Chapter 1

Ancestral Reconstruction of Bat Echolocation Calls

J.P. Meagher*

Department of Statistics,
University of Warwick,
J.Meagher@Warwick.ac.uk†

The stated aim of the Statistical Data Science workshop jointly organised by the Department of Mathematics and Data Science Institute at Imperial College London, and Winton Global investment management, is "exploring the nature of the relationship between statistics and data science".

1. Introduction

1.1. *Motivation*

Data science, the " study of the generalisable extraction of knowledge from data"[1] is motivated by availability of interesting datasets. One field producing such datasets is Bioacoustics, where data is collected, often through citizen science[2] initiatives, for monitoring and conservation purposes.[3] Bats (order *Chiroptera*) have been of particular interest.[4]

Bats form the second most speciose order of mammals, behind rodents, with over 1200 species.[5] Bats have been identified as ideal bioindicators for monitoring climate change and habitat quality.[6] Good bioindicators are easy to identify and sample, well distributed geographically, respond to changes in habitat in a manner correlated with other taxa, and have a well understood natural history.[7] Bats have the potential to satisfy these characteristics excellently.

Bats are, for the most part, nocturnal, flying echolocators.[8] Echolocation refers to the bats use of, typically ultrasonic,[9] calls and their echoes to

---

*Author footnote.
†Affiliation footnote.

forage and navigate in the night sky.[10] Thus, bats leak information about themselves into the environment, allowing acoustic monitoring.[4] Work towards the development of automatic acoustic monitoring algorithms for bats[11][12] is ongoing. Similar algorithms have been developed for insects[13] and birds,[14] reflecting the level of interest in bioacoustic monitoring.

The growing database of bat call recordings[15] also opens up other avenues for research. Bat echolocation calls are diverse and can generally be sorted into categories according to the duration, bandwidth and use of harmonics in the call.[16] It has been observed that closely related species have similar call structures indicating that some of the variation is due to a shared evolutionary history.[17] Comparative analysis of bat echolocation call parameters for ancestral reconstruction[18] have been performed.[15] However, an approach based Statistical Data Science principles may shed further light on the evolution of echolocation in bats, without relying heavily on domain specific knowledge.

1.2. *Literature review*

Echolocation, or biosonar, is a process whereby sound is produced and the echoes that return from objects are used to perceive the environment.[19] Animals that use echolocation include bats, toothed whales, some birds, and some shrews and rats.[19] Bat echolocation calls, ranging from 9 to 212 kHz, differ from those of other echolocators in that they are often laryngeal, complex, and of a relatively long duration.[20] The evolutionary dynamics that resulted in bats being nocturnal, flying echolocators have resulted in a wide variety of call structures across species.[17] Echolocation in bats demonstrates convergent evolution, and bat species have been arranged in 'guilds' based on habitat and foraging strategy, with similar call structures observed within a 'guild'.[21][22][23] However, it has been acknowledged that "echolocation signals reflect a phylogenetically determined basic call structure shaped by specific ecological conditions."[24] Ancestral Reconstruction is a key aspect of understanding the variation is call structure that is due to phylogenetic relationships between species.

Ancestral Reconstruction involves the extrapolation back in time from measured characteristics of current populations to the ancestral state, the estimate of the same characteristic in common ancestors.[18] Effective ancestral reconstruction relies on an accurate model for evolution, that is, both the dynamic evolutionary process and the phylogenetic relationships between species.[18] There have been maximum parsimony,[25] maximum like-
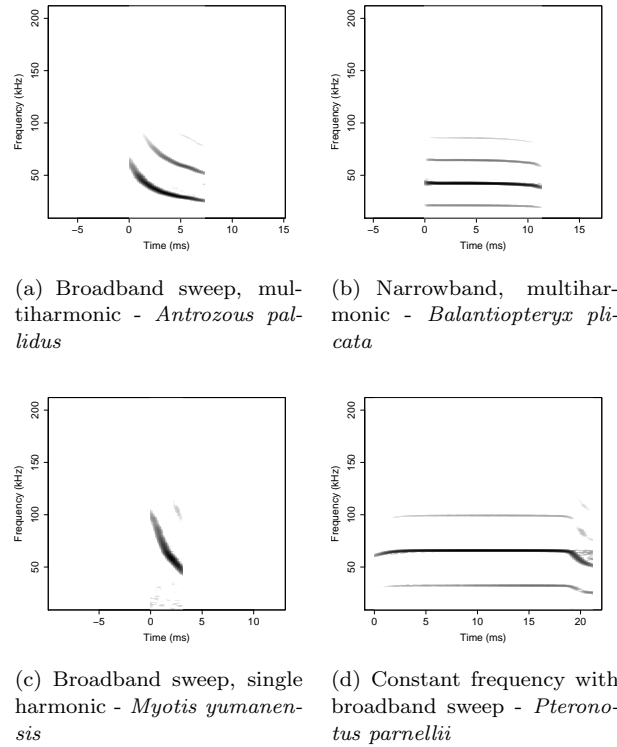
(a) Broadband sweep, multiharmonic - *Antrozous pallidus*

(b) Narrowband, multiharmonic - *Balantiopteryx plicata*

(c) Broadband sweep, single harmonic - *Myotis yumanensis*

(d) Constant frequency with broadband sweep - *Pteronotus parnellii*

Fig. 1.   Representative call spectrograms for some bat species, demonstrating echolocation call structures.

lihood,[26] and Bayesian[27] approaches taken to ancestral reconstruction.

Other studies have considered bat echolocation calls for ancestral reconstruction.[28][15] Fenton[28] hypothesised that early bats used short, broadband clicks for echolocation based on the observation that this is the method of echolocating employed by all animals except bats. Collen[15] proposed instead that early bats used short, multi-harmonic, narrowband laryngeal calls for echolocation, based on a comparative analysis of echolocation call parameters. An approach to the reconstruction of early bat echolocation calls which is not so selective in the information it employs may shed further light on this topic. Gaussian Process Regression on Phylogenies for function valued traits provide a promising path to explore.[29] Jones & Moriarty proposed an extension to Gaussian Process Regression[30] which allowed the generalisation of Ornstein-Uhlenbeck[31] models of continuous-time charac-

ter evolution for traits considered as functional data objects,[32] through a phylogeny. This study represents an effort to apply these methods to the ancestral reconstruction of bat echolocation calls.

2. Methods

2.1. *Data Description*

The post processed echolocation call data accompanying Stathopoulos et al.[11] was used in this analysis. Echolocation calls were recorded across north and central Mexico with a Pettersson 1000x bat detector (Pettersson Elektronik AB, Uppsala, Sweden). Live trapped bats were measured and identified to species level using field keys.[33][34] Bats were recorded either while released from the hand or while tied to a zip line. The bat detector was set to record calls manually in real time, full spectrum at 500 kHz. In total the dataset consists of 22 species in five families, 449 individual bats and 1816 individual echolocation call recordings.

Phylogenetic relationships between species are described and dated according to the Bat super-tree produced by Collen.[15]
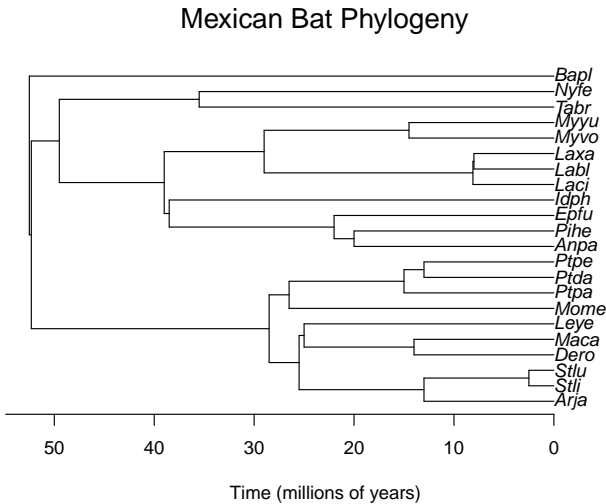
Mexican Bat Phylogeny

Table 1.   Echolocation Call Dataset Statistics

Species	Key	Individuals	Calls
Family: Emballonuridae			
1 *Balantiopteryx plicata*	Bapl	16	100
Family: Molossidae			
2 *Nyctinomops femorosaccus*	Nyfe	16	100
3 *Tadarida brasiliensis*	Tabr	49	100
Family: Vespertilionidae			
4 *Antrozous pallidus*	Anpa	58	100
5 *Eptesicus fuscus*	Epfu	74	100
6 *Idionycteris phyllotis*	Idph	6	100
7 *Lasiurus blossevillii*	Labl	10	90
8 *Lasiurus cinereus*	Laci	5	42
9 *Lasiurus xanthinus*	Laxa	8	100
10 *Myotis volans*	Myvo	8	100
11 *Myotis yumanensis*	Myyu	5	89
12 *Pipistrellus hesperus*	Pihe	85	100
Family: Mormoopidae			
13 *Mormoops megalophylla*	Mome	10	100
14 *Pteronotus davyi*	Ptda	8	100
15 *Pteronotus parnellii*	Ptpa	23	100
16 *Pteronotus personatus*	Ptpe	7	51
Family:Phyllostomidae			
17 *Artibeus jamaicensis*	Arja	11	82
18 *Desmodus rotundus*	Dero	6	38
19 *Leptonycteris yerbabuenae*	Leye	26	100
20 *Macrotus californicus*	Maca	6	53
21 *Sturnira ludovici*	Stlu	8	51
22 *Sturnira lilium*	Stli	4	20

2.2. *Recordings*

Describe the Data. Get started with notation. Worth looking at vassilos stuff here.

2.3. *Power spectral Density*

Define PSD, get references, Fourier transform of the data.

2.4. Functional Data Objects

Justification of PSD as a functional data object, smoothing, regularisation etc.

2.5. Component Analysis

PCA and ICA execution. Tuning of components.

2.6. Phylogenetic Gaussian Processes

Gaussian Processes, Matern / OU kernel, interpretation of parameters proceedings

3. Results

3.1. Simulation Study

given hypers, what happens along the given tree

3.2. Mexican Bat Echolocation call Dataset

What happens when the method is applied to the Mexican bat calls?

4. Discussion

4.1. What do we learn from this?

Phylogenetic / Non-phylogenetic Signal. Not just noise. Although speculative and probably limited going back to root may be useful for more recent ancestors.

4.2. Future Work in this Area

Implement for Spectrograms, produce some recordings.

References

1. V. Dhar, Data science and prediction, *Communications of the ACM.* **56**(12), 64–73 (2013).
2. C. Kullenberg and D. Kasperowski, What is citizen science?–a scientometric meta-analysis, *PloS one.* **11**(1), e0147152 (2016).

3. P. E. Allen and C. B. Cooper. Citizen science as a tool for biodiversity monitoring. (2006).
4. N. Pettorelli, J. E. Baillie, and S. M. Durant, Indicator bats program: a system for the global acoustic monitoring of bats (2013).
5. N. B. Simmons, Order chiroptera, *Mammal species of the world: a taxonomic and geographic reference*. **1**, 312–529 (2005).
6. G. Jones, D. S. Jacobs, T. H. Kunz, M. R. Willig, and P. A. Racey, Carpe noctem: the importance of bats as bioindicators, *Endangered species research*. **8**(1-2), 93–115 (2009).
7. C. E. Moreno, G. Sánchez-Rojas, E. Pineda, and F. Escobar, Shortcuts for biodiversity evaluation: a review of terminology and recommendations for the use of target groups, bioindicators and surrogates, *International Journal of Environment and Health*. **1**(1), 71–86 (2007).
8. T. Kunz and E. Pierson, Bats of the world: an introduction, *Walkers Bats of the World*. pp. 1–46 (1994).
9. J. F. Corso, Bone-conduction thresholds for sonic and ultrasonic frequencies, *The Journal of the Acoustical Society of America*. **35**(11), 1738–1743 (1963).
10. D. R. Griffin, Echolocation by blind men, bats and radar, *Science*. **100**(2609), 589–590 (1944).
11. V. Stathopoulos, V. Zamora-Gutierrez, K. E. Jones, and M. Girolami, Bat echolocation call identification for biodiversity monitoring: a probabilistic approach, *Journal of the Royal Statistical Society: Series C (Applied Statistics)* (2017).
12. C. L. Walters, R. Freeman, A. Collen, C. Dietz, M. Brock Fenton, G. Jones, M. K. Obrist, S. J. Puechmaille, T. Sattler, B. M. Siemers, et al., A continental-scale tool for acoustic identification of european bats, *Journal of Applied Ecology*. **49**(5), 1064–1074 (2012).
13. D. Chesmore, Automated bioacoustic identification of species, *Anais da Academia Brasileira de Ciências*. **76**(2), 436–440 (2004).
14. F. Briggs, B. Lakshminarayanan, L. Neal, X. Z. Fern, R. Raich, S. J. Hadley, A. S. Hadley, and M. G. Betts, Acoustic classification of multiple simultaneous bird species: A multi-instance multi-label approach, *The Journal of the Acoustical Society of America*. **131**(6), 4640–4650 (2012).
15. A. Collen. *The evolution of echolocation in bats: a comparative approach.* PhD thesis, UCL (University College London) (2012).
16. A. Maltby, K. E. Jones, and G. Jones, .4-understanding the evolutionary origin and diversification of bat echolocation calls, *Handbook of Behavioral Neuroscience*. **19**, 37–47 (2010).
17. G. Jones and E. C. Teeling, The evolution of echolocation in bats, *Trends in Ecology & Evolution*. **21**(3), 149–156 (2006).
18. J. B. Joy, R. H. Liang, R. M. McCloskey, T. Nguyen, and A. F. Poon, Ancestral reconstruction, *PLoS Comput Biol*. **12**(7), e1004763 (2016).
19. G. Jones, Echolocation, *Current Biology*. **15**(13), R484–R488 (2005).
20. J. A. Thomas, *Echolocation in bats and dolphins*. University of Chicago Press (2004).
21. H. Aldridge and I. Rautenbach, Morphology, echolocation and resource par-

titioning in insectivorous bats, *The Journal of Animal Ecology*. pp. 763–778 (1987).

22. G. Neuweiler, Auditory adaptations for prey capture in echolocating bats, *Physiological reviews*. **70**(3), 615–641 (1990).

23. H.-U. Schnitzler and E. K. Kalko, 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 (2001).

24. H. Schnitzler, E. Kalko, and A. Denzinger, Evolution of echolocation and foraging behavior in bats, *Echolocation in bats and dolphins*. pp. 331–339 (2004).

25. W. M. Fitch, Toward defining the course of evolution: minimum change for a specific tree topology, *Systematic Biology*. **20**(4), 406–416 (1971).

26. T. Pupko, I. Pe, R. Shamir, and D. Graur, A fast algorithm for joint reconstruction of ancestral amino acid sequences, *Molecular Biology and Evolution*. **17**(6), 890–896 (2000).

27. M. Pagel, A. Meade, and D. Barker, Bayesian estimation of ancestral character states on phylogenies, *Systematic biology*. **53**(5), 673–684 (2004).

28. M. Fenton, D. Audet, M. K. Obrist, and J. Rydell, Signal strength, timing, and self-deafening: the evolution of echolocation in bats, *Paleobiology*. **21** (02), 229–242 (1995).

29. N. S. Jones and J. Moriarty, Evolutionary inference for function-valued traits: Gaussian process regression on phylogenies, *Journal of The Royal Society Interface*. **10**(78), 20120616 (2013).

30. C. E. Rasmussen, Gaussian processes for machine learning (2006).

31. G. E. Uhlenbeck and L. S. Ornstein, On the theory of the brownian motion, *Physical review*. **36**(5), 823 (1930).

32. J. O. Ramsay, *Functional data analysis*. Wiley Online Library (2006).

33. G. Ceballos and G. Oliva. Los mamíferos silvestres de méxico. comisión nacional para el conocimiento y uso de la biodiversidad (2005).

34. R. Medellín and H. Arita, O. sánchez h. 2008. identificación de los murciélagos de méxico. claves de campo, *Revista Mexicana de Mastozoología*. **2**, 1–83 .