**Testing the accuracy of BirdNET and potential applications**

**Yi-Chin Tseng**

# Outline

Spatiotemporal variation of avian biodiversity is a commonly used indicator of environmental change. Conventionally, such information was derived with human observers (e.g., point counts and mist net), while passive acoustic monitoring (PAM) with autonomous recording units (ARU) is rapidly emerging as an alternative survey method. Given the large amount of acoustic data PAM can potentially collect, effort has been made to develop algorithms to automatically transform acoustic data into interpretable form, such as species identification or detection. Recently, one of the most successful attempts is a deep neural network (DNN) called BirdNET, that is able to identify 984 North American and European bird species by their sound. The source code of BirdNET was published in 2021 on Github. In this study, we will test the performance of BirdNET on an independent dataset collected in John Prince Research Forest, Canada. Furthermore, we will access the accuracy metrics for individual species and imply future applications of such acoustic data.

**Keywords**: Pacific North West, forest birds, migration, BirdNET

# Objectives

* Determine the best confidence threshold for selected focal bird species and evaluate the species-wise performance
* Discuss where the error might come from (e.g., time of the day, location)

# Method and materials

## Audio data

The audio data was collected in John Prince Research Forest (54° 27'N, 124° 10'W, 700 m a.s.l) in 2020 breeding season. A total of 66 AudioMoth (Open Acoustic Devices, 2020) were evenly distributed across the region (Fig. 1). All AudioMoth was under an identical recording schedule, repeating daily from four to seven am, one minute on, followed by four minutes off. The working days were ~30 – 60 days between May – July. Given this, 67,301 one-minute recordings were collected and used in the study. All recordings were formatted into a 48 kHz sampling rate and the mono pulse code modulation WAV.

## BirdNET

An automatic bird sound classifier, BirdNET (Kahl et al., 2021), was developed given the demand of new analytical approaches to efficiently process PAM data (Shonfield & Bayne, 2017). The training data of BirdNET came from two datasets: Macaulay Library, part of the Cornell Laboratory of Ornithology containing over 750,000 avian acoustic recordings (Macaulay, 2021), and Xeno-Canto, a community-curated collection featuring more than 500,000 avian acoustic recordings (Xeno-canto, 2021). BirdNET is capable of identifying 984 North American and European bird species by their sound. Using BirdNET, species presence confidences can be retrieved for each three-second sound segment. BirdNET was deployed as a website (<https://birdnet.cornell.edu/>) with its source code provided in a GitHub repository (<https://github.com/kahst/BirdNET>). In this study, the BirdNET analysis will be run on the Google Colaboratory (Colab), an online platform allowing users to get free access to graphics processing units (GPUs).

|  |  |
| --- | --- |
| **(a)** | **(b)** |

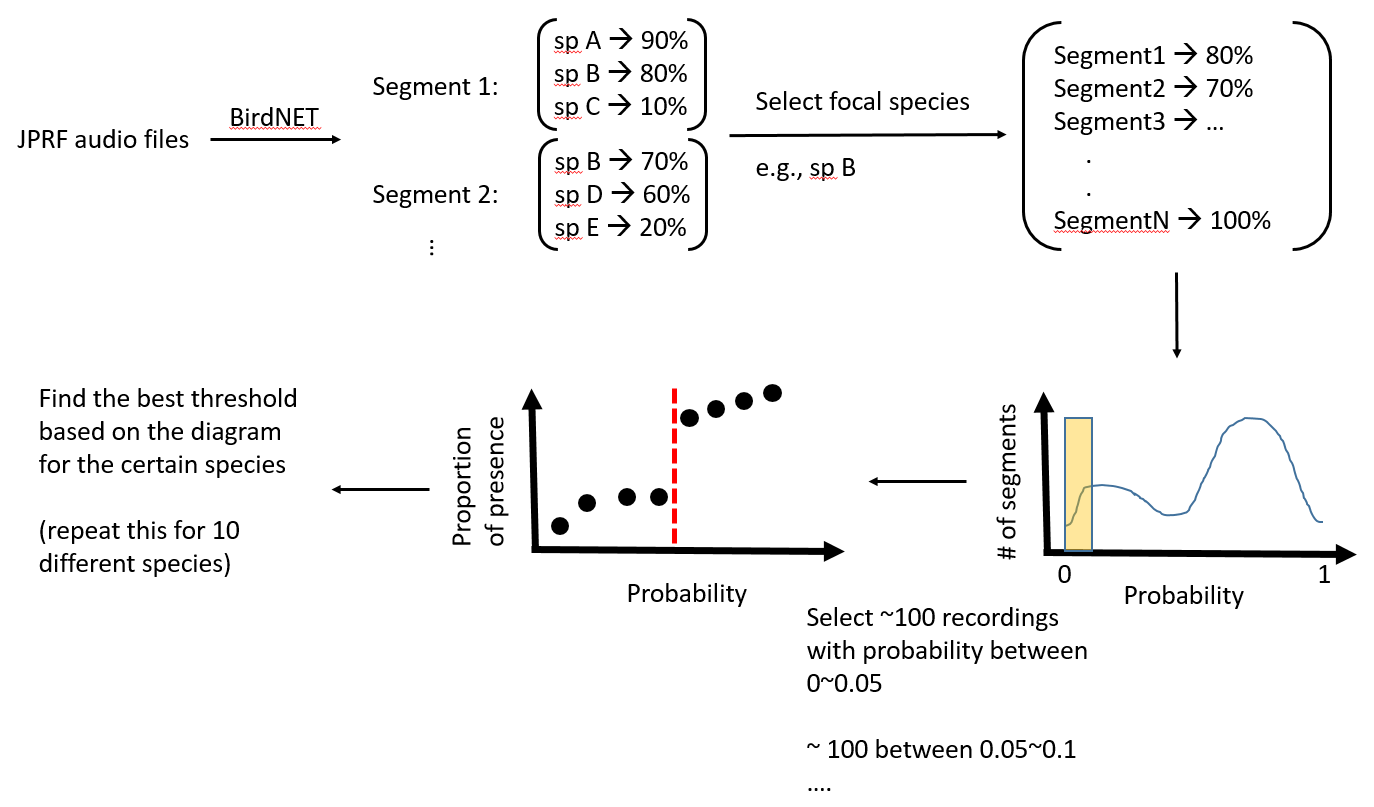
Fig. 1. Locations of Audio Moth recorders in the research forest.

## Species-wise accuracy assessment

Given that each bird species might have different best confidence threshold, we will run the evaluation process separately for each of the focal species. The 11 focal species were selected by different levels of vocal activity in two different habitats, namely early-seral forest and matured forest (cite). In early-seral forest habitat, seven species were selected: White-throated Sparrow (*Zonotrichia albicollis*), Least Flycatcher (*Empidonax minimus*), Olive-sided Flycatcher (*Contopus cooperi*), Alder Flycatcher (*Empidonax alnorum*), Western Wood-PeWee (*Contopus sordidulus*), Dusky Flycatcher (*Empidonax oberholseri*), and Willow Flycatcher (*Empidonax traillii*). In the matured forest, four species were selected: Swainson’s Thrush (*Catharus ustulatus*), Pacific Wren (*Troglodytes pacificus*), Western Tanager (*Piranga ludoviciana*), and Hammond’s Flycatcher (*Empidonax hammondii*).

To determine the best confidence threshold, positive detections of each focal species will be extracted. Each detection will be categorized into one of the 20 groups (i.e., 0 – 0.05, 0.05 – 0.1, 0.1 – 0.15, etc.) based on its confidence score. Then, ground truth will be done by randomly selecting 100 recordings from each of the confidence group with the percentage of the true positive calculated. The best confidence threshold can be determined based on the balance of false positive (the lower the better) and the true positive (the higher the better).

# Workflow



# References

Kahl, S., Wood, C. M., Eibl, M., & Klinck, H. (2021). BirdNET: A deep learning solution for avian diversity monitoring. *Ecological Informatics*, *61*, 101236. https://doi.org/10.1016/J.ECOINF.2021.101236

Macaulay. (2021). *The World’s Premier Scientific Archive of Natural History Audio, Video, and Photographs*. https://www.macaulaylibrary.org/

Shonfield, J., & Bayne, E. M. (2017). Autonomous recording units in avian ecological research: current use and future applications. *Avian Conservation and Ecology, Published Online: May 26, 2017 | Doi:10.5751/ACE-00974-120114*, *12*(1), 14. https://doi.org/10.5751/ACE-00974-120114

Xeno-canto. (2021). *Sharing Bird Sounds from Around the World*. https://xeno-canto.org/