For R code and dataset please visit: https://github.com/YueLiuUFPU/FNR642_SemesterProjectYue.git

Investigating the Presence of Dialects in Carolina Chickadee's call

1 Introduction

Dialects are unique subdivisions of a particular language characterized by differences in pronunciation, vocabulary, grammar, and syntax (Chambers & Trudgill, 1998). While this term is often associated with human language, many non-human animals also have communication systems that evolve through similar processes. Bird dialect formation, for example, refers to the process by which different populations or subgroups of birds develop distinct variations in their calls or songs (Baker & Cunningham, 1985; Date & Lemon, 1993; Eliot A. Brenowitz et al., 1985). Similar to human language, the formation of dialects in bird communication can result from a complex interplay of various factors. The geographic isolation theory posits that physical barriers or distance can lead to the evolution of distinct communication systems. In this project, I focus on Carolina chickadees (*Poecile carolinensis*), a common species in the southeastern United States, to investigate the formation of dialects in their vocal system, particularly in their call system. My prediction is that there are significant differences in sound parameters between the chick-a-dee calls produced by the Carolina Chickadee populations in two different habitats, Ross Nature Reserve and Martel Reserve.

2 Methods

2-1 Call collection. Beginning in January 2023, I conducted field recordings of Carolina Chickadees at two separate study sites: Ross Nature Reserve and Martel Reserve. Each recording was then processed in a professional audio editing software, Adobe Audition (version 2023), and saved as an individual .wav file. Carolina Chickadee calls consist of four distinct note types (a, b, c, and d), and the number of each note type within a call can vary. To accurately analyze these calls, it was necessary to isolate each note type into its own individual .wav file. Therefore, I carefully extracted each note from the calls and saved them as separate files for further analysis. Since the sample sizes of the note a, b, and c are small and unbalanced, I decided to use note "d" only in further analysis to test the hypothesis.

2-2 Extract sound parameter. For comparing the similarity of the "d" note produced by Carolina Chickadee populations in Ross Nature Reserve and Martel Reserve, I utilized Cluster Analysis. Prior to conducting the cluster analysis, it was necessary to quantify the audio files. I imported all 25 "d" note audio files into the R programming language and utilized the "WarbleR" bioacoustics analysis package to generate spectrograms and extract sound parameters (Araya-Salas & Smith-Vidaurre, 2017; Erbe & Thomas, 2022). "WarbleR" automatically extracts 28 sound parameters, but not all of them are necessary for analyzing bird calls, so I refined the parameters to 12 for the following analyses. These parameters were saved in a .csv file for further analysis.

2-3 Principal Component. Having 12 sound parameter makes our dataset a high-dimensional dataset. I conducted PCA to reduce the number of variables in the dataset, while identify the most important variables and patterns in the data, and to simplify complex datasets(Hassan & Ramli, 2018).

2-4 Cluster Analysis. Given the relatively small sample size, I utilized four different means to run cluster analysis on the dataset and determine the optimal number of clusters: (1) K-means clustering, (2) PAM clustering, (3) Model-based clustering, and (4) Hierarchical Clustering on Principal Component (Calabrese et al., n.d.; Koh et al., 2022; Mcloughlin et al., 2019).

For R code and dataset please visit: https://github.com/YueLiuUFPU/FNR642_SemesterProjectYue.git

3 Results

3-1 PCA. Two principal components were kept (Figure 1). The first component explains 99.1% of the variance, and the most important sound parameter in the 1st principal component is kurt. Kurt is the kurtosis of the distribution of the sound signal, which usually is used as a predictor of biological impacts from noise exposure.

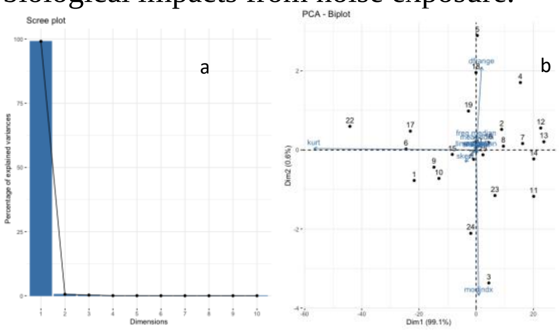

Figure 1: Visualizations of the PCA results. (1a) The screeplot which shows the percentages of explained variance by each principal component. (1b) The biplot which shows all data points in the reduced-dimensional space defined by the first two principal components.

3-2 cluster analysis. Both of the 4 different cluster analysis suggested 2 as the optimal number of cluster analysis (Figure 2).

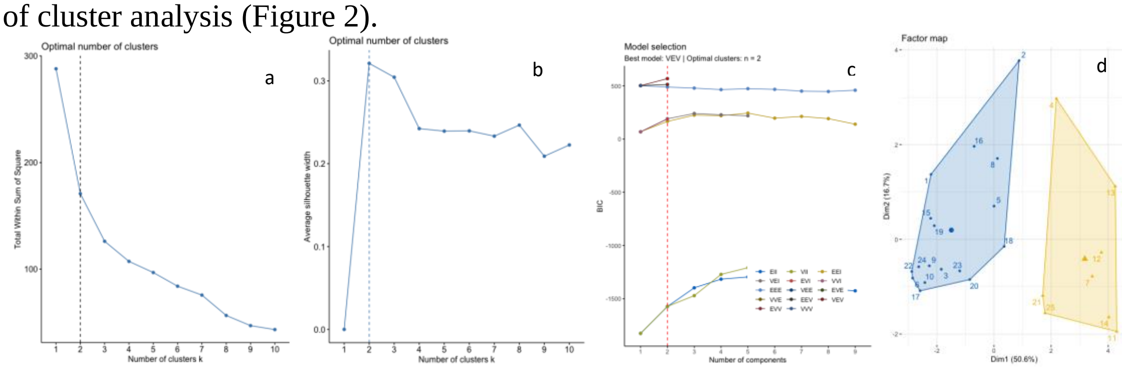

Figure 2: Visualizations of the results of the four cluster analyses. (2a) Plot for K-means clustering; (2b) plot for PAM clustering; (2c) plot for Model-based clustering; and (2d) Plot for Hierarchical-clustering on Principal Component.

4 Discussion

The four cluster analyses conducted using various means consistently suggested that two was the optimal number of clusters, confirming my prediction that there are significant differences in sound parameters between the "d" notes produced by the Carolina Chickadee populations in Ross Nature Reserve and Martel Reserve. And the difference is mainly related to different noise exposure levels in the two study sites. Based on these findings, we can conclude that there is indeed a dialect present in the Carolina Chickadee's call, at least in the sound parameters of "d" notes. Furthermore, our analysis suggests that the most significant factor contributing to the formation of this dialect is the level of noise in the environment. These results provide valuable insights into the communication behavior of Carolina Chickadee populations in different environments and highlight the potential impact of noise pollution on their vocal communication.

5 References

For R code and dataset please visit: https://github.com/YueLiuUFPU/FNR642_SemesterProjectYue.git

Araya-Salas, M., & Smith-Vidaurre, G. (2017). warbleR: An R package to streamline analysis of animal acoustic signals. *Methods in Ecology and Evolution*, *8*(2), 184–191.

Baker, M. C., & Cunningham, M. A. (1985). The Biology of Bird-Song Dialects. *Behavioral and Brain Sciences*, *8*(1), 85–100.

Calabrese, L., Campanella, G., & Proverbio, E. (n.d.). *Use Of Cluster Analysis Of Acoustic Emission Signals In Evaluating Damage Severity In Concrete Structures*.

Chambers, J. K., & Trudgill, P. (1998). *Dialectology*. Cambridge University Press.

Date, E. M., & Lemon, R. E. (1993). Sound Transmission: A Basis for Dialects in Birdsong? *Behaviour*, *124*(3), 291–312.

Eliot A. Brenowitz, Eliot A. Brenowitz, & Brenowitz, E. A. (1985). Bird-song dialects: Filling in the gaps. *Behavioral and Brain Sciences*, *8*(1), 101–102.

Erbe, C., & Thomas, J. A. (Eds.). (2022). *Exploring Animal Behavior Through Sound: Volume 1: Methods*. Springer International Publishing. https://doi.org/10.1007/978-3-030-97540-1

Hassan, N., & Ramli, D. A. (2018). A Comparative Study of Blind Source Separation for Bioacoustics Sounds based on FastICA, PCA and NMF. *Procedia Computer Science*, *126*, 363–372.

Koh, K.-Y., Ahmad, S., Lee, J., Suh, G.-H., & Lee, C.-M. (2022). Hierarchical Clustering on Principal Components Analysis to Detect Clusters of Highly Pathogenic Avian Influenza Subtype H5N6 Epidemic across South Korean Poultry Farms. *Symmetry*, *14*(3), 598.

Mcloughlin, M. P., Stewart, R., & McElligott, A. G. (2019). Automated bioacoustics: Methods in ecology and conservation and their potential for animal welfare monitoring. *Journal of The Royal Society Interface*, *16*(155), 20190225.