	13	14
Hipposideros ridleyi	61	63	11	1

[a] Lowest frequency of the dominant harmonic; [b] Highest frequency of dominant harmonic; [c] Number of genuine bat microphone triggers in riparian reserves. [d] It is possible that the call frequency range for *R. trifoliatus* overlaps with that of *R. francisii*, a recently named species in Sabah. Since the latter species is very rare and not known from the site, we assign all calls to *R.trifoliatus.*

3.3.4 Environmental Characteristics of Riparian Sites

Potential covariates were extracted at each point on every river, providing eight data points per river, i.e. 128 over the landscape. Riparian reserve width (m), applicable to only sites in the oil palm, was estimated at each point from LiDAR imagery at 5m resolution (described in Swinfield *et al*. 2020). The width of the river channel, measured in the field using a laser rangefinder (Leica Rangemaster CRF 1000), was subtracted from this estimate to determine the actual land surface within the reserve. The covariate used, *RiparianWidth*, is the mean width on a single side of the river, since this is typically referenced in environmental policy documents.

Previous research in the project area highlighted forest structural and landscape characteristics as important predictors of bat community composition (Struebig *et al.* 2013), and so two additional covariates were also extracted from each sample point. Average topographic ruggedness (*Topography*; 0-1) was measured within 50m of each point using Shuttle Radar Topography Mission data, according to Wilson *et al.* (2007). Mean canopy height (*CanopyHeight*: m), also sourced from the LiDAR data, was used as a proxy for forest quality. Taller forests in the landscape tend to have more trees, higher biomass, and greater vertical and horizontal complexity, all measures of forest structure and quality that are highly correlated (Deere *et al.* 2020). Both variables were extracted to points at a spatial scale of 50 m as this matched those used for a study on birds at the site (Mitchell *et al.* 2018) as we assumed bats and birds would be using the landscape in similar ways.

3.3.5 Statistical Analysis

For each point, we summed total activity or foraging activity for all call data combined, as well as grouping by call-type. Due to the unequal sample effort between rivers (Table 3.1) and high numbers of zeros in the dataset (which could not be remedied via zero-inflation models), this was calculated as a mean value across visits for each sampling point, e.g. sum of total bat passes recorded at the point per night divided by the number of nights recorded. We then applied generalized linear mixed effect models (GLMMs) with Poisson error distributions to explore the potential determinants of bat activity across sites. In addition, for a subset of CF bats that could be identified to species level, we applied hierarchical Bayesian multi-species occupancy modelling to determine whether there were patterns in species occurrence between rivers. Occupancy modelling was better suited to species-level analyses given the propensity for zero counts.

3.3.5.1 Activity

We first applied GLMMs to determine whether bat activity differed between forest and oil palm sites, and whether this was associated with the covariates. We then performed a second analysis using the subset of the data from oil palm riparian reserves to determine whether the width of the reserve influenced bat activity relative to the other covariates. Both analyses were conducted for all bat activity, CF calls, and FMqCF calls, totalling six GLMMs (Table S3.3). We did not work with data for the other call types due to insufficient numbers of bat passes.

In order to avoid collinearity in the GLMMs, we only included non-correlated variables (rs <0.5). The final fixed covariates were *HabitatType* (‘forest’ vs. ‘oil palm’, categorical with two levels), *TimeafterSunset* (continuous), *Topography* (continuous), *CanopyHeight* (continuous) and *RiparianWidth* (continuous). To incorporate the dependency among observations of the same river, we used *River-ID* as a random intercept. *HabitatType* was used only in the first GLMM to assess covariates across sites collectively [Equation (1:3)]. As

RiparianWidth is only applicable to oil palm reserves, we only included this covariate in the models focusing on those reserves [Equation (4:7)] (Table S3.3).

We used the package "lme4" (Bates *et al.* 2015) in the software R to fit the models in Equation (1:7) (Table S3.3) using a Nelder-Mead optimizer from the package "optimx" (Nash 2014). We checked dispersion using the package "blmeco" (Korner-Nievergelt *et al.* 2015; overdispersion < 0.75; > 1.4; Table S4). We centred and standardised covariates prior to analysis. We then undertook likelihood ratio tests to determine which covariates from the top five models were statistically important (Table S3.5). Considering the results of the likelihood ratio tests, the best fit model was chosen for each call group to analyse the relationship between activity including all covariates identified as significant (Table S3.3; Table S3.5). We then used the package "multcomp" (Hothorn, Bretz and Westfall 2008) to perform a multiple comparison test on these final models.

Data were too zero-inflated to explore foraging activity using feeding buzzes further using GLMMs. Therefore, we used the non-parametric Mann-Whitney U test to compare foraging activity between the forest and oil palm, as these data were not normally distributed (riparian forests: $W = 0.843$, $p < 0.001$). Foraging activity was examined as a proportion of total activity at any given point.

3.3.5.2 Occupancy

We used hierarchical Bayesian multi-species occupancy models adapted from Deere *et al.* (2017) to investigate species-specific occurrences in relation to environmental conditions, while accounting for imperfect detection (MacKenzie *et al.* 2002). This was only possible for seven species that could be identified from the acoustic data. Our framework linked occurrence and detection models for individual species via a hierarchical component that modelled regression coefficients as realisations from a common community-level distribution using hyper-parameters, which assume species respond to environmental conditions in a similar way. Since some species (e.g. *Rhinolophus trifoliatus*) were much more abundant than rare or elusive species (e.g. *Hipposideros ridleyi*), their occupancy estimates can be used to improve the estimation precision of the latter where detection records are limited (Pacifici *et al.* 2014).

We modelled the occurrence of species i at site j using the binary variable, $z_{i,j}$, where 1 indicates species presence and 0 denotes species absence. The occurrence state was specified as the outcome of a Bernoulli process, $z_{i,j} \sim \text{Bern}(\psi i,j)$, where $\psi i,j$ represents the latent occupancy state. The true occurrence state is imperfectly observed so we incorporated a second Bernoulli process, $x_{i,j,k} \sim \text{Bern}(p_{i,j,k} * z_{i,j})$, where $x_{i,j,k}$ represents the observed detection histories, and $p_{i,j,k}$ is the detection probability of species i, at site j for temporal replicate k. Within this framework, detection probability is conditional on species presence ($z_{i,j} = 1$), therefore it is not possible to estimate detection probability at a site where a species is absent ($zi,j = 0$). Our approach assumes abundance of a species does not affect the probability of detection (Royle and Dorazio 2008). To conduct the analysis, we specified models of the form:

Model 1
$$logit(\psi_{i,j}) = \mu_i River_j + \ \alpha_{1,i} + \ \alpha_{2,i} HabitatType_j \\ + \alpha_{3,i} CanopyHeight_j + \alpha_{4,i} Topography_j$$

$$logit(p_{i,j,k}) = v_i TimeafterSunset_j$$

Model 2
$$logit(\psi_{i,j}) = \mu_i River_j + \ \alpha_{1,i} + \ \alpha_{2,i} HabitatType_j$$

$$logit(p_{i,j,k}) = v_i TimeafterSunset_j$$

Model 3
$$logit(\psi_{i,j}) = \mu_i River_j + \ \alpha_{1,i} + \ \alpha_{2,i} CanopyHeight_j$$

$$logit(p_{i,j,k}) = v_i TimeafterSunset_j$$

Model 4
$$logit(\psi_{i,j}) = \mu_i River_j + \ \alpha_{1,i} + \ \alpha_{2,i} Topography_j$$

$$logit(p_{i,j,k}) = v_i TimeafterSunset_j$$

Model 5
$$logit(\psi_{i,j}) = \mu_i River_j + \ \alpha_{1,i} + \ \alpha_{2,i} CanopyHeight_j \\ + \alpha_{3,i} Topography_j$$

$$logit(p_{i,j,k}) = v_i TimeafterSunset_j$$

Occupancy and detection probabilities were modelled with intercepts on the logit scale, specific for each species. Five models were used to compare the fit of singular habitat characteristics in explaining species occurrence [Models 2, 3 & 4], against cumulative habitat characteristics [Models 1 & 5]. Covariate selection was partially informed by the results of the GLMM (Table S3.4). Covariates were centred and standardised prior to analysis to place them on a comparable scale and improve model convergence.

Models were specified using uninformative priors for intercept, variance, and slope parameters. We specified *River* as a random spatial effect for each species specifically. For each model we ran three parallel MCMC chains for 100,000 iterations each, discarding the first half during the burn-in process and thinning posterior samples by 50. The Gelman-Ruben statistic was assessed to ensure convergence (<1.1 for all parameters; Gelman and Hill 2007). Widely applicable AIC was subsequently calculated for each model, following Gelman, Hwang, and Vehtari (2013), and used for model selection [Model 5] (< 2 ΔAIC; Table S3.6). This model was then used to describe occupancy and detectability.

3.4 Results

Over 49 survey nights, we recorded 24,228 sound files, which were manually identified as 5,696 genuine bat triggers. Bats with frequency modulating quasi-constant frequency calls (FMqCF) were the most active amongst the community, constituting 70 % of total activity recorded. Seven CF taxa were identified and constituted 23.7 % of all calls (Table 3.1). The remaining call-types (FMb and MHFM) comprised 6.3 % of activity, and so are not explored further due to insufficient data.

3.4.1 Activity

3.4.1.1 Total activity and foraging activity between habitat types

We found no significant between activity in forest and oil palm rivers. This includes total activity, CF activity, and FMqCF activity (Figure 3.2a-c). Similarly, there was no overall difference in the proportion of foraging activity for activity overall (W = 442, p = 0.964) or for FMqCF species (W =465.5, p = 0.763; Figure 3.2d-f). However, the foraging activity of CF species (expressed as a proportion of total activity at a given point) was significantly greater alongside the forest rivers than the rivers in the oil palm (W = 345, p = 0.017).

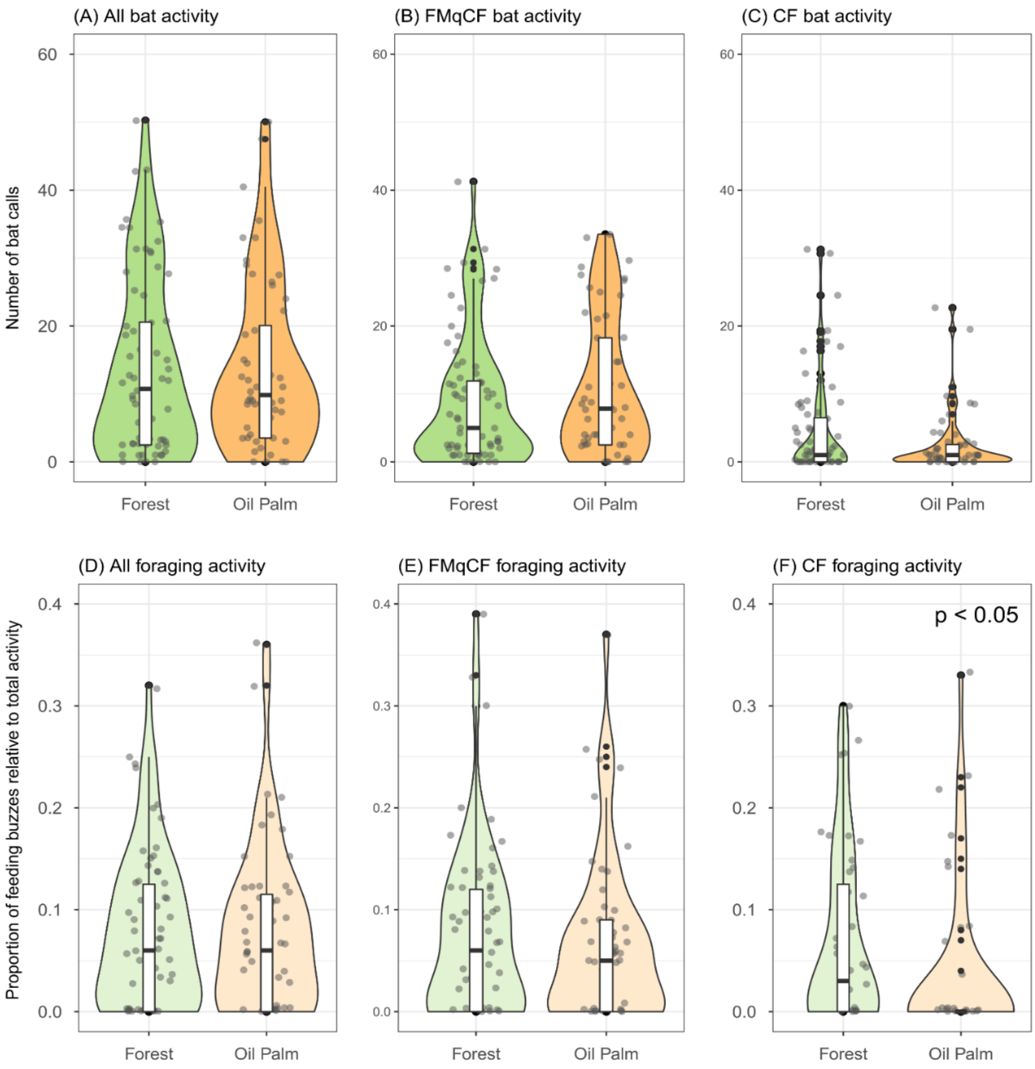


Figure 3.2 Bat activity and foraging activity between rivers and habitats

Violin plots and boxplots demonstrating variation in bat activity (A-C) and foraging activity (D-F) between rivers in contiguous forest versus those in oil palm for all species recorded in the study. Activity is represented as total activity (A, D), FMqCF activity (B, E), and CF activity (C, F). Points represent bat activity per sampling point (Forest = 72 points across 9 sites; Oil palm = 56 points across 7 sites), and are jittered to improve the presentation of the data distribution.

3.4.1.2 Total activity in relation to habitat covariates

Considering all study sites and total activity assessed together, the GLMMs revealed bats to be more active at sites with a higher forest canopy (Figure 3.3-3.4; Table S3.7), a pattern that was also evident for the activity of FMqCF-calling bats, but not for CF-calling bats when assessed separately (Table S3.5). Detection was negatively affected by time-of-recording, with greater activity detected closer to sunset for CF bats, FMqCF bats, and all call groups combined. Habitat type (i.e. forest vs. oil palm) was not identified as an important factor driving activity for any call group.

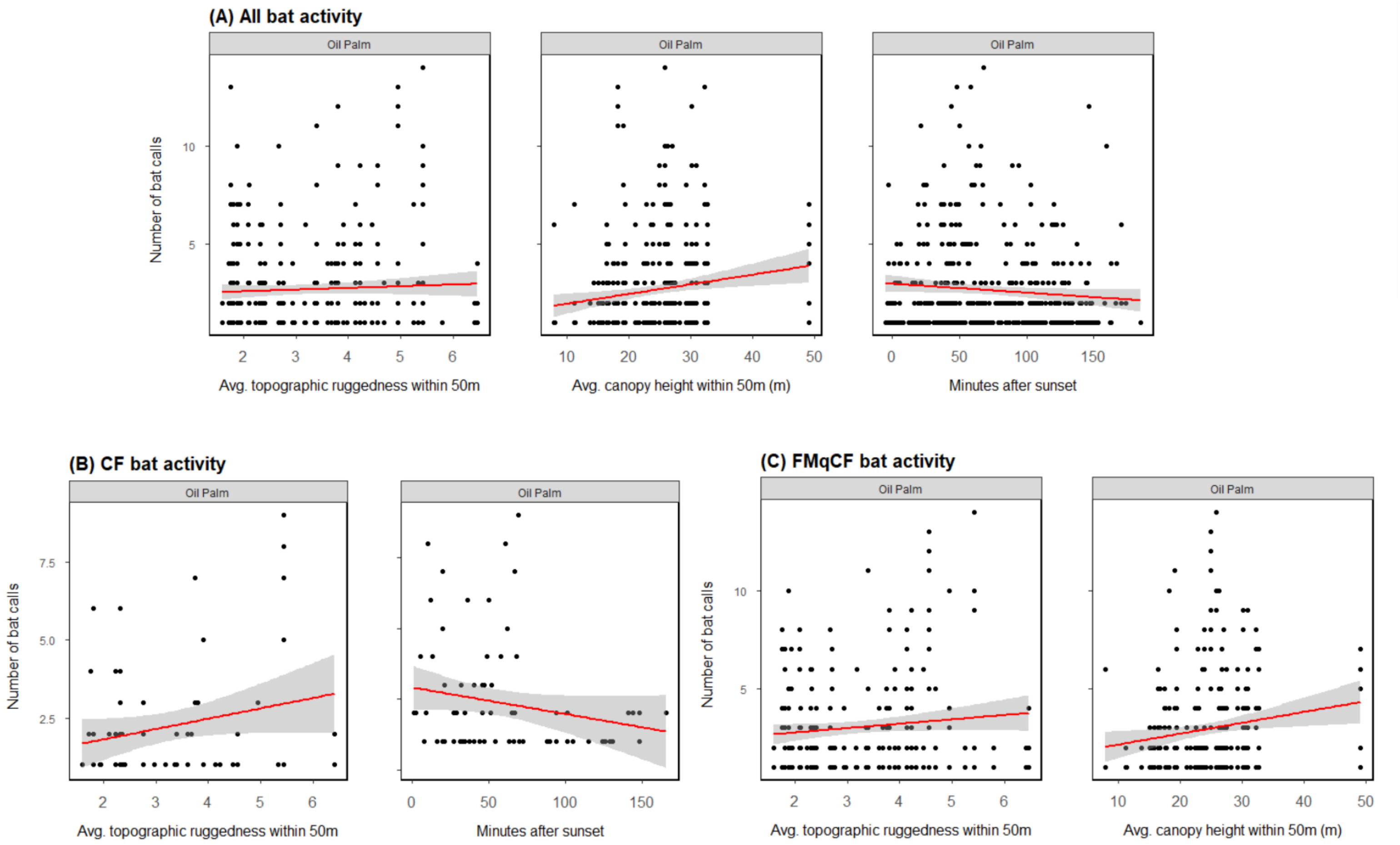


Figure 3.3 The relationship between bat activity and significant habitat covariates in oil palm

The relationship between bat activity (per sample point) and significant habitat covariates as determined by GLMM analysis for for (A) total activity, (B) CF species, and (C) FMqCF species. Shading represents 95% confidence interval.

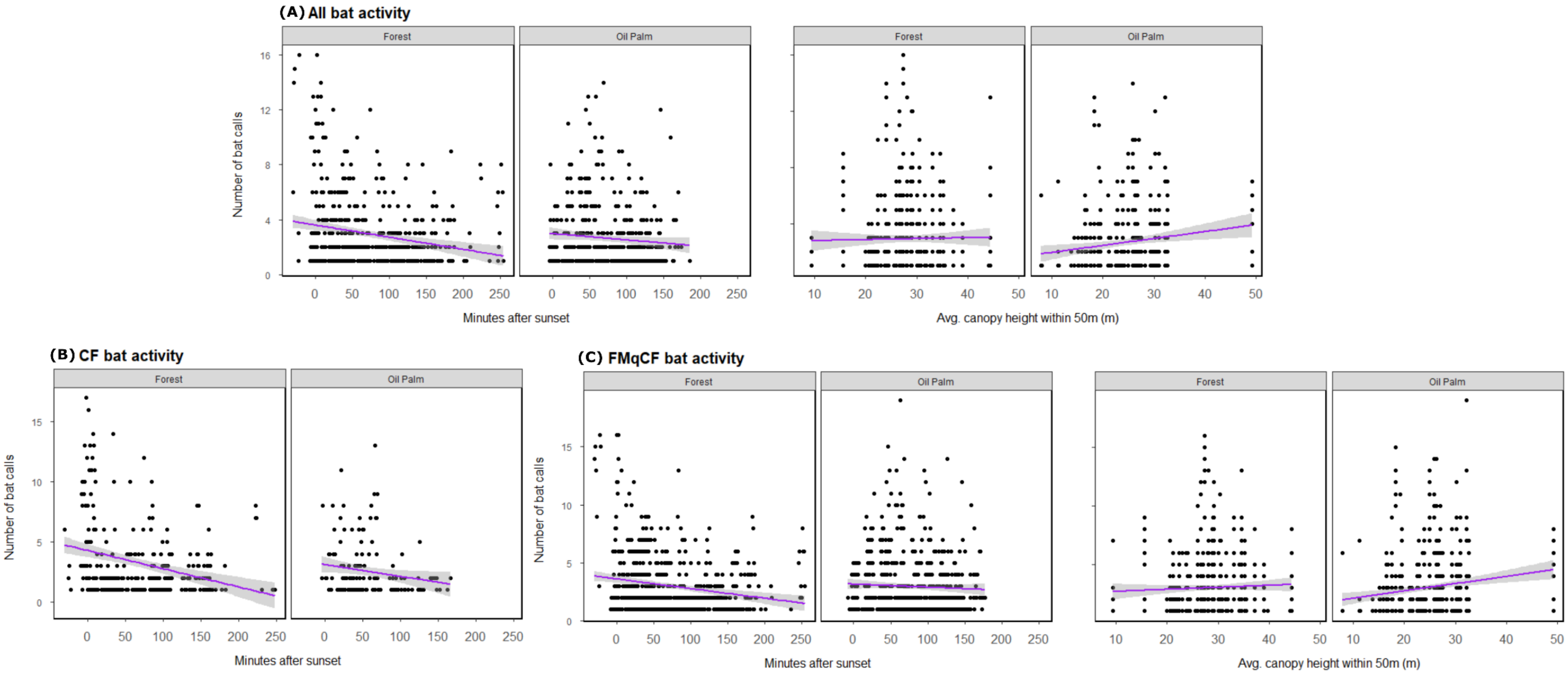


Figure 3.4 The relationship between bat activity and significant habitat covariates within both forest and oil palm

The relationship between bat activity (per sample point) and significant habitat covariates as determined by GLMM analysis for (A) total activity, (B) CF species, and (C) FMqCF species. Shading represents 95% confidence interval.

When the GLMMs were repeated for the oil palm subset of these data to explore the role of riparian reserve width, bats were again shown to be more active in oil palm reserves with a higher forest canopy, but also those with greater topographic ruggedness. Different habitat covariates were identified to be important for different call groups in oil palm reserves (Figure 3.3-3.4; Table S3.7). However, activity appeared independent of the width of oil palm reserves for all call groups assessed (Table S3.5).

3.4.2 Site occupancy

Occupancy modelling revealed a significant response to the covariates for five of the seven taxa that could be identified to the species level (Figure 3.5). Two species, *Hipposideros cervinus* and *H. galeritus*, had a greater occupancy probability at sites that were more rugged topographically. The detectability of *H. cervinus,* as well as three Rhinolophus species, also declined as time progressed after sunset. The other species exhibited no clear occupancy pattern across the wider landscape or in relation to the covariates investigated.

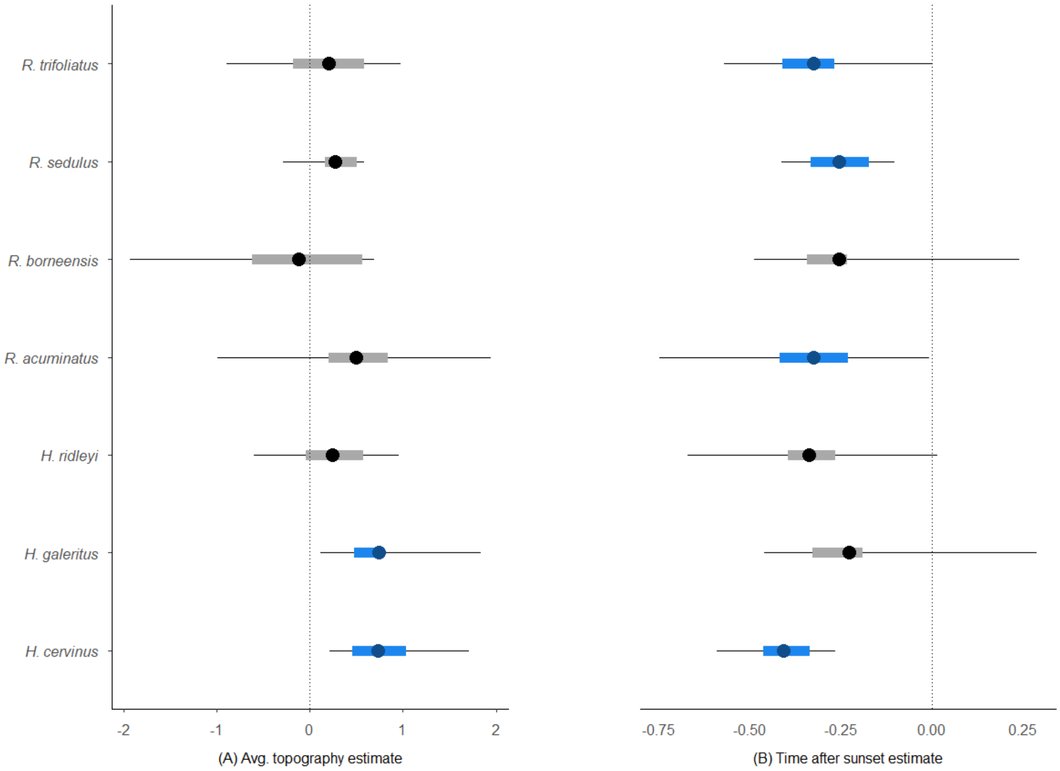


Figure 3.5 Caterpillar plots of the output from the hierarchical Bayesian multi-species occupancy model

Caterpillar plots of the output from the hierarchical Bayesian multi-species occupancy model [4] (including 95% Bayesian credible interval). Graphs show species-specific baseline occupancy estimates for environmental covariates (A) *Topography*, and detection covariate *TimeafterSunset.* Significant associations are highlighted in blue.

3.5 Discussion

We demonstrate that riparian reserves in oil palm plantations support comparable levels of insectivorous bat activity to riparian areas in neighbouring continuous forest. However, our results demonstrate variation in the importance of specific habitat covariates for different groups of bats, as defined by their echolocation call type. The timing of the survey was an important predictor in all of our analyses. In addition, bat activity was driven by canopy height (i.e. a measure of forest quality), and appeared especially important for bats using FMqCF calls. Topography also had an important influence on activity of bats using CF calls in the oil palm riparian reserves. We also note lower foraging activity recorded in oil palm compared to forested riparian sites for CF species. Overall, our analyses reveal that topography and forest quality may be more important than the width of the riparian reserve in driving bat activity alongside rivers in oil palm reserves. These findings support a previous study of captured bats that demonstrated the role of canopy height in structuring the bat community of disturbed tropical forest (Struebig *et al.* 2013).

3.5.1 Bat Activity in Riparian Reserves

Bats characterised by FMqCF calls are suited to forest edges and more open habitats, and were the most active group across all study sites, indicating that these bats are quite tolerant to extensive habitat disturbance. Indeed, Struebig *et al*. (2013) studied changes in bat assemblages across a forest disturbance gradient and found that 85% of edge/open species were captured in repeatedly logged forest. Elsewhere in the tropics, Meyer, Schwarz, and Fahr (2004) also found that FMqCF bats are more abundant/active in open, partially cluttered habitat, such as that typical of heavily logged forest. We also found that CF bats, those using flutter detection and often perceived to be specialists of cluttered forest environments, were also at high densities in riparian reserves in oil palm compared to the rivers in continuous forest sites. It is possible that CF activity in riparian reserves is inflated as bats use these landscape features to commute to other sites while avoiding the agricultural land. Differences in foraging activity between the habitat types support this assertion, but we require further recordings in adjacent oil palm estates to be more certain. Forests that are repeatedly logged also tend to comprise fewer and smaller trees with scarcer roosting opportunities in tree cavities (Struebig *et al.* 2013); making it more likely that tree-roosting CF bats (e.g. *R. sedulus, H. ridleyi;* Table S3.2) are 'passing through' rather than resident in riparian reserves. In theory, the low wing loading and aspect ratios typical of CF bats make flying long distances energetically expensive (Kingston, Jones, *et al*. 2003), although some cave-roosting CF taxa (e.g. *H. cervinus*) have narrower wings and have been recaptured in forest fragments 10-12 km from continuous forest sites (Struebig *et al.* 2008). Therefore, landscape connectivity may also have a role to play in the long-term effectiveness of riparian reserves for these species.

3.5.2 Foraging Activity in Riparian Reserves

There was high variability in foraging activity across riparian reserves as well as between forest and oil palm (Figure 3.2; Table 3.1). Importantly, we found that constant-frequency (CF) foraging activity was greater in forest than in oil palm. Oil palm plantations typically have negative consequences on invertebrates (Turner and Foster 2009), including the reduction in moth diversity (VunKhen 2006), Diptera and Hemiptera (Edwards, Magrach, *et al.* 2014), which are known to be prey for insectivorous bats in Sabah (Hemprich-Bennett *et al.* 2021). Therefore, while CF bat species may be commuting through oil palm plantations, our results suggest these habitats do not provide adequate food resources to support Rhinolophid and Hipposiderid bat species. Previous studies in this landscape indicate that riparian reserve assemblages of some invertebrate groups, namely ants and dung beetles, are somewhat similar to those in nearby forest provided riparian reserves are sufficiently wide (Luke *et al.* 2019). It is therefore feasible that the prey-base in riparian reserves may be sufficient to support some bat species at least, particularly those using FMqCF calls.

3.5.3 Importance of Habitat Structure

Changes to different structural components of riparian reserve habitats have differing implications for bat species. The long CF component of Rhinolophid and Hipposiderid bats generates a continuous input of environmental information (Russo, Ancillotto and Jones 2018). Consequently, these taxa are able to navigate more cluttered environments of lower canopy forests (Monadjem and Reside 2008), and therefore may be less affected by canopy height than bats using other call types. On the other hand, FMqCF bats are restricted to less cluttered environments; predominantly foraging along the edge, in gaps, or above forest (Altringham 1996; Table S3.2). Monadjem and Reside (2008) found the inverse relationship between canopy height and FMqCF bats in *Acacia*-dominated savanna in Swaziland, where higher activity was recorded in riparian reserves with lower canopies. This may be due to differences in structural complexity between savannah and rainforest. However, Monadjem & Reside (2008) also proposed that, unlike birds, bats discriminate between microhabitats rather than larger-scale habitat features. Topographic complexity and canopy height are known to predict the variation in local microclimate within the SAFE project (Jucker, Hardwick, *et al.* 2018), e.g. taller canopies reduce diurnal variation in microclimate. This could benefit aerial invertebrates for which abundance positively correlates with canopy height in repeatedly logged forest in Malaysia (Akutsu, Khen and Toda 2007). Species, such as the Lesser bamboo bat (*Tylonycteris pachypus*), which utilise quasi-constant frequency (FMqCF) calls may adjust the relative dominance of each component on environmental demand, e.g. reduce the FM component and increase the CF component in open spaces (Kalko and Schnitzler 1993; Altringham 1996). This adaptability allows bats that employ this echolocation technique to exploit multiple habitat types, thereby reducing their susceptibility to disturbance. As bat activity is correlated with prey availability (Kusch *et al.* 2004; Müller *et al.* 2012), higher prey abundance in higher canopy forest may be indirectly driving the increase in FMqCF bat activity observed within our study. However, it was not possible to test this hypothesis in our study.

Areas with higher topographic ruggedness are more difficult to access for logging. As such, it is likely that rugged areas provide an indirect benefit to wildlife by providing a refuge for species that require mature forest, rather than directly affecting their survival. Although riparian reserve width did not directly affect bat activity or species occurrence at river sites, forest extent can be important for maintaining a taller forest canopy. Edge effects can lead to increased mortality of mature trees, which in turn leads to a decrease in canopy height (Laurance *et al.* 2018). Therefore, riparian reserves should still be as wide as possible to reduce the long-term risk from edge-effects. The minimum 40m width threshold on each riverbank recommended for birds may be appropriate in this regard, because reserves of this size also represent greater tree biomass and higher forest canopies (Mitchell *et al.* 2018). This seems all the more important when considering that bat activity was highest in the most topographically-rugged oil palm reserve sites, which were typically rocky and steep terrain, prone to tree falls and landslides.

3.5.4 Detection

Both the GLMM and occupancy analysis highlighted the importance of considering the influence of time in acoustic survey designs. The emergence time and hours of peak activity vary between bat species and different foraging guilds (Jones and Rydell 1994). Our study confirms sampling effort may not be comparable if surveys are not controlled for temporal variation. Different call-types are also not sampled equally using acoustic monitoring due to the differences in intensity of their echolocation calls (Hayes 2000). For example, species utilising FMqCF calls, which are high intensity, were recorded abundantly, but other call-types may not have been detected as easily. Forest-dwelling species recorded by capture surveys in the forest, such as *Megaderma spasma* and *Nycteris tragata* have low intensity (MHFM) calls, whilst Vespertilionidae (*Kerivoula* and *Myotis* spp.) have low to intermediate intensity calls (FMb) (Waters and Jones 1995). More cluttered environments also reduce the detection distance for echolocation calls (Russo, Ancillotto and Jones 2018). Therefore, both their call-type and their preferred habitat are likely to contribute to these species being under-represented in our sampling. While this somewhat contravenes a key assumption of the occupancy models applied here (i.e., that detection is not influenced by abundance), we note that the conclusions of these models supported those of the GLMMs, and so we consider the outcome robust. Live trapping could be used to complement acoustic sampling to better represent these species, although we maintain that this remains logistically prohibitive in many tropical landscapes, particularly in and around rivers.

3.6 Conclusions

There is growing recognition that riparian reserves have a positive role to play in biodiversity conservation within tropical agricultural landscapes, but the evidence base remains weak (Luke *et al.* 2019). Our study confirms that these reserves in oil palm support comparable levels of bat activity to otherwise forested river sites. As palm oil production continues to rise, it is important to use evidence-based decision-making to inform more sustainable practices. With this in mind, our study highlights the importance of considering simple measures of habitat structure in the design of riparian reserves since taller vegetation supported higher levels of activity. If reflective of other taxa, this indicates that the designation and protection of riparian reserves should continue to be promoted on biodiversity grounds, with the onus on taller, more structurally complex forest, being retained alongside rivers within plantations.

3.7 Acknowledgements

We thank the Sabah Biodiversity Council, Sabah Forest Department, Yayasan Sabah, and Benta Wawasan Sdn Bhd. for research permissions and site access, and the many research assistants at the SAFE project site that helped with the bat surveys. We also thank Nick Deere for statistical advice regarding the occupancy analysis, and the Coomes Lab at the University of Cambridge for access to LiDAR information.

3.8 Supplementary information

Table S3.1 Habitat characteristics of riparian sites

Site description, habitat characteristics of riparian sites within 50 m (± standard deviation)

River	Canopy height (m)	Reserve width (m)	Topography (m)
Forested rivers			
0	36.248 (±3.113)	-	6.307 (±1.545)
5	30.532 (±2.167)	-	7.158 (±1.115)
15	22.816 (±7.076)	-	6.120 (±1.201)
30	30.229 (±3.758)	-	5.016 (±0.857)
60	29.465 (±4.913)	-	4.217 (±1.365)
120	23.406 (±2.486)	-	3.058 (±1.582)
RVJR	38.596 (±6.477)	-	4.671 (±1.887)
RLFE	34.356 (±3.296)	-	4.826(±2.175)
Oil palm riparian reserves			
RR2	15.765 (±4.344)	20.19 (±16.80)	3.530 (±1.557)
RR3	24.023 (±3.619)	61.88 (±18.34)	3.424 (±1.427)
RR7	28.762 (±2.819)	48.06 (±8.37)	2.543 (±0.805)
RR12	23.230 (±4.705)	42.69 (±21.16)	3.898 (±0.869)
RR14	31.370 (±13.349)	64.63 (±26.64)	3.044 (±1.265)
RR16	22.110 (±3.984)	74.88 (±41.39)	3.719 (±1.539)
OP2	15.606 (± 1.744)	0 (±0)	2.769 (±0.706)

Table S3.2 Roost preference and foraging guild for insectivorous bat species present in Borneo

Roost preference and foraging guild for insectivorous bat species present in Borneo. Call types represent the four call types used in the study: constant frequency (CF), frequency modulating quasi-constant frequency or quasi-constant frequency (FMqCF), broad-band frequency modulating (FMb), and multi-harmonic frequency modulating (MHFM). Species utilizing CF, FMb, or MHFM calls are classified as narrow-space foragers.

Family	Species	Primary roost types	Foraging Guild	Call type
Emballonuridae	*Taphozous melanopogon*	Caves, man-made structures	Open space	FMqCF
	Taphozous longimanus	Caves, man-made structures	Open space	FMqCF
	Taphozous theobaldi	Caves, man-made structures	Open space	FMqCF
	Saccolaimus saccolaimus	Caves, man-made structures	Open space	FMqCF
	Emballonura monticola	Rock crevices, fallen trees, man-made structures	Edges	FMqCF
	Emballonura alecto	Rock crevices, fallen trees, man-made structures	Edges	FMqCF
Nycteridae				
	Nycteris tragata	Tree hollows, fallen trees	Narrow space (forest interior)	MHFM
Megadermatidae				
	Megaderma spasma	Tree hollows, fallen trees, caves	Narrow space (forest interior)	MHFM
Rhinolophidae	*Rhinolophus luctus*	Tree hollows, fallen trees, caves	Narrow space (forest interior)	CF
	Rhinolophus sedulus	Tree hollows, fallen trees	Narrow space (forest interior)	CF
	Rhinolophus trifoliatus	Foliage	Narrow space (forest interior)	CF
	Rhinolophus francisi	Unknown, but likely trees	Narrow space (forest interior)	CF
	Rhinolophus philippinensis	Caves	Narrow space (forest interior)	CF
	Rhinolophus creaghi	Caves	Narrow space (forest interior)	CF

	Rhinolophus acuminatus	Caves	Narrow space (forest interior)	CF
	Rhinolophus arcuatus	Caves	Narrow space (forest interior)	CF
	Rhinolophus affinis	Caves	Narrow space (forest interior)	CF
	Rhinolophus pusillus	Caves	Narrow space (forest interior)	CF
	Rhinolophus borneensis	Caves, tree hollows	Narrow space (forest interior)	CF
Hipposideridae	*Hipposideros diadema*	Caves	Narrow space (forest interior)	CF
	Hipposideros galeritus	Caves, rock crevices	Narrow space (forest interior)	CF
	Hipposideros cineraceus	Caves	Narrow space (forest interior)	CF
	Hipposideros bicolor	Caves	Narrow space (forest interior)	CF
	Hipposideros ridleyi	Tree hollows, fallen trees	Narrow space (forest interior)	CF
	Hipposideros doriae	Tree hollows, fallen trees	Narrow space (forest interior)	CF
	Hipposideros coxi	Caves	Narrow space (forest interior)	CF
	Hipposideros larvatus	Caves	Narrow space (forest interior)	CF
	Hipposideros cervinus	Caves, rock crevices	Narrow space (forest interior)	CF
	Hipposideros dyacorum	Caves, tree hollows	Narrow space (forest interior)	CF
	Hipposideros ater	Caves	Narrow space (forest interior)	CF
	Coelops robinsoni	Caves, tree hollows	Narrow space (forest interior)	CF
Miniopteridae				
	Miniopterus magnater	Caves	Open space	FMqCF
	Miniopterus schreibersii	Caves	Open space	FMqCF
	Miniopterus pusillus	Caves	Open space	FMqCF
	Miniopterus medius	Caves	Open space	FMqCF
	Miniopterus australis	Caves	Open space	FMqCF
	Miniopterus paululus	Caves	Open space	FMqCF

Vespertilionidae				
Kerivoulinae				
	Kerivoula intermedia	Foliage	Narrow space (forest interior)	FMb
	Kerivoula minuta	Foliage	Narrow space (forest interior)	FMb
	Kerivoula hardwickii	Foliage	Narrow space (forest interior)	FMb
	Kerivoula lenis	Foliage	Narrow space (forest interior)	FMb
	Kerivoula whiteheadi	Foliage	Narrow space (forest interior)	FMb
	Kerivoula pellucida	Foliage	Narrow space (forest interior)	FMb
	Kerivoula papiliosa	Foliage	Narrow space (forest interior)	FMb
	Phoniscus jagorii	Foliage	Narrow space (forest interior)	FMb
	Phoniscus atrox	Foliage	Narrow space (forest interior)	FMb
Murininae				
	Murina rozendaali	Foliage	Narrow space (forest interior)	FMb
	Murina suilla	Foliage	Narrow space (forest interior)	FMb
	Murina aenea	Foliage	Narrow space (forest interior)	FMb
	Murina peninsularis	Foliage	Narrow space (forest interior)	FMb
	Harpiocephalus harpia	Foliage	Narrow space (forest interior)	FMb
Vespertilioninae				
	Pipistrellus javanicus	Man-made structures, caves	Open space / edge	FMqCF
	Pipistrellus tenuis	Man-made structures, caves	Open space / edge	FMqCF
	Falsistrellus petersi	Man-made structures, caves	Open space / edge	FMqCF
	Pipistrellus stenopterus	Man-made structures, caves	Open space / edge	FMqCF
	Arielulus cuprosus	Man-made structures, caves	Open space / edge	FMqCF
	Pipistrellus ceylonicus	Man-made structures, caves	Open space / edge	FMqCF
	Hypsugo kitcheneri	Man-made structures, caves	Open space / edge	FMqCF

	Hypsugo vordermanni	Man-made structures, caves	Open space / edge	FMqCF
	Hypsugo imbricatus	Man-made structures, caves	Open space / edge	FMqCF
	Tylonycteris pachypus	Bamboo	Edge	FMqCF
	Tylonycteris robustula	Bamboo	Edge	FMqCF
	Glischropus tylopus	Bamboo	Edge	FMqCF
	Philetor brachypterus	Man-made structures, caves	Open space / edge	FMqCF
	Scotophilus collinus	Man-made structures, caves	Open space / edge	FMqCF
	Hesperoptenus doriae	Man-made structures, caves	Open space / edge	FMqCF
	Hesperoptenus blanfordi	Man-made structures, caves	Open space / edge	FMqCF
	Hesperoptenus tomesi	Man-made structures, caves	Open space / edge	FMqCF
	Myotis ater	Caves	Edge	FMb
	Myotis horsfieldii	Man-made structures, caves, foliage	Edge	FMb
	Myotis muricola	Caves, foliage	Edge	FMb
	Myotis ridleyi	Foliage	Edge	FMb
	Myotis gomantongensis	Caves	Edge	FMb
	Myotis adversus	Caves	Edge	FMb
	Myotis borneensis	Caves	Edge	FMb
	Myotis hasseltii	Caves	Edge	FMb
	Myotis siligorensis	Caves	Edge	FMb
	Myotis macrotarsus	Caves	Edge	FMb
Molossidae				
	Mops Mops	Caves, tree hollows	Open/edge space	FMqCF
	Chaerephon plicatus	Caves	Open/edge space	FMqCF
	Cheiromeles torquatus	Caves	Open/edge space	FMqCF

Table S3.3 Equations used for the final Generalised linear mixed effect models

Equations used for the final Generalised linear mixed effect models for total activity (all call groups, CF, and FMqCF) in both habitats (eqn 1:3) and in oil palm reserves (eqn 4:7) compared for best fit - where NCalls_{ij} is the jth observation in River_i, and i=1,...,8, and River_i is the random intercept, which is assumed to be normally distributed with mean 0 and variance σ2.

$$\begin{aligned}&\mathrm{NCalls}_{ij} \sim \mathrm{Poisson}(\mu_{ij})\\&\mu_{ij} = Topography_{ij} + CanopyHeight_{ij} + TimeafterSunset_{ij} + HabitatType_{ij} + River_i\\&River_i \sim N(0, \sigma 2)\end{aligned} \qquad \text{(eqn 1)}$$

$$\begin{aligned}&\mathrm{NCalls}_{ij} \sim \mathrm{Poisson}(\mu_{ij})\\&\mu_{ij} = CanopyHeight_{ij} + TimeafterSunset_{ij} + River_i\\&River_i \sim N(0, \sigma 2)\end{aligned} \qquad \text{(eqn 2)}$$

$$\begin{aligned}&\mathrm{NCalls}_{ij} \sim \mathrm{Poisson}(\mu_{ij})\\&\mu_{ij} = TimeafterSunset_{ij} + River_i\\&River_i \sim N(0, \sigma 2)\end{aligned} \qquad \text{(eqn 3)}$$

$$\begin{aligned}&\mathrm{NCalls}_{ij} \sim \mathrm{Poisson}(\mu_{ij})\\&\mu_{ij} = Topography_{ij} + CanopyHeight_{ij} + TimeafterSunset_{ij} + RiparianWidth_{ij} + River_i\\&River_i \sim N(0, \sigma 2)\end{aligned} \qquad \text{(eqn 4)}$$

$$\begin{aligned}&\mathrm{NCalls}_{ij} \sim \mathrm{Poisson}(\mu_{ij})\\&\mu_{ij} = Topography_{ij} + CanopyHeight_{ij} + TimeafterSunset_{ij} + River_i\\&River_i \sim N(0, \sigma 2)\end{aligned} \qquad \text{(eqn 5)}$$

$$\begin{aligned}&\mathrm{NCalls}_{ij} \sim \mathrm{Poisson}(\mu_{ij})\\&\mu_{ij} = Topography_{ij} + TimeafterSunset_{ij} + River_i\\&River_i \sim N(0, \sigma 2)\end{aligned} \qquad \text{(eqn 6)}$$

$$\begin{aligned}&\mathrm{NCalls}_{ij} \sim \mathrm{Poisson}(\mu_{ij})\\&\mu_{ij} = CanopyHeight_{ij} + TimeafterSunset_{ij} + River_i\\&River_i \sim N(0, \sigma 2)\end{aligned} \qquad \text{(eqn 7)}$$

Table S3.4 GLMM model comparisons for each group

GLMM model comparisons for each group. The top five models are presented based on AIC & BIC using all covariates listed in Equation (1) for both habitats and Equation (6) for oil palm exclusively.

Species	Model	AIC	BIC	Loglik	Deviance	df	disp.
Forest & Oil Palm							
All	CH+Time	7005.3	7026.8	-3498.6	6997.3	1599	1.295
	Rug+CH+Time	7005.3	7032.2	-3497.7	6995.3	1598	1.294
	Rug+CH+Time+Hab	7006.6	7038.9	-3497.3	6994.6	1597	1.294
	CH+Time+Hab	7007.3	7033.9	-3498.5	6997.0	1598	1.295
	Rug+Time	7036.0	7057.5	-3514.0	7028.0	1599	1.302
CF	Time	1700.0	1711.9	-847.0	1694.0	383	1.287
	Time+Hab	1701.3	1717.1	-846.6	1693.3	382	1.287
	Rug+Time+Hab	1701.3	1723.0	-846.6	1693.2	381	1.287
	Rug+Time	1702.0	1717.8	-847.0	1694.0	382	1.287
	CH+Time	1702.0	1717.9	-847.0	1694.0	382	1.287
FM	CH+Time+Hab	4572.3	4596.9	-2281.1	4562.3	1023	1.291
	CH+Time	4572.5	4592.3	-2282.3	4564.5	1024	1.291
	Rug+CH+Time+Hab	4573.4	4603.1	-2280.7	4561.4	1022	1.291
	Rug+CH+Time	4574.1	4598.8	-2282.1	4564.1	1023	1.291
	Time	4611.6	4626.4	-2302.8	4605.6	1025	1.308
Oil Palm exclusively							
All	CH+Time+Rip	3252.8	3275.9	-1621.4	3242.8	748	1.284
	Rug+CH+Time+Rip	3254.2	3282.0	-1621.1	3242.2	747	1.283
	Rug+CH	3256.7	3275.2	-1624.3	3248.7	749	1.286
	Rug+CH+Rip	3258.4	3281.6	-1624.2	3248.4	748	1.286

	CH+Time	3260.5	3279.0	-1626.2	3252.5	749	1.289
CF	Rug+Time+CH	525.8	540.2	-257.9	515.8	126	1.147
	Rug+Time	526.1	537.6	-259.1	518.1	127	1.155
	Rug+Time+Rip	526.6	540.9	-258.3	516.6	126	1.147
	Rug+Time+CH+Rip	527.8	545.0	-257.9	515.8	125	1.145
	Rug	531.9	540.5	-263.0	525.9	128	1.190
FMqCF	Rug+CH+Time+Rip	2315.1	2340.5	-1151.6	2303.1	499	1.332
	Rug+CH+Rip	2315.9	2337.0	-1153.0	2305.9	500	1.334
	Rug+CH+Time	2316.8	2337.9	-1153.4	2306.8	500	1.336
	CH+Rug	2317.4	2334.3	-1154.7	2309.4	501	1.338
	CH+ Time	2318.4	2335.3	-1155.2	2310.4	501	1.339

Rug = *Topography*; CH =*CanopyHeight;* Time = *TimeafterSunset;* Rip =*RiparianWidth;* Hab =*HabitatType;* CF = *constant frequency call type*; FMqCF = *frequency modulated quasi-constant frequency call type*

Table S3.5 Likelihood ratio results for comparisons between top GLMM models for each group

Likelihood ratio results for comparisons between top GLMM models for each group. Significant *P*-values indicate the inclusion of the covariate had a significant impact on the model.

	Covariate	Chi-Squ	DF	*P* value	
Forest & Oil Palm					
All calls	*HabitatType*	0.7328	1	0.3920	
	Topography	2.3826	1	0.1227	
	CanopyHeight	33.346	1	<0.001	***
	TimeafterSunset	108.23	1	<0.001	***
Constant frequency (CF) calls	*HabitatType*	0.741	1	0.3894	
	Topography	0.000	1	0.9960	
	CanopyHeight	0.015	1	0.9015	
	TimeafterSunset	72.537	1	<0.001	***
Frequency-modulated quasi-constant frequency (FMqCF) calls	*HabitatType*	2.663	1	0.1027	
	Topography	0.824	1	0.3641	
	CanopyHeight	43.087	1	<0.001	***
	TimeafterSunset	47.798	1	<0.001	***
Oil Palm exclusively					
All calls	*RiparianWidth*	0.509	1	0.4756	
	Topography	10.236	1	0.0013	**
	CanopyHeight	16.294	1	<0.001	***
	TimeafterSunset	6.200	1	0.0128	*
Constant frequency (CF) calls	*RiparianWidth*	0.083	1	0.7739	
	Topography	10.143	1	0.0014	**
	CanopyHeight	0.780	1	0.3718	
	TimeafterSunset	7.083	1	0.0078	**
Frequency-modulated quasi-constant frequency (FMqCF) calls	*RiparianWidth*	3.685	1	0.0549	
	Topography	6.009	1	0.0142	*
	CanopyHeight	20.740	1	<0.001	***
	TimeafterSunset	2.793	1	0.0947	

Significance values *p<0.05; **p<0.01; ***p<0.001

Table S3.6 Measures for the predictive quality of each hierarchical Bayesian multispecies occupancy model

Measures for the predictive quality of each hierarchical Bayesian multispecies occupancy model. *MinutesafterSunset* was included as a detection covariate, as well as *River* as a random effort, in all models. **Bold** indicates model of best fit.

Model [no.]		Dev.	Lppd	WAIC	*p*D	ΔAIC
[4]	***Topography***	**1168.993**	**1127.443**	**1213.404**	**42.980**	**0.00**
[5]	*CanopyHeight* + *Topography*	1160.779	1119.325	1217.353	49.014	3.949
[2]	*HabitatType*	1167.501	1130.470	1217.777	43.654	4.373
[3]	*CanopyHeight*	1168.993	1128.942	1221.665	46.362	8.261
[1]	All co-variates	1167.993	1127.276	1225.146	48.935	11.742

*Lppd = log pointwise predictive density; *p*D = effective number of parameters; CPO = conditional predictive ordinance (Formulas from Gelman *et al.* (2013)

Table S3.7 Estimated regression parameters, standard errors, z-values and P-values for the Poisson GLMM

Estimated regression parameters, standard errors, *z*-values and *P*-values for the Poisson GLMM presented in eqn (2:3) for forest and oil, and eqn (5:7) for oil palm reserves. The estimated variance for River is 0.024 (± 0. 0.154).

		Estimate	Std. error	*z* value	*P* value	
Forest & Oil Palm						
All species	Intercept	0.941	0.509	18.500	< 0.001	***
	CanopyHeight	-0.197	0.185	-10.660	< 0.001	***
	TimeafterSunset	0.118	0.021	5.760	< 0.001	***
Constant frequency species	Intercept	0.824	0.107	7.723	< 0.001	***
	TimeafterSunset	-0.339	0.039	-8.743	< 0.001	***
Frequency-modulated quasi-constant frequency species	Intercept	1.007	0.062	16.225	< 0.001	***
	CanopyHeight	0.157	0.025	6.259	< 0.001	***
	TimeafterSunset	-0.157	0.022	-7.048	< 0.001	***
Oil Palm exclusively						
All species	Intercept	1.053	0.721	14.614	< 0.001	***
	Topography	0.124	0.040	3.125	<0.001	***
	CanopyHeight	0.156	0.029	5.312	< 0.001	***
	TimeafterSunset	-0.072	0.030	-2.434	0.058	
Constant frequency species	Intercept	0.921	0.134	6.850	< 0.001	***
	Topography	0.271	0.090	3.023	0.018	*
	TimeafterSunset	-0.262	0.096	-2.733	0.007	**
Frequency-modulated quasi-constant frequency species	Intercept	1.156	0.084	13.763	< 0.001	***
	CanopyHeight	0.161	0.032	5.039	< 0.001	***
	TimeafterSunset	0.092	0.046	2.011	0.125	

Significance values *p<0.05; **p<0.01; ***p<0.001

Chapter 4 A machine learning framework to classify Southeast Asia echolocating bats

Natalie Yoh*[1], Tigga Kingston[2,3], Ellen McArthur[4], Oliver E. Aylen[5], Joe Chun-Chia Huang[6], Emy Ritta Jinggong[4], Faisal Ali Anwarali Khan[4], Benjamin P. Y.H. Lee[7], Simon L. Mitchell[1], Jake E. Bicknell[1], Matthew J. Struebig1[1]

1. Durrell Institute of Conservation and Ecology (DICE), School of Anthropology and Conservation, University of Kent, Canterbury, United Kingdom
2. Department of Biological Sciences, Texas Tech University, Lubbock, Texas, USA
3. Southeast Asian Bat Conservation Research Unit, Lubbock City, Texas, USA
4. Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia
5. University of Otago, Department of Zoology, Dunedin, Otago, New Zealand
6. Taiwan Forestry Research Institute, Taipei, Taiwan
7. Wildlife Management Division, National Parks Board, Singapore

*Correspondence: tallyyoh@gmail.com

This chapter is based on the following publication:

Yoh, N., *et al*. (2022). A machine learning framework to classify Southeast Asian echolocating bats. Ecological Indicators, 136. doi: 10.1016/j.ecolind.2022.108696

Author Contributions: Natalie Yoh: Conceptualization, Data Curation, Methodology, Software, Writing - Original Draft, Tigga Kingston: Conceptualization, Resources, Writing - Review & Editing, Ellen McArthur: Resources, Writing - Review & Editing, Oliver E. Aylen: Resources, Joe Chun-Chia Huang: Resources, Writing - Review & Editing, Emy Ritta Jinggong: Resources, Faisal Ali Anwarali Khan: Resources, Writing - Review & Editing, Benjamin P. Y.-H. Lee: Resources, Simon L. Mitchell: Methodology, Data curation, Resources, Jake E. Bicknell: Supervision, Writing - Review & Editing, Matthew J. Struebig: Conceptualization, Supervision, Resources, Writing - Review & Editing

4.1 Abstract

Bats comprise a quarter of all mammal species, provide key ecosystem services, and serve as effective bioindicators. Automated methods for classifying echolocation calls of free-flying bats are useful for monitoring but are not widely used in the tropics. This is particularly problematic in Southeast Asia, which supports more than 388 bat species. Here, sparse reference call databases and significant overlap among species call characteristics makes the development of automated processing methods complex. To address this, we outline a semi-automated framework for classifying bat calls in Southeast Asia and demonstrate how this can reliably speed up manual data processing. We implemented the framework to develop a classifier for the bats of Borneo and tested this at a landscape in Sabah. Borneo has a relatively well-described bat fauna, including reference calls for 52% of all 81 known echolocating species on the island. We applied machine learning to classify calls into one of four call types that serve as indicators of dominant ecological ensembles: frequency-modulated (FM; forest-specialists), constant frequency (CF; forest-specialists and edge/gap foragers), quasi-constant frequency (QCF; edge/gap foragers), and frequency-modulated quasi constant frequency (FMqCF; edge/gap and open-space foragers) calls. Where possible, we further identified calls to species/sonotype. Each classification is provided with a confidence value and a recommended threshold for manual verification. Of the 245,991 calls recorded in our test landscape, 85% were correctly identified to call type and only 10% needed manual verification for three of the call types. The classifier was most successful at classifying CF calls, reducing the volume of calls to be manually verified by over 95% for three common species. The most difficult bats to classify were those with FMqCF calls, with only a 52% reduction in files. Our framework allows users to rapidly filter acoustic files for common species and isolate files of interest, cutting the total volume of data to be processed by 86%. This provides an alternative method where species-specific classifiers are not yet feasible and enables researchers to expand non-invasive monitoring of bat species. Notably, this approach incorporates aerial insectivorous ensembles that are regularly absent from field datasets despite being important components of the bat community, thus improving our capacity to monitor bats remotely in tropical landscapes.

Keywords: Acoustic monitoring, Chiroptera, echolocation, Southeast Asia, machine learning, supervised algorithm

4.2 Introduction

Biodiversity monitoring is critical to informing conservation practice. Still, multi-taxon assessments are frequently constrained by resources, time, and survey bias (Gardner *et al.* 2008). Focusing survey efforts on biological indicators is one way to ameliorate these challenges so long as these species or groups reflect the needs of others in the system, particularly in the way they respond to environmental change and other conservation threats.

Bats can be effective bioindicators of ecosystem health (Jones *et al*., 2009). There is a growing literature on responses of bat assemblages to various anthropogenic pressures such as forest loss and fragmentation (Gardner *et al.* 2008; Meyer, Struebig and Willig 2016; Park 2015). The use of multiple sampling techniques provides the best way to monitor the whole bat assemblage, including live-capture methods and acoustic monitoring (Russo, Ancillotto and Jones 2018). However, in the tropics, monitoring has been largely confined to live-capture methods (e.g., harp traps and mist-nets). Whilst they can be highly effective at monitoring bat species in the forest understory (Tanshi and Kingston 2021), these methods can be labour intensive, invasive, and are often logistically challenging (Fisher-Phelps *et al.* 2017). Moreover, insectivorous bat species that forage in open spaces above forests, or around forest gaps or edges, are difficult to catch using these methods. Therefore, key components of bat assemblages that could serve as potential indicator taxa are often absent from or are underrepresented in field datasets when only one approach is used (Kingston, 2013, 2016).

Acoustic monitoring, whereby call signatures of biological sounds are compared to reference libraries, offers an alternative to bat capture techniques (Walters *et al.* 2013). Passive acoustic monitoring (PAM) techniques can be used to quantify a range of ecological metrics, including species diversity (López-Baucells *et al.* 2019), animal movement and activity (Furmankiewicz and Kucharska 2009), population dynamics (particularly for roost monitoring; Revilla-Martín *et al.*, 2020), and responses to anthropogenic change (Meyer, Struebig and Willig 2016; Yoh, Azhar, *et al.* 2020). It is used to monitor a range of terrestrial species including birds, amphibians, insects, and terrestrial mammals, but is most extensively applied to insectivorous bat monitoring (Sugai *et al.* 2019).

Two major shortfalls of PAM are the time required to process the large volume of acoustic data generated, as well as the availability of reference libraries (Gibb *et al.* 2019). Individual echolocating bats adjust their call structure in response to different habitats, foraging space, and stages of prey pursuit (Kalko and Schnitzler 1993). This within-individual and within-species variability are coupled with morphological, phylogenetic, and habitat constraints on adaptive call structure, and thus many species calls overlap in structure (Pham *et al.* 2021; Russo, Ancillotto and Jones 2018; Walters *et al.* 2013). There are likewise technical challenges when using PAM for bats compared to other taxa. Most terrestrial mammal species produce infrasonic vocalisations (< 20 kHz) whereas most bats produce ultrasonic calls (> 20 kHz) which can be over 200 kHz (Fenton and Bell 1981). As frequency increases, so too does atmospheric attenuation, which can lead to incomplete sampling of the call structure (loss of higher frequencies) and reduce detection distances. Both can lead to a sampling bias in favour of low-frequency species (Lawrence and Simmons 1982; Russo, Ancillotto and Jones 2018). Bat species that do not rely on echolocation for foraging cannot be monitored using acoustic surveys (Russo, Ancillotto and Jones 2018). In Borneo, this includes 18% of bat fauna (family Pteropodidae*;* 18 species from 11 genera; Phillipps & Phillipps, 2016). For these taxa live-capture methods remain an essential monitoring tool.

To help mitigate some of the challenges associated with monitoring bats acoustically, there has been a rise in the development of automated or semi-automated classifiers (Kwok 2019; Tabak *et al.* 2019). Still, between 1990 and 2018, just ca. 19% of studies based on PAM in terrestrial environments processed their data using fully automated classifiers, and a further 15% used a semi-automated classifier in combination with manual identification (Sugai *et al.* 2019). Such classifiers, built using supervised machine learning algorithms, can determine classifications through pattern recognition of call characteristics, and provide a quick and repeatable method of distinguishing between species calls. Classifiers can therefore help reduce the processing burden of high volumes of acoustic recordings (Valletta *et al.* 2017).

Global attempts to assess how bats are impacted by environmental change using acoustic monitoring networks (e.g., iBats; Jones *et al.* 2013) remain constrained by the availability of reference calls needed to encapsulate call plasticity within and across species when training these algorithms. As such, acoustic classifiers are largely concentrated in Europe (e.g., Parsons & Jones, 2000), North America (e.g., Clement *et al*., 2014), and Japan (e.g., Kobayashi *et al.*, 2021), where bat assemblages comprise relatively few species that are intensively studied compared to other regions of the world. This therefore hinders our ability to monitor bats effectively in species-rich areas, where the costs of establishing local call reference libraries are prohibitively high (Kershenbaum *et al.* 2016). Consequently, there remain important gaps in our understanding of how large numbers of bat species respond to environmental changes across the Central African and Asian tropics in particular (Meyer, Struebig and Willig 2016).

Several developments in recent years show promise for the automated classification of tropical bat calls. Software such as *Waveman* (Chen *et al*., 2020) demonstrates machine learning can be a viable technique for differentiating calls. However, attempts to classify species from Thailand and Vietnam highlight how limited training data can restrict confidence in identifications (Hughes *et al.* 2011; Pham *et al.* 2021). This illustrates the importance of manual post-validation when using automated classifiers, in order to minimise the risk of incorrect identifications (Russo and Voigt, 2016). Recently, López-Baucells *et al*. (2019) proposed a semi-automated approach that combines automated classification with targeted post-validation of files. This provides a low risk, efficient method for automating the processing of bat calls in areas with limited reference call libraries.

Southeast Asia is a global hotspot for bat diversity with at least 388 species (Simmons and Cirranello 2021b). However, this diversity is highly threatened by rapid land-use changes, with at least 23% of Southeast Asia's bats predicted to be extirpated by 2100 (Lane, Kingston and Lee 2006). So far, bat research has been dominated by live-capture studies, and PAM is rarely applied. Bat research is also spatially biased (Fisher-Phelps *et al.* 2017), and as a result, there remain major gaps in our understanding of species' responses to anthropogenic threats (Kingston, 2010; Pham *et al.*, 2021). This creates a circular problem whereby the lack of tools limits research capacity, which further restricts the ability to improve tools. Meanwhile, the International Union for the Conservation of Nature (IUCN) reports that at least 97 of these insectivorous bat species are declining (IUCN 2021).

A way to fast-track the development of bat call classifiers for Southeast Asia is to shift the emphasis from species-level identification to identifying call type. Insectivorous bats can be divided into three broad foraging ensembles defined by the acoustic and flight challenges of foraging in different environments (Denzinger & Schnitzler, 2013; Schnitzler & Kalko, 2001): forest interior, edge/gap, and open space.

Bats foraging in the forest interior must distinguish target echoes of potential prey from those coming from surrounding vegetation. Bats in the families Hipposideridae and Rhinolophidae have evolved a strategy that enables them to detect insect wing movement against static vegetation. Sound energy is focused into a very narrow range of frequencies, almost a single "note". These are referred to as constant frequency or CF calls (Denzinger & Schnitzler, 2013; Schnitzler & Kalko, 2001). As an alternative strategy, other species foraging in the forest interior (mainly within the families Vespertilionidae, Nycteridae, Megadermatidae) use low-intensity calls that cover a wide range of frequencies in a short time, these can be a single harmonic (frequency-modulated calls; FM) or comprise multiple harmonics that sweep down (multi-harmonic FM sweeps; Denzinger & Schnitzler, 2013; Schnitzler & Kalko, 2001).

Edge/gap foragers (including Emballonuridae and Vespertilionidae) often represent the greatest number of calls recorded during acoustic surveys and are adapted to foraging in areas near background vegetation, such as the forest edge, where the background can be used for orientation but can mask the presence of insects (Schnitzler, Moss and Denzinger 2003). The calls of these species are typically mid to high intensity dominated by a narrow-band FM component followed by a short quasi-CF component (FMqCF) and are often highly flexible, which allows these bats to maximise their sensory input for a range of environmental conditions and to minimise masking effects. Whereas most species calls begin with a narrow-band FM component followed by a quasi-CF component, several Emballonurid species (e.g., *Emballonura monticola*) use calls characterised by a downward sweeping FM or quasi-CF component (QCF) to hawk insects in less cluttered spaces (Pottie *et al.* 2005).

Open-space foragers hawk airborne prey across large, open spaces, such as above the forest canopy (Denzinger and Schnitzler 2013). Their prey is more widely dispersed than within the forest interior, and consequently, they need an increased range of prey detection. Therefore, they use narrowband, high intensity calls with a long call duration and typically emit frequencies below 30 kHz (Denzinger and Schnitzler 2013; Jung, Molinari and Kalko 2014). They also use FMqCF calls, consisting of a long quasi-CF component (8-25 ms). In Southeast Asia, this includes species from the families Molossidae and Emballonuridae. Although certain bats can adapt their foraging strategy to different environments, there are limits to this behavioural flexibility. Therefore, echolocating bats are assigned to a foraging ensemble according to which habitat their echolocation call design is best adapted to (Denzinger and Schnitzler 2013; Siemers and Schnitzler 2004).

Here we present a semi-automated method for identifying echolocation calls of bats in Southeast Asia. We developed a rapid, autonomous framework for assigning echolocation calls to species or into call types/sonotypes representative of different ecological ensembles present in the region (Fig. 4.1-4.2). These call types/sonotypes serve as indicators for lesser-known or less conspicuous species. We apply the technique to Borneo’s bat fauna, which is relatively well described taxonomically (Simmons and Cirranello 2021b). We emphasise how our framework can be applied elsewhere in Southeast Asia with comparable bat assemblage composition as reference calls become available. By applying this framework to acoustic datasets, more comprehensive information can be generated regarding how tropical bats utilise landscapes and respond to environmental change

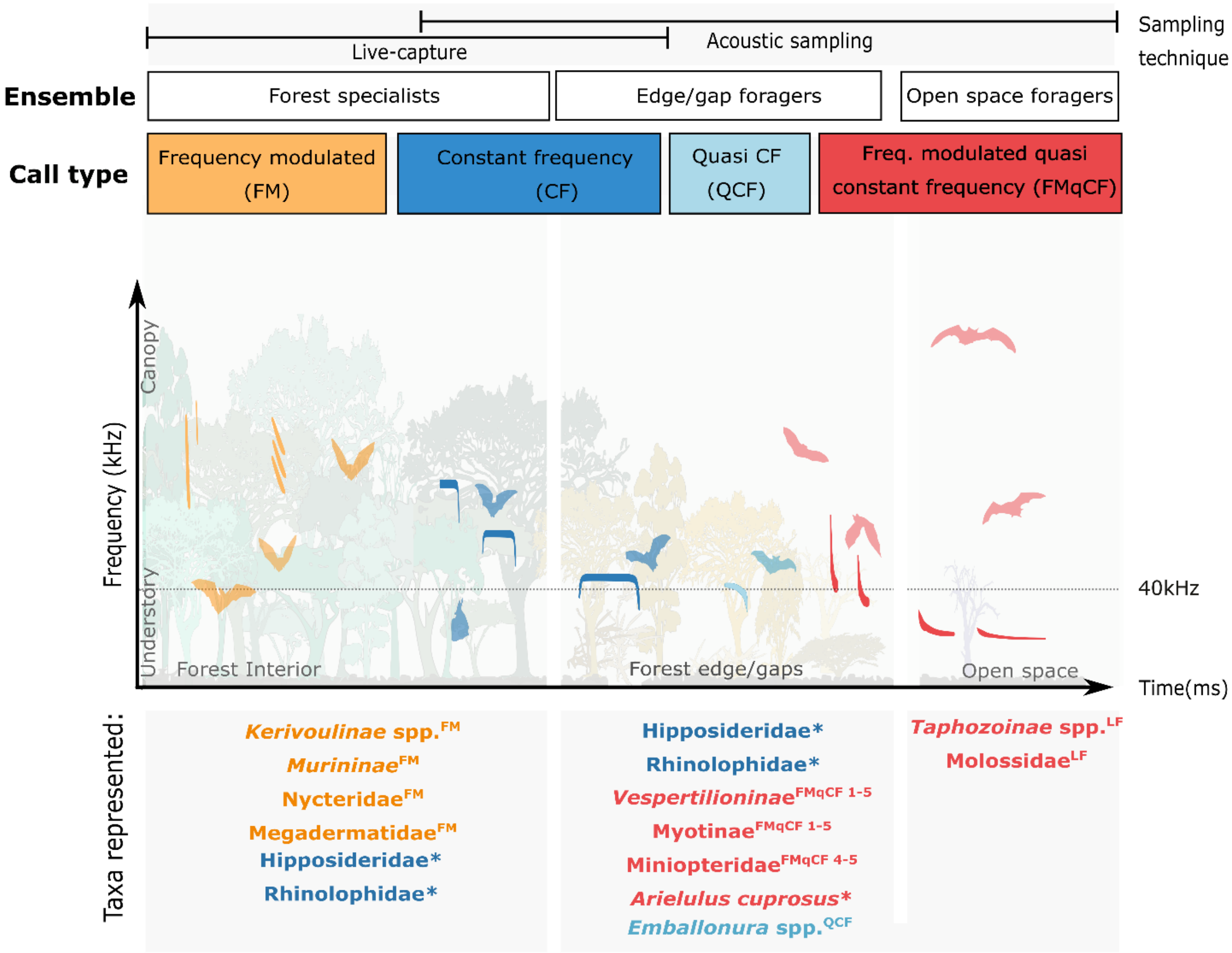


Figure 4.1 The four call types used in the Borneo bat classifier

Representation of bat ensembles in Borneo, their corresponding call types, and species/sonotypes used to train the bat call classifier. (*Identified to species; ^{FM}Identified to call type ‘frequency modulated’; ^{FMqCF1}Identified to ‘frequency modulated quasi constant frequency sonotype 1’; ^{LF}Identified to sonotype type ‘low frequency’).

.

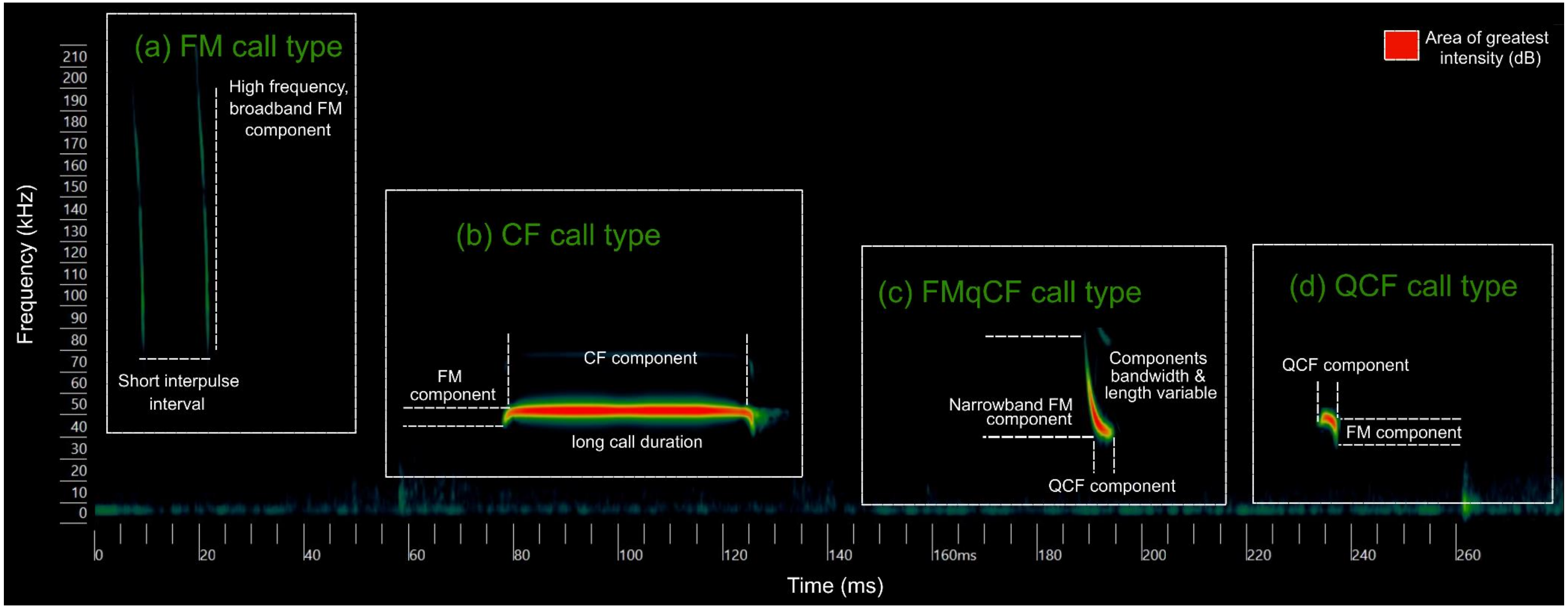


Figure 4.2 Sonograms for the four call types used in the Borneo bat classifier

FM – frequency modulated, CF – constant frequency, FMqCF – frequency modulated quasi constant frequency, QCF – quasi constant frequency.

4.3 Materials and methods

Current reference databases are typically insufficient for training species-specific classifiers. Hierarchical classifiers help alleviate this problem by classifying calls to ensemble/call type, predefined call groups, or to species level when sufficient training material exists. Ultimately, such classification limits the volume of acoustic data that requires manual identification by a user. Our Borneo Bat Classifier (BBC) incorporates two hierarchical classification stages with three components, each trained using bat calls from Borneo. First, calls are identified to one of four broad call types. Second, depending on the call type identified, a call may be further classified to species (if an identity can be inferred from a call database or the literature), or sonotype (a taxonomic unit described only by its acoustic parameters and lacking a referent species identity in databases or the literature). A corresponding confidence value is provided for each assignment. This tiered approach maximises the classification accuracy for the data available, by prioritising specific call parameters within the machine learning algorithms for distinguishing between species of the same call types. Each species/sonotype is provided with a recommended confidence threshold beyond which manual verification is required. This approach minimises the manual workload while preserving the overall confidence in identifications.

4.3.1 Input data

We collated reference calls from 687 captured bats of 42 species from 23 sites across the three countries of Borneo (see Supplementary Notes for the complete methodology for collecting reference calls; Fig. 4.3). To enhance the variability encapsulated within the training data, we also included calls of free-flying bats recorded by static detectors (Song Meter 2 BAT, Wildlife Acoustics) in a typical forest-farmland landscape in Tawau district, Sabah (at the Stability of Altered Forest Ecosystems Project, SAFE; www.safeproject.net), hereafter referred to as field recordings. Unlike the reference calls, which are tied to the identification of captured bats, calls from field recordings are not linked to in-hand identifications. Static detectors were set at 26 locations in the SAFE landscape for 862 hours in multiple habitat types (Supplementary Notes). Reference calls and field recordings were collected as WAV sound files. The calls from Sarawak contributed to the Asian Bat Call Database and are available from the Chirovox library (Görföl *et al*., 2022; McArthur & Khan, 2021).

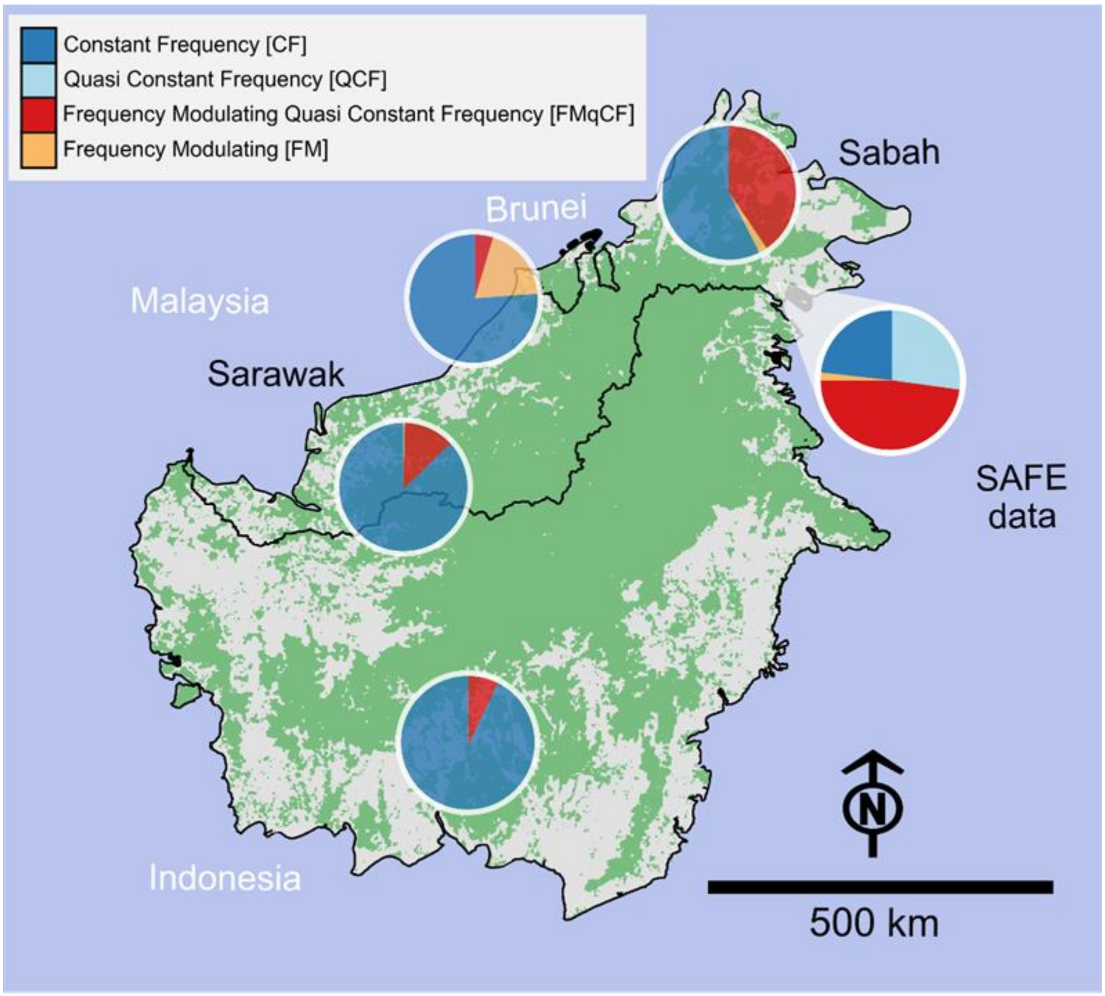


Figure 4.3 Map for the reference calls used to train the Borneo bat classifier

Reference calls collected from 23 sites in Borneo, aggregated to political administrative units. In the Malaysian states, 17 species were recorded in Sabah, 35 in Sarawak. Ten species were recorded in the Indonesian provinces of West and Central Kalimantan, and 21 in the Nation of Brunei. Calls collected from Tawau district Sabah are field data (SAFE landscape). Forest cover shown in green is for 2015 from https://earthenginepartners.appspot.com/science-2013-global-forest. White areas represent non-forest cover.

4.3.1.1 File Processing

All WAV files were split into sequences of five seconds with a minimum of two recognisable echolocation calls per species/sonotype using Kaleidoscope v.5.1.9 (Wildlife Acoustics Inc, USA). This was used to define a bat pass as a measure of activity (Torrent *et al.* 2018). Files were also filtered within a target frequency range between 8 and 250 kHz and call durations between 1 and 500ms to reduce the amount of non-bat ‘noise’ in samples. In each of the total 34,792 bat passes (each five seconds long), there were ≥ 2 calls from at least one species/sonotype present.

Using Kaleidoscope Viewer (FFT size 256, window size 128, Hamming window, and cache size 256 MB), call parameters within each recording were compared to the relevant literature and against the reference calls to determine the species identification (Table 4.1; Supplementary Notes). The parameters included: the frequencies (in kHz) at the start and end of the call, the maximum, minimum, and frequency of maximum energy (peak freq.), call duration (ms; from start to end frequencies), pulse interval (ms), duty cycle (%), and measures of call shape based on slope of the call (see Supplementary Notes and Supplementary Table 4.2). Many species in the region produce calls in which parameters overlap (e.g., *Hipposideros cineraceus* and *H. dyacorum*). Therefore, we grouped species into sonotypes or identified them only to call type when there was a risk of misidentification (Table 4.1). For the call characteristics for the FMqCF sonotypes see Supplementary Table 4.3.

Table 4.1 Number of bat passes used to train and test the Borneo Bat Classifier

The total number of bat passes/files (5-second-long sequences which include multiple calls) that were available for training and testing the Borneo Bat Classifier per call type/sonotype/species, along with the number of calls extracted. Values represent both reference calls and field recordings and where only one species was present in the sequence.

		Field		**Sabah**		**Sarawak**		**Brunei**		**Kalimantan**	
ID category	**Code**	Files	Calls	Files	Calls	Files	Calls	Files	Calls	Files	Calls
Constant frequency [CF]											
H. cineraceus/*dyacorum*	H140	1	2	5	48	12	338	8	49	10	140
Hipposideros ater	Hate					5	124			2	31
Hipposideros bicolor	Hbic					2	35				
Hipposideros cervinus	Hcer	18	54	14	397	14	487			24	423
Hipposideros coxi	Hcox					2	132				
Hipposideros diadema	Hdia	25	313	3	138	3	67				
Hipposideros galeritus	Hgal	78	218	2	146	8	217			6	139
Hipposideros larvatus	Hlar					3	67				
Hipposideros ridleyi	Hrid			9	276	2	50			2	24
Rhinolophus acuminatus	Racu	1	4	3	35	6	699				
Rhinolophus affinis	Raff	4	14			5	216				
Rhinolophus borneensis	Rbor	143	613	13	545	7	371			14	180
Rhinolophus creaghi	Rcre					6	113				

Rhinolophus luctus	Rluc	645				2	71				
Rhinolophus philippinensis	Rphi					10	390				
Rhinolophus sedulus	Rsed	2893	25723	8	154	3	80	24	354	12	119
Rhinolophus trifoliatus	Rtri	4312	32576	47	3051	1	41	8	103	14	82
Frequency modulated quasi constant frequency [FMqCF]											
FMqCF sonotype 1	FMqCF1	261	5520								
FMqCF sonotype 2	FMqCF2	567	9779								
FMqCF sonotype 3	FMqCF3	1368	14948								
FMqCF sonotype 4	FMqCF4	2082	24116		2953		225				
Glischropus tylopus				36		1					
Myotis horsfieldii						4					
Tylonycteris robustula				55							
FMqCF sonotype 5	FMqCF5	3350	65102		4869		1090		31		86
Miniopterus australis						1		1			
Myotis muricola						7		2			
Myotis ridleyi						2		4		4	
Tylonycteris pachypus				10							
Low frequency sonotype	LF	1215	6742				1293				
Chaerephon plicatus						35					
Saccolaimus saccolaimus						1					
Arielulus cuprosus	Acup	4	29			1	92				

Quasi-constant frequency* [QCF]		5288	71854		9	
Emballonura alecto	QCF				1	
Emballonura monticola	QCF				1	
Frequency modulated [FM]*	FM	1006	4898	174	420	127
Kerivoula hardwickii				5		18
Kerivoula intermedia					2	7
Kerivoula lenis						1
Kerivoula minuta					1	7
Kerivoula papillosa				4	5	21
Kerivoula pellucida						8
Murina peninsularis					1	1
Murina suilla					3	7
Megaderma spasma					2	
Nycteris tragata					3	
Phoniscus atrox				1		2
Phoniscus jagorii				2		2

Only WAV files with a single species present were used in the classifier design to ensure there was no misidentification between calls. Adobe Audition (Adobe Systems) was used to scrub non-target bat species from the reference call files that comprised calls from multiple species (e.g., where there was a flyby). Calls that were obscured or faint (< 20dB), feeding buzzes, and social calls were also excluded. Field recordings were subset to those in which only one species was identified in the manual identification process (28,831 of 34,792 files).

4.3.1.2 Call parameters

We used the threshold function in the R package "Bioacoustics" (Marchal, Fabianek and Scott 2020) to measure 26 call parameters (Supplementary Table 4.2; settings: minimum duration 1.5ms, maximum duration 80ms, FFT size = 512, FFT overlap = 0.875, extraction threshold (sensitivity of which extraction is triggered) = 4dB, signal to noise ratio threshold (SNR; sensitivity threshold at which the extraction stops) = 4dB, and Hanning window). Threshold and SNR threshold parameters were calibrated to determine which provided the greatest proportion of calls extracted with the smallest rates of noise/error introduced.

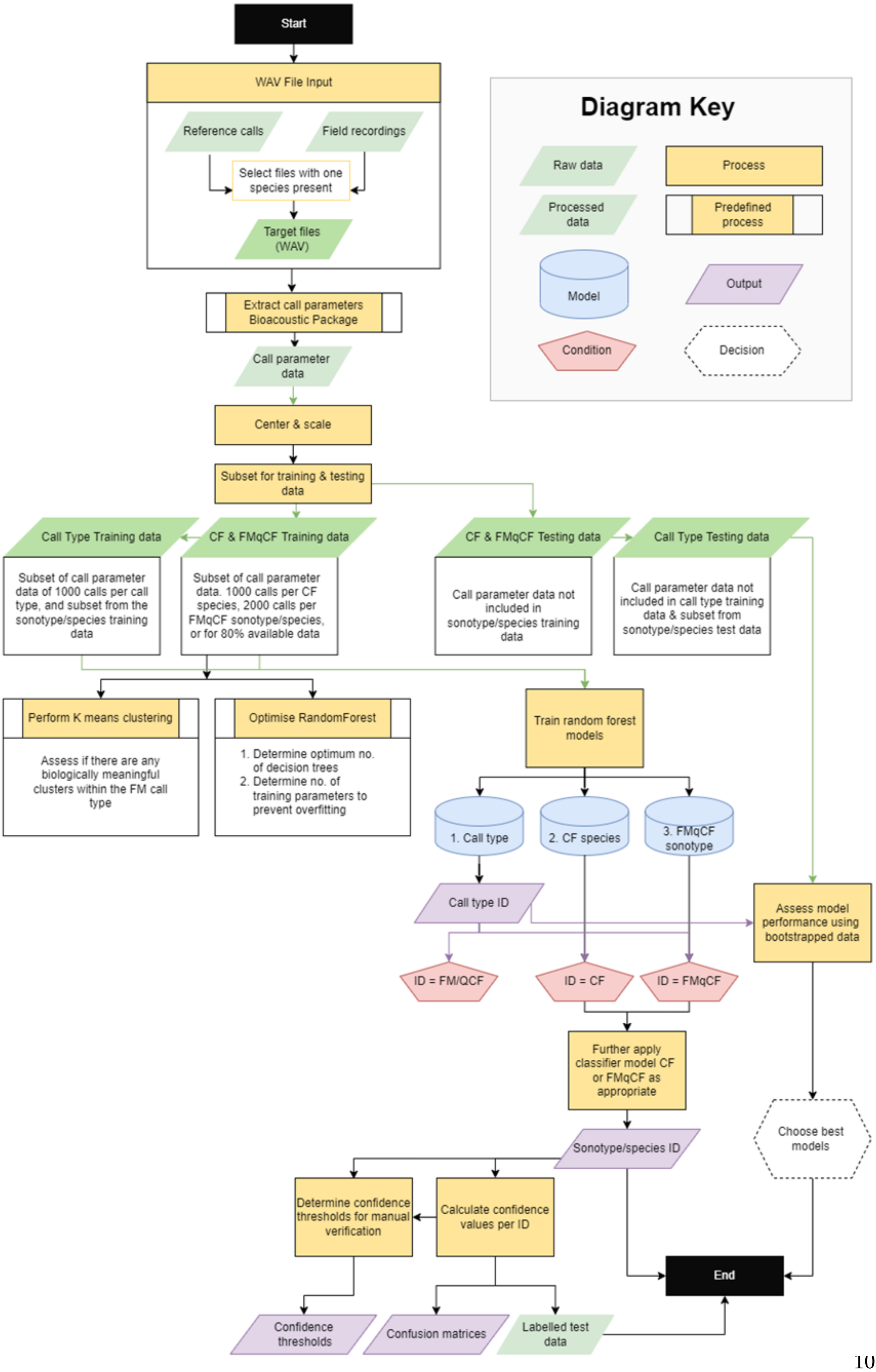

Start
WAV File Input
Reference calls
Field recordings
Select files with one species present
Target files (WAV)
Extract call parameters Bioacoustic Package
Call parameter data
Center & scale
Subset for training & testing data
Call Type Training data
Subset of call parameter data of 1000 calls per call type, and subset from the sonotype/species training data
CF & FMqCF Training data
Subset of call parameter data. 1000 calls per CF species, 2000 calls per FMqCF sonotype/species, or for 80% available data
CF & FMqCF Testing data
Call parameter data not included in sonotype/species training data
Call Type Testing data
Call parameter data not included in call type training data & subset from sonotype/species test data
Perform K means clustering
Assess if there are any biologically meaningful clusters within the FM call type
Optimise RandomForest
1. Determine optimum no. of decision trees
2. Determine no. of training parameters to prevent overfitting
Train random forest models
1. Call type
2. CF species
3. FMqCF sonotype
Call type ID
Assess model performance using bootstrapped data
ID = FM/QCF
ID = CF
ID = FMqCF
Further apply classifier model CF or FMqCF as appropriate
Sonotype/species ID
Choose best models
Determine confidence thresholds for manual verification
Calculate confidence values per ID
End
Confidence thresholds
Confusion matrices
Labelled test data
Diagram Key
Raw data
Process
Processed data
Predefined process
Model
Output
Condition
Decision

Figure 4.4 Pipeline for designing the semi-automated classifier

Pipeline for building the classifier framework to identify bat calls first to call type and then to species/sonotype where appropriate.

4.3.1.3 Call classification via machine learning

The BBC comprised two hierarchical stages including three components, each based on a separate random forest model. In the first stage, calls were classified into one of four broad call-types (‘frequency modulated’, FM; ‘constant frequency’, CF; ‘frequency-modulated quasi constant frequency’, FMqCF; and ‘quasi-constant frequency’, QCF; Table 4.1). Calls identified as CF or FMqCF underwent an additional classification stage. Those classified as CF were classified to species using a second model. Similarly, calls classified as FMqCF were subsequently classified into sonotype/species using a third model. See Fig. 4.4 for full pipeline.

4.3.1.4 Subset for training and testing data

The first model (i.e., call type) was trained using 1000 random calls per call type, as was the second model that identified calls to CF species (1000 calls per species). The third model (i.e., FMqCF) was trained using 2000 calls per sonotype/species. We compared five training data sizes (250-5000 calls) per model to determine the optimum size in terms of accuracy (the percentage of overall correct classifications out of the total number of classifications performed) and kappa (accuracy normalised for random chance per classification class; Harrell Jr, 2015). The remaining calls not used for training were used for testing. Where a call type or sonotype/species had insufficient calls to meet these training thresholds, 80% of the available data were used for training to set aside 20% for testing.

4.3.1.5 Constructing the models

For the BBC, we used random forest supervised machine learning algorithms as these performed the best amongst five other algorithms tested (Supplementary Fig. 4.1). A random forest is an ensemble of an arbitrary number of decision trees randomly built using bootstrapped samples of a training dataset which is used to assign the classification of highest likelihood (Breiman 2001). Due to their repetitive structure, these supervised machine learning algorithms are robust to outliers and can incorporate mixed variable datasets (Olden, Lawler and Poff 2008). As a result, they provide the highest certainty for the lowest resource requirements and have previously been used to classify echolocation calls in multiple species-rich regions, including Amazonia and Central America (López-Baucells *et al.* 2019; Zamora-Gutierrez *et al.* 2016). Models were constructed using the R package “caret” (Classification and Regression Training; Kuhn *et al.* 2020).

To determine the optimum number of call parameters to be included in each random forest, we tested for overfitting (the process by which too many parameters included in a model reduces its performance) using 10-fold cross-validations for models containing between 1 to 26 call parameters (James *et al.* 2013). We also calculated the error rate for the models using between 1 and 500 decision trees to determine which provided the least error for the lowest computational power. The optimum number of parameters with the mean lowest error rate was 15. However, there was no evidence of overfitting when using up to 26 parameters, and the error rate plateaued at approximately 100 decision trees. We therefore used 26 parameters and 100 decision trees across all models for consistency.

4.3.1.6 *k-means clustering*

To check whether it was possible to discriminate individual species within the FM call type we applied k-means clustering, an unsupervised machine learning approach used to cluster observations without prior information of species identity (Hartigan and Wong 1979). However, there was no distinction between the species assigned to each of these clusters and therefore we did not classify FM calls beyond the call type.

4.3.2 Performance testing

4.3.2.1 Testing success rate on bootstrapped data

We evaluated classification accuracy and predictive power of the models on the testing dataset using accuracy and kappa performance metrics, with acceptable agreement determined as > 0.41 (McHugh 2012). We chose recall (percentage of true positives, e.g., number of correct classifications per class out of total classification per class) and precision (probability given the class that the classification is correct) as metrics to evaluate the classification success for each call type. Unlike metrics such as specificity (percentage of true negatives) or negative predictive value (probability, considering each class, that it is correctly identified as not a given class), these metrics highlight the true positives in the classification process and are therefore considered the most reliable and conservative performance metrics for multicategory acoustic classifiers (Jennings, Parsons and Pocock 2008).

We assessed the relative importance of call parameters using variable importance scores (James *et al.* 2013) and the system runtime required to train the models. This was measured on an Intel i5 2.50 GHz core processor with 8 GB RAM. The "best" models were defined as those at each stage with the greatest accuracy and predictive power for the lowest computational expense.

4.3.2.2 Accuracy thresholds for manual verification

Each classification was assigned both an automatic identification label and the corresponding accuracy of that identification as a percentage (Fig. 4.5). Following López-Baucells *et al.* (2019), we estimated the percentage of bat passes that would need to be manually verified using the BBC depending on eight classification accuracy thresholds between 60-95% in 5% steps. Files with < 60% accuracy were discounted.

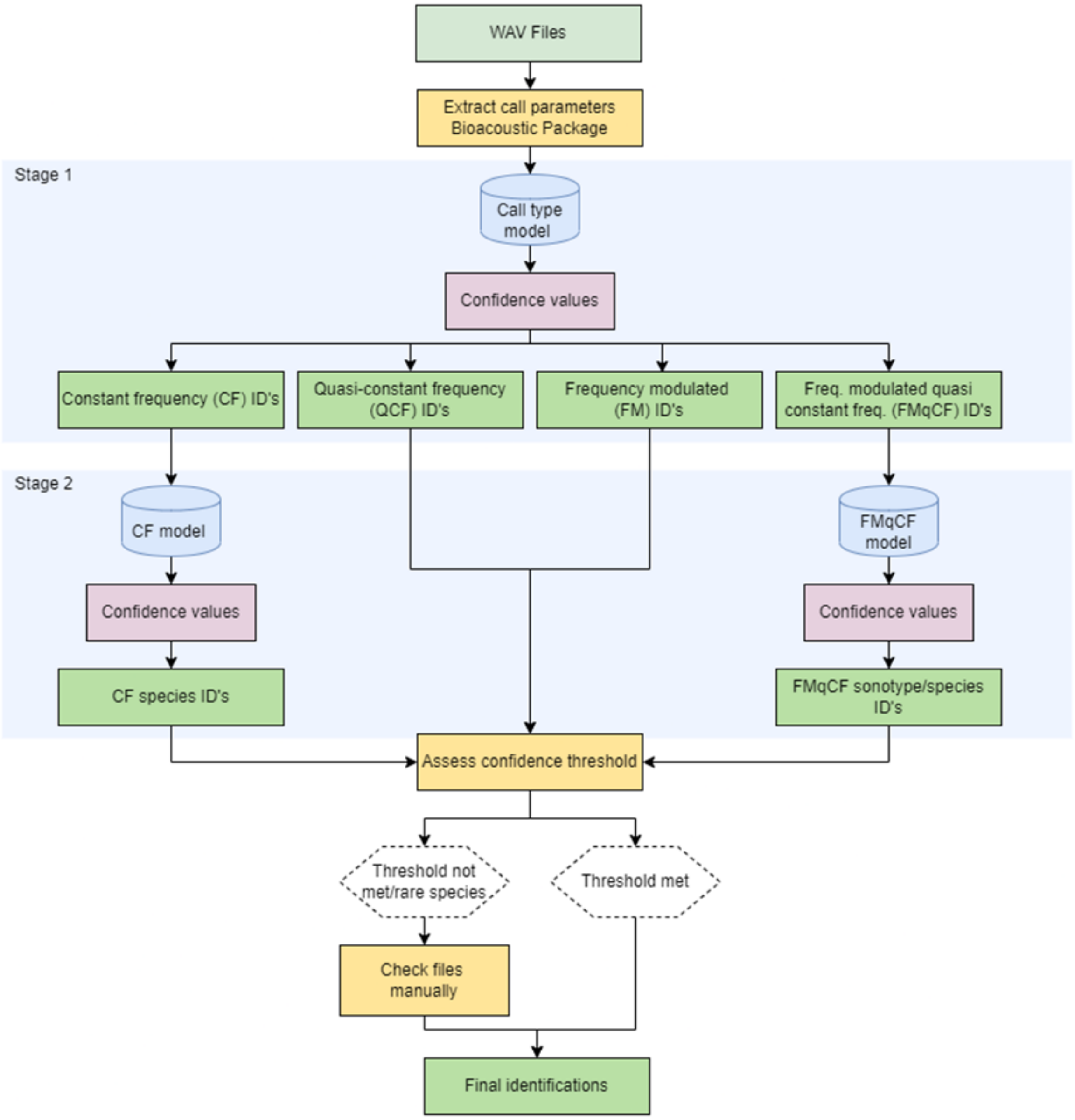


Figure 4.5 Diagram for the user application of the classifier framework

Diagram for the user application of the classifier framework. WAV files (pre-processed) are imported into R to extract their call parameters, these are then classified to call type. Depending on the call type identified, this may be the final assignment or it may be further classified to species/sonotype. Each assignment will have a corresponding confidence value. These are compared against confidence thresholds for each call type/sonotype/species to determine if the file requires manual verification.

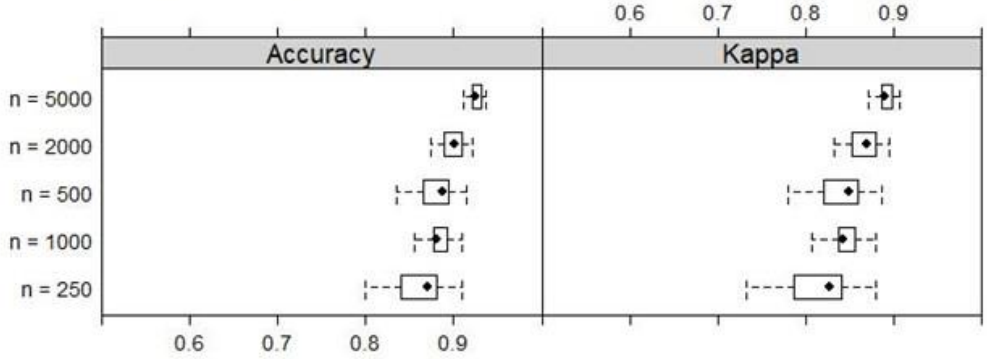
The optimal threshold was determined as where the F1-score (harmonic mean of precision and recall) was > 0.9 (Kuhn *et al.* 2020). The optimal threshold was accepted for identification to call type, for common species, and where there was adequate testing data in the classifier (> 250 calls). We advise manual verification of all rare species, where the threshold was not met, or for species where test data was < 250 calls.

4.4 Results

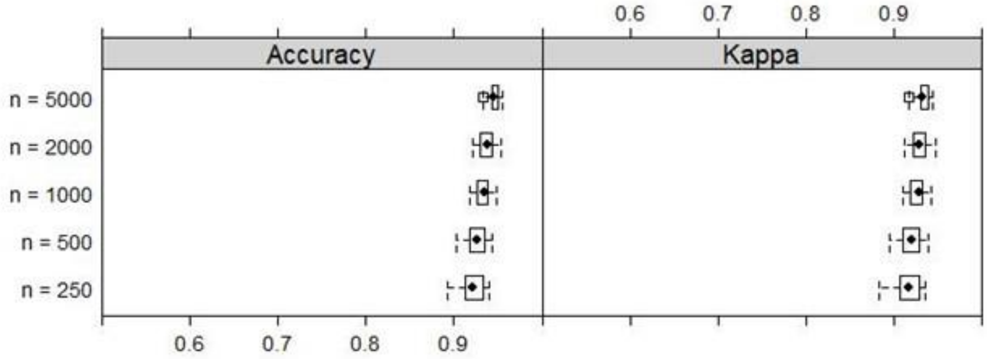
4.4.1 Minimum training dataset size

We found > 85% accuracy in identifying calls to call type using 1000 calls per category (Fig. 4.6), with a kappa greater than 0.81 indicating almost perfect agreement. Accuracy was still high (~80%) at smaller training data sizes, however there was greater variability in performance with accuracy varying up to ~10% and kappa by ~15%. There was little improvement by increasing 1000 calls to 2000 or 5000 calls compared to the increase in computational power (which ranged from ~3-fold to > 10-fold depending on the measure of power and training data size; Supplementary Table 4.4). The second model, classifying calls to CF species, showed a similar trend. Training data sizes ≥ 1000 ensured accuracy and kappa > 0.9. For the third model that classified FMqCF calls to sonotype/species, performance increased incrementally with increasing training data input. To achieve a kappa that would be considered substantial agreement (0.61-0.8), the model required 2000 calls per sonotype/species.

(A) Classification to call type



(B) Classification to CF species



(C) Classification to FMqCF sonotype/species

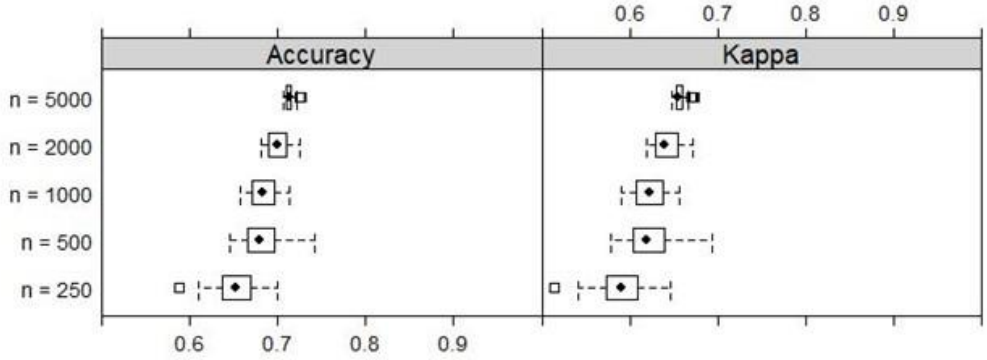


Figure 4.6 Comparison of different training data sizes for each classification algorithm

(A) Classification to call type, (B) classification to CF species, and (C) classification to FMqCF sonotype/species. Dots are medians, boxes 25% and 75% quartiles and whiskers denote the range.

4.4.2 Call parameter importance

The random forests prioritised different parameters when classifying to either call type or sonotype/species (Fig. 4.7). “Maximum frequency” was the most important parameter for determining call type and differentiating between CF calls. The CF model also shared eight of the ten most important parameters as used for classifying to call type, though their importance between models varied. While “raw slope estimate” (slope) and “smoothed slope estimate after Kalman filtering” (slope smoothed) were not in the top ten for the CF model, they regained importance in the FMqCF model where they were the seventh and eighth most important parameters respectively. The “characteristic frequency/frequency at which the slope is the flattest” (char. freq.) was the parameter of most importance for distinguishing between FMqCF calls and the second most important for determining call type.

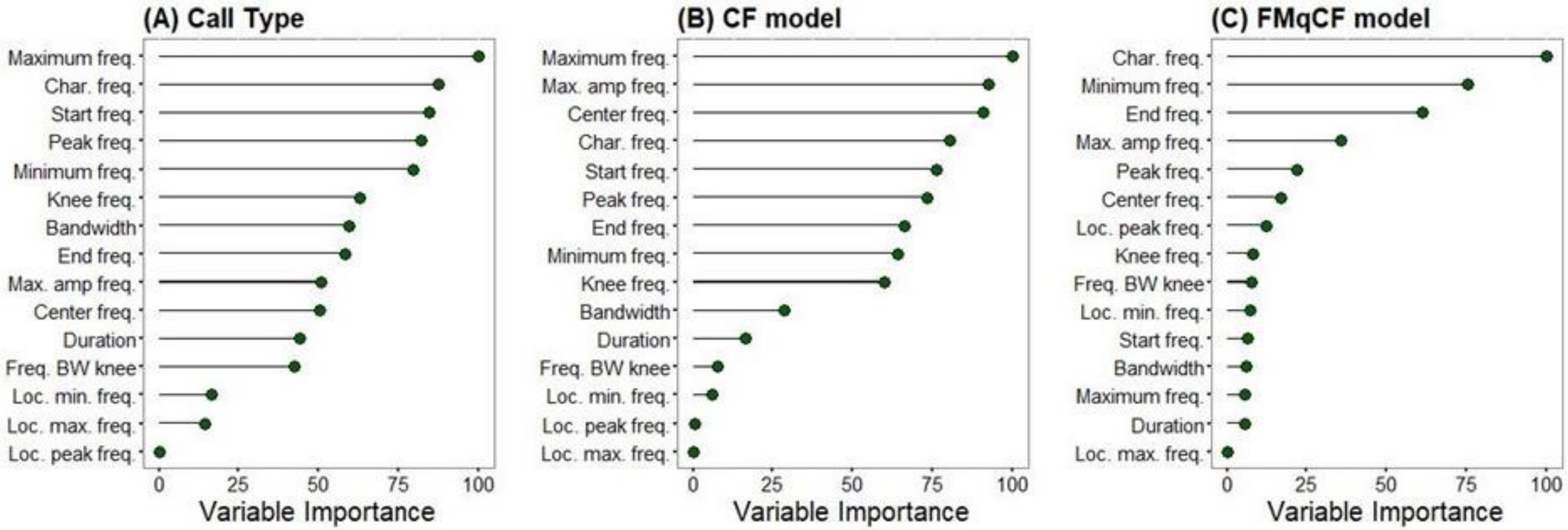


Figure 4.7 Comparison of variable importance in each classifier stage

The importance of each call parameter used in the final random forest classifier - the Borneo Bat Classifier. High values indicate greater parameter importance. See supplementary table 4.2 for parameter acronyms. (A) Classification to call type, (B) classification to CF species, and (C) classification to FMqCF sonotype/species

4.4.3 Success in performing classifications

To call type, the classifier achieved > 90% balanced accuracy for all call types (Table 4.2). Using the bootstrapped data > 85% of the calls were identified correctly to call type (Table 4.3). FM calls were correctly identified in 88.5% of cases. Where they were misclassified, they were most commonly reported as FMqCF type calls (7.9% of cases) (Supplementary Table 4.5). QCF calls were correctly identified in 91.38% of cases but were misidentified predominantly as CF calls (4.05% of cases).

Across both the second (CF classification) and third model (FMqCF classification), all sonotypes/species achieved a balanced accuracy outcome ≥ 80%. A balanced accuracy > 90% was achieved for all species included in the CF model, except for *Hipposideros bicolor* and *Rhinolophus creaghi*, which had a balanced accuracy score of 80% and 86% respectively (Table 4.2). FMqCF sonotypes showed the greatest rates of misclassification. However, they were still correctly identified in 65.90-73.55% of cases (Supplementary Table 4.6).

Table 4.2 Performance (%) of random forest models in each stage of the Borneo Bat Classifier

Performance (%) of random forest models in each stage of the Borneo Bat Classifier. The optimal confidence threshold was defined as an F1-score > 0.9 (NA = testing data < 250 calls; NR = values did not reach > 0.9 F1-score threshold). The percentage of each call type/sonotype/species that requires manual verification is given according to this optimal threshold. Total % of calls to verify - all calls identified to call type/sonotype/species with given confidence 60-100%.

ID category	**Code**	**Recall**	**Specificity**	**Precision**	**Balanced Accuracy**	**F1-Score**	**Optimal confidence threshold**	**Total % of calls to verify**
Model 1: Classification to call type								
Constant frequency	CF	0.88	0.98	0.93	0.93	0.96	60	0
Frequency modulated quasi constant frequency	FMqCF	0.88	0.97	0.97	0.93	0.97	60	0
Quasi-constant frequency	QCF	0.91	0.95	0.88	0.93	0.94	60	0
Frequency modulated	FM	0.88	0.95	0.21	0.92	0.93	80	53
Model 2: Classification to CF species								
H. cineraceus/dyacorum	H140	0.97	1.00	0.58	0.98	0.97	75	20
Hipposideros ater	Hate	0.92	1.00	0.61	0.96	0.95	65	16
Hipposideros bicolor	Hbic	0.60	1.00	0.75	0.8	NA	NA	100
Hipposideros cervinus	Hcer	0.99	1.00	0.88	0.99	0.95	60	0
Hipposideros coxi	Hcox	1.00	1.00	0.08	1.00	NA	NA	100
Hipposideros diadema	Hdia	0.97	0.99	0.19	0.98	NR	NR	100
Hipposideros galeritus	Hgal	0.94	1.00	0.57	0.97	0.98	60	0
Hipposideros larvatus	Hlar	1.00	1.00	1.00	1.00	NA	NA	100

Hipposideros ridleyi	Hrid	0.84	1.00	0.76	0.92	0.94	60	0
Rhinolophus acuminatus	Racu	0.99	1.00	0.38	0.99	0.92	60	0
Rhinolophus affinis	Raff	0.89	1.00	0.94	0.94	NA	NA	100
Rhinolophus borneensis	Rbor	0.96	0.99	0.58	0.98	0.91	95	62
Rhinolophus creaghi	Rcre	0.71	1.00	0.36	0.86	NA	NA	100
Rhinolophus luctus	Rluc	0.99	0.99	0.39	0.99	0.94	80	31
Rhinolophus philippinensis	Rphi	1.00	1.00	0.22	1.00	0.95	85	49
Rhinolophus sedulus	Rsed	0.96	0.98	0.97	0.97	1.00	60	0
Rhinolophus trifoliatus	Rtri	0.99	0.9	0.91	0.95	0.96	60	0
Model 3: Classification to FMqCF sonotype/species								
FMqCF sonotype 1	FMqCF1	0.66	0.95	0.33	0.80	NR	NR	100
FMqCF sonotype 2	FMqCF2	0.69	0.96	0.57	0.83	0.90	75	54
FMqCF sonotype 3	FMqCF3	0.74	0.98	0.8	0.86	0.92	60	0
FMqCF sonotype 4	FMqCF4	0.69	0.95	0.76	0.82	NR	NR	100
FMqCF sonotype 5	FMqCF5	0.94	0.95	0.96	0.95	0.98	60	0
Low frequency sonotype	LF	0.93	0.98	0.65	0.95	0.92	95	52
Arielulus cuprosus	Acup	0.83	1.00	0.19	0.92	NA	NA	100

Precision was generally lower for high-frequency calls (> 120 kHz) and/or where there was limited training data (e.g., *Arielulus cuprosus*). There were several exceptions for CF species which show less variability than FMqCF calls. *Rhinolophus acuminatus* was trained using 590 calls and was correctly identified in 98.7% of cases. A further three CF species achieved a balanced accuracy of 100% (all calls correctly identified). However, all three had small test data sizes (26-78 calls) constituting only a few individuals. Therefore, it is likely errors would occur if a larger testing data with more individuals were available for these species.

Table 4.3 Confusion matrix for correct and incorrect classifications to call type

Confusion matrix demonstrating the percentage of correct and incorrect bat identifications made for bootstrapped test data for call type 1000 model. Grey = correct species identification.

	True Call Type Identification			
Prediction	CF	FMqCF	QCF	FM
CF	88.5	1.2	4.05	1.7
FMqCF	2.4	87.9	2.86	7.9
QCF	3.8	5	91.38	2
FM	5.3	5.9	1.7	88.5
N calls	62,170	125,398	69,854	3,619

4.4.4 Manual verification

Manual verification was only needed for a minority of calls for the main classifier: CF, FMqCF, and QCF call types all reached an F1 score > 0.9 at the 60% confidence threshold (Table 4.2; Supplementary Table 4.7), and while this threshold was higher (80%) for FM calls the classifier still reduced the number of calls for manually processing by almost half (Table 4.2, Fig. 4.8). As FM is the least common call type present, this means that < 1% of the total sum of calls identified to call type need manual verification.

Six CF species also did not require any manual processing, having reached the necessary F1-score at the 60% confidence threshold. This included the three most common CF species, *Hipposideros cervinus, R. sedulus* and *R. trifoliatus* (Table 4.2). Of the remaining species, a further four reduced processing demands by over 50%, *Hipposideros ater* (15.4%), *H. cineraceus/dyacorum* (20%), *R. luctus* (30.8%), and *Rhinolophus philippinensis* (48.7%). Six species required all calls to be checked, either because of low sample sizes/rarity (n < 500;

Hipposideros larvatus, H. bicolor, H. coxi, R. creaghi, and *R. affinis*) or because they did not achieve a satisfactory F1-score (e.g., *Hipposideros diadema*). Overall, this reduced the number of CF calls to manually verify to < 1% (536 out of 54,900 calls).

Two of the seven FMqCF sonotypes, FMqCF3 and FMqCF5, did not require any manual processing. For two additional sonotypes, low frequency and FMqCF2, the model reduced processing requirements by almost half, with 51.5% and 53.5% of calls requiring manual verification respectively. However, two sonotypes did not achieve an F1-score > 0.9 and therefore FMqCF1 and FMqCF4 require all files to be manually checked. We also advise this for *A. cuprosus* due to the small training data size (n = 121). In total, this reduced the number of FMqCF calls to manually verify down to 27.5% (30,259 out of 110,232). Across all models, this means the BBC reduces the number of calls to check by 86.18% (34,006 out of 245,991).

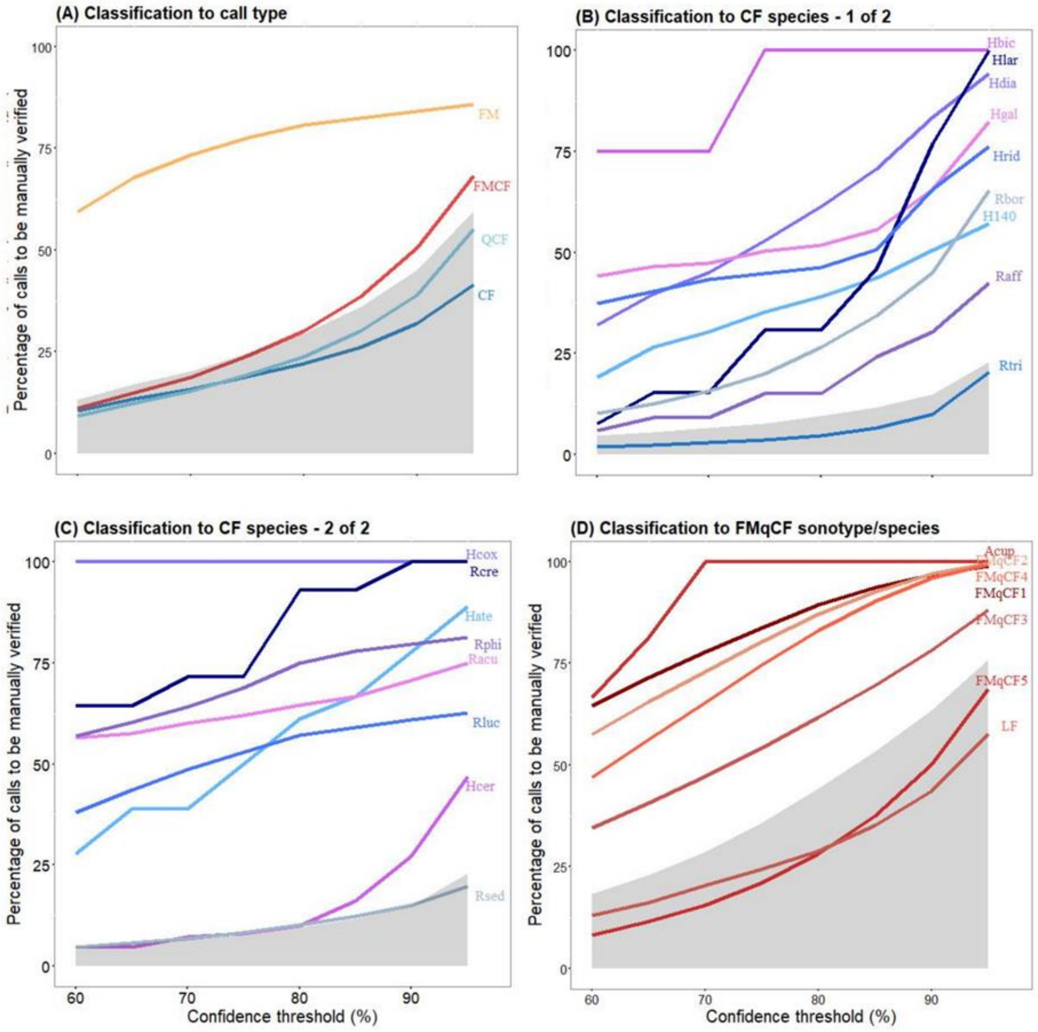


Figure 4.8 Percentage of echolocation calls requiring manual verification using the Borneo Bat Classifier

Percentage of echolocation calls requiring manual verification dependent on model and sonotype/species by confidence threshold. (A) For classification to call type, (B and C) for classification to constant frequency (CF) species divided across two panels, and (C) for classification to frequency modulated quasi constant frequency (FMqCF) sonotype. Shaded area – the overall proportion of files per model.

4.5 Discussion

We developed a hierarchical classification framework that can be used to develop classifiers to greatly reduce the processing of bat echolocation calls, particularly in localities where species-specific training data may be limited. Our framework is intended to be applicable to bat faunas across Southeast Asia and has demonstrated utility with the bats of Borneo.

Our approach substantially reduces human input and demonstrates how information on different call types and species call characteristics can lead to meaningful classifications of acoustic data that represent different ecological ensembles and indicators for lesser-known species. The ultimate aspiration for acoustic monitoring and automated classification is to differentiate among all species present. However, the urgency for monitoring data, even if only a subset of the total community, has never been higher. In many regions of Southeast Asia, heavy deforestation is expected to result in over 40% of regional bat species to be lost by the end of the century (Lane, Kingston and Lee 2006). The Borneo Bat Classifier introduced here provides a means to document populations of some individual taxa (e.g., CF-calling bats), while also resolving several ensembles of bats from acoustic data.

A key benefit of the BBC is that it performs best for the most common species and sonotypes. For example, over 99% of CF calls recorded in our field dataset could be identified to species level. The FMqCF call type, which represents the greatest volume of calls, also displays the greatest call plasticity, resulting in both within-sequence variability and interspecific overlap in call parameters which makes differentiating between species/sonotypes more challenging than CF calls. However, our approach reduces overall processing of this call type by 70% (30,259 to check out of 110,232). By grouping calls in this way, future users can rapidly and reliably discriminate between edge/gap, open-space bats, and other groups without relying on species-specific identifications. Therefore, this tool can vastly reduce the manual processing demands of acoustic projects where this level of classification is appropriate. It is important to state that this tool is not designed to replace species-specific monitoring and should not be used on its own to inform species-specific conservation efforts. It is designed to assess how the insectivorous bat community more broadly varies between habitats, management strategies, or over time. In this way, it can help inform monitoring efforts where species-specific monitoring is not possible or can only be weakly undertaken.

Our results support previous recommendations against differentiating broadband FM calls to species in Asian bat assemblages (Hughes *et al.*, 2011; Kingston *et al.*, 1999). Even under controlled conditions, calls can be difficult to discriminate between species due to overlap in call variation (see Schmieder *et al.*, 2012). Species utilising these FM calls (e.g. Kerivoulinae, Murininae) are typically forest specialists that are well sampled using live-capture methods, but are poor candidates for acoustic monitoring due to the low-intensity and high frequencies of their echolocation calls, making reliable field recordings very difficult to obtain (Kingston 2013; Russo, Ancillotto, and Jones 2018). For example, although Kerivoulinae and Murininae bats are relatively common in Southeast Asian landscapes, including in our test landscape in Sabah (Struebig *et al.* 2013), their FM calls only comprise a very small portion (4898 calls, 2%) of files collected by the acoustic recorders simply because they are too quiet and too high pitched to be reliably recorded. Therefore, we do not recommend acoustic approaches be used to monitor these taxa.

An additional benefit of our approach is that we were able to create a classifier without access to a complete reference call library of Borneo’s echolocating bats. Species that were absent from our training data (e.g., the FM echolocator, *Kerivoula whiteheadi;* FMqCF echolocator, *Miniopterus medius*) will still be encapsulated in the broad call types. A key outcome of using representative call types or sonotypes in this way is that it allows classifiers to be developed in other regions in Southeast Asia where bat call inventories are less complete. While this methodology shows promising results for Borneo, this approach needs further development and testing to be applied to other regions in Southeast Asia. Such classifiers should be straightforward to develop as echolocating bats in this region share a common community structure, dominated by the same families found in Borneo.

Currently, random forests provide our classifier with the best performing algorithm, but this could change as the number of reference calls increases in Southeast Asia. Deep neural networks are known to provide the highest classification accuracies for the automated identification of bats in Europe (Parsons and Jones 2000) but these methods are computationally intensive; requiring extensive reference libraries that are not currently available for most species in the tropics (Walters *et al.* 2013). Kobayashi *et al.* (2021) required 54,525 calls to train a classifier to recognise 30 species in Japan. The Hungarian Natural History Museum and the Southeast Asian Bat Conservation and Research Unit recently launched the Asian Bat Call Database, a repository dedicated to making acoustic recordings of bats in Asia more accessible to acoustic researchers (Görföl *et al.* 2022). Chen *et al.* (2020) have already demonstrated the potential for neural networks to be used for identifying Asian species and reference banks such as this would expand their application in the future. Considering these future developments, our framework can be easily updated to a neural network algorithm as such call databases become widely available.

There remain important considerations when interpreting results generated by this tool. The current classifier framework determines sonotypes/species identification by individual calls (rather than from a string of calls), therefore it cannot recognise call alternation (i.e. alternating between two or more call structures, *Pipistrellus stenopterus*) (Kingston, Jones, *et al.* 2003). Rather, the classifier treats alternate calls as potentially belonging to different species, which can inflate the number of calls for verification. Two alternator sonotypes, FMqCF1 and FMqCF2, constituted approximately 10% of FMqCF calls. An option for users who have limited time to manually process files would be to re-group FMqCF1 and FMqCF5 post-classification. As the species from each of these sonotypes are producing calls of a similar structure above 50 kHz, they are likely to all be utilising their environment in an ecologically similar way, and both represent edge/gap foragers. However, FMqCF2, which produces calls of a similar frequency to FMqCF4, are more likely to be open space hawkers (Kingston, Jones, *et al.* 2003) and therefore these two sonotypes should not be grouped.

There are also species not currently included in the training dataset. While most of these species are represented within the current call types, our classifier may omit a fifth call type typical of open-space foragers of the family Emballonuridae. This includes calls from three species in Borneo belonging to the genus *Taphozous,* all of which are open-space hawking insectivores (Wei *et al.* 2008), and have calls characterised as multi-harmonic, low frequency QCF. These calls share similar properties with both the QCF call type and low frequency sonotype, therefore it remains to be seen whether they would be classified into these categories or whether a new call type would be needed once reference calls became available. Nevertheless, it is notable that in our sample landscape none of the calls were manually attributed to this group, suggesting that they may be infrequently detected in typical forest-farmland surveys.

Only three species utilising CF calls are yet to be included in our classifier (*Rhinolophus francisi, R. pusillus,* and *Hipposideros doriae*). These are either very rare or have a patchy distribution, and thus are rarely captured. *R. francisi* was only described in 2015, and is only reported from five localities in Borneo (Soisook *et al.* 2015), producing overlapping frequencies with the common species *R. trifoliatus*. *Hipposideros doriae* and *Coelops robinsoni* are also similarly rare and patchily distributed, and produce very high frequency, broadband calls with a very abbreviated or absent CF component (Kingston 2016). On the other hand, the CF calls of *R. pusillus* and *H. larvatus* should be relatively simple to discriminate by the classifier but so far there are no available recordings for these species since they are highly localised to karst outcrops (Phillipps and Phillipps 2016). Another consideration is CF species are known to express geographic variation in call frequency (Chen *et al.*, 2009), however we did not observe substantial variation in call frequencies across the geographic range included in this study. Nonetheless, increasing the extent of data used in this tool to date would help its efficacy in other localities. Where the framework is developed for other regions, users should assess whether there is evidence of geographic variation in species' call parameters for that area.

4.6 Conclusions

Our acoustic classification framework and subsequent classifier for Borneo greatly expands the capacity for monitoring bats in Southeast Asia - reducing the need for manual processing of bat calls in Borneo by seven-fold. Our framework incorporates aerial insectivorous ensembles that are regularly absent from biodiversity studies despite being important components of bat assemblages. It can be used to design additional classifiers in Southeast Asia and the palaeotropics more broadly, where species-specific classifiers are not yet possible. Therefore, improving the potential to use bat assemblages as bioindicators in tropical environments.

4.7 Availability

The BBC classifier user script is available open access from the GitHub repository github.com/TallyYoh/BorneoBatCalls (doi:10.5281/zenodo.4725680) in the programming language R version 3.6.3. The BBC classifier user script v2.0 is provided in Appendix II.

4.8 Acknowledgements

Research permit issued by Universiti Brunei Darussalam (UBD) (same ref no. for 2019 and 2020): UBD/AVC-RI/1.21.1(a). Ethics approval issued by UBD: Ubd/FOS/E2(g) Forest Entry Permit issued by Forestry Department, Ministry of Primary Resources and Tourism:[277]/JPH/UND/17 PT.1 We would also like to acknowledge Kuala Belalong Field Studies Centre (UBD), Prof. Phil Bishop (University of Otago) & Prof. Ulmar Grafe for facilitating fieldwork in Brunei. We are grateful to the Sabah Biodiversity Council for granting the access licenses (permit JKM/MBS.1000-2/2 JLD.6 (53) and JKM/MBS.1000-2/2 (374)), Sabah Parks for permission to work in the CRBR, and Sarawak Forest Department for granting permit no. NCCD.907.4.4 (JLD. 12)-3 for collecting reference calls in Gunung Mulu National Park. Specifically, Madam Rimi Repin, Deputy Director (Research and Education) of Sabah Parks, Thomas Yussop, Park Manager of Crocker Range Park, Yassin Miki, Park Manager of Kinabalu Park, and Simon Limbawang, Head of Inobong Research Station (Kinabalu Park). We thank James Jupikely, Monica Suleiman, Jaya Seelan, Azniza Mahyudin, and Kalsum Yusah, Universiti Malaysia Sabah, Isham Azhar, Katheryn V. Fitzgerald, Tenaja Smith-Butler, Michael Farmer, and Sarah Fritts, Texas Tech University, and Rieka Yu, University of Missouri, for logistical and participatory support of the project. Additional colleagues and students participated in the fieldwork and we thank them for their contribution: Gábor Csorba, Tamás Görföl, Amirah Amat, Ummu Safiyyah Daud, Nurhani Adila Abdul Halim, Su Suzaine Sukiman and Nurkujayati Martin, Alex Troutman, and Tyler Test. We would also like to acknowledge Kyle Armstrong for his technical support and guidance at the initial conception stage. Finally, we would like to sincerely thank the three anonymous reviewers who dedicated their time and expertise towards this publication.

4.9 Supplementary information

Table S4.1 Survey information relating to reference call data

Survey information relating to reference collection of full-spectrum call recordings for the Borneo Bat Classifier. CF – Constant-frequency call type

Country	Location	GPS Lat	GPS Long	Trapping type	Detector and microphone used (sampling rate kHz)	CF recorded in hand (Y/N)	Flight tent specification	Habitat for recording upon release	Associated DOI	Chirovox dataset and UID
Malaysia										
	Crocker Range National Park	5.838	116.147	Mist-nets, Harp traps	Anabat Walkabout, with uni-directional adapter (500)	Y	NA	Forest clutter, Forest edge/gap, Open-space	doi:10.3390/d12020060	
	Mulu National Park	4.023	114.821	Mist-nets	Echo Meter EM3+ internal microphone (384)	NA	NA	Open-space	doi:10.14709/BarbJ.14.1.2021.11	Dataset: DS04, UID: A000401 - A000713
	Mulu National Park	4.092	114.895	Mist-nets, Harp traps	Echo Meter EM3+ internal microphone (384)	Y	Polyester mesh: 2x4x2 m	Forest clutter, Forest edge/gap, Open-space	doi:10.14709/BarbJ.14.1.2021.11	Dataset: DS04, UID: A000401

									- A000713
Bako National Park	1.716	110.466	Mist-nets, Harp traps	Echo Meter EM3+ internal microphone (384)	Y	NA	Forest clutter, Forest edge/gap, Open-space	doi:10.14709/BarbJ.14.1.2021.11	Dataset: DS22, UID: A003786 - A003812
Samunsam Wildlife Sanctuary	1.954	109.607	Mist-nets, Harp traps	Echo Meter EM3+ internal microphone (384)	Y	NA	Forest clutter, Forest edge/gap, Open-space		Dataset: DS11, UID: A001513 - A001536, A001802
Kubah National Park	1.6128	110.196	Mist-nets, Harp traps	Advanced electret (384)	N	Polyester mesh: 100(W) ×200(L)× 195(H) (CM)	Forest clutter, Forest edge/gap, Open-space		Dataset: DS11, UID: A001537 - A001566
Sadong Jaya Nature Reserve	1.557	110.694	Mist-nets, Harp traps	NA	N	NA	Forest clutter, Forest edge/gap, Open-space		Dataset DS11, UID: A001567

								- A001568
Mount Silabur Cave	0.969	110.519	Mist-nets, Harp traps	Advanced electret (384)	N	Polyester mesh: 100(W)× 200(L)× 195(H) (CM)	Open-space	Dataset DS11, UID: A001569 - A001647
Kumpang Langgir	1.493	111.689	Mist-nets, Harp traps	Echo Meter EM3+ internal microphone (384)	Y	NA	Open-space	Dataset DS11, UID: A001648 - A001667
Lambir Hills National Park	4.198	114.042	Mist-nets, Harp traps	Echo Meter EM3+ internal microphone (384)	Y	NA	Forest clutter, Forest edge/gap,	Dataset DS11, UID: A001668 - A001690
Niah National Park	3.801	113.784	Mist-nets, Harp traps	Echo Meter EM3+ internal microphone (384)	Y	NA	Forest clutter, Forest edge/gap, Open-space	Dataset DS11, UID: A001691 - A001704

Bukit Kana National Park	2.666	111.866	Mist-nets, Harp traps	Module's Ultra-quiet microphone (384)	N	Polyester mesh: 100(W) ×200(L)× 195(H) (CM)	Forest edge/gap, Open-space	Dataset: DS11, UID: A001705 - A001720
Tanjung Datu National Park	2.055	109.642	Mist-nets, Harp traps	Echo Meter EM3+ internal microphone (384)	NA	NA	Forest edge/gap, Open-space	Dataset: DS11, UID: A001721 - A001751
Mount Penrissen	1.116	110.216	Mist-nets, Harp traps	Echo Meter EM3+ internal microphone (384)	NA	NA	Open-space	Dataset: DS11, UID: : A001752 - A001788
Long Banga	3.207	115.396	Mist-nets, Harp traps	Echo Meter EM3+ internal microphone (384)	NA	NA	Forest clutter, Forest edge/gap, Open-space	Dataset: DS11, UID: A001789 - A001801
Gunung Gading	1.690	109.845	Mist-nets, Harp traps	Echo Meter EM3+ internal microphone (384)	NA	NA	Forest clutter, Forest edge/gap,	Dataset: DS11, UID:

	National Park									A001803 - A001812
Brunei										
	Ulu Temburong National Park	4.448	115.212	Mist-nets, Harp traps	Anabat Walkabout, with uni-directional adapter (500)	Y	3.5m x 3m	Forest clutter, Forest edge/gap, Open-space		
	Ulu Temburong National Park	4.448	115.212	Mist-nets, Harp traps	Pettersson D1000x, advanced electret microphone (500)	Y	NA	Forest clutter, Forest edge/gap	doi:10.1007/s10531-012-0393-0	
	Andulau Forest Reserve	4.65	114.516	Mist-nets, Harp traps	Pettersson D1000x, advanced electret microphone (500)	Y	NA	Forest clutter, Forest edge/gap	doi:10.1007/s10531-012-0393-0	
	Tasek Merimbun Heritage Park	4.583	114.666	Mist-nets, Harp traps	Pettersson D1000x, advanced electret microphone (500)	Y	NA	Forest edge/gap Open-space	doi:10.1007/s10531-012-0393-0	
	Bukit Teraja Protection Forest	4.3	114.416	Mist-nets, Harp traps	Pettersson D1000x, advanced electret microphone (500)	Y	NA	Forest clutter, Forest edge/gap	doi:10.1007/s10531-012-0393-0	

	Sungai Ingei Conservation Forest	4.15	114.7	Mist-nets, Harp traps	Pettersson D1000x, advanced electret microphone (500)	Y	NA	Forest clutter, Forest edge/gap	doi:10.1007/s10531-012-0393-0
	Peradayan, Forest Reserve	4.75	115.166	Harp traps	Pettersson D1000x, advanced electret microphone (500)	Y	NA	Forest clutter, Forest edge/gap	doi:10.1007/s10531-012-0393-0
Indonesia	Murung Raya	0.086	114.884	Harp trap	Pettersson D1000x, advanced electret microphone (500)	Y	NA	Forest clutter, Forest edge/gap	
	Nanga Tayap	-1.509	11.706	Harp trap	Pettersson D1000x, advanced electret microphone (500)	Y	NA	Forest clutter, Forest edge/gap	

Table S4.2 Call parameters used in Borneo Bat Classifier

Call parameters extracted using the threshold function of “Bioacoustics” R package used to train the Borneo Bat Classifier.

Call parameter	Description (unit)
Peak freq.	Frequency of maximum energy (Hz)
Maximum freq.	Maximum frequency (Hz)
Minimum. freq.	Minimum frequency (Hz)
Bandwidth	Total bandwidth (Hz)
Start freq.	Frequency at the start of the call (Hz)
Center freq.	Frequency at the mid-point of the call (Hz)
End freq.	Frequency at the end of the call (Hz)
Freq. knee	Knee/frequency at which the slope is the steepest (Hz)
Characteristic freq.	The characteristic frequency/frequency at which the slope is the flattest (Hz)
Freq. BW knee	Frequency bandwidth between the knee (Hz)
Max. amp freq.	Frequency of maximum energy within the characteristic frequency
Loc. peak freq.	Location of frequency of maximum energy (%)
Loc. min. freq.	Location of minimum energy (%)
Loc. max. freq	Location of maximum energy (%)
Loc. knee	Location of the knee (%)
BW knee FC	Temporal bandwidth between knee & characteristic frequency (Hz)
Slope	Raw slope estimate (Hz/ms)
Smoothed slope	Smoothed slope estimate after Kalman filtering (Hz/ms)
Est. slope start	Slope estimate at the beginning of the call (Hz/ms)
Est. slope end	Slope estimate at the end of the call (Hz/ms)
Slope neg.	Slope negative antropy (Hz/ms)
Mid offset	Mid-offset (dB)
SNR	Signal to noise ration (dB)
Harm. distortion	Level of harmonic distortion (dB)
Smoothness	Time/frequency regularity
Duration	Call duration from the start frequency to the end frequency (ms)

Table S4.3 Sonotype descriptions

Descriptions of the call characteristics for the sonotypes classified as the call type "Frequency modulated quasi constant frequency" (FMqCF) in the Borneo Bat Classifier.

Sonotype	Description
FMqCF sonotypes	These sonotypes include species from the families Vespertilioninae, Miniopteridae, and Emballonuridae. Where species straddle a threshold (e.g. if two calls have a peak frequency of 40.5 & 39.5) round the FME up. If there are only two calls in the file & it is not clear whether it is an alternator, assume so or leave blank as one species presenting two calls is more likely than two species only presenting one call each.
FMqCF Type 1 (Edge/gap)	Call peak frequency alternates: Peak frequency - call type 1 > 40kHz & Peak frequency - call type 2 > 50kHz
FMqCF Type 2 (Open space or Edge/gap)	Call peak frequency alternates: Peak frequency - Call type 1 < 40kHz & Peak frequency - Call type 2 < 50kHz
FMqCF Type 3 (Open space)	Call peak frequency does not alternate: Peak frequency ≈ 35 kHz (31 ≤ 40 kHz)
FMqCF Type 4 (Open space or Edge/gap)	Call peak frequency does not alternate: Peak frequency ≈ 45 kHz (41 ≤ 50 kHz)
FMqCF Type 5 (Edge/gap)	Call peak frequency does not alternate: Peak frequency ≈ 55 kHz (≥ 51 kHz)

Low frequency sonotype	This sonotypes represents open-space insectivores and includes species from the families Vespertilioninae, Molossidae, and Emballonuridae. Species within this sonotype may produce FMqCF, multi-harmonic qCF calls, or calls that use a combination of call structures. The distinguishing characteristic of this sonotype is the low peak frequency ($\lessapprox$ 30 kHz).

Table S4.4 Comparison of computational power for each random forest model

Comparison of the computational power required to execute each random forest model. User central processing unit (CPU) time represents the time spent by the R session whereas the system CPU time provides the time spent by the operating system.

Model	Training data size	User CPU (secs)	System CPU (secs)	Elapsed time (secs)
Call type				
	250	9.70	0.92	12.49
	500	18.44	0.22	20.03
	1000	38.67	0.31	39.78
	2000	88.67	2.73	92.31
	5000	204.36	12.05	221.54
CF model				
	250	67.48	38.28	138.42
	500	121.30	20.56	167.78
	1000	184.22	2.99	190.38
	2000	329.34	90.70	481.18
	5000	621.67	2.77	690.07
FMqCF model				
	250	35.51	1.68	50.44
	500	73.61	0.58	77.57
	1000	163.53	3.08	169.20
	2000	377.39	7.66	392.82
	5000	1227.67	3.03	1265.59

Table S4.5 Confusion matrix for correct and incorrect identifications for the CF classifier model

Confusion matrix demonstrating the percentage of correct and incorrect identifications made for bootstrapped test data for CF 1000 model. Green = correct species identification; grey = calls originally misidentified to CF call type.

	FM	H140	Hate	Hbic	Hcer	Hcox	Hdia	Hgal	Hlar	Hrid	Racu	Raff	Rbor	Rcre	Rluc	Rphi	Rsed	Rtri	FMq CF1	FMq CF2	FMq CF3	FMq CF4	FMq CF5	LF	Acup	QCF
FM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	0
H140	0	96.8	8.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.1	0	0	0	0.6	0	0	NA	0.3
Hate	0	0	91.7	0	0	0	0	0	0	0	0	0	0.2	0	0	0	0	0	0	0	0	0	0	0	NA	0.2
Hbic	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	0
Hcer	6.6	0	0	0	98.9	0	0	0	0	0	0	0	0.2	0	0	0	0.1	0	0	0	0.7	0	0.1	0	NA	0.1
Hcox	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	1.2	0	1.1	NA	0
Hdia	0	0	0	0	0	0	96.6	5	0	11.48	0	2.9	0	14.3	0	0	2	0	0	3.9	0	0	0.1	0	NA	0.1
Hgal	3.3	1.6	0	0	0.5	0	0	93.8	0	1.64	0	0	0	0	0.2	0	0	0	8.6	0	0	6.8	1.3	0.8	NA	0.5
Hlar	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	0
Hrid	0	0	0	0	0	0	2.3	0	0	83.6	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0	NA	0
Racu	0	0	0	0	0	0	0	0	0	1.6	98.7	0	0.9	0	0	0	0	0.1	0	1.9	0.7	1.2	5.3	0	NA	1.2
Raff	0	0	0	0	0	0	0	1.3	0	0	0	88.6	0.2	0	0	0	0	0	0	0	0	0	0	0	NA	0
Rbor	0	0	0	40	0	0	0	0	0	0	1.3	5.7	95.8	14.3	0	0	1.1	0.1	0	1.9	3.5	1.2	0.8	1.4	NA	0.1
Rcre	0	0	0	0	0	0	0	0	0	0	0	0	0	71.4	0	0	0	0	0	0	0	0.6	0	0	NA	0.2
Rluc	42.6	0	0	0	0	0	0	0	0	0	0	0	0.2	0	99.5	0	0.8	0.4	25.7	51.9	47.6	25.8	10.4	42.6	NA	5.1
Rphi	1.6	0	0	0	0	0	0	0	0	0	0	0	0.2	0	0.2	100	0	0	5.7	0	1.4	1.8	0.4	28.7	NA	0.4
Rsed	42.6	1.6	0	0	0.5	0	1.1	0	0	1.6	0	2.9	0	0	0.2	0	96.2	0.2	51.4	30.8	35.7	42.3	39.4	19.1	NA	4.9
Rtri	3.3	0	0	0	0	0	0	0	0	0	0	0	2.2	0	0	0	0.1	99.1	8.6	9.6	9.8	18.4	41.4	6.3	NA	87
FMqCF1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	0
FMqCF2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	0
FMqCF3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	0
FMqCF4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	0
FMqCF5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	0
LF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	0
Acup	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	0
QCF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	0
N calls	61	63	12	5	185	1	89	80	13	61	75	35	455	7	545	38	22650	30681	35	52	143	163	753	366	0	2832

Table S4.6 Confusion matrix for correct and incorrect identifications for the FMqCF classifier model

Confusion matrix demonstrating the percentage of correct and incorrect identifications made for bootstrapped test data for FMqCF 2000 model. Green = correct species identification; grey = calls originally misidentified to FMqCF call type.

	FM	H140	Hate	Hbic	Hcer	Hcox	Hdia	Hgal	Hlar	Hrid	Racu	Raff	Rbor	Rcre	Rluc	Rphi	Rsed	Rtri	FMq CF1	FMq CF2	FMq CF3	FMq CF4	FMq CF5	LF	Acup	QCF
BB	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H140	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hate	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hbic	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hcer	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hcox	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hdia	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hgal	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hlar	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hrid	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Racu	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Raff	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rbor	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rcre	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rluc	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rphi	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rsed	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rtri	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FMq-CF1	12.3	55.6	NA	NA	NA	0	NA	16.7	NA	0	4.8	NA	4.4	0	20.3	0	4.9	4	65.9	12.2	4.85	13.5	1.4	0.4	0	8.5
FMq-CF2	4.9	0	NA	NA	NA	0	NA	0	NA	0	0	NA	2.2	75	26	0	1.6	5	13.1	69.4	14.41	6.9	0.2	1.7	0	3.5
FMq-CF3	8.8	0	NA	NA	NA	0	NA	16.7	NA	0	0	NA	15.6	25	8.2	0	4.4	13.8	4.2	9.2	73.55	3.1	0.6	4.3	0	5.7
FMq-CF4	24.6	11.1	NA	NA	NA	0	NA	16.7	NA	20	57.1	NA	28.9	0	0	0	13	17.2	10.2	8.3	4.27	69.2	3.1	0.5	0	27
FMq-CF5	22.8	33.3	NA	NA	NA	0	NA	16.7	NA	60	38.1	NA	40	0	0	0	19.7	12.4	6	0.8	0.18	5.8	94.4	0.2	0	34.8
LF	26.7	0	NA	NA	NA	100	NA	33.3	NA	20	0	NA	8.9	0	45.6	100	56.4	46.8	0.6	0.1	2.74	1.4	0.4	93	16.7	20.1
Acup	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0.9	0.1	0	0	0	0	0.1	83.3	0.5
QCF	0	0	NA	NA	NA	0	NA	0	NA	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N calls	285	9	0	0	0	25	0	6	0	5	21	0	45	4	158	20	631	581	4029	7846	12395	19859	61887	4210	6	2001

Table S4.7 Overall classifier performance

Classifier performance values for each confidence threshold in 5% increments. *Lowest threshold at which an F1-score > 0.9 was achieved; NR – threshold not achieved.

Model	ID	Threshold	Precision	Recall	F1-score	
Call type						
	FM	60	0.498	0.943	0.652	
		65	0.583	0.951	0.722	
		70	0.712	0.962	0.818	
		75	0.790	0.968	0.870	
		80	0.889	0.975	0.930	*
		85	0.929	0.979	0.953	
		90	0.970	0.986	0.978	
		95	0.988	0.990	0.989	
	CF	60	0.973	0.956	0.964	*
		65	0.978	0.963	0.970	
		70	0.983	0.971	0.977	
		75	0.987	0.975	0.981	
		80	0.992	0.980	0.986	
		85	0.995	0.983	0.989	
		90	0.997	0.988	0.993	
		95	0.999	0.993	0.996	
	FMqCF	60	0.990	0.943	0.966	*
		65	0.992	0.951	0.971	
		70	0.994	0.961	0.977	
		75	0.995	0.966	0.981	
		80	0.997	0.974	0.985	
		85	0.998	0.977	0.988	
		90	0.999	0.982	0.990	
		95	0.999	0.986	0.993	

	QCF	60	0.919	0.967	0.942	*
		65	0.926	0.973	0.949	
		70	0.935	0.980	0.957	
		75	0.941	0.985	0.962	
		80	0.951	0.991	0.970	
		85	0.957	0.994	0.975	
		90	0.968	0.997	0.982	
		95	0.977	0.998	0.988	
CF						
	H140	60	0.682	1	0.811	
		65	0.753	1	0.859	
		70	0.795	1	0.885	
		75	0.838	1	0.912	*
		80	0.873	1	0.932	
		85	0.879	1	0.936	
		90	0.940	1	0.969	
		95	0.956	1	0.977	
	Hate	60	0.769	1	0.870	
		65	0.909	1	0.952	*
		70	0.909	1	0.952	
		75	1	1	1	
		80	1	1	1	
		85	1	1	1	
		90	1	1	1	
		95	1	1	1	
	Hbic	60	1	1	1	*
		65	1	1	1	
		70	1	1	1	
		75	NA	NA	NA	
		80	NA	NA	NA	

	85	NA	NA	NA	
	90	NA	NA	NA	
	95	NA	NA	NA	
Hcer	60	0.915	0.995	0.953	*
	65	0.914	0.995	0.953	
	70	0.912	0.994	0.952	
	75	0.917	0.994	0.954	
	80	0.914	1	0.955	
	85	0.908	1	0.952	
	90	0.900	1	0.947	
	95	0.873	1	0.932	
Hdia	60	0.272	0.988	0.426	
	65	0.291	1	0.451	
	70	0.296	1	0.457	
	75	0.316	1	0.480	
	80	0.347	1	0.516	
	85	0.344	1	0.512	
	90	0.456	1	0.626	
	95	0.538	1	0.700	NR
Hgal	60	0.959	1	0.979	*
	65	0.957	1	0.978	
	70	0.957	1	0.978	
	75	0.969	1	0.984	
	80	1	1	1	
	85	1	1	1	
	90	1	1	1	
	95	1	1	1	
Hlar	60	1	1	1	*
	65	1	1	1	

	70	1	1	1	
	75	1	1	1	
	80	1	1	1	
	85	1	1	1	
	90	1	1	1	
	95	NA	NA	NA	
Hrid	60	0.976	0.909	0.941	*
	65	1.000	0.975	0.987	
	70	1.000	0.974	0.987	
	75	1	1	1	
	80	1	1	1	
	85	1	1	1	
	90	1	1	1	
	95	1	1	1	
Racu	60	0.847	1	0.917	*
	65	0.878	1	0.935	
	70	0.897	1	0.946	
	75	0.945	1	0.972	
	80	0.971	1	0.985	
	85	1	1	1	
	90	1	1	1	
	95	1	1	1	
Raff	60	1	1	1	*
	65	1	1	1	
	70	1	1	1	
	75	1	1	1	
	80	1	1	1	
	85	1	1	1	
	90	1	1	1	
	95	1	1	1	

Rbor	60	0.622	0.986	0.762	
	65	0.623	0.985	0.763	
	70	0.634	0.985	0.771	
	75	0.644	0.987	0.779	
	80	0.666	0.989	0.796	
	85	0.715	0.988	0.830	
	90	0.763	0.987	0.861	
	95	0.840	0.986	0.907	*
Rcre	60	1	1	1	*
	65	1	1	1	
	70	1	1	1	
	75	1	1	1	
	80	1	1	1	
	85	1	1	1	
	90	NA	NA	NA	
	95	NA	NA	NA	
Rluc	60	0.630	0.998	0.773	
	65	0.693	1	0.819	
	70	0.753	1	0.859	
	75	0.818	1	0.900	
	80	0.885	1	0.939	*
	85	0.927	1	0.962	
	90	0.953	1	0.976	
	95	0.977	1	0.988	
Rphi	60	0.487	1	0.655	
	65	0.529	1	0.692	
	70	0.587	1	0.740	
	75	0.667	1	0.800	
	80	0.795	1	0.886	

		85	0.897	1	0.946	*
		90	0.944	1	0.971	
		95	1	1	1	
	Rsed	60	0.990	0.974	0.982	*
		65	0.991	0.976	0.984	
		70	0.993	0.978	0.986	
		75	0.995	0.981	0.988	
		80	0.997	0.984	0.990	
		85	0.998	0.987	0.993	
		90	0.999	0.991	0.995	
		95	1.000	0.996	0.998	
	Rtri	60	0.920	0.998	0.958	*
		65	0.921	0.999	0.959	
		70	0.923	0.999	0.959	
		75	0.924	1.000	0.960	
		80	0.925	1.000	0.961	
		85	0.927	1.000	0.962	
		90	0.934	1.000	0.966	
		95	0.947	1.000	0.973	
FMqCF						
	FMqCF1	60	0.700	0.790	0.742	
		65	0.766	0.806	0.786	
		70	0.809	0.814	0.812	
		75	0.825	0.811	0.818	
		80	0.847	0.799	0.822	NR
		85	0.862	0.776	0.817	
		90	0.884	0.765	0.820	
		95	0.896	0.742	0.812	
	FMqCF2	60	0.817	0.827	0.822	

	65	0.859	0.850	0.855	
	70	0.892	0.868	0.880	
	75	0.927	0.881	0.903	*
	80	0.946	0.885	0.914	
	85	0.969	0.887	0.927	
	90	0.993	0.856	0.919	
	95	1.000	0.700	0.824	
FMqCF3	60	0.938	0.903	0.920	*
	65	0.949	0.924	0.936	
	70	0.956	0.937	0.947	
	75	0.965	0.950	0.957	
	80	0.973	0.959	0.966	
	85	0.980	0.965	0.972	
	90	0.985	0.968	0.976	
	95	0.988	0.967	0.977	
FMqCF4	60	0.884	0.837	0.860	
	65	0.901	0.849	0.874	
	70	0.916	0.853	0.884	
	75	0.930	0.852	0.889	NR
	80	0.945	0.840	0.889	
	85	0.962	0.814	0.882	
	90	0.969	0.748	0.844	
	95	0.983	0.532	0.690	
FMqCF5	60	0.969	0.987	0.978	*
	65	0.972	0.991	0.981	
	70	0.975	0.993	0.984	
	75	0.979	0.994	0.987	
	80	0.982	0.996	0.989	
	85	0.986	0.997	0.991	
	90	0.990	0.997	0.993	

	95	0.993	0.997	0.995	
LF	60	0.717	0.970	0.825	
	65	0.732	0.974	0.836	
	70	0.746	0.979	0.847	
	75	0.759	0.985	0.857	
	80	0.774	0.989	0.869	
	85	0.795	0.993	0.883	
	90	0.817	0.995	0.897	
	95	0.845	0.997	0.915	*
Acup	60	0.556	0.833	0.667	NR
	65	0.500	0.667	0.571	
	70	NA	NA		
	75	NA	NA		
	80	NA	NA		
	85	NA	NA		
	90	NA	NA		
	95	NA	NA		

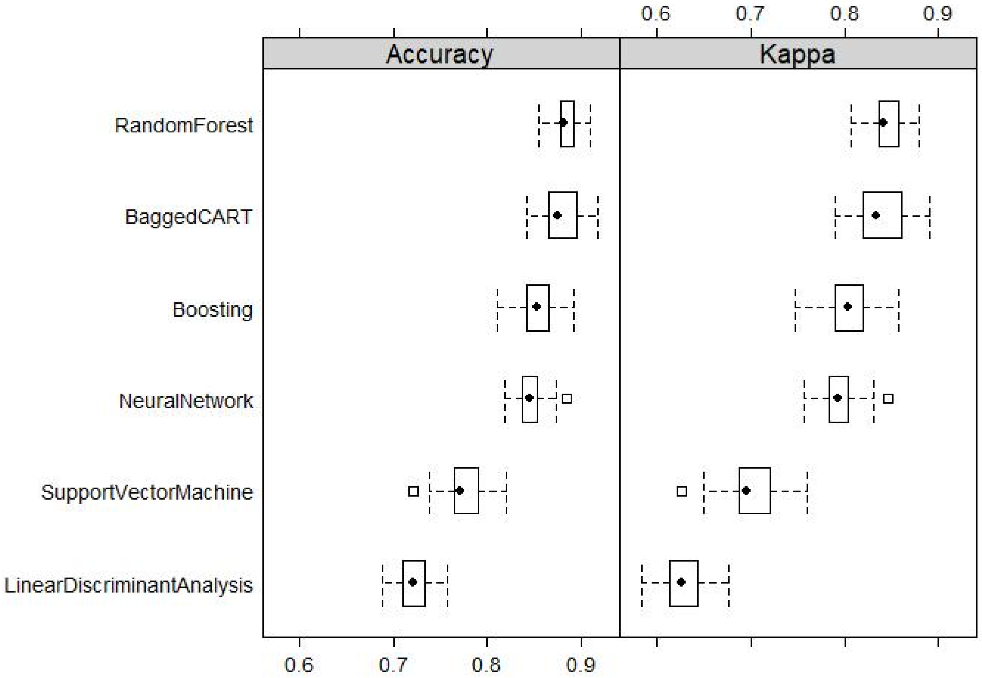


Figure S4.1 Comparison of the performance of six supervised machine learning algorithms Comparison of the performance of six supervised machine learning algorithms (bagging, boosting, artificial neural networks support vector machines, linear discriminant analysis, random forests) using training data of 5000 calls. There was little difference between random forests and bagging so either algorithm would have been equally viable for building the classifier. However, the overall variance was lower for random forests so this algorithm was chosen for the final design. Dots are medians, boxes 25% and 75% quartiles and whiskers denote the range.

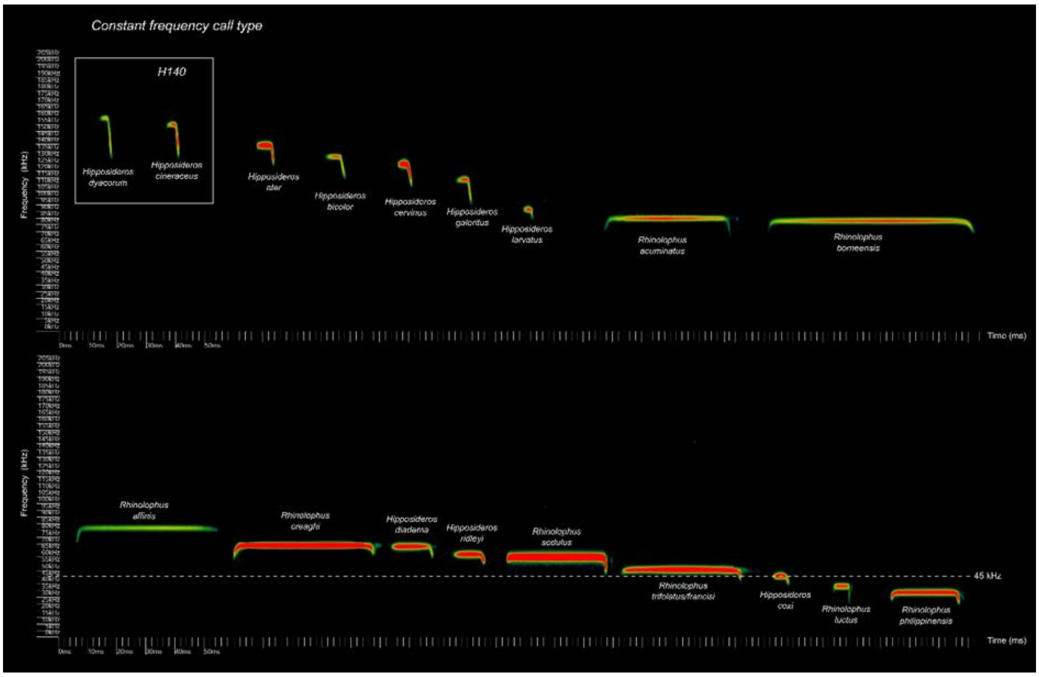


Figure S4.2 Sonograms for "constant frequency" species

Echolocation call design from each species included in the call type "constant frequency". Call duration and frequency have been scaled to allow comparison between species. Sonogram image is for illustrative purposes and should not be used in isolation for species identification. Call amplitude represented here is not standardised between species and therefore should not be used as a diagnostic feature.

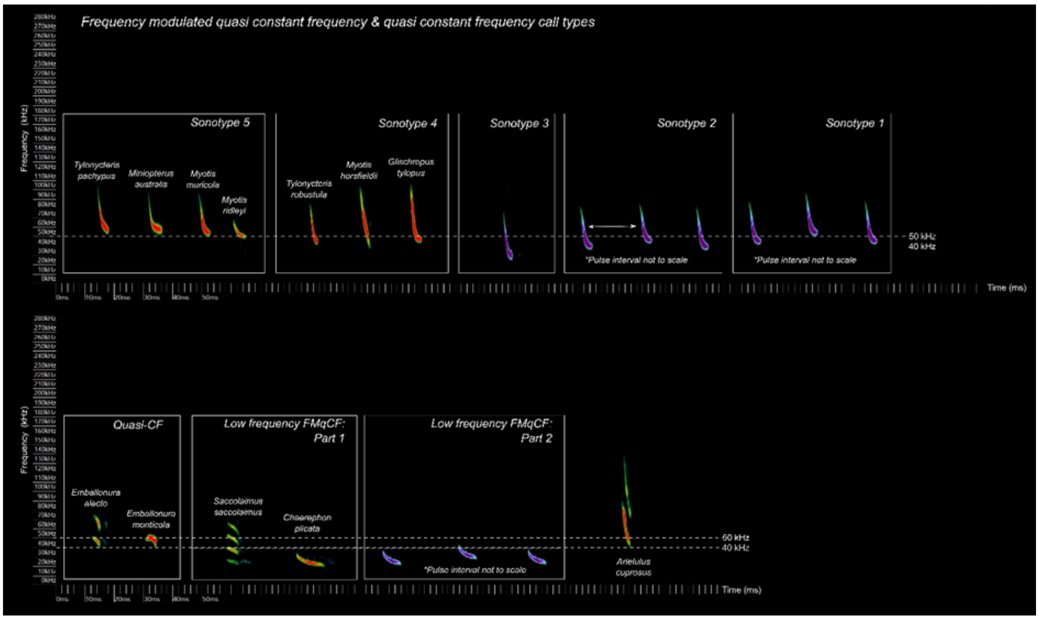


Figure S4.3 Sonograms for "frequency modulated quasi constant frequency" and "quasi constant frequency" sonotypes/species

Echolocation call design from each species/sonotype included in the call types "frequency modulated quasi constant frequency" and "quasi constant frequency". Call duration and frequency have been scaled to allow comparison between species/sonotypes. The sonogram image is for illustrative purposes and should not be used in isolation for species/sonotype identification. Call amplitude represented here is not standardised between species and therefore should not be used as a diagnostic feature. Calls displayed in blue/purple are artist representations of sonotype calls, not true calls

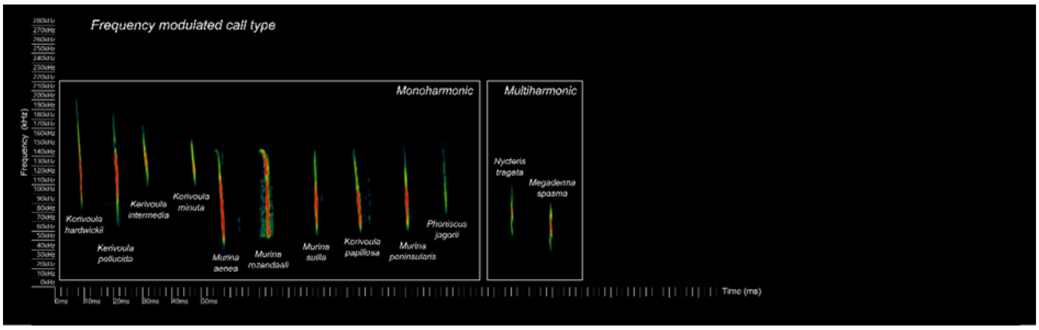


Figure S4.4 Sonograms for “frequency modulated” call type

Echolocation call design from each species included in the call type “frequency modulated”. Calls have been separated by whether they are monoharmonic or multi-harmonic. Call duration and frequency have been scaled to allow comparison between species. The sonogram image is for illustrative purposes and should not be used in isolation for species identification. Call amplitude represented here is not standardised between species and therefore should not be used as a diagnostic feature.

Supplementary Notes. Information for collecting training dataset for Borneo bat acoustic classifier

Reference calls from captured identified bats were collected from in-hand recordings, upon release, or from within a flight tent (Supplementary Table 4.1). For release calls, detectors were positioned towards each bat as it was released, and recordings continued until the bat was out of range (typically 2 seconds). The first 5 seconds or first ten calls in each recording were discarded to remove non-natural calls. Rhinolophid and hipposiderid species were recorded in-hand to preclude Doppler Shift Compensation and capture the resting frequency of their echolocation signals. Field recordings were conducted in 26 locations in old-growth forest (3 locations; 74 hours), twice-logged forest (5 locations; 152 hours), heavily logged forest (11 locations; 345 hours), forest remnants (5 sites; 247 hours), as well as oil palm (1 site; 39 hours) and acacia plantations (1 site; 25 hours). Sites were surveyed acoustically using SongMeter 2 (SM2BAT) full spectrum detectors (Wildlife Acoustics Inc, USA). Calls were collected in the evening and early morning (18:00-06:20). Detectors were set to automatically record bats in real-time with a full spectrum resolution of 16 bit, with a sample rate of 384 kHz, on an adaptive trigger level relative to noise floor of 18 SNR across consecutive nights per location.

Literature list used for the species identifications:

Francis, C.M., Habersetzer, J. (1998). Interspecific and intraspecific variation in echolocation frequency and morphology of horse-shoe bats, *Rhinolophus* and *Hipposideros*, in: Kunz, T.H., Racey, P.A. (Eds.), *Bat Biology and Conservation*. Smithsonian Institution Press, Washington, DC, pp. 169–179.

Hughes, A.C., Satasook, C., Bates, P.J.J., Soisook, P., Sritongchuay, T., Jones, G., Bumrungsri, S. (2011). Using Echolocation Calls to Identify Thai Bat Species: Vespertilionidae, Emballonuridae, Nycteridae and Megadermatidae. *Acta Chiropterologica 13*, 447–455.

Kingston, T., Jones, G., Akbar, Z., Kunz, T.H. (2003). Alternation of Echolocation Calls in 5 Species of Aerial-Feeding Insectivorous Bats from Malaysia. *Journal of Mammalogy 84*, 205–215.

Kingston, T., Jones, G., Akbar, Z., Kunz, T.H. (1999). Echolocation signal design in Kerivoulinae and Murininae (Chiroptera: Vespertilionidae) from Malaysia. *Journal of Zoology 249*, 359–374.

Kingston, T., Jones, G., Zubaid, A., Kunz, T.H. (2000). Resource Partitioning in Rhinolophoid Bats Revisited. *Oecologia 124*, 332–342.

McKenzie, N.L., Bullen, R.D. (2019). What can echolocation recordings reveal about the foraging ecology of *Saccolaimus saccolaimus* (Emballonuridae) in north-western Australia? *Australian Journal of Zoology 66*, 326–334.

Rajasegaran, P. (2019). Population Genetic and Ecological Studies of the Bornean Endemic Cox's Leaf-nosed Bat (Hipposideridae: *Hipposideros coxi*) in Sarawak. Masters thesis, Universiti Malaysia Sarawak (UNIMAS).

Soisook, P., Struebig, M.J., Noerfahmy, S., Bernard, H., Maryanto, I., Chen, S.-F., Rossiter, S.J., Kuo, H.-C., Deshpande, K., Bates, P.J.J., Sykes, D., Miguez, R.P. (2015). Description of a New Species of the *Rhinolophus trifoliatus*-Group (Chiroptera: Rhinolophidae) from Southeast Asia. *Acta Chiropterologica 17*, 21–36.

Stoffberg, S., Jacobs, D.S., Matthee, C.A. (2011). The Divergence of Echolocation Frequency in Horseshoe Bats: Moth Hearing, Body Size or Habitat? *Journal of Mammalian Evolution 18*, 117–129.

Thabah, A., Rossiter, S.J., Kingston, T., Zhang, S., Parsons, S., Mya, K.M., Zubaid, A., Jones, G.W. (2006). Genetic divergence and echolocation call frequency in cryptic species of *Hipposideros larvatus* (Chiroptera: Hipposideridae) from the Indo Malayan region. *Biological Journal of the Linnean Society 88*(1), 119–130.

Yoon, K.B., Park, Y.C. (2016). Echolocation Call Structure and Intensity of the Malaysian *Myotis muricola* (Chiroptera: Vespertilionidae). *Journal of Forest and Environmental Science 32*, 99–102.

Chapter 5 Benign effects of logging on aerial insectivorous bats in the palaeotropics revealed by remote sensing technologies

Natalie Yoh[1], Dave J. I. Seaman[1], Nicolas J. Deere[1], Henry Bernard[2], Jake E. Bicknell[1] and Matthew J. Struebig[1]

[1.] Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, Canterbury, Kent, CT2 7NR, UK

[2.] Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah, Kota Kinabalu, Malaysia

Manuscript under review with the Journal of Applied Ecology

Author Contributions: MS, HB, and JB conceived and designed the study. MS led the fieldwork. NY processed the bat calls. NY, DS, and NJD performed the statistical analysis. NY, JB, and MS wrote the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

5.1 Abstract

1. Logging is the most widespread disturbance in tropical forests, altering ecological communities, functions, and services. However, many species can persist in logged forests, particularly where disturbance levels are comparatively low. Despite a growing knowledge of logging effects on wildlife, there remains little information for palaeotropical bats, in part due to major challenges in monitoring, particularly for aerial insectivorous species.

2. We integrated remote sensing data from passive acoustic bat detectors with LiDAR-derived measures of forest structure from a human-modified landscape in Sabah, Borneo. Our appraisal benefitted from a semi-automated classifier of bat calls that vastly speeds up the analysis of acoustic recording data. Across a gradient of habitat disturbance comprising old-growth forest, repeatedly logged forest, and tree plantations, we recorded 105,576 bat passes from 21 taxa.

3. We show that logging pressure (as depicted by changes to habitat quality e.g., canopy height or forest structure) had negligible impact on the acoustic diversity and activity of common bats. However, the extent of high-biomass forest was positively associated with bat activity, as was local topography. Logged forest supported higher levels of activity for several common taxa compared to old-growth forest. Across the landscape, plantations supported the lowest levels of bat activity, and several species were not recorded from this habitat.

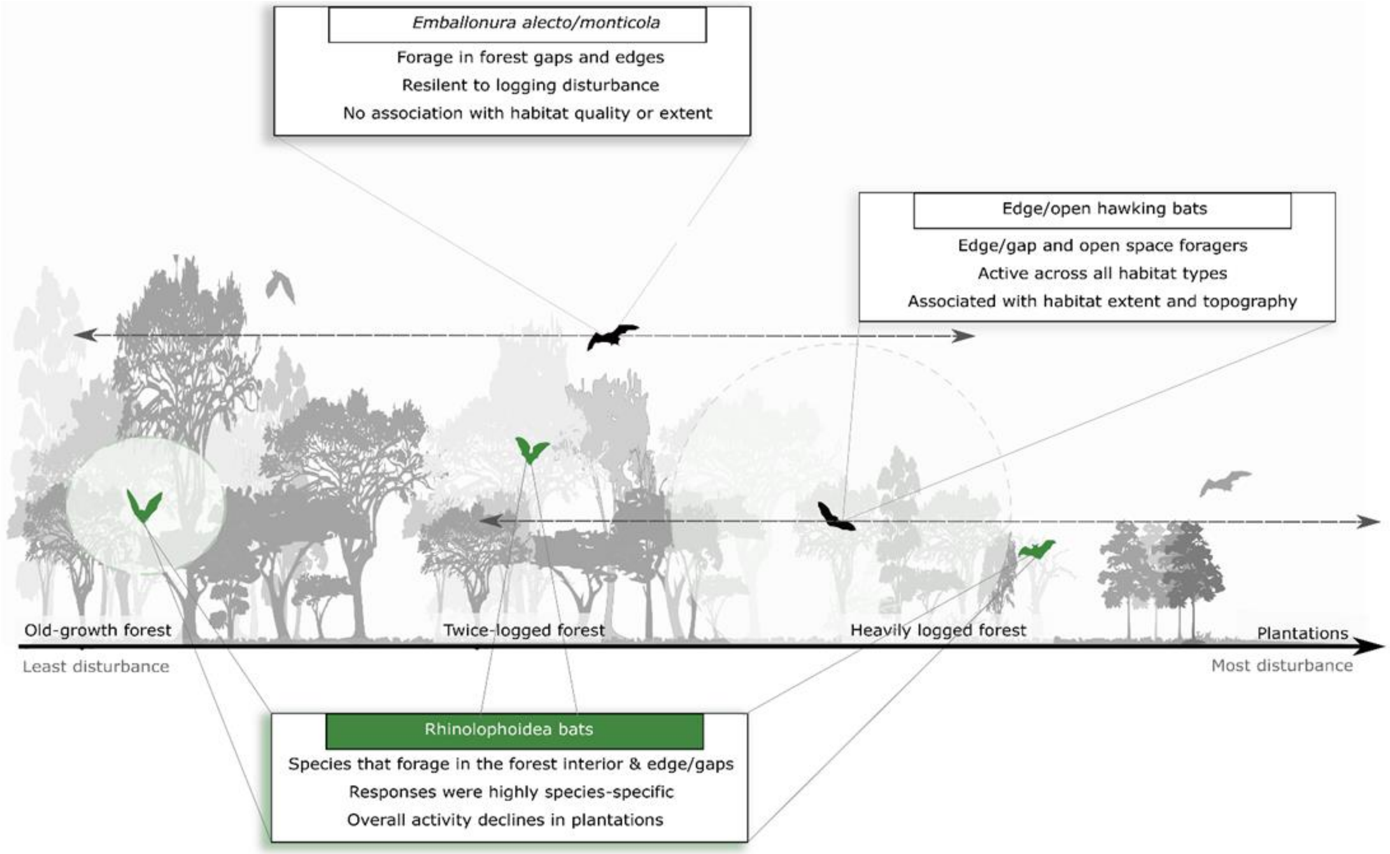
4. We found different call groups demonstrated different responses to forest disturbance. Sheath-tailed bats (*Emballonura* spp.) were active across all habitat types and were the most resilient to logging. Edge hawking bats were more prevalent in highly forested and topographically-rugged areas. Horseshoe and leaf-nosed bats (Rhinolophoidea) demonstrated idiosyncratic responses to logging but were consistently absent from plantations.

5. *Synthesis and applications.* Our findings demonstrate how coupling information from acoustic detectors, automatic classification algorithms, and LiDAR data can be used to assess the impacts of land-use change on underrepresented taxa. In doing so, we show that logged areas can provide an important refuge for many common bat species in palaeotropical forest, but do not capture the full breadth of

forest-specialist species found in old-growth forests. Nevertheless, logged forests provided a substantially better habitat for bats than industrial tree plantations.

Keywords: Chiroptera; Southeast Asia; acoustic monitoring; land-use change, biodiversity; remote sensing

5.2 Graphical abstract



5.3 Introduction

Tropical forests provide valuable habitat for some of the highest levels of biodiversity in the world, while contributing to the livelihoods of millions of people (Edwards, Tobias, *et al.* 2014; Meijaard *et al.* 2005; Putz *et al.* 2012). More than 25% of the world's tropical forests are allocated to logging (Edwards, Tobias, *et al.* 2014), making this activity the most widespread disturbance for these globally-important ecosystems (Blaser *et al.* 2011; Costantini, Edwards and Simons 2016). As tropical forests continue to be logged, and often subsequently converted for other land-uses (primarily agriculture), it is important to understand how these developments could impact biodiversity.

Logged forests are increasingly recognised as valuable wildlife habitats. A pantropical meta-analysis comprising 48 studies showed that lightly logged forests (e.g. < 10 m^3 ha^{-1}) can support more species of mammals, amphibians, and invertebrates than old-growth forests (Burivalova, Şekercioğlu and Koh 2014). This implies that logged forests should be retained or better protected for biodiversity conservation rather than being cleared for other land uses (Meijaard and Sheil, 2007; Putz *et al.*, 2012). However, logging disturbance is not uniform. The extent of disturbance - and hence impact on biodiversity - depends on the number of logging cycles, logging intensity, and extraction techniques (Bicknell *et al.*, 2014). Light Detection and Ranging (LiDAR) remote sensing provides a method through which the habitat degradation caused by logging activities can be characterised. Using LiDAR, several studies have shown how logging can impact tropical forest structure, including reducing above-ground biomass and impacting canopy structure (e.g., increasing canopy gap area, reducing canopy height) (d'Oliveira *et al.* 2012; Kent *et al.* 2015; Rangel Pinagé *et al.* 2019). While logging can be much less detrimental than the conversion of forest to other land uses (Gibson *et al.* 2011; Wearn *et al.* 2017), it can still cause substantial population declines among specialist taxa (Burivalova, Şekercioğlu and Koh 2014; Thorn *et al.* 2018).

Despite the extensive literature on the effects of logging on biodiversity, there have been relatively few studies investigating how logging impacts tropical bats. This is important because bats represent the second most diverse mammalian order, with over 1,400 species globally (Simmons and Cirranello 2021a). Bats also provide valuable ecosystem services as seed dispersers, pollinators, and suppressors of insect populations (Kunz *et al.* 2011). As such, bats are considered important bioindicators of ecosystem health (Jones *et al.* 2009). Presley *et al.* (2008) outlined three potential demographic responses for bats in logged forests: (1) common species remain common but decrease in abundance; (2) common species decline in abundance, becoming rare in logged forests; or (3) rare species are locally absent from logged forest. Studies so far have found that bat responses are highly variable between different ensembles of species, particularly dietary guilds (Bicknell *et al.,* 2015; Castro *et al.*, 2021; Castro-Arellano *et al.*, 2009; Clarke *et al.*, 2005; Presley *et al.*, 2008). Small, common phytophagous species (i.e. frugivorous and nectarivorous species) that rely on pioneering plants for foraging opportunities often respond positively to logging disturbance (Clarke, Rostant and Racey 2005; Presley *et al.* 2008). On the other hand, larger, insectivorous or animalivorous phyllostomid species are more vulnerable to logging, as well as habitat disturbance more generally, and experience declines in diversity and abundance (Brändel *et al.*, 2020; Clarke *et al.*, 2005; Farneda *et al.,* 2015).

The study of bats in tropical regions has been restricted by limitations in sampling techniques. Most studies rely on live-capture techniques (primarily mist-netting or harp trapping) which are logistically challenging and introduce bias in the taxonomic and ecological coverage of focal species sampled. Mist-netting is the main method applied in the neotropics as it is highly effective for sampling phyllostomid bats. As a result, most logging-effect studies involving bats have been undertaken in the American tropics (e.g., Bicknell *et al.,* 2015; Castro *et al.,* 2021; Castro-Arellano *et al.,* 2009; Clarke *et al.,* 2005). In contrast, much less research has been undertaken in the Asian and African tropics as the bat fauna is dominated by aerial insectivorous taxa (i.e. those that use ultrasonic calls for orientation and foraging insect prey), which are poorly represented using nets (Castro & Michalski, 2014; Meyer *et al.,* 2016). Of the studies undertaken, most have been conducted in Malaysia and Indonesia using mist-nets (Danielsen and Heegaard 1995; Zubaid 1993) or harp traps (Joann Christine, Fletcher and Abd Rahman 2013; Matthew J. Struebig *et al.* 2013). These capture-based studies have demonstrated that logging impacts community composition, reduces species richness, and causes abundance declines (Danielsen and Heegaard 1995; Joann Christine, Fletcher and Abd Rahman 2013; Zubaid 1993). The most comprehensive study of bat responses to logging in the region examined bat assemblages across a habitat disturbance gradient, comparing old-growth forest to logged forest areas that had undergone various extraction cycles (Struebig *et al.* 2013). There was no discernible effect of logging on bat species richness, but species composition differed significantly between old-growth and repeatedly logged forest sites. This shift in community composition mirrors similar changes observed in the neotropical bat fauna in response to logging (Castro-Arellano *et al.* 2009; Clarke, Rostant and Racey 2005). However, even harp traps are biased to a subset of the insectivorous bat fauna that can be readily captured in cluttered environments such as forests, and so miss many of the echolocating fauna utilising other habitats. Therefore, more research is needed to determine how different logging disturbances impact bat assemblages in palaeotropical regions.

Here, we quantify bat diversity and activity along a disturbance gradient from protected, old-growth forest, through different degrees of logging pressure, to plantations. Our study is based in Sabah, Borneo, which supports high levels of vertebrate biodiversity, and high-value timber species (Brookfield and Byron 1990). The forest landscape has a legacy of high logging intensity (cumulative extraction rates between 150 and 179 m^3 ha^{-1}) (Matthew J. Struebig *et al.* 2013). As 72% of Borneo’s bat species are echolocating insectivores (Phillipps & Phillipps, 2016), we use acoustic monitoring coupled with a novel classification algorithm as an unintrusive alternative to traditional capture methods. We subsequently examine how bats respond to landscape changes using airborne LiDAR of forest metrics. Specifically, we sought to use remote sensing technologies to characterise (i) acoustic richness, (ii) community composition, and (iii) bat activity in each disturbance type, while also (iv) assessing possible relationships between habitat extent and quality and bat activity.

5.4 Materials and methods

5.4.1 Study area and sampling design

Fieldwork was undertaken within the Kalabakan and Ulu Segama Forest Reserves in and around the Stability of Altered Forest Ecosystems Project (SAFE; www.safeproject.net) (443’N, 117^{o}35’E) in Sabah, Malaysian Borneo (Figure 51). The landscape is characterised by lowland and hill dipterocarp forest, which was initially logged in the 1970s followed by two subsequent rotations in the late 1990s and early 2000s. The forest in Ulu Segama was left to recover (‘twice-logged forest’; 3 sampling blocks; LFE, LF2, LF3) (Supplementary Table 5.1). However, except for designated fragments and riparian reserves, the forest in Kalabakan was subject to salvage logging (the removal of all commercially valuable trees: ‘heavily logged forest’; 6 blocks; A-F). This highly disturbed forest has a high density of roads and skids, few emergent trees, and is dominated by pioneer and invasive vegetation (Struebig *et al.,* 2013). The landscape adjoins ca. 1 million ha of continuous forest, including old-growth forest in Maliau Basin Conservation Area (‘Old Growth Forest’; 2 blocks; OG2, OG3). The forest is surrounded by a mosaic of tree plantations, mostly oil palm (*Elaeis guineensis*), but also Acacia (*Acacia mangium*; ‘Tree Plantation; 1 block; AC1) planted around 2000.

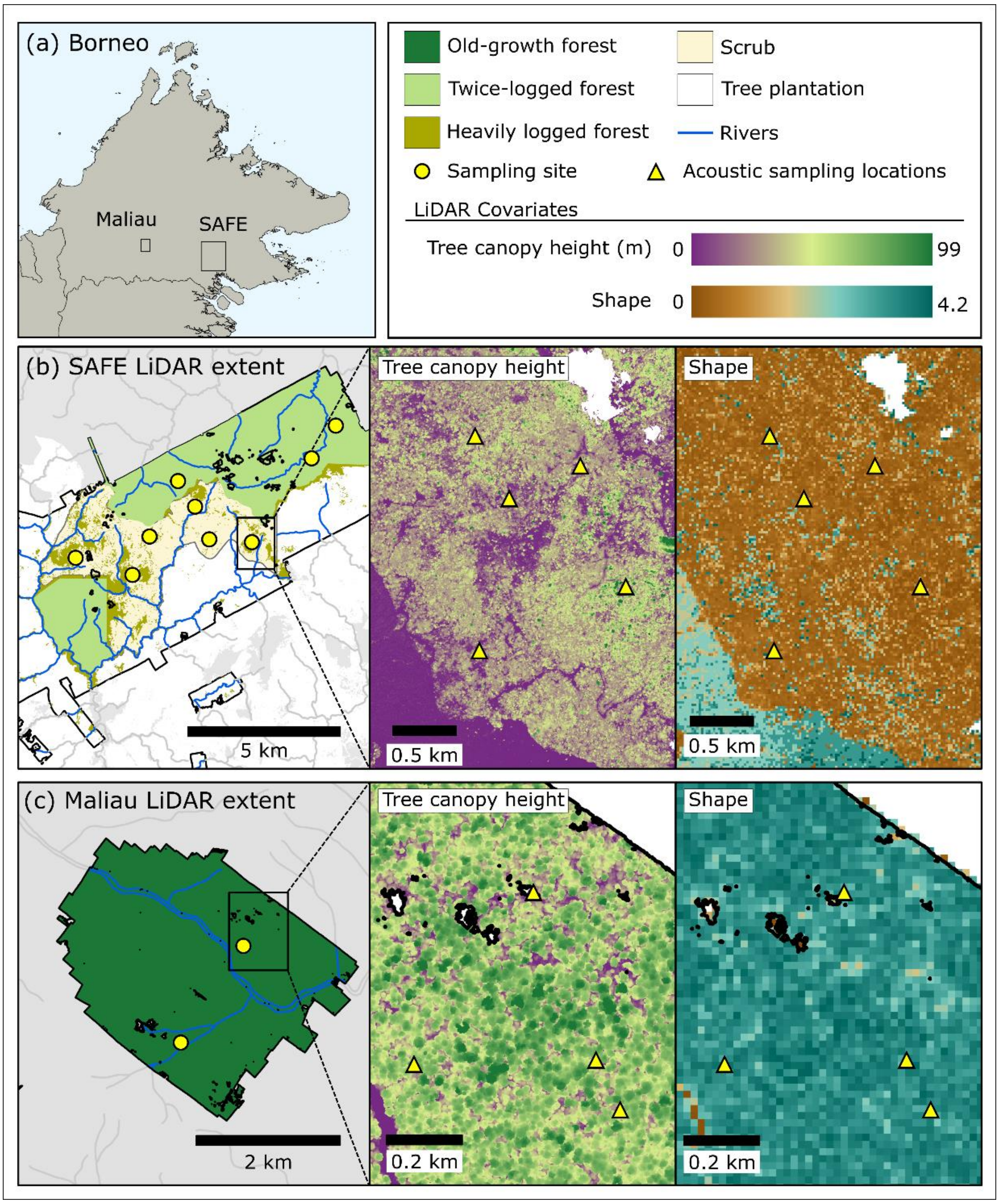


Figure 5.1 Map of the Stability of Altered Forest Ecosystem Project (SAFE) and sampling sites in Maliau Basin, Malaysian Borneo

(a) SAFE and Maliau in Malaysian Borneo, (b) sample blocks across SAFE and examples of two LiDAR variables across multiple survey sites, and (c) the two sampling blocks in Maliau with the corresponding LiDAR coverage.

5.4.2 Defining the forest disturbance gradient

Logging activity varied substantially over the landscape. We characterised the resulting structural disturbance to forests using airborne light detection and ranging (LiDAR) products to quantify the extent (i.e., habitat quantity) and structure (i.e., habitat quality) of forest, and hence define a disturbance gradient around our sampling sites. Nine covariates were extracted from 30 m resolution LiDAR layers produced in November 2014 from a Leica ALS50-II sensor (Jucker *et al.*, 2018). Of these, three were ultimately included in the modelling. From these data, we utilised information on aboveground vegetation biomass (AGB), which was calculated via a Borneo-specific model (see Jucker *et al.* 2018 for full details). To distinguish forest from non-forest, and hence estimate *ForestExtent* around sample sites (i.e., habitat quantity; binary; forest vs not forest), we applied a threshold to the AGB maps above which cells were classified as forest. Our definition of forest excluded areas of plantation and highly degraded Young Regenerating Forest, and was defined as vegetation biomass ≥160 t ha^{-1} (i.e. equivalent to above-ground carbon ≥ 75 $tCha^{-1}$; Martin and Thomas, 2011) following the High Carbon Stock Approach (HCS; The HCS Approach Toolkit, 2017). We chose a more conservative forest categorisation, as areas < 160 t ha^{-1} biomass were subject to salvage logging and therefore, more closely resemble areas of scrub and forest clearance.

We also extracted two measures of forest structure as proxies of habitat quality: mean canopy height (*CanopyHeight;* a continuous variable) and plant area density shape (*Shape;* continuous; morphological measurement of the relative distribution of vegetation within the canopy) (Jucker, Asner, *et al.* 2018; Swinfield, Milodowski, *et al.* 2020). There is little research regarding bats and LiDAR metrics, but *Shape* is known to be an important predictor of other forest-dependent mammal species (Nicolas J. Deere *et al.* 2020). Previous research highlighted that *CanopyHeight* and average topographic ruggedness (*Topography;* continuous) can be important predictors for bat occurrence in the project area (Mullin *et al.* 2020). Therefore, after the initial exploratory analysis, we included *Topography* which was measured within 50 m of each point using Shuttle Radar Topography Mission data, and converted from a digital elevation model to a topographic ruggedness index, according to Wilson, O'Connell, Brown, Guinan, and Grehan (2007). All covariates were extracted from the maps at a 100, 200, and 500 m buffer radius. Final models were analysed using 100 m as these models best fit the data. Spatial data processing was implemented in ArcGIS 10.2.1 (ESRI).

5.4.3 Acoustic sampling, processing, and bat identification

Bats were recorded using SongMeter-2 automated recorders (Wildlife Acoustics) fitted with an omnidirectional ultrasonic SMX-US microphone between April and May 2011 and April and June 2012 (Supplementary Table 5.1). Recorders were stationed across 47 sample points at 1.5-2.0 m height and set to record nightly bat activity (18:15-06:15) between one and three consecutive nights. Each recorder was set to record by triggers of high frequency sounds appropriate for echolocating bats in the region (sampling rate 384 kHz, 16-bit resolution; high pass filter 12 kHz (fs/32); trigger level 18 signal to noise ratio).

All preprocessing of the recordings was undertaken using Kaleidoscope v.5.1.9g software (Wildlife Acoustics) following procedures outlined in Yoh *et al*. (2022). Files were divided into 5-second long sequences to define a bat pass as a unit of activity (Torrent *et al.* 2018). Calls were automatically assigned to call type, sonotype (a group of species whose calls are indistinguishable; López-Baucells *et al.*, 2021), or species, and any calls that failed to meet specific confidence thresholds were manually verified (Yoh *et al*. 2022). Through this process, calls were first identified to one of four call types (Supplementary Figure 5.2): frequency modulated (hereafter referred to as whispering bats), constant frequency (Rhinolophoidea bats), frequency modulated quasi constant frequency (edge/open hawking bats), and quasi-constant frequency (*Emballonura alecto/monticola*). Those identified as an edge/open hawking bat were automatically classified further to one of six sonotypes. Calls classified as a Rhinolophoid species were further differentiated to 17 taxa.

The classifier identified 158,563 files containing bat passes. Of these files, 71,482 included bat passes that required manual verification. All bat passes that did not meet the specified confidence thresholds (26,351 out of 71,482), excluding those for whispering bats, were manually identified in Kaleidoscope Viewer by a single researcher (NY) to reduce potential bias. Whispering bats represented the largest proportion of files to be manually verified (45,131 files). We checked 50% of whispering bat passes and found this call type was easily mistaken for environmental noise at low confidence and < 1% contained true bat passes. Therefore, we discarded whispering bat calls that did not meet the confidence threshold. Activity (the sum of bat passes per sonotype/species) was aggregated to date to standardise activity across the different sampling points.

5.4.4 Statistical analysis

5.4.4.1 Acoustic diversity along a disturbance gradient

We constructed species/sonotype accumulation curves from activity data to confirm that acoustic sampling was adequate for each habitat. Curves were rarefied via 9999 permutations in the R package iNEXT (Hsieh, Ma and Chao 2016), and repeated for each site to generate richness values and associated 95% confidence intervals. As a sonotype does not necessarily represent the same unit of diversity as a species (e.g., a sonotype can represent 2-10 species), we use call/sonotype richness to represent the acoustic diversity in the landscape, not the overall number of species.

5.4.4.2 Bat community composition and habitat association

We used non-metric multidimensional scaling (NMDS) ordinations based on Bray-Curtis dissimilarity coefficients of $\log_{10}$-transformed bat activity data to determine whether the structure of bat communities reflected the various habitat type or the disturbance gradient. One taxon (*Hipposideros cineraceus/dyacorum*) was removed from the analysis as it was only recorded on a single occasion. Ordinations were based on 9999 permutations in the R package "vegan" (Oksanen *et al.* 2020) specifying two dimensions. We used PERMANOVA from the package "RVAideMemoire" (Hervé 2021) to test for differences in bat community composition between habitats.

To determine if there were differences in bat activity between forest sites and tree plantations, we first compared activity between habitats using Kruskal-Wallis tests followed by Dunn post-hoc tests. Tests were implemented separately for each call type and also for overall levels of bat activity, and all *p*-values were adjusted using the Benjamini-Hochberg method to account for the risk of false positives in call detection (Haynes 2013). To determine if there was a relationship between habitat covariates and bat activity, we constructed Generalized Additive Models for Location, Scale and Shape (GAMLSS) appropriate for the Weibull distribution of these data (Stasinopoulos *et al.* 2020) - using the "GAMLSS" package. LiDAR datasets were not available for tree plantation sites and so were omitted from these analyses. Covariates with high collinearity ($r \geq 0.7$) were not included in the same models. Covariates were centred and scaled to one unit standard deviation. Models were performed separately for total bat activity, call type, and the four dominant taxa. Habitat type (*HabitatType*; categorical; three levels) was included to assess if there were differences in activity not explained by the other measures of habitat quality or extent. We then model-averaged the final global models using the dredge function from package "MumIn" to identify the best models (ΔAIC < 2; Supplementary Table 5.2). We inspected the final models for the goodness of fit using residual diagnostics (Burnham & Anderson, 2002). All analysis was performed using R version 4.1.1 (2021-08-10) statistical software (R Core Team 2017).

5.5 Results

From our data of remotely sensed bat activity, we identified 105,576 bat passes from 21 taxa across the landscape. More than 76% of all bat activity was represented by five taxa: *Emballonura alecto/monticola* (18.6% of calls), Sonotype 6 (17.3%), *Rhinolophus sedulus* (15.1%), Sonotype 5 (13.2%), and *Rhinolophus trifoliatus* (12.2%; Table 1). Two species (*H.*

cineraceus/dyacorum and *Rhinolophus philippinensis*) were only recorded in heavily logged forest, once and on three occasions respectively. For activity, the composition of the top five taxa was similar across habitats, consisting mostly of common, generalist species. Sonotype 6 and *E. alecto/monticola* were among the top five most active taxa across all habitat types (Table 5.1).

Table 5.1 Total bat passes per sonotype/species in each habitat along the disturbance gradient

For each habitat type, relative bat passes (total bat passes/total nights surveyed) are provided in brackets for comparison.

	Old-growth forest	Twice-logged forest	Heavily logged forest	Tree Plantation
Total nights surveyed	23	36	73	12
Total detector sites	8	12	23	4
Whispering bats (e.g., *Kerivoula* spp.)	628 (27.3)	1,988 (55.2)	2,556 (35.0)	49 (4.1)
Rhinolophoidea bats				
Hipposideros cineraceus/dyacorum	0	0	1 (< 0.1)	0
Hipposideros cervinus	5 (0.2)	54 (1.5)	9 (0.1)	22 (1.8)
Hipposideros diadema	248 (10.8)	36 (1.0)	33 (0.5)	7 (0.6)
Hipposideros galeritus	37 (1.6)	162 (4.5)	36 (0.5)	5 (0.4)
Hipposideros ridleyi	3 (0.1)	3 (0.1)	0	0
Rhinolophus acuminatus	229 (10.0)	165 (4.6)	262 (3.6)	91 (7.6)
Rhinolophus affinis	0	0	2 (< 0.1)	2 (0.2)
Rhinolophus borneensis	68 (3.0)	822 (22.8)	143 (2.0)	1 (0.1)

Rhinolophus creaghi	1	443	4	8
	(< 0.1)	(12.3)	(0.1)	(0.1)
Rhinolophus luctus	503	80	181	5
	(21.9)	(2.2)	(2.5)	(0.4)
Rhinolophus philippinensis	0	0	3	0
			(< 0.1)	
Rhinolophus sedulus	196	9,558	6,069	86
	(8.5)	(265.5)	(83.1)	(7.2)
Rhinolophus trifoliatus	1,703	1,611	9,542	42
	(74.0)	(44.8)	(130.7)	(3.5)
Edge/open hawking bats				
Sonotype 1	10	39	271	95
	(0.4)	(1.1)	(3.7)	(7.9)
Sonotype 2	241	6,151	980	108
(e.g., *Hesperoptenus blanfordi*)	(10.5)	(170.9)	(1.3)	(9.0)
Sonotype 3	2,373	504	2,053	100
(e.g., *Pipistrellus stenopterus*)	(103.2)	(14.0)	(28.1)	(8.3)
Sonotype 4	770	302	499	534
(e.g., *Glischropus tylopus*)	(3.3)	(8.9)	(6.8)	(44.5)
Sonotype 5	3,911	2,883	6,180	961
(e.g., *Myotis muricola*)	(170.0)	(80.1)	(84.7)	(80.1)
Sonotype 6	9,249	775	7,019	1,256
(e.g., *Chaerephon plicatus*)	(402.1)	(21.5)	(96.2)	(104.7)
Emballonura alecto/monticola	2,590	4,634	12,410	463
	(112.6)	(128.7)	(170.0)	(38.6)

5.5.1 Acoustic diversity

In three of the four habitats, acoustic richness reached or approached an asymptote after 10,000 bat calls were recorded (Figure 5.2). However, in the heavily logged forest, diversity was expected to continue increasing for at least double this volume of activity before levelling off. Overall, the acoustic diversity between sites was very similar, ranging between 15.8 and 17.2 taxa (Supplementary Table 5.3).

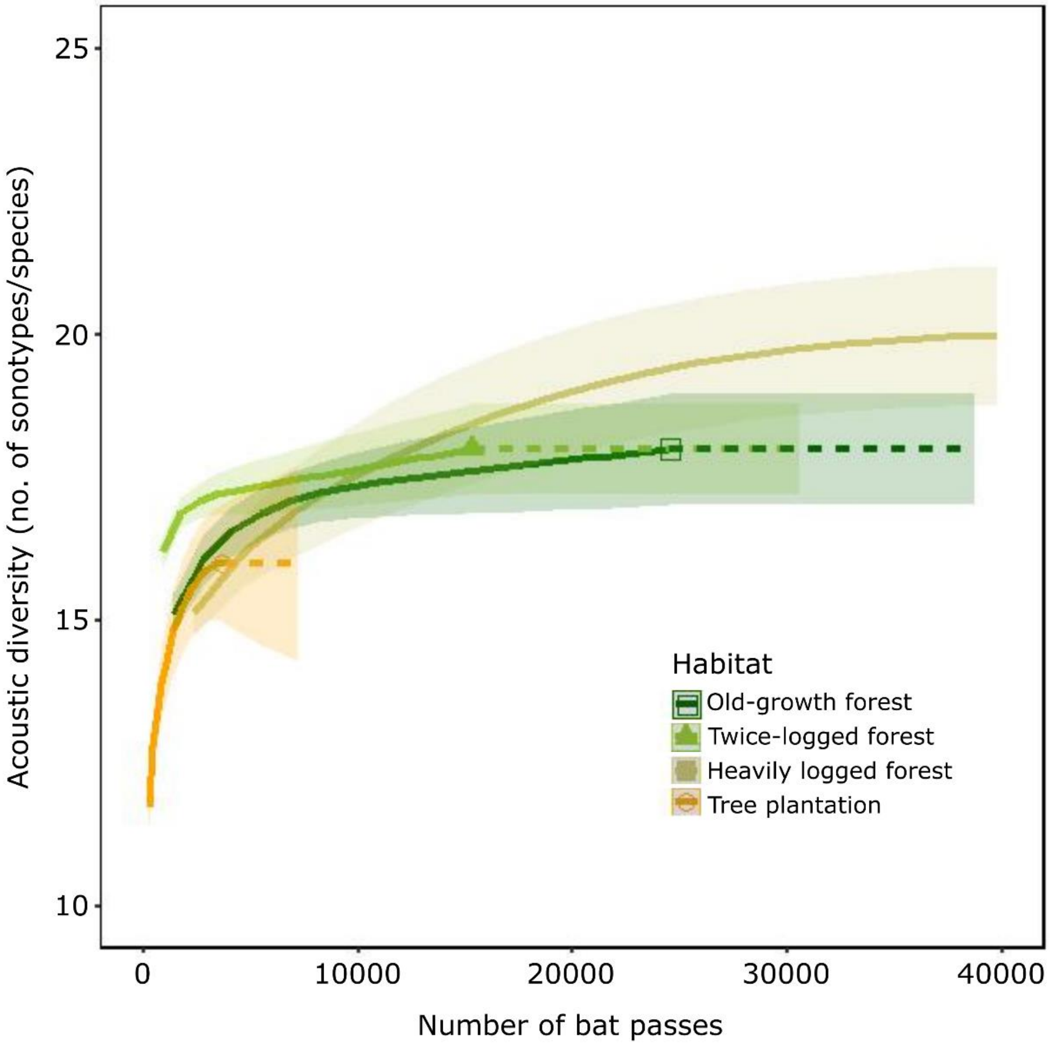


Figure 5.2 Species/sonotype diversity along the disturbance gradient

Rarefied accumulation curves of species/sonotype diversity for each of four habitat types based on 9999 permutations. Dashed lines indicate extrapolation beyond the bounds of observed data. Shading indicates the lower and upper 95% confidence limits.

5.5.2 Bat community composition and association with habitats

The NMDS ordination of bat activity revealed substantial variation in taxonomic composition across the landscape and overlap between the various habitat types (Figure 5.3). The greatest variation occurred among twice-logged sites and the least among tree plantation sites. Across the landscape, there was a significant difference in bat community composition between habitats (PERMANOVA: $R^2 = 0.17$, $p = 0.003$), driven by differences between the heavily logged forest and the twice-logged forest (pairwise comparisons $p = 0.02$).

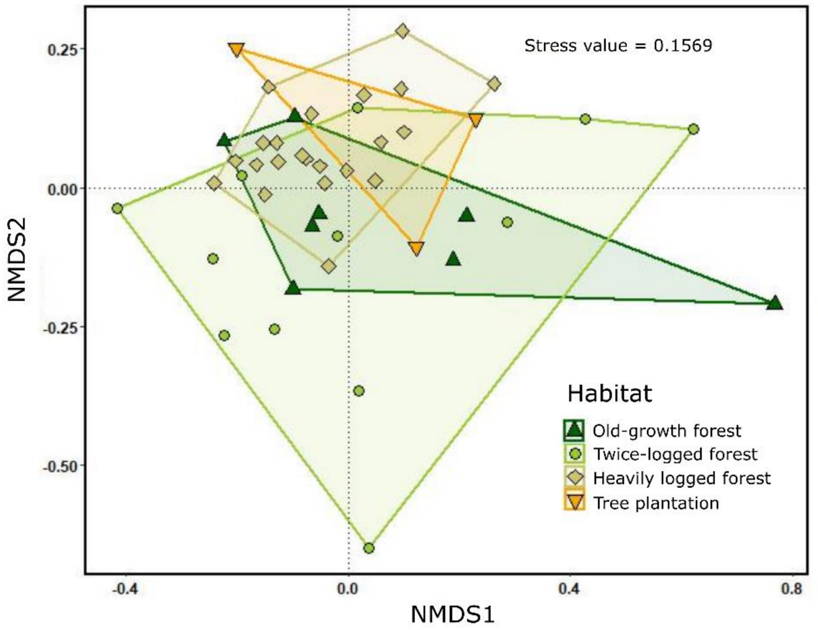


Figure 5.3 Nonmetric multi-dimensional scaling showing the (dis)similarities between bat communities across the four habitat types.

We found significant differences in bat activity between habitats, for total activity ($H(3) = 10.126$, $p = 0.018$) and for each call type (whispering bat activity - $H(3) = 25.483$, $p < 0.001$; Rhinolophoid activity - $H(3) = 45.251$, $p < 0.001$; edge/open hawking bats - $H(3) = 15.763$, $p = 0.001$; *Emballonura* spp. activity - $H(3) = 10.039$, $p = 0.018$) (Figure 4). Activity for all call types was lower in the plantations compared to forest (Supplementary Table 5.4; Figure 5.4). GAMLSS analyses revealed that bat activity levels were similar amongst the three forest types (old-growth forest, twice-logged forest, heavily logged forest; Figures 5.4-5.5). Overall bat activity was positively associated with *ForestExtent* but was not associated with any measure of habitat quality (Figure 5.5). The activity of edge/open hawking bats was also positively associated with *ForestExtent* as well as *Topography*. Both edge/open hawking bats and whispering bats were positively associated with *Topography*. Whispering bats were the only call type to show a significant difference between habitat types, and neither Rhinolophoid activity nor *Emballonura* spp. activity exhibited a response to any habitat covariate. We found differences between the responses of individual taxa and these did not reflect responses at the call type level (Figure 5.6).

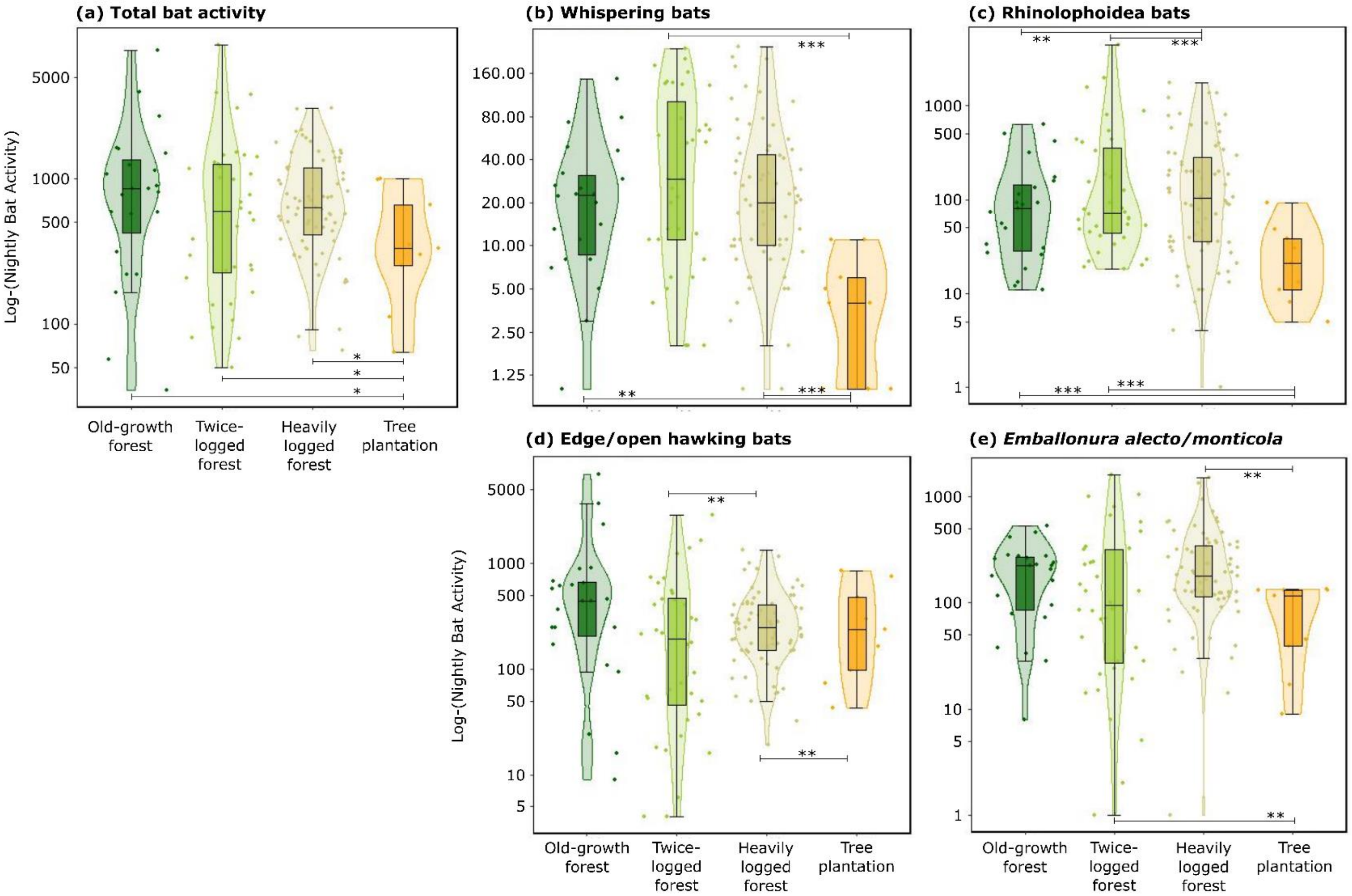


Figure 5.4 A comparison of overall nightly bat activity across each of the four habitat types.

Each point represents an individual recording location. (a) Total bat activity, (b) whispering bats activity, (b) Rhinolophoid bat activity, (c) edge/open bat activity, and (d) Emballonura alecto/monticola activity. Horizontal lines are median values, the boxes are between the 25th and 75th percentiles, and the whiskers represent the interquartile range. Statistical significance of the comparisons is according to the Dunn test results - * < 0.05, ** < 0.01, *** < 0.001.

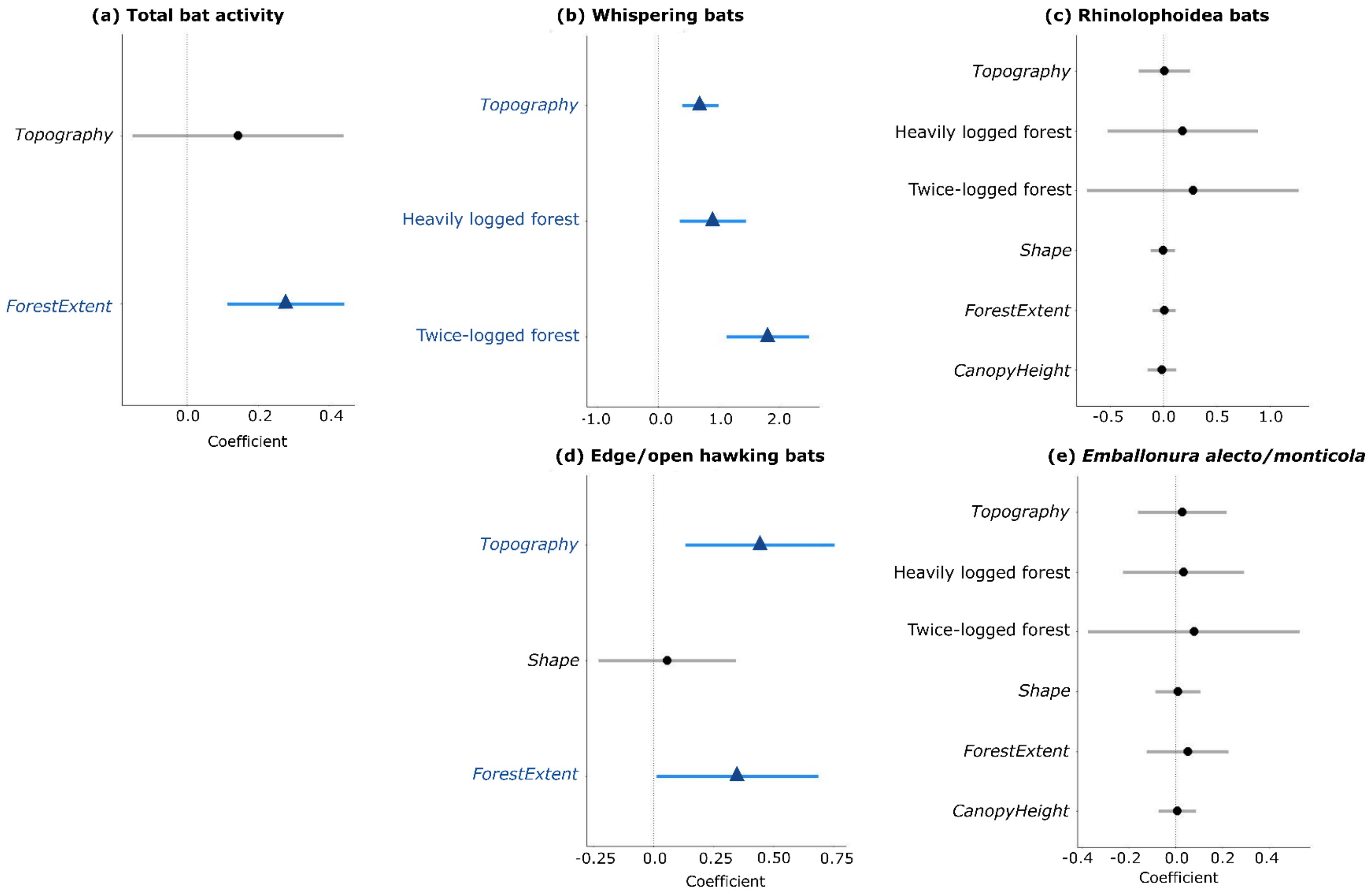


Figure 5.5 Caterpillar plots from the General Additive Models for Location, Shape, and Space (including 95% confidence interval)

(a) Total activity, (b) whispering bats activity, (b) Rhinolophoidea bat activity, (c) edge/open bat activity, and (d) Emballonura alecto/monticola activity. Significant associations are shown with a triangle point and highlighted in blue. Models are listed in Supplementary Table 5.2.

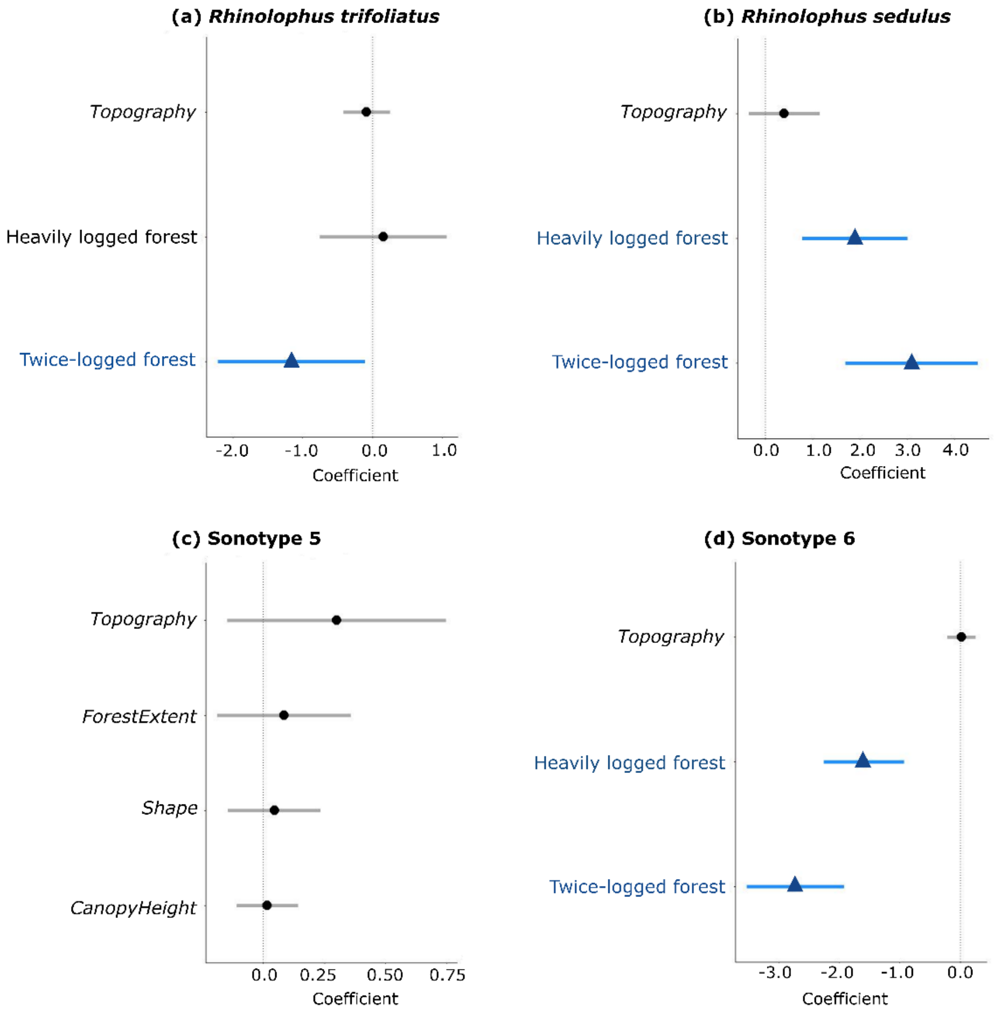


Figure 5.6 Caterpillar plots from the General Additive Models for Location, Shape, and Space (including 95% confidence interval) for the dominant sonotypes/species

(a) *Rhinolophus trifoliatus* activity, (b) *Rhinolophus sedulus* activity, (c) Sonotype 5 activity, and (d) Sonotype 6 activity. Significant associations are highlighted in blue. Models are listed in Supplementary Table 5.2.

5.6 Discussion

Using a novel approach, pairing a semi-automated classifier of bat calls with LiDAR-derived forest structural metrics, we found little evidence that logging disturbance negatively affected the diversity or composition of bat communities across a heavily modified landscape. Although we observed discernible differences in bat community composition between twice-logged forest and heavily logged forest, both these disturbed forest types supported similar bat communities to old-growth forest. Although bats persisted in highly disturbed logged forests, bat activity declined in tree plantations and several taxa were not recorded in this habitat. These findings support previous studies of other vertebrate groups that show logged forests provide better quality habitat for many forest species compared to plantations, such as oil palm and Acacia (Edwards, Magrach, *et al.* 2014; Gibson *et al.* 2011). Species most adversely affected by logging tend to have narrow niche breadth and are adapted to the stable conditions usually found in the forest interior. This includes species dependent on old-growth trees for feeding or nesting (including cavity-nesting birds and saproxylic insects), large-bodied/long-lived species with low fecundity, insectivorous/animalivorous species, or target species for poachers (Bicknell & Peres, 2010; Costantini *et al.*, 2016; Edwards *et al.*, 2014; Thorn *et al.*, 2018).

Logging pressure (e.g. as reflected by timber extraction rate) is an important determinant of how selective logging affects biodiversity (Burivalova, Şekercioğlu and Koh 2014; Nicolas J. Deere *et al.* 2020). Although high numbers of species are reported from twice-logged tropical forests (Putz *et al.* 2012), further logging cycles deteriorate habitat conditions to such an extent that forest-specialist species decline and taxa associated with forest edges or gaps proliferate to take their place (Cleary *et al.*, 2007; Edwards *et al.*, 2014). These changes in community composition have been observed for multiple taxonomic groups and collectively affect ecosystem functioning (Burivalova *et al.*, 2014; Cleary *et al.*, 2009; Edwards *et al.*, 2014). In contrast, our acoustic study of palaeotropical bats implies that logging pressure does not necessarily reduce the habitat value of logged forest for common bat taxa. We found no indication that deterioration of habitat quality (e.g., reductions in canopy height) negatively impacted acoustic diversity or overall bat activity. Unlike previous studies of riparian forest remnants in the landscape (Mullin *et al.* 2020), we found habitat extent was more important than habitat quality for determining bat activity. However, we utilised a more conservative measure of forest extent based on LiDAR data that represents high-quality forest and as such, excludes young regenerating forests. Therefore, our results imply that once a minimum threshold is reached, habitat extent becomes more pivotal in predicting overall bat activity.

Species adapted to foraging in more open spaces, such as sheath-tailed bats (*E. alecto/monticola*), were the most resilient to disturbance, as reflected by comparable levels of activity within plantations and old-growth forest. Edge/open hawking bats also maintained similar activity levels in plantations, but this was positively associated with the availability of good quality habitat, as determined by greater forest extent. Bats of all other call types declined substantially in tree plantations. We observed one of the demographic responses outlined by Presley *et al.* (2008), the decline of rare species in logged forests. For example, *H. ridleyi* (a forest-specialist species classified as Vulnerable by the IUCN, 2022) was absent from both the heavily logged forest and tree plantations, though persisted in twice-logged forest. However, we also observed positive demographic responses to logging. Several common taxa appeared to benefit from logging disturbance, in line with findings from Joann Christine, Fletcher and Abd Rahman (2013). Positive responses to logging have also been observed for terrestrial mammals in the same study system (Wearn *et al.* 2017). We, therefore, suggest an additional demographic response, namely that common species remain common and increase in abundance.

Overall levels of bat activity provided a poor representation of how each call type and taxon responded to habitat disturbance. Edge/open hawking bat activity was dominated by two sonotypes (sonotype 5 and 6) which together constitute 67% of activity for this call type. Overall, edge/open hawking bat activity was positively associated with increased habitat extent and topography ruggedness. However, it is unclear which sonotype drives the relationship at the call type level as neither sonotype exhibited a relationship between habitat extent or topography. Different Rhinolophoid species indicated different susceptibility to logging disturbance, as indicated by changes in their activity. This aligns with the known ecology of this group. Unlike edge/open hawking bats, the call type Rhinolophoidea includes forest specialists as well as species adapted to more disturbed areas, such as forest edge/gaps (Kingston 2013; Furey and Racey 2016). Therefore, it is to be expected that there would be more heterogeneity between the responses of species in this call type, compared to more ecologically similar species in the other call types. *R. sedulus* and *R. trifoliatus* collectively represented 89% of activity for this call type and exhibited contrasting responses to logging, consistent with those observed by Struebig *et al.* (2013). Repeated logging cycles reduce roosting opportunities for forest-roosting species, such as woodpeckers (*Picinae* spp.). As such, forest roosting species are considered to be more susceptible to logging disturbance than those that dwell elsewhere (Costantini, Edwards and Simons 2016; Struebig *et al.* 2013).

Logging appeared to positively influence the activity of forest-specialist whispering bats (e.g., woolly bats, *Kerivoula* spp.), though acoustic sampling remains problematic to draw robust conclusions for this group because the low-intensity, high-frequency calls typical of whispering bats remain difficult to detect and record (Kingston, 2013; Russo *et al.,* 2018). Capture studies using harp traps demonstrate whispering bats are highly sensitive to logging and forest fragmentation (Kingston, Francis, *et al.* 2003; Struebig *et al.* 2008; Struebig *et al.* 2013). This difficulty in detection suggests that for whispering bat species, acoustic monitoring is unlikely to fully capture their responses to habitat disturbance and degradation. However, for the wider palaeotropical bat species community, our study demonstrates acoustic monitoring and semi-automated classification of calls is an effective survey method.

5.7 Conclusions

Logging continues to be the most extensive disturbance affecting tropical forests worldwide. Using remote sensing technologies, we show that even heavily logged forest areas can support comparable levels of palaeotropical bat activity to old-growth forest. However, old-growth forest remains an important habitat for several rarer, forest-specialist species which do not persist in disturbed habitats. We demonstrate how acoustic monitoring methods can be used to incorporate aerial insectivores into biodiversity studies, and therefore combining acoustic data with live-trapping information could help better evaluate the more subtle, species-specific impacts of logging disturbance.

5.8 Acknowledgements

We thank the Sabah Biodiversity Council, Sabah Forest Department, Yayasan Sabah, and Benta Wawasan Sdn Bhd. for research permissions and site access, and the many research assistants at the SAFE project site that helped with the bat surveys. We also thank the Coomes Lab at the University of Cambridge for access to LiDAR information, and Anthony Turner and Felicia Lasmana who assisted with setting up acoustic equipment. The acoustic monitoring campaign was supported by a Leverhulme Trust Early Career Fellowship awarded to MJS as well as research grants from Bat Conservation International and the Mohamed bin Zayed Species Conservation Fund. NY was supported by the UK’s Natural Environmental Research Council (NERC) via a EnvEast DTP scholarship (NE/L002582/1). The LiDAR campaign and subsequent data processing was funded via NERC’s Human Modified Tropical Forests programme (NE/K016261/1; NE/K016377/1).

.

5.9 Supplementary information

Table S5.1 Site descriptions for each of the acoustic survey points in Sabah, Borneo

Block refers to the Stability of Altered Forest Ecosystem's (SAFE) site codes during 2011-2012. Supplementary information

Habitat type	Site Description	Block	Point	No. of nights	Total minutes	Latitude	Longitude	Start Date	End date
Old-growth forest	The OG2 block of control sites consists of old-growth forest which has undergone no logging	OG2	1	2	1480	4.746902	116.9683	2012-Jun-01	2012-Jun-03
			2	2	1480	4.751981	116.9663	2011-Apr-15	2011-Apr-17
			3	2	1480	4.751261	116.9662	2011-Apr-19	2011-Apr-21
			4	2	1480	4.748933	116.9661	2011-Apr-21	2011-Apr-23
	The OG3 block of control sites in Maliau Basin has	OG3	1	1	740	4.735662	116.9546	2011-May-24	2011-May-25

	undergone limited selective logging to increase access								
			2	2	1480	4.734725	116.9542	2011-Apr-26	2011-Apr-28
			3	2	1480	4.735012	116.9548	2011-Apr-28	2011-Apr-30
			4	2	1480	4.735926	116.9568	2011-Apr-30	2011-May-02
			5	1	740	4.735926	116.9568	2011-May-03	2011-May-04
Twice-logged forest	Twice logged forest to the Northeast of the SAFE experimental area. Forest has been recovering since the early 2000’s.	LF2	1	2	1480	4.77019	117.698	2012-May-05	2012-May-07
			2	2	1480	4.768901	117.6982	2012-May-07	2012-May-09

			3	2	1480	4.769667	117.6983	2012-May-09	2012-May-11
			4	2	1480	4.768477	117.7001	2012-May-11	2012-May-13
		LF3	1	2	1480	4.753733	117.691	2012-May-05	2012-May-07
			2	2	1480	4.756727	117.6903	2012-May-07	2012-May-09
			3	2	1480	4.758399	117.6913	2012-May-09	2012-May-11
			4	1	740	4.757936	117.6932	2012-May-11	2012-May-12
	Twice logged forest to the north of the SAFE experimental area and at the edge of the continuous forest. Forest has been	LFE	1	2	1480	4.724952	117.5969	2012-Jun-12	2012-Jun-14

	recovering since the early 2000’s.								
			2	2	1480	4.726703	117.5962	2012-Jun-14	2012-Jun-16
			3	4	2960	4.724116	117.597	2012-May-29	2012-Jun-02
			4	2	1480	4.724393	117.5971	2012-Jun-02	2012-Jun-04
Heavily logged forest	Fragment of forest in a matrix of repeatedly logged forest.	Frag. A	1	2	1480	4.709926	117.6463	2012-May-22	2012-May-24
			2	2	1480	4.710397	117.6538	2012-May-20	2012-May-22
			3	2	1480	4.708927	117.653	2011-May-24	2011-May-26
			4	3	2220	4.710989	117.6508	2011-May-19	2011-May-22

			5	1	740	4.709926	117.6463	2011-May-26	2011-May-27
Heavily logged forest	Fragment of forest in a matrix of repeatedly logged forest.	Frag. B	1	2	1480	4.72994	117.6197	2012-Apr-26	2012-Apr-28
			2	2	1480	4.729936	117.6197	2012-Apr-21	2012-Apr-23
			3	2	1480	4.73543	117.6182	2012-Jun-13	2012-Jun-15
			4	2	1480	4.730977	117.6185	2012-Apr-23	2012-Apr-25
		Frag. C	1	1.5	930	4.709915	117.6249	2011-May-17	2011-May-19
			2	4	2960	4.709754	117.623	2011-May-12	2011-May-17
			3	2	1480	4.708323	117.6241	2012-Apr-20	2012-Apr-22

	4	2	1480	4.708793	117.6238	2012-Apr-22	2012-Apr-24
Frag. D	1	2	1480	4.71509	117.584	2011-May-07	2011-May-09
Frag. E	2	2	1480	4.71169	117.5926	2011-May-09	2011-May-11
Frag. F	3	2	1480	4.71107	117.5875	2012-Apr-19	2012-Apr-21
	4	1	740	4.71234	117.5859	2011-May-11	2011-May-12
	1	2	1480	4.693655	117.5814	2012-Apr-26	2012-Apr-28
	2	2.5	1551	4.690838	117.5759	2012-Apr-23	2012-Apr-26
	3	2	1480	4.687996	117.5904	2012-Apr-19	2012-Apr-21

			4	2	1480	4.695058	117.5783	2012-Apr-21	2012-Apr-23
			1	2	1480	4.694425	117.5412	2012-May-21	2012-May-23
			2	2	1480	4.697215	117.5432	2012-May-24	2012-May-26
			3	2	1480	4.697581	117.5372	2012-May-16	2012-May-18
			4	2	1480	4.696135	117.5401	2012-May-27	2012-May-29
Tree plantation	Acacia plantations	Acacia	1	2	1480	4.66408	117.6519	2012-May-20	2012-May-22
			2	2	1480	NA	NA	2012-May-22	2012-May-24
			3	2	1480	NA	NA	2012-Jun-12	2012-Jun-14

	4	2	1480	NA	NA	2012-May-24	2012-May-26
	5	2	1480	NA	NA	2012-Jun-14	2012-Jun-16

Table S5.2 Candidate models used to characterise bat activity at the 100m scale. Shading represents top models (ΔAICc < 2).

Model	df	logLik‡	AICc§	ΔAICc¶	wt¥
Total Bat Activity ~					
ForestExtent + Topography	4	-870.36	1748.70	0.00	0.47
ForestExtent	3	-871.91	1749.80	1.10	0.27
Shape	3	-873.38	1752.80	4.03	0.06
Shape + Topography	4	-872.56	1753.10	4.40	0.05
HabitatType + Topography	5	-871.59	1753.20	4.46	0.05
CanopyHeight	3	-874.11	1754.20	5.49	0.03
HabitatType	4	-873.20	1754.40	5.67	0.03
CanopyHeight + Topography	4	-873.37	1754.70	6.02	0.02
Topography	3	-875.53	1757.10	8.34	0.01
Null Model	2	-876.84	1757.70	8.95	0.01
Whispering Bat Activity ~					
HabitatType + Topography	5	-508.75	1027.50	7.21	1
Treatment	4	-518.43	1044.90	24.55	0
Topography	3	-520.42	1046.80	26.54	0
CanopyHeight + Topography	4	-519.67	1047.30	27.03	0
Null model	2	-521.91	1047.80	27.51	0
ForestExtent + Topography	4	-519.91	1047.80	27.52	0
Shape + Topography	4	-520.39	1048.80	28.47	0
CanopyHeight	3	-521.41	1048.80	28.52	0
ForestExtent	3	-521.72	1049.40	29.13	0
Shape	3	-521.89	1049.80	29.48	0
Rhinolophoidea Bat Activity ~					
Null model	2	-714.39	1432.80	0.00	0.23
HabitatType	4	-712.57	1433.10	0.36	0.19
CanopyHeight	3	-714.16	1434.30	1.53	0.11
HabitatType + Topography	5	-712.26	1434.50	1.73	0.10
ForestExtent	3	-714.28	1434.60	1.77	0.10
Shape	3	-714.36	1434.70	1.93	0.09
Topography	3	-714.37	1434.70	1.95	0.09
CanopyHeight + Topography	4	-714.16	1436.30	3.52	0.04
ForestExtent + Topography	4	-714.26	1436.50	3.73	0.04
Shape + Topography	4	-714.34	1436.70	3.90	0.03
Edge/open Hawking Bat Activity ~					
ForestExtent + Topography	4	-757.62	1523.20	0.00	0.64
Shape + Topography	4	-759.37	1526.70	3.50	0.11
Topography	4	-759.74	1527.50	4.25	0.08
HabitatType + Topography	5	-758.79	1527.60	4.33	0.07
ForestExtent	3	-761.73	1529.50	6.22	0.03
HabitatType	4	-760.73	1529.50	6.23	0.03
CanopyHeight	3	-762.03	1530.10	6.82	0.02
Shape	3	-762.06	1530.10	6.88	0.02
Topography	3	-765.21	1536.40	13.17	0.00
Null model	2	-769.00	1542.00	18.76	0.00

***Emballonura alecto/monticola* Activity ~**

ForestExtent	3	-707.94	1421.90	0.00	0.24
Null model	2	-709.21	1422.40	0.54	0.18
ForestExtent + Topography	4	-707.78	1423.60	1.68	0.10
Shape	3	-708.87	1423.70	1.87	0.10
CanopyHeight	3	-708.98	1424.00	2.09	0.09
HabitatType	4	-708.05	1424.10	2.23	0.08
Topography	3	-709.14	1424.30	2.40	0.07
HabitatType + Topography	5	-707.21	1424.40	2.54	0.07
Shape + Topography	4	-708.81	1425.60	3.74	0.04
CanopyHeight + Topography	4	-708.93	1425.90	3.99	0.03

***Rhinolophus trifoliatus* Activity ~**

HabitatType	4	-569.04	1146.10	3.98	0.59
HabitatType + Topography	5	-568.58	1147.20	5.06	0.34
Null model	2	-574.31	1152.60	10.52	0.02
Shape	3	-574.18	1154.40	12.26	0.01
Topography	3	-574.20	1154.40	12.29	0.01
ForestExtent	3	-574.24	1154.50	12.37	0.01
CanopyHeight	3	-574.27	1154.50	12.44	0.01
CanopyHeight + Topography	4	-574.09	1156.20	14.07	0.00
Shape + Topography	4	-574.13	1156.30	14.15	0.00
ForestExtent + Topography	4	-574.15	1156.30	14.20	0.00

***Rhinolophus sedulus* Activity ~**

HabitatType + Topography	5	-527.53	1065.10	0.78	0.68
HabitatType	4	-529.31	1066.60	2.33	0.31
CanopyHeight	3	-535.46	1076.90	12.65	0.00
Null model	2	-536.66	1077.30	13.04	0.00
Shape	3	-536.12	1078.20	13.95	0.00
CanopyHeight + Topography	4	-535.45	1078.90	14.62	0.00
ForestExtent	3	-536.58	1079.20	14.88	0.00
Topography	3	-536.66	1079.30	15.04	0.00
Shape + Topography	4	-536.11	1080.20	15.94	0.00
ForestExtent + Topography	4	-536.58	1081.20	16.87	0.00

Sonotype 5 Activity ~

Shape + Topography	4	-628.49	1265.00	5.15	0.27
ForestExtent + Topography	4	-628.76	1265.50	5.69	0.21
Topography	3	-630.01	1266.00	6.18	0.16
CanopyHeight + Topography	4	-629.46	1266.90	7.10	0.10
Shape	3	-630.66	1267.30	7.50	0.08
ForestExtent	3	-631.21	1268.40	8.59	0.05
Null model	2	-632.31	1268.60	8.79	0.04
CanopyHeight	3	-631.48	1269.00	9.12	0.04
HabitatType + Topography	5	-629.66	1269.30	9.50	0.03
HabitatType	4	-630.98	1270.00	10.13	0.02

Sonotype 6 Activity ~

HabitatType	4	-629.11	1266.20	0.00	0.71
HabitatType + Topography	5	-628.99	1268.00	1.76	0.29
CanopyHeight + Topography	4	-641.32	1290.60	24.43	0.00

Topography	3	-643.11	1292.20	26.02	0.00
Shape + Topography	4	-642.39	1292.80	26.57	0.00
ForestExtent + Topography	4	-642.92	1293.80	27.63	0.00
CanopyHeight	3	-645.23	1296.50	30.25	0.00
Shape	3	-646.83	1299.70	33.45	0.00
Null model	2	-648.84	1301.70	35.46	0.00
ForestExtent	3	-648.06	1302.10	35.90	0.00

‡ The log-likelihood value, an alternative measure of model performance and denotes the plausibility of the model. § Akaike’s Information Criterion AICc. ¶ Delta AICc score, the difference between the AICc score of each model against the best performing model. ¥ Akaike weight.

Table S5.3 Biodiversity estimates of acoustic diversity (sonotype) richness summarised at for each habitat.

	t	**Method**	**Sample coverage est.**	**Acoustic diversity**
Habitat				
Old-growth forest	20	Interpolated	0.999	16.402
Twice-logged forest	20	Interpolated	0.999	17.217
Heavily logged forest	46	Interpolated	0.999	15.804
Tree plantation	10	Observed	0.999	16.000

t = reference sample size

Table S5.4 Results of the post-hoc Dunn test comparing bat activity between the four habitat types
Significance - * < 0.05, ** < 0.01, *** < 0.001.

Habitat	**Z**	***p*-value**	
Total activity			
Heavily logged forest - Old-growth forest	-0.526	0.599	
Heavily logged forest - Tree plantation	2.956	0.019	*
Old-growth forest - Tree plantation	2.946	0.010	*
Heavily logged forest - Twice-logged forest	0.766	0.533	
Old-growth forest - Twice-logged forest	1.060	0.434	
Tree plantation - Twice-logged forest	-2.298	0.043	*
Whispering bats			
Heavily logged forest - Old-growth forest	-1.767	0.116	
Heavily logged forest - Tree plantation	3.818	< 0.001	***
Old-growth forest - Tree plantation	3.548	0.001	**
Heavily logged forest - Twice-logged forest	-1.526	0.152	
Old-growth forest - Twice-logged forest	1.139	0.255	
Tree plantation - Twice-logged forest	-4.659	< 0.001	***
Rhinolophoidea bats			
Heavily logged forest - Old-growth forest	-3.280	0.002	**
Heavily logged forest - Tree plantation	1.909	0.068	
Old-growth forest - Tree plantation	3.987	< 0.001	***
Heavily logged forest - Twice-logged forest	-5.035	< 0.001	***
Old-growth forest - Twice-logged forest	1.254	0.210	
Tree plantation - Twice-logged forest	-5.281	< 0.001	***
Edge/open hawking bats			
Heavily logged forest - Old-growth forest	0.563	0.574	
Heavily logged forest - Tree plantation	3.137	0.005	**
Old-growth forest - Tree plantation	1.056	0.582	
Heavily logged forest - Twice-logged forest	3.219	0.008	**
Old-growth forest - Twice-logged forest	0.706	0.576	
Tree plantation - Twice-logged forest	-0.739	0.690	
Emballonura alecto/monticola			
Heavily logged forest - Old-growth forest	-0.532	0.595	
Heavily logged forest - Tree plantation	3.003	0.016	**
Old-growth forest - Tree plantation	1.999	0.137	
Heavily logged forest - Twice-logged forest	1.222	0.332	
Old-growth forest - Twice-logged forest	1.000	0.381	
Tree plantation - Twice-logged forest	-1.992	0.093	**

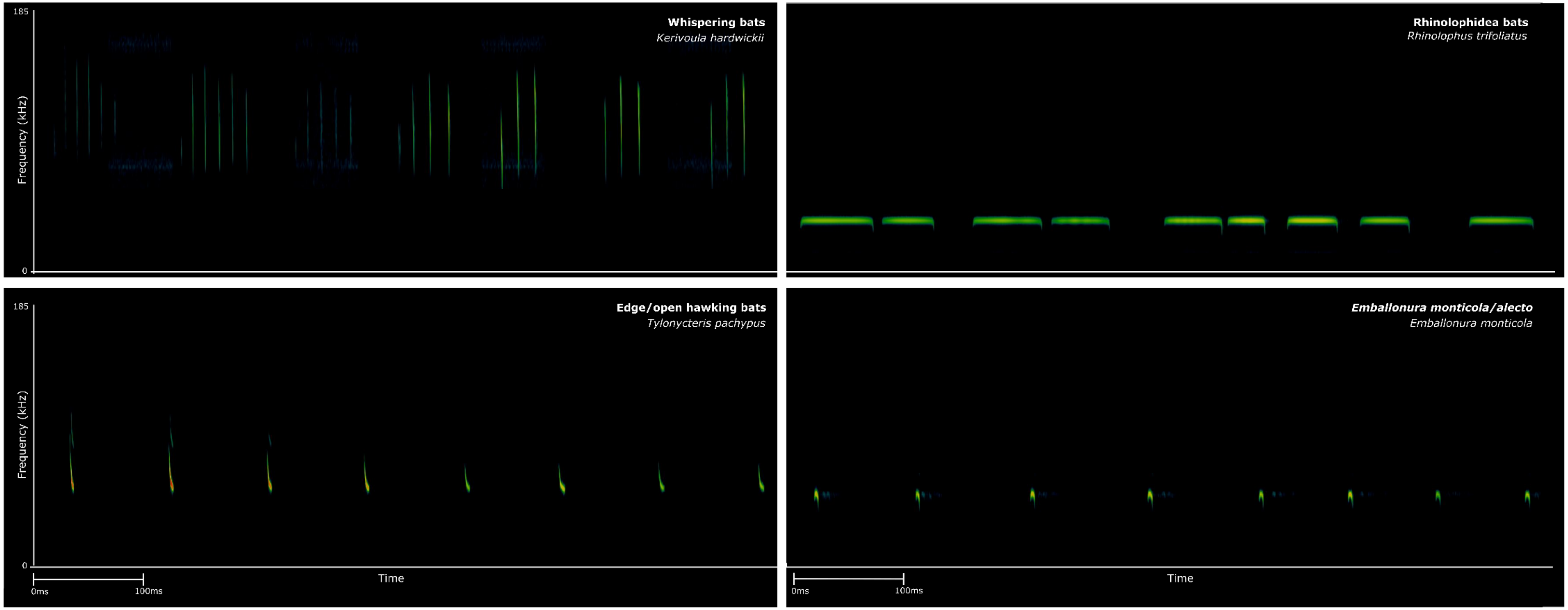


Figure S5.1 Example sonograms for each of the four call types used in the study

Call amplitude has not been standardised between species and therefore this image is for illustrative purposes

Chapter 6 Discussion

In this thesis, I demonstrate the utility of acoustic monitoring for characterising bat biodiversity in a region of the world threatened by accelerating land-use change. I introduce a new algorithm to expedite the processing of large acoustic datasets for bats in Southeast Asia. The application of this new algorithm was subsequently demonstrated by applying this approach to assess how bat species are responding to logging disturbance. Novel technological and statistical approaches, including passive acoustic monitoring, machine learning, and LiDAR, were used throughout this thesis to review the effectiveness of different management strategies for protecting biodiversity, as well as to improve the capacity for ongoing monitoring in the future. In this way, this thesis supports species monitoring efforts in Southeast Asia and the tropics more broadly.

6.1 Contributions to the research field

6.1.1.1 Impacts of land-use change

Studies investigating how tropical bats are impacted by land-use change are increasing (Meyer, Struebig and Willig 2016; Frick, Kingston and Flanders 2020). However, to date, almost all land-cover change studies involving palaeotropical bats have relied on live trapping techniques which miss a substantial proportion of bat communities (Kingston 2013; Meyer 2015). **Chapters 2, 3,** and **5** demonstrate how acoustic studies can be used to assess bat responses to land-use change and how they can complement knowledge from live-trapping studies. To our knowledge, this thesis represents the largest acoustic monitoring study of bats in Southeast Asia. In total, there were 1,115 hours of field recordings (**Chapters 2, 3,** and **5**) across a spectrum of land-use disturbance types. In this section, I highlight some of the empirical contributions made by this thesis to the research field.

This research demonstrated how habitat disturbance influences bat community composition. Although disturbed forests can still support comparable bat activity to old-growth forests, this activity does not necessarily represent the same taxa (Figure 6.1). Echolocation call type is one of several interconnected morphological traits that determine a species' ecological niche and therefore its sensitivity to disturbance (Farneda *et al.* 2015; Núñez *et al.* 2019; Norberg, Rayner and Lighthill 1987). Call structures associated with the forest interior (short, high frequency, broadband calls) decrease in response to increased disturbance, in exchange for an increase of calls associated with edge or open foragers. This indicates that the activity of forest specialist species declines, such as high-frequency Hipposiderids (e.g., *Hipposideros ridleyi*), and is replaced by more generalist species. This shift in community composition is consistent with results from previous studies investigating how disturbance impacts bats (Rocha *et al.* 2017; Brändel *et al.* 2020; Estrada-Villegas, Meyer and Kalko 2010; Struebig *et al.* 2008) and several other taxonomic groups (Stork *et al.* 2017; Woodcock *et al.* 2011; Cleary *et al.* 2007; Ng *et al.* 2021; Cleary *et al.* 2009). However, this study found that bat communities can largely persist under low-intensity disturbance (buffer zones in **Chapter 2**, logged forest in **Chapter 5**), but specialist species are lost as disturbance intensity increases (transition zones in **Chapter 2**, plantations in **Chapter 5**). A global meta-analysis outlined 38 out of 45 studies (84%) that found that agricultural intensification had a negative impact on bat abundance/activity or species richness (Williams-Guillén *et al.* 2016). Similarly, the abundance of terrestrial mammals in Borneo has been shown to increase in logged forests compared to old-growth forests, but substantially decline in oil palm plantations (Wearn *et al.* 2017). Low-intensity agroforestry has been shown to support other taxonomic groups, including amphibians and reptiles (e.g., Wanger *et al.* 2010; Indonesia), and bird and ant diversity (e.g., Philpott *et al.* 2008; Latin America) in natural-shade coffee plantations. Burivalova, Şekercioğlu and Koh (2014) also showed that most taxonomic groups are resilient to low-intensity selective logging (< 10 m^3 ha^{-1}). However, the consensus between all of these studies is that as disturbance intensity increases, diversity declines.

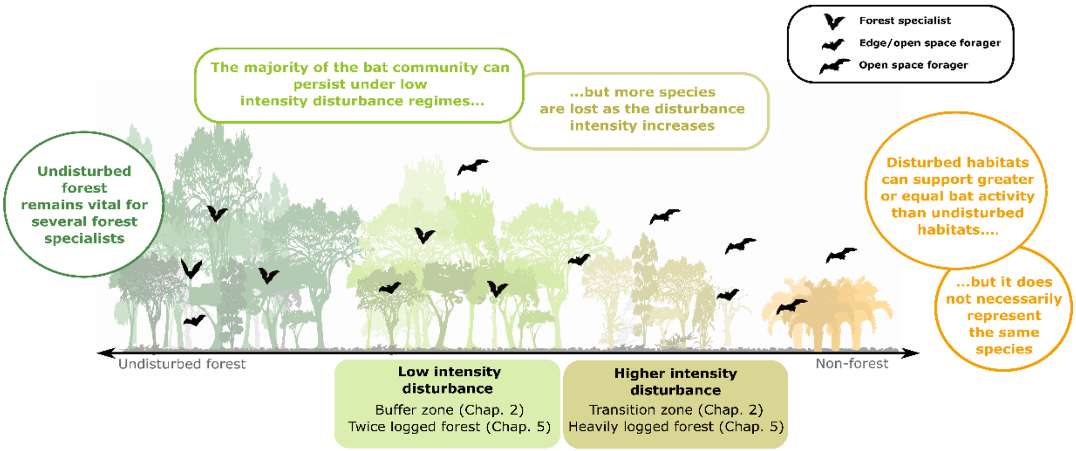


Figure 6.1 Visual summary of the main thesis findings**.**

Summary of how bats are affected by land-use disturbance in Southeast Asia, as identified in this thesis.

I did not examine the role of source-sink dynamics or periodic movement between disturbed habitats and neighbouring old-growth forest. Maintaining areas of undisturbed forest is essential for the long-term persistence of many forest species in disturbed areas, including within low-intensity agriculture (birds and invertebrates; Tscharntke *et al.* 2008; Gilroy *et al.* 2014; Sekercioglu *et al.* 2007). **Chapter 3** indicated evidence of movement between habitats as the foraging activity of constant frequency species was significantly lower in oil palm than in other habitat types. Therefore, it is likely these species can only persist due to foraging opportunities in neighbouring forests. Very few studies have examined source-sink phenomena in the tropics, particularly in Southeast Asia and Africa (Gilroy and Edwards 2017). Hence, more research is needed to establish how landscape configuration impacts species in human-modified tropical landscapes. Nonetheless, it is well recognised that undisturbed forest remains essential for highly sensitive, forest specialists (Barlow *et al.* 2010; Gibson *et al.* 2011; Chazdon *et al.* 2009; Edwards *et al.* 2017; Gilroy *et al.* 2014). Preserving forest remnants and maintaining connectivity between patches is therefore vital to conserving tropical biodiversity at the landscape level.

LiDAR-derived vegetation metrics are recognised as good indicators of ecosystem structure and habitat quality (Simonson, Allen and Coomes 2014). Despite this, there is a lack of ecological studies in Asia, Africa, and South America that incorporate these measures (Acebes, Lillo and Jaime-González 2021). In this thesis, I used LiDAR-derived metrics of forest structure to empirically assess the influence of habitat quantity and habitat quality on echolocating bat communities (**Chapters 3** and **5**). This included deriving high-resolution measures of forest extent (Riparian Width in **Chapter 3**; Above Ground Biomass in **Chapter 5**) and forest quality (Topography and Canopy Height in **Chapter 3**; Median Biomass and SD Biomass in **Chapter 5**). The findings further corroborate the importance of habitat quality in determining species distribution and persistence in fragmented landscapes (Laurance *et al.* 2018; Haddad *et al.* 2015). This research demonstrates the importance of structural complexity for certain bats (e.g., *Emballonura* spp. and frequency modulated quasi-constant frequency (FMqCF) bats), thus highlighting how preserving forest integrity is essential for conserving bat assemblages as a whole, not just forest specialists. Similar studies have been conducted for birds, large mammals, anurans, invertebrates, and freshwater fish in the same fragmentation system (Mitchell *et al.* 2018; Deere *et al.* 2022; Wearn *et al.* 2017). For example, Deere *et al.* (2022) showed that forest extent had a consistently positive relationship with species richness or abundance for each of these taxonomic groups. In contrast, the effect of habitat quality is more variable between taxonomic groups. For mammals, habitat degradation has been shown to reduce species richness in riparian forests in Borneo and Brazil (Deere *et al.* 2022; Zimbres, Peres and Machado 2017). In other fragmented landscapes, habitat quality has a greater impact than habitat extent for determining habitat use by mammal and reptile species in Brazil (peccary - Regolin *et al.* 2021; lizards - Silva *et al.* 2022). However, the relative importance of extent and quality varies between guilds (Ye, Skidmore and Wang 2013; Ewers and Didham 2006; Mattos, Zimbres and Marinho-Filho 2021). Therefore, the results presented in this thesis support previous conclusions that both habitat quality and extent need to be considered in conservation planning. Combining these cross-taxa studies provides a more comprehensive foundation for informing conservation policies.

6.1.1.2 Survey design and tool development

The different survey approaches used in the thesis generated different quantities of acoustic data. In **Chapters 2** and **3**, data were collected using walked transects (where a surveyor operates the detector in the field). However, static detectors require less intensive survey effort and can be deployed over longer timeframes. As such, > 88% of the field recording hours across this thesis were collected from static detectors (**Chapter 5**). The main constraint of large acoustic datasets is the time taken to manually identify taxa within them. The semi-automated approach was over seven times faster than manual identification alone: 6 weeks to process 60 hours of acoustic recordings (Mullin, *pers. comm.*; 1 week = 10 hours; **Chapter 3**), compared to 13 weeks to process 987 hours (1 week = 75.9 hours; **Chapter 5**). The Borneo Bat Classifier introduced in **Chapter 4** reduces the time demands of processing acoustic data, thereby making acoustic monitoring more accessible to researchers. The fastest data to process was for **Chapter 2** (68 hours of recording) as these recordings were not identified as species/sonotypes. This study used acoustic signatures to examine patterns of response. Although it was not possible to investigate taxon-specific responses using this approach, it still demonstrated the shift from short calls, high-frequency calls in forest habitats to longer calls with higher curvature (FMqCF) calls in more open spaces. Therefore, it can still be possible to monitor the impacts of land-use change on bat communities even when it is not possible to assign calls to sonotype/species.

It is important to acknowledge that acoustic monitoring cannot be used to monitor all bat responses to land-use change. Across the field recordings, calls from interior, forest specialist species Kerivoulinae, Murininae, Nycteridae, and Megadermatidae were consistently low despite several species being relatively common (e.g., *Murina suilla*; Struebig *et al.* 2008; Azhar and Rossiter 2018). In **Chapter 3**, this group only represented 6.3% of activity (FMb and MHFM) and 5.2% in **Chapter 5** (whispering bats). However, this group can represent over 30% of captures in harp trapping studies (Kingston *et al.* 1999; Francis 1990). It is well recognised that low-intensity, high-frequency calls are difficult to detect using acoustic monitoring techniques (Russo, Ancillotto and Jones 2018; Russo and Voigt 2016; Walters *et* *al.* 2013). In Swaziland, Monadjem *et al.* (2017) demonstrated clear calls could not be reliably recorded for *Kerivoula* and *Nycteris* species even within 0.1m. Therefore, acoustic monitoring using current recording technologies should not be used to monitor these species, and live capture techniques should be used instead, at least for the time being. Live capture techniques are also important to facilitate tagging or marking individuals or collecting tissue samples to investigate population dynamics, physiology, or confirmed species' records (Gannon, Sikes, and The Animal Care and Use Committee of the American Society of Mammalogists 2007). **Chapter 2** reported 8 new species records for Crocker Range Biosphere Reserve using mist nets and harp traps. This would not have been possible using acoustic monitoring as one species does not echolocate (*Cynopterus minutus*) and at least four cannot be acoustically differentiated into species. Therefore, combining acoustic monitoring with live capture methods provides the most comprehensive way to monitor bats across the whole community (as shown for Brazil; Silva and Bernard 2017; Appel *et al.* 2021).

6.1.1.3 Borneo as a case study

The northern part of Borneo provides an ideal case study for the effects of land-cover change on bats as the region supports high species diversity (that is relatively well understood) and has undergone rapid land-use change over the last fifty years, but still retains a high proportion of old-growth forest compared to other parts of Southeast Asia (Hansen *et al.* 2013; Struebig *et al.* 2010; Meyer, Struebig and Willig 2016; Kingston 2013). However, Southeast Asia is politically, culturally, and economically a very diverse region (Dunning 2022; Samek *et al.* 2004; Dayley 2019). Differences in topography and climate, as well as anthropogenic differences, such as population size, economic growth, and political systems, have all impacted the relative rates of land-use change across Southeast Asia (Imai *et al.* 2018; Dunning 2022; Samek *et al.* 2004). Human-wildlife interactions also vary widely. Asia-Pacific cultures have many positive cultural associations with bats. However, hunting for consumption remains common in many communities (Low *et al.* 2021). Southeast Asia also has extensive karst landscapes, around 10% of the land area, which support a high diversity of cave-roosting species (Day and Urich 2000; Suyanto and Struebig 2007; Furey, Mackie and Racey 2010; Sedlock *et al.* 2014). Cave disturbance (e.g., persecution/hunting, mining/mineral extraction, and tourism) now affects over a third of threatened bat species globally (Frick, Kingston and Flanders 2020). The rates of disturbance and conservation regulations for these landscapes vary substantially across Southeast Asian countries (Day and Urich 2000). All of these threats are interrelated and so the threats to bats - and biodiversity more broadly - are not homogenous across the region. Ongoing multi-national research, by networks such as the

Southeast Asian Bat Research Unit (SEABCRU), will provide future insights into how best to protect species considering these differences.

6.2 Implications for policy and practice

The recent development of low-cost, open-source hardware has massively improved the affordability of acoustic monitoring initiatives (Hill *et al.* 2018). The development of automatic or semi-automatic processing tools is the next frontier for the large-scale application of acoustic monitoring. In Europe, North America, and Australia these tools are already widely available to researchers and hobbyists. These include classifiers that work on the user's smartphone, such as the free Merlin Bird ID app by Cornell Lab of Ornithology (TheCornellLab 2021) and the FrogID app by the Australian Museum Research Institute (Rowley and Callaghan 2020). However, in the tropics, the development of processing tools has been impeded due to a lack of available species call databases (Gibb *et al.* 2019). The framework presented in this thesis demonstrates how acoustic classifiers can be built in regions where training databases or taxonomic knowledge may be limited, thereby making acoustic monitoring more accessible in palaeotropical regions. It should be noted that machine learning in conservation is still in its infancy. There are concerns as to how well machine learning accuracy metrics accurately reflect an algorithm's ability to predict situations in practice (Wearn, Freeman and Jacoby 2019) and we are still learning more about the calls of species considered to be well-studied (as demonstrated by Montauban *et al.* 2021). Therefore, these tools need to be continually assessed and revised. It is also pivotal that the limitations of such techniques are made transparent and widely acknowledged to ensure they do not lead to unintended misidentification, and subsequently, negative consequences for wildlife. Nevertheless, machine learning has the potential to be hugely beneficial for monitoring and conservation research.

Southeast Asia is a global epicentre for threatened mammal species (Frick, Kingston and Flanders 2020; Schipper *et al.* 2008). Bats use the landscape differently from both non-volant mammals and birds. For example, they have a higher vagility than volant mammals and many bird species. Thus bat species are often less impacted by fragment isolation and are better able to exploit areas of high food availability (Pichler *et al.* 2022; Helbig-Bonitz *et al.* 2015). Therefore, it is important to include them alongside studies for other taxonomic groups when designing management or intervention options. A key finding in this thesis is the importance of protecting key landscape features (i.e. riparian reserves) for promoting bat diversity, especially where they preserve taller, more structurally complex forest. As such, in **Chapter**

3, I stated that conservation set-asides should be at least 40m wide on each side of the river, as recommended for birds (Mitchell *et al.* 2018). However, new research indicates that edge effects can reduce canopy heights up to 92m from the forest edge (Ordway and Asner 2020) and reserve widths of 40-100m may be necessary to conserve biodiversity across taxa (Deere *et al.* 2022). Therefore, it is likely reserves greater than 40m are needed to ensure the long-term integrity of forest and their communities in riparian reserves. Our results also support previous findings that conclude areas of high-quality forest are important for conserving highly specialised, forest species within human-modified landscapes (Barlow *et al.* 2010; Gibson *et al.* 2011; Chazdon *et al.* 2009; Edwards *et al.* 2017; Gilroy *et al.* 2014). Hence the need to establish heterogeneous, ecological landscapes. In this way, conservationists can help design landscapes that conserve tropical biodiversity while still supporting sustainable resource use.

Establishing effective conservation practices is time-critical. There are many new and ongoing threats to wildlife in Southeast Asia. For example, the global demand and price of palm oil have surged since 2019 (IMARC Group 2022). Promisingly, Gaveau *et al.* (2022) showed that deforestation for these plantations in Malaysia and Indonesia had reduced between 2012 and 2017 in response to the reduced market value and new deforestation legislation. However, rising prices and the relaxation of several forest regulations during the COVID-19 pandemic have cast concern that there will be a greater expansion of new oil palm plantations (Shigetomi, Ishimura and Yamamoto 2020; Gaveau *et al.* 2022). In addition, there are grave concerns about the potential environmental damage of the planned move of Indonesia's capital from Jakarta to East Kalimantan, Borneo (Teo *et al.* 2020; Normile 2022). The new city, named Nusantara, aims to be "the greenest, smartest and sustainable capital city in the world" and to learn from previous relocations of capital cities and established green cities (de Vries 2021). However, it is yet to be seen how Indonesia is incorporating environmental and sustainability assessments into the city planning. Both Indonesia and Malaysia have recently restated their commitment to collaborate with other Southeast Asian countries to protect biodiversity in this megadiverse region (Association of Southeast Asian Nations 2021). Therefore, research such as presented in this thesis can be used to inform strategies to meet conservation and sustainable development goals, such as agreed in the Post-2020 Global Biodiversity Framework of the Convention on Biological Diversity and the 2030 Agenda for Sustainable Development (United Nations 2015; UN Convention on Biological Diversity 2021). Southeast Asia's future commitment to environmental best practices will have global consequences for species conservation and climate regulation.

6.3 Future research

6.3.1.1 Activity versus abundance

Acoustic monitoring allows us to monitor many species where it may not otherwise be possible. Nevertheless, a challenge of this approach is that activity is often used as a surrogate for relative abundance but it does not provide a true measure of the number of individuals (Hayes 2000). In addition, it can be difficult to contextualise different rates of activity between different sites (e.g., what constitutes an area of high activity). Ecobat (Lintott *et al.* 2018) provides the first web-based tool to standardise assessments of activity rates across the UK There is still a lot more development needed before this initiative can be expanded internationally. Even so, acoustic activity data can inform conservation because it provides a relative activity metric, as well as confirmation of species occurrence (e.g., it has been used to document the occurrence of the near-threatened Yellow-bellied Glider *Petaurus australis*; Whisson *et al.* 2021). It is therefore an important future avenue of research that acoustic activity measures better support population metrics, and that these are used to standardise the interpretation of high vs low activity (and everything in between). Needless to say, this will not be possible where reference calls are not available or where species calls cannot be differentiated.

6.3.1.2 Other landscape factors

One aspect of conservation areas that was not consistently investigated during this thesis was the importance of elevation and topography. Studies investigating the relationship between elevation and bat diversity are limited in the tropics. However, research suggests bat diversity and activity decrease at higher elevations (Arias-Aguilar *et al.* 2020; Curran *et al.* 2012; Cisneros *et al.* 2014). Lok *et al.* (2021) demonstrated a negative relationship between bat species diversity and elevation at Crocker Range Biosphere Reserve, the same survey area as in **Chapter 2**. In **Chapter 2,** surveys were conducted in hill dipterocarp forest (300-800m) and upper hill dipterocarp forest (800-1,300m; Saw 2010). Lok *et al.* (2021) also included lower montane (1,300-1,800m) and upper montane (> 1,800m; Saw 2010). Neither study included lowland forest. Therefore, it is still not possible to fully discern whether elevation affected the diversity or activity of bats observed between habitat types. By comparison, **Chapter 3** found that more elevated sites at the Stability of Altered Forest Ecosystems Project were positively associated with bat activity, particularly for *H. galeritus* and *H. cervinus*. Landscape topography is a key factor in determining whether a site is converted for plantations and the impact of conservation set-asides (Bicknell *et al.* 2021). However, rather than being a natural pattern of distribution, these sites are less accessible and thus are less disturbed.

Therefore, to ensure the benefits of conservation initiatives are not over or under-inflated, assessments need to adequately assess how accessible a site was to disturbance before the implementation of conservation actions.

6.3.1.3 Technological advancements

There have been several significant developments in bioacoustic monitoring in Southeast Asia since commencing this thesis in 2018. Most notably has been the recent launch of ChiroVox - the largest open-access bat call repository (Görföl *et al.*, 2022). At the time of writing (March 2022), ChiroVox hosts 3,902 recordings representing 223 species, including nearly half of Southeast Asia's echolocating bats (*ChiroVox* 2022; Görföl *et al.* 2022). Unlike previous databases, such as Echobank (Collen 2012), these calls are available to the public as an open-access or consensual-access resource and each recording is assigned a unique identifier that can be cited across publications. This provides a global resource for those working with bat bioacoustics and improves the repeatability of acoustic analyses. Community sharing of acoustic data such as this will ensure it is much easier to develop acoustic classifiers for new regions in the future.

There has also been a change in the way we apply acoustic classifiers to bat calls. Instead of classifying calls to species (e.g., Parsons and Jones, 2000), there are new tools that focus on classifying calls into representative call groups. Recently, Roemer, Julien and Bas (2021) used random forests to develop the first global bat classifier which assigns all echolocating bat species into one of eight universal call types. As different sonotypes can be separated by their acoustic strategy and spatial niche (Denzinger and Schnitzler 2013), this classifier aims to assess the state of bat communities anywhere in the world, including regions without local call libraries (Roemer, Julien and Bas 2021). The classifier presented in **Chapter 4** provides similar benefits (e.g., includes species with no reference calls) but is more regionally specific. For example, it combines classification to call type with more refined species identifications where possible. One key aspect of the work by Roemer, Julien and Bas (2021) is that the authors provide nomenclature that can be used to describe call types universally. Throughout the literature and across this thesis (Table 1.1), the nomenclature used for describing different call groups or guilds varies substantially. This can make cross-study comparisons problematic (as different call types/species are grouped differently) and makes it difficult for non-technical audiences to process the information presented. The eight sonotypes presented by Roemer, Julien and Bas (2021) provide a solution to the former, but they still do not provide easy, recognisable nomenclature for non-technical audiences. Establishing universal names for these sonotypes that are ecologically or taxonomically meaningful would be hugely beneficial for science communication concerning bat conservation.

6.4 Conclusion

Land-use change continues to threaten global biodiversity and the global climate. To ensure successful conservation initiatives, species data must not only be available but also be incorporated into conservation decision-making. Such action is needed urgently to curtail the sixth mass extinction. New technological methods, such as passive acoustic monitoring and machine learning, can complement existing surveying methods and facilitate more extensive and rapid species monitoring. Further sharing of resources, such as open-access call databases, will enable the future development of these approaches. Studies so far demonstrate how different human-modified habitats can be valuable to conservation efforts and offer ways to mitigate species losses in habitats that would be otherwise unable to support diversity. Only by addressing the underlying mechanisms of anthropogenic disturbance can we ensure the long-term protection of remaining biodiversity in a human-modified world.

References

Acebes, P., Lillo, P. and Jaime-González, C. (2021). Disentangling LiDAR Contribution in Modelling Species–Habitat Structure Relationships in Terrestrial Ecosystems Worldwide. A Systematic Review and Future Directions. *Remote Sensing* **13**:3447.

Acharya, P.R. *et al.* (2015). Feeding behaviour of the dawn bat (Eonycteris spelaea) promotes cross pollination of economically important plants in Southeast Asia. *Journal of Pollination Ecology* **15**:44–50.

Akutsu, K., Khen, C.V. and Toda, M.J. (2007). Assessment of higher insect taxa as bioindicators for different logging-disturbance regimes in lowland tropical rain forest in Sabah, Malaysia. *Ecol. Res.* **22**:542–550.

Aleixandre-Benavent, R. *et al.* (2018). Trends in global research in deforestation. A bibliometric analysis. *Land Use Policy* **72**:293–302.

Altringham, J. (1996). *Bats Biology & Behaviour*. Oxford, UK: Oxford University Press.

Anthony, E.L.P. (1988). Age Determination in Bats. In: Kunz, T. H. ed. *Ecological and Behavioural Methods for the Study of Bats*. Washington, DC, USA: Smithsonian Institute Press, pp. 47–58.

Appel, G. *et al.* (2021). Use of complementary methods to sample bats in the Amazon. *Acta Chiropterologica* **23**:499–511.

Arias-Aguilar, A. *et al.* (2020). Diversity and activity patterns of aerial insectivorous bats along an altitudinal gradient in a tropical forest in Costa Rica. *Hystrix, the Italian Journal of Mammalogy* **31**:58–63.

Arlettaz, R., Jones, G. and Racey, P.A. (2001). Effect of acoustic clutter on prey detection by bats. *Nature* **414**:742–745.

Arneth, A. *et al.* (2019). Framing and Context. In: Shukla, P. R. et al. eds. *Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*. IPCC.

Asner, G.P. *et al.* (2009). A Contemporary Assessment of Change in Humid Tropical Forests. *Conservation Biology* **23**:1386–1395.

Association of Southeast Asian Nations (2021). *ASEAN JOINT STATEMENT TO THE FIFTEENTH MEETING OF THE CONFERENCE OF THE PARTIES TO THE CONVENTION ON BIOLOGICAL DIVERSITY (CBD COP 15)*. [Online]. 15th Meeting of the Conference of the Parties to the Convention on Biological Diversity. Available at: https://asean.org/asean-joint-statement-to-the-fifteenth-meeting-of-the-conference-of-the-parties-to-the-convention-on-biological-diversity-cbd-cop-15/.

Azam, C. *et al.* (2016). Disentangling the relative effect of light pollution, impervious surfaces and intensive agriculture on bat activity with a national-scale monitoring program. *Landscape Ecology* **31**:2471–2483.

Azhar, B. *et al.* (2015). Effects of in situ habitat quality and landscape characteristics in the oil palm agricultural matrix on tropical understory birds, fruit bats and butterflies. *Biodiversity and Conservation* **24**:3125–3144.

Azhar, M.I. and Rossiter, S.J. (2018). IUCN Red List of Threatened Species: Murina suilla. *IUCN Red List of Threatened Species* [Online]. Available at: https://www.iucnredlist.org/en [Accessed: 15 March 2022].

Balmford, A. (2021). Concentrating vs. spreading our footprint: how to meet humanity's needs at least cost to nature. *Journal of Zoology* **315**:79–109.

Barlow, J. *et al.* (2010). Measuring the Conservation Value of Tropical Primary Forests: The Effect of Occasional Species on Estimates of Biodiversity Uniqueness. *PLOS ONE* **5**:e9609.

Barona, E. *et al.* (2010). The role of pasture and soybean in deforestation of the Brazilian Amazon. *Environmental Research Letters* **5**:024002.

Bas, Y., Bas, D. and Julien, J.-F. (2017). Tadarida: A Toolbox for Animal Detection on Acoustic Recordings. *Journal of Open Research Software* **5**:6.

Bates, D. *et al.* (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**:1–48.

Bianco, M.J. *et al.* (2019). Machine learning in acoustics: Theory and applications. *The Journal of the Acoustical Society of America* **146**:3590–3628.

Bicknell, J. *et al.* (2021). Enhancing the ecological value of tropical agriculture through set-aside. *Nature Portfolio. [Preprint]*. [Online]. Available at: https://doi.org/10.21203/rs.3.rs-272185/v1 [Accessed: 16 February 2022].

Bicknell, J. and Peres, C.A. (2010). Vertebrate population responses to reduced-impact logging in a neotropical forest. *Forest Ecology and Management* **259**:2267–2275.

Bicknell, J.E. *et al.* (2014). Improved timber harvest techniques maintain biodiversity in tropical forests. *Current Biology* **24**:1119–1120.

Bicknell, J.E., Struebig, M.J. and Davies, Z.G. (2015). Reconciling timber extraction with biodiversity conservation in tropical forests using reduced-impact logging. *Journal of Applied Ecology* **52**:379–388.

Blaser, J. *et al.* (2011). *Status of Tropical Forest Management 2011*. [Online]. ITTO Technical Series. Yokohama, Japan. Available at: https://www.itto.int/direct/topics/topics_pdf_download/topics_id=2660&no=0&disp=inline.

Boyles, J.G. and Storm, J.J. (2007). The perils of picky eating: dietary breadth is related to extinction risk in insectivorous bats. *PLoS One* **2**:e672.

Bradshaw, C.J.A., Sodhi, N.S. and Brook, B.W. (2009). Tropical turmoil: a biodiversity tragedy in progress. *Front. Ecol. Environ.* **7**:79–87.

Brändel, S.D. *et al.* (2020). Consequences of fragmentation for Neotropical bats: The importance of the matrix. *Biological Conservation* **252**:108792.

Breiman, L. (2001). Random Forests. *Machine Learning* **45**:5–32.

Broken-Brow, J. (2018). Anabat Insight Analysis Software. [Online]. Available at: https://www.titley-scientific.com/uk/downloads/dl/file/id/51/product/0/anabat_insight_user_manual_v1_7.pdf.

Broken-Brow, J. and Thompson, D. (2017). Anabat Walkabout. [Online]. Available at: https://www.titley-scientific.com/uk/downloads/dl/file/id/18/product/0/anabat_walkabout_user_manual_v1_4.pdf.

Brookfield, H. and Byron, Y. (1990). Deforestation and timber extraction in Borneo and the Malay Peninsula: The record since 1965. *Global Environmental Change* **1**:42–56.

Browning, E. *et al.* (2017). *Passive Acoustic Monitoring in Ecology and Conservation.* [Online]. WWF-UK. Available at: https://repository.oceanbestpractices.org/handle/11329/1370 [Accessed: 1 October 2020].

de Bruyn, M. *et al.* (2014). Borneo and Indochina are Major Evolutionary Hotspots for Southeast Asian Biodiversity. *Systematic Biology* **63**:879–901.

Burgin, C.J. *et al.* (2018). How many species of mammals are there? *Journal of Mammalogy* **99**:1–14.

Burivalova, Z., Şekercioğlu, Ç.H. and Koh, L.P. (2014). Thresholds of Logging Intensity to Maintain Tropical Forest Biodiversity. *Current Biology* **24**:1893–1898.

Burnham, K.P. and Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. [Online]. 2nd ed. New York: Springer-Verlag. Available at: https://www.springer.com/gp/book/9780387953649 [Accessed: 12 October 2021].

Buxton, R.T. *et al.* (2018). Pairing camera traps and acoustic recorders to monitor the ecological impact of human disturbance. *Global Ecology and Conservation* **16**:e00493.

Cardinale, B.J. *et al.* (2012). Biodiversity loss and its impact on humanity. *Nature* **486**:59–67.

Cardinale, B.J. *et al.* (2006). Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* **443**:989–992.

Carignan, V. and Villard, M.-A. (2002). Selecting Indicator Species to Monitor Ecological Integrity: A Review. *Environmental Monitoring and Assessment* **78**:45–61.

Carroll, C. *et al.* (2004). Extinction Debt of Protected Areas in Developing Landscapes. *Conserv. Biol.* **18**:1110–1120.

Castro, A.B. *et al.* (2021). Influence of reduced-impact logging on Central Amazonian bats using a before-after-control-impact design. *Animal Conservation* [Online]. Available at: https://zslpublications.onlinelibrary.wiley.com/doi/abs/10.1111/acv.12739 [Accessed: 6 September 2021].

Castro, I.J. and Michalski, F. (2014). Effects of logging on bats in tropical forests. *Natureza & Conservação* **12**:99–105.

Castro-Arellano, I. *et al.* (2009). Reduced-impact logging and temporal activity of understorey bats in lowland Amazonia. *Biological Conservation* **142**:2131–2139.

Ceballos, G., Ehrlich, P.R. and Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences* **114**:E6089–E6096.

Chao, A. *et al.* (2014). Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs* **84**:45–67.

Chapman, P.M. *et al.* (2018). Inter-annual dynamics and persistence of small mammal communities in a selectively logged tropical forest in Borneo. *Biodiversity and Conservation* **27**:3155–3169.

Chatterjee, D. *et al.* (2014). Ranging pattern and habitat use of Sambar (Rusa unicolor) in Sariska Tiger Reserve, Rajasthan, Western India. *Deer specialist group news*:60.

Chazdon, R.L. *et al.* (2009). The Potential for Species Conservation in Tropical Secondary Forests. *Conservation Biology* **23**:1406–1417.

Chen, S.F., Jones, G. and Rossiter, S.J. (2009). Determinants of echolocation call frequency variation in the Formosan lesser horseshoe bat (Rhinolophus monoceros). *Proceedings of the Royal Society B: Biological Sciences* **276**:3901–3909.

Chen, X. *et al.* (2020). Automatic standardized processing and identification of tropical bat calls using deep learning approaches. *Biological Conservation* **241**:108269.

Chhaya, V. *et al.* (2021). Community Bioacoustics: Studying Acoustic Community Structure for Ecological and Conservation Insights. *Frontiers in Ecology and Evolution* [Online] **9**. Available at: https://www.frontiersin.org/article/10.3389/fevo.2021.706445 [Accessed: 4 February 2022].

ChiroVox (2022). [Online]. Available at: https://www.chirovox.org/ [Accessed: 15 March 2022].

Cisneros, L.M. *et al.* (2014). Multiple dimensions of bat biodiversity along an extensive tropical elevational gradient. *J. Anim. Ecol.* **83**:1124–1136.

Clarke, F., Rostant, L. and Racey, P. (2005). Life after logging: Post-logging recovery of a neotropical bat community. *Journal of Applied Ecology* **42**:409–420.

Cleary, D. *et al.* (2009). Butterfly species and traits associated with selectively logged forest in Borneo. *Basic and Applied Ecology* **10**:237–245.

Cleary, D.F.R. *et al.* (2007). Bird species and traits associated with logged and unlogged forest in Borneo. *Ecological Applications* **17**:1184–1197.

Clement, M.J. *et al.* (2014). The effect of call libraries and acoustic filters on the identification of bat echolocation. *Ecology and Evolution* **4**:3482–3493.

Clink, D.J. and Klinck, H. (2021). Unsupervised acoustic classification of individual gibbon females and the implications for passive acoustic monitoring. *Methods in Ecology and Evolution* **12**:328–341.

Coe, M.T. *et al.* (2013). Deforestation and climate feedbacks threaten the ecological integrity of south–southeastern Amazonia. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**:20120155.

Collen, A.L. (2012). *The Evolution of Echolocation in Bats: A Comparative Approach*. UCL (University College London). Available at: http://discovery.ucl.ac.uk/id/eprint/1370574 [Accessed: 26 November 2019].

Cortés-Delgado, N. and Sosa, V.J. (2014). Do Bats Roost and Forage in Shade Coffee Plantations? A Perspective from the Frugivorous Bat Sturnira hondurensis. *Biotropica* **46**:624–632.

Costantini, D., Edwards, D.P. and Simons, M.J.P. (2016). Life after logging in tropical forests of Borneo: A meta-analysis. *Biological Conservation* **196**:182–188.

Cowie, R.H., Bouchet, P. and Fontaine, B. (2022). The Sixth Mass Extinction: fact, fiction or speculation? *Biological Reviews* [Online] **n/a**. Available at:

https://onlinelibrary.wiley.com/doi/abs/10.1111/brv.12816 [Accessed: 11 February 2022].

Crooks, K.R. and Sanjayan, M. eds. (2006). *Connectivity Conservation*. [Online]. Conservation Biology. Cambridge: Cambridge University Press. Available at: https://www.cambridge.org/core/books/connectivity-conservation/4601E1BD535CC098BD375786F85B718D [Accessed: 13 March 2022].

Curran, L.M. *et al.* (2004). Lowland forest loss in protected areas of Indonesian Borneo. *Science* **303**:1000–1003.

Curran, M. *et al.* (2012). Species diversity of bats along an altitudinal gradient on Mount Mulanje, southern Malawi. *Journal of Tropical Ecology* **28**:243–253.

Curtis, P.G. *et al.* (2018). Classifying drivers of global forest loss. *Science* **361**:1108–1111.

Danielsen, F. and Heegaard, M. (1995). Impact of logging and plantation development on species diversity: a case study from Sumatra. In: *Management of Tropical Forests*. pp. 73–92.

Das, I. (2006). Crocker Range National Park, Sabah, as a refuge for Borneo’s montane herpetofauna. *Amphib. Reptile Conserv.* **4**:3–11.

Day, M. and Urich, P. (2000). An assessment of protected karst landscapes in Southeast Asia. *Cave and Karst Science* **27**:61–70.

Dayley, R. (2019). *Southeast Asia in the New International Era*. 7th ed. New York: Routledge.

Deere, N.J. *et al.* (2017). High Carbon Stock forests provide co-benefits for tropical biodiversity. *Journal of Applied Ecology* **55**:997–1008.

Deere, N J *et al.* (2020). Implications of zero-deforestation commitments: forest quality and hunting pressure limit mammal persistence in fragmented tropical landscapes. *Conservation Letters*.

Deere, Nicolas J. *et al.* (2020). Maximizing the value of forest restoration for tropical mammals by detecting three-dimensional habitat associations. *Proceedings of the National Academy of Sciences* **117**:26254–26262.

Deere, N.J. *et al.* (2022). Riparian buffers can help mitigate biodiversity declines in oil palm agriculture. *Frontiers in Ecology and the Environment* [Online]. Available at: https://kar.kent.ac.uk/90058/ [Accessed: 29 March 2022].

DeFries, R. *et al.* (2007). Land Use Change Around Protected Areas: Management to Balance Human Needs and Ecological Function. *Ecological Applications* **17**:1031–1038.

Denzinger, A. and Schnitzler, H.-U. (2013). Bat guilds, a concept to classify the highly diverse foraging and echolocation behaviors of microchiropteran bats. *Frontiers in Physiology* **4**:164.

Derkarabetian, S. *et al.* (2019). A demonstration of unsupervised machine learning in species delimitation. *Molecular Phylogenetics and Evolution* **139**:106562.

Dodgin, S.R., Hall, C.L. and Howard, D.R. (2020). Eavesdropping on the dead: heterogeneity of tallgrass prairie soundscapes and their relationship with invertebrate necrophilous communities. *Biodiversity* **21**:28–40.

Doohan, B. *et al.* (2019). The sound of management: Acoustic monitoring for agricultural industries. *Ecological Indicators* **96**:739–746.

Dunning, K. (2022). Biodiversity conservation policy in megadiverse countries: Comparing policy systems for 2020 targets to inform management in the coming decades. *Journal of Environmental Management* **302**:113815.

Edwards, D.P., Tobias, J.A., *et al.* (2014). Maintaining ecosystem function and services in logged tropical forests. *Trends in Ecology and Evolution* **29**:511–520.

Edwards, D.P., Magrach, A., *et al.* (2014). Selective-logging and oil palm: multitaxon impacts, biodiversity indicators, and trade-offs for conservation planning. *Ecological Applications* **24**:2029–2049.

Edwards, D.P. *et al.* (2017). Tropical secondary forest regeneration conserves high levels of avian phylogenetic diversity. *Biological Conservation* **209**:432–439.

Edwards, D.P. *et al.* (2010). Wildlife-friendly oil palm plantations fail to protect biodiversity effectively. *Conservation Letters* **3**:236–242.

Estoque, R.C. *et al.* (2019). The future of Southeast Asia’s forests. *Nature Communications* **10**:1829.

Estrada-Villegas, S., Meyer, C.F.J. and Kalko, E.K.V. (2010). Effects of tropical forest fragmentation on aerial insectivorous bats in a land-bridge island system. *Biological Conservation* **143**:597–608.

Evelyn, M.J. and Stiles, D.A. (2003). Roosting Requirements of Two Frugivorous Bats (Sturnira lilium and Artibeus intermedius) in Fragmented Neotropical Forest. *Biotropica* **35**:405–418.

Ewers, R.M. *et al.* (2011). A large-scale forest fragmentation experiment: the Stability of Altered Forest Ecosystems Project. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**:3292–3302.

Ewers, R.M. *et al.* (2015). Logging cuts the functional importance of invertebrates in tropical rainforest. *Nat. Commun.* **6**:6836.

Ewers, R.M. and Didham, R.K. (2006). Confounding factors in the detection of species responses to habitat fragmentation. *Biological Reviews* **81**:117–142.

Fahrig, L. (2017). Ecological Responses to Habitat Fragmentation Per Se. *Annual Review of Ecology, Evolution, and Systematics* **48**:1–23.

FAO (2018). *FAOSTAT*. [Online]. Rome, Italy: Food and Agriculture Organization of the United Nations. Available at: http://www.fao.org/faostat/en/.

FAO (2021). Statistical database: Forests data. License: CC BY-NC-SA 3.0 IGO.

FAO & UNEP (2020). *The State of the World's Forests 2020: Forests, biodiversity and people*. [Online]. The State of the World's Forests (SOFO). Number 2020. Rome, Italy: FAO & UNEP. Available at: https://www.fao.org/documents/card/en/c/ca8642en [Accessed: 19 January 2022].

Faria, D. and Baumgarten, J. (2007). Shade cacao plantations (Theobroma cacao) and bat conservation in southern Bahia, Brazil. *Biodiversity and Conservation* **16**:291–312.

Farneda, F.Z. *et al.* (2015). Trait-related responses to habitat fragmentation in Amazonian bats. *Journal of Applied Ecology* **52**:1381–1391.

Fenton, M.B. and Bell, G.P. (1981). Recognition of Species of Insectivorous Bats by Their Echolocation Calls. *Journal of Mammalogy* **62**:233–243.

Ferreira, D.F. *et al.* (2017). Season-modulated responses of Neotropical bats to forest fragmentation. *Ecology and Evolution* **7**:4059–4071.

Filgueiras, B.K.C. *et al.* (2016). Spatial replacement of dung beetles in edge-affected habitats: biotic homogenization or divergence in fragmented tropical forest landscapes? *Diversity and Distributions* **22**:400–409.

Fischer, R. *et al.* (2021). Accelerated forest fragmentation leads to critical increase in tropical forest edge area. *Science Advances* **7**:eabg7012.

Fisher-Phelps, M. *et al.* (2017). Protecting bias: Across time and ecology, open-source bat locality data are heavily biased by distance to protected area. *Ecological Informatics* **40**:22–34.

Fitzherbert, E.B. *et al.* (2008). How will oil palm expansion affect biodiversity? *Trends in Ecology & Evolution* **23**:538–545.

Fleming, T.H. and Holland, J.N. (2018). Nectar Bat-Plant Interactions in North American Deserts. *Hystrix, the Italian Journal of Mammalogy* **29**:33–39.

Francis, C.M. (1989). A Comparison of Mist Nets and Two Designs of Harp Traps for Capturing Bats. *Journal of Mammalogy* **70**:865–870.

Francis, C.M. (1990). Trophic Structure of Bat Communities in the Understorey of Lowland Dipterocarp Rain Forest in Malaysia. *Journal of Tropical Ecology* **6**:421–431.

Francis, C.M. and Hill, J.E. (1986). A review of the Bornean Pipistrellus (Mammalia : Chiroptera). **50**:43–56.

Frasier, K.E. *et al.* (2017). Automated classification of dolphin echolocation click types from the Gulf of Mexico. *PLOS Computational Biology* **13**:e1005823.

Frick, W. (2013). Acoustic monitoring of bats, considerations of options for long-term monitoring. *Therya* **4**:69–78.

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.

Fukuda, D., Tisen, O., *et al.* (2009). Bat diversity in the vegetation mosaic around a lowland dipterocarp forest of Borneo. *The Raffles Bulletin of Zoology* **57**:213–221.

Fukuda, D., Tisen, O.B., *et al.* (2009). Bat diversity in the vegetation mosaic around a lowland dipterocarp forest of Borneo. *Raffles Bull. Zool.* **57**:213–221.

Furey, N.M., Mackie, I.J. and Racey, P.A. (2010). Bat diversity in Vietnamese limestone karst areas and the implications of forest degradation. *Biodiversity and Conservation* **19**:1821–1838.

Furey, N.M. and Racey, P.A. (2016). Can wing morphology inform conservation priorities for Southeast Asian cave bats? *Biotropica* **48**:545–556.

Furmankiewicz, J. and Kucharska, M. (2009). Migration of Bats along a Large River Valley in Southwestern Poland. *Journal of Mammalogy* **90**:1310–1317.

Galindo-González, J. and Sosa, V.J. (2003). Frugivorous bats in isolated trees and riparian vegetation associated with human-made pastures in a fragmented tropical landscape. *Southwest. Nat.* **48**:579–590.

Gannon, W.L., Sikes, R.S., and The Animal Care and Use Committee of the American Society of Mammalogists (2007). Guidelines of the American Society of Mammalogists for the Use of Wild Mammals in Research. *Journal of Mammalogy* **88**:809–823.

García-Morales, R., Badano, E.I. and Moreno, C.E. (2013). Response of Neotropical Bat Assemblages to Human Land Use. *Conservation Biology* **27**:1096–1106.

Gardner, T.A. *et al.* (2008). The cost-effectiveness of biodiversity surveys in tropical forests. *Ecology Letters* **11**:139–150.

Gaveau, D.L.A. *et al.* (2013). Reconciling Forest Conservation and Logging in Indonesian Borneo. *PLOS ONE* **8**:e69887.

Gaveau, D.L.A. *et al.* (2022). Slowing deforestation in Indonesia follows declining oil palm expansion and lower oil prices. *PLOS ONE* **17**:e0266178.

Gelman, A. and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. [Online]. Cambridge University Press. Available at: https://nyuscholars.nyu.edu/en/publications/data-analysis-using-regression-and-multilevelhierarchical-models [Accessed: 10 December 2019].

Gelman, A., Hwang, J. and Vehtari, A. (2013). Understanding predictive information criteria for Bayesian models. *arXiv [stat.ME]* [Online]. Available at: http://arxiv.org/abs/1307.5928.

Ghanem, S.J. and Voigt, C.C. (2012). Increasing Awareness of Ecosystem Services Provided by Bats. In: Brockmann, H. J. et al. eds. *Advances in the Study of Behavior*. Academic Press, pp. 279–302. Available at: https://www.sciencedirect.com/science/article/pii/B9780123942883000071 [Accessed: 4 February 2022].

Giam, X. (2017). Global biodiversity loss from tropical deforestation. *Proceedings of the National Academy of Sciences* **114**:5775–5777.

Giam, X. *et al.* (2012). Reservoirs of richness: least disturbed tropical forests are centres of undescribed species diversity. *Proceedings of the Royal Society B: Biological Sciences* **279**:67–76.

Gibb, R. *et al.* (2019). Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in Ecology and Evolution* **10**:169–185.

Gibson, L. *et al.* (2011). Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature* **478**:378–381.

Gilroy, J.J. *et al.* (2014). Surrounding habitats mediate the trade-off between land-sharing and land-sparing agriculture in the tropics. *Journal of Applied Ecology* **51**:1337–1346.

Gilroy, J.J. and Edwards, D.P. (2017). Source-Sink Dynamics: a Neglected Problem for Landscape-Scale Biodiversity Conservation in the Tropics. *Current Landscape Ecology Reports* **2**:51–60.

Görföl, T. *et al.* (2022). ChiroVox: a public library of bat calls. *PeerJ* **10**:e12445.

Grantham, H.S. *et al.* (2020). Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity. *Nature Communications* **11**:5978.

Grantham, H.S. *et al.* (2009). Delaying conservation actions for improved knowledge: how long should we wait? *Ecology Letters* **12**:293–301.

Haddad, N.M. *et al.* (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances* **1**:e1500052.

Hamer, K. *et al.* (2014). Impacts of selective logging on insectivorous birds in Borneo: The importance of trophic position, body size and foraging height. *Biological Conservation* **188**.

Hansen, M.C. *et al.* (2013). High-resolution global maps of 21st-century forest cover change. *Science* **342**:850–853.

Harrell Jr, F.E. (2015). *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. 2nd ed. 2015 edition. Cham Heidelberg New York: Springer.

Hartigan, J.A. and Wong, M.A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Applied Statistics* **28**:100.

Hayes, J.P. (2000). Assumptions and practical considerations in the design and interpretation of echolocation-monitoring studies. *Acta Chiropt.* **2**:225–236.

Haynes, W. (2013). Benjamini–Hochberg Method. In: Dubitzky, W. et al. eds. *Encyclopedia of Systems Biology*. New York, NY: Springer, pp. 78–78. Available at: https://doi.org/10.1007/978-1-4419-9863-7_1215 [Accessed: 29 November 2021].

Helbig-Bonitz, M. *et al.* (2015). Bats are Not Birds – Different Responses to Human Land-use on a Tropical Mountain. *Biotropica* **47**:497–508.

Hemprich-Bennett, D.R. *et al.* (2021). Altered structure of bat-prey interaction networks in logged tropical forests revealed by metabarcoding. *Molecular Ecology* **30**:5844–5857.

Henry, M., Cosson, J.F. and Pons, J.M. (2010). Modelling multi-scale spatial variation in species richness from abundance data in a complex neotropical bat assemblage. *Ecological Modelling* **221**:2018–2027.

Hervé, M. (2021). RVAideMemoire: Testing and Plotting Procedures for Biostatistics. [R package]. Available at: https://CRAN.R-project.org/package=RVAideMemoire.

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

Hill, J.L. and Curran, P.J. (2003). Area, shape and isolation of tropical forest fragments: effects on tree species diversity and implications for conservation. *Journal of Biogeography* **30**:1391–1403.

Holl, K.D., Luong, J.C. and Brancalion, P.H.S. (2022). Overcoming biotic homogenization in ecological restoration. *Trends in Ecology & Evolution* [Online]. Available at:

https://www.sciencedirect.com/science/article/pii/S0169534722001136 [Accessed: 13 July 2022].

Holt, E.A. and Miller, S.W. (2010). Bioindicators: Using Organisms to Measure Environmental Impacts. *Nature Education Knowledge* **3**:8.

Hothorn, T., Bretz, F. and Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal* **50**:346–363.

Hsieh, T.C., Ma, K.H. and Chao, A. (2016). iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution* **7**:1451–1456.

Huang, J.C.-C. *et al.* (2014). A Recent Bat Survey Reveals Bukit Barisan Selatan Landscape as a Chiropteran Diversity Hotspot in Sumatra. *Acta Chiropterologica* **16**:413–449.

Huang, J.C.-C. *et al.* (2019). Echolocation and roosting ecology determine sensitivity of forest-dependent bats to coffee agriculture. *Biotropica* **51**:757–768.

Huang, J.C.-C. (2020). Rhinolophus trifoliatus. *The IUCN Red List of Threatened Species* **2020**:e.T19574A21990821.

Hughes, A.C. *et al.* (2011). Using Echolocation Calls to Identify Thai Bat Species: Vespertilionidae, Emballonuridae, Nycteridae and Megadermatidae. *Acta Chiropterologica* **13**:447–455.

Hylander, K. and Ehrlén, J. (2013). The mechanisms causing extinction debts. *Trends in Ecology & Evolution* **28**:341–346.

Imai, N. *et al.* (2018). Factors affecting forest area change in Southeast Asia during 1980-2010. *PLOS ONE* **13**:e0197391.

IMARC Group (2022). *Palm Oil Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2022-2027.* [Online]. Available at: https://www.imarcgroup.com/palm-oil-processing-plant.

Ishwaran, N., Persic, A. and Tri, N.H. (2008). Concept and practice: the case of UNESCO biosphere reserves. *International Journal of Environment and Sustainable Development* **7**:118.

IUCN (2022). *IUCN Red List of Threatened Species. Version 2021-3.* [Online]. Available at: http://www.iucnredlist.org. [Accessed: 11 April 2022].

IUCN (2019). The IUCN Red List of Threatened Species. [Online]. Available at: https://www.iucnredlist.org.

IUCN (2021). The IUCN Red List of Threatened Species. Version 2021-1. [Online]. Available at: https://www.iucnredlist.org/en [Accessed: 9 April 2021].

James, G. *et al.* (2013). *An Introduction to Statistical Learning: With Applications in R*. Springer International Publishing.

Jarvis, R.M. *et al.* (2020). Navigating spaces between conservation research and practice: Are we making progress? *Ecological Solutions and Evidence* **1**:e12028.

Jayaraj, V.K. (2022). Rhinolophus sedulus. *The IUCN Red List of Threatened Species* **2020**:e.T19565A21994153.

Jenkins, C.N. and Joppa, L. (2009). Expansion of the global terrestrial protected area system. *Biological Conservation* **142**:2166–2174.

Jennings, N., Parsons, S. and Pocock, M.J.O. (2008). Human vs. machine: identification of bat species from their echolocation calls by humans and by artificial neural networks. *Canadian Journal of Zoology* **86**:371–377.

Joann Christine, L., Fletcher, C. and Abd Rahman, K. (2013). Spatial effects of Virgin Jungle Reserve (VJR) on the community of insectivorous bats in peninsular Malaysia. *Journal of Tropical Forest Science* **25**:118–130.

Jones, Gareth *et al.* (2009). Carpe noctem: the importance of bats as bioindicators. *Endangered Species Research* **8**:93–115.

Jones, G *et al.* (2009). Carpe noctem: the importance of bats as bioindicators. *Endanger. Species Res.* **8**:93–115.

Jones, G. (1999). Scaling of echolocation call parameters in bats. *J. Exp. Biol.* **202**:3359–3367.

Jones, G. and Rydell, J. (1994). Foraging strategy and predation risk as factors influencing emergence time in echolocating bats. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **346**:445–455.

Jones, K.E. *et al.* (2013). Indicator Bats Program: A System for the Global Acoustic Monitoring of Bats. In: Collen, B. et al. eds. *Biodiversity Monitoring and Conservation: Bridging the Gap between Global Commitment and Local Action*. 1st ed. Online: John Wiley & Sons, pp. 2013–247. Available at: https://onlinelibrary.wiley.com/doi/10.1002/9781118490747.ch10 [Accessed: 28 September 2020].

Jucker, T., Hardwick, S.R., *et al.* (2018). Canopy structure and topography jointly constrain the microclimate of human-modified tropical landscapes. *Glob. Chang. Biol.* **24**:5243–5258.

Jucker, T., Asner, G.P., *et al.* (2018). Estimating aboveground carbon density and its uncertainty in Borneo’s structurally complex tropical forests using airborne laser scanning. *Biogeosciences* **15**:3811–3830.

Jung, K., Molinari, J. and Kalko, E.K.V. (2014). Driving Factors for the Evolution of Species-Specific Echolocation Call Design in New World Free-Tailed Bats (Molossidae). *PLOS ONE* **9**:e85279.

Kalko, E.K.V. and Schnitzler, H.-U. (1993). Plasticity in echolocation signals of European pipistrelle bats in search flight: implications for habitat use and prey detection. *Behavioral Ecology and Sociobiology* **33**:415–428.

Kent, R. *et al.* (2015). Airborne LiDAR Detects Selectively Logged Tropical Forest Even in an Advanced Stage of Recovery. *Remote Sensing* **7**:8348–8367.

Kershenbaum, A. *et al.* (2016). Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biological Reviews* **91**:13–52.

Kingston, T., Jones, G., *et al.* (2003). Alternation of Echolocation Calls in 5 Species of Aerial-Feeding Insectivorous Bats from Malaysia. *Journal of Mammalogy* **84**:205–215.

Kingston, T. (2016). Bats. In: Larsen, T. H. ed. *Core Standardized Methods for Rapid Biological Field Assessment*. Arlington, VA: Conservation International, pp. 59–82.

Kingston, T. *et al.* (1999). Echolocation signal design in Kerivoulinae and Murininae (Chiroptera: Vespertilionidae) from Malaysia. *Journal of Zoology* **249**:359–374.

Kingston, T. (2010a). Research priorities for bat conservation in Southeast Asia: a consensus approach. *Biodivers. Conserv.* **19**:471–484.

Kingston, T. (2010b). Research priorities for bat conservation in Southeast Asia: a consensus approach. *Biodiversity and Conservation* **19**:471–484.

Kingston, T. *et al.* (2000). Resource Partitioning in Rhinolophoid Bats Revisited. *Oecologia* **124**:332–342.

Kingston, T. (2013). Response of Bat Diversity to Forest Disturbance in Southeast Asia: Insights from Long-Term Research in Malaysia. In: Adams, R. A. and Pedersen, S. C. eds. *Bat Evolution, Ecology, and Conservation*. New York, NY: Springer New York, pp. 169–185. Available at: https://doi.org/10.1007/978-1-4614-7397-8_9.

Kingston, T., Francis, C.M., *et al.* (2003). Species richness in an insectivorous bat assemblage from Malaysia. *Journal of Tropical Ecology* **19**:67–79.

Kingston, T., Lim, B.L. and Akbar, Z. (2006). *Bats of Krau Wildlife Reserve*. Penerbit Universiti Kebangsaan Malaysia.

Kitayama, K. (1992). An altitudinal transect study of the vegetation on Mount Kinabalu, Borneo. *Vegetatio* **102**:149–171.

Klingbeil, B.T. and Willig, M.R. (2010). Seasonal differences in population-, ensemble- and community-level responses of bats to landscape structure in Amazonia. *Oikos* **119**:1654–1664.

Kobayashi, K. *et al.* (2021). Development of a species identification system of Japanese bats from echolocation calls using convolutional neural networks. *Ecological Informatics* **62**:101253.

Koh, L.P. and Wilcove, D.S. (2008). Is oil palm agriculture really destroying tropical biodiversity? *Conservation Letters* **1**:60–64.

Korner-Nievergelt, F. *et al.* (2015). *Bayesian Data Analysis in Ecology Using Linear Models with R, BUGS, and Stan*. Academic Press.

Kuhn, M. *et al.* (2020). caret: Classification and Regression Training. [Online]. Available at: https://CRAN.R-project.org/package=caret [Accessed: 1 October 2020].

Kunz, T.H. *et al.* (2011). Ecosystem services provided by bats. *Annals of the New York Academy of Sciences* **1223**:1–38.

Kunz, T.H. (1982). Roosting Ecology of Bats. In: Kunz, T. H. ed. *Ecology of Bats*. Boston, MA: Springer US, pp. 1–55. Available at: https://doi.org/10.1007/978-1-4613-3421-7_1 [Accessed: 13 March 2022].

Kupfer, J.A., Malanson, G.P. and Franklin, S.B. (2006). Not seeing the ocean for the islands: the mediating influence of matrix-based processes on forest fragmentation effects. *Global Ecology and Biogeography* **15**:8–20.

Kusch, J. *et al.* (2004). Foraging habitat preferences of bats in relation to food supply and spatial vegetation structures in a western European low mountain range forest. *Folia Zoologica* **53**:113–128.

Kusuminda, T. and Yapa, W.B. (2017). First record of a Wrinkle-lipped Free-tailed Bat Chaerephon plicatus Buchannan, 1800 (Mammalia: Chiroptera: Molossidae) colony in Sri Lanka, with notes on echolocation calls and taxonomy. *Journal of Threatened Taxa* **9**:10115–10120.

Kwok, R. (2019). AI empowers conservation biology. *Nature* **567**:133–134.

Lane, D.J.W., Kingston, T. and Lee, B.P.Y.H. (2006). Dramatic decline in bat species richness in Singapore, with implications for Southeast Asia. *Biological Conservation* **131**:584–593.

Laurance, W.F. *et al.* (2018). An Amazonian rainforest and its fragments as a laboratory of global change: Amazonian fragments and global change. *Biol Rev* **93**:223–247.

Laurance, W.F. *et al.* (2012). Averting biodiversity collapse in tropical forest protected areas. *Nature* **489**:290–294.

Lawrence, B.D. and Simmons, J.A. (1982). Measurements of atmospheric attenuation at ultrasonic frequencies and the significance for echolocation by bats. *The Journal of the Acoustical Society of America* **71**:585.

Lazarus, B.A. *et al.* (2021). Impact of natural salt lick on the home range of Panthera tigris at the Royal Belum Rainforest, Malaysia. *Scientific Reports* **11**:10596.

Leblois, A., Damette, O. and Wolfersberger, J. (2017). What has Driven Deforestation in Developing Countries Since the 2000s? Evidence from New Remote-Sensing Data. *World Development* **92**:82–102.

Leelapaibul, W., Bumrungsri, S. and Pattanawiboon, A. (2005). Diet of wrinkle-lipped free-tailed bat (Tadarida plicata Buchannan, 1800) in central Thailand: insectivorous bats potentially act as biological pest control agents. *Acta Chiropterologica* **7**:111–119.

Licona, M. *et al.* (2011). Using ungulate occurrence to evaluate community-based conservation within a biosphere reserve model: Ungulate occurrence in Amazon. *Anim. Conserv.* **14**:206–214.

Linden, V.M.G. *et al.* (2014). Changes of Bat Activity, Species Richness, Diversity and Community Composition Over an Altitudinal Gradient in the Soutpansberg Range, South Africa. *Acta Chiropterologica* **16**:27–40.

Lintott, P.R. *et al.* (2018). Ecobat: An online resource to facilitate transparent, evidence-based interpretation of bat activity data. *Ecology and Evolution* **8**:935–941.

Liu, S. *et al.* (2008). Characteristics and Analysis of Echolocation Calls by Coelops frithi. [Online]. Available at: https://tspace.library.utoronto.ca/handle/1807/64044 [Accessed: 1 December 2020].

Lok, Y.C. *et al.* (2021). Bat species diversity trend along an elevation gradient: A study in Crocker Range Park, Sabah, Borneo. *Biodiversity Data Journal* **9**:e72651.

López-Baucells, A. *et al.* (2021). Optimizing bat bioacoustic surveys in human-modified Neotropical landscapes. *Ecological Applications* **31**:e02366.

López-Baucells, A. *et al.* (2019). Stronger together: Combining automated classifiers with manual post-validation optimizes the workload vs reliability trade-off of species identification in bat acoustic surveys. *Ecological Informatics* **49**:45–53.

López-González, C. *et al.* (2015). Activity of Insectivorous Bats Associated with Cattle Ponds at La Michilía Biosphere Reserve, Durango, Mexico: Implications for Conservation. *Acta Chiropterologica* **17**:117–129.

Low, M.-R. *et al.* (2021). Bane or Blessing? Reviewing Cultural Values of Bats across the Asia-Pacific Region. *Journal of Ethnobiology* **41**:18–34.

Luke, S.H. *et al.* (2019). Riparian buffers in tropical agriculture: Scientific support, effectiveness and directions for policy. *J. Appl. Ecol.* **56**:85–92.

Luyssaert, S. *et al.* (2014). Land management and land-cover change have impacts of similar magnitude on surface temperature. *Nature Climate Change* **4**:389–393.

MacKenzie, D.I. *et al.* (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology* **83**:2248–2255.

Magrach, A. and Sanz, M.J. (2020). Environmental and social consequences of the increase in the demand for 'superfoods' world-wide. *People and Nature* **2**:267–278.

Malhi, Y. *et al.* (2008). Climate Change, Deforestation, and the Fate of the Amazon. *Science* [Online]. Available at: https://www.science.org/doi/abs/10.1126/science.1146961 [Accessed: 27 January 2022].

Malhi, Y. *et al.* (2014). Tropical Forests in the Anthropocene. *Annu. Rev. Environ. Resour.* **39**:125–159.

Marchal, J., Fabianek, F. and Scott, C. (2020). Bioacoustics: Analyse Audio Recordings and Automatically Extract Animal Vocalizations. [Online]. Available at: https://CRAN.R-project.org/package=bioacoustics.

Martin, A.R. and Thomas, S.C. (2011). A Reassessment of Carbon Content in Tropical Trees. *PLOS ONE* **6**:e23533.

Mattos, I. de, Zimbres, B. and Marinho-Filho, J. (2021). Habitat Specificity Modulates the Response of Small Mammals to Habitat Fragmentation, Loss, and Quality in a Neotropical Savanna. *Frontiers in Ecology and Evolution* [Online] **9**. Available at: https://www.frontiersin.org/article/10.3389/fevo.2021.751315 [Accessed: 1 April 2022].

McArthur, E. and Khan, F.A.A. (2021). Towards a regional call library: Classifying calls of a species-rich bat assemblage in a Bornean karst rainforest. *Journal of Bat Research & Conservation* **14**:95–117.

McHugh, M.L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica* **22**:276–282.

Medway, Lord and Harrisson, T. (1963). MAMMALS OF BORNEO: Field Keys and an Annotated checklist. *Journal of the Malayan Branch of the Royal Asiatic Society* **36**:i–193.

Meijaard, E. *et al.* (2020). Coconut oil, conservation and the conscientious consumer. *Current biology: CB* **30**:3274–3275.

Meijaard, E. *et al.* (2005). *Life After Logging: Reconciling Wildlife Conservation and Production Forestry in Indonesian Borneo*.

Meijaard, E. *et al.* (2018). *Oil Palm and Biodiversity: A Situation Analysis by the IUCN Oil Palm Task Force*. [Online]. Gland, Switzerland. Available at: https://www.cifor.org/library/6940/ [Accessed: 30 October 2019].

Meijaard, E. and Sheil, D. (2007). A logged forest in Borneo is better than none at all. *Nature* **446**:974–974.

Mello, K. de *et al.* (2018). Impacts of tropical forest cover on water quality in agricultural watersheds in southeastern Brazil. *Ecological Indicators* **93**:1293–1301.

Meyer, C.F.J. (2015). Methodological challenges in monitoring bat population- and assemblage-level changes for anthropogenic impact assessment. *Mammalian Biology* **80**:159–169.

Meyer, C.F.J. and Kalko, E.K.V. (2008). Assemblage-level responses of phyllostomid bats to tropical forest fragmentation: land-bridge islands as a model system. *Journal of Biogeography* **35**:1711–1726.

Meyer, C.F.J., Schwarz, C.J. and Fahr, J. (2004). Activity patterns and habitat preferences of insectivorous bats in a West African forest–savanna mosaic. *Journal of Tropical Ecology* **20**:397–407.

Meyer, C.F.J., Struebig, M.J. and Willig, M.R. (2016). Responses of tropical bats to habitat fragmentation, logging, and deforestation. In: *Bats in the Anthropocene: Conservation of Bats in a Changing World*. Springer, pp. 63–103. Available at: https://www.oapen.org/download?type=document&docid=1001864#page=68.

Miettinen, J., Shi, C. and Liew, S.C. (2011). Deforestation rates in insular Southeast Asia between 2000 and 2010. *Global Change Biology* **17**:2261–2270.

Mitchell, Simon L *et al.* (2018). Riparian reserves help protect forest bird communities in oil palm dominated landscapes. *J. Appl. Ecol.* **55**:2744–2755.

Mitchell, Simon L. *et al.* (2018). Riparian reserves help protect forest bird communities in oil palm dominated landscapes. *Journal of Applied Ecology* **55**:2744–2755.

Monadjem, A. *et al.* (2017). Acoustic Call Library and Detection Distances for Bats of Swaziland. *Acta Chiropterologica* **19**:175–187.

Monadjem, A. and Reside, A. (2008). The influence of riparian vegetation on the distribution and abundance of bats in an African savanna. *Acta Chiropterologica* **10**:339–348.

Montauban, C. *et al.* (2021). Bat echolocation plasticity in allopatry: a call for caution in acoustic identification of Pipistrellus sp. *Behavioral Ecology and Sociobiology* [Online] **75**. Available at: https://link.springer.com/epdf/10.1007/s00265-021-03002-7 [Accessed: 8 April 2021].

Mora, C. *et al.* (2011). How Many Species Are There on Earth and in the Ocean? *PLOS Biology* **9**:e1001127.

Mora, C. and Sale, P. (2011). Ongoing global biodiversity loss and the need to move beyond protected areas: A review of the technical and practical shortcomings of protected areas on land and sea. *Marine Ecology-Progress Series* **434**:251–266.

Motoko, F. *et al.* (2012). *Mammals & Birds of Bukit Batu Area of Giam Siak Kecil-Bukit Batu Biosphere Reserve, Raiu, Indonesia*. Kyoto Working Papers on Area Studies (G-COE Series).

Müller, J. *et al.* (2012). Aggregative response in bats: prey abundance versus habitat. *Oecologia* **169**:673–684.

Mullin, K.E. *et al.* (2020). Riparian Reserves Promote Insectivorous Bat Activity in Oil Palm Dominated Landscapes. *Frontiers in Forests and Global Change* **3**:73.

Murphy, D.J., Goggin, K. and Paterson, R.R.M. (2021). Oil palm in the 2020s and beyond: challenges and solutions. *CABI Agriculture and Bioscience* **2**:39.

Naiman, R.J., Decamps, H. and McClain, M.E. (2010). *Riparia: Ecology, Conservation, and Management of Streamside Communities*. [Online]. Elsevier. Available at: https://play.google.com/store/books/details?id=n6i_2G2f2KAC.

Nash, J.C. (2014). On Best Practice Optimization Methods in R. *Journal of Statistical Software* **60**:1–14.

Neuweiler, G. (2000). *The Biology of Bats*. Oxford University Press.

Newbold, T. *et al.* (2015). Global effects of land use on local terrestrial biodiversity. *Nature* **520**:45–50.

Newton, P. *et al.* (2020). The Number and Spatial Distribution of Forest-Proximate People Globally. *One Earth* **3**:363–370.

Ng, W.P. *et al.* (2021). Mammal species composition and habitat associations in a commercial forest and mixed-plantation landscape. *Forest Ecology and Management* **491**:119163.

Norberg, U.M., Rayner, J.M.V. and Lighthill, M.J. (1987). Ecological morphology and flight in bats (Mammalia; Chiroptera): wing adaptations, flight performance, foraging strategy and echolocation. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* **316**:335–427.

Normile, D. (2022). Indonesia’s utopian new capital may not be as green as it looks. *Science* [Online] **375**. Available at: https://www.science.org/content/article/indonesia-s-utopian-new-capital-may-not-be-green-it-looks [Accessed: 1 April 2022].

Numa, C., Verdú, J.R. and Sánchez-Palomino, P. (2005). Phyllostomid bat diversity in a variegated coffee landscape. *Biological Conservation* **122**:151–158.

Núñez, S.F. *et al.* (2019). Echolocation and Stratum Preference: Key Trait Correlates of Vulnerability of Insectivorous Bats to Tropical Forest Fragmentation. *Frontiers in Ecology and Evolution* **7**:373.

Nurul-Ain, E., Rosli, H. and Kingston, T. (2017). Resource availability and roosting ecology shape reproductive phenology of rain forest insectivorous bats. *Biotropica* **49**:382–394.

Oksanen, J. *et al.* (2020). vegan: Community Ecology Package. [Online]. Available at: https://CRAN.R-project.org/package=vegan [Accessed: 2 November 2021].

Olden, J.D., Lawler, J.J. and Poff, N.L. (2008). Machine learning methods without tears: a primer for ecologists. *The Quarterly Review of Biology* **83**:171–193.

d’Oliveira, M.V.N. *et al.* (2012). Estimating forest biomass and identifying low-intensity logging areas using airborne scanning lidar in Antimary State Forest, Acre State, Western Brazilian Amazon. *Remote Sensing of Environment* **124**:479–491.

Ordway, E.M. and Asner, G.P. (2020). Carbon declines along tropical forest edges correspond to heterogeneous effects on canopy structure and function. *Proceedings of the National Academy of Sciences* **117**:7863–7870.

Pacifici, K. *et al.* (2014). Guidelines for a priori grouping of species in hierarchical community models. *Ecol. Evol.* **4**:877–888.

Padfield, R. *et al.* (2019). Co-producing a research agenda for sustainable palm oil. *Frontiers in Forests and Global Change* [Online]. Available at: https://kar.kent.ac.uk/73349/ [Accessed: 18 December 2019].

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**:1178–1190.

Park, K.J. (2015). Mitigating the impacts of agriculture on biodiversity: bats and their potential role as bioindicators. *Mammalian Biology* **80**:191–204.

Parsons, S. and Jones, G. (2000). Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks. *The Journal of Experimental Biology* **203**:2641–2656.

Pawar, S. (2003). Taxonomic Chauvinism and the Methodologically Challenged. *BioScience* **53**:861–864.

Payne, J. (2000). *Panduan Lapangan Mamalia Di Kalimantan, Sabah, Sarawak Dan.*. The Sabah Society.

Payne, J., Francis, C.M. and Phillipps, K. (1985). *A Field Guide to the Mammals of Borneo*. 1st edition. Kota Kinabalu, Malaysia: The Sabah Society and World Wildlife Fund.

Pearson, T.R.H. *et al.* (2017). Greenhouse gas emissions from tropical forest degradation: an underestimated source. *Carbon Balance and Management* **12**:3.

Peña-Cuéllar, E. de la *et al.* (2015). Structure and diversity of phyllostomid bat assemblages on riparian corridors in a human-dominated tropical landscape. *Ecology and Evolution* **5**:903–913.

Phalan, B. *et al.* (2013). Crop Expansion and Conservation Priorities in Tropical Countries. *PLoS ONE* **8**:e51759.

Pham, L.K. *et al.* (2021). Description of Echolocation Call Parameters for Urban Bats in Vietnam as a Step Towards a More Integrated Acoustic Monitoring of Urban Wildlife in Southeast Asia. *Diversity* **13**:18.

Phelps, K. *et al.* (2018). Assemblage and Species Threshold Responses to Environmental and Disturbance Gradients Shape Bat Diversity in Disturbed Cave Landscapes. *Diversity* **10**:55.

Phillipps, Q. and Phillipps, K. (2016). *Phillipps' Field Guide to the Mammals of Borneo*. 1st ed. Oxford, UK: John Beaufoy Publishing Ltd.

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

Phommexay, P. *et al.* (2011). The impact of rubber plantations on the diversity and activity of understorey insectivorous bats in southern Thailand. *Biodiversity and Conservation* **20**:1441–1456.

Pichler, T.R. *et al.* (2022). Comparative biogeography of volant and nonvolant mammals in a temperate island archipelago. *Ecosphere* **13**:e3911.

Pineda, E. *et al.* (2005). Frog, Bat, and Dung Beetle Diversity in the Cloud Forest and Coffee Agroecosystems of Veracruz, Mexico. *Conservation Biology* **19**:400–410.

Pinto, N. and Keitt, T.H. (2008). Scale-Dependent Responses to Forest Cover Displayed by Frugivore Bats. *Oikos* **117**:1725–1731.

Pirker, J. *et al.* (2016). What are the limits to oil palm expansion? *Glob. Environ. Change* **40**:73–81.

Pottie, S.A. *et al.* (2005). The microchiropteran bat fauna of Singapore. *Acta Chiropterologica* **7**:237–247.

Presley, S.J. *et al.* (2008). Effects of reduced-impact logging and forest physiognomy on bat populations of lowland Amazonian forest. *Journal of Applied Ecology* **45**:14–25.

Putz, F.E. *et al.* (2012). Sustaining conservation values in selectively logged tropical forests: the attained and the attainable. *Conservation Letters* **5**:296–303.

R Core Team (2017). R: A Language and Environment for Statistical Computing.

Raghuram, H. *et al.* (2011). Foraging ecology of pteropodid bats: Pollination and seed dispersal. *Bats: Biology, Behavior and Conservation*:177–188.

Rai, N.D. *et al.* (2021). Beyond fortress conservation: The long-term integration of natural and social science research for an inclusive conservation practice in India. *Biological Conservation* **254**:108888.

Ramírez-Fráncel, L.A. *et al.* (2022). Bats and their vital ecosystem services: a global review. *Integrative Zoology* **17**:2–23.

Rangel Pinagé, E. *et al.* (2019). Long-Term Impacts of Selective Logging on Amazon Forest Dynamics from Multi-Temporal Airborne LiDAR. *Remote Sensing* **11**:709.

Raven, P.H. (1988). Our Diminishing Tropical Forests. In: Wilson, E. O. and Peter, F. M. eds. *Biodiversity*. National Academies Press (US). Available at: https://www.ncbi.nlm.nih.gov/books/NBK219329/ [Accessed: 4 October 2019].

Read, J.L., Parkhurst, B. and Delean, S. (2015). Can Australian bush birds be used as canaries? Detection of pervasive environmental impacts at an arid Australian mine site. *Emu - Austral Ornithology* **115**:117–125.

Reddy, S. and Davalos, L. (2003). Geographic sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography* **30**.

Reed, M.G. (2019). The contributions of UNESCO Man and Biosphere Programme and biosphere reserves to the practice of sustainability science. *Sustainability Science* **14**:809–821.

Regolin, A.L. *et al.* (2021). Habitat quality, not habitat amount, drives mammalian habitat use in the Brazilian Pantanal. *Landscape Ecology* **36**:2519–2533.

Revilla-Martín, N. *et al.* (2020). Monitoring cave-dwelling bats using remote passive acoustic detectors: a new approach for cave monitoring. *Bioacoustics*:1–16.

Rice, J.C. and Rochet, M.-J. (2005). A framework for selecting a suite of indicators for fisheries management. *ICES Journal of Marine Science* **62**:516–527.

Ricketts, T.H. (2001). The matrix matters: effective isolation in fragmented landscapes. *The American Naturalist* **158**:87–99.

Ritchie, H. (2021). The world has lost one-third of its forest, but an end of deforestation is possible. *Published online at OurWorldInData.org* [Online]. Available at: https://ourworldindata.org/world-lost-one-third-forests.

Rocha, R. *et al.* (2017). Consequences of a large-scale fragmentation experiment for Neotropical bats: disentangling the relative importance of local and landscape-scale effects. *Landscape Ecology* **32**:31–45.

Rocha, R., López-Baucells, A. and Fernández-Llamazares, Á. (2021). Ethnobiology of Bats: Exploring Human-Bat Inter-Relationships in a Rapidly Changing World. *Journal of Ethnobiology* **41**:3–17.

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.

Rosa, I.M.D. *et al.* (2016). The Environmental Legacy of Modern Tropical Deforestation. *Curr. Biol.* **26**:2161–2166.

Rowley, J.J.L. and Callaghan, C.T. (2020). The FrogID dataset: expert-validated occurrence records of Australia’s frogs collected by citizen scientists. *ZooKeys* **912**:139–151.

Royle, J.A. and Dorazio, R.M. (2008). *Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations and Communities*. [Online]. San Diego, California: Academic Press. Available at: http://dx.doi.org/10.1016/B978-0-12-374097-7.50001-5.

RSPO Principles & Criteria for the Production of Sustainable Palm Oil 2018 (2018). Roundtable on Sustainable Palm Oil.

Runyan, C.W., D’Odorico, P. and Lawrence, D. (2012). Physical and biological feedbacks of deforestation. *Reviews of Geophysics* [Online] **50**. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1029/2012RG000394 [Accessed: 27 January 2022].

Russo, D. *et al.* (2021). Do We Need to Use Bats as Bioindicators? *Biology* **10**:693.

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. and Jones, G. (2015). Bats as bioindicators: an introduction. *Mammalian Biology* **80**:157–158.

Russo, D. and Voigt, C.C. (2016). The use of automated identification of bat echolocation calls in acoustic monitoring: A cautionary note for a sound analysis. *Ecological Indicators* **66**:598–602.

Sagot, M. and Chaverri, G. (2015). Effects of roost specialization on extinction risk in bats. *Conservation Biology* **29**:1666–1673.

Salafsky, N. *et al.* (2019). Defining and using evidence in conservation practice. *Conservation Science and Practice* **1**:e27.

Samek, J. *et al.* (2004). Land Use and Land Cover Change in Southeast Asia. In: pp. 111–122.

Saw, L.G. (2010). Vegetation of Peninsular Malaysia. In: pp. 21–300.

Schipper, J. *et al.* (2008). The status of the world's land and marine mammals: diversity, threat, and knowledge. *Science* **322**:225–230.

Schmieder, D.A. *et al.* (2012). Sensory constraints on prey detection performance in an ensemble of vespertilionid understorey rain forest bats. *Functional Ecology* **26**:1043–1053.

Schnitzler, H.-U. and Denzinger, A. (2011). Auditory fovea and Doppler shift compensation: adaptations for flutter detection in echolocating bats using CF-FM signals. *Journal of Comparative Physiology A* **197**:541–559.

Schnitzler, H.-U. and Kalko, E.K.V. (2001). Echolocation by Insect-Eating Bats: We define four distinct functional groups of bats and find differences in signal structure that correlate with the typical echolocation tasks faced by each group. *BioScience* **51**:557–569.

Schnitzler, H.-U., Moss, C.F. and Denzinger, A. (2003). From spatial orientation to food acquisition in echolocating bats. *Trends in Ecology & Evolution* **18**:386–394.

Schulte to Bühne, H. *et al.* (2021). Improving Predictions of Climate Change–Land Use Change Interactions. *Trends in Ecology & Evolution* **36**:29–38.

Sedlock, J.L. *et al.* (2014). A Survey of Bats in a Karst Landscape in the Central Philippines. *Acta Chiropterologica* **16**:197–211.

Sekercioglu, C.H. *et al.* (2007). Persistence of Forest Birds in the Costa Rican Agricultural Countryside. *Conservation Biology* **21**:482–494.

Seltmann, A. *et al.* (2017). Habitat disturbance results in chronic stress and impaired health status in forest-dwelling paleotropical bats. *Conservation Physiology* **5**:cox020.

Senawi, J. and Kingston, T. (2019). Clutter negotiating ability in an ensemble of forest interior bats is driven by body mass. *Journal of Experimental Biology* [Online] **222**. Available at: http://dx.doi.org/10.1242/jeb.203950.

Shafie, N. *et al.* (2011). Diversity Pattern of Bats at Two Contrasting Habitat Types along Kerian River, Perak, Malaysia. *Tropical life sciences research* **22**:13–22.

Sheherazade, Ober, H.K. and Tsang, S.M. (2019). Contributions of bats to the local economy through durian pollination in Sulawesi, Indonesia. *Biotropica* [Online]. Available at: https://onlinelibrary.wiley.com/doi/10.1111/btp.12712 [Accessed: 5 September 2021].

Shigetomi, Y., Ishimura, Y. and Yamamoto, Y. (2020). Trends in global dependency on the Indonesian palm oil and resultant environmental impacts. *Scientific Reports* **10**:20624.

Siemers, B.M. and Schnitzler, H.-U. (2004). Echolocation signals reflect niche differentiation in five sympatric congeneric bat species. *Nature* **429**:657–661.

Silva, C.A. and Lima, M. (2018). Soy Moratorium in Mato Grosso: Deforestation undermines the agreement. *Land Use Policy* **71**:540–542.

Silva, C.R. and Bernard, E. (2017). Bioacoustics as an Important Complementary Tool in Bat Inventories in the Caatinga Drylands of Brazil. *Acta Chiropterologica* **19**:409–418.

Silva, D.J. *et al.* (2022). Habitat Quality, Not Patch Size, Modulates Lizard Responses to Habitat Loss and Fragmentation in the Southwestern Amazon. *Journal of Herpetology* **56**:75–83.

Silva Junior, C.H.L. *et al.* (2021). The Brazilian Amazon deforestation rate in 2020 is the greatest of the decade. *Nature Ecology & Evolution* **5**:144–145.

Simmons, J.A. *et al.* (1977). Simmons JA, Altes RA, Beuter KJ, Bullock TH, Capranica RR, Goldstein JL, Griffin DR, Konishi M, Neff WD, Neuweiler G, Schnitzler HU, Schuller G, Sovijarvi ARA, Suga N (1977) Localization and identification of acoustic signals, with refer- ence to echolocation. In: Bullock TH (ed), . Aberkon, Berlin, pp 239-277. In: Bullock, T. ed. *Recognition of Complex Acoustic Signals*. Berlin: Aberkon, pp. 239–277.

Simmons, N.B. and Cirranello, A.L. (2021a). *Bat Species of the World: A Taxonomic and Geographic Database* [Online]. Available at: https://batnames.org/home.html [Accessed: 9 April 2021].

Simmons, N.B. and Cirranello, A.L. (2021b). *Bat Species of the World: A Taxonomic and Geographic Database*.

Simonson, W.D., Allen, H.D. and Coomes, D.A. (2014). Applications of airborne lidar for the assessment of animal species diversity. *Methods in Ecology and Evolution* **5**:719–729.

Sodhi, N. *et al.* (2011). Effects of Habitat Fragmentation on Tropical Birds. In: *Conservation of Tropical Birds*. New Jersey, USA: Blackwell Publishing Ltd, pp. 27–44.

Sodhi, N.S., Brook, B.W. and Bradshaw, J. (2013). *Tropical Conservation Biology*. first. Oxford, UK.: John Wiley & Sons.

Soisook, P. *et al.* (2015). Description of a New Species of the Rhinolophus trifoliatus-Group (Chiroptera: Rhinolophidae) from Southeast Asia. *Acta Chiropterologica* **17**:21–36.

Song, X.-P. *et al.* (2018). Global land change from 1982 to 2016. *Nature* **560**:639–643.

SonoBat-4 (2017). [Online]. Available at: https://peersonic.co.uk/sonobat [Accessed: 8 April 2021].

Start, A.N. (1975). *The Feeding Biology in Relation to Food Sources of Nectarivorous Bats (Chiroptera: Macroglossinae) in Malaysia*. Ph.D. Thesis, University of Aberdeen. Available at: https://www.proquest.com/openview/91a08a3bcf65041c8976668ac731cb64/1?pq-origsite=gscholar&cbl=51922&diss=y.

Stasinopoulos, M.D. *et al.* (2020). *Flexible Regression and Smoothing: Using GAMLSS in R*. 1st edition. S.l.: Chapman and Hall/CRC.

Stephenson, P.J. *et al.* (2017). Unblocking the flow of biodiversity data for decision-making in Africa. *Biological Conservation* **213**:335–340.

Stork, N. e. *et al.* (2017). Consistency of effects of tropical-forest disturbance on species composition and richness relative to use of indicator taxa. *Conservation Biology* **31**:924–933.

Struebig, M.J. *et al.* (2012). Bat diversity in the lowland forests of the Heart of Borneo. *Biodiversity and Conservation* **21**:3711–3727.

Struebig, M.J. *et al.* (2010). Bats of Borneo: diversity, distributions and representation in protected areas. *Biodiversity and Conservation* **19**:449–469.

Struebig, M.J. *et al.* (2008). Conservation value of forest fragments to Palaeotropical bats. *Biological Conservation* **141**:2112–2126.

Struebig, M.J. *et al.* (2011). Parallel declines in species and genetic diversity in tropical forest fragments. *Ecol. Lett.* **14**:582–590.

Struebig, Matthew J *et al.* (2013). Quantifying the Biodiversity Value of Repeatedly Logged Rainforests. *Global Change in Multispecies Systems: Part 3*:183–224.

Struebig, Matthew J. *et al.* (2013). Quantifying the Biodiversity Value of Repeatedly Logged Rainforests: Gradient and Comparative Approaches from Borneo. In:

Woodward, G. and O’Gorman, E. J. eds. *Advances in Ecological Research*. Academic Press, pp. 183–224. Available at: https://www.sciencedirect.com/science/article/pii/B9780124171992000033 [Accessed: 7 October 2021].

Sueur, J. *et al.* (2014). Acoustic Indices for Biodiversity Assessment and Landscape Investigation. *Acta Acustica united with Acustica* **100**:772–781.

Sugai, L.S.M. *et al.* (2019). Terrestrial Passive Acoustic Monitoring: Review and Perspectives. *BioScience* **69**:15–25.

Suyanto, A. and Struebig, M.J. (2007). Bats of the Sangkulirang limestone karst formations, East Kalimantan — a priority region for Bornean bat conservation. *Acta Chiropterologica* **9**:67–95.

Swinfield, T., Both, S., *et al.* (2020). Imaging spectroscopy reveals the effects of topography and logging on the leaf chemistry of tropical forest canopy trees. *Glob. Chang. Biol.* **26**:989–1002.

Swinfield, T., Milodowski, D., *et al.* (2020). LiDAR canopy structure 2014 [Data set]. [Online]. Available at: https://zenodo.org/record/4020697 [Accessed: 10 May 2022].

Tabacchi, E. *et al.* (2000). Impacts of riparian vegetation on hydrological processes. *Hydrol. Process.* **14**:2959–2976.

Tabak, M.A. *et al.* (2021). Automated classification of bat echolocation call recordings with artificial intelligence. *bioRxiv*:2021.06.23.449619.

Tabak, M.A. *et al.* (2019). Machine learning to classify animal species in camera trap images: Applications in ecology. *Methods in Ecology and Evolution* **10**:585–590.

Tabeni, S., Spirito, F. and Ojeda, R.A. (2013). Conservation of small and medium-sized mammals following native woodland regrowth: A case study in a long-term UNESCO Biosphere Reserve, Argentina. *Journal of Arid Environments* **88**:250–253.

Tanshi, I. and Kingston, T. (2021). Introduction and Implementation of Harp Traps Signal a New Era in Bat Research. In: Lim, B. K. et al. eds. *50 Years of Bat Research: Foundations and New Frontiers*. Cham: Springer International Publishing, pp. 255–270. Available at: https://doi.org/10.1007/978-3-030-54727-1_16 [Accessed: 9 April 2021].

Taylor, P.J. *et al.* (2013). Diversity of bats in the Soutpansberg and Blouberg Mountains of northern South Africa: complementarity of acoustic and non-acoustic survey methods. *S. Afr. J. Wildl. Res.* **43**:12–26.

Teo, H.C. *et al.* (2020). Environmental Impacts of Planned Capitals and Lessons for Indonesia’s New Capital. *Land* **9**:438.

Terborgh, J. and Peres, C. (2002). The Problem of People in Parks. In: Terborgh, John et al. eds. *Making Parks Work: Strategies for Preserving Tropical Nature*. Island Press, pp. 307–319. Available at: https://play.google.com/store/books/details?id=6sQtZ7XRGUwC.

The HCS Approach Toolkit. (2017). Singapore: HCS Approach Steering Group.

TheCornellLab (2021). *Identify the Birds You See or Hear with Merlin Bird ID* [Online]. Available at: https://merlin.allaboutbirds.org/ [Accessed: 31 March 2022].

Thorn, S. *et al.* (2018). Impacts of salvage logging on biodiversity: A meta-analysis. *Journal of Applied Ecology* **55**:279–289.

Tilman, D. *et al.* (2017). Future threats to biodiversity and pathways to their prevention. *Nature* **546**:73–81.

Torrent, L. *et al.* (2018). The importance of lakes for bat conservation in Amazonian rainforests: an assessment using autonomous recorders. *Remote Sensing in Ecology and Conservation* **4**:339–351.

Troudet, J. *et al.* (2017). Taxonomic bias in biodiversity data and societal preferences. *Scientific Reports* **7**:9132.

Tscharntke, T. *et al.* (2008). Landscape Constraints on Functional Diversity of Birds and Insects in Tropical Agroecosystems. *Ecology* **89**:944–951.

Tuen, A.A., Rahman, M.A. and Salleh, M.A. (2002). Preliminary survey of mammals at Crocker Range Park, Sabah, Malaysia. *ASEAN Review of Biodiversity and Environmental Conservation (ARBEC)* [Online]. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.507.6672&rep=rep1&type=pdf [Accessed: 4 October 2019].

Turner, E.C. and Foster, W.A. (2009). The impact of forest conversion to oil palm on arthropod abundance and biomass in Sabah, Malaysia. *Journal of Tropical Ecology* **25**:23–30.

UN Convention on Biological Diversity (2021). *FIRST DRAFT OF THE POST-2020 GLOBAL BIODIVERSITY FRAMEWORK*. [Online]. Available at: https://www.cbd.int/doc/c/abb5/591f/2e46096d3f0330b08ce87a45/wg2020-03-03-en.pdf.

UNESCO (2017). *Biosphere Reserves—Learning Sites for Sustainable Development* [Online]. Available at: http: //www.unesco.org/new/en/natural-sciences/environment/ecological-sciences/biosphere-reserves/ [Accessed: 8 October 2019].

UNESCO (2014). *Crocker Range* [Online]. Available at: http://www.unesco.org/new/en/natural-sciences/ environment/ecological-sciences/biosphere-reserves/asia-and-the-pacific/malaysia/crocker-range/ [Accessed: 16 December 2019].

UNESCO (2019). *Southeast Asian Biosphere Reserve Network* [Online]. Available at: http://www.unesco.org/new/ en/natural-sciences/environment/ecological-sciences/man-and-biosphere-programme/networks/seabrnet/ [Accessed: 8 October 2019].

United Nations (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*. [Online]. Available at: https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E.

Valletta, J.J. *et al.* (2017). Applications of machine learning in animal behaviour studies. *Animal Behaviour* **124**:203–220.

Vaughan, N., Jones, G. and Harris, S. (1996). Effects of sewage effluent on the activity of bats (Chiroptera: Vespertilionidae) foraging along rivers. *Biological Conservation* **78**:337–343.

Voigt, C.C. and Kingston, T. eds. (2016). *Bats in the Anthropocene: Conservation of Bats in a Changing World*. [Online]. Springer International Publishing. Available at: https://www.springer.com/gp/book/9783319252186 [Accessed: 5 September 2021].

de Vries, W.T. (2021). Urban Greening for New Capital Cities. A Meta Review. *Frontiers in Sustainable Cities* [Online] **3**. Available at: https://www.frontiersin.org/article/10.3389/frsc.2021.670807 [Accessed: 1 April 2022].

VunKhen, C. (2006). Impacts of forest conversion on biodiversity as indicated by moths. *Malay. Nat. J.* **57**:383–418.

Walters, C.L. *et al.* (2013). Challenges of Using Bioacoustics to Globally Monitor Bats. In: Adams, R. A. and Pedersen, S. C. eds. *Bat Evolution, Ecology, and Conservation*. New York, NY: Springer, pp. 479–499. Available at: https://doi.org/10.1007/978-1-4614-7397-8_23 [Accessed: 29 September 2020].

Wanger, T.C. *et al.* (2010). Effects of Land-Use Change on Community Composition of Tropical Amphibians and Reptiles in Sulawesi, Indonesia. *Conservation Biology* **24**:795–802.

Waters, D.A. and Jones, G. (1995). Echolocation call structure and intensity in five species of insectivorous bats. *J. Exp. Biol.* **198**:475–489.

Wearn, O.R. *et al.* (2017). Mammalian species abundance across a gradient of tropical land-use intensity: A hierarchical multi-species modelling approach. *Biological Conservation* **212**:162–171.

Wearn, O.R., Freeman, R. and Jacoby, D.M.P. (2019). Responsible AI for conservation. *Nature Machine Intelligence* **1**:72–73.

Weatherbase (2019). *Tambunan, Malaysia Travel Weather Averages (Weatherbase)* [Online]. Available at: http://www.weatherbase.com/weather/weather.php3?s=131694&cityname=Tambunan-Sabah-Malaysia&units=metric [Accessed: 3 October 2019].

Wei, L. *et al.* (2008). Wing morphology, echolocation calls, diet and emergence time of black-bearded tomb bats (Taphozous melanopogon, Emballonuridae) from southwest China. *Acta Chiropterologica* **10**:51–59.

Weier, S.M. *et al.* (2017). Changes of bat species composition over altitudinal gradients on northern and southern aspects of the Soutpansberg mountain range, South Africa. *Mammalia* **81**:49–60.

Welch, J.N. and Beaulieu, J.M. (2018). Predicting Extinction Risk for Data Deficient Bats. *Diversity* **10**:63.

Whisson, D.A. *et al.* (2021). Passive acoustic monitoring for detecting the Yellow-bellied Glider, a highly vocal arboreal marsupial. *PLOS ONE* **16**:e0252092.

Wildlife Acoustics Inc (2019). Kaleidoscope Viewer. [Online]. Available at: http://condor.wildlifeacoustics.com/Kaleidoscope.pdf.

Williams-Guillén, K. *et al.* (2016). Bats in the Anthropogenic Matrix: Challenges and Opportunities for the Conservation of Chiroptera and Their Ecosystem Services in Agricultural Landscapes. In: Voigt, C. C. and Kingston, T. eds. *Bats in the Anthropocene: Conservation of Bats in a Changing World*. Cham: Springer International Publishing, pp. 151–186. Available at: https://doi.org/10.1007/978-3-319-25220-9_6 [Accessed: 29 March 2022].

Williams-Guillén, K. and Perfecto, I. (2011). Ensemble composition and activity levels of insectivorous bats in response to management intensification in coffee agroforestry systems. *PloS One* **6**:e16502.

Wilson, D.E. and Mittermeier, R.A. eds. (2013). Bats. In: *Handbook of the Mammals of the World*. Barcelona, Spain: Lynx Edicions, pp. 1–1008. Available at: https://www.lynxeds.com/product/handbook-of-the-mammals-of-the-world-volume-9/ [Accessed: 5 September 2021].

Wilson, M.F.J. *et al.* (2007). Multiscale Terrain Analysis of Multibeam Bathymetry Data for Habitat Mapping on the Continental Slope. *Mar. Geod.* **30**:3–35.

Woldegeorgis, G. and Wube, T. (2012). A SURVEY ON MAMMALS OF THE YAYU FOREST IN SOUTHWEST ETHIOPIA. *SINET: Ethiopian Journal of Science* **35**:135–138.

Woodcock, P. *et al.* (2011). The conservation value of South East Asia’s highly degraded forests: evidence from leaf-litter ants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**:3256–3264.

Wordley, C.F.R. *et al.* (2017). Bats in the Ghats: Agricultural intensification reduces functional diversity and increases trait filtering in a biodiversity hotspot in India. *Biological Conservation* **210**:48–55.

Wordley, C.F.R. *et al.* (2018). Heard but not seen: Comparing bat assemblages and study methods in a mosaic landscape in the Western Ghats of India. *Ecology and Evolution* **8**:3883–3894.

Wrege, P.H. *et al.* (2017). Acoustic monitoring for conservation in tropical forests: examples from forest elephants. *Methods in Ecology and Evolution* **8**:1292–1301.

Wright, S.J. (2010). The future of tropical forests. *Annals of the New York Academy of Sciences* **1195**:1–27.

Wu, X. *et al.* (2022). A survey of human-in-the-loop for machine learning. *Future Generation Computer Systems* **135**:364–381.

Yasuma, S. *et al.* (2003). *Identification Keys to the Mammals of Borneo*. Park Management Component BBEC Programme. Kota Kinabalu, Malaysia: Sabah Parks and JICA.

Ye, X., Skidmore, A.K. and Wang, T. (2013). Within-patch habitat quality determines the resilience of specialist species in fragmented landscapes. *Landscape Ecology* **28**:135–147.

Yoh, N. *et al.* (2022). A machine learning framework to classify Southeast Asian echolocating bats. *Ecological Indicators* **136**:108696.

Yoh, N., Azhar, I., *et al.* (2020). Bat Ensembles Differ in Response to Use Zones in a Tropical Biosphere Reserve. *Diversity* **12**:60.

Yoh, N., Syme, P., *et al.* (2020). Echolocation of Central Amazonian ‘whispering’ phyllostomid bats: call design and interspecific variation. *Mammal Research* **65**:583–597.

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* **7**:1082–1091.

Zimbres, B., Peres, C.A. and Machado, R.B. (2017). Terrestrial mammal responses to habitat structure and quality of remnant riparian forests in an Amazonian cattle-ranching landscape. *Biological Conservation* **206**:283–292.

Zubaid, A. (1993). A comparison of the bat fauna between a primary and fragmented secondary forest in Peninsular Malaysia. *Mammalia* **57**:201–206.

Appendix I Co-authored publications

Peer-reviewed journal articles that were contributed to, that supplement the research conducted within this thesis. Here, I provide the titles and abstracts in reverse chronological order, full text copies are available online.

Research article: Edge effects and vertical stratification of aerial insectivorous bats across the interface of primary-secondary Amazonian rainforest

PLOS One: In press (September 2022)

Natalie Yoh, James Clarke, Adrià López-Baucells, Maria Mas, Ricardo Rocha, Paulo E.D. Bobrowiec, Christoph F.J. Meyer

Abstract

Edge effects - abiotic and biotic changes associated with habitat boundaries - are key drivers of community change in fragmented landscapes. Their influence is heavily modulated by matrix composition. With over half of the world's tropical forests predicted to become forest edge by the end of the century, it is paramount that conservationists gain a better understanding of how tropical biota is impacted by edge gradients. Bats comprise a large fraction of tropical mammalian fauna and are demonstrably sensitive to habitat modification. Yet, knowledge about how bat assemblages are affected by edge effects remains scarce. Capitalizing on a whole-ecosystem manipulation in the Central Amazon, the aims of this study were to i) assess the consequences of edge effects for twelve aerial insectivorous bat species across the interface of primary and secondary forest and ii) investigate if the activity levels of these species differed between the understory and canopy and if they were modulated by distance from the edge. Acoustic surveys were conducted along four 2-km transects each traversing equal parts of primary and ca. 30-year-old secondary forest. Five models were used to assess the changes in the relative activity of forest specialists (three species), flexible forest foragers (three species), and edge foragers (six species). Modelling results revealed no evidence of edge effects, except for forest specialists in the understory. No significant differences in activity were found between the secondary or primary forest but most species exhibited pronounced vertical stratification. Our study highlights that forest specialist bats are more edge-sensitive than both flexible forest and edge foraging bats and suggests that the influence of edge effects on aerial insectivorous bats may exceed 2 km. The absence of pronounced edge effects and the comparable activity levels between primary and old secondary forests indicates that old secondary forest can help ameliorate the consequences of fragmentation on tropical aerial insectivorous bats.

Research article: Could Nintendo's Animal Crossing be a tool for conservation messaging?

People and Nature: 2021, DOI: 10.1002/pan3.10240

Jessica C. Fisher*, **Natalie Yoh*,** Takahiro Kubo, Danielle Rundle
(*authors contributed equally)

Abstract

1. The current extinction crisis demands worldwide commitment to conservation across all sectors of society. By transcending the traditional disciplinary boundaries, conservationists can reach new audiences to communicate pro-conservation knowledge, education and awareness messages.

2. There are approximately 2.7 billion video-gamers worldwide, with millions more joining as a result of global lockdowns. In March 2020, *Animal Crossing: New Horizons* was released by Nintendo, fast becoming the second-best selling video game ever in Japan, and selling over 26.4 million units worldwide. Unlike many popular video games, its unique premise involves players creating an island, growing vegetation, catching wildlife and donating fossils and species to a museum. The game has been praised for its positivity, escapism and measurable benefits to mental well-being.

3. Here, we articulate how different features of the game, including the islands, their biodiversity and inhabitants, encourage players to exhibit pro-conservation behaviours and attitudes (e.g. recycling litter, or planting a diversity of flowers), as well as improving players' knowledge about the diversity of relatively little known taxa (marine and freshwater fishes and invertebrates). We also highlight where pitfalls exist (e.g. encouraging the collection of threatened species). We principally frame these discussions in the context of Japan's cultural relationship with the natural world, including its history of insect-collecting and its management of green spaces. We conclude by outlining some recommendations about potential improvements to future releases, or for similar games, that could further promote conservation messaging.

4. This perspective sheds light on the avenues through which *Animal Crossing: New Horizons* encourages pro-conservation knowledge, attitudes and behaviours of its international audience, with potential for these experiences to translate into real-world conservation actions. During a critical time in humanity's history, video gaming could therefore provide a huge opportunity for communicating conservation messages to billions of people worldwide.

Research article: Optimising bat bioacoustic surveys in human-modified neotropical landscapes

Ecological Applications: 2021, Volume 36(6), e02366. DOI: 10.1002/eap.2366

Adrià López-Baucells, **Natalie Yoh**, Ricardo Rocha, Paulo E. D. Bobrowiec, Jorge M. Palmeirim, Christoph F. J. Meyer

Abstract

Edge effects, abiotic and biotic changes associated with habitat boundaries, are key drivers of community change in fragmented landscapes. Their influence is heavily modulated by matrix composition. With over half of the world's tropical forests predicted to become forest edge by the end of the century, it is paramount that conservationists gain a better understanding of how tropical biota is impacted by edge gradients. Bats comprise a large fraction of tropical mammalian fauna and are demonstrably sensitive to habitat modification. Yet, knowledge about how bat assemblages are affected by edge effects remains scarce. Capitalizing on a whole-ecosystem manipulation in the Central Amazon, the aims of this study were to i) assess the consequences of edge effects for twelve aerial insectivorous bat species across the interface of primary and secondary forest, and ii) investigate if the activity levels of these species differed between the understory and canopy and if they were modulated by distance from the edge. Acoustic surveys were conducted along four 2-km transects each traversing equal parts of primary and ca. 30-year-old secondary forest. Five models were used to assess the changes in the relative activity of forest specialists (three species), flexible forest foragers (three species), and edge foragers (six species). Modelling results revealed limited evidence of edge effects, except for forest specialists in the understory. No significant differences in activity were found between the secondary or primary forest but almost all species exhibited pronounced vertical stratification. Previously defined bat guilds appear to hold here as our study highlights that forest bats are more edge-sensitive than edge foraging bats. The absence of pronounced edge effects and the comparable activity levels between primary and old secondary forests indicates that old secondary forest can help ameliorate the consequences of fragmentation on tropical aerial insectivorous bats.

Research article: Echolocation of Central Amazonian ‘whispering’ phyllostomid bats: call design and interspecific variation (2020)

Mammal Research: 2020, Volume 65, 583–597. DOI: 10.1007/s13364-020-00503-0

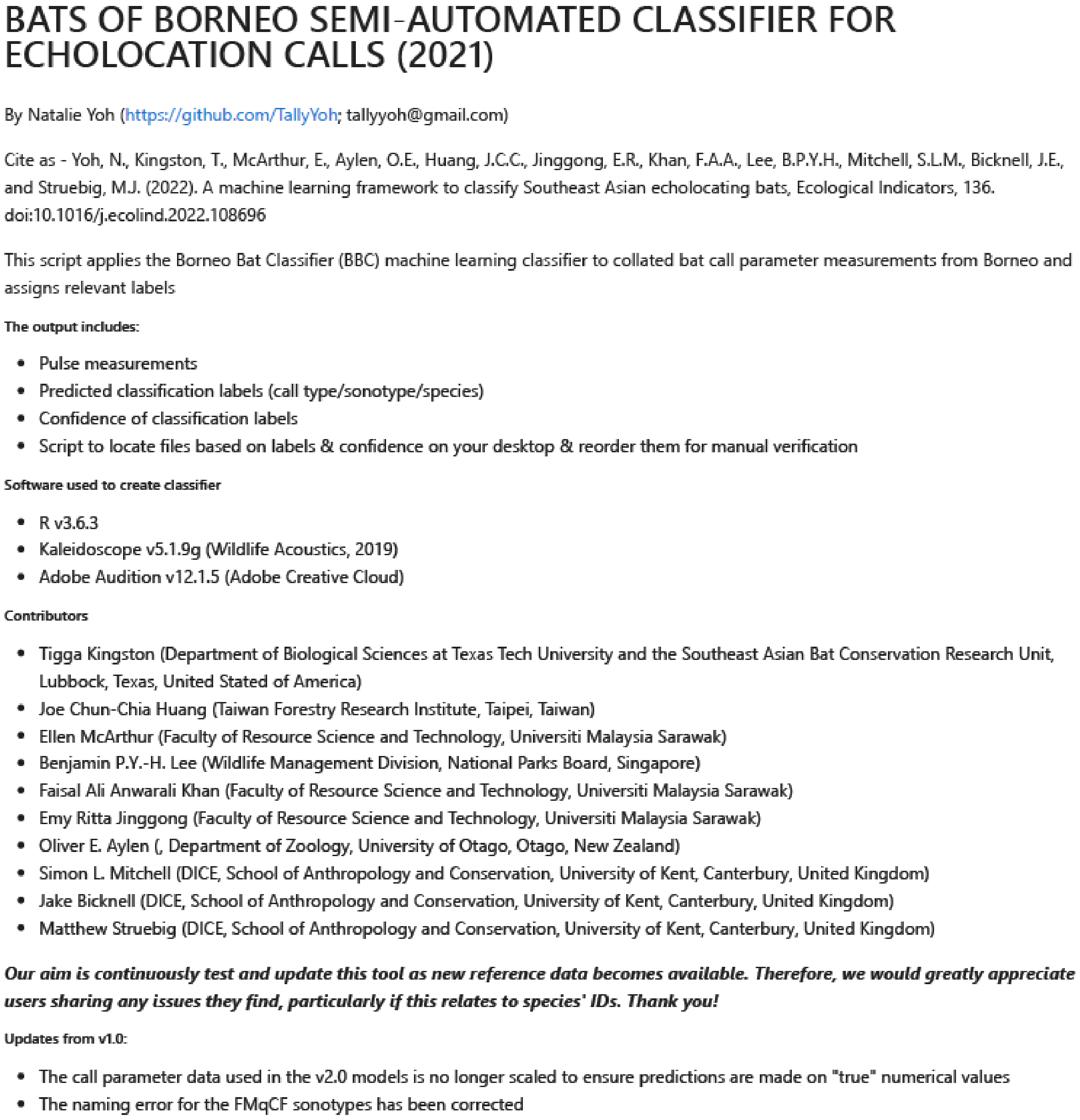
Natalie Yoh, Peter Syme, Ricardo Rocha, Christoph F. J. Meyer, Adrià López-Baucells

Abstract

Phyllostomids (New World leaf-nosed bats) are the ecologically most diverse bat family and have undergone the most extensive adaptive radiation of any mammalian family. However comprehensive, multi-species studies regarding phyllostomid echolocation are scarce in the literature despite abundant ecological research. In this study, we describe the call structure and interspecific variation in call design of 40 sympatric phyllostomid species from the Central Brazilian Amazon, focussing on general patterns within genera, subfamilies and between feeding guilds. All but one species utilized short, broadband FM calls consisting of multiple harmonics. As reported for other bat families, peak frequency was negatively correlated with body mass and forearm length. Twenty-five species alternated the harmonic of maximum energy, principally between the second and third harmonic. Based on PCA, we were unable to detect any significant differences in echolocation call parameters between genera, subfamilies or different feeding guilds, confirming that acoustic surveys cannot be used to reliably monitor these species. We present *Ametrida centurio* as an exception to this generalized phyllostomid structure, as it is unique in producing a mono-harmonic call. Finally, we discuss several hypotheses regarding the evolutionary pressures influencing phyllostomid call structure.

Appendix II Borneo Bat Classifier user script

The script for processing bat calls using the Borneo Bat Classifier (V2.0), available at https://github.com/TallyYoh/BorneoBatCalls



BATS OF BORNEO SEMI-AUTOMATED CLASSIFIER FOR ECHOLOCATION CALLS (2021)

By Natalie Yoh (https://github.com/TallyYoh; tallyyoh@gmail.com)

Cite as - Yoh, N., Kingston, T., McArthur, E., Aylen, O.E., Huang, J.C.C., Jinggong, E.R., Khan, F.A.A., Lee, B.P.Y.H., Mitchell, S.L.M., Bicknell, J.E., and Struebig, M.J. (2022). A machine learning framework to classify Southeast Asian echolocating bats, Ecological Indicators, 136. doi:10.1016/j.ecolind.2022.108696

This script applies the Borneo Bat Classifier (BBC) machine learning classifier to collated bat call parameter measurements from Borneo and assigns relevant labels

The output includes:

- Pulse measurements
- Predicted classification labels (call type/sonotype/species)
- Confidence of classification labels
- Script to locate files based on labels & confidence on your desktop & reorder them for manual verification

Software used to create classifier

- R v3.6.3
- Kaleidoscope v5.1.9g (Wildlife Acoustics, 2019)
- Adobe Audition v12.1.5 (Adobe Creative Cloud)

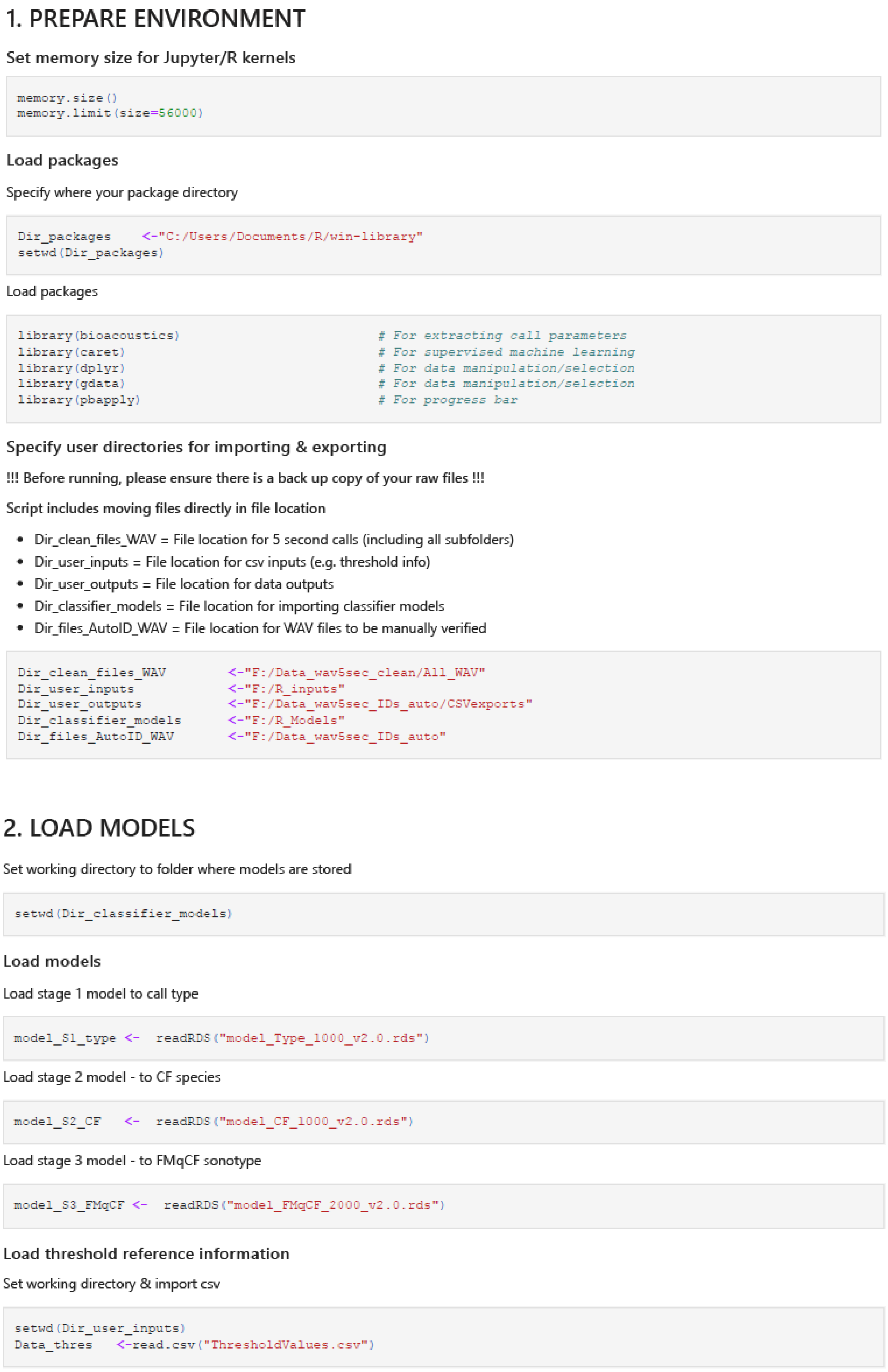
Contributors

- Tigga Kingston (Department of Biological Sciences at Texas Tech University and the Southeast Asian Bat Conservation Research Unit, Lubbock, Texas, United Stated of America)
- Joe Chun-Chia Huang (Taiwan Forestry Research Institute, Taipei, Taiwan)
- Ellen McArthur (Faculty of Resource Science and Technology, Universiti Malaysia Sarawak)
- Benjamin P.Y.-H. Lee (Wildlife Management Division, National Parks Board, Singapore)
- Faisal Ali Anwarali Khan (Faculty of Resource Science and Technology, Universiti Malaysia Sarawak)
- Emy Ritta Jinggong (Faculty of Resource Science and Technology, Universiti Malaysia Sarawak)
- Oliver E. Aylen (, Department of Zoology, University of Otago, Otago, New Zealand)
- Simon L. Mitchell (DICE, School of Anthropology and Conservation, University of Kent, Canterbury, United Kingdom)
- Jake Bicknell (DICE, School of Anthropology and Conservation, University of Kent, Canterbury, United Kingdom)
- Matthew Struebig (DICE, School of Anthropology and Conservation, University of Kent, Canterbury, United Kingdom)

Our aim is continuously test and update this tool as new reference data becomes available. Therefore, we would greatly appreciate users sharing any issues they find, particularly if this relates to species' IDs. Thank you!

Updates from v1.0:

- The call parameter data used in the v2.0 models is no longer scaled to ensure predictions are made on "true" numerical values
- The naming error for the FMqCF sonotypes has been corrected



1. PREPARE ENVIRONMENT

Set memory size for Jupyter/R kernels

```
memory.size()
memory.limit(size=56000)
```

Load packages

Specify where your package directory

```
Dir_packages    <-"C:/Users/Documents/R/win-library"
setwd(Dir_packages)
```

Load packages

```
library(bioacoustics)                         # For extracting call parameters
library(caret)                                # For supervised machine learning
library(dplyr)                                # For data manipulation/selection
library(gdata)                                # For data manipulation/selection
library(pbapply)                              # For progress bar
```

Specify user directories for importing & exporting

!!! Before running, please ensure there is a back up copy of your raw files !!!

Script includes moving files directly in file location

- Dir_clean_files_WAV = File location for 5 second calls (including all subfolders)
- Dir_user_inputs = File location for csv inputs (e.g. threshold info)
- Dir_user_outputs = File location for data outputs
- Dir_classifier_models = File location for importing classifier models
- Dir_files_AutoID_WAV = File location for WAV files to be manually verified

```
Dir_clean_files_WAV         <-"F:/Data_wav5sec_clean/All_WAV"
Dir_user_inputs             <-"F:/R_inputs"
Dir_user_outputs            <-"F:/Data_wav5sec_IDs_auto/CSVexports"
Dir_classifier_models       <-"F:/R_Models"
Dir_files_AutoID_WAV        <-"F:/Data_wav5sec_IDs_auto"
```

2. LOAD MODELS

Set working directory to folder where models are stored

```
setwd(Dir_classifier_models)
```

Load models

Load stage 1 model to call type

```
model_S1_type <-  readRDS("model_Type_1000_v2.0.rds")
```

Load stage 2 model - to CF species

```
model_S2_CF   <-  readRDS("model_CF_1000_v2.0.rds")
```

Load stage 3 model - to FMqCF sonotype

```
model_S3_FMqCF <-  readRDS("model_FMqCF_2000_v2.0.rds")
```

Load threshold reference information

Set working directory & import csv

```
setwd(Dir_user_inputs)
Data_thres    <-read.csv("ThresholdValues.csv")
```

3. IMPORT & EXTRACT CALL PARAMETERS

Extracts call parameters from WAV files for classification using the Bioacoustic.R package

WAV files for import should first have been subset to 5 second fragments to quantify a bat pass & be filtered for noise in Kaliedoscope or other sound analysis software. See Yoh et al. (2021) for more information

Select file directories for where files are stored. This will perform extractions in two batches

```
files_P1        <- dir(Dir_clean_files_WAV, recursive = TRUE, full.names = TRUE, pattern = "[.]wav$")
```

Filter files for those identified as noise in Kaliedoscope

```
# convert to dataframe
files_P1        <-as.data.frame(files_P1)

# remove files listed as "noise"
files_P1_crop   <-as.character(files_P1[!grepl("NOISE", files_P1$files_P1),])
```

Detect & extract pulse measurements

Extractions conducted using the Bioacoustics.R package threshold function (https://rdrr.io/cran/bioacoustics/) Extractions can be performed for time expansion 1 or 10 as necessary (use "time_exp = 10" if necessary)

```
TDP1   <- setNames(
           pblapply(
                 files_P1_crop,
                 threshold_detection,
                 time_exp = 1,
                 threshold = 4,
                 SNR_thr = 4,
                 FFT_size = 512,), basename(files_P1_crop))
```

Collate measurements

Remove filenames where no values were extracted (e.g. only noise)

```
TDP1    <- TDP1[lapply(TDP1, function(x) length(x$data)) > 0]
```

Keep the extracted features and merge in a single data frame for further analysis

```
Data_WAV_raw  <- do.call("rbind", c(lapply(TDP1, function(x) x$data$event_data),
                                           list(stringsAsFactors = FALSE)))
```

Remove file extention from filenames

```
Data_WAV_raw$filename <-sub(pattern = "(.*)\\..*$", replacement = "\\1", basename(Data_WAV_raw$filename))
head(Data_WAV_raw)
```

Include filename location information

Extract filename from file locations

```
filename      <-sub(pattern = "(.*)\\..*$", replacement = "\\1", basename(files_P1_crop))
```

Create dataframe with full file location & filename

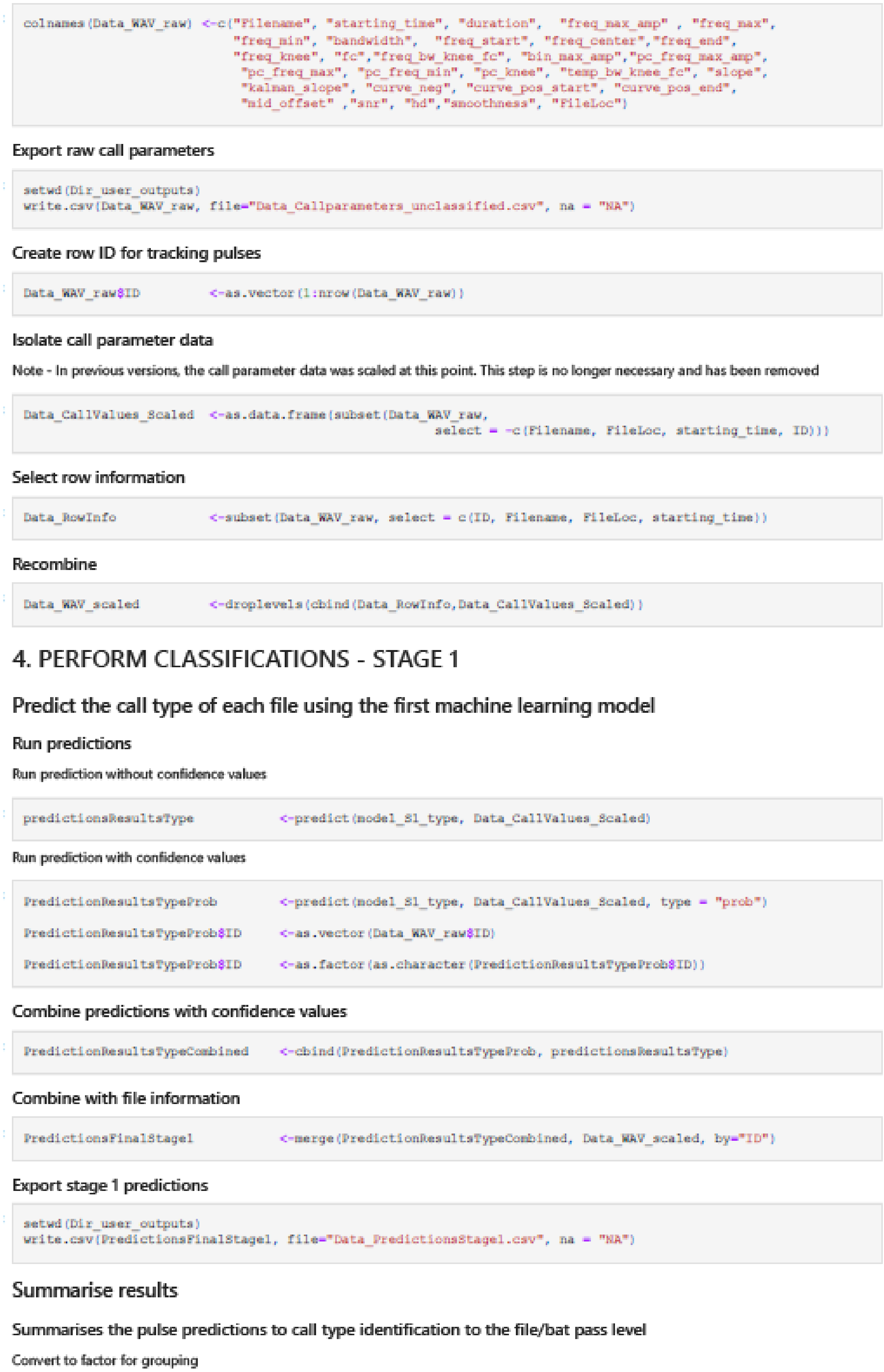
```
FileLoc       <-data.frame(FileLoc=totalfileloc, filename=filename)
```

Add to main dataframe

```
Data_WAV_raw <-merge(Data_WAV_raw, FileLoc, by="filename")
```

Clean & export pulse measurements

Rename columns - Include/remove additionals where applicable



```
colnames(Data_WAV_raw) <-c("Filename", "starting_time", "duration",  "freq_max_amp" , "freq_max",
                           "freq_min", "bandwidth",  "freq_start", "freq_center","freq_end",
                           "freq_knee", "fc","freq_bw_knee_fc", "bin_max_amp","pc_freq_max_amp",
                            "pc_freq_max", "pc_freq_min", "pc_knee", "temp_bw_knee_fc", "slope",
                            "kalman_slope", "curve_neg", "curve_pos_start", "curve_pos_end",
                            "mid_offset" ,"snr", "hd","smoothness", "FileLoc")
```

Export raw call parameters

```
setwd(Dir_user_outputs)
write.csv(Data_WAV_raw, file="Data_Callparameters_unclassified.csv", na = "NA")
```

Create row ID for tracking pulses

```
Data_WAV_raw$ID         <-as.vector(1:nrow(Data_WAV_raw))
```

Isolate call parameter data

Note - In previous versions, the call parameter data was scaled at this point. This step is no longer necessary and has been removed

```
Data_CallValues_Scaled  <-as.data.frame(subset(Data_WAV_raw,
                                                     select = -c(Filename, FileLoc, starting_time, ID)))
```

Select row information

```
Data_RowInfo            <-subset(Data_WAV_raw, select = c(ID, Filename, FileLoc, starting_time))
```

Recombine

```
Data_WAV_scaled         <-droplevels(cbind(Data_RowInfo,Data_CallValues_Scaled))
```

4. PERFORM CLASSIFICATIONS - STAGE 1

Predict the call type of each file using the first machine learning model

Run predictions

Run prediction without confidence values

```
predictionsResultsType           <-predict(model_S1_type, Data_CallValues_Scaled)
```

Run prediction with confidence values

```
PredictionResultsTypeProb        <-predict(model_S1_type, Data_CallValues_Scaled, type = "prob")

PredictionResultsTypeProb$ID     <-as.vector(Data_WAV_raw$ID)

PredictionResultsTypeProb$ID     <-as.factor(as.character(PredictionResultsTypeProb$ID))
```

Combine predictions with confidence values

```
PredictionResultsTypeCombined    <-cbind(PredictionResultsTypeProb, predictionsResultsType)
```

Combine with file information

```
PredictionsFinalStage1           <-merge(PredictionResultsTypeCombined, Data_WAV_scaled, by="ID")
```

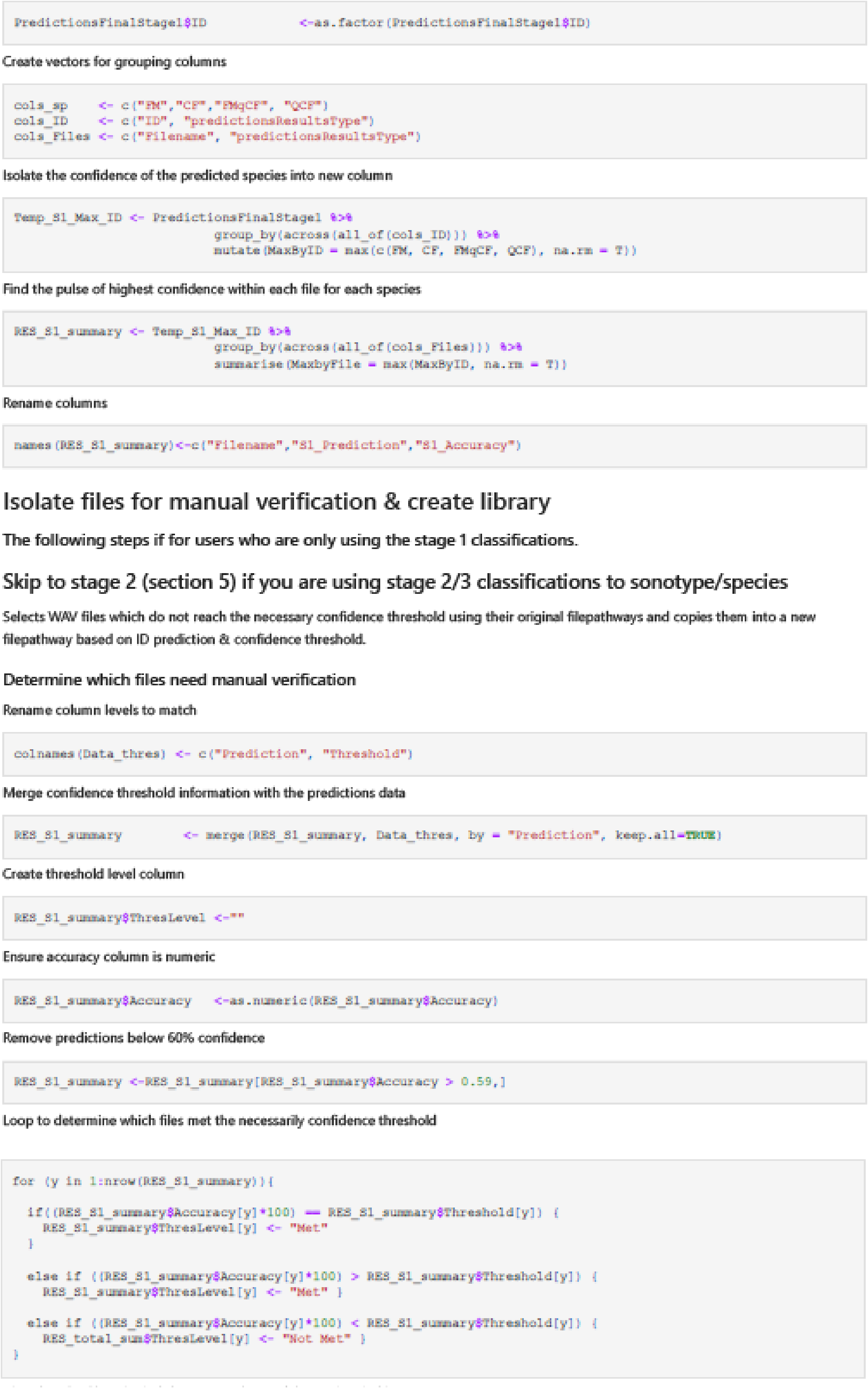
Export stage 1 predictions

```
setwd(Dir_user_outputs)
write.csv(PredictionsFinalStage1, file="Data_PredictionsStage1.csv", na = "NA")
```

Summarise results

Summarises the pulse predictions to call type identification to the file/bat pass level

Convert to factor for grouping



```
PredictionsFinalStage1$ID            <-as.factor(PredictionsFinalStage1$ID)
```

Create vectors for grouping columns

```
cols_sp    <- c("FM","CF","FMqCF", "QCF")
cols_ID    <- c("ID", "predictionsResultsType")
cols_Files <- c("Filename", "predictionsResultsType")
```

Isolate the confidence of the predicted species into new column

```
Temp_S1_Max_ID <- PredictionsFinalStage1 %>%
                          group_by(across(all_of(cols_ID))) %>%
                          mutate(MaxByID = max(c(FM, CF, FMqCF, QCF), na.rm = T))
```

Find the pulse of highest confidence within each file for each species

```
RES_S1_summary <- Temp_S1_Max_ID %>%
                          group_by(across(all_of(cols_Files))) %>%
                          summarise(MaxbyFile = max(MaxByID, na.rm = T))
```

Rename columns

```
names(RES_S1_summary)<-c("Filename","S1_Prediction","S1_Accuracy")
```

Isolate files for manual verification & create library

The following steps if for users who are only using the stage 1 classifications.

Skip to stage 2 (section 5) if you are using stage 2/3 classifications to sonotype/species

Selects WAV files which do not reach the necessary confidence threshold using their original filepathways and copies them into a new filepathway based on ID prediction & confidence threshold.

Determine which files need manual verification

Rename column levels to match

```
colnames(Data_thres) <- c("Prediction", "Threshold")
```

Merge confidence threshold information with the predictions data

```
RES_S1_summary        <- merge(RES_S1_summary, Data_thres, by = "Prediction", keep.all=TRUE)
```

Create threshold level column

```
RES_S1_summary$ThresLevel <-""
```

Ensure accuracy column is numeric

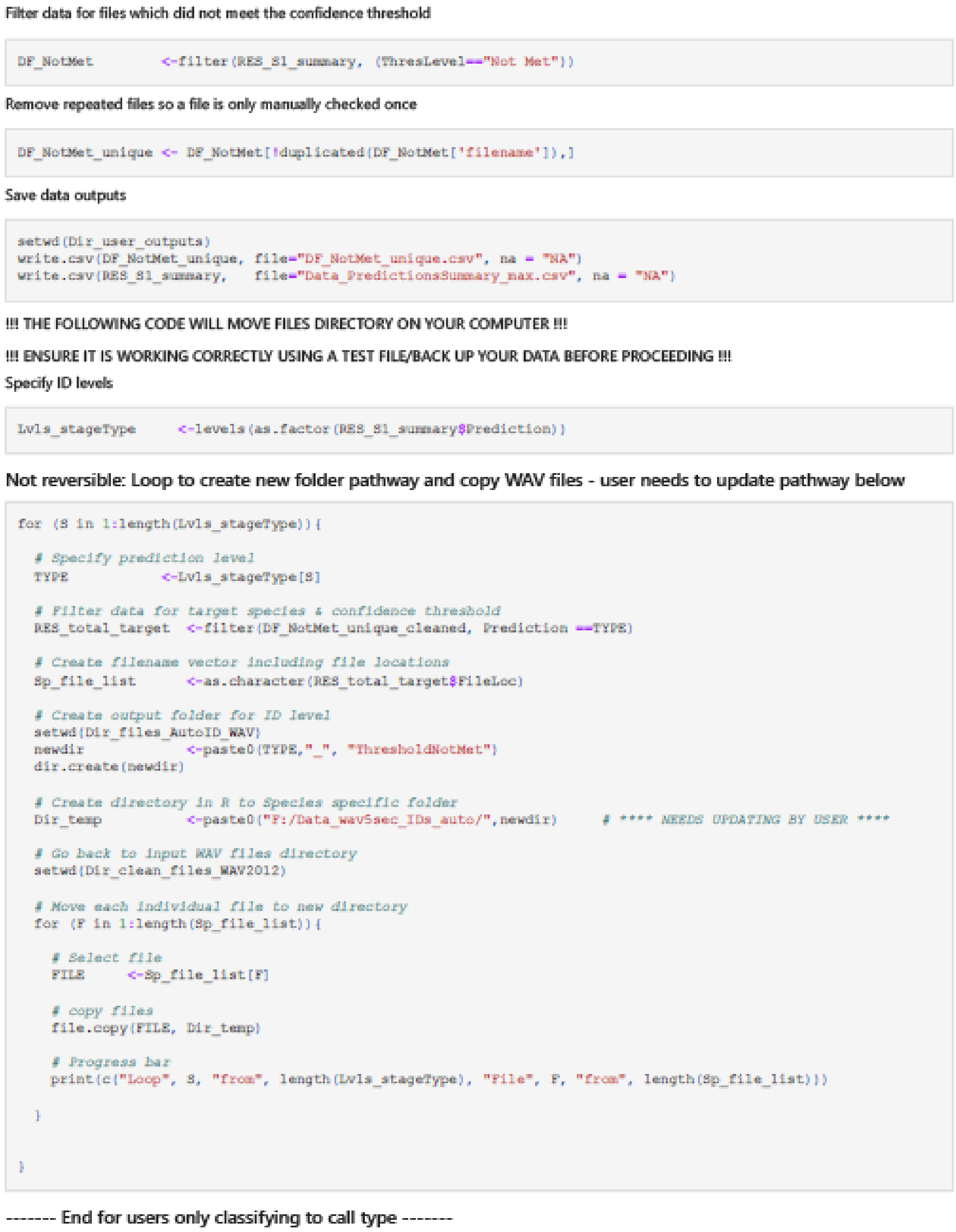
```
RES_S1_summary$Accuracy   <-as.numeric(RES_S1_summary$Accuracy)
```

Remove predictions below 60% confidence

```
RES_S1_summary <-RES_S1_summary[RES_S1_summary$Accuracy > 0.59,]
```

Loop to determine which files met the necessarily confidence threshold

```
for (y in 1:nrow(RES_S1_summary)){

  if((RES_S1_summary$Accuracy[y]*100) == RES_S1_summary$Threshold[y]) {
    RES_S1_summary$ThresLevel[y] <- "Met"
  }

  else if ((RES_S1_summary$Accuracy[y]*100) > RES_S1_summary$Threshold[y]) {
    RES_S1_summary$ThresLevel[y] <- "Met" }

  else if ((RES_S1_summary$Accuracy[y]*100) < RES_S1_summary$Threshold[y]) {
    RES_total_sum$ThresLevel[y] <- "Not Met" }
}
```



Filter data for files which did not meet the confidence threshold

```
DF_NotMet        <-filter(RES_S1_summary, (ThresLevel=="Not Met"))
```

Remove repeated files so a file is only manually checked once

```
DF_NotMet_unique <- DF_NotMet[!duplicated(DF_NotMet['filename']),]
```

Save data outputs

```
setwd(Dir_user_outputs)
write.csv(DF_NotMet_unique, file="DF_NotMet_unique.csv", na = "NA")
write.csv(RES_S1_summary,   file="Data_PredictionsSummary_max.csv", na = "NA")
```

!!! THE FOLLOWING CODE WILL MOVE FILES DIRECTORY ON YOUR COMPUTER !!!

!!! ENSURE IT IS WORKING CORRECTLY USING A TEST FILE/BACK UP YOUR DATA BEFORE PROCEEDING !!!

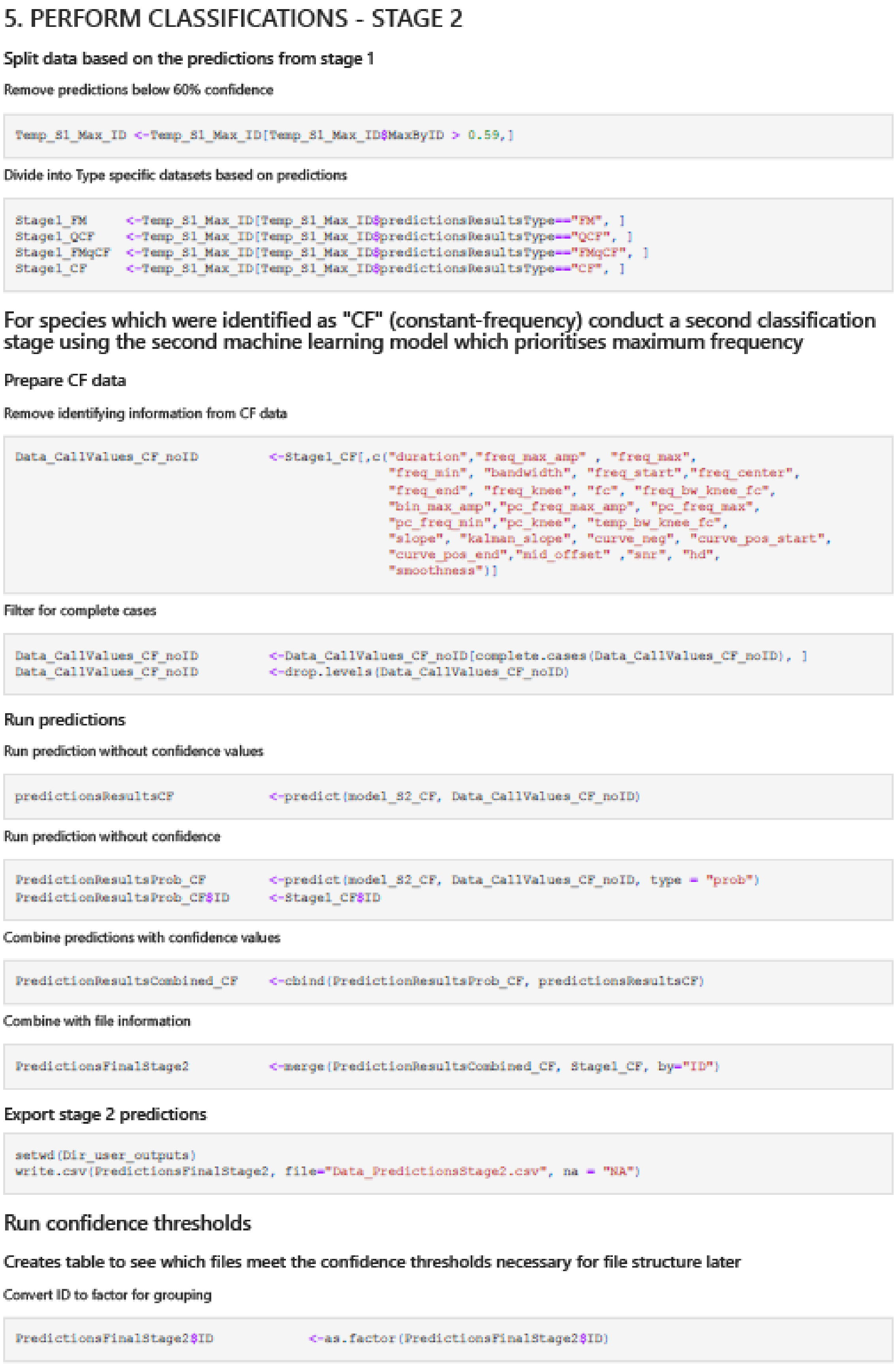
Specify ID levels

```
Lvls_stageType     <-levels(as.factor(RES_S1_summary$Prediction))
```

Not reversible: Loop to create new folder pathway and copy WAV files - user needs to update pathway below

```
for (S in 1:length(Lvls_stageType)){

  # Specify prediction level
  TYPE           <-Lvls_stageType[S]

  # Filter data for target species & confidence threshold
  RES_total_target  <-filter(DF_NotMet_unique_cleaned, Prediction ==TYPE)

  # Create filename vector including file locations
  Sp_file_list      <-as.character(RES_total_target$FileLoc)

  # Create output folder for ID level
  setwd(Dir_files_AutoID_WAV)
  newdir            <-paste0(TYPE,"_", "ThresholdNotMet")
  dir.create(newdir)

  # Create directory in R to Species specific folder
  Dir_temp          <-paste0("F:/Data_wav5sec_IDs_auto/",newdir)     # **** NEEDS UPDATING BY USER ****

  # Go back to input WAV files directory
  setwd(Dir_clean_files_WAV2012)

  # Move each individual file to new directory
  for (F in 1:length(Sp_file_list)){

    # Select file
    FILE     <-Sp_file_list[F]

    # copy files
    file.copy(FILE, Dir_temp)

    # Progress bar
    print(c("Loop", S, "from", length(Lvls_stageType), "File", F, "from", length(Sp_file_list)))

  }


}
```

------- End for users only classifying to call type -------



5. PERFORM CLASSIFICATIONS - STAGE 2

Split data based on the predictions from stage 1

Remove predictions below 60% confidence

```
Temp_S1_Max_ID <-Temp_S1_Max_ID[Temp_S1_Max_ID$MaxByID > 0.59,]
```

Divide into Type specific datasets based on predictions

```
Stage1_FM     <-Temp_S1_Max_ID[Temp_S1_Max_ID$predictionsResultsType=="FM", ]
Stage1_QCF    <-Temp_S1_Max_ID[Temp_S1_Max_ID$predictionsResultsType=="QCF", ]
Stage1_FMqCF  <-Temp_S1_Max_ID[Temp_S1_Max_ID$predictionsResultsType=="FMqCF", ]
Stage1_CF     <-Temp_S1_Max_ID[Temp_S1_Max_ID$predictionsResultsType=="CF", ]
```

For species which were identified as "CF" (constant-frequency) conduct a second classification stage using the second machine learning model which prioritises maximum frequency

Prepare CF data

Remove identifying information from CF data

```
Data_CallValues_CF_noID         <-Stage1_CF[,c("duration","freq_max_amp" , "freq_max",
                                               "freq_min", "bandwidth", "freq_start","freq_center",
                                               "freq_end", "freq_knee", "fc", "freq_bw_knee_fc",
                                               "bin_max_amp","pc_freq_max_amp", "pc_freq_max",
                                               "pc_freq_min","pc_knee", "temp_bw_knee_fc",
                                               "slope", "kalman_slope", "curve_neg", "curve_pos_start",
                                               "curve_pos_end","mid_offset" ,"snr", "hd",
                                               "smoothness")]
```

Filter for complete cases

```
Data_CallValues_CF_noID         <-Data_CallValues_CF_noID[complete.cases(Data_CallValues_CF_noID), ]
Data_CallValues_CF_noID         <-drop.levels(Data_CallValues_CF_noID)
```

Run predictions

Run prediction without confidence values

```
predictionsResultsCF            <-predict(model_S2_CF, Data_CallValues_CF_noID)
```

Run prediction without confidence

```
PredictionResultsProb_CF        <-predict(model_S2_CF, Data_CallValues_CF_noID, type = "prob")
PredictionResultsProb_CF$ID     <-Stage1_CF$ID
```

Combine predictions with confidence values

```
PredictionResultsCombined_CF    <-cbind(PredictionResultsProb_CF, predictionsResultsCF)
```

Combine with file information

```
PredictionsFinalStage2          <-merge(PredictionResultsCombined_CF, Stage1_CF, by="ID")
```

Export stage 2 predictions

```
setwd(Dir_user_outputs)
write.csv(PredictionsFinalStage2, file="Data_PredictionsStage2.csv", na = "NA")
```

Run confidence thresholds

Creates table to see which files meet the confidence thresholds necessary for file structure later

Convert ID to factor for grouping

```
PredictionsFinalStage2$ID            <-as.factor(PredictionsFinalStage2$ID)
```

Convert ID to factor for grouping

```
PredictionsFinalStage2$ID            <-as.factor(PredictionsFinalStage2$ID)
```

Identify which CF are present in the data

```
Levels_CF                             <-levels(as.factor(PredictionsFinalStage2$predictionsResultsCF))
```

Create vector for grouping species (***user needs to edit depending on the species listed in Levels_CF***)

```
cols_sp    <- c( "CF_H140", "CF_Hate","CF_Hbic","CF_Hcer","CF_Hcox" ,
                 "CF_Hdia","CF_Hgal",  "CF_Hlar",  "CF_Hrid",  "CF_Racu" , "CF_Raff",
                 "CF_Rbor"  , "CF_Rcre","CF_Rluc" ,  "CF_Rphi", "CF_Rsed","CF_Rtri" )
```

Create vectors for grouping

```
cols_ID    <- c("ID", "predictionsResultsCF")
cols_Files <- c("Filename", "predictionsResultsCF")
```

Isolate the confidence of the predicted species into new column (***user needs to edit depending on the species listed in Levels_CF***)

```
Temp_S2_Max_ID <- PredictionsFinalStage2 %>%
                           group_by(across(all_of(cols_ID))) %>%
                           mutate(MaxByID = max(c(CF_H140, CF_Hate,CF_Hbic,CF_Hcer,CF_Hcox ,
                                                  CF_Hdia,CF_Hgal,  CF_Hlar,  CF_Hrid, CF_Racu ,
                                                  CF_Raff,CF_Rbor, CF_Rcre,CF_Rluc ,
                                                  CF_Rphi, CF_Rsed,CF_Rtri), na.rm = T))
```

Find the pulse of highest confidence within each file for each species

```
RES_S2_summary <- Temp_S2_Max_ID %>%
                           group_by(across(all_of(cols_Files))) %>%
                           summarise(MaxbyFile = max(MaxByID, na.rm = T))
```

Rename columns

```
names(RES_S2_summary)<-c("Filename","S2_Prediction","S2_Accuracy")
```

6. PERFORM CLASSIFICATIONS - STAGE 3

For species which were identified as "FMqCF" (frequency modulated quasi-constant frequency) conduct a second classification stage using the third machine learning model which prioritises call shape

Prepare FMqCF data

Remove identifying information from FMqCF data

```
Data_CallValues_FMqCF_noID         <-Stage1_FMqCF[,c("duration","freq_max_amp" , "freq_max",
                                                     "freq_min", "bandwidth","freq_start",
                                                     "freq_center","freq_end", "freq_knee", "fc",
                                                     "freq_bw_knee_fc", "bin_max_amp",  "pc_freq_max_amp",
                                                     "pc_freq_max",  "pc_freq_min", "pc_knee",
                                                     "temp_bw_knee_fc", "slope", "kalman_slope",
                                                     "curve_neg", "curve_pos_start", "curve_pos_end",
                                                     "mid_offset" ,"snr", "hd", "smoothness")]
```

Filter for complete cases

```
Data_CallValues_FMqCF_noID         <-Data_CallValues_FMqCF_noID[complete.cases(Data_CallValues_FMqCF_noID), ]
Data_CallValues_FMqCF_noID         <-drop.levels(Data_CallValues_FMqCF_noID)
```

Run predictions

Run predication without confidence

```
predictionsResultsFMqCF            <-predict(model_S3_FMqCF, Data_CallValues_FMqCF_noID)
```

Run predictions with confidence

```
PredictionResultsProb_FMqCF        <-predict(model_S3_FMqCF, Data_CallValues_FMqCF_noID, type = "prob")
PredictionResultsProb_FMqCF$ID     <-Stage1_FMqCF$ID
```

Combine predictions with confidence values

```
PredictionResultsCombined_FMqCF     <-cbind(PredictionResultsProb_FMqCF, predictionsResultsFMqCF)
```

Combine with file information

```
PredictionsFinalStage3          <-merge(PredictionResultsCombined_FMqCF, Stage1_FMqCF, by="ID")
```

Export stage 3 predictions

```
setwd(Dir_user_outputs)
write.csv(PredictionsFinalStage3, file="Data_PredictionsStage3.csv", na = "NA")
```

Run confidence thresholds

Creates table to see which files meet the confidence thresholds necessary for file structure later

Convert ID to factor for grouping

```
PredictionsFinalStage3$ID            <-as.factor(PredictionsFinalStage3$ID)
```

Identify which FMqCF are present in the data

```
Levels_FMqCF                          <-levels(as.factor(PredictionsFinalStage3$predictionsResultsFMqCF))
```

Create vector for grouping species (***user may need to edit depending on the species listed in Levels_FMqCF***)

```
cols_sp    <- c("FMqCF1" , "FMqCF2","FMqCF3","FMqCF4" ,"FMqCF5","LF"  , "LF_Acup")
```

Create vectors for grouping

```
cols_ID    <- c("ID", "predictionsResultsFMqCF")
cols_Files <- c("Filename", "predictionsResultsFMqCF")
```

Isolate the confidence of the predicted species into new column

```
Temp_S3_Max_ID <- PredictionsFinalStage3 %>%
                          group_by(across(all_of(cols_ID))) %>%
                          mutate(MaxByID = max(c(FMqCF1 , FMqCF2,FMqCF3,FMqCF4 ,
                                                 FMqCF5,LF  , LF_Acup), na.rm = T))
```

Find the pulse of highest confidence within each file for each species

```
RES_S3_summary <- Temp_S3_Max_ID %>%
                          group_by(across(all_of(cols_Files))) %>%
                          summarise(MaxbyFile = max(MaxByID, na.rm = T))
```

Rename columns

```
names(RES_S3_summary)<-c("Filename","S3_Prediction","S3_Accuracy")
```

7. COMBINE FINAL PREDICTIONS

Collate predictions from each classification stage & isolate files for manual verification

Create summaries

Select relevant columns

```
SummaryQCF             <-Stage1_QCF[, c("ID","predictionsResultsType","MaxByID", "Filename", "FileLoc")]

SummaryFM              <-Stage1_FM[, c("ID","predictionsResultsType","MaxByID", "Filename", "FileLoc")]

SummaryCF              <-Temp_S2_Max_ID[, c("ID","predictionsResultsCF","MaxByID", "Filename", "FileLoc")]

SummaryFMqCF           <-Temp_S3_Max_ID[, c("ID","predictionsResultsFMqCF","MaxByID", "Filename", "FileLoc")]
```

Rename column names to match

```
colnames(SummaryQCF)      <-c("ID","Prediction","Accuracy","Filename", "FileLoc")
colnames(SummaryFM)        <-c("ID","Prediction","Accuracy","Filename", "FileLoc")
colnames(SummaryCF)        <-c("ID","Prediction","Accuracy","Filename", "FileLoc")
colnames(SummaryFMqCF)     <-c("ID","Prediction","Accuracy","Filename", "FileLoc")
```

Join predictions

```
RES_total_raw      <-rbind(SummaryQCF, SummaryFM, SummaryCF, SummaryFMqCF)
```

Subset again for files above 60% confidence

This will subset for files identified in stages 2 and 3 to sonotype/species

```
RES_total_raw      <-RES_total_raw[RES_total_raw$Accuracy>0.59,]
```

Export final predictions

```
setwd(Dir_user_outputs)
write.csv(RES_total_raw, file="Data_PredictionsSummary.csv", na = "NA")
```

Summarise results

Summarises the pulse predictions to sonotype/species identification to the file/bat pass level

Convert to factor for grouping

```
RES_total_raw$ID            <-as.factor(RES_total_raw$ID)
```

Create vectors for grouping columns

```
cols_sp    <- levels(as.factor(RES_total_raw$Prediction))
cols_Files <- c("Filename", "Prediction")
```

Find the pulse of highest confidence within each file for each species

```
RES_total_sum <- RES_total_raw %>%
                         group_by(across(all_of(cols_Files))) %>%
                         summarise(MaxbyFile = max(Accuracy, na.rm = T))
```

Rename columns

```
names(RES_total_sum)<-c("filename","Prediction","Accuracy")
```

Readd file location

```
RES_total_sum <-merge(RES_total_sum, FileLoc, by="filename", all= FALSE)
```

Isolate files for manual verification & create library

The following steps if for users who are using stage 2/3 classifications.

Selects WAV files which do not reach the necessary confidence threshold using their original filepathways and copies them into a new filepathway based on ID prediction & confidence threshold.

Determine which files need manual verification

Rename column levels to match

```
colnames(Data_thres) <- c("Prediction", "Threshold")
```

Merge confidence threshold information with the predictions data

```
RES_total_sum        <- merge(RES_total_sum, Data_thres, by = "Prediction", keep.all=TRUE)
```

Rename column levels to match

```
colnames(Data_thres) <- c("Prediction", "Threshold")
```

Merge confidence threshold information with the predictions data

```
RES_total_sum        <- merge(RES_total_sum, Data_thres, by = "Prediction", keep.all=TRUE)
```

Create threshold level column

```
RES_total_sum$ThresLevel <-""
```

Ensure accuracy column is numeric

```
RES_total_sum$Accuracy   <-as.numeric(RES_total_sum$Accuracy)
```

Remove predictions below 60% confidence

```
RES_total_sum <-RES_total_sum[RES_total_sum$Accuracy > 0.59,]
```

Loop to determine which files met the necessarily confidence threshold

```
for (y in 1:nrow(RES_total_sum)){

  if((RES_total_sum$Accuracy[y]*100) == RES_total_sum$Threshold[y]) {
    RES_total_sum$ThresLevel[y] <- "Met"
  }

  else if ((RES_total_sum$Accuracy[y]*100) > RES_total_sum$Threshold[y]) {
    RES_total_sum$ThresLevel[y] <- "Met" }

  else if ((RES_total_sum$Accuracy[y]*100) < RES_total_sum$Threshold[y]) {
    RES_total_sum$ThresLevel[y] <- "Not Met" }
}
```

Filter data for files which did not meet the confidence threshold

```
DF_NotMet        <-filter(RES_total_sum, (ThresLevel=="Not Met"))
```

Remove repeated files so a file is only manually checked once

```
DF_NotMet_unique <- DF_NotMet[!duplicated(DF_NotMet['filename']),]
```

Save data outputs

```
setwd(Dir_user_outputs)
write.csv(DF_NotMet_unique, file="DF_NotMet_unique.csv", na = "NA")
write.csv(RES_total_sum,    file="Data_PredictionsSummary_max.csv", na = "NA")
```

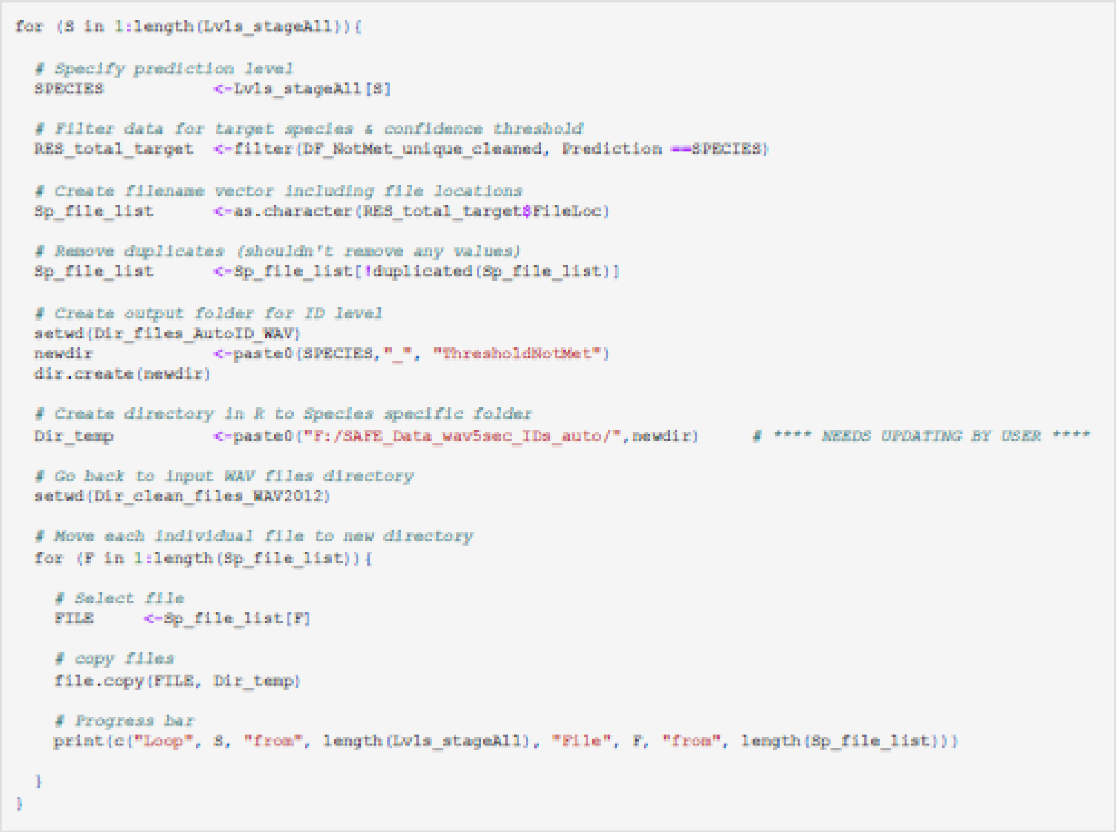
!!! THE FOLLOWING CODE WILL MOVE FILES DIRECTORY ON YOUR COMPUTER !!!

!!! ENSURE IT IS WORKING CORRECTLY USING A TEST FILE/BACK UP YOUR DATA BEFORE PROCEEDING !!!

Specify ID levels

```
Lvls_stageAll     <-levels(as.factor(RES_total_sum$Prediction))
```

Not reversible: Loop to create new folder pathway and copy WAV files - user needs to update pathway below



```
for (S in 1:length(Lvls_stageAll)){

  # Specify prediction level
  SPECIES             <-Lvls_stageAll[S]

  # Filter data for target species & confidence threshold
  RES_total_target    <-filter(DF_NotMet_unique_cleaned, Prediction ==SPECIES)

  # Create filename vector including file locations
  Sp_file_list        <-as.character(RES_total_target$FileLoc)

  # Remove duplicates (shouldn't remove any values)
  Sp_file_list        <-Sp_file_list[!duplicated(Sp_file_list)]

  # Create output folder for ID level
  setwd(Dir_files_AutoID_WAV)
  newdir              <-paste0(SPECIES,"_", "ThresholdNotMet")
  dir.create(newdir)

  # Create directory in R to Species specific folder
  Dir_temp            <-paste0("F:/SAFE_Data_wav5sec_IDs_auto/",newdir)     # **** NEEDS UPDATING BY USER ****

  # Go back to input WAV files directory
  setwd(Dir_clean_files_WAV2012)

  # Move each individual file to new directory
  for (F in 1:length(Sp_file_list)){

    # Select file
    FILE     <-Sp_file_list[F]

    # copy files
    file.copy(FILE, Dir_temp)

    # Progress bar
    print(c("Loop", S, "from", length(Lvls_stageAll), "File", F, "from", length(Sp_file_list)))

  }
}
```

---- End for all users -----

Note - Recordings will differ depending on the recording equipment and conditions. Please ensure you manually check a subset of the calls that meet the confidence threshold to ensure the classifier works adequately on your data before analysing the data