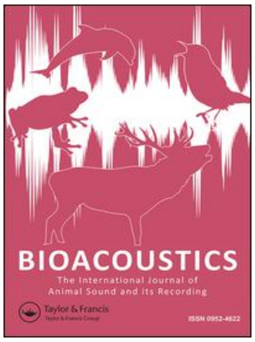

Taylor & Francis
Taylor & Francis Group

Bioacoustics
The International Journal of Animal Sound and its Recording

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tbio20

Random forest is the best species predictor for a community of insectivorous bats inhabiting a mountain ecosystem of central Mexico

Jorge Ayala-Berdon , Kevin I. Medina-Bello , Issachar L. López-Cuamatzi , Rommy Vázquez-Fuerte , M. Cristina MacSwiney G. , Lorena Orozco-Lugo , Ignacio Iñiguez-Dávalos , Antonio Guillén-Servent & Margarita Martínez-Gómez

To cite this article: Jorge Ayala-Berdon , Kevin I. Medina-Bello , Issachar L. López-Cuamatzi , Rommy Vázquez-Fuerte , M. Cristina MacSwiney G. , Lorena Orozco-Lugo , Ignacio Iñiguez-Dávalos , Antonio Guillén-Servent & Margarita Martínez-Gómez (2020): Random forest is the best species predictor for a community of insectivorous bats inhabiting a mountain ecosystem of central Mexico, Bioacoustics, DOI: 10.1080/09524622.2020.1835539

To link to this article: https://doi.org/10.1080/09524622.2020.1835539

Published online: 26 Oct 2020.

BIOACOUSTICS
https://doi.org/10.1080/09524622.2020.1835539

Random forest is the best species predictor for a community of insectivorous bats inhabiting a mountain ecosystem of central Mexico

Jorge Ayala-Berdon [a], Kevin I. Medina-Bello [a], Issachar L. López-Cuamatzi [a], Rommy Vázquez-Fuerte [b], M. Cristina MacSwiney G. [c], Lorena Orozco-Lugo [d], Ignacio Iñiguez-Dávalos [e], Antonio Guillén-Servent [f] and Margarita Martínez-Gómez [a,g]

[a]CONACYT, Universidad Autónoma de Tlaxcala, Tlaxcala, México; [b]Escuela Nacional de Estudios Superiores, Unidad Morelia, Universidad Nacional Autónoma de México, Michoacán, México; [c]Centro de Investigaciones Tropicales, Universidad Veracruzana, Veracruz, México; [d]Centro de Investigación de Biodiversidad y Conservación, Universidad Autónoma del Estado de Morelos, Morelos, México; [e]Departamento de Ecología y Recursos Naturales - IMECBIO, Universidad de Guadalajara, Jalisco, Mexico; [f]Instituto de Ecología, A.C. (INECOL), Veracruz, México; [g]Departamento de Biol. Celular y Fisiología, Universidad Nacional Autónoma de México, México

ABSTRACT

Bats are nocturnal animals that can be identified by recording and analysing quantitatively their echolocation calls. For this task, many studies have used both parametric and non-parametric approximations with a variety of results. This urges the necessity of developing more call libraries, that should be analysed using the different statistical approaches to test their performance. This could be relevant in countries holding high biodiversity where the knowledge of the variation in the call structure among species is still scarce. We constructed and validated a call library from bats inhabiting a mountain ecosystem of central Mexico using the Linear Discriminant Function, Artificial Neural Network and Random Forest approaches. We recorded and analysed 2,325 pulses from 114 individuals and 16 bat species of the families Vespertilionidae, Mormoopidae, Molossidae, and Natalidae. The Random forest model (81.3%) was the better species predictor over the artificial neural network and the discriminant function analysis (69% and 62.1%, respectively). Our work is one of the few attempts to do this exercise that has been conducted in Mexico. The library can be useful as a starting point of research in other regions of the highlands in central Mexico where the information is still scarce.

ARTICLE HISTORY
Received 16 June 2020
Accepted 24 September 2020

KEYWORDS
Bats; call library; classification; central Mexico; echolocation; statistical approaches

Introduction

Monitoring biodiversity is fundamental to understand ecosystem processes, both at regional and global levels (Ahlén and Baagøe 1999; Ochoa et al. 2000; Welsh and Droege 2001). Nevertheless, for some species, this task is challenging because both individual capture and tracking are difficult in the field. Bats are nocturnal flying animals

CONTACT Jorge Ayala-Berdon jorgeayalaberdon@gmail.com

difficult to observe and identify without capturing individuals with the use of mist nets or other trapping devices (Kunz and Kurta 1988; Rydell and Speakman 1995; Speakman 2001; Kunz et al. 2009). Additionally, some species can detect and avoid the nets, and many tend to forage beyond the places where they can be caught. In aerial insectivorous bats, capture instances are fortuitous, and its rates of captures do not reflect well the local abundance of the species (Kalko et al. 1996; O'Farrell 1997; Kingston et al. 2003; MacSwiney G et al. 2008). Luckily, the discovery of bat echolocation, and the development of ultrasound detectors, which can record and store calls, have allowed researchers to detect and identify the bat species by using the ultrasound pulses that animals broadcast when they are flying (Griffin et al. 1960; Fenton and Bell 1981; O'Farrell 1997; Ahlén and Baagøe 1999; O'Farrell and Miller 1999; Britzke et al. 2013; Waters and Gannon 2004; MacSwiney G et al. 2008; among others).

Acoustic identification is usually conducted by analysing the temporal and spectral structure of the pulses produced by bats and classifying them according to the characteristics of reference recordings (Brigham and Cebek 1989; Vaughan et al. 1997; Parsons and Jones 2000; Britzke et al. 2002, 2011; Redgwell et al. 2009). Reference recordings are obtained from well-identified individuals that have been successfully captured and recorded when they are flying as in typical natural conditions. Calls recorded from the wild animals can be assigned to species by comparing them to the reference calls by visual inspection in case of species with idiosyncratic echolocation calls, or by quantitative methods that reduce the bias associated with the researcher performing the identification (Russo and Jones 2002; Waters and Gannon 2004; Parsons and Szewczak 2009). In this regard, many studies have used parametric classification methods to complete the task (Krusic and Neefus 1996; Britzke et al. 2002; Vaughan et al. 1997; Russo and Jones 2002; Biscardi et al. 2004; among others), and some others have explored other non-parametric machine-learning approaches, with different results (Herr et al. 1997; Burnett and Masters 1999; Parsons and Jones 2000; Parsons and Obrist 2000; Broders et al. 2004; Skowronski and Harris 2006; Jennings et al. 2008; Redgwell et al. 2009; among others). Because call parameters tend to vary among individuals of the same species due to age, size, gender, presence of conspecifics and geographical distribution (Brigham and Cebek 1989; Jones et al. 1992; Kalko and Schnitzler 1993; Obrist 1995; Barclay 1999; Kazial et al. 2001; Russo et al. 2001), and the performance of the statistical methods may differ in regard of the algorithms they use (Biscardi et al. 2004), high differences in the probability of accurate identification of species have been reported. This urges the necessity of developing more regional call libraries that should be analysed using different statistical approaches to test their performance. This could be especially relevant in countries with a high diversity of bat species, where monitoring has been traditionally conducted with the use of mist nets (O'Farrell and Miller 1999), and the knowledge of the variation in the call structure among species is still scarce (but see Orozco-Lugo et al. 2013; Rivera-Parra and Burneo 2013; Rodríguez-San Pedro and Simonetti 2013; Zamora-Gutierrez et al. 2016).

Here, we developed and validated a call library from 16 bat species inhabiting a mountain ecosystem of Central Mexico. To do this, we used both parametric (a Discriminant Function

Analysis -DFA-) and two non-parametric (Artificial Neural Network -ANN- and Random Forest -RF-) approaches. Because it has been reported that most of the call parameters of insectivorous bats tend to present a non-parametric structure (Waters and Gannon 2004), we predicted that ANN and RF would have better accuracy for species identification than the DFA.

Material and methods

Study site

Our study was conducted at La Malinche National Park (hereafter LMNP) (19°13′34.08”N, 98° 1′28.92”W; 4100 m a.s.l.), a mountain ecosystem located in Central Mexico (Acosta and Kong 1991). The site is a natural protected area that is mainly composed of crops in the lowlands, and Pine, fir forests, and mountain prairie areas at the middle and high elevations (Villers et al. 2006). Climate is temperate sub-humid with a rainy season in summer (INEGI 1987), and the average annual ambient temperature is 15°C (Lara 2006).

Because we were interested in being able to classify the calls of all species that could be present at the park, we first constructed a potential list of the bats that might occur in the area. To do this, we consulted maps of the species presented by Medellín et al. (2008) and the IUCN List for Threatened Species (https://www.iucnredlist.org). Nevertheless, we just were able to capture half of the species of the list with our mist-netting effort in the study site (see results). So, we completed our library with calls obtained from the nearest localities available from LMNP (Table 1).

Bat captures

We captured bats from 2014 to 2018 as part of continuous monitoring that has been conducted in LMNP, and the other localities present in the vicinities of the study site (Table 1). Bats were captured with 3 or 6 m long and 2 m high mist nets set in the forest, out of the caves when animals were emerging from their roosts or in waterbodies that the bats visited for drinking or foraging. Nets were open at dusk, checked every 20–30 minutes, and closed at ~01:00 am. We also obtained recordings of bats emerging from roosts from whose specific identity we knew after previous inspections. Captured bats were identified to species level with the use of Mexican field guides (Medellín et al. 2008), and their age and reproductive condition were registered. For taxonomic names, we followed Ramírez-Pulido et al. (2014). Body mass was obtained either with the use of an electronic balance to the nearest 0.2 g (Ohaus®) or a spring balance to the nearest 0.5 g. Age of bats (i.e., either juvenile or full-grown) was assessed by checking the presence of the epiphyseal gap of the fourth metacarpal bone of the wings (Kunz et al. 1996). After taking measurements, bats were released at their capture site, and the echolocation calls that the animals broadcasted on their departure were recorded. To collect as much call variation as possible, we recorded both males and females as well as juvenile bats (Britzke et al. 2010, Table 1). Animals were captured and handled under permission of the Mexican Department of Wildlife Management (SEMARNAT 07019, and FAUT-0251 granted to our institutions).

Table 1. Species used for the construction and validation of the call library of bats living at La Malinche National Park, a mountain ecosystem located in central Mexico. We constructed a potential list of 16 species that might occur in the area. Nevertheless, we were just able to capture half of them with the use of mist nets, so, we completed our library with calls obtained from other localities near our study site.

Bat species by family	*(n)*	Sex proportion M/F	Age proportion A/J	Reproductive condition (proportion) R/NR	Locality	Place of capture	Recording technique
Vespertilionidae							
Corynorhinus mexicanus	(10)	6/4	10/0	4/6	LMNP (19º13′34.08″N; 97° 59′25.9″O)	Water body	Hand release
Eptesicus fuscus	(13)	10/3	13/0	0/13	LMNP (19º13′34.08″N; 97°59′25.9″O)	Water body	Hand release
E. furinalis	(5)	5/0	0/5	0/5	(21º04′53.80″N, 87º12′11.84″W)	Water body	Hand release
	–	–	–	–	(21º12′36.65″N, 87º12′33.61″W)	Medium semideciduous forest	Flying freely
	–	–	–	–	(21º13′37.98″N, 88º40′45.34″W)	Water body	Flying freely
	–	–	–	–	(20º54′33.16″N, 88º52′00.87″W)	Water body	Flying freely
	–	–	–	–	(21°17′53.80″N, 88°36′22.08″W)	Pastureland	Flying freely
Lasiurus cinereus	(6)	6/0	6/0	0/6	LMNP (19º13′34.08″N; 97° 59′25.9″O)	Water body	Hand release
L. ega	(1)	1/0	1/0	0/1	na	Within the forest	Zip line
Myotis californicus	(10)	9/1	10/0	0/10	LMNP (19º13′34.08″N; 97° 59′25.9″O)	Water body	Hand release
M. melanorhinus	(8)	7/1	8/0	0/8	LMNP (19º13′34.08″N; 97°59′25.9″O)	Water body	Hand release
M. thysanodes	(3)	3/0	3/0	1/2	LMNP (19º13′34.08″N; 97° 59′25.9″O)	Water body and exit of a cave	Hand release
M. velifer	(10)	5/5	10/0	0/10	LMNP (19º13′34.08″N; 97°59′25.9″O)	Water body	Hand release
M. volans	(14)	3/11	14/0	1/13	LMNP (19º13′34.08″N; 97°59′25.9″O)	Water body	Hand release
Molossidae							
Nyctinomops macrotis	(4)	4/0	4/0	0/4	(18°55′15.93″ N, 99°14′15.60″ O)	Water body	Hand release
Tadarida brasiliensis	(8)	8/0	8/0	4/4	(18°55′15.93″ N, 99°14′15.60″ O)	Within the forest	Hand release

(*Continued*)

Table 1. (Continued).

Bat species by family	*(n)*	Sex proportion M/F	Age proportion A/J	Reproductive condition (proportion) R/NR	Locality	Place of capture	Recording technique
Mormoopidae							
Mormoops megalophylla	(8)	–	–	–	(20°33′03.89″N, 89°54′44.90″W)	Water body and exit of a cave	Flying freely
Pteronotus davyi	(9)	9/0	9/0	0/9	(19°21′34.9″N; 96°41′41.3″O)	Exit of a cave	Hand release
P. parnellii	(4)	4/0	4/0	0/4	(19°21′34.9″N; 96°41′41.3″O)	Exit of a cave	Hand release
Natalidae							
Natalus mexicanus	(1)	na	na	na	(20° 9′31.01″, 103°54′46.48″)	Water body	Zip line

(*n*) = number of bats recorded; M = male; F = female; A = adult; J = juvenile; R = reproductive; NR = non-reproductive; na = no data.

Call records

We recorded search echolocation calls from bats that were released near the places where individuals were captured: a) in a zip-line or b) from the hand. Both methods have demonstrated to be quite effective to record calls to build reference call libraries of insectivorous bats around the world (Szewczak 2004). We also recorded some individuals of *Mormoops megalophylla* and *Eptesicus furinalis* when bats were flying freely (Table 1). The zip-line where the bats were recorded had an extension of ~10 m. When bats were released either in the zip line or from the hand, the person holding an ultrasound detector (either Petterson -models 1000x and D980-, Pettersson Elektronik AB, Uppsala, Sweden, or an Avisoft UltraSoundGate model 116 H; Avisoft Bioacoustics, Glienicke, Germany) was positioned from 10 to 30 m in front of the bats' flying trajectory, and the calls emitted by the animals were recorded. This was similar to the recordings obtained from the free-flying bats, but the person doing the recordings was positioned out of the cave or within the vegetation where animals were flying freely. Calls were digitally recorded at a sampling rate of 300 kHz, which allowed us to record sounds of up to 150 kHz well above the maximum frequency of the calls, including functional harmonics broadcasted by all the species considered in the library (Rydell et al. 2002).

Data analyses

Recordings

Recordings were analysed using the Sonobat® software ver. 3.1.5 (Szewczak 2010). For the analysis, we used the sequences that had the highest signal-to-noise ratio (i.e., those with a quality higher than 85%). We chose 15 variables (3 temporal and 12 spectral) that have been reported in the bibliography to be most useful for call identification of insectivorous bats (Kalko and Schnitzler 1989; Vaughan et al. 1997; Redgwell et al. 2009) (see appendix). In all cases, we measured the harmonic with the most energy. Using parameters universally recognised to identify bat calls may enhance the repeatability of data among researchers (Britzke et al. 2010).

Building and validating the classification tools

For the validation of calls (i.e., the evaluation of the capacity of the algorithms to correctly identify the known calls), we used each pulse as the unit of measurement. To do the acoustical identification, we used three different methods, one parametric: DFA (Sokal and Rohlf 1981); and two non-parametric: ANN (Haykin 1999) and the RF algorithm (Breiman 2001).

We choose these methods because it has been proved that they provide good classification results for insectivorous bats (Britzke et al. 2010; Zamora-Gutierrez et al. 2016). While DFA constructs discriminant functions based on the linear combination of variables that maximises the differences of the featured means to allow predictions (Poulsen and French 2008), NNA is a non-linear adaptive machine-learning algorithm that trains and correct itself to optimise the model to perform the identifications, and the RF algorithm constructs a series of decision trees to predict and classify the variables to make classifications (Cutler et al. 2007).

Because data manipulation previous to analysis tends to be time-consuming, we trained the data for the models in the simplest way we could (i.e., we did not: i) eliminate any correlated variable, and ii) separate the frequency modulated – FM – from the quasi-constant frequency-modulated – QCF-FM – and the constant frequency-modulated – CF-FM – calls from the data set). In this way, the classification tools allowed us to analyse the data once they were obtained. In all models, we randomly assigned 50% of the calls of each bat species for the training data set and evaluated the correct identifications in the remaining 50% of calls. Before being split, data were randomised within species to avoid pseudo-replication due to: 1) the recording method and the place where the animals were recorded, and 2) the data associated with the individuals as their identity, sex, or age. The ANN had a very simple structure and consisted of three layers: 1) a layer of 15 node inputs (i.e., the call variables we chose, see appendix), 2) a hidden layer with 16 nodes, and 3) an output layer with 16 nodes that corresponded to the number of bat species we intended to classify. For this model, we used a preset decay value of 0.001. We chose 16 nodes in the hidden layer because although there is not a clear rule for assigning this parameter, it is highly recommended that the number of nodes would be between the nodes of the input and the output layers (Samarasinghe 2007). The RF classifier consisted of 500 trees, and the number of variables tried at each split was three. All analyses were performed in R ver 3.5.0 (function *lda* from the *Car* library for the DFA, function *nnet* from the *nnet* library for the ANN analysis, and function *RandomForest* from the *RandomForest* library for the RF analysis) (R Core Team 2018). All values presented in the results section are showed as means with their respective standard deviation unless noted otherwise.

Evaluation of the models' performance

We evaluated the models' performance by calculating the receiver operating characteristic curve (ROC) for each predictive algorithm. ROC is a graphical representation (where true positive rate -TPR- is plotted in the Y-axis and false-positive rate -FPR- is plotted in the X-axis) which helps to illustrate the diagnostic capacity of classifiers (Fawcett 2006). ROC's have been used in diverse scientific fields of science to evaluate and compare models (e.g., Bradley 1997; Goldbaum et al. 2002; Hobson et al. 2014). For binary classifiers, ROC is represented by a single point in the ROC space (Fawcett 2006). A perfect classifier is that placed in the coordinates 0,1 of the ROC space. This point represents a model with both no false negatives and no false positives (Fawcett 2006).

To calculate the ROC curve, we first generated the confusion matrix outputted from each model. We then counted the total number of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). We calculated TPR and FPR according to Fawcett (2006) as:

$$TPR = \frac{TP}{(TP + FN)}$$

where TP and FN are the total scores by the model true positives and false negatives, respectively, and:

$$FPR = \frac{FP}{(FP + TN)}$$

where FP and TN are the total scores by the model of false positives and true negatives, respectively. We also calculated the area under the curve (AUC) from each model. Data were obtained from the confusion matrix we previously obtained from each model. The AUC is an effective and combined measure of TPR and FPR that describes the inherent validity of classifying models (Bradley 1997). Maximum AUC = 1 means that the diagnostic test is perfect in the differentiation between the correct and incorrect classifications. AUC = 0.5 means that classifications occur by a random process, while AUC = 0 indicates incorrect classifications in all subjects (Bradley 1997). In binary models, the use of AUC has been controversial because the linear function calculated by common software tends to overestimate the AUC. To overcome this overestimation, 1) we used the step function interpolation to generate the ROC curve, and 2) calculated the AUC manually following Muschelli (2019). AUC was estimated using the formula:

$$AUC = (TPR)(1 - FPR)$$

Results

We obtained a total of 2,325 pulses from 114 individuals and 16 bat species of the families Vespertilionidae, Mormoopidae, Molossidae, and Natalidae. Eight species were captured in LMNP and eight of them caught nearby the study site. We recorded both males and females as well as juvenile bats. 12 species were recorded with the hand released method, two with the zip-line technique, and two when bats were flying-freely (Table 1).

Description of the echolocation calls

Bats presented differences in the parameters of their echolocation calls. These parameters were highly variable, where the start frequency was the most variable feature, while the mean of the third quartile amplitude was the feature that varied less (Table 2). Bats of the families Vespertilionidae (i.e., those of the *Myotis, Eptesicus, Corynorhinus* and *Lasiurus* genera) (n = 10 spp.), Molossidae (i.e., bats of *the Nyctinomops* and *Tadarida* genera) (n = 2 spp.), and Natalidae (i.e., bats of the genus *Natalus*) (n = 1 spp.) presented typically FM echolocation calls (Figure 1). In this mode of echolocation, bats of the *Myotis* genus presented calls that were very much alike. On the other hand, bats of the family Mormoopidae showed QC-FM calls in bats of the genus *Mormoops* (n = 1 spp.), and CF-FM calls (n = 2 spp.) in bats of the *Pteronotus* genus (Figure 1). Except for *N. macrotis* which presented less pronounced echolocation calls, the rest of the FM echolocators showed steep downward echolocation pulses that varied from 2.8 ± 0.7 ms in *M. velifer* to 9.0 ± 2.6 ms in *N. macrotis*. The rest of the bats which showed a mixture of components in their pulses (i.e., either QC-FM or CF-FM) exhibited more even calls, where *M. megalophylla* presented the shortest call duration (4.1 ± 1.2 ms) while *P. parnellii* displayed the longest one (21.2 ± 5.4 ms). Finally, *Nyctinomops macrotis* and *N. mexicanus* presented the minimum and maximum values in the lowest frequency (17.2 ± 5.0 kHz and 77.6 ± 22.0 kHz, respectively), as well as the frequency characteristic (18.4 ± 5.1 kHz and 80.4 ± 23.0 kHz, respectively) from all the bats measured in the call library (Table 2).

Table 2. Parameter values from the variables used to construct and validate a call library of 16 bat species inhabiting La Malinche National Park, a mountain ecosystem of central Mexico. To do the acoustical identification, we used three different methods, one parametric: discriminant function analysis; and two non-parametric: neural network and the random forest algorithm. Data are presented as mean ± SD. See appendix one for the meaning of each variable.

Bat species by family	Call variable													
(n)	*Dc*	*Fc*	*Hf*	*Lf*	*Efm*	*Fmp*	*Pma*	*Sf*	*Ef*	*Pc*	*Sc*	*Tc*	*Cc*	*cft*	*Casl*
Vespertilionidae															
Come (305)	4.6	29.0	49.1 ± 5.7	26.5	22.6 ± 3.5	37.1 ± 5.1	49.5	49.1	26.5 ± 5.3	0.8	0.9	0.9	0.7	2.6±	5.4
	±1.7	±6.4		±5.3			±15.6	±5.7		±0.1	±0.1	±0.1	±0.1	4.9	±5.9
Epfu (261)	5.5	31.4	61.1	29.5	31.6 ± 9.5	35.2	62.8	61.0	29.7 ± 3.4	0.6	0.8	1.0	0.8	1.6	6.3
	±2.0	±4.0	±10.2	±2.9		±4.9	±14.2	±10.3		±0.1	±0.1	±0.1	±0.1	±1.2	±2.6
Epfi (54)	6.6	37.6	52.8	37.2	15.6 ± 6.4	38.9 ± 1.4	58.0	52.8	37.7 ± 1.6	0.8	0.9	1.0	0.9	1.0	2.5
	±1.5	±1.4	±6.9	±1.4			±15.1	±6.9		±0.1	±0.0	±0.0	±0.1	±0.6	±1.2
Laci (177)	4.3	26.9	50.1	24.9	25.2	29.8 ± 3.8	58.9	50.1	25.2 ± 2.4	0.7	0.8	0.9	0.8	9.8	14.6
	±2.2	±2.8	±11.1	±2.3	±10.1		±16.6	±11.1		±0.2	±0.1	±0.0	±0.1	±8.1	±7.6
Laeg (32)	4.7	32.2	58.5	30.0	28.5	36.5 ± 2.0	55.8	58.5	30.0	0.7	0.1	0.1	0.8	1.9	6.4
	±1.0	±1.5	±3.8	±0.8	±3.8		±10.3	±3.8	±0.8	±0.1	±0.1	±0.0	±0.1	±0.9	±1.6
Myca (234)	3.2	47.5	87.9	42.6	45.3	52.2	66.3	87.9	42.6	0.6	0.8	0.9	0.8	8.1	15.0
	±1.0	±4.7	±14.3	±5.0	±12.2	±7.2	±12.8	±14.3	±5.0	±0.1	±0.1	±0.1	±0.1	±3.4	±4.9
Myme (239)	3.3	44.1	78.9	38.3	40.7	48.6	62.2	78.9	38.2	0.5	0.8	0.9	0.8	6.2	13.1
	±1.0	±3.4	±10.5	±3.7	±10.8	±4.4	±12.4	±10.5	±3.7	±0.1	±0.1	±0.1	±0.1	±4.6	±4.0
Myth (77)	3.0	26.3	66.1	19.7	46.4	30.9	70.5	66.1	19.7	0.6	0.8	0.9	0.8	8.1	30.8 ± 19.6
	±1.4	±12.2	±13.4	±2.4	±13.8	±5.1	±13.0	±13.4	±2.4	±0.1	±0.1	±0.1	±0.1	±9.9	
Myve(242)	2.8	43.0	51.9	39.3	42.6	48.5	67.3	81.9	39.3	0.6	0.8	0.9	0.8	7.8	15.6
	±0.7	±3.2	±13.9	±2.7	±13.1	±5.2	±14.2	±13.9	±2.8	±0.1	±0.1	±0.1	±0.1	±5.6	±4.8
Myvo (420)	4.0	41.1	76.4	36.7	39.6	45.6	67.2	76.4	36.8	0.6	0.7	0.9	0.8	4.8	11.3
	±1.4	±4.7	±14.3	±4.1	±12.4	±6.6	±13.1	±14.3	±4.1	±0.1	±0.2	±0.1	±0.1	±5.3	±5.3
Molossidae															
Nyma	9.0	18.4	24.1	17.2	6.9	20.0	40.2	24.1	17.3	0.9	1.0	0.9	0.8	0.2	0.8
(19)	±2.6	±5.1	±4.6	±5.0	±3.1	±4.7	±13.1	±4.6	±5.0	±0.1	±0.0	±0.0	±0.1	±0.1	±0.4
Tabr	5.3	26.0	43.3	22.6	20.7	51.8	43.3	22.6	0.8	0.8	0.9	0.9	0.8	1.6	4.2
(67)	±1.6	±2.6	±6.6	±2.1	±6.9	±6.1	±19.5	±6.6	±2.1	±0.1	±0.1	±0.0	±0.1	±1.6	±1.6

(*Continued*)

Table 2. (Continued).

Bat species by family	Call variable													
(n)	*Dc*	*Fc*	*Hf*	*Lf*	*Efm*	*Fmp*	*Pma*	*Sf*	*Ef*	*Pc*	*Sc*	*Tc*	*Cc*	*cft*	*Casl*
Mormoopidae															
Mome (35)	4.1	44.7	50.2	37.0	13.3	44.4	65.5	50.2	37.0	0.8	0.9	0.9	0.9	3.3	3.6
	±1.2	±10.3	±10.3	±9.0	±4.8	±9.9	±19.3	±10.3	±9.0	±0.1	±0.1	±0.1	±0.1	±1.9	±2.0
Ptda	5.7	54.3	65.8	52.2	13.5	58.6	60.0	65.7	52.3	0.8	0.9	0.9	0.9	2.6	2.7
(48)	±1.6	±11.3	±13.0	±10.8	±3.1	±12.7	±20.0	±12.9	±10.8	±0.1	±0.1	±0.1	±0.1	±2.3	±1.4
Ptpa (94)	21.2	63.5 ± 1.6	64.2	52.7	11.4	61.4	65.3	62.9	52.8	0.8	0.9	0.9	0.9	0.1	0.6
	±5.4		±1.6	±2.2	±2.7	±3.7	±19.6	±2.6	±2.2	±0.1	±0.1	±0.1	±0.0	±0.1	±0.3
Natalidae															
Name (21)	5.1	80.4	121.6 ± 36.0	77.6	44.0	94.7	55.5	121.2	77.6	0.7	0.9	0.9	0.9	4.4	9.0
	±1.0	±23.0		±22.0	±14.5	±21.1	±19.1	±36.5	±22.0	±0.1	±0.1	±0.1	±0.1	±3.5	±4.9

(n) = total number of calls obtained per bat species, Dc = call duration, Fc = frequency characteristic, Hf = highest frequency, Lf = lowest Frequency, Efm = Maximum frequency extension, Fmp = Frequency of maximum call amplitude, Pma = Percentage of the duration of the call in which the maximum amplitude occurs, Sf = Start frequency, Ef = End frequency, Pc = Mean of the first quartile amplitude, Sc = Mean of the second quartile amplitude, Tc = Mean of the third quartile amplitude, Cc = Mean of the fourth quartile amplitude.

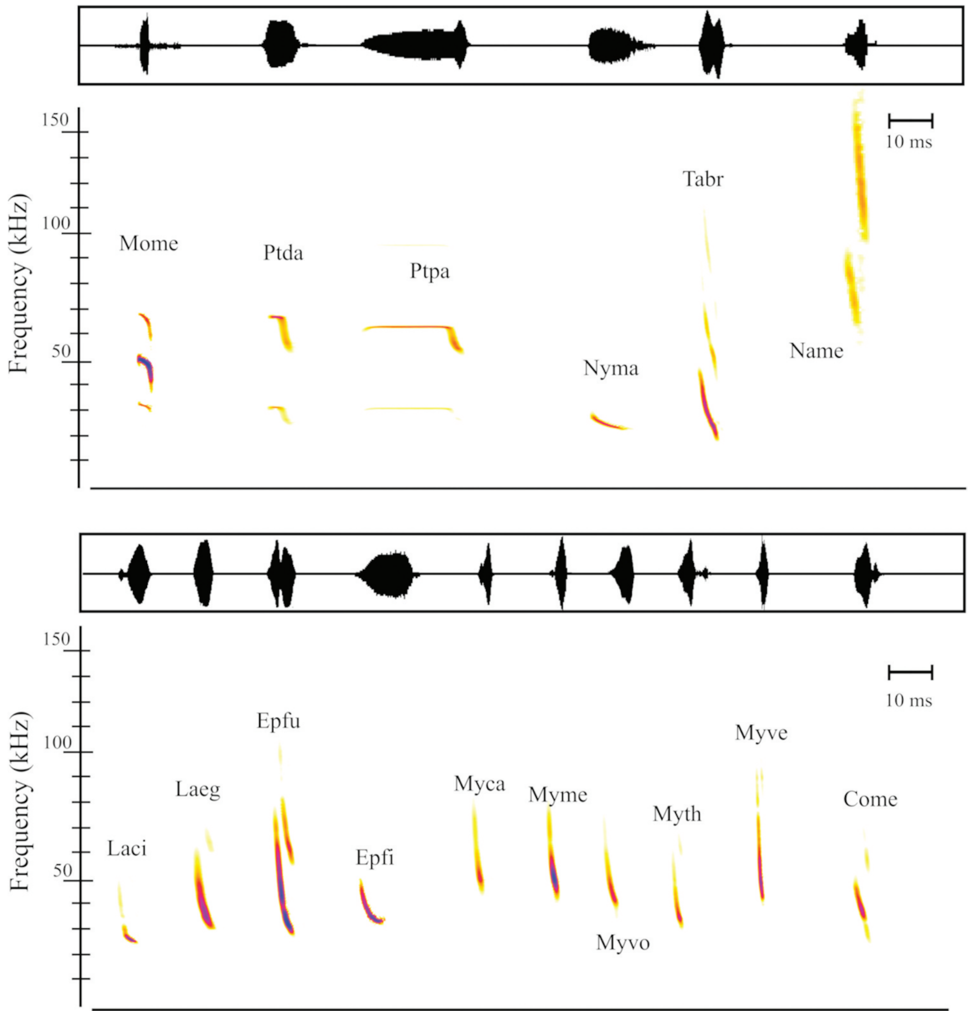

Figure 1. Spectrogram of the echolocation calls of 16 insectivorous bat species composing the community of La Malinche National Park, a mountain ecosystem of central Mexico. Bats of the families Vespertilionidae: Come (*Corynorhinus mexicanus*), Epfu (*Eptesicus fuscus*), Epfi (*E. furinalis*), Laci (*Lasiurus cinereus*), Laeg (*L. ega*), Myca (*Myotis californicus*), Myme (*M. melanorhinus*), Myth *(M. thysanodes*), Myve (*M. velifer*) and Myvo (*M. volans*); Molossidae: Nyma (*Nyctinomops macrotis*) and Tabra (*Tadarida brasiliensis*); and Natalidae: Name (*Natalus mexicanus*) presented FM echolocation calls, while bats of the family Mormoopidae: Mome (*Mormoops megalophylla*) Ptda (*Pteronotus davyi*) and Ptpa (*Pteronotus parnellii*) showed QC-FM and CF-FM echolocation calls.

Performance of models in species identification

For the development of the models to evaluate the species identification, we obtained a variable number of pulses which ranged from 19 from *N. macrotis* to 420 from *E. fuscus* (Table 3). In the LDA model, the three first linear discriminants explained 82.68% of the total variation, while the most explainable variables where call duration, the lowest frequency, and the end of the frequency. On the other hand, the mean of the second,

Table 3. Percentage of correct identifications of the three different statistical approaches of bats inhabiting LMNP, a mountain ecosystem of central Mexico. In all approaches, we used 50% of the calls to train the models and the remaining 50% to make classifications.

Bat species by family	*(n)*	DFA (%)	ANN (%)	RF (%)
Vespertilionidae				
Corynorhinus mexicanus	(305)	83.4	84.8	89.5
Eptesicus fuscus	(261)	77.4	77.7	85.1
E. furinalis	(54)	73.1	92.8	88.4
Lasiurus cinereus	(177)	66.3	83.7	93.4
L. ega	(32)	5	26.3	38.4
Myotis californicus	(234)	47.5	72.8	70
M. melanorhinus	(239)	38.2	34.1	59
M. thysanodes	(77)	66.7	54.7	85.7
M. velifer	(242)	39.1	44.9	69
M. volans	(420)	61.7	63.6	78
Molossidae				
Nyctinomops macrotis	(19)	100	75	90.9
Tadarida brasiliensis	(67)	71.9	50	66.6
Mormoopidae				
Mormoops megalophylla	(35)	56.3	84.2	93.7
Pteronotus davyi	(48)	86.2	67.8	95
P. parnellii	(94)	90.7	100	98
Natalidae				
Natalus mexicanus	(21)	30	90.9	100
Overall accuracy		62.1	69	81.3

(n) = total number of calls obtained per bat species

third, and fourth quartile amplitude, and the frequency characteristic, the lowest frequency, and the end frequency were the most explainable variables for ANN and the RF approaches, respectively.

The three algorithms we used performed differently. In all models *L. ega* presented the lowest percentage of correct identifications, nevertheless, this value was lower in the DFA than the ANN and the RF algorithms tested (5 %, 26.3, and 38.4% respectively). The DFA identified 100% correctly *L. ega*, while the ANN and the RF did it in *P. parnellii* and *N. mexicanus* respectively (Table 3). Finally, the overall percentage of correct classification differed among the three methods tested, where the RF model (81.3% of correct classifications) was the best species predictor over the ANN and the DFA approaches (69% and 62.1 of overall correct classifications, respectively) (Table 3). This same pattern was shown by the ROC and the AUC curves (Figure 2), where the single cut point of the RF model was the closest to the coordinate 0,1 and with the higher AUC value (0.78) respect to ANN (AUC = 0.63) and DFA (AUC = 0.62) respectively.

Discussion

Here we present the structure and the statistical validation (which was addressed with the use of three different statistical approximations), of the echolocation calls collected from a community of insectivorous bats inhabiting a National Park located in Central Mexico. Although many authors have compared a variety of statistical approaches to identify the bats in multiple call libraries collected around the world (e.g., Parsons and Jones 2000;

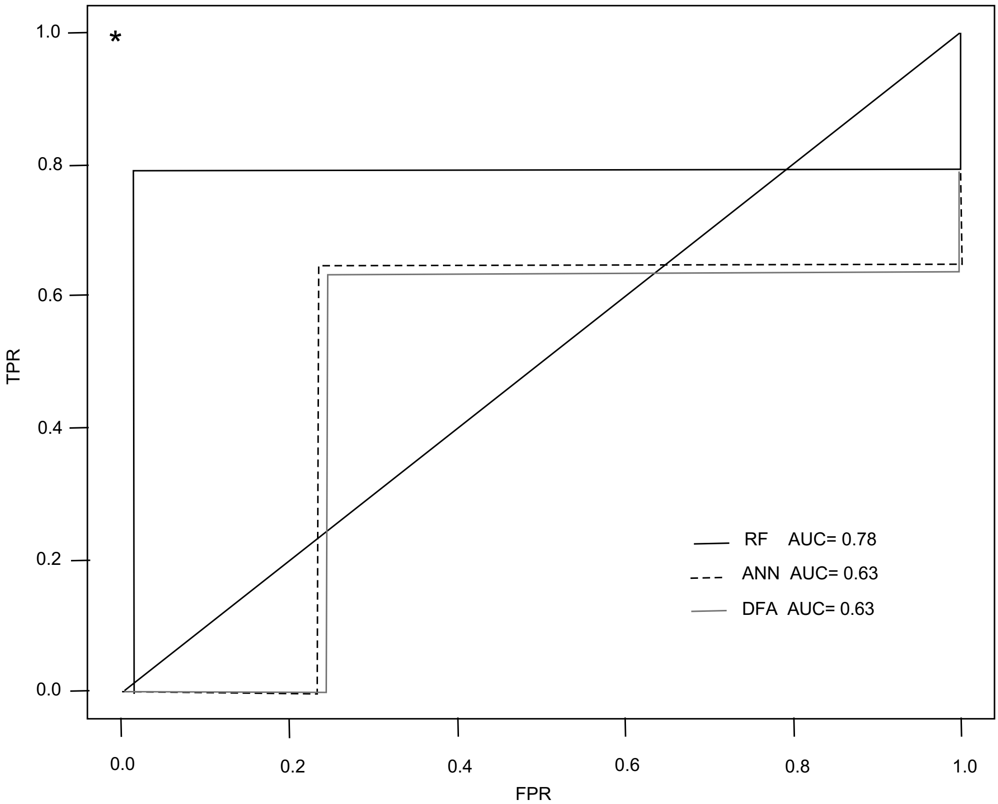

Figure 2. Receiver operating characteristic (ROC) and area under the curve (AUC) to evaluate the diagnostic capacity of Discriminant Function Analysis (DFA), Artificial Neural Network (ANN), and Random Forest (RF) to evaluate the capacity of the algorithms to correctly identify the calls of bats inhabiting La Malinche National Park, a montane ecosystem of central Mexico. The black star indicates a perfect classification model with values of false-positive rate = 0 and true-positive rate = 1. Black diagonal line shows the ROC space when classification follows a random process (AUC = 0.50).

Biscardi et al. 2004; Armitage and Ober 2010; Britzke et al. 2011), to our knowledge, this is the first attempt to do this exercise that has been conducted in Mexico. This information is useful for supporting acoustic monitoring in LMNP, and other areas nearby, as a complementary method to the traditional mist-netting (Ayala-Berdon et al. 2017). The library can be useful as a starting point of research in other regions of the highlands in central Mexico, where the information is still scarce.

Because the compilation of call libraries is aimed at the acoustic identification of bats detected in the wild habitat, usually for monitoring purposes, the correct classification of the echolocation pulses recorded from the animals is crucial. Several investigations have suggested that the success of acoustical identification of bat species is determined by: 1) the number of species present in a given ecosystem, and how similar they are in terms of their morphology and the structure of their echolocation pulses, and 2) the way that data is constructed and analysed quantitatively (Vaughan et al. 1997; Parsons and Jones 2000; Russo and Jones 2002; Britzke et al. 2002; Biscardi et al. 2004; Brigham et al. 2004; Redgwell et al. 2009). We assess these assumptions with the results we obtained below.

Community composition of bats at LMNP

In this work, we found that LMNP may hold up to 16 insectivorous bat species belonging to the families Vespertilionidae, Mormoopidae, Molossidae, and Natalidae. Temperate forests are characterised by having a high dominance of insectivorous bat species (Patten 2004). These species tend to show certain foraging strategies that have moulded their morphological and echolocation traits (Norberg and Rayner 1987). For example, the species which tend to forage within vegetation or along the borders normally present short and wide wings and emit FM pulses, while bats that usually forage in open spaces have long narrow wings and emit CF or a mixture of components in their echolocation calls (Norberg and Rayner 1987; Denzinger and Schnitzler 2013). In our study site, 75% of the species belonged to the families Vespertilionidae and Molossidae (Table 1), which have wide and narrow wings and emit FM echolocation pulses (Barrios-Gómez et al. 2019). Several investigations have shown that the number of bat species and the way they echolocate could affect the performance of the call libraries. For example, Biscardi et al. (2004) and Russo and Jones (2002) found that while CF echolocators are easy to identify, the misclassifications tend to be high in bats emitting FM calls. According to these findings, in this work, we found that the percentage of accuracy of identification for the CF-FM echolocators (i.e., *P.davyi* and *P. mexicanus*) was high (from 84.2 to 98% for the whole algorithms tested) (Table 3). Conversely, we found high variability in the percentage of correct classifications (from 5 to 100%) among bats emitting FM pulses (Table 3).

In our models, some species showed a low accuracy in the classification performed by the three algorithms used. For example, *L. ega* presented the lowest percentages of call identifications in the DFA, ANN and RF models (i.e., from 5 to 38%), while *N. mexicanus* and *T. brasiliensis* showed 30% and 50% of accuracy in the DFA and ANN approaches, respectively (Table 3). For bats of the *Myotis* genus, the percentage of correct classification ranged from 34.1 to 85.7% for all tested models. One possible explanation of these results may rely on the low sample size we obtained from these species, which would have reduced the estimates of the confidence of the models. Additionally, it has been reported that these species tend to show high plasticity in their echolocation pulses. For example, *L. ega* normally displays a short narrow-band tail in its echolocation calls when flying near obstacles, but the tail becomes longer when the bats fly in open spaces (Rydell et al. 2002). On the other hand, it has been reported that *N. mexicanus* generally emit calls at very variable intervals, which are useful to animals either to catch prey airborne or in the surface of vegetation (Rydell et al. 2002). This is similar to what it has been found for *T. brasiliensis*, which can modify their echolocation calls (from FC to FM) depending on the structure of its environment, the presence of conspecifics, or when animals confront environmental noise (Gillam and McCracken 2007). In regard of the *Myotis* genus, several studies have found that the calls of the species comprising this group are very similar (Vaughan et al. 1997; Russo and Jones 2002). We believe that this is a consequence of their relatively recent diversification (Ruedi et al. 2013); and the close relatedness and similar ecology among the species (Parsons and Jones 2000). Then, although we standardised the recordings to increase the success of the algorithms tested, the plasticity of echolocation calls and the relatedness among the species could have affected the percentage of correct classification we found in our call library.

Accuracy of the algorithms used to identify the bat species at LMNP

In this work, we found that the RF was the most successful algorithm for the species classification followed by the ANN and the DFA approximations. Similar results have been observed for other researchers in the past (e.g., Herr et al. 1997; Armitage and Ober 2010; Britzke et al. 2010; Nuñez et al. 2018). The success of non-parametric over parametric methods may be related to the non-parametric structure of the variables we measured in the calls (Britzke et al. 2010). Although it has been reported that DFA is quite robust to departures from normality (Mardia et al. 1994), it seems that these departures work well in low sample sizes (i.e., less than eight species according to Armitage and Ober 2010). However, the performance of DFA can be lower in cases with more classification groups, as in our study. Additionally, our models included some call parameters that were highly correlated (Hair et al. 2006). This would have added bias rather than increase the performance of the discriminant power of the analysis (Russo and Jones 2002; Armitage and Ober 2010). In this sense, it has been reported that RF has a better power to handle correlated variables even better than ANN (Armitage and Ober 2010). Both ANN and RF are non-parametric machine-learning methods that have been observed performing optimally in the identification of insectivorous bats (Veelenturf 1995; Breiman 2001; Archer and Kimes 2007; Samarasinghe 2007). Artificial Neural Networks, by one hand, can be taught to recognise patterns of the structure from the input data and they could minimise the errors caused by misclassifications using the back-propagation algorithm (White 1992). This characteristic improves the ability of the model to make better predictions over the unknown calls (Parsons and Jones 2000). Nevertheless, while ANN and RF are two non-parametric methods that are designed to deal with large samples, ANN tends to be sensitive to imbalanced data, while RF is not (Chen et al. 2004). Here, the number of pulses varied from 420 calls from *E. fuscus* to 19 in *N. macrotis*, and this could explain the lower accuracy showed by the ANN compared with the RF algorithm. In this condition, balancing data (either by weighting or resampling by over-sampling or under-sampling the data) is recommended (Buda et al. 2018). Nevertheless, this manipulation could be avoided by the use of algorithms, as the RF, that can handle such imbalances (Chen et al. 2004).

In our library, we found that the overall accuracy of species identification from the different methods used ranged from 62 to 81% (Table 3). These percentages may seem to be low compared with those that have been shown by other studies (e.g., MacSwiney G et al. 2008; Walters et al. 2012; Rodríguez-San Pedro and Simonetti 2013). Nevertheless, in our study, we included calls from bats of different sexes, gender, and locations. We also performed the analyses in the simplest way we could, this is, we did not manipulate the data previous to the analyses, we did not eliminate any correlated variable and we did not separate the FM from the CF-FM calls from the data set. This gave our call library the advantage that it can be used in an automatic mode just immediately after the calls have been obtained. This could give an advantage to the researchers using the library, because the identification of bat species may be conducted once that the monitoring has been performed. Finally, we propose the development of more call libraries that should be evaluated to test their performance, especially in those places as the mountains were the information is still scarce. This may enhance the understanding of the bat fauna composing the ecosystems, where the monitoring has been traditionally performed with the use of mist nets.

Acknowledgements

We thank the program *Por Amor al Planeta* 2013, granted by the Volkswagen Company to MMG for financing the fieldwork, and La Malinche Biological station for logistical support. We also thank A. Soto, N Rodríguez, and all the students, partially or fully involved in the development of this project.

Disclosure statement

The authors declare they do not have a conflict of interest.

Ethical statement

The authors declare that the study followed the institutional and national ethical guidelines for scientific research in the sites where the research was conducted.

Funding

This work was supported by the Por Amor al Planeta 2013 granted by the Volkswagen Company to MMG.

ORCID

Jorge Ayala-Berdon http://orcid.org/0000-0003-2344-1565
M. Cristina MacSwiney G. http://orcid.org/0000-0002-9007-4622
Ignacio Iñiguez-Dávalos http://orcid.org/0000-0002-9559-4950
Margarita Martínez-Gómez http://orcid.org/0000-0002-3534-1265

References

Acosta R, Kong A. 1991. Guía de las excursiones botánicas y micológicas al Cerro El Peñón y Cañada Grande del estado de Tlaxcala. Mexico: Gobierno del estado de Tlaxcala.

Ahlén I, Baagøe HJ. 1999. Use of ultrasound detectors for bat studies in Europe: experiences from field identification, surveys, and monitoring. Acta Chiropt. 1:137–150.

Archer KJ, Kimes RV. 2007. Empirical characterization of random forest variable importance measures. Comput Stat Data Anal. 52:2249–2260. doi:10.1016/j.csda.2007.08.015.

Armitage DW, Ober HK. 2010. A comparison of supervised learning techniques in the classification of bat echolocation calls. Ecol Inform. 5:465–473. doi:10.1016/j.ecoinf.2010.08.001.

Ayala-Berdon J, Vázquez-Fuerte R, Rodríguez-Peña N, Martínez-Gómez. 2017. Bat fauna associated with artificial ponds in La Malinche National Park, a mountain ecosystem of Mexico. Mammalia. 81:573–581. doi:10.1515/mammalia-2016-0055.

Barclay RM. 1999. Bats are not birds—a cautionary note on using echolocation calls to identify bats: a comment. J Mammal. 80:290–296. doi:10.2307/1383229.

Barrios-Gómez KM, López-Wilchis R, Díaz-Larrea J, Guevara-Chumacero LM. 2019. Spatial distribution of bat richness in Mexico at different taxonomic levels: biogeographical and conservation implications. Therya. 10:11–23. doi:10.12933/therya-19-611.

Biscardi S, Orprecio J, Tsoar A, Ratcliffe JM. 2004. Data, sample sizes and statistics affect the recognition of species of bats by their echolocation calls. Acta Chiropterol. 6:347–363. doi:10.3161/001.006.0212.

Bradley AP. 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit. 30:1145–1159.

Breiman L. 2001. Random Forests. Mach Learn. 45:5–32. doi:10.1023/A:1010933404324.

Brigham RM, Cebek JE. 1989. Intraspecific variation in the echolocation calls of two species of insectivorous bats. J Mammal. 70:426–428. doi:10.2307/1381534.

Brigham RM, Kalko EKV, Jones G, Parsons S, Limpens HJGA. 2004. Bat echolocation research: tools, techniques and analysis. Austin (Texas): Bat Conservation International.

Britzke ER, Murray KL, Heywood JS, Robbins LW, Kurta A, Kennedy J. 2002. Acoustic identification. In: Kurta A, Kennedy J, editors. The Indiana bat: biology and management of an endangered species. Austin (TX): Bat Conservation International; p. 221–225.

Britzke ER, Duchamp JE, Murray KL, Swihart RK, Robbins LW. 2011. Acoustic identification of bats in the eastern United States: a comparison of parametric and nonparametric methods. J Wildl Manage. 75:660–667. doi:10.1002/jwmg.68.

Britzke ER, Gillam EH, Murray KL. 2013. Current state of understanding of ultrasonic detectors for the study of bat ecology. Acta Theriol. 58:109–117. doi:10.1007/s13364-013-0131-3.

Britzke ER, Slack BA, Armstrong MP, Loeb C. 2010. Effects of orientation and weatherproofing on the detection of bat echolocation calls. J Fish Wild Manage. 1:136–141. doi:10.3996/072010-JFWM-025.

Broders H, Findlay C, Zheng L. 2004. Effects of clutter on echolocation call structure of *Myotis septentrionalis* and *M. lucifugus*. J Mammal. 85:273–281. doi:10.1644/BWG-102.

Buda M, Maki A, Mazurowski MA. 2018. A systematic study of the class imbalance problem in convolutional neural networks. Neural Netw. 106:249–259. doi:10.1016/j.neunet.2018.07.011.

Burnett SC, Masters WM. 1999. The use of neural networks to classify echolocation calls of bats. J Acoust Soc Am. 106:2189. doi:10.1121/1.427419.

Chen C, Liaw A, Breiman L. 2004. Using random forest to learn imbalanced data. Technical Report. Berkeley (CA): University of California.

Cutler DR, Edwards TC Jr, Beard K, Cutler A, Hess KT, Gibson J, Lawler JJ. 2007. Random forests for classification in ecology. Ecology. 88:2783–2792. doi:10.1890/07-0539.1.

Denzinger A, Schnitzler HU. 2013. Bat guilds, a concept to classify the highly diverse foraging and echolocation behaviors of microchiropteran bats. Front Physiol. 4:164.

Fawcett T. 2006. An introduction to ROC analysis. Pattern Recognit Lett. 27:861–874. doi:10.1016/j.patrec.2005.10.010.

Fenton MB, Bell GP. 1981. Recognitionofspeciesofinsectivorous bats by their echolocation calls. J Mammal. 62:233–243. doi:10.2307/1380701.

Gillam EH, McCracken GF. 2007. Variability in the echolocation of Tadarida brasiliensis: effects of geography and local acoustic environment. Anim Behav. 74:277–286. doi:10.1016/j.anbehav.2006.12.006.

Goldbaum MH, Sample PA, Chan K, Williams J, Lee TW, Blumenthal E, Girkin CA, Zangwill LM, Bowd C, Sejnowski T, et al. 2002. Comparing machine learning classifiers for diagnosing glaucoma from standard automated perimetry. Invest Ophthalmol Vis Sci. 43:162–169.

Griffin DR, Webster FA, Michael CR. 1960. The echolocation of flying insects by bats. Animal Behav. 8:141–151.

Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL. 2006. Multivariate data analysis. Upper Saddle River (New Jersey): Pearson Education Inc.

Haykin S. 1999. Neural networks. New Jersey: Prentice-Hall.

Herr A, Klomp NI, Atkinson JS. 1997. Identification of bat echolocation calls using a decision tree classification system. Comput Complex. 4:1–9.

Hobson M, Graff P, Feroz F, Lasenby A. 2014. Machine-learning in astronomy. Proc Int Astronom Union. 10:279–287. doi:10.1017/S1743921314013672.

INEGI. 1987. Anexo cartográfico del estado de Tlaxcala. México: Instituto Nacional de Estadística y Geografía.

Jennings N, Parsons S, Pocock MJO. 2008. Human vs. Machine: identification of bat species from their echolocation calls by humans and by artificial neural networks. Can J Zool. 86:371–377. doi:10.1139/Z08-009.
Jones G, Gordon T, Nightingale J. 1992. Sex and age differences in the echolocation calls of the lesser horseshoe bat, *Rhinolophus hipposideros*. Mammalia. 56:189–193. doi:10.1515/mamm-1992-0202.
Kalko EK, Schnitzler HU. 1989. The echolocation and hunting behavior of Daubenton's bat, *Myotis daubentoni*. Behav Ecol Sociobiol. 24:225–238. doi:10.1007/BF00295202.
Kalko EK, Schnitzler HU. 1993. Plasticity in echolocation signals of European pipistrelle bats in search flight: implications for habitat use and prey detection. Behav Ecol Sociobiol. 33:415–428. doi:10.1007/BF00170257.
Kalko EKV, Handley CO Jr., Handley D. 1996. Organization, diversity, and long-term dynamics of a Neotropical bat community. In: Cody M, Smallwood J, editors. Long-term studies in vertebrate communities. Los Angeles: Academic Press; p. 503–553.
Kazial KA, Burnett SC, Masters WM. 2001. Individual and group variation in echolocation calls of big brown bats, *Eptesicus fuscus* (Chiroptera: Vespertilionidae). J Mammal. 82:339–351. doi:10.1644/1545-1542(2001)082<0339:IAGVIE>2.0.CO;2.
Kingston T, Francis CM, Akbar Z, Kunz TH. 2003. Species richness in an insectivorous bat assemblage from Malaysia. J Trop Ecol. 19:1–12. doi:10.1017/S0266467403003080.
Krusic RA, Neefus CD. 1996. Habitat associations of bat species in the White Mountain National Forest. In: Barclay RMR, Brigham RM, editors. Bats and forest symposium. Victoria (Canada): British Columbia Ministry of Forests; p. 185–198.
Kunz TH, Kurta A. 1988. Methods of capturing and holding bats. In: Kunz TH, editor. Ecological and behavioral methods for the study of bats. Washington (EE.UU): Smithsonian Institution Press; p. 1–30.
Kunz TH, Wemmer C, Hayssen V. 1996. Sex, age, and reproductive condition of mammals. In: Wilson DE, Cole FR, Nichols JD, Rudran R, Foster MS, editors. Measuring and monitoring biological diversity: standard methods for mammals. Washington (DC): Smithsonian Institution Press; p. 279–290.
Kunz TH, Hodgkison R, Weise CD. 2009. Methods of capturing and handling bats. In: Kunz TH, Parsons S, editors. Ecological and behavioral methods for the study of bats. Baltimore (MD): The Johns Hopkins University Press; p. 3–35.
Lara C. 2006. Temporal dynamics of flower use by hummingbirds in a highland temperate forest in Mexico. Ecoscience. 13:23–29. doi:10.2980/1195-6860(2006)13[23:TDOFUB]2.0.CO;2.
MacSwiney G MC, Clarke FM, Racey PA. 2008. What you see is not what you get: the role of ultrasonic detectors in increasing inventory completeness in Neotropical bat assemblages. J Appl Ecol. 45:1364–1371. doi:10.1111/j.1365-2664.2008.01531.x.
Mardia KV, Kent JT, Bibby JM. 1994. Multivariate Analysis. London: Academic Press.
Medellín RA, Arita TH, Sánchez HO. 2008. Identificación de los murciélagos de México. Clave de campo. 2a edición ed. México: Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad de México.
Muschelli J. 2019. ROC and AUC with a binary predictor: a potentially misleading metric. J Classif. doi:10.1007/s00357-019-09345-1.
Norberg UM, Rayner JM. 1987. Ecological morphology and flight in bats (Mammalia; Chiroptera): wing adaptations, flight performance, foraging strategy and echolocation. Philos Trans R Soc Lond B Biol Sci. 316:335–427.
Nuñez GB, Lemus G, Wolf MM, Rodales AL, González EM, Crisci C. 2018. The first artificial intelligence algorithm for identification of bat species in Uruguay. Ecol Inform. 46:97–102. doi:10.1016/j.ecoinf.2018.05.005.
O'Farrell MJ. 1997. Use of echolocation calls for the identification of free-flying bats. Trans West Sect Wilde Soc. 33:1–8.
O'Farrell MJ, Miller BW. 1999. Use of vocal signatures for the inventory of free-flying Neotropical Bats. Biotropica. 31:507–516. doi:10.1111/j.1744-7429.1999.tb00394.x.

Obrist MK. 1995. Flexible bat echolocation: the influence of individual, habitat and conspecifics on sonar signal design. Behav Ecol Sociobiol. 36:207–219. doi:10.1007/BF00177798.

Ochoa GJ, O'Farrell MJ, Miller BW. 2000. Contribution of acoustic methods to the study of insectivorous bat diversity in protected areas from northern Venezuela. Acta Chiropt. 2:171–183.

Orozco-Lugo L, Guillén-Servent A, Valenzuela-Galván D, Arita HT. 2013. Descripción de los pulsos de ecolocalización de once especies de murciélagos insectívoros aéreos de una selva baja caducifolia en Morelos, México. Therya. 4:33–46. doi:10.12933/therya-13-103.

Parsons S, Obrist MK. 2000. Recent methodological advances in the recording and analysis of chiropteran biosonar signals in the field. In: Thomas JA, Moss C, Vater M, editors. Echolocation in Bats and Dolphins. Chicago (IL): University of Chicago; p. 468–478.

Parsons S, Szewczak JM. 2009. Detecting, recording and analyzing the vocalizations of bats. In: Kunz TH, Parsons S, editors. Ecological and behavioral methods for the study of bats. Baltimore (MD): The Johns Hopkins University Press; p. 91–111.

Parsons S, Jones G. 2000. Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks. J Exp Biol. 203:2641–2656.

Patten MA. 2004. Correlates of species richness in North American bat families. J Biogeogr. 31:975–985. doi:10.1111/j.1365-2699.2004.01087.x.

Poulsen J, French A. 2008. Discriminant function analysis. San Francisco (CA): San Francisco State University.

R Core Team. 2018. R Foundation for Statistical Computing; Vienna, Austria: 2015. R: A language and environment for statistical computing, 2013.

Ramírez-Pulido J, González-Ruiz N, Gardner AL, Arroyo-Cabrales J. 2014. List of recent land mammals of Mexico, 2014. Spec Publ Mus Tex Tech Univ. 63:1–69.

Redgwell RD, Szewczak JM, Jones G, Parsons S. 2009. Classification of echolocation calls from 14 species of bat by support vector machines and ensembles of neural networks. Algorithms. 2:907–924. doi:10.3390/a2030907.

Rivera-Parra P, Burneo SF. 2013. Primera biblioteca de llamadas de ecolocalización de murciélagos del Ecuador. Therya. 4:79–88.

Rodríguez-San Pedro A, Simonetti JA. 2013. Acoustic identification of four species of bats (Order Chiroptera) in central Chile. Bioacoustics. 22:165–172. doi:10.1080/09524622.2013.763384.

Ruedi M, Stadelmann B, Gager Y, Douzery EJ, Francis CM, Lin LK, Guillén-Servent A, Cibois A. 2013. Molecular phylogenetic reconstructions identify East Asia as the cradle for the evolution of the cosmopolitan genus *Myotis* (Mammalia, Chiroptera). Mol Phylogenet Evol. 69:437–449. doi:10.1016/j.ympev.2013.08.011.

Russo D, Jones G. 2002. Identification of twenty-two bat species (Mammalia: Chiroptera) from Italy by analysis of time-expanded recordings of echolocation calls. J Zool. 258:91–103. doi:10.1017/S0952836902001231.

Russo D, Jones G, Mucedda M. 2001. Influence of age, sex and body size on echolocation calls of Mediterranean (*Rhinolophus euryale*) and Mehely's (*Rhinolophus mehelyi*) horseshoe bats (Chiroptera: Rhinolophidae). Mammalia. 65:429–436. doi:10.1515/mamm.2001.65.4.429.

Rydell J, Arita HT, Santos M, Granados J. 2002. Acoustic identification of insectivorous bats (Order Chiroptera) of Yucatan, Mexico. J Zool. 257:27–36. doi:10.1017/S0952836902000626.

Rydell J, Speakman JR. 1995. Evolution of nocturnality in bats: potential competitors and predators during their early history. Biol J Linn Soc. 54:183–191. doi:10.1111/j.1095-8312.1995.tb01031.x.

Samarasinghe S. 2007. Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition. Boca Raton (FL): CRC Press.

Skowronski MD, Harris JG. 2006. Acoustic detection and classification of microchiroptera using machine learning: lessons learned from automatic speech recognition. J Acoust Soc Am. 119:1817–1833. doi:10.1121/1.2166948.

Sokal RR, Rohlf F. 1981. Biometry. New York: W. H. Freeman & Co.

Speakman JR. 2001. The evolution of flight and echolocation in bats: another leap in the dark. Mammal Rev. 31:111–130. doi:10.1046/j.1365-2907.2001.00082.x.

Szewczak JM. 2004. Advanced analysis techniques for identifying bat species. In: Brigham RM, Kalko EKV, Jones G, Parson S, Limpens HJGA, editors. Bat echolocation research: tools, techniques, and analysis. Austin (TX): Bat Conservation International; p. 121–126.
Szewczak JM 2010. Sonobat. Version 3.0. Accessed 2019 Jul 1. www.sonobat.com.
Vaughan N, Jones G, Harris S. 1997. Identification of British bat species by multivariate analysis of echolocation call parameters. Bioacoustics. 7:189–207. doi:10.1080/09524622.1997.9753331.
Veelenturf LPJ. 1995. Analysis and applications of artificial neural networks. Hertfordshire (UK): Prentice Hall International.
Villers RLF, Rojas F, Tenorio P. 2006. Guía Botánica del Parque Nacional Malinche, Tlaxcala-Puebla. México: Centro de Ciencias de la Atmósfera e Instituto de Biología. Ciudad de México.
Walters CL, Freeman R, Collen A, Dietz C, Brock Fenton M, Jones G, Obrist MK, Puechmaille SJ, Sattler T, Siemers B, et al. 2012. A continental-scale tool for acoustic identification of European bats. J Appl Ecol. 49:1064–1074. doi:10.1111/j.1365-2664.2012.02182.x.
Waters DA, Gannon WL. 2004. Bat call libraries: management and potential use. In: Brigham RM, Kalko EKV, Jones G, Parson S, Limpens HJGA, editors. Bat echolocation research: tools, techniques, and analysis. Austin (TX): Bat Conservation International; p. 150–157.
Welsh HH, Droege S. 2001. A case for using plethodontid salamanders for monitoring biodiversity and ecosystem integrity of north american forests. Conserv Biol. 15:558–569. doi:10.1046/j.1523-1739.2001.015003558.x.
White H. 1992. Artificial neural networks: approximation and learning theory. Cambridge (MA): Blackwell.
Zamora-Gutierrez V, Lopez-Gonzalez C, MacSwiney Gonzalez MC, Fenton B, Jones G, Kalko EK, Puechmille SJ, Stathopoulos V, Jones KE. 2016. Acoustic identification of Mexican bats based on taxonomic and ecological constraints on call design. Methods Ecol Evol. 7:1082–1091. doi:10.1111/2041-210X.12556.

Appendix. Call parameters used for the species identification of bats inhabiting La Malinche National Park, a mountain ecosystem of central Mexico

Call parameter	Description	Biological interpretation	References
Call duration	Call duration (milliseconds).	Short calls give bats better localisation while long ones give them more temporal information.	Vaughan et al. 1997; Russo and Jones 2002;
Highest frequency	Highest apparent frequency of the call (kHz).	Lower frequencies travel further through the air while higher frequencies give bats better resolution.	Zamora-Gutierrez et al. 2016
Lowest frequency	Lowest apparent frequency of the call (kHz).	Lower frequencies travel further through the air while higher frequencies give bats better resolution.	Zamora-Gutierrez et al. 2016
Bandwidth	Total frequency spread of the call. Calculated from the difference between the highest and lowest frequency (kHz).	High bandwidths allow greater resolution to bats, while low bandwidths allow concentration of energy and increased range detection.	Zamora-Gutierrez et al. 2016
Characteristic frequency	Characteristic frequency of the call. Determined by finding the point in the final 40% of the call having the lowest slope or exhibiting the end of the main trend of the body of the call (kHz).	Lower frequencies travel further through the air, while higher frequencies give better resolution to bats.	Zamora-Gutierrez et al. 2016
Frequency with maximum amplitude	The frequency of the maximum amplitude of the call (kHz).	Lower frequencies travel further while higher frequencies give better resolution to bats.	Vaughan et al. 1997; Russo and Jones 2002
Percentage of maximum amplitude	Percentage of the entire call duration at which the maximum amplitude occurs (%).	Higher amplitude produces stronger echoes which are resistant to atmospheric attenuation.	Zamora-Gutierrez et al. 2016
Start frequency	Frequency of the start of the call. Typically, the same point as the highest frequency, but different if the call initially rises in frequency (kHz).	Lower frequencies travel further while trough the air, while higher frequencies give greater resolution to bats.	Vaughan et al. 1997; Russo and Jones 2002; Zamora-Gutierrez et al. 2016
End frequency	Frequency of the end of the call. Typically, the same point as the lowest frequency, but different if the call ends with a rise in frequency.	Lower frequencies travel further while trough the air, while higher frequencies give greater resolution to bats.	Vaughan et al. 1997; Russo and Jones 2002; Zamora-Gutierrez et al. 2016
Amplitude of the first quartile	Total amplitude of the first quartile of the call (relative units).	Higher amplitude produces stronger echoes and are resistant to atmospheric attenuation.	Redgwell et al. 2009; Zamora-Gutierrez et al. 2016
Amplitude of the second quartile	Total amplitude of the second quartile of the call (relative units).	Higher amplitude produces stronger echoes and resistant to atmospheric attenuation.	Redgwell et al. 2009; Zamora-Gutierrez et al. 2016
Amplitude of the third quartile	Total amplitude of the third quartile of the call (relative units).	Higher amplitude produces stronger echoes and resistant to atmospheric attenuation.	Redgwell et al. 2009; Zamora-Gutierrez et al. 2016
Amplitude of the fourth quartile	Total amplitude of the fourth quartile of the call (relative units).	Higher amplitude produces stronger echoes and resistant to atmospheric attenuation.	Redgwell et al. 2009; Zamora-Gutierrez et al. 2016
Rate of change	The second order derivative of the frequency-time course of each call expressed at kHz/ms2 (relative units).	Determinate the shape of call. FM calls are gives greater resolution and localisation, and CF calls gives greater range and more temporal information.	Redgwell et al. 2009; Zamora-Gutierrez et al. 2016
Call slope	The slope of the call, calculated from the difference in frequency and time from the point of highest frequency to the point of the lowest frequency (relative units).	Greater slope gives greater resolution and localisation while lower slope gives greater range and more temporal information to bats.	Present research.