Big Data Analysis Final Project

Lauren Michael

2023-05-01

## How Has Bird Abundance Changed Over Time on Block Island, RI?

This analysis was completed for the final project of the BIO539 - Big Data Analysis class at The University of Rhode Island. The work completed in this project serves as exploratory analysis that will be integrated into my masters thesis. For this project, I will be looking at the long-term changes in bird abundance of various bird species on Block Island, RI using a 56 year-long continuous bird banding dataset.

This dataset, collected at the Block Island Banding Station (BIBS) on Block Island, RI., is publicly accessible via the United States Geological Survey Bird Banding Lab (USGS BBL). This data was collected from 1967 to 2023, though due to some database issues, not all years are present in this dataset. Most notably, 1967 and 2004 are both missing. The dataset issues will be remedied this summer and the analysis will be rerun with the complete dataset. Currently, the net hours dataset has only been entered up until 2016, which means that any analysis that requires effort calculations will only include years up until 2016.

To begin, we need to download the necessary packages: tidyverse and lubridate. The first step of the process involves putting together all of the data I have, which come from different sources. Prior to 2005, the banding data was recording to paper datasheets, entered into a database in a program called Filemaker, and then printed from Filemaker and mailed to the Bird Banding Lab (BBL) to submit the data. Upon inspection of the data, it became apparent that the BBL had lost some of this data, which means that the Filemaker datafiles were more accurate than those downloaded from the BBL. The curator of the database created an excel file where they compared the BBL records and the Filemaker records and matched them up, with the hope of fixing any issues. This is the datafile labeled “matchedBIBSdata.” After 2005, the banding data was collected on paper datasheets and then entered straight into the BBL website. This is the data labeled “banddata.” To end up with the most accurate datafile from 1967-2022, I then needed to combine the two datasets. This code is all commented out so that I can avoid uploading the larger dataset to github, but you can see the process below. In short, here was my process:

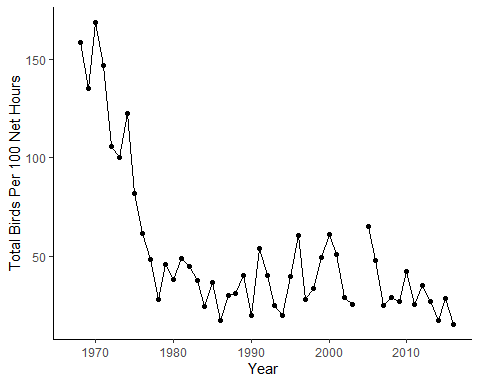
1. Filter the BBL data (which included every bird ever banded on Block Island) down to what I wanted it to be. This means only the BIBS banding site (filtered using the lead bander’s permit number). I then also filtered the BBL data so that it only included data after 2006.
2. Ensure that both datasets had the same columns. The matched dataset had many extra columns that were unnecessary, so those were removed.
3. Use the rbind function to add all of the rows together to make one complete dataset.
4. Filter this combined dataset to only include fall data (what I’m interested in). Fall in this case means any data collected in September, October, or November. This is the dataset we will be working with, titled “falldataBIBS.”

Here is where I began the real analysis now that the data was all ready to work with. Begin by importing the falldataBIBS file as well as the BIBSnethours file. As mentioned, falldata BIBS contains all of the fall data from the BIBS from 1967-present. BIBSnethours is a document containing daily net hour information from 1967-2016. Net hours are calculated as the number of nets open each day and the number of hours they are open. For example, if 10 nets are open for 10 hours, this totals to 100 net hours for that day. We can use net hours to account for effort in our analysis. Net hours have not been entered beyond 2016, so all analysis involving net hours stops at 2016 for this reason.

Next I needed to pair down the net hours data so that it matched the capture data and could be used to account for effort. To do this, I filtered net hours down to just include fall and made sure that the date was reading in as a date in R. I then summarized it down to net hours per year.

## Figure 1. How have total captures changed over time?

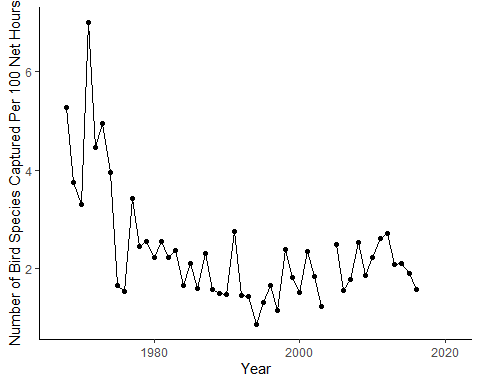
For my first figure, I wanted to look at how total numbers of captures have changed over time. First I needed to filter the fall dataset down to only data before 2016 (to match the net hours) and then summarized by year. I then merged the net hours dataset into the yearly capture dataset by the BANDING\_YEAR column. To calculate effort, I then divided the number of yearly captures by net hours and multiplied by 100. This is what I graphed for this figure.



I expected a slow decline over time between 1968 to present, but was not expecting such a steep decrease in bird captures in the late 1970s. I did not run a linear model on this figure because the data was far from linear and will require a more complex model to determine what is causing these decreases over time.

## Figure 2. How has the number of species captured per year changed over time?

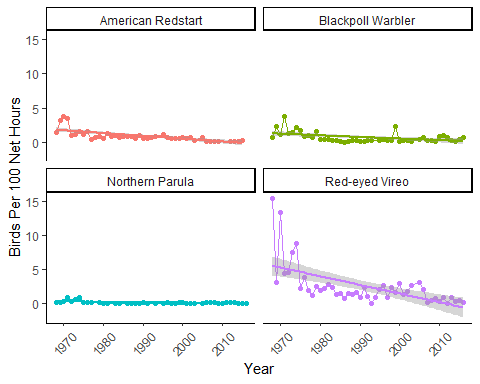
For this figure, I simply grouped the fall data by banding year and summarized the number of distinct species per year. I accounted for effort by dividing the number of species per year by the total yearly net hours and multiplying by 100. I then plotted it with year on the x axis and number of species per 100 net hours on the y axis.



I did not account for effort originally in this analysis but after seeing the amount of variation, I went back and divided the number of species by the net hours and multiplied by 100 to make the numbers more round. Once you account for effort, we end up seeing a pattern that is quite similar to the first figure, with a steep drop in the 1970s. This made me curious to dig more into specific species to see if this pattern continued.

## Figure 3. How does the abundance of long distance migrants change over time?

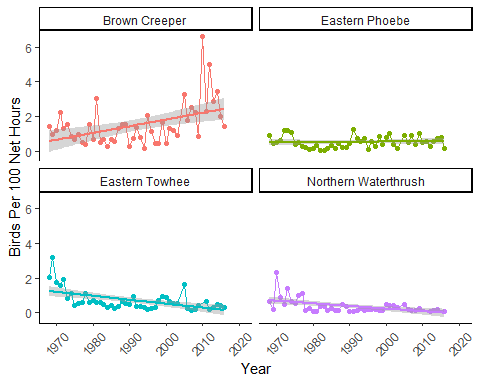
The next few figures all use the same approach. I took a subset of the species we catch on Block Island, based on migratory strategy, and put them next to each other to see how different common migrants have changed over time. Migratory strategies were taken from a recent paper by Kamm et al. 2019. All of these figures include data that was filtered to only include the species of interest. Effort was calculated as the number of birds of a given species/effort \*100. I classified long distance migrants as birds that winter in Central America or South America.



Interestingly, Red-eyed vireos seem to show a very similar pattern to the two previous figures, with a sharp decrease in the mid-1970s. It is challenging to see any real pattern with the other three birds because capture numbers of these species are quite low every year. That said, it appears that populations of redstarts, parulas, and blackpolls are relatively stable on Block Island.

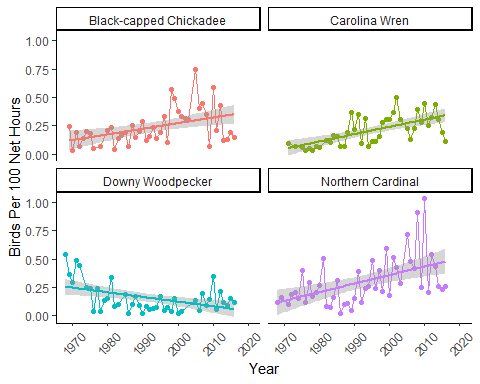
## Figure 4. How does the abundance of short distance migrants change over time?

This figure used the exact same approach as Figure 4, except I only focused on short distance migrants. Short distance migrants were classified as birds that overwinter in the southern US.



Eastern Towhees and Northern Waterthrushes seem to also mirror Figure 1 and 2 with a drop in numbers in the 1970s. Eastern Phoebes do not have a distinct pattern over time. Interestingly, Brown Creepers seem to be increasing over time. Further analysis is necessary to determine why (outside of the scope of this project), but my prediction is that this may be due to the higher vegetation height on Block Island that has been noted by the banders in recent years.Further analysis will be needed to look more into this prediction.

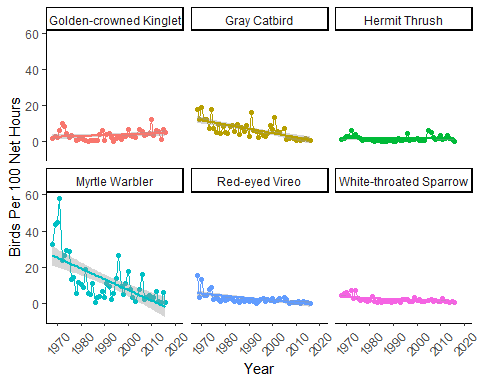
## Figure 5. How does the abundance of resident birds change over time?

This figure used the exact same approach as Figure 4 and 5, with that exception that it is focused on resident birds, or birds that are present on Block Island year-round. I was curious about resident birds because they do not have to undertake long, strenuous migrations, and I predicted that they would likely not experience the same declines that other more vulnerable species experience. 

Three out of the four resident species that I looked at seemed to be largely increasing overtime, though there is still a great deal of variation. Interesting we still seem to see cyclical boom and bust years, even as bird abundance is increasing. All of these figures are not easily represented by a linear model because of all of the variation in the data, but I decided to include the linear model simply to help visualize whether populations are increasing or decreasing over time, and if so, how rapidly.

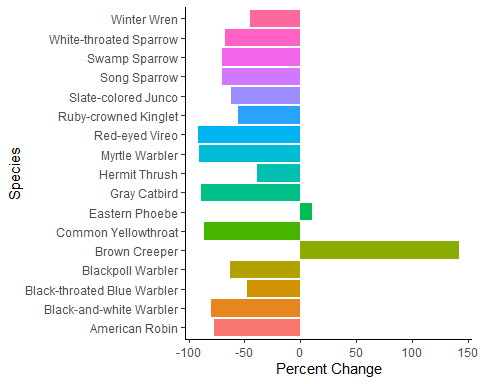
## Figure 6. How as the abundance of the six most common birds on Block Island changed over time?

In this case, “most common” simply means that bird species that are most commonly caught on Block Island. This was determined by tallying the number of captures of each species across years and sorting so that I could pull out the top six species. From there, I was able to follow the same methods from previous figures. I filtered for the species I was interested in, counted up yearly capture totals, merged that data with the net hours, and calculated birds per 100 net hours to account for effort.



## Figure 7. What is the percent change of bird abundance from the 1970s to present?

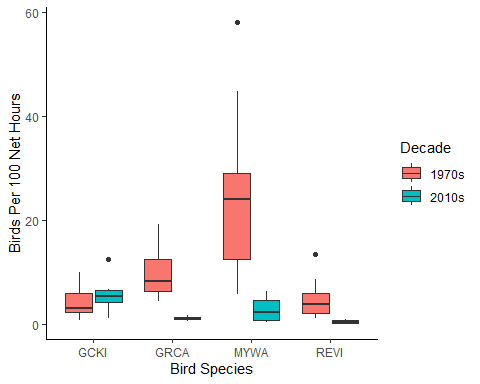
For this analysis, I created two separate dataframes. The first was filtered to just be 1970 (the beginning of my data) to 1980 to cover the span on one decade. The second was filtered to only include 2010 to 2020 (the most recent decade of data that I have). Once I had both dataframes filtered, I was able to count the number of observations in the dataframes by banding year. After this, I then took the averages of those counts across each time interval. Once both dataframes were complete, I merged them together based on species name, removed all NAs, and calculated the percent change.



This is consistent with the other results I found, where Brown Creepers are the only species increasing for some reason. Really interesting. I’d be curious to look more at vegetation changes over time in my future analysis. Creepers are tree-dwelling species and Block Island traditionally has very few trees, though the vegetation is getting a lot taller and might be more conducive to foraging for a Brown Creeper. Really interesting!

## Figure 8: Has the number of bird captures of the four most common bird species on Block Island changed significantly from the 1970s until present?

For my last figure, I made a boxplot comparing bird captures in the 70s and bird captures in the 2010S. This is similar to my percent change analysis but it’ll allow us to make more direct comparisons and run stats! I only ran this analysis on the four most common birds: GRCA, GCKI, REVI, AND MYWA. To do this, I created a dataframe that was filtered down to only the two decades I was interested in. I then added a “decades” column that designated which decade it was from. All data was accounted for effort. I could then plot a boxplot with “decade” as the fill. Finally, I was able to create a dataframe with the same data (accounted for effort) but with each decade in a different column. This allowed me to run a paired t-test on my four species to see if the differences were significant.



All four species had significant differences between decades. REVIs, MYWAs, and GRCAs all significantly decreased over time, while GCKIs increased slightly (but significantly) over time.

## References

Kamm, M. D., T. L. Lloyd-Evans, M. Handmaker, and J. M. Reed. 2019. A half-century of changes in migratory landbird numbers along coastal Massachusetts. PLOS ONE 14:1-18.