

**Yoh, Natalie, Kingston, Tigga, McArthur, Ellen, Aylen, Oliver E, Huang, Joe Chun-Chia, Jinggong, Emy Ritta, Khan, Faisal Ali Anwarali, Lee, Benjamin P.Y.-H., Mitchell,, Simon L., Bicknell, Jake E. and and others (2022) *A machine learning framework to classify Southeast Asian echolocating bats*. Ecological Indicators, 136 . ISSN 1470-160X.**

## Downloaded from

<https://kar.kent.ac.uk/93211/> The University of Kent's Academic Repository KAR

## The version of record is available from

<https://doi.org/10.1016/j.ecolind.2022.108696>

## This document version

Publisher pdf

## DOI for this version

## Licence for this version

CC BY-NC-ND (Attribution-NonCommercial-NoDerivatives)

## Additional information

## Versions of research works

### Versions of Record

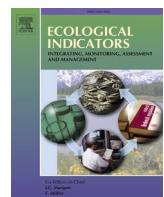
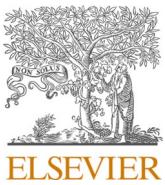
If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

### Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal**, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

### Enquiries

If you have questions about this document contact [ResearchSupport@kent.ac.uk](mailto:ResearchSupport@kent.ac.uk). Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).



## A machine learning framework to classify Southeast Asian echolocating bats



Natalie Yoh <sup>a,\*</sup>, Tigga Kingston <sup>b,c</sup>, Ellen McArthur <sup>d</sup>, Oliver E. Aylen <sup>e</sup>, Joe Chun-Chia Huang <sup>f</sup>, Emy Ritta Jinggong <sup>d</sup>, Faisal Ali Anwarali Khan <sup>d</sup>, Benjamin P.Y.H. Lee <sup>g</sup>, Simon L. Mitchell <sup>a</sup>, Jake E. Bicknell <sup>a</sup>, Matthew J. Struebig <sup>a</sup>

<sup>a</sup> Durrell Institute of Conservation and Ecology (DICE), School of Anthropology and Conservation, University of Kent, Canterbury, United Kingdom

<sup>b</sup> Department of Biological Sciences, Texas Tech University, Lubbock, TX, USA

<sup>c</sup> Southeast Asian Bat Conservation Research Unit, Lubbock City, TX, USA

<sup>d</sup> Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

<sup>e</sup> University of Otago, Department of Zoology, Dunedin, Otago, New Zealand

<sup>f</sup> Taiwan Forestry Research Institute, Taipei, Taiwan

<sup>g</sup> Wildlife Management Division, National Parks Board, Singapore

### ARTICLE INFO

#### Keywords:

Acoustic monitoring  
Chiroptera  
Echolocation  
Southeast Asia  
Machine learning  
Supervised algorithm

### ABSTRACT

Bats comprise a quarter of all mammal species, provide key ecosystem services and serve as effective bio-indicators. 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.

### 1. 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

\* Corresponding author.

E-mail address: [tallyoh@gmail.com](mailto:tallyoh@gmail.com) (N. Yoh).

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 et al., 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 et al., 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 & 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 in 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 under-represented 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). PAM techniques can be used to quantify a range of ecological metrics, including species diversity (López-Baucells et al., 2019), animal movement and activity (Furmanekiewicz & Kucharska, 2009), population dynamics (particularly for roost monitoring; Revilla-Martín et al., 2020), and responses to anthropogenic change (Meyer et al., 2016; Yoh 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 & Schnitzler, 1993). This within-individual and within-species variability is coupled with morphological, phylogenetic, and habitat constraints on adaptive call structure, and thus many species calls overlap in structure (Pham et al., 2021; Russo et al., 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 & 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 & Simmons, 1982; Russo et al., 2018). Bat species that do not rely on echolocation for foraging cannot be monitored using acoustic surveys (Russo et al., 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 (Kershnerbaum 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 et al., 2016).

Several developments in recent years show promise for automated classification of tropical bat calls. Software such as *Waveman* (Chen et al., 2020) demonstrates how 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 identification (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 & Cirranello, 2021). 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 et al., 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 et al., 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 maximize their sensory input for a range of environmental conditions

and to minimize 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).

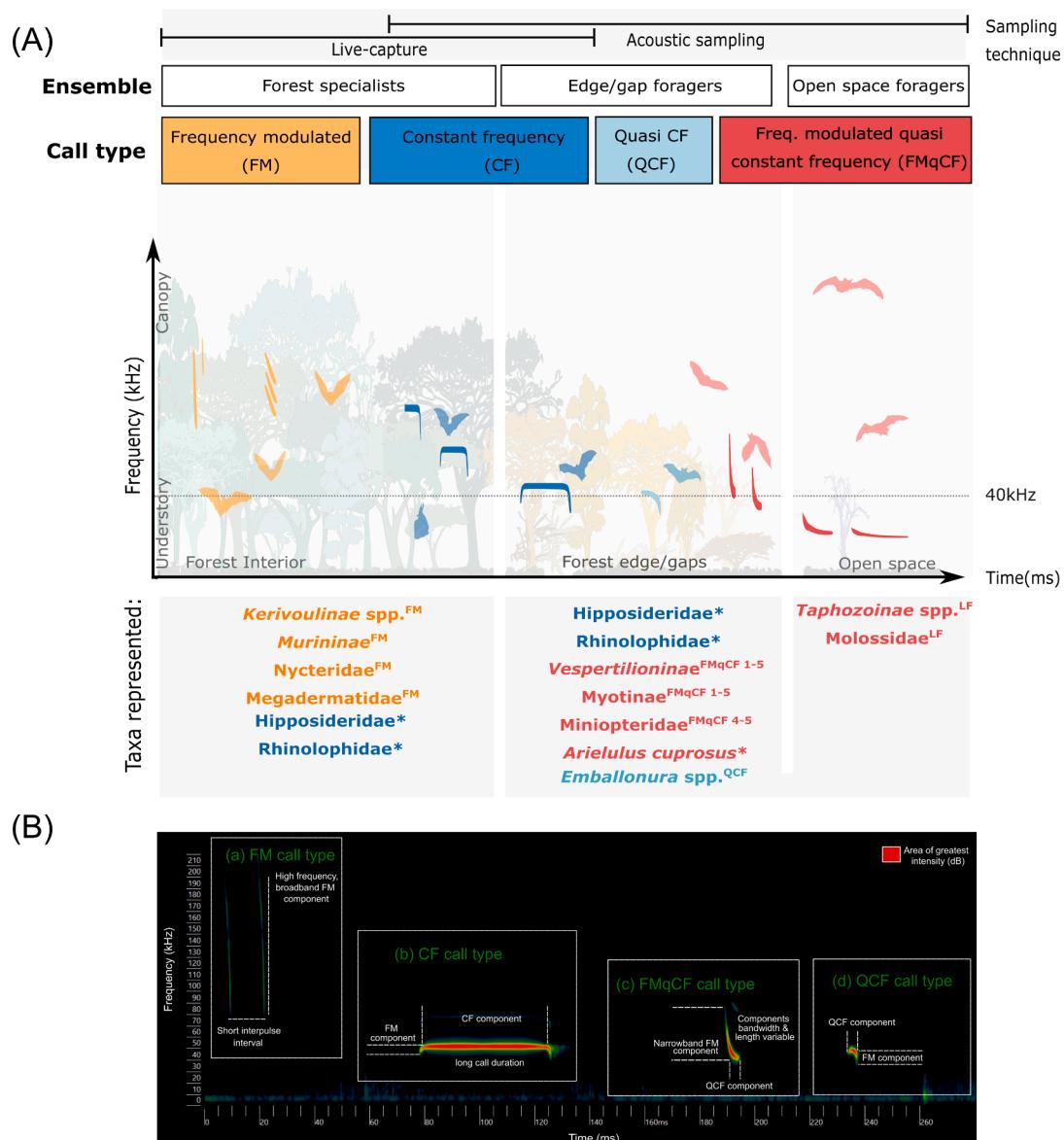
Open-space foragers hawk airborne prey across large, open spaces, such as above the forest canopy (Denzinger & Schnitzler, 2013). Their prey is more widely dispersed than within the forest interior, and consequently, they need an increased range of prey detection. Therefore, they use narrowband, high intensity calls with a long call duration and typically emit frequencies below 30 kHz (Denzinger & Schnitzler, 2013; Jung et al., 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 & Schnitzler, 2013; Siemers &

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. 1). 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 & Cirranello, 2021). 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.

## 2. Materials & methods

Current reference databases are typically insufficient for training



**Fig. 1.** The four call types used in the Borneo bat classifier. (A) Representation of bat ensembles in Borneo, their corresponding call types, and species/sonotypes used to train the bat call classifier. (\*Identified to species; <sup>FM</sup>Identified to call type 'frequency modulated'; <sup>FMqCF</sup>Identified to 'frequency modulated quasi constant frequency sonotype 1'; <sup>LF</sup>Identified to sonotype type 'low frequency'); (B) Example sonogram for each of the four call types. FM – frequency modulated, CF – constant frequency, FMqCF – frequency modulated quasi constant frequency, QCF – quasi constant frequency.

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.

## 2.1. 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. 2). 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](http://www.safeproject.net)), hereafter referred to as field recordings. Unlike the reference calls, which are tied to 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

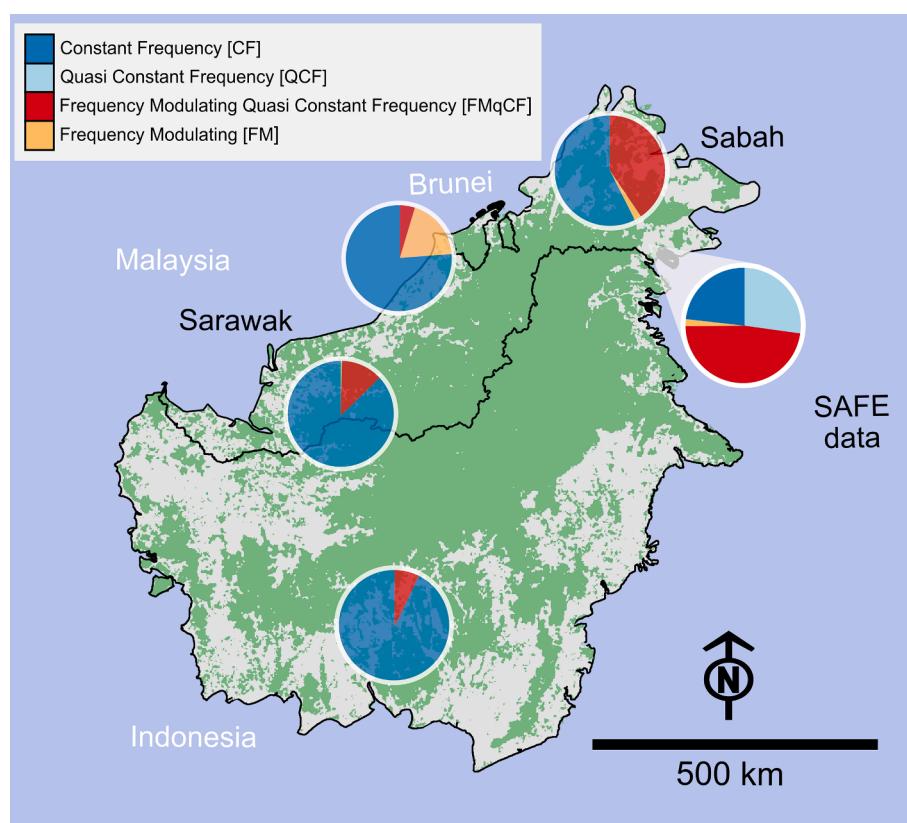
landscape for 862 hours in multiple habitat types (Supplementary Notes). Reference calls and field recordings were collected as WAV sound files. The calls from Sarawak contributed to the Asian Bat Call Database and are available from the Chirovox library (Görföl et al., 2022; McArthur & Khan, 2021).

### 2.1.1. 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 500 ms 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  $\geq 2$  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 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 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 1). For the call characteristics for the FMqCF sonotypes see Supplementary Table 3.

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



**Fig. 2.** 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.

**Table 1**

The total number of bat passes (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.

ID category	Code	Field		Sabah		Sarawak		Brunei		Kalimantan	
		Files	Pulses	Files	Pulses	Files	Pulses	Files	Pulses	Files	Pulses
<b>Constant frequency [CF]</b>											
<i>H. cinereus/dyacorum</i>	H140	1	2	5	48	12	338	8	49	10	140
<i>Hipposideros ater</i>	Hate			5	124			2		2	31
<i>Hipposideros bicolor</i>	Hbic			2	35						
<i>Hipposideros cervinus</i>	Hcer	18	54	14	397	14	487			24	423
<i>Hipposideros coxi</i>	Hcox			2	132						
<i>Hipposideros diadema</i>	Hdia	25	313	3	138	3	67				
<i>Hipposideros galeritus</i>	Hgal	78	218	2	146	8	217			6	139
<i>Hipposideros larvatus</i>	Hlar					3	67				
<i>Hipposideros ridleyi</i>	Hrid			9	276	2	50			2	24
<i>Rhinolophus acuminatus</i>	Racu	1	4	3	35	6	699				
<i>Rhinolophus affinis</i>	Raff	4	14			5	216				
<i>Rhinolophus borneensis</i>	Rbor	143	613	13	545	7	371			14	180
<i>Rhinolophus creaghi</i>	Rcre					6	113				
<i>Rhinolophus luctus</i>	Rluc	645				2	71				
<i>Rhinolophus philippinensis</i>	Rphi					10	390				
<i>Rhinolophus sedulus</i>	Rsed	2893	25,723	8	154	3	80	24	354	12	119
<i>Rhinolophus trifoliatus</i>	Rtri	4312	32,576	47	3051	1	41	8	103	14	82
<b>Frequency modulated constant frequency [FMCF]</b>											
FMCF sonotype 1	FMCF1	261	5520								
FMCF sonotype 2	FMCF2	567	9779								
FMCF sonotype 3	FMCF3	1368	14,948								
FMCF sonotype 4	FMCF4	2082	24,116		2953		225				
<i>Glischropus tylopus</i>				36		1					
<i>Myotis horsfieldii</i>						4					
<i>Tylonycteris robustula</i>				55							
FMCF sonotype 5	FMCF5	3350	65,102		4869		1090			31	86
<i>Miniopterus australis</i>						1					
<i>Myotis muricola</i>						7					
<i>Myotis ridleyi</i>						2					
<i>Tylonycteris pachypus</i>				10							
Low frequency sonotype	LF	1215	6742				1293				
<i>Chaerephon plicatus</i>						35					
<i>Saccopteryx saccolaimus</i>						1					
<i>Ariolulus cuprosus</i>	Acup	4	29			1	92				
Quasi-constant frequency* [QCF]		5288	71,854			9					
<i>Emballonura alecto</i>	QCF					1					
<i>Emballonura monticola</i>	QCF					1					
<b>Frequency modulated [FM] *</b>											
<i>Kerivoula hardwickii</i>	FM	1006	4898		174		420			127	
<i>Kerivoula intermedia</i>				5				18			
<i>Kerivoula lenis</i>						2		7			
<i>Kerivoula minuta</i>						1		1			
<i>Kerivoula papillosa</i>				4		5		21			
<i>Kerivoula pellucida</i>								8			
<i>Murina peninsulae</i>						1		1			
<i>Murina suilla</i>						3		7			
<i>Megaderma spasma</i>						2					
<i>Nycterus tragata</i>						3					
<i>Phoniscus atrox</i>				1				2			
<i>Phoniscus jagorii</i>				2				2			

species (e.g., where there was a flyby). Calls that were obscured or faint (< 20 dB), 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).

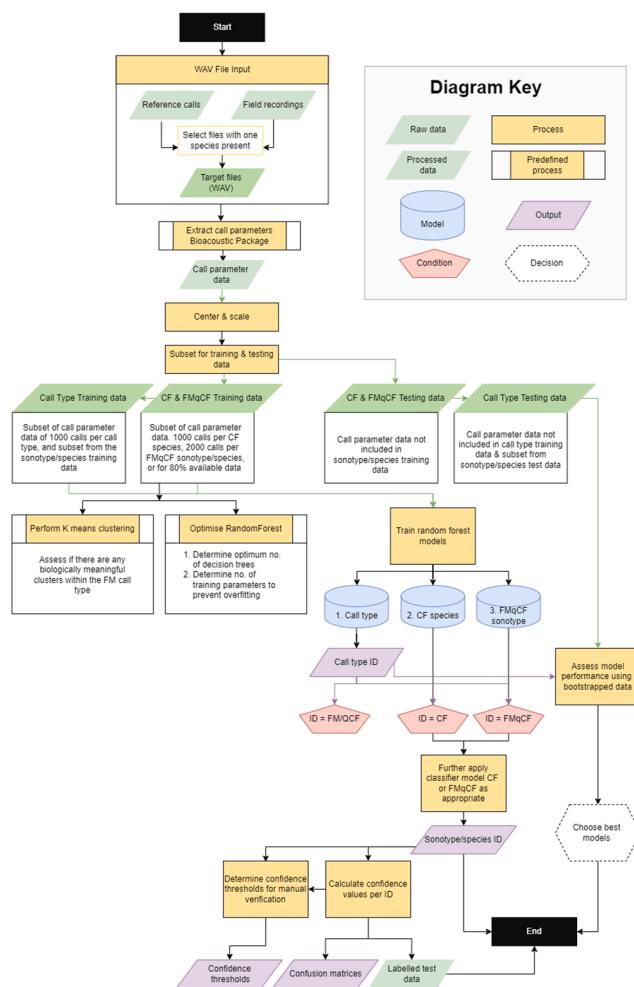
#### 2.1.2. Call parameters

We used the threshold function in the R package “Bioacoustics” (Marchal et al., 2020) to measure 26 call parameters (Supplementary Table 2; settings: minimum duration 1.5 ms, maximum duration 80 ms, FFT size = 512, FFT overlap = 0.875, extraction threshold (sensitivity of which extraction is triggered) = 4 dB, signal to noise ratio threshold (SNR; sensitivity threshold at which the extraction stops) = 4 dB, 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. All the

numerical call measurement values were subsequently centred and scaled to normalise the data (James et al., 2013).

#### 2.2. 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 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. 3 for full pipeline.



**Fig. 3.** Pipeline for building the classifier framework to identify bat calls first to call type and then to species/sonotype where appropriate.

#### 2.2.1. 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, 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.

#### 2.2.2. 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. 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 et al., 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 and 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 decision trees across all models for consistency.

#### 2.2.3. 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 & 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.

#### 2.3. Performance testing

##### 2.3.1. Testing success rate on bootstrapped data

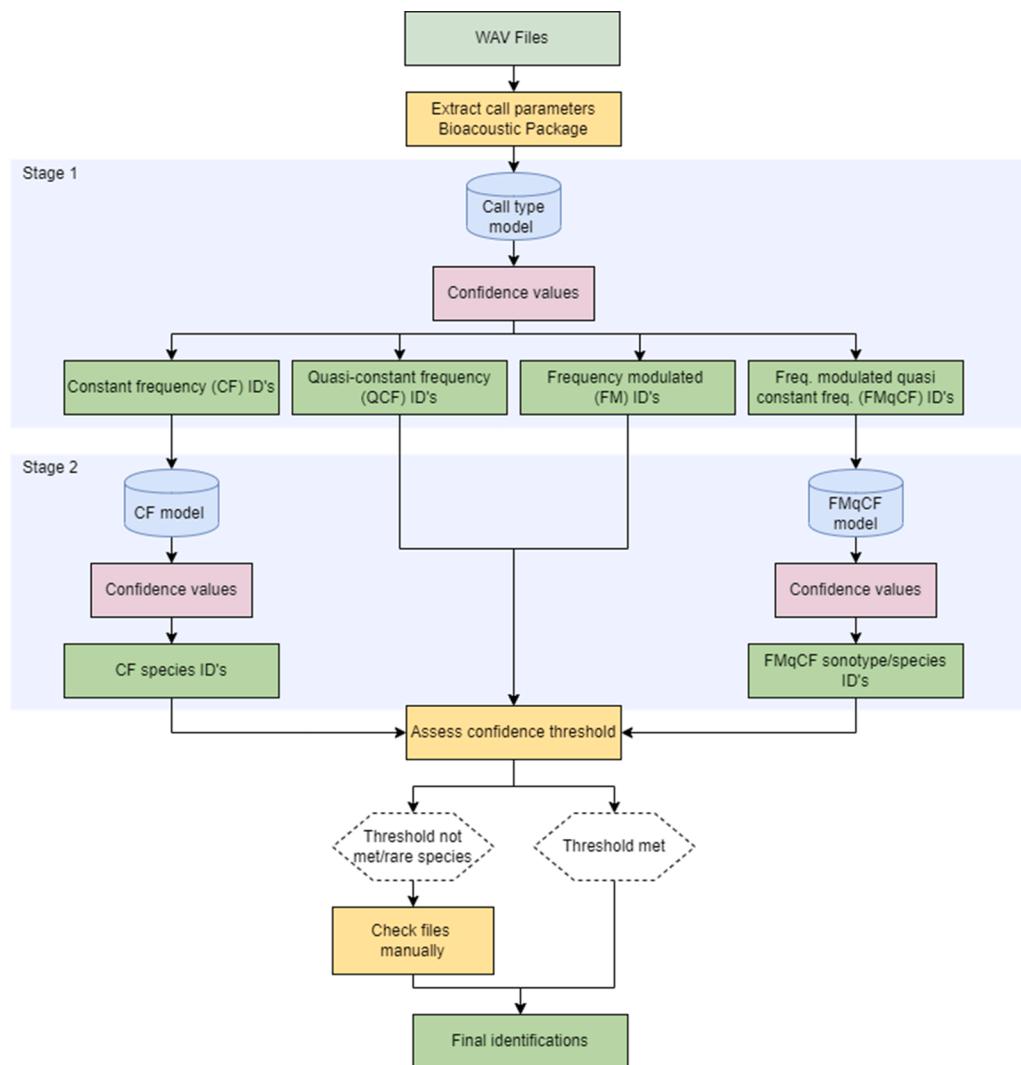
We evaluated classification accuracy and predictive power of the models on the testing dataset using accuracy and kappa performance metrics, with acceptable agreement determined as  $> 0.41$  (McHugh, 2012). We chose recall (percentage of true positives, e.g., number of correct classifications per class 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 multiclass acoustic classifiers (Jennings et al., 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.

##### 2.3.2. 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). 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 and 95% in 5% steps. Files with  $< 60\%$  accuracy were discounted.

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.



**Fig. 4.** 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.

### 3. Results

#### 3.1. Minimum training dataset size

We found  $> 85\%$  accuracy in identifying calls to call type using 1000 calls per category (Fig. 5), with a kappa  $> 0.81$  indicating almost perfect agreement. Accuracy was still high ( $\sim 80\%$ ) at smaller training data sizes, however there was greater variability in performance with accuracy varying up to  $\sim 10\%$  and kappa by  $\sim 15\%$ . There was little improvement by increasing 1000 calls to 2000 or 5000 calls compared to the increase in computational power, (which ranged from  $\sim 3$ -fold to  $> 10$ -fold depending on the measure of power and training data size; Supplementary Table 4). The second model, classifying calls to CF species, showed a similar trend. Training data sizes  $\geq 1000$  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.

#### 3.2. Call parameter importance

The random forest prioritized different parameters when classifying to either call type or sonotype/species (Fig. 6). “Maximum frequency” was the most important parameter for determining call type and

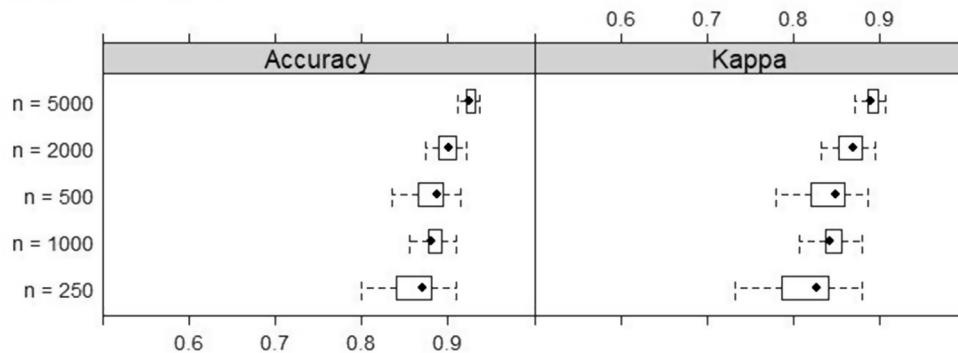
differentiating between CF calls. The CF model also shared eight of the ten most important parameters as used for classifying to call type, though their importance between models varied. While “raw slope estimate” (slope) and “smoothed slope estimate after Kalman filtering” (slope smoothed) were not 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 second most important for determining call type.

#### 3.3. Success in performing classifications

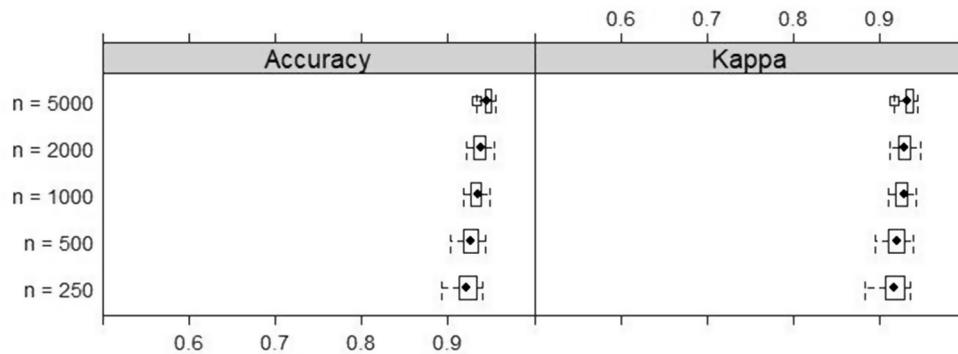
To call type, the classifier achieved  $> 90\%$  balanced accuracy for all call-types (Table 2). Using the bootstrapped data  $> 85\%$  of the calls were identified correctly to call type (Table 3). FM calls were correctly identified in 88.5% of cases. Where they were misclassified, they were most commonly reported as FMqCF type calls (7.9% cases) (Supplementary Table 5). QCF calls were correctly identified in 91.38% of cases but were misidentified predominantly as CF calls (4.05% cases).

Across both the second (CF classification) and third models (FMqCF classification), all sonotypes/species achieved a balanced accuracy outcome  $\geq 80\%$ . 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%

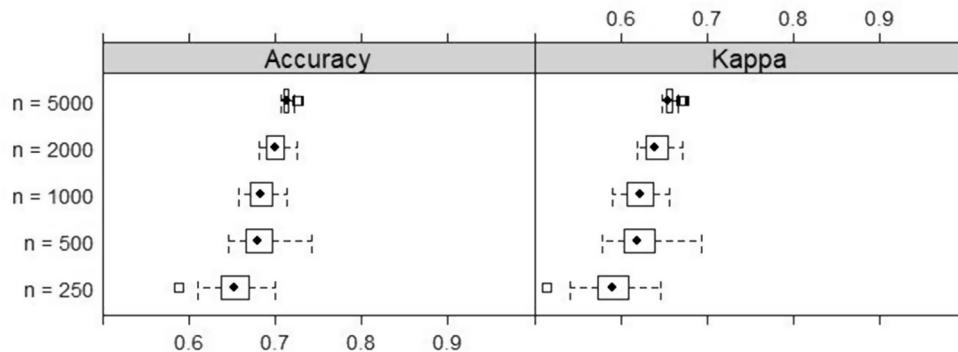
## (A) Classification to call type



## (B) Classification to CF species



## (C) Classification to FMqCF sonotype/species



**Fig. 5.** Comparison of each model performance using 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.

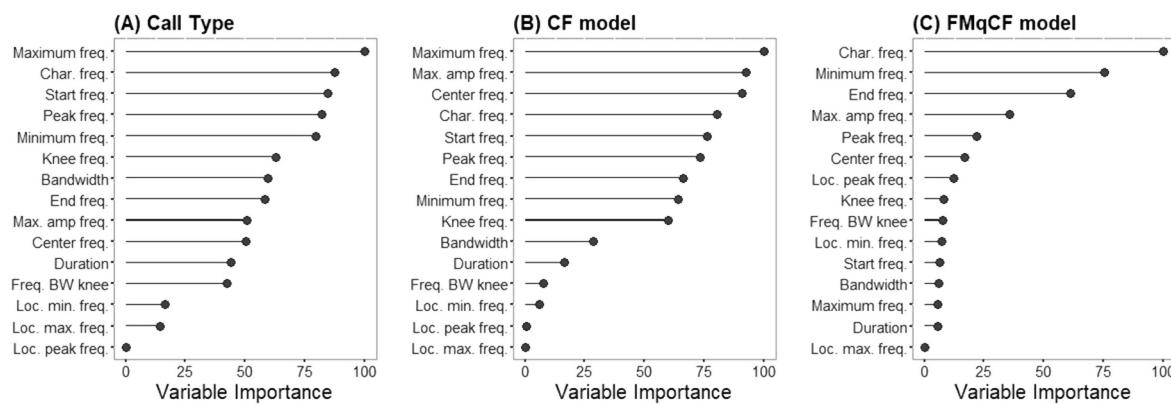
respectively (Table 2). FMqCF sonotypes showed the greatest rates of misclassification. However, they were still correctly identified in 65.9–73.55% of cases (Supplementary Table 6).

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.

### 3.4. 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 2; Supplementary Table 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 2, Fig. 7). 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 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



**Fig. 6.** 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 2 for parameter acronyms. (A) Classification to call type, (B) classification to CF species, and (C) classification to FMqCF sonotype/species.

**Table 2**

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
<b>Model 1: Classification to call type</b>								
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
<b>Model 2: Classification to CF species</b>								
<i>H. cineraceus/dyacorum</i>	H140	0.97	1.00	0.58	0.98	0.97	75	20
<i>Hipposideros ater</i>	Hate	0.92	1.00	0.61	0.96	0.95	65	16
<i>Hipposideros bicolor</i>	Hbic	0.60	1.00	0.75	0.8	NA	NA	100
<i>Hipposideros cervinus</i>	Hcer	0.99	1.00	0.88	0.99	0.95	60	0
<i>Hipposideros coxi</i>	Hcox	1.00	1.00	0.08	1.00	NA	NA	100
<i>Hipposideros diadema</i>	Hdia	0.97	0.99	0.19	0.98	NR	NR	100
<i>Hipposideros galeritus</i>	Hgal	0.94	1.00	0.57	0.97	0.98	60	0
<i>Hipposideros larvatus</i>	Hlar	1.00	1.00	1.00	1.00	NA	NA	100
<i>Hipposideros ridleyi</i>	Hrid	0.84	1.00	0.76	0.92	0.94	60	0
<i>Rhinolophus acuminatus</i>	Racu	0.99	1.00	0.38	0.99	0.92	60	0
<i>Rhinolophus affinis</i>	Raff	0.89	1.00	0.94	0.94	NA	NA	100
<i>Rhinolophus borneensis</i>	Rbor	0.96	0.99	0.58	0.98	0.91	95	62
<i>Rhinolophus creaghi</i>	Rcre	0.71	1.00	0.36	0.86	NA	NA	100
<i>Rhinolophus luctus</i>	Rluc	0.99	0.99	0.39	0.99	0.94	80	31
<i>Rhinolophus philippinensis</i>	Rphi	1.00	1.00	0.22	1.00	0.95	85	49
<i>Rhinolophus sedulus</i>	Rsed	0.96	0.98	0.97	0.97	1.00	60	0
<i>Rhinolophus trifoliatus</i>	Rtri	0.99	0.9	0.91	0.95	0.96	60	0
<b>Model 3: Classification to FMqCF sonotype/species</b>								
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
<i>Arielulus cuprosus</i>	Acup	0.83	1.00	0.19	0.92	NA	NA	100

*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 pulses 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).

**Table 3**

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.

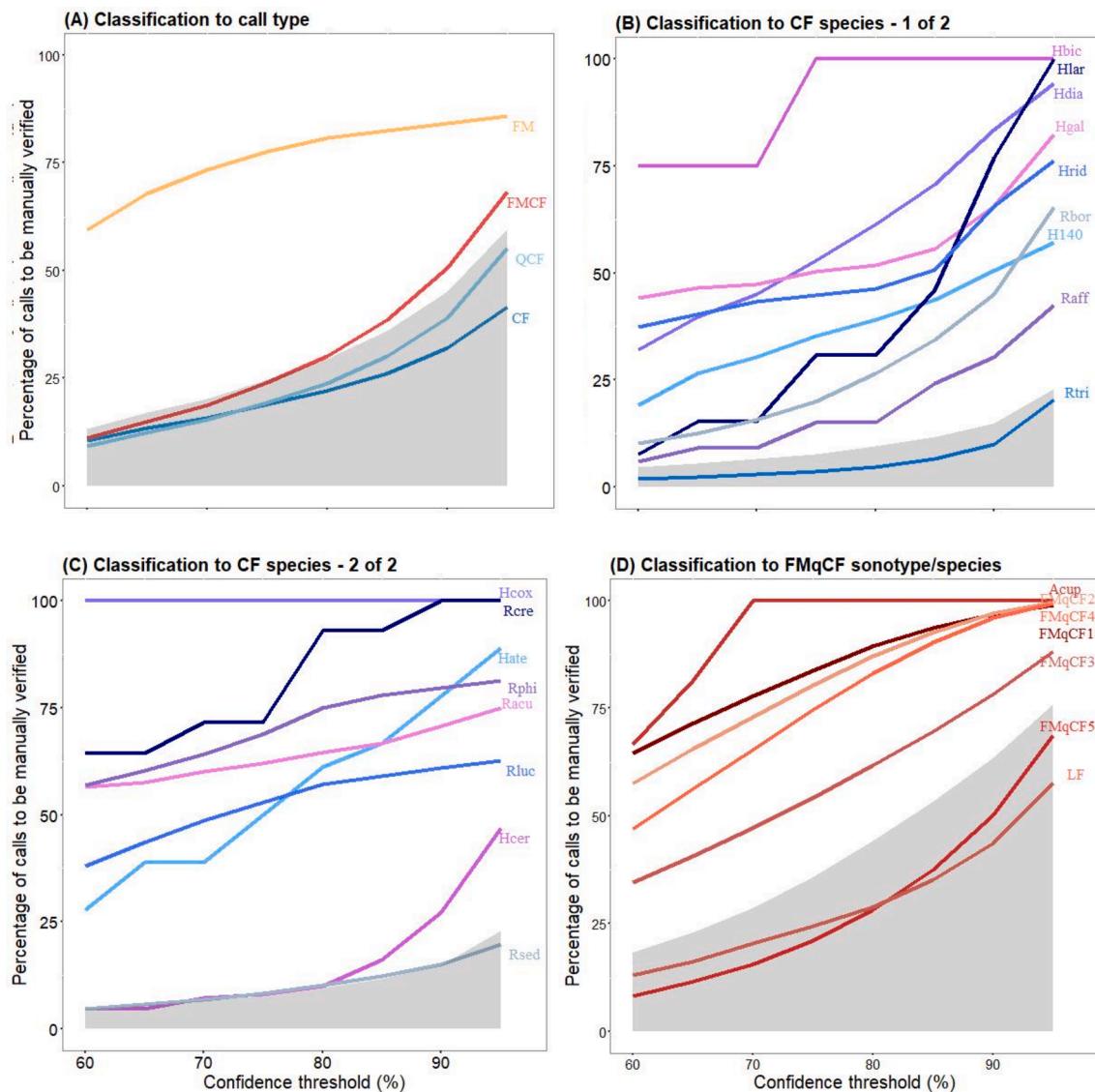
Prediction	True Call Type Identification			
	CF	FMqCF	QCF	FM
CF	88.5	1.2	4.05	1.7
FMqCF	2.4	87.9	2.86	7.9
QCF	3.8	5	91.38	2
FM	5.3	5.9	1.7	88.5
N calls	62,170	125,398	69,854	3619

#### 4. 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 et al., 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



**Fig. 7.** Percentage of echolocation calls requiring manual verification using the Borneo Bat Classifier, dependant on model and sonotype/species by confidence threshold. For clarity, the model for constant frequency (CF) species has been divided across two panels (B and C). FMqCF – frequency modulated quasi constant frequency. Shaded area – the overall proportion of files per model.

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 et al., 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 generated by 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 & 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 (Görböl 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 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, constitute 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 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 will 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 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 (Philipp & Philipp, 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.

## 5. Conclusion

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 paleotropics more broadly, where species-specific classifiers are not yet possible. Therefore, improving the potential to use bat assemblages as bioindicators in tropical environments.

## 6. Availability

The BBC classifier user script is available open access from the GitHub repository [github.com/TallyYoh/BorneoBatCalls](https://github.com/TallyYoh/BorneoBatCalls) (<https://doi.org/10.5281/zenodo.4725680>) in the programming language R version 3.6.3.

## Funding

NY was funded by a Natural Environmental Research Council (NERC) EnvEast DTP scholarship (grant number NE/L002582/1). NERC also funded the acoustic surveys in Sabah (NE/K016407/1; <https://lom.bok.nerc-hmtf.info/>) along with the Mohamed bin Zayed species Conservation Fund (11253049). FAAK and ERJ were supported by the Ministry of Higher Education Malaysia through F07/FRGS/1878/2019. JCCH was supported by the Global Biodiversity Information Facility through BIFA grant. BPYHL was supported by a Ministry of National Development EDGE Scholarship and the Wildlife Reserves Singapore Conservation Fund. TK was funded by the US National Science Foundation (165871). We thank the Sabah Biodiversity Council, Sabah Forest Department, Yayasan Sabah, and Benta Wawasan Sdn Bhd. for research permissions in Sabah.

## CRediT authorship contribution statement

**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.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## 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 Grawe 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 Göröföl, Amira 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.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ecolind.2022.108696>.

## References

Breiman, L., 2001. Random forests. *Mach. Learning* 45 (1), 5–32. <https://doi.org/10.1023/A:1010933404324>.

Chen, S.F., Jones, G., Rossiter, S.J., 2009. Determinants of echolocation call frequency variation in the Formosan lesser horseshoe bat (*Rhinolophus monoceros*). *Proc. R. Soc. B: Biol. Sci.* 276 (1674), 3901–3909. <https://doi.org/10.1098/rspb.2009.1185>.

Chen, X., Zhao, J., Chen, Y.-H., Zhou, W., Hughes, A.C., 2020. Automatic standardized processing and identification of tropical bat calls using deep learning approaches. *Biol. Conserv.* 241, 108269. <https://doi.org/10.1016/j.biocon.2019.108269>.

Clement, M.J., Murray, K.L., Solick, D.I., Gruver, J.C., 2014. The effect of call libraries and acoustic filters on the identification of bat echolocation. *Ecol. Evol.* 4 (17), 3482–3493. <https://doi.org/10.1002/ece3.1201>.

Denzinger, A., Schnitzler, H.-U., 2013. Bat guilds, a concept to classify the highly diverse foraging and echolocation behaviors of microchiropteran bats. *Front. Physiol.* 4, 164. <https://doi.org/10.3389/fphys.2013.00164>.

Fenton, M.B., Bell, G.P., 1981. Recognition of species of insectivorous bats by their echolocation calls. *J. Mammal.* 62 (2), 233–243. <https://doi.org/10.2307/1380701>.

Fisher-Phelps, M., Cao, G., Wilson, R.M., Kingston, T., 2017. Protecting bias: Across time and ecology, open-source bat locality data are heavily biased by distance to protected area. *Ecol. Inf.* 40, 22–34. <https://doi.org/10.1016/j.ecoinf.2017.05.003>.

Furmankiewicz, J., Kucharska, M., 2009. Migration of Bats along a Large River Valley in Southwestern Poland. *J. Mammal.* 90 (6), 1310–1317. <https://doi.org/10.1644/09-MAMM-S-099R1.1>.

Gardner, T.A., Barlow, J., Araujo, I.S., Ávila-Pires, T.C., Bonaldo, A.B., Costa, J.E., Esposito, M.C., Ferreira, L.V., Hawes, J., Hernandez, M.I.M., Hoogmoed, M.S., Leite, R.N., Lo-Man-Hung, N.F., Malcolm, J.R., Martins, M.B., Mestre, L.A.M., Miranda-Santos, R., Overal, W.L., Parry, L., Peters, S.L., Ribeiro-Junior, M.A., Da Silva, M.N.F., Da Silva Motta, C., Peres, C.A., 2008. The cost-effectiveness of biodiversity surveys in tropical forests. *Ecol. Lett.* 11 (2), 139–150. <https://doi.org/10.1111/j.1461-0248.2007.01133.x>.

Gibb, R., Browning, E., Glover-Kapfer, P., Jones, K.E., Börger, L., 2019. Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods Ecol. Evol.* 10 (2), 169–185. <https://doi.org/10.1111/2041-210X.13101>.

Göröföl, T., Huang, J.-C., Csorba, G., Györössy, D., Estók, P., Kingston, T., Szabadi, K.L., McArthur, E., Senawi, J., Furey, N.M., Tu, V.T., Thong, V.D., Khan, F.A.A., Jinggong, E.R., Donnelly, M., Kumaran, J.V., Liu, J.-N., Chen, S.-F., Tuanmu, M.-N., Ho, Y.-Y., Chang, H.-C., Elias, N.-A., Abdullah, N.-I., Lim, L.-S., Squire, C.D., Zsebők, S., 2022. ChiroVox: A public library of bat calls. *PeerJ* 10, e12445. <https://doi.org/10.7717/peerj.12445>.

Harrell Jr, F.E., 2015. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression and Survival Analysis*, 2nd ed. Springer.

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. *Appl. Stat.* 28 (1), 100. <https://doi.org/10.2307/2346830>.

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: *Vesperilionidae, Emballonuridae, Nycteridae and Megadermatidae*. *Acta Chiropterol.* 13 (2), 447–455. <https://doi.org/10.3161/150811011X624938>.

IUCN, 2021. The IUCN Red List of Threatened Species. Version 2021-1. <https://www.iucnredlist.org/en>.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. *An Introduction to Statistical Learning: With Applications in R*. Springer International Publishing.

Jennings, N., Parsons, S., Pocock, M.J.O., 2008. Human vs. machine: Identification of bat species from their echolocation calls by humans and by artificial neural networks. *Can. J. Zool.* 86 (5), 371–377. <https://doi.org/10.1139/Z08-009>.

Jones, G., Jacobs, D.S., Kunz, T.H., Willig, M.R., Racey, P.A., 2009. Carpe noctem: The importance of bats as bioindicators. *Endanger. Species Res.* 8, 93–115. <https://doi.org/10.3354/esr00182>.

Jones, K.E., Russ, J.A., Bashta, A.-T., Bilhari, Z., Catto, C., Csósz, I., Gorbachev, A., Györfi, P., Hughes, A., Ivashkiv, I., Koryagina, N., Kurali, A., Langton, S., Collen, A., Margiean, G., Pandourski, I., Parsons, S., Prokofev, I., Szodoray-Paradi, A., Szodoray-Paradi, F., Tilova, E., Walters, C.L., Weatherill, A., Zavarzin, O., 2013. Indicator Bats Program: A System for the Global Acoustic Monitoring of Bats. In: Collen, B., Pettorelli, N., Baillie, J.E.M., Durant, S.M. (Eds.), *Biodiversity Monitoring and Conservation: Bridging the Gap between Global Commitment and Local Action*. John Wiley & Sons, pp. 211–247. <https://doi.org/10.1002/9781118490747.ch10>.

Jung, K., Molinari, J., Kalko, E.K.V., Ratcliffe, J.M., 2014. Driving Factors for the Evolution of Species-Specific Echolocation Call Design in New World Free-Tailed Bats (*Molossidae*). *PLoS ONE* 9 (1), e85279. <https://doi.org/10.1371/journal.pone.0085279>.

Kalko, E.K.V., Schnitzler, H.-U., 1993. Plasticity in echolocation signals of European pipistrelle bats in search flight: Implications for habitat use and prey detection. *Behav. Ecol. Sociobiol.* 33 (6), 415–428. <https://doi.org/10.1007/BF00170257>.

Kershenbaum, A., Blumstein, D.T., Roch, M.A., Akçay, C., Backus, G., Bee, M.A., Bohn, K., Cao, Y., Carter, G., Cäsar, C., Coen, M., DeRuiter, S.L., Doyle, L., Edelman, S., Ferrer-i-Cancho, R., Freeberg, T.M., Garland, E.C., Gustison, M., Harley, H.E., Huetz, C., Hughes, M., Hyland Bruno, J., Ilany, A., Jin, D.Z., Johnson, M., Ju, C., Karnowski, J., Lohr, B., Manser, M.B., McCowan, B.,

Mercado, E., Narins, P.M., Piel, A., Rice, M., Salmi, R., Sasahara, K., Sayigh, L., Shiu, Y.u., Taylor, C., Vallejo, E.E., Waller, S., Zamora-Gutierrez, V., 2016. Acoustic sequences in non-human animals: A tutorial review and prospectus. *Biol. Rev.* 91 (1), 13–52. <https://doi.org/10.1111/brv.12160>.

Kingston, T., 2010. Research priorities for bat conservation in Southeast Asia: A consensus approach. *Biodivers. Conserv.* 19 (2), 471–484. <https://doi.org/10.1007/s10531-008-9458-5>.

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

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

Kingston, T., Jones, G., Akbar, Z., Kunz, T.H., 1999. Echolocation signal design in *Kerivoulinae* and *Murininae* (Chiroptera: *Vesperilionidae*) from Malaysia. *J. Zool.* 249 (3), 359–374. <https://doi.org/10.1111/j.1469-7998.1999.tb0071x>.

Kingston, T., Jones, G., Akbar, Z., Kunz, T.H., 2003. Alteration of Echolocation Calls in 5 Species of Aerial-Feeding Insectivorous Bats from Malaysia. *J. Mammal.* 84 (1), 205–215. [https://doi.org/10.1644/1545-1542\(2003\)084<205:AOECIS>2.0.CO;2](https://doi.org/10.1644/1545-1542(2003)084<205:AOECIS>2.0.CO;2).

Kobayashi, K., Masuda, K., Haga, C., Matsui, T., Fukui, D., Machimura, T., 2021. Development of a species identification system of Japanese bats from echolocation calls using convolutional neural networks. *Ecol. Inf.* 62, 101253. <https://doi.org/10.1016/j.ecolinf.2021.101253>.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kunkel, B., R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., & Hunt, T., 2020. caret: Classification and Regression Training (6.0-86) [Computer software]. <https://CRAN.R-project.org/package=caret>.

Kwok, R., 2019. AI empowers conservation biology. *Nature* 567 (7746), 133–134. <https://doi.org/10.1038/d41586-019-00746-1>.

Lane, D.J.W., Kingston, T., Lee, B.P.Y.H., 2006. Dramatic decline in bat species richness in Singapore, with implications for Southeast Asia. *Biol. Conserv.* 131 (4), 584–593. <https://doi.org/10.1016/j.biocon.2006.03.005>.

Lawrence, B.D., Simmons, J.A., 1982. Measurements of atmospheric attenuation at ultrasonic frequencies and the significance for echolocation by bats. *J. Acoust. Soc. Am.* 71 (3), 585–590. <https://doi.org/10.1121/1.387529>.

López-Baucells, A., Torrent, L., Rocha, R., Bobrowiec, P.E.D., Palmeirim, J.M., Meyer, C. F.J., 2019. Stronger together: combining automated classifiers with manual post-validation optimizes the workload vs reliability trade-off of species identification in bat acoustic surveys. *Ecol. Inf.* 49, 45–53. <https://doi.org/10.1016/j.ecolinf.2018.11.004>.

Marchal, J., Fabianek, F., & Scott, C., 2020. Bioacoustics: Analyse Audio Recordings and Automatically Extract Animal Vocalizations (R package version 0.2.3) [Computer software]. <https://CRAN.R-project.org/package=bioacoustics>.

McArthur, E., Khan, F.A.A., 2021. Towards a regional call library: classifying calls of a species-rich bat assemblage in a Bornean karst rainforest. *J. Bat Res. Conserv.* 14 (1), 95–117. <https://doi.org/10.14709/BarbJ.14.1.2021.11>.

McHugh, M.L., 2012. Interrater reliability: The kappa statistic. *Biochem. Med.* 22 (3), 276–282.

Meyer, C.F.J., Struebig, M.J., Willig, M.R., 2016. Responses of tropical bats to habitat fragmentation, logging, and deforestation. In: Voigt, C.C., Kingston, T. (Eds.), *Bats in the Anthropocene: Conservation of Bats in a Changing World*. Springer, pp. 63–103. <https://www.oapen.org/download?type=document&docid=1001864#page=68>. [https://doi.org/10.1007/978-3-319-25220-9\\_4](https://doi.org/10.1007/978-3-319-25220-9_4).

Olden, J.D., Lawler, J.J., Poff, N.L., 2008. Machine learning methods without tears: A primer for ecologists. *Q. Rev. Biol.* 83 (2), 171–193. <https://doi.org/10.1086/587826>.

Park, K.J., 2015. Mitigating the impacts of agriculture on biodiversity: bats and their potential role as bioindicators. *Mamm. Biol.* 80 (3), 191–204. <https://doi.org/10.1016/j.mambio.2014.10.004>.

Parsons, S., Jones, G., 2000. Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks. *J. Exp. Biol.* 203 (17), 2641–2656. <https://doi.org/10.1242/jeb.203.17.2641>.

Pham, L.K., Tran, B.V., Le, Q.T., Nguyen, T.T., Voigt, C.C., 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 (1), 18. <https://doi.org/10.3390/d13010018>.

Phillipps, Q., Phillipps, K., 2016. *Phillipps' Field Guide to the Mammals of Borneo*, 1st ed. John Beaufoy Publishing Ltd.

Pottie, S.A., Lane, D.J.W., Kingston, T., Lee, B.P.Y.H., 2005. The microchiropteran bat fauna of Singapore. *Acta Chiropterol.* 7 (2), 237–247. [https://doi.org/10.3161/1733-5329\(2005\)7\[237:TMBFOS\]2.0.CO;2](https://doi.org/10.3161/1733-5329(2005)7[237:TMBFOS]2.0.CO;2).

Revilla-Martín, N., Budinski, I., Puig-Montserrat, X., Flaquer, C., López-Baucells, A., 2020. Monitoring cave-dwelling bats using remote passive acoustic detectors: a new approach for cave monitoring. *Bioacoustics* 30 (5), 527–542. <https://doi.org/10.1080/09524622.2020.1816492>.

Russo, D., Ancillotto, L., Jones, G., 2018. Bats are still not birds in the digital era: Echolocation call variation and why it matters for bat species identification. *Can. J. Zool.* 96 (2), 63–78. <https://doi.org/10.1139/cjz-2017-0089>.

Russo, D., Voigt, C.C., 2016. The use of automated identification of bat echolocation calls in acoustic monitoring: a cautionary note for a sound analysis. *Ecol. Ind.* 66, 598–602. <https://doi.org/10.1016/j.ecolind.2016.02.036>.

Schmieder, D.A., Kingston, T., Hashim, R., Siemers, B.M., Patek, S., 2012. Sensory constraints on prey detection performance in an ensemble of vespertilionid understorey rain forest bats. *Funct. Ecol.* 26 (5), 1043–1053. <https://doi.org/10.1111/j.1365-2435.2012.02024.x>.

Schnitzler, H.-U., 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 (7), 557–569. [https://doi.org/10.1641/0006-3568\(2001\)051\[0557:EBIEB\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0557:EBIEB]2.0.CO;2).

Schnitzler, H.-U., Moss, C.F., Denzinger, A., 2003. From spatial orientation to food acquisition in echolocating bats. *Trends Ecol. Evol.* 18 (8), 386–394. [https://doi.org/10.1016/S0169-5347\(03\)00185-X](https://doi.org/10.1016/S0169-5347(03)00185-X).

Siemers, B.M., Schnitzler, H.-U., 2004. Echolocation signals reflect niche differentiation in five sympatric congeneric bat species. *Nature* 429 (6992), 657–661. <https://doi.org/10.1038/nature02547>.

Simmons, N. B., & Cirranello, A. L., 2021. Bat Species of the World: A taxonomic and geographic database.

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 Chiropterol.* 17 (1), 21–36. <https://doi.org/10.3161/15081109ACC2015.17.1.002>.

Struebig, M.J., Turner, A., Giles, E., Lasmana, F., Tollington, S., Bernard, H., Bell, D., 2013. Quantifying the Biodiversity Value of Repeatedly Logged Rainforests. *Global Change Multiplespecies Syst.: Part 3* 183–224. <https://doi.org/10.1016/B978-0-12-417199-2.00003-3>.

Sugai, L.S.M., Silva, T.S.F., Ribeiro Jr, J.W., Llusia, D., 2019. Terrestrial passive acoustic monitoring: review and perspectives. *Bioscience* 69 (1), 15–25. <https://doi.org/10.1093/biosci/biy147>.

Tabak, M.A., Norouzzadeh, M.S., Wolfson, D.W., Sweeney, S.J., Vercauteren, K.C., Snow, N.P., Halseth, J.M., Di Salvo, P.A., Lewis, J.S., White, M.D., Teton, B., Beasley, J.C., Schlichting, P.E., Boughton, R.K., Wight, B., Newkirk, E.S., Ivan, J.S., Odell, E.A., Brook, R.K., Lukacs, P.M., Moeller, A.K., Mandeville, E.G., Clune, J., Miller, R.S., Photopoulou, T., 2019. Machine learning to classify animal species in camera trap images: applications in ecology. *Methods Ecol. Evol.* 10 (4), 585–590. <https://doi.org/10.1111/2041-210X.13120>.

Tanshi, I., Kingston, T., 2021. Introduction and Implementation of Harp Traps Signal a New Era in Bat Research. In: Lim, B.K., Fenton, M.B., Brigham, R.M., Mistry, S., Kurta, A., Gillam, E.H., Russell, A., Ortega, J. (Eds.), *50 Years of Bat Research: Foundations and New Frontiers*. Springer International Publishing, pp. 255–270. [https://doi.org/10.1007/978-3-030-54727-1\\_16](https://doi.org/10.1007/978-3-030-54727-1_16).

Torrent, L., López-Baucells, A., Rocha, R., Bobrowiec, P.E.D., Meyer, C.F.J., Pettorelli, N., Merchant, N., 2018. The importance of lakes for bat conservation in Amazonian rainforests: an assessment using autonomous recorders. *Remote Sens. Ecol. Conserv.* 4 (4), 339–351. <https://doi.org/10.1002/rse2.83>.

Valletta, J.J., Torney, C., Kings, M., Thornton, A., Madden, J., 2017. Applications of machine learning in animal behaviour studies. *Anim. Behav.* 124, 203–220. <https://doi.org/10.1016/j.anbehav.2016.12.005>.

Walters, C.L., Collen, A., Lucas, T., Mroz, K., Sayer, C.A., Jones, K.E., 2013. Challenges of using bioacoustics to globally monitor bats. In: Adams, R.A., Pedersen, S.C. (Eds.), *Bat Evolution, Ecology, and Conservation*. Springer, pp. 479–499. [https://doi.org/10.1007/978-1-4614-7397-8\\_23](https://doi.org/10.1007/978-1-4614-7397-8_23).

Wei, L., Han, N., Zhang, L., Helgen, K.M., Parsons, S., Zhou, S., Zhang, S., 2008. Wing morphology, echolocation calls, diet and emergence time of black-bearded tomb bats (*Taphozous melanopogon*, *Emballonuridae*) from southwest China. *Acta Chiropterol.* 10 (1), 51–59. <https://doi.org/10.3161/15081108X331081>.

Yoh, N., Azhar, I., Fitzgerald, K.V., Yu, R., Smith-Butler, T., Mahyudin, A., Kingston, T., 2020. Bat Ensembles Differ in Response to Use Zones in a Tropical Biosphere Reserve. *Diversity* 12 (2), 60. <https://doi.org/10.3390/d12020060>.

Zamora-Gutierrez, V., Lopez-Gonzalez, C., MacSwiney Gonzalez, M.C., Fenton, B., Jones, G., Kalko, E.K.V., Puechmaille, S.J., Stathopoulos, V., Jones, K.E., Freckleton, R., 2016. Acoustic identification of Mexican bats based on taxonomic and ecological constraints on call design. *Methods Ecol. Evol.* 7 (9), 1082–1091. <https://doi.org/10.1111/2041-210X.12556>.