DISSERTATION

SPATIAL AND TEMPORAL PATTERNS OF ABUNDANCE OF SWALLOW AND MARTIN COMMUNAL ROOSTS

Submitted by

Maria Carolina Tiburcio Dias Belotti

Department of Fish, Wildlife, and Conservation Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2025

Doctoral Committee:

Advisor: Kyle G. Horton

David Koons
Larissa Bailey
Matthew Ross

ABSTRACT

SPATIAL AND TEMPORAL PATTERNS OF ABUNDANCE OF SWALLOW AND MARTIN COMMUNAL ROOSTS

During their non-breeding period, many species of swallows and martins (family: Hirundinidae) congregate in large communal roosts, which can gather hundreds of thousands of individuals. These roosts are well-known within local birdwatching communities; however, monitoring them at large spatial scales and with day-to-day temporal resolution is challenging. Due to the high densities of birds within these aggregations, however, their early morning dispersals can be systematically detected by nearby weather radars, which can be used to collect data about roost timing, size, and location. Nonetheless, finding roost signatures amongst the millions of rendered reflectivity images is extremely time-consuming — a fact that has limited the spatial and temporal scopes of previous radar-based studies. We leveraged the recent advances in computer vision and high-performance computing to partially automate this task and build a dataset of 22 years of roost detections captured by 12 radar stations in the Great Lakes region. We first verified that these detections correspond to swallow and martin roost dispersals, comparing the phenology of our findings with that obtained from eBird data. We then describe changes in the roost size distribution throughout the season and discuss the relationship between a roost's size and its persistence in the landscape. We also obtained trends of roost activity, comparing those with breeding population trends from eBird and the North American Breeding Bird Survey. Lastly, we used two years of data (2014-2015) collected by an operational weather radar in Manaus to describe, for the first time, the phenology of swallow and martin aggregations in the Low Negro-Solimões region of the Amazon Rainforest. We compared roosting behavior in the Amazon with what we observed 6,000 km away, in the Great Lakes region, regarding the daily number of birds and the roost size distribution.

ACKNOWLEDGEMENTS

Chapters 1 and 2

The conceptualization, data collection, and analysis of these chapters was done in collaboration with Wenlong Zhao, Yuting Deng, Victoria F. Simons, Gustavo Perez, Jeffrey F. Kelly, Subhransu Maji, Daniel Sheldon, and Kyle G. Horton.

Chapter 3

This chapter used the same dataset created and analyzed in co-authorship with Wenlong Zhao, Yuting Deng, Victoria F. Simons, Gustavo Perez, Jeffrey F. Kelly, Subhransu Maji, Daniel Sheldon, and Kyle G. Horton. The analysis of trends of swallow and martin populations was conceptualized and implemented in collaboration with Brian Gerber.

We also thank the developers of `ebirdst` and `bbsBayes2` for making their code and data available and for writing excellent documentation for their packages.

Chapter 4

The conception, data collection, and analysis of this paper was done in collaboration with Carlos Augusto Moreira Rodriguez, Victória de Souza Wojahn, and Kyle G. Horton.

We thank the Amazon Protection Service for making the radar data from Manaus available.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Pairing weather radars with eBird data for large-scale monitoring of swallow and martin communal roosts

1.1 Introduction

Swallows and martins (family: Hirundinidae) are highly specialized aerial insectivores which are mostly migratory, with some species traveling up to 20,000 km between their breeding and nonbreeding grounds. During the nonbreeding period, many species of this family congregate in large communal roosts that can gather hundreds of thousands of individuals (Ward & Zahavi, 1973; Winkler, 2006). These large aggregations can offer protection against predators and may act as information-sharing centers, where birds gather to learn about optimal foraging locations. The latter theory about communal roosting behavior is particularly well-suited for swallows and martins, since these central place foragers rely on ephemeral but locally abundant emergences of aerial insects (Lalla et al., 2022; Ward & Zahavi, 1973; Winkler, 2006). In the Americas, the timing and location of these aggregations are often predictable, with roosts occurring on the same site year after year. Roosts gathering Bank Swallows (*Riparia riparia*), Tree Swallows (*Tachycineta bicolor*), and Barn Swallows (*Hirundo rustica*), for example, have been reported as early as 1918 in Black River Bay, on Lake Ontario, where our data suggests they are still observed to this day (Belotti et al., 2023; Bent, 1942).

Monitoring these aggregations is crucial for the conservation of these species which, like other aerial insectivores, have been facing widespread breeding population declines for yet unknown reasons (Rosenberg et al., 2019; Spiller & Dettmers, 2019). Furthermore, understanding the drivers and constraints that dictate roost size and timing could help us elucidate questions related to the ecological function of these roosts and their downstream effects on regional food

webs. However, in spite of the size of these aggregations, finding them can be challenging. Birds usually arrive at the roosting location in relatively low concentrations at dusk, when lighting conditions are limited. Additionally, their much denser dispersal occurs early at dawn, before most human observers typically start their daily activities (Arruda, 2023; Russell & Gauthreaux, 1999). In densely populated areas, roost observations have grown in the past century, particularly with the rise of community science initiatives. In remote areas of the Amazon Rainforest, however, it is likely that some of these large gatherings of swallows and martins remain all but unnoticed by human eyes.

The Purple Martin Conservation Association (PMCA) hosts the most comprehensive dataset of roost locations in North America. Through their project MartinRoost, they collect reports of Purple Martin roosts submitted by volunteers throughout the United States. Each report corresponds to roost observations from one season, with start and end date, location, species composition, notes about observations made on the ground, and science dissemination opportunities associated with that roost — some roosts have become regional events, with roost watch parties gathering hundreds of families. In the past twenty years, the community science platform eBird has grown as yet another roost monitoring tool in North America (Sullivan et al., 2009). Birdwatchers on this platform submit checklists of observed birds at a given location, at a specific time and date. They also have the option of reporting the number of individuals observed for each species. Even though community science data can be used to obtain valuable ecological insights, analyzing it poses some important challenges. Because the sampling process is heterogeneous, there is significant variation in observation time, day of the year, and effort. Additionally, since remote regions can be difficult to access, community science datasets usually have a strong spatial bias towards regions with denser human populations (see Johnston et al., 2021, for a review).

Weather radars can address some of these limitations. Due to the high densities observed in roosts and the characteristic dispersal shape of swallows and martins, their early morning departures produce a distinctive doughnut-shaped signature on data collected by nearby weather

radars, which can be used to quantify aggregation size, location, date, and timing (Russell & Gauthreaux, 1998, 1999; Russell et al., 1998). Previous efforts to find these signatures and collect data from them have allowed us to map 33 roost locations throughout the southeastern United States for the first time (Russell et al., 1998). We have also confirmed, by pairing radar-detected roosts with community science reports, that swallows and martins are the main species found in these aggregations (Kelly & Pletschet, 2017; Kelly et al., 2012). Additionally, a study found that roosts in the United States occur on four main land cover types, namely urban, forest, water bodies, and croplands, and that urban roosts were typically more persistent throughout the years when compared to roosts observed in other land cover types (Bridge et al., 2015).

Until recently, however, our ability to use radar data to obtain fine-scale spatio-temporal data on roosting behavior was limited by the fact that finding roost signatures on millions of rendered reflectivity scans is extremely time-consuming. Nevertheless, recent advances in computer vision and high-performance computing have allowed us to streamline some of this effort. A machine learning pipeline developed by Cheng et al. (2020), capable of detecting roosts on radar data and tracking them across successive scans, allows us to leverage the broad spatial and temporal coverage of the network of weather radars to understand the communal roosting behavior. By employing this technique to find roosts, human supervision is limited to removing false positives and labeling tracks according to how much contamination from precipitation and other sources of radar noise is caught within them. This model has since been improved, further reducing the human supervision effort by improving the model's accuracy (Perez et al., 2022), and new techniques have been developed to estimate cumulative curves of roosting activity throughout the season with less manual screening (Perez et al., 2024).

These new advances allow us to gain spatial and temporal reach in our study of swallow and martin communal roosting behavior. However, they also require that we match our conceptual understanding of this behavior with what the technology we use to study it is observing and measuring. This is particularly true with regards to species identification. In our labeling protocol, we defined a series of morphological characteristics of roosts collected from the litera-

ture (Kelly & Pletschet, 2017; Kelly et al., 2012; Laughlin et al., 2013; Russell & Gauthreaux, 1998, 1999; Russell et al., 1998). We aim to understand if roosts selected with this protocol could really correspond to observations made on the ground by community scientists. We do so by comparing the timing of radar-detected roosts with that observed on eBird data for seven species known to form large aggregations (Common Grackles (*Quiscalus quiscula*), Red-winged Blackbirds (*Agelaius phoeniceus*), Cliff Swallows (*Petrochelidon pyrrhonota*), Purple Martins (*Progne subis*), Tree Swallows, Barn Swallows, Bank Swallows). We also wanted to assess the quality of the spatial coverage of these two sources, which we measured by how many roosts they can detect per day.

1.2 Materials and Methods

1.2.1 Radar Data Processing

In the Great Lakes region, swallow and martin roosts typically occur in the summer, between July and August (Kelly & Pletschet, 2017; Russell, 1996). We downloaded and rendered all radar scans available between June 1st and October 31st for the 12 radar stations within our study region, from 2000 to 2022. We focused on the period between 30 minutes before local sunrise and 90 minutes afterwards, when we typically observe roost departures. These data are publicly available in an Amazon Web Services server (https://registry.opendata.aws/noaa-nexrad). Radial velocity and reflectivity in these scans were used as input to the machine learning pipeline developed by Cheng et al. (2020), which leverages the capabilities of a Faster R-CNN model coupled with a Kalman Filter to detect and track roost dispersals in radar data (Cheng et al., 2020). The pipeline outputs collections of bounding boxes (so-called tracks), each collection representing a single dispersal event. We designed a manual screening protocol used by three human annotators (MB, YD, and VS) to label each track according to the intensity and type of noise contamination within each track, if there was any. The screeners also labeled days when precipitation, anomalous propagation, or other sources of radar noise were so intense that we wouldn't have been able to see the roosts if they were present. Since we were interested in the

presence/absence of roosts for this study, we did not remove tracks with noise-contaminated dispersals.

Some of the weather radars selected for this study have overlapping ranges and it often occurred that the same roost was detected by two radars. We account for that by grouping overlapping roosts occurring on the same day. We did this by first obtaining the union of all the bounding boxes for each dispersal. We then calculated the ratio between the area of the intersection and the area of the union of each pair of overlapping tracks occurring on the same day. If this ratio was greater than 0.2, the tracks were considered to represent the same roost dispersal. We refer to these grouped tracks as unique detections or unique radar tracks. We determined the location of each roost dispersal by obtaining the center of the first bounding box on each unique detection.

1.2.2 eBird Data Processing

We downloaded all the eBird checklists made from 2000 to 2022 that reported at least one of the five swallow and martin species known to join roosts in the United States, namely: Tree Swallows, Barn Swallows, Bank Swallows, Cliff Swallows, and Purple Martins. Additionally, we obtained eBird checklists for Common Grackles and Red-winged Blackbirds since these two species also form large aggregations and are sometimes reported to join swallow and martin roosts. Our initial filtering process removed checklists with the same unique sampled identifier, indicating that they were made by groups of observers. We selected checklists made after 16:00, when roosts start to form, and before 10:00, when roosts disperse. These were further filtered to keep only checklists reporting more than a thousand individuals, indicating high aggregations of birds. Lastly, we removed checklists that were outside the range of 150 km of the twelve radar stations in the Great Lakes region.

If many individuals observed the same roost and were not part of a single observation group (reported on eBird with a unique sample identifier), then that single roost was associated with multiple checklists. Because we were interested in unique roost detections, we grouped check-

lists made at the same time within 1 km of each other. We also grouped observations made in the morning or in the evening of the previous day in the same location (within 1 km of each other), to reflect the fact that the same roost can be detected on two different days. We refer to each observation resulting from this grouping procedure as a unique detection or a unique roost observation.

1.2.3 Comparison

In our comparison of the timing of radar observations and eBird checklists, we extracted the number of unique detections per day of the year from each source and for each species in the case of eBird data. With this, we could extract the peak roost detection dates for both sources. Additionally, we calculated the cumulative number of unique detections per day for each source between June 1st and October 31st, grouping checklists for all species of swallows and martins and for Common Grackles and Red-winged Blackbirds. We filtered eBird data from years after 2010 in order to ensure we had enough days with detections for each year. We standardized the cumulative curves dividing them by the summed count per year. We fit a logistic smoother to these data to find a single representation of the phenology of radar, swallow and martin, and grackle and blackbird detections. Finally, we performed Pearson correlation tests on the detrended (differenced) smooth curves to assess the similarity between radar detection phenology and the phenology of swallow and martin detections and blackbird and grackle detections.

When assessing the quality of the spatio-temporal coverage of both sources, we compared the distribution of the number of unique detections per day for each source within the same region, during the same periods. We did this by further filtering the eBird dataset to remove observations made before June 1st or after October 31st, when the radar data were not processed. We also removed checklists submitted when the radars were not sampling due to heavy precipitation, anomalous propagation, other sources of unidentified noise, or files missing from the archive. Finally, when roosts occur within 5 km of the radar station, they do not appear with the

characteristic ring shape, because the radar samples them at very low elevations. This causes the machine learning pipeline to fail to detect them. Because of this, we omitted unique eBird detections made within 5 km of the radar station.

1.3 Results

After our post-processing of eBird data, we were left with 788 unique detections of swallow and martin aggregations made throughout the year, from 2000 to 2022, regardless of radar sampling. Of those, 94% were associated with checklists reporting high counts of a single species: 256 (32% of total) reported roosts gathering only Purple Martin, 211 (27%) detections were of Tree Swallow aggregations, 164 (21%) were Bank Swallow roosts, 86 (11%) were roosts containing exclusively Barn Swallows, and 19 (2%) unique detections reported aggregations of Cliff Swallows alone. The remaining 52 (7%) unique detections referred to checklists reporting high counts of multiple swallow and martin species. Most of the swallow and martin checklists with high counts were made during our sampling window from June 1st to October 31st when the radars were sampling, that is, on days where the data was available and not fully contaminated by weather or anomalous propagation (n = 583, 74%, see Figure 1.1).

The seasonality of the number of detections per day for both sources was similar for these species, however eBird detected some large numbers of swallows and martins between April and June, before the start of the roosting season in July. Large counts of Common Grackles and Red-winged Blackbirds have a very different seasonal pattern from that observed for swallows and martins, with two peaks of detections, one before and one after their breeding season (see Figure 1.2). The analysis of the cumulative detection curves for radar data and for our two groups of study species shows how the phenology of radar detections matches what we observe for swallow and martin high count checklists (Pearson correlation coefficient 0.79, p-value < 0.00001) and is significantly different from the phenology of Common Grackles and Red-winged Blackbirds (Pearson correlation coefficient -0.29, p-value = 0.00024, see Figure 1.3).

We found that the average number of unique detections per day for the radars was 8.31 (sd: 7.10), while the average number of eBird unique detections per day was 1.55 (sd: 1.00, see Figure 1.4). The maximum number of unique detections made by the radars on a single day (38) was 4.75 times larger than the same metric for eBird (8). The radars detected more roosts than eBird on 95% of the sampled days (see Figure 1.4). Each year, the maximum number of unique eBird detections reported on days sampled by the radars (the peak detection date per year) was reached between July 16 and August 29 across all years with more than 2 unique detections per day (n = 10). The median peak detection date occurred on August 16 (see Figure 1.1). The peak of detections across all years (n = 23) for the radars varied between August 4 and August 24, with an average number of detections at the peak dates of 27.4 (sd: 4.2). The median peak date across all years occurred on August 15 and 17 (see Figure 1.1).

With regards to the last report of the roosting season, we found some variation across species on eBird data. The last report of the season for Bank Swallows occurred between July 29 and August 28, with the median happening between August 20 and 21. For Barn Swallows, the last high-number report happened between August 4 and September 16, with the median occurring on August 26. The last report of Purple Martin roosts occurred between August 27 and September 15 across all years, with the median happening in early September, between September 6 and 8. Lastly, for Tree Swallows, the last report occurred between September 6 and October 26, with the median happening between October 10 and 12 (see Figure 1.1).

1.4 Discussion

The phenological similarities between radar detections and swallow and martin eBird detections suggest that these species are likely the ones joining the radar-detected aggregations, indicating that the machine learning model, paired with human supervision, is actually capturing the communal roosting behavior observed in previous studies (Kelly & Pletschet, 2017; Kelly et al., 2012). It is likely, however, that some of the activity observed in the later part of the season is driven by Red-winged Blackbirds and Common Grackles, which would also explain why

the radar detections per day never return to the zero baseline after the roosting season peaks. Further studies need to take this into account, particularly when examining phenology changes at the end of the roosting season. A thorough study of the morphology of the dispersals of different species including with regards to other data products such as radial velocity, differential reflectivity, and differential phase, could help us understand if we can further refine the radar's taxonomic resolution.

Interestingly, we observed large counts of all species but Purple Martins before the main post-breeding roosting pulse. The nature and function of these aggregations is not yet well-established. It is unlikely that these are breeding colonies, both because of the large counts and because of the brief pause in occurrence before the post-breeding roosts start. On the other end of the season, we notice that Tree Swallows remain roosting until late October, unlike all other species of swallows and martins. In fact, eBird data suggests that Bank Swallows are the first to disappear from the roosts at the end of August, followed by Barn Swallows, who also stopped being detected in late August. Purple Martins are the next to leave, in mid-September, and finally Tree Swallows typically stay for another month, and only stop being detected in mid-October. On eBird data, the end of the roosting season for swallows and martins is not confounded with the start of the aggregations of Common Grackles and Red-winged Blackbirds. For that reason, we suggest that the two data sources could be integrated in phenology models to obtain a more accurate definition of when communal roosts of swallows and martins stop occurring every year, much like what was done in Horton et al. (2019) and Horton et al. (2018).

However, establishing the true peak roosting date from eBird data alone might be challenging. On most years, the peak number of unique detections did not exceed 3, which leaves a narrow margin for over or underestimation of the total number of roosts due to sampling biases. This was not the case with the radar dataset, in which the average number of roosts detected at the peak date was 27.4 with a relatively low standard deviation of 4.2. Probably because of this, we also observed more variation on peak detection date throughout the years on eBird — 47 days from July 16 until September 1st — when compared to the range of peak detection dates

on radars — 20 days between August 4 and August 24. Additionally, the low standard deviation in the number of detections on peak dates seen in the radar data suggests that they are indeed detecting most roosts within their range.

Lastly, the difference between the number of roosts detected per day by eBird and the weather radars suggests that the latter can offer a much broader coverage of the airspace, helping us locate rare aggregations that are hard to find if observers are not directly looking for them. Additionally, as will be shown in the rest of this dissertation, with the methodology established by (Chilson et al., 2012),weather radars allow us to systematically estimate the size of each communal roost without running into some of the biases associated with visual counts of high numbers of birds (Davis et al., 2022).

1.5 Figures

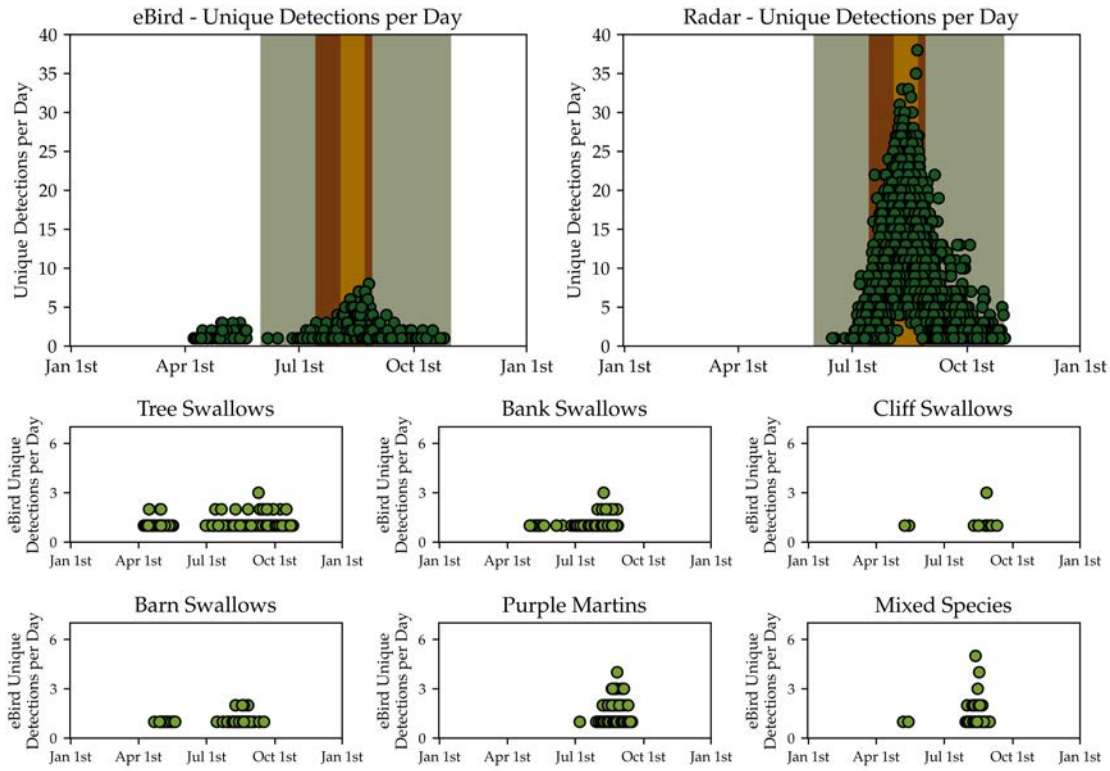


Figure 1.1: Plots representing the number of unique detections of swallow and martin roosts per day of the year, across all years, from 2000 to 2022 (eBird on the top left, radar on the top right). The plot on the top left aggregates the eBird detections per species represented on the plots bottom. The plot with "Mixed Species" shows numbers of unique detections per day reporting at least two of our target swallow and martins species. The shaded areas represent the days sampled by the radars (grey), the range of peak detection dates across all years on eBird (brown) and the range of peak detection dates across all years on radars (in yellow).

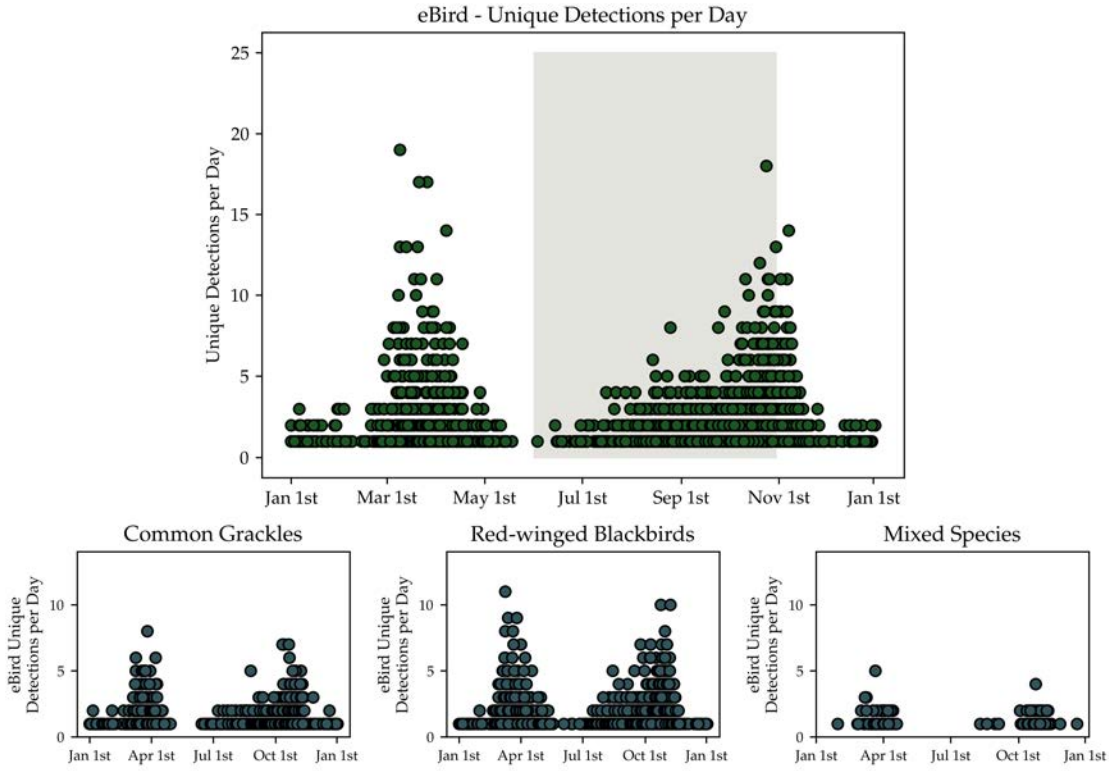


Figure 1.2: Plots representing the aggregated number of unique detections of Common Grackles and Red-winged Blackbirds per day of the year, across all years, from 2000 to 2022. The two plots on the bottom represent the same variable for each species and for unique detections reporting both species (mixed species).

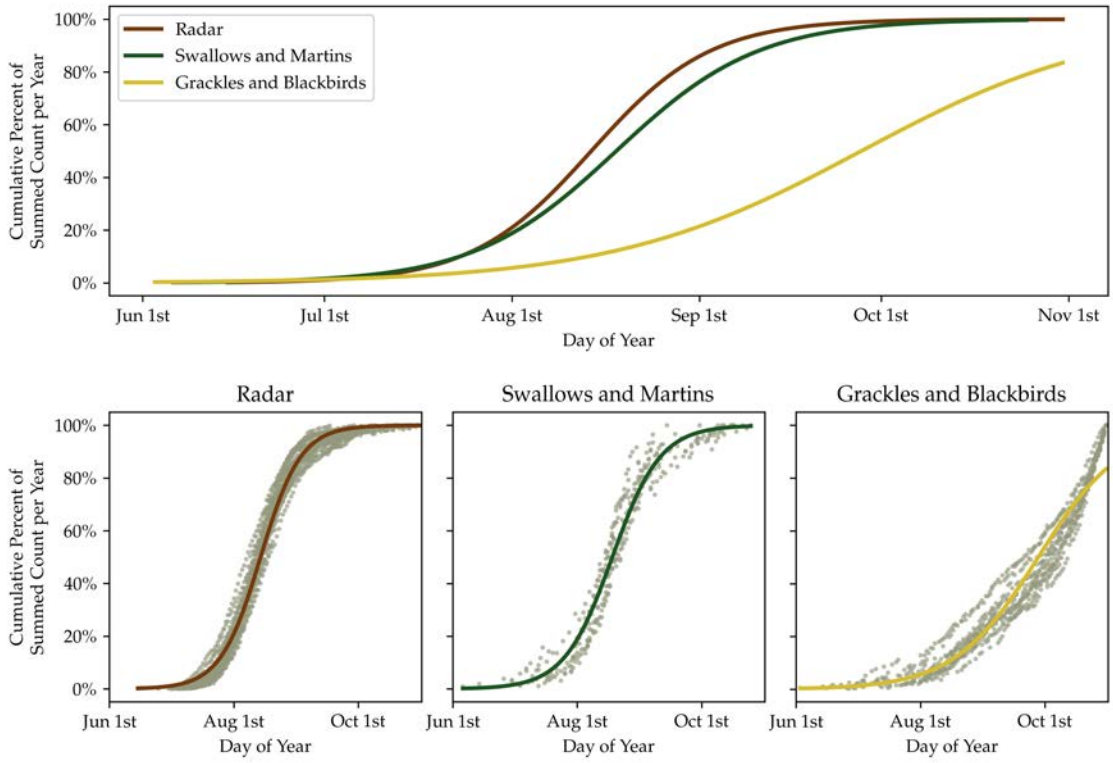


Figure 1.3: Plots at the bottom represent standardized cumulative detections per day of the year across all years (gray background points) and the smooth curve obtained fom fitting a logistic regression to the points. The upper plot contains just the smooth curves for comparison. The Pearson correlation coefficient between the detrended curves obtained for swallow and martin eBird detections and the radar detections was 0.79 (p-value < 0.00001), whereas the correlation coefficient between radar detections and grackle and blackbird eBird detections was -0.29 (p-value = 0.00023).

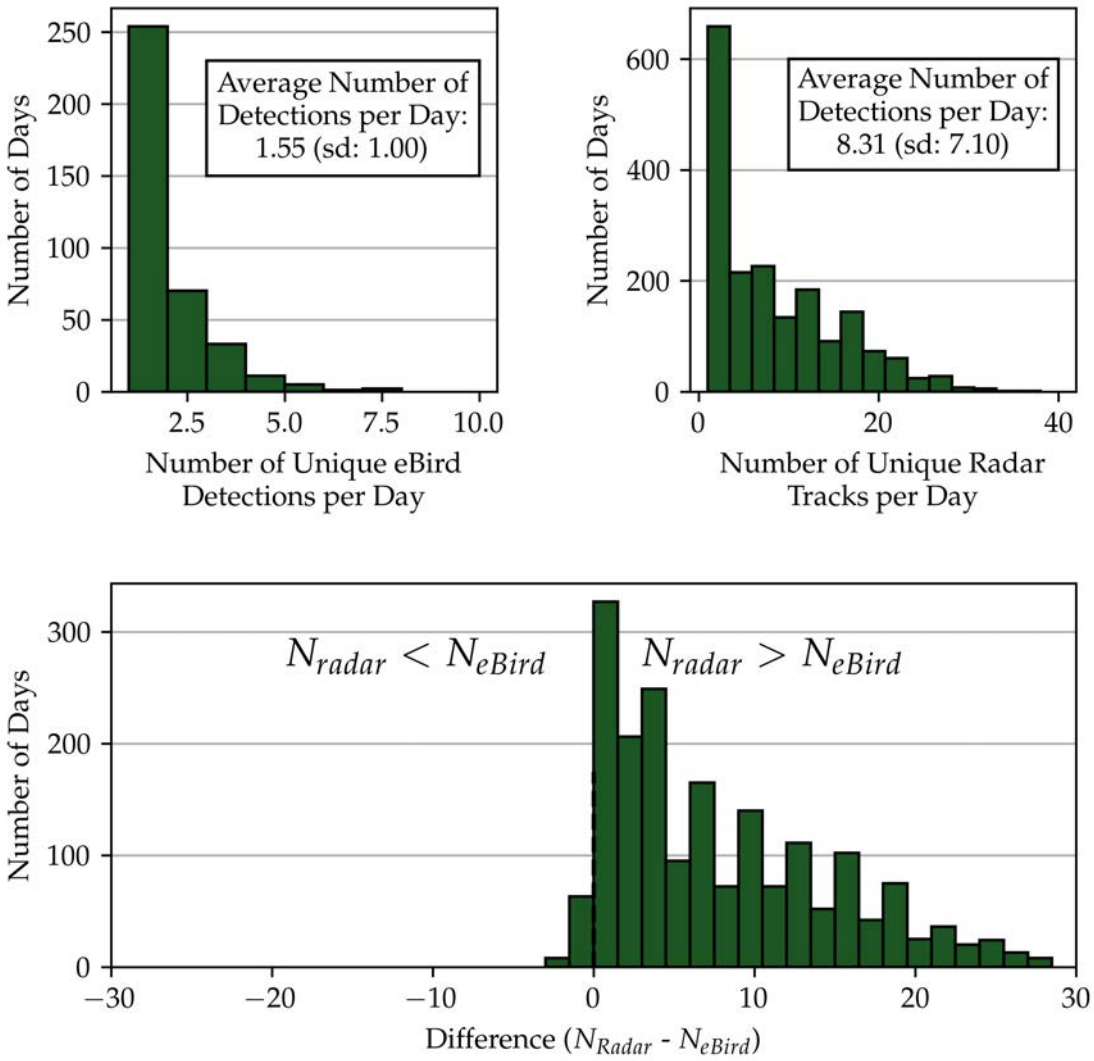


Figure 1.4: Plots at the top are histograms of the number of unique detections per day for the radars (left) and for eBird (right). At the bottom, we plot the histogram of the difference between the number of unique tracks per day and number of unique checklists per day. The number of radar detections per day was larger than the number of eBird detections per day on 95% of days.

Chapter 2

Long-term analysis of persistence and size of swallow and martin roosts in the US Great Lakes[1]

2.1 Introduction

Swallows and martins (family: Hirundinidae) are highly specialized aerial insectivores that, like other species that rely on insects as their main energy source, have been facing regional and continental-wide declines in their breeding populations in the United States and Canada (Rosenberg et al., 2019; Smith et al., 2015). Some of the potential drivers of these declines include habitat loss, decreases in prey abundance, impacts from environmental contaminants, and phenological mismatches due to climate change (Cox et al., 2019; Hallmann et al., 2014; Kardynal et al., 2020; Spiller & Dettmers, 2019). Particularly in the case of the migratory species in this family, it is still a challenge to understand how these drivers of decline affect different stages of their annual cycles and how these effects interact between seasons to produce the overall pattern observed in their breeding populations (Holmes, 2007; Marra et al., 2015; Webster et al., 2002). Gathering metrics of population size and the variation therein across broad spatial and temporal scales would have important conservation implications, while also providing insights into the fundamental ecology of these species.

Recent advances in tracking technologies have allowed us to address some of these questions, but they are limited to individual-level data over relatively short periods (Supp et al., 2021; Wikelski et al., 2007). Due to these limitations, tools such as community science data (iNaturalist; Wikiaves; Sullivan et al., 2009), stable isotope analysis (Imlay et al., 2018), and weather radars (Dokter et al., 2010, 2018; Horton et al., 2019) have recently been ascending as alterna-

[1]This chapter was published in: Belotti, M.C.T.D., Deng, Y., Zhao, W., Simons, V.F., Cheng, Z., Perez, G., Tielens, E., Maji, S., Sheldon, D., Kelly, J.F. and Horton, K.G. (2023), Long-term analysis of persistence and size of swallow and martin roosts in the US Great Lakes. *Remote Sensing in Ecology and Conservation*, 9: 469-482. https://doi.org/10.1002/rse2.323.

tives to observe population-wide patterns at large spatial and temporal scales. Particularly in the case of swallows and martins, weather radars are an invaluable tool to collect data about their communal roosting behavior, though until recently we lacked the tools to leverage the potential of this data source.

After their breeding season, from mid-July until the end of September, species of the Hirundinidae family congregate in large communal pre-migratory roosts, which can tally more than 100,000 birds at the end of each day (Allen & Nice, 1952; Bent, 1942). The early departure of thousands of birds in synchrony from these roosting places is so conspicuous that they can easily be identified in weather radars, where they appear as donut-shaped rings that disperse outwards in the hours near sunrise (Russell et al., 1998). While the primary ecological importance of these roosts is largely unknown, especially prior to a long-distance migration, these behaviors offer a unique and important snapshot to monitor population changes. The main challenge of using weather surveillance radars to obtain data about these aggregations is manually finding the roosts in the collection of millions of radar scans generated across decades. Additionally, establishing systematic criteria for classifying the multiple sources of noise and weather contamination inside the roosting signature is complicated by the vast diversity of patterns captured by the radars, especially at large spatial scales.

However, machine learning approaches have the potential to alleviate this barrier by disentangling the roost signatures and automatically discerning their locations and the boundaries of their dispersal (Cheng et al., 2020). With such tools, we can now analyze a dataset that spans two decades and can provide high spatial and temporal sampling frequency. We decided to test this approach in the Great Lakes region, where multiple species of Hirundinidae have been shown to roost. Large congregations of Purple Martins (*Progne subis*), Tree Swallows (*Tachycineta bicolor*), Barn Swallows (*Hirundo rustica*), Bank Swallows (*Riparia riparia*), and Cliff Swallows (*Petrochelidon pyrrhonota*) are regularly seen in this region between July and September (Bridge et al., 2015; Kelly et al., 2012) Burney, 2002. Additionally, we have found historical accounts tracing back to 1910 of roosts gathering Tree Swallows and Bank Swallows at

the Black River Bay, in Lake Ontario, near Toledo, Ohio, and at Cedar Point, also in Ohio (Bent, 1942, pages 396 and 420). With more than two decades of radar remote sensing data available throughout this region, we set out to use these newly available tools to identify the locations of Hirundinidae roosts and provide some of the first estimates of macroscale abundance.

With these data, we set out to quantify how consistently the population in our study region returned to the same regions to roost over the years (which we refer to as "population-level site fidelity"), and whether site usage consistency is related to roost size. Investigations of the spatial and temporal patterns of site fidelity can help us understand the role played by competition and social dynamics as drivers of nonbreeding habitat selection in migratory bird species (for an example see Marra & Holmes, 2001; Sherry & Holmes, 1996). Additionally, if site fidelity is driven by adults who return to good habitats based on experiences from previous seasons, then examining site fidelity patterns can provide indirect information about adult survival and overall population dynamics (Cresswell, 2014). Lastly, populations that strongly rely on a specific region during their migratory journey might be particularly vulnerable to anthropogenic impacts on that site, making them a species of conservation concern (Warkentin & Hernández, 1996).

Finally, an important step in understanding the ecological implications of such post-breeding aggregations is characterizing within-season variations in their size distribution. Up until now, the techniques available to quantify the sizes of these large roosts suffered from observer variability, and thus were hard to reproduce from year to year, or even on a daily basis (Erwin, 1982; Frederick et al., 2003). With the use of weather radars, we can not only estimate the number of birds within each roost, but also examine within and between-season changes. We describe the wide range of roost sizes found in the Great Lakes region and how this distribution changes as the overall population increases throughout the season.

2.2 Methods

2.2.1 Data Source

Within the contiguous United States, the National Weather Service, the Federal Aviation Administration, and the United States Air Force jointly operate a network of 143 S-band weather radars that conduct volume scans of the atmosphere at regular intervals (from 5 to 10 minutes). Each scan is composed of multiple 360° horizontal sweeps of the radar antenna taken at different elevation angles. Scans are organized in a 3D array following the CfRadial convention in which cells, indexed by the horizontal angle of the antenna (each angle is called an azimuth), by distance from the radar (called range), and by elevation angle corresponding to measurements of the returned signal from a voxel of the volume covered by the radar beam (see Dixon & Lee, 2021, for a detailed specification of the CfRadial standard).

For the purposes of this paper, we downloaded all radar scans from the Great Lakes region (12 radar stations) available between June 1st and October 31st across 21 years (from 2000 to 2020, see Figure 2.1a and Appendix B, Figure B.3, for data completeness information). Each day, we processed the data in the period starting 30 minutes before local sunrise and ending 90 minutes afterwards (an average of 16.2 scans per day). Reflectivity factor and radial velocity from the first sweep of each scan were rendered in Cartesian grids with 600 x 600 pixels (1 pixel = 0.5 km). These matrices were then used as input to the pre-trained model described below. Each station's local sunrise time was calculated using `PyEphem` and `pytz` Python packages.

2.2.2 Automated detection and screening

The machine learning system developed by Cheng et al. (2020) operates in two steps – detection and tracking – following the paradigm proposed by Ren et al. (2015) (see Figure 2.1b). In the detection step, given some candidate proposal regions defined by positions in a radar scan and reference roost sizes, we trained a Faster R-CNN (acronym for Region-based Convolutional Neural Network) detector to predict the probability that a roost of each candidate size exists near each reference position (Ren et al., 2015). A high predicted probability indicates a roost

detection according to the model. The detector also predicts bounding boxes for the detected roosts. To ease implementation, we leveraged the "R101-FPN 3x" Faster R-CNN model from Detectron2, an open-source library in PyTorch (Wu et al., 2019). We picked 45 anchors at each position and otherwise used the default model configurations from the library. We loaded the model checkpoint trained for the COCO detection task from Detectron2 and fine-tuned it using the annotated training dataset developed for the paper by Laughlin et al. (2016). Since different annotators of the training data differ in their annotation styles, we adopted the factors from Cheng et al. (2020) to scale bounding box annotations, thus alleviating annotator biases. We used an Adam optimizer with a learning rate of 0.001 to train the detector (Kingma & Ba, 2014). For each scan, we kept 100 top-scoring roost detections with predicted probability scores of at least 0.05.

In the tracking stage, since the same roost is expected to show up in temporally consecutive scans, we reproduce the strategy described by Cheng et al. (2020) to predict the association among roost detections from individual scans. Finally, when possible, we removed rain false positives using dual-polarization data (specifically correlation coefficient) and wind farms according to the U.S. Wind Turbine Database (Hoen et al., 2018). While polarimetric products provide a robust filter for the detection of weather contamination, they are not available throughout the entire temporal extent of our study (only from 2013 onward), and thus additional manual screening was necessary for removing potential non-biological contaminants.

We manually picked a low score threshold (0.05) to ensure high recall regardless of the low precision of the predictions since we wanted to make sure that the pipeline would capture all roosts. To remove the false positives and detections with weather or noise contamination from our dataset, we developed a screening protocol (Figure 2.1c) used by three different screeners (MB, YD, VS) to evaluate all the predicted positive detections (screening one year of data for a given station takes on average approximately 45 minutes, see A for details of the screening protocol). We classified each true positive track, composed of a series of consecutive detections, as either clear, weather contaminated, having more than 20% anomalous propagation, containing

other sources of unknown noise, or tracking failures, in which the tracking algorithm failed to follow the roosts (see 2.2 for examples). Days that had more than 50% of the reflectivity renderings filled with weather or anomalous propagation were discarded from the analysis.

2.2.3 Overlap Correction and Mean Shift Clustering

Some of the weather radars selected for this study have clear overlapping ranges (see Figure 2.3a). If a roost occurred inside this overlapping region, it could be detected by more than one station, which would result in duplicated tracks and inflated estimates of the number of birds. To remove this bias, for each track we computed the union of all its bounding boxes. We then grouped our dataset by day and calculated, for each pair of tracks, the overlap ratio, which we defined as the ratio between the area of the intersection of each pair of tracks and the area of their union (Figure 2.1d). If the overlap ratio was greater than 0.2, the tracks were considered to be in the same group (see Appendix B for details about the selection of the ratio threshold).

To establish regions of repeated use throughout the years, we captured the first detection of each group described above and used it as input to the mean shift clustering algorithm (implementation by Pedregosa et al., 2011, Figure 2.1e), which finds density-based clusters while making minimal assumptions about cluster shape (see Appendix B for detailed methods of parameter choice). With this approach, we assume that regions consistently used by birds to roost would have a higher density of points and thus be included in the same cluster. We chose to use the first detection of each group because we observed that these were the most precise representatives of the actual roost's location. The cluster labels were then extrapolated to other points within each group, which is why the final clusters overlap.

Up until this post-processing stage, we included the first detection of all the tracks classified by the screeners as true positives (bounding boxes that contained a verified roost dispersal), even the ones that were noise-contaminated or that the algorithm failed to track. These so-called "contaminated" tracks provide useful information about roost location, but they cannot be used to calculate estimates of number of birds, since it can't be assumed that all the scatterers

in them are exclusively birds. For this reason, from this step onwards they were removed from the analysis.

2.2.4 Reflectivity Extraction and Post-Processing

Following manual screening to remove contaminants, we used coordinates of the center of each roost and its radius, both estimated by the detector, to obtain the bounding boxes of each departure event. We used these data to extract the raw reflectivity factor values from Level II radar data in polar coordinates (each voxel's location is defined by its range and azimuth) across all elevation angles of each full scan. To estimate the number of birds per roost, we followed the approach established by Chilson et al. (2012): we converted reflectivity factor (Ze) measured at each range interval sample, originally in dBZ, to linear scale mm^6/m^3, and then transformed it into reflectivity, which can be interpreted as the density of scatterers in the atmosphere (units are cm^2/km^3). We can then multiply this result by the theoretical volume sampled by the radar beam at each range interval, and divide it by the radar cross section (RCS) of the scatterers assumed to be found in the roost (Figure 2.1f). In summing the results of these operations in the region inside each bounding box, we obtain a count estimate for that sweep.

We opted to obtain conservative estimates of the number of birds by choosing Purple Martins as our benchmark radar cross section, since they are the largest species of Hirundinidae found in North America. Though previous studies suggest that the majority of roosts in the region are Purple Martins (Bridge et al., 2015; Kelly et al., 2012), our ability to ascertain the species identification is still somewhat limited using radar data alone. The radar cross section of Purple Martins can be obtained from their mass – 51g, see Dunning (2008) – adopting the relationship between mass and RCS proposed by Horton et al. (2019): $\log(\text{RCS}) = 0.699x\log(\text{mass})$. Bank Swallows, the smallest species to participate in such aggregations, would yield counts approximately 2.6 times higher, since their mass is 13 g according to Dunning (2008).

The volume sampled by the radar is assumed to be shaped like a truncated cone with an axis aligned with the antenna's peak power axis, cut by two parallel planes at each range gate.

For data before the Super Resolution upgrade, the cone's apex angle was assumed to be 1° in width. After the 2007-2008 upgrade, we assumed an elliptical cone of 1° vertical beam width and 0.5° horizontal beam width (Torres, 2007). In order to further clean the data from spurious contamination, we removed voxels inside the detection's bounding box if the reflectivity factor in them exceeded 30 dBZ (Horton et al., 2019).

To prevent non-biological scatterers from the upper levels of the atmosphere from artificially increasing our estimate of the number of birds, we only analyzed sweeps that had heights below 5000 meters at the 150 km range. We counted birds across the entire vertical structure of the roost departure by grouping the multiple sweeps of each detection in bins of 1° intervals according to their elevation angle. Afterward, we took the mean bird count of each bin and summed the counts from all available bins. With this procedure, we hoped to avoid summing regions in the airspace sampled twice by consecutive sweeps that had elevation angles with less than 1° interval between them.

2.2.5 Analysis

To describe the peak activity in each group of tracks obtained as described above, we extracted the maximum number of birds in each series of detections. We then aggregated groups of tracks by day and summed the maximum bird counts obtained from roosts occurring on the same day. We did this assuming that we can sum the maximum bird count for each group without double counting since we have corrected most overlapping tracks. In doing so, we have a measurement of the maximum number of birds per day in the study region. We further aggregate this by cluster and year and compute the maximum number of birds per cluster per year, which we interpret as the peak number of birds throughout the region each year. To obtain a cluster's persistence, we sum the number of years when each cluster was detected (Figure 2.1g).

Because radar stations had incomplete archives for some years, we have restricted the dataset to the 225 pairs of station-years, out of a total of 246 (91.4%), in which we had at least 130 days of sampling (see Appendix B, Figure B.3 for data completeness information). All the analyses

conducted in the rest of the paper will describe results from this filtered dataset. On Figure 2.6, however, we show clusters detected every year when their nearest stations were operational, regardless of the number of days, to highlight the full temporal scope of the dataset.

We used the peak activity in each group of tracks to measure changes in the distribution of roost sizes throughout the season. We tested the null hypothesis of no change in mean roost size between the months of July, August, September and October with the Kruskal-Wallis test (implemented in base R version 4.1.3). Since the Kruskal-Wallis test suggested that the means were in fact different (p-value < 0.001), we performed pairwise comparisons with the Wilcoxon Rank Sum test using the Holm method to adjust the p-values for multiple comparisons (implemented in the base R version 4.1.3).

2.3 Results

2.3.1 Detections, tracks, and groups of tracks

Across a total of 940,641 radar scans from 2000 to 2020, we obtained 67,273 detections, or snapshots of departure events, which the machine learning pipeline grouped into 16,222 tracks representing sequences of snapshots of the same dispersal. Of those tracks, 85.32% (13,841) were considered clear, 0.62% (100) had at least 20% of anomalous propagation inside the bounding boxes, 10.27% (1,667) were contaminated by weather or other unknown sources of noise, and 3.78% (614) were bad-tracks. After correcting for radar range overlaps, we had a total of 14,118 groups of tracks, each one corresponding to a unique dispersal event. Each track group had an average of 4.76 detections (SD = 3.45). The first detection of each group of tracks was then used as input to the mean shift clustering algorithm, which yielded 104 clusters with silhouette coefficient of 0.703 (represented in Figure 2.3). After removal of tracks contaminated by noise or poorly captured by the machine learning pipeline, we had 12,248 groups of clear tracks, with an average of 4.40 detections (SD = 3.12).

Every year in which we had more than 130 sample days per station, we extracted the day when the maximum number of martins and swallows was detected in the area covered by the

NEXRAD network in the Great Lakes region. This number was highest in 2007 (on August 17th), when 536,836 birds across 19 clusters were detected by the radars. In contrast, the peak number of birds was lowest in 2013, with a total 143,625 birds spread across 13 clusters on August 13th. In Figure 2.4, we show, in each month, the frequency in days of each bird count bin (across all 21 years). The largest single roost in our series occurred on September 14th, 2005, in Fond du Lac, Wisconsin, when we estimated a departure of 187,736 birds.

2.3.2 Roost size distribution

The data suggests that small roost sizes are more common than larger ones since the daily size distribution throughout the years is heavy tailed (see Figure 2.5a). Our results indicate that, in August, when the overall population is higher, there is an increase in the frequency of smaller-sized roosts paired with the formation of roosts of size larger than 100,000 birds (Figure 2.5c). Pairwise comparisons showed statistically significant (p-value < 0.001) differences in the mean group size of all pairs of months except for September and October (p-value = 0.11). These tests reveal an interesting asymmetry between the beginning and the end of the season, with the average group size in July (mean 5,388, SD = 7,998) being smaller than the average observed in September (mean 9,011, SD = 14,123).

During July, September and October, more than 60% of the population can be found in smaller roosts, ranging from a few individuals up to 25,000 birds (Figure 2.5b). This pattern changes in August, when roosts of larger size start to gather larger percentages of the population. Even though roosts of more than 100,000 individuals occur in every year of our 21-year series, they do not gather the majority of the Hirundinidae population in the Great Lakes region, since the proportion of the population aggregated in smaller roosts is larger.

2.3.3 Persistence and roost size analysis

Of 104 clusters, which we interpret as regions of frequent use by swallows and martins, 98 of them were detected in more than one year (94.2%), 11 of which occurred every year from 2000 to 2021 (10.6%, see Figure 2.6). We found three clusters that were active on all years when

the radar stations closer to them were active: two of them occurred in east Lake Ontario for 17 years (near KTYX), and one in northern Minnesota (close to KDLH) for 20 years (see Figure 2.6). We used a simple linear regression to test if persistence was a significant predictor of cluster size, measured as the peak number of birds per cluster per year. To account for years when the entire season was unavailable on the archive, we used the number of persistent years over the total number of fully sampled (more than 130 days) years as a predictor variable. The fitted model was $\log y = 1.00x + 3.33$, and it was statistically significant with p-value < 0.001, and R-squared = 0.441 (see Figure 2.7), indicating that areas that were consistently used throughout the years hosted larger roosts.

2.4 Discussion

Weather surveillance radars have proven to be an invaluable source of data about bird migration ecology. Particularly in the case of some species of the Hirundinidae family, due to their specific post-breeding roost formation behavior and timing, weather radars offer a rare opportunity to monitor and inventory their populations with macroscale coverage (Kelly & Horton, 2016). The main challenge currently faced in the effort to reach such broad spatial and temporal scales is the detection of roosts, which is further complicated by the diverse sources of biological and non-biological scattering that are typically captured by the radars. By integrating machine learning and human supervision, we have addressed both challenges, extending the temporal range from previous studies like Bridge et al. (2015) and Kelly et al. (2012), but more importantly extracting critical indices of the magnitude of each communal roost throughout the years.

An important caveat of working with weather radar data is their lack of ability to sample lower heights above ground level, especially at greater distances to the radar station. When we can assume that the observed biomass density does not change with radar range, the systematic under-estimation of density at greater distances to the radar station can be corrected by using vertical profiles of reflectivity (Buler & Diehl, 2009). In the case of communal roosts, however, we can not assume that biomass density is constant at varying ranges, and thus techniques to

adjust density estimates in this scenario based on radar data alone have not yet been developed. We remark, however, that in this study we found little correlation between distance to the radar station and roost size or number of birds (see Appendix B, Figure B.4), indicating that this issue likely has not influenced our conclusions.

We decided to use Purple Martins as our benchmark Hirundinidae because they are the largest species in this family, and thus produce conservative estimates of the number of birds per roost. Conservative abundance estimates can often be enough to diagnose population declines, monitor year-to-year abundance variation, and inform conservation decision-making (for an example, see Hunter et al., 2010), yet they limit our ability to understand the full complexity of the Hirundinidae communal roosting behavior. Further studies pairing radar roost detections with on-the-ground species identification would allow us to characterize species-specific radar signatures in terms of their dispersal time, flight speed, and polarimetric measurements (Stepanian et al., 2016). Moreover, verified data could eventually be used as training data for a new model layer that assigns a species to each dispersal event.

Our results show 11 subregions (clusters) in the Great Lakes which were used year-after-year for the past 21 years, demonstrating how consistent the selection of pre-migratory roosting habitat is for these species (Allen & Nice, 1952; Bent, 1942; Bridge et al., 2015; Russell et al., 1998). This finding, together with the discovery that more persistent clusters were also larger in size, suggests that these regions might successfully support this many birds, and prompts further questions about the role swallows and martins have in those ecosystems. A study conducted by Powell et al. (1991) with much smaller avian aggregations demonstrated, for example, how seabirds in Florida Bay have an important role in translocating and concentrating nutrients in the regions close to their roost/colony sites, ultimately leading to changes in the adjacent seagrass communities.

We further exemplify the potential impact of a Purple Martin roost on the Great Lakes trophic network by using the approach proposed by Kelly et al. (2013) to calculate the number of insects consumed on the day of peak activity in 2007 (August 17th). Assuming an adult Purple Martin

needs 137 kJ per day to maintain regular metabolic functions and that each gram of an insect's dry mass provides 23 kJ, we can derive the total dry mass needed to fulfill the energetic requirements of a single adult. With this, we can employ the relationship between insect length and dry mass proposed by Sage (1982), and assume an average prey will have 20 mm, to determine how many insects would be needed to satisfy that daily dry mass requirement. In doing so, we estimate that on August 17th, 2007, Purple Martins may have consumed more than 33 million insects. Examining summed daily martin activity, an average season may see as many as 338 million insects consumed throughout the Great Lakes region during the roosting period. While these simple calculations make many assumptions, they might serve as basis for deeper investigations into the trophic role played by aerial insectivores across broad regions, and what the impacts of this guild's recent declines could be.

A secondary finding of our approach is related to how birds of the Hirundinidae family occupy the airspace. Our data suggests that, during their dawn departure, they might reach up to 4.5 km above ground level, which is higher than the maximum heights found by Helms et al. (2016) but similar to results from Dreelin et al. (2018) for swallows and martins during their breeding season. Combining altitude loggers and radar studies to explore demographic and seasonal species-specific patterns of flight height might provide valuable insight into how these aerial insectivores use different levels of the aerosphere during their annual cycle. Additionally, because the machine learning pipeline detects and tracks each roost's dispersal, it could be possible to rebuild a three-dimensional model of the departure that improves our understanding of how the birds coordinate their collective take off, and might eventually be used to correct the range bias mentioned above.

Empirical studies of animal aggregations in fish, ungulates, and House Sparrows (*Passer domesticus*) have found group size distributions of all these species to be typically heavy-tailed, with smaller groups being significantly more common than larger ones (for examples, see Bonabeau et al., 1999; Gerard et al., 2002; Griesser et al., 2011). Theoretical studies have demonstrated that these distributions typically arise from the combined effects of animal movement and group

splitting and amalgamation (also called fission and fusion) processes (Niwa, 2004; Okubo, 1986). Though these models accurately describe animals in movement, their translation to the communal roosting scenario is not straightforward, since roosts are not a direct result of movement dynamics, but rather of central place foragers daily gathering in a central place from which they depart the next day to forage (Lalla, 2022). Nonetheless, we have shown that the same pattern of group size distribution occurs in our study system. This result prompts further investigation into models that can describe the mechanisms responsible for generating the specific dynamic of roost formation behavior.

As expected, we found that the daily distribution of roost sizes changes throughout the season, yet we have, for the first time, quantified the progression of these within-season changes. Further studies of this pattern relating it to environmental variables and investigating its change throughout the years may reveal important factors influencing roost formation behavior and driving the roost size distribution. Since most of the Hirundinidae population in the Great Lakes seem to gather in groups with smaller sizes, it would be interesting to investigate why larger roosts are more persistent throughout the years, and what is the relationship between group size, roost persistence, and fitness.

2.5 Figures

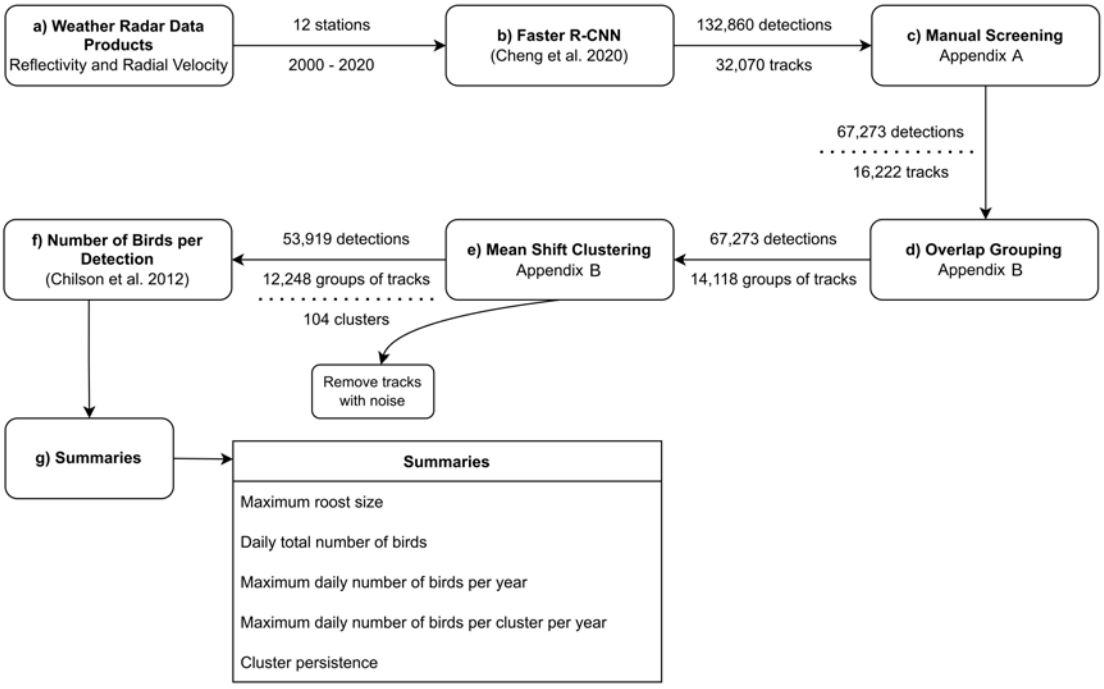


Figure 2.1: Diagram representing the data processing steps, beginning with the radar data download from Amazon Web Services and ending at the extraction of biological summaries from the bird counts on each roost. At each step, we report the number of detections (single frames found in one scan), tracks (sequences of detections in subsequent scans), and groups of tracks (groups of overlapping tracks that correspond to a single roost). Each group of tracks corresponds to one roost's morning dispersal event.

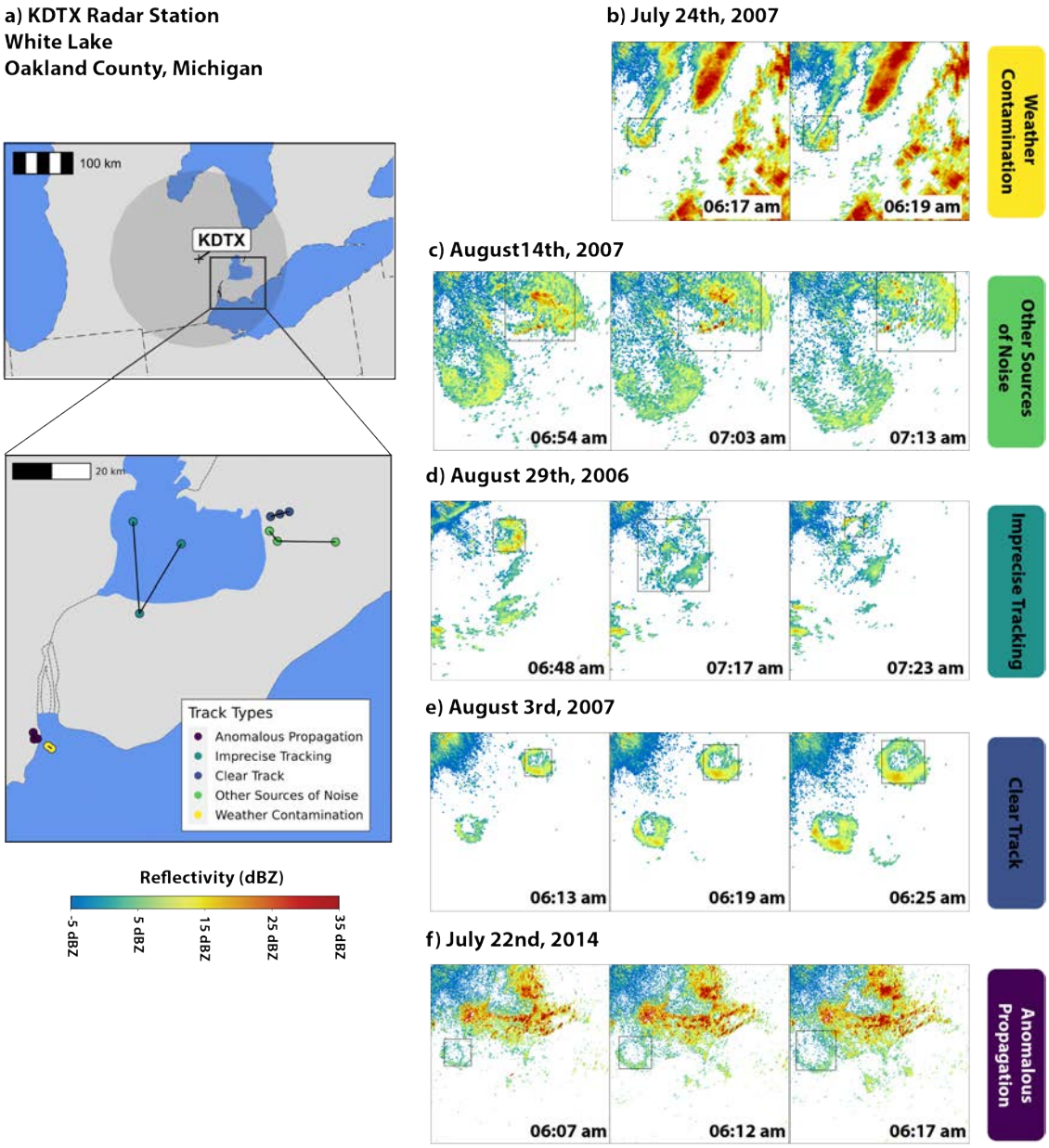


Figure 2.2: (a) A map of the 150 km buffer (shaded gray) around the KDTX station, located in Oakland County, Michigan, and an inset representing the locations of each detection shown on the right side of the panel (b) a weather-contaminated track; (c) a track containing sources of noise other than weather and anomalous propagation (in this case, ground clutter); (d) a subset of detections from a track where the machine learning pipeline did not build the bounding boxes correctly; (e) a clear track that captures the roost without noise; (f) a track contaminated with noise from anomalous propagation in the upper right corner.

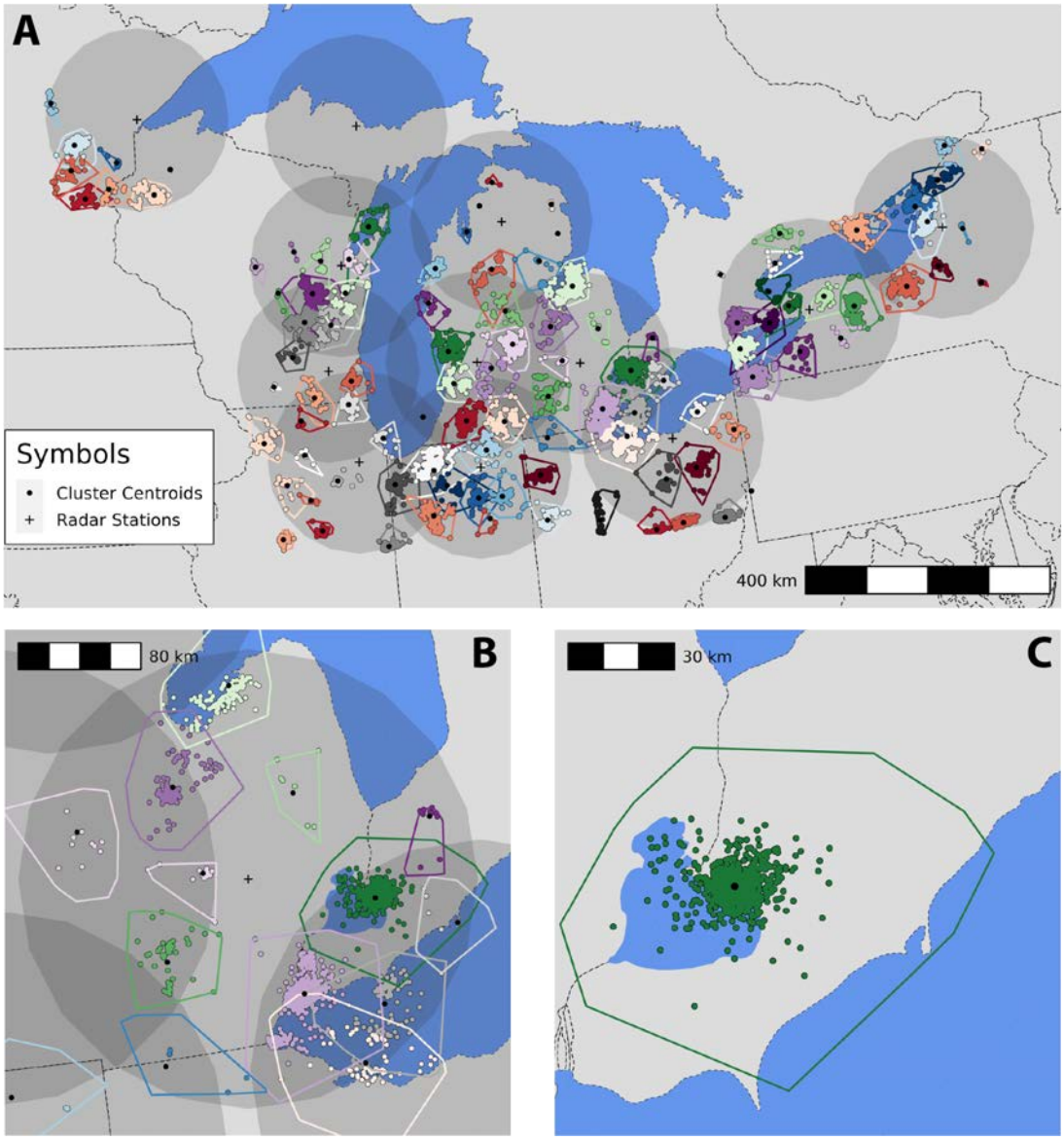


Figure 2.3: (a) Map of the study region, with circular radar ranges in shaded gray (150 km radius). Clusters and respective convex hulls are represented in the same color. (b) Inset showing clusters in the extent of a single station (KDTX) located in White Lake, Michigan. (c) Inset with a single cluster located on the Northeast of Lake Saint Clair. This region saw martins and swallows across all 21 years of our series.

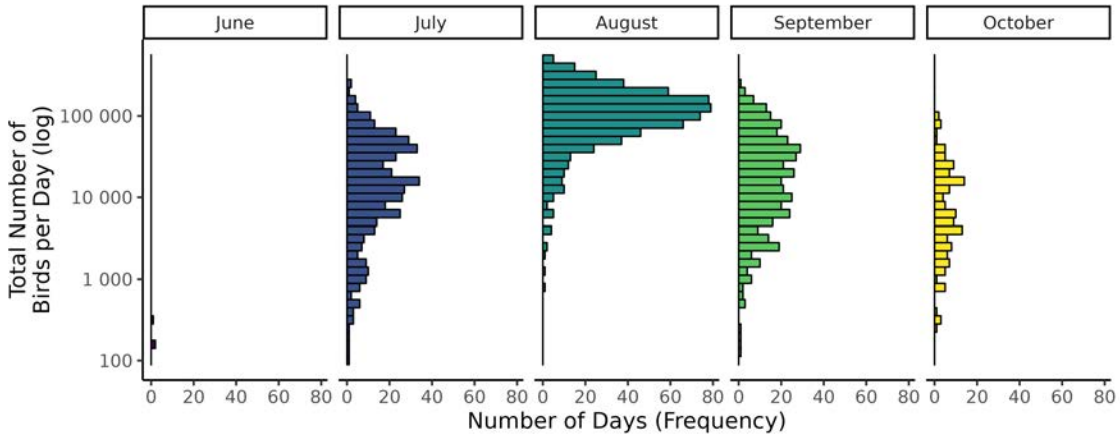


Figure 2.4: Histograms of the daily sum of the number of birds throughout the Great Lakes study region across 21 years, from 2000 to 2020, grouped by month. Note how the total count increases in August.

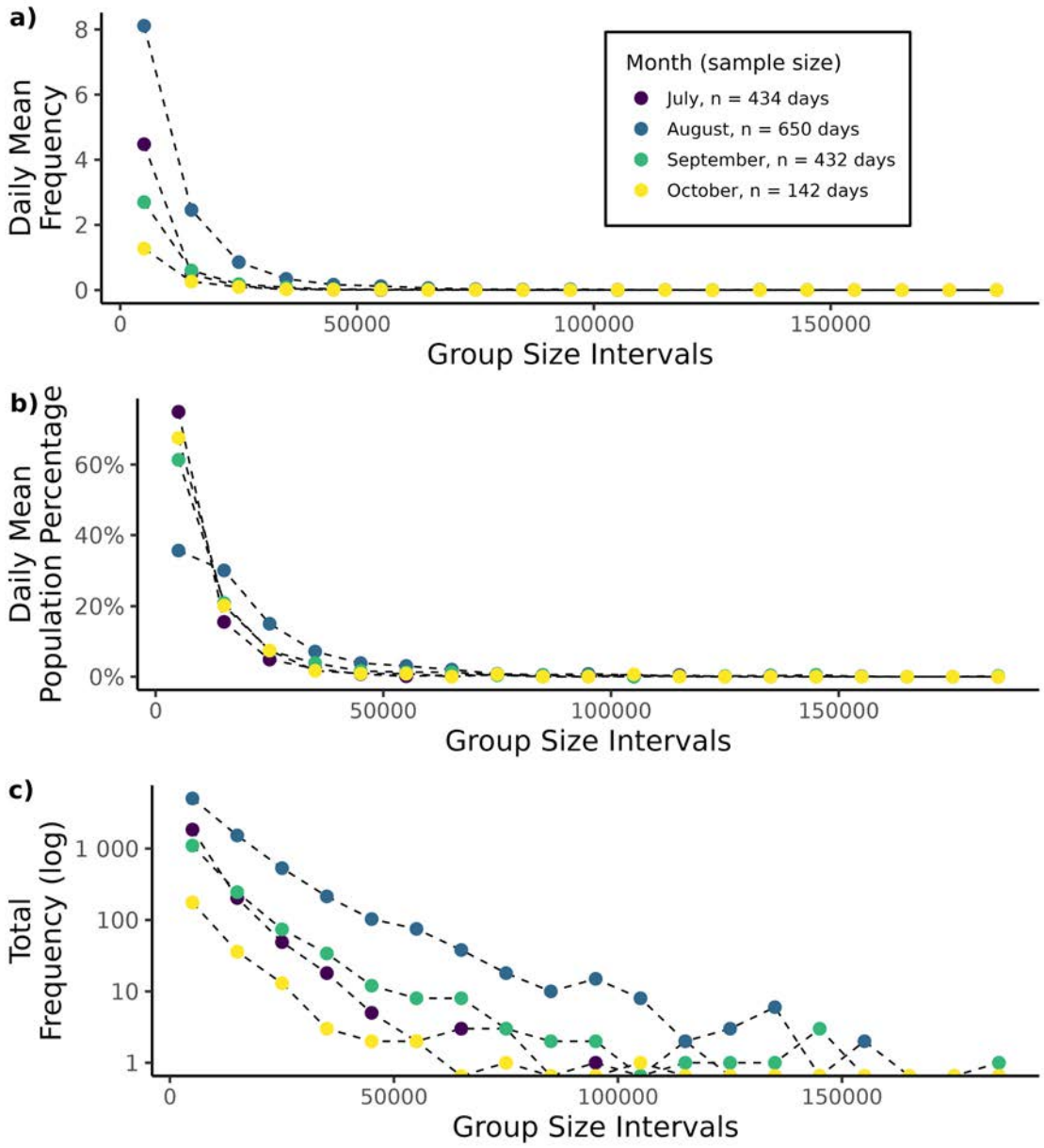


Figure 2.5: (a) Daily mean frequency of each roost size bin of 10,000 individuals, colored by month; (b) Daily average percentage of the population found in each group size bin of 10,000 individuals, colored by month; (c) Total frequency of each roost size across the 21 years of our series (from 2000 to 2021).

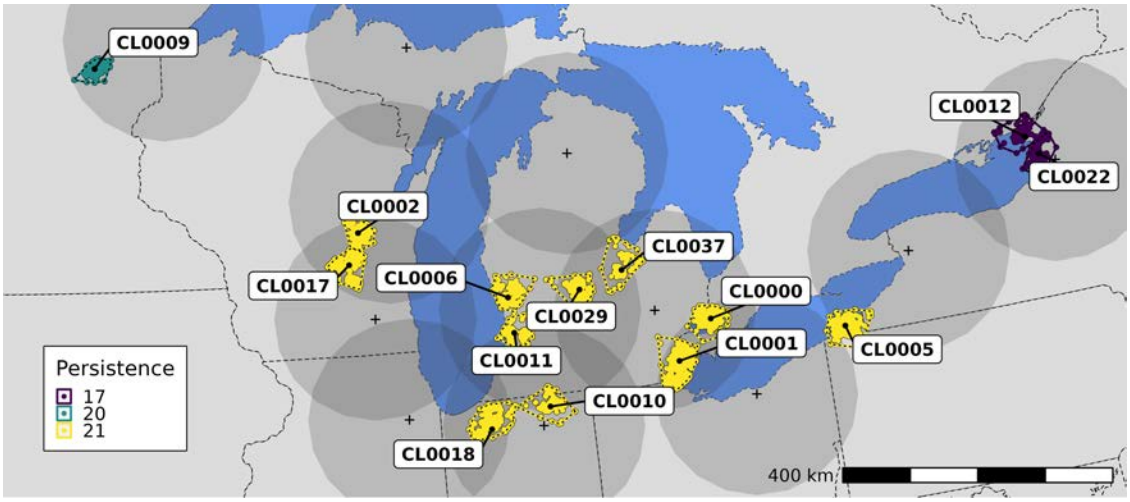


Figure 2.6: Throughout the Great Lakes region, the eleven roost clusters represented above were detected every year from 2000 through 2020

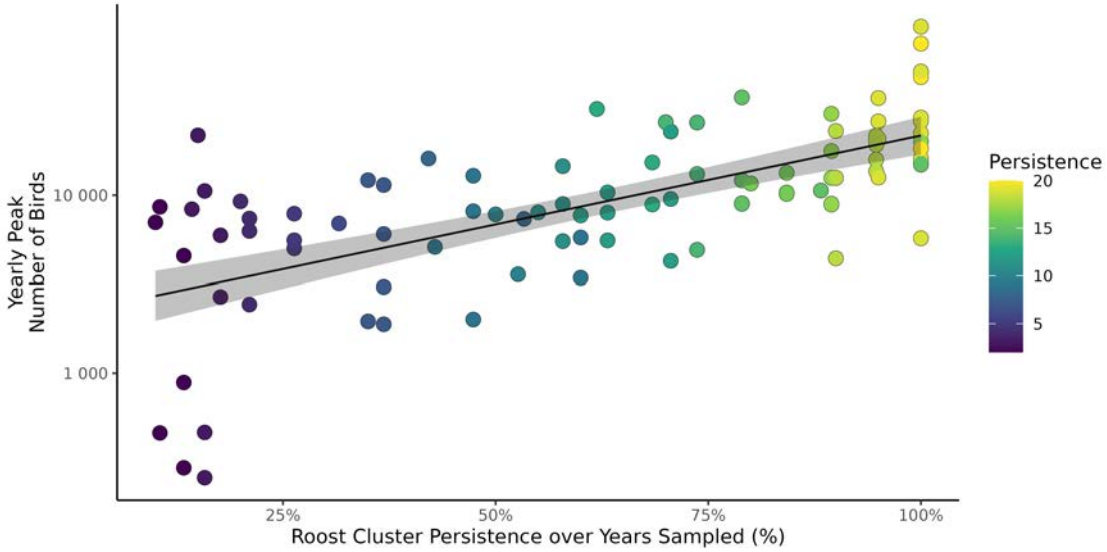


Figure 2.7: Relationship between yearly peak number of birds and roost persistence. The fitted model was: $log(y) = 0.0474x + 3.373$, and it was statistically significant with p-value < 0.0001, $F(1,96) = 86.38$, and $R^2 = 0.4736$.

Bibliography

Allen, R. W., & Nice, M. M. (1952). A Study of the Breeding Biology of the Purple Martin (Progne subis). *American Midland Naturalist*, *47*(3), 606. https://doi.org/10.2307/2422034

Bent, A. C. (1942). *Life Histories of the North American Flycatchers, Larks, Swallows and their Allies*. Bulletin of the United States National Museum. https://doi.org/10.5479/si.03629236.179.i

Bonabeau, E., Dagorn, D., & Fréon, P. (1999). Scaling in animal group-size distributions. *Proceedings of the National Academy of Sciences*, *96*(8), 4472–4477. https://doi.org/10.1073/pnas.96.8.4472

Bridge, E. S., Pletschet, S. M., Fagin, T., Chilson, P. B., Horton, K. G., Broadfoot, K. R., & Kelly, J. F. (2015). Persistence and habitat associations of Purple Martin roosts quantified via weather surveillance radar. *Landscape Ecology*, *31*(1), 43–53. https://doi.org/10.1007/s10980-015-0279-0

Buler, J. J., & Diehl, R. H. (2009). Quantifying Bird Density During Migratory Stopover Using Weather Surveillance Radar. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(8), 2741–2751. https://doi.org/10.1109/tgrs.2009.2014463

Cheng, Z., Gabriel, S., Bhambhani, P., Sheldon, D., Maji, S., Laughlin, A., & Winkler, D. (2020). Detecting and Tracking Communal Bird Roosts in Weather Radar Data. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(01), 378–385. https://doi.org/10.1609/aaai.v34i01.5373

Chilson, P. B., Frick, W. F., Stepanian, P. M., Shipley, J. R., Kunz, T. H., & Kelly, J. F. (2012). Estimating animal densities in the aerosphere using weather radar: ToZor not toZ? *Ecosphere*, *3*(8), art72. https://doi.org/10.1890/es12-00027.1

Cox, A. R., Robertson, R. J., Lendvai, Á. Z., Everitt, K., & Bonier, F. (2019). Rainy springs linked to poor nestling growth in a declining avian aerial insectivore (*Tachycineta bicolor*). *Pro-*

ceedings of the Royal Society B: Biological Sciences, *286*(1898), 20190018. https://doi.org/10.1098/rspb.2019.0018

Dixon, M., & Lee, W.-C. (2021, January). CfRadial Data File Format. Retrieved February 22, 2022, from https://github.com/NCAR/CfRadial/blob/master/docs/CfRadialDoc.v1.5.20211201.pdf

Dokter, A. M., Desmet, P., Spaaks, J. H., Hoey, S. v., Veen, L., Verlinden, L., Nilsson, C., Haase, G., Leijnse, H., Farnsworth, A., Bouten, W., & Shamoun-Baranes, J. (2018). bioRad: biological analysis and visualization of weather radar data. *Ecography*, *42*(5), 852–860. https://doi.org/10.1111/ecog.04028

Dokter, A. M., Liechti, F., Stark, H., Delobbe, L., Tabary, P., & Holleman, I. (2010). Bird migration flight altitudes studied by a network of operational weather radars. *Journal of The Royal Society Interface*, *8*(54), 30–43. https://doi.org/10.1098/rsif.2010.0116

Dreelin, R. A., Shipley, J. R., & Winkler, D. W. (2018). Flight Behavior of Individual Aerial Insectivores Revealed by Novel Altitudinal Dataloggers. *Frontiers in Ecology and Evolution*, *6*. Retrieved June 23, 2022, from https://www.frontiersin.org/article/10.3389/fevo.2018.00182

Dunning, J. (2008). CRC Handbook of Avian body masses. *CRC handbook of avian body masses, 2nd ed*, *384*. https://doi.org/10.1201/9781420064452

Erwin, M. (1982). Observer variability in estimating numbers. *Journal of Field Ornithology*, *53*(2), 159–167. https://sora.unm.edu/sites/default/files/journals/jfo/v053n02/p0159-p0167.pdf

Frederick, P. C., Hylton, B., Heath, J. A., & Ruane, M. (2003). Accuracy and variation in estimates of large numbers of birds by individual observers using an aerial survey simulator. *Journal of Field Ornithology*, *74*(3), 281–287. Retrieved May 31, 2022, from https://www.jstor.org/stable/4131030

Gerard, J.-F., Bideau, E., Maublanc, M.-L., Loisel, P., & Marchal, C. (2002). Herd Size in Large Herbivores: Encoded in the Individual or Emergent? *Biological Bulletin*, *202*(3), 275–282. https://doi.org/10.2307/1543479

Griesser, M., Ma, Q., Webber, S., Bowgen, K., & Sumpter, D. J. T. (2011). Understanding Animal Group-Size Distributions. *Plos One*, *6*(8), e23438. https://doi.org/10.1371/journal.pone.0023438

Hallmann, C. A., Foppen, R. P. B., Turnhout, C. A. M. v., Kroon, H. d., & Jongejans, E. (2014). Declines in insectivorous birds are associated with high neonicotinoid concentrations. *Nature*, *511*(7509), 341–343. https://doi.org/10.1038/nature13531

Helms, J. A., Godfrey, A. P., Ames, T., & Bridge, E. S. (2016). Predator foraging altitudes reveal the structure of aerial insect communities. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep28670

Hoen, B., Diffendorfer, J. E., Rand, J., Kramer, L. A., Garrity, C. P., Roper, A. D., & Hunt, H. (2018). United States Wind Turbine Database. https://doi.org/10.5066/f7tx3dn0

Holmes, R. T. (2007). Understanding population change in migratory songbirds: long-term and experimental studies of Neotropical migrants in breeding and wintering areas. *Ibis*, *149*(s2), 2–13. https://doi.org/https://doi.org/10.1111/j.1474-919X.2007.00685.x

Horton, K. G., Doren, B. M. V., Sorte, F. A. L., Cohen, E. B., Clipp, H. L., Buler, J. J., Fink, D., Kelly, J. F., & Farnsworth, A. (2019). Holding steady: Little change in intensity or timing of bird migration over the Gulf of Mexico. *Global Change Biology*, *25*(3), 1106–1118. https://doi.org/10.1111/gcb.14540

Hunter, C. M., Caswell, H., Runge, M. C., Regehr, E. V., Amstrup, S. C., & Stirling, I. (2010). Climate change threatens polar bear populations: a stochastic demographic analysis. *Ecology*, *91*(10), 2883–2897. https://doi.org/10.1890/09-1641.1

Imlay, T. L., Mills Flemming, J., Saldanha, S., Wheelwright, N. T., & Leonard, M. L. (2018). Breeding phenology and performance for four swallows over 57 years: relationships with temperature and precipitation. *Ecosphere*, *9*(4), e02166. https://doi.org/10.1002/ecs2.2166

Kardynal, K. J., Jardine, T. D., Génier, C. S. V., Bumelis, K. H., Mitchell, G. W., Evans, M., & Hobson, K. A. (2020). Mercury exposure to swallows breeding in Canada inferred from feathers grown on breeding and non-breeding grounds. *Ecotoxicology*, *29*(7), 876–891. https://doi.org/10.1007/s10646-020-02249-6

Kelly, J. F., Bridge, E. S., Frick, W. F., & Chilson, P. B. (2013). Ecological Energetics of an Abundant Aerial Insectivore, the Purple Martin (R. M. Brigham, Ed.). *PLoS ONE*, *8*(9), e76616. https://doi.org/10.1371/journal.pone.0076616

Kelly, J. F., & Horton, K. G. (2016). Toward a predictive macrosystems framework for migration ecology. *Global Ecology and Biogeography*, *25*(10), 1159–1165. https://doi.org/10.1111/geb.12473

Kelly, J. F., Shipley, J. R., Chilson, P. B., Howard, K. W., Frick, W. F., & Kunz, T. H. (2012). Quantifying animal phenology in the aerosphere at a continental scale using NEXRAD weather radars. *Ecosphere*, *3*(2), art16. https://doi.org/10.1890/es11-00257.1

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Lalla, K. (2022). *Purple martin (Progne subis) movement ecology during three stages of the annual cycle* [Master's thesis, McGill University].

Laughlin, A. J., Sheldon, D. R., Winkler, D. W., & Taylor, C. M. (2016). Quantifying non-breeding season occupancy patterns and the timing and drivers of autumn migration for a migratory songbird using Doppler radar. *Ecography*, *39*(10), 1017–1024. https://doi.org/10.1111/ecog.01988

Marra, P. P., Cohen, E. B., Loss, S. R., Rutter, J. E., & Tonra, C. M. (2015). A call for full annual cycle research in animal ecology. *Biology Letters*, *11*(8), 20150552. https://doi.org/10.1098/rsbl.2015.0552

Marra, P. P., & Holmes, R. T. (2001). Consequences of Dominance-Mediated Habitat Segregation in American Redstarts During the Nonbreeding Season. *The Auk*, *118*(1), 92–104. https://doi.org/10.1093/auk/118.1.92

Niwa, H.-S. (2004). Space-irrelevant scaling law for fish school sizes. *Journal of Theoretical Biology*, *228*(3), 347–357. https://doi.org/10.1016/j.jtbi.2004.01.011

Okubo, A. (1986). Dynamical aspects of animal grouping: Swarms, schools, flocks, and herds. *Advances in Biophysics, 22*, 1–94. https://doi.org/10.1016/0065-227x(86)90003-1

Powell, G. V. N., Fourqurean, J. W., Kenworthy, W. J., & Zieman, J. C. (1991). Bird colonies cause seagrass enrichment in a subtropical estuary: Observational and experimental evidence. *Estuarine, Coastal and Shelf Science*, *32*(6), 567–579. https://doi.org/10.1016/0272-7714(91)90075-m

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 28). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf

Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., Stanton, J. C., Panjabi, A., Helft, L., Parr, M., & Marra, P. P. (2019). Decline of the North American avifauna. *Science*, *366*(6461), 120–124. https://doi.org/10.1126/science.aaw1313

Russell, K. R., Mizrahi, D. S., & Gauthreaux, S. A. (1998). Large-Scale Mapping of Purple Martin Pre-Migratory Roosts Using WSR-88D Weather Surveillance Radar. *Journal of Field Ornithology*, *69*(2), 316–325. http://www.jstor.org/stable/4514321

Sage, R. D. (1982). Wet and Dry-weight Estimates of Insects and Spiders Based on Length. *The American Midland Naturalist*, *108*(2), 407–411. https://doi.org/10.2307/2425505

Sherry, T. W., & Holmes, R. T. (1996). Winter Habitat Quality, Population Limitation, and Conservation of Neotropical-Nearctic Migrant Birds. *Ecology*, *77*(1), 36–48. Retrieved May 23, 2022, from https://esajournals-onlinelibrary-wiley-com.ezproxy2.library.colostate.edu/doi/abs/10.2307/2265652

Smith, A. C., Hudson, M.-A. R., Downes, C. M., & Francis, C. M. (2015). Change Points in the Population Trends of Aerial-Insectivorous Birds in North America: Synchronized in Time

across Species and Regions (R. M. Brigham, Ed.). *Plos One*, *10*(7), e0130768. https://doi.org/10.1371/journal.pone.0130768

Spiller, K. J., & Dettmers, R. (2019). Evidence for multiple drivers of aerial insectivore declines in North America. *The Condor*, *121*(2), duz010. https://doi.org/10.1093/condor/duz010

Stepanian, P. M., Horton, K. G., Melnikov, V. M., Zrnić, D. S., & Gauthreaux, S. A. (2016). Dual-polarization radar products for biological applications. *Ecosphere*, *7*(11). https://doi.org/10.1002/ecs2.1539

Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., & Kelling, S. (2009). eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, *142*(10), 2282–2292. https://doi.org/10.1016/j.biocon.2009.05.006

Supp, S. R., Bohrer, G., Fieberg, J., & La Sorte, F. A. (2021). Estimating the movements of terrestrial animal populations using broad-scale occurrence data. *Movement Ecology*, *9*(1), 60. https://doi.org/10.1186/s40462-021-00294-2

Torres, S. M. (2007). Initial implementation of super-resolution data on the NEXRAD network. Retrieved March 20, 2022, from https://ams.confex.com/ams/87ANNUAL/techprogram/paper_116240.htm

Warkentin, I. G., & Hernández, D. (1996). The conservation implications of site fidelity: A case study involving nearctic-neotropical migrant songbirds wintering in a Costa Rican mangrove. *Biological Conservation*, *77*(2-3), 143–150. https://doi.org/10.1016/0006-3207(95)00146-8

Webster, M. S., Marra, P. P., Haig, S. M., Bensch, S., & Holmes, R. T. (2002). Links between worlds: unraveling migratory connectivity. *Trends in Ecology & Evolution*, *17*(2), 76–83. https://doi.org/https://doi.org/10.1016/S0169-5347(01)02380-1

Wikelski, M., Kays, R. W., Kasdin, N. J., Thorup, K., Smith, J. A., & Swenson, G. W., Jr. (2007). Going wild: what a global small-animal tracking system could do for experimental biologists. *Journal of Experimental Biology*, *210*(2), 181–186. https://doi.org/10.1242/jeb.02629

Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., & Girshick, R. (2019). Detectron2. https://github.com/facebookresearch/detectron2

Chapter 3

Aggregating three sources of long-term trends of swallows and martins to identify priority conservation areas in the Great Lakes region

3.1 Introduction

Relative abundance estimates and trends are a crucial source of information for conservation decision-making, determining land use policies and allocation of conservation funds (Hudson et al., 2017). Long-term monitoring programs that yield such trends are even more important in the face of a changing climate, when there is an urgent need to understand if and how animals are changing their use of space and time to adapt to new environmental conditions (Magurran et al., 2010; Parmesan et al., 2005; Wolkovich et al., 2014). Additionally, due to their straightforward interpretation, relative abundance trends act as headline indicators, packing what can be a complex biological process into a metric that is easily understood by the public and by policy-makers, gathering support for conservation and promoting discussion about the biodiversity crisis (Gregory et al., 2008; Loh et al., 2005; van Strien et al., 2012).

In recent years, there has been growing concern about observed declines of avian aerial insectivores in North America, with some authors estimating a decrease of 31.8% (95% confidence intervals: 26% and 36.4%) in breeding population abundance of 26 species of this guild between 1970 and 2017 (Nebel et al., 2010; Rosenberg et al., 2019; Smith et al., 2015). Causes for these declines are complex and hard to disentangle, particularly for migratory species, but include habitat loss, impacts from environmental contaminants, and phenological mismatches due to climate change (Cox et al., 2019; Hallmann et al., 2014; Kardynal et al., 2020; Spiller & Dettmers, 2019). In Canada and the United States, two sources of large-scale relative abundance trends

of songbirds are available: the North American Breeding Bird Survey (hereon, BBS) and, more recently, the data products from eBird Status and Trends (hereon, eBird).

The BBS has been conducted annually since 1966 (with an exception in 2020 due to the COVID pandemic) across 2500 to 3000 roadside routes where expert volunteers conduct fifty 3-minute point counts (Droege, 1990). The surveys are coordinated by the United States Geological Survey, the Canadian Wildlife Service of Environment and Climate Change, and the Mexican National Commission for the Knowledge and Use of Biodiversity (the latter since 2008). Initially, there was concern about potential roadside biases associated with habitat types that might be missed because they do not often occur close to roads (an idea proposed, for example, in Keller & Scallan, 1999). Recent studies, however, concluded that within the conterminous United States and in Southern Canada, this bias is likely minimal (Lawler & O'Connor, 2004; Veech et al., 2017). Another significant criticism of the BBS data is that it does not take probability of detection into account, which makes results less comparable between species if their breeding behavior results in variation in detectability (Somershoe et al., 2006). Additionally, due to challenges with counting the number of birds in breeding colonies, there is more inter-year variation in counts of social breeders such as Cliff Swallows (*Petrochelidon pyrrhonota*), Purple Martins (*Progne subis*), Bank Swallows (*Riparia riparia*), and Barn Swallows (*Hirundo rustica*) (Droege, 1990; Johnson et al., 2017).

While BBS has served as the keystone data for assessing relative abundance trends of bird populations throughout North America for many years, it is not the only data source that allows this type of inference. Recently, the Cornell Lab of Ornithology published a dataset of breeding population trends of hundreds of avian species from 2012 to 2022 which can be downloaded and manipulated with the R package `ebirdst` (Strimas-Mackey et al., 2023). One of the main issues of obtaining trends with eBird data is the pervasive confounding between real relative abundance changes and year-to-year variation in the observation process, particularly due to the growing number of checklists received by eBird every year. To address this and other sources

of temporal confounding, Fink et al. (2023) developed a double machine learning model that yields trend estimates while explicitly accounting for spatial dependency between trends.

In addition to these two approaches, we propose a complementary source of trends in the overall seasonal impact of swallow and martin communal roosts. After these aerial insectivores finish breeding in North America, before starting their southward migration, they aggregate in large communal roosts (Russell & Gauthreaux, 1998). Due to the high density and flight altitudes reached during the early morning roost dispersal, roosts are systematically detected by weather surveillance radars, which can be used to estimate roost size, timing, and location (see Belotti et al., 2023; Bridge et al., 2015; Deng et al., 2023; Kelly & Pletschet, 2017; Kelly et al., 2012; Russell, 1996). We focus our study on the Great Lakes region, where roosting occurs mainly from late July to early September, gathering adults and juveniles of five species of Hirundinidae: Purple Martins, Tree Swallows (*Tachycineta bicolor*), Barn Swallows, Bank Swallows, and Cliff Swallows. The roost size estimates obtained from the radars reflect daily relative abundances of Hirundinidae, but because the roost turnover rate is unknown, we cannot compose them into a regional relative abundance across the roosting season (e.g., individuals could be double counted). Instead, we define a metric of overall seasonal impact of swallow and martin roosts, which is the sum of the daily number of birds found between June 1st and October 31st for each season, balancing changes in population size and season duration.

Having three different sources of trends gives rise to the challenge of comparing different estimates and correctly interpreting them in light of their strengths and weaknesses. A commonly used approach treats the BBS estimates as a benchmark against which to test other sources of data, as done by Walker and Taylor (2017), Horns et al. (2018), and Saracco et al. (2008). This approach is justified by the fact that the BBS is the longest running structured survey of breeding species in North America, and both the survey design and statistical models used to analyze these data have been widely discussed within the scientific community following the publication of foundational works, such as Link and Sauer (2002) and Link et al. (2020). However, because of the different sampling biases and process-specific characteristics of each source, it

is unlikely that the average trend estimates will be pointwise similar or highly correlated, which we will illustrate with our results. Additionally, a pointwise match with BBS trends is even less likely in estimates from data collected at different moments of the annual cycle, like radar data or data from the Christmas Bird Count, yet each of these metrics can offer valuable information about the state of bird populations.

Rather than prioritizing one data source to the detriment of the others, we aim to maximize the use of available information across sources by spatially combining trend estimates to identify regions of conservation concern, using the Great Lakes region of the United States as a testbed for our analyses. We will present a method that accomplishes this by leveraging the fact that the models used to analyze the three data sources output realizations of the distribution of estimated trends: i.e., posterior distributions, in the case of the Bayesian models employed for both radar and BBS data, and resampling folds, in the case of eBird. We employ this technique to identify regions that accumulate evidence of important declines across our three data sources and could, thus, be considered priority conservation areas. We will use species-specific estimates from eBird and BBS to complement the results from the radar data, which lacks the same taxonomic resolution. We will discuss the significance of these findings for aerial insectivores and comment on the applicability of this technique to other species, highlighting its advantages in structured conservation decision-making frameworks.

3.2 Materials and Methods

3.2.1 Spatial Sampling

In order to spatially combine and compare the three different sources of trends, we needed to establish a spatial partitioning of the Great Lakes region that matched the sampling requirements of all three sources. Inference on the BBS data is possible in any spatial partition if at least one route per cell exists. The original survey design includes sampling routes in a grid with cells of 1° in longitude and 1° in latitude, equivalent in area to a square with a width of 100 km (Droege, 1990). The eBird trends are released as model predictions within a grid of square

cells with a width of 27 km. Lastly, the radar data can be partitioned into cells with any radius larger than approximately 10 km, to account for possible error in the estimation of roost location (Kelly & Pletschet, 2017). We accommodated these different prediction frames by aggregating eBird cells until each larger cell had the same average area as the original BBS cell in the Great Lakes region (see Appendix C for maps of all the sampling schemes). This procedure resulted in a square grid with cells of approximately 100 km by 100 km. While the area covered by BBS and eBird spans the entire Great Lakes region, radar coverage is limited by the radar range of 150 km (roosts farther away are likely missed due to higher beam heights, see Appendix C, Figure C.1 for a radar coverage map). Our analyses will focus on 65 cells that were sampled by all three sources.

3.2.2 Radar Data Filtering and Processing

In what follows, we describe specific aspects of the data filtering and processing relevant to this study, but refer the reader to Belotti et al. (2023) for additional details. The machine learning pipeline is thoroughly described and evaluated in Cheng et al. (2020).

Swallow and martin roosts typically occur in the summer, after these species have finished breeding and before they start their southwards migration (Kelly & Pletschet, 2017; Russell, 1996). We captured the roosting period by downloading all radar scans available between June 1st and October 31st for the Great Lakes region (12 radar stations) from 2000 to 2022. We restricted our sample to scans made between 30 minutes before local sunrise and 90 minutes afterwards (an average of 16.2 scans per day), when we typically observe roost departures. Radial velocity and reflectivity in these scans were used as input to a machine learning pipeline that leverages the capabilities of a Faster R-CNN model coupled with a Kalman Filter to detect and track roost dispersals in radar data (Cheng et al., 2020). The pipeline outputs collections of bounding boxes (so-called tracks), each collection representing a single dispersal event.

The machine learning pipeline can identify roosts with reasonable accuracy; however, the differentiation between roost dispersals and precipitation still requires quality control (Cheng

et al., 2020). Because of this, we designed a manual screening protocol used by three independent screeners to examine all the scans overlaid with bounding boxes to evaluate the predicted positive tracks, classifying them according to the predominant type of noise, if there was any, or labeling them as clear roosts if they were not contaminated. The screeners also labeled days when precipitation, anomalous propagation, or other sources of radar noise were so intense that we could not see roosts in the scan if they were present. These days were considered unsampled.

After the manual screening, we used the coordinates of the bounding box centroid and the radius of the bounding box to extract radar reflectivity measurements for each detection. Under conditions specified in Chilson et al. (2012), radar reflectivity can be interpreted as the density of scatterers of a given radar cross-section within each radar sampling volume. We make the assumption that all scatterers within the roosts are Purple Martins — which yields conservative estimates of the number of birds, since they are the largest Hirundinidae in the Americas — and use the relationship established by Horton et al. (2019) to infer their radar cross section of 15.62 cm^2 from their average mass of 51 g (Dunning, 2008). By dividing radar reflectivity (measured in linear scale) within each pixel by the target radar cross section and multiplying it by the radar sampling volume, we obtain estimates of the number of birds within each volume of the airspace sample by the radar.

We summarized counts across radar sweeps to obtain estimates of the numbers of birds across the entire vertical structure of each dispersal. To do this, we grouped the sweeps in each detection in bins of 1°, according to their elevation angle. We then took the mean bird count of each bin and summed counts across altitudinal bins. Doing so helps minimize the double counting of birds in airspace regions sampled twice by two consecutive sweeps with less than a 1° angle change between them. The last post-processing step was to handle roost departures found in regions sampled by more than one radar. We accounted for that by grouping overlapping tracks so that each group of tracks corresponds to a single dispersal event.

At this stage in the data processing workflow, we generated estimates of the number of birds per clear dispersal, and we could then aggregate those to get daily estimates of the number of birds within any geographical subregion of our study area. We did so by summing dispersal size estimates per day for each grid cell. This yielded counts per day within a season. We further summarized those by summing the counts across each season, which gave us an estimate of the total population impact (in units of number of birds × day) of swallow and martin roosts on each cell for each year. This metric describes the potential influence roosts could have on the landscapes that support them, and has been used in studies of water quality to calculate nutrient deposition rates of waterfowl flocks (Gremillion & Malone, 1986; Manny et al., 1994).

We extracted two metrics of the quality of our sampling effort per year. The first was the percentage of unobserved days within each season for each cell. This reflected the fact that some scans were missing from the Amazon Web Services archive, where all the radar data collected in the United States is stored, or that weather and/or noise were so intense on a given day that we would not have been able to see the roosts if they were present. The second metric was the percentage of groups of tracks that were classified as noise by the screeners. This metric was used to account for the fact that some dispersals had to be removed from our daily count summaries because they were contaminated with noise. We emphasize that low values of both metrics reflect poor observation quality of a given cell-year, whereas high values reflect better observation quality.

3.2.3 Population Impact Trends

We obtained trends of population impacts by adapting the hierarchical generalized additive model with cell-year random effects established by Knape (2016) for the Swedish Breeding Bird Survey and employed in Smith and Edwards (2020) to analyze BBS data. Our model accounts for spatial autocorrelation in the residuals by using an intrinsic conditional autoregressive (iCAR) structure detailed in Morris et al. (2019) and implemented in Smith et al. (2023) and Bled et al. (2013) for the BBS data.

Our estimates of population impact were either zero or a very large number, with a mean of 93.415 and a standard deviation of 179.646. In order to model such large variation, we divided the impacts for each cell and year by the maximum impact observed on that cell across all years. This rescaling mapped our large numbers to the interval between 0 and 1 with a cell-specific scaling factor, where 1 corresponds to the maximum value observed within a cell across all years of our study. However, our model should be capable of estimating values that exceed the maximum value of each cell, so our response should not be constrained to the unit interval. Because of this, we used a gamma likelihood with an exponential link to connect the nonlinear model expression to our adjusted response $Y_{i,t}$, where $i = 1...$I indexes cells and $t = 1...$T indexes years (scaled so that 2000 is year 1, 2001 is year 2, etc.). The model can then be written as:

$$\begin{aligned} Y_{i,t} &\sim \text{gamma}(a, \frac{a}{\mu_{i,t}}) \\ \log(\mu_{i,t}) &= \psi_i + S_i(t) + \eta \times \text{unobserved}_{i,t} + \delta \times \text{noise}_{i,t} \end{aligned} \tag{3.1}$$

where the rate parameter of the gamma distribution can be written as a function of the distribution's mean ($\mu_{i,t}$) and shape (a).

The parameters ψ_i are random intercepts representing initial population size in the baseline year. The smooth function $S_i(t)$ describes the nonlinear relationship between impact and year within each cell. We attempted to add cell-year effects to describe fluctuations around the mean smooth for each cell, as was done in Link et al. (2020) and Smith and Edwards (2020); however, our data did not allow estimation of both the smooth and the fluctuations. We account for incomplete observation of a given season and for the presence of noise-contaminated tracks not included in the measured total impact by adding the percentage of unobserved days and the percentage of noise-contaminated tracks as covariates in the model. The coefficients for these covariates are η and δ, respectively.

The influence of both effort correction covariates on the scaled total impact is mediated by the exponential function. We designed these covariates so that when they are equal to zero, a season was completely observed, and their effect on the total impact will be equal to

1 ($\exp(0) = 1$). As the covariates move towards 1, our observation of the season becomes increasingly incomplete, requiring increasingly larger corrections, that is, we need to multiply the total impact by a number greater than 1. Because of this, we restrict η and δ to be positive to reflect the fact that lower observation of a season requires a greater correction factor (recalling that $\exp(x) > 1 \Leftrightarrow x > 0$).

We used thin plate regression splines to build a basis of K knots $B_i = \{b_{i,1}(t), ..., b_{i,K}(t)\}$ for each cell of our grid (Wood, 2017). We can then write the smooth function for each cell as a linear combination of the elements of the basis and a series of cell-specific coefficients $\beta_{k,i}$:

$$S_i(t) = \sum_{k=1}^{K} \beta_{k,i} b_{i,k}(t) \tag{3.2}$$

where each $\beta_{k,i}$ is drawn from a normal distribution with standard deviation σ_β and a cell-specific mean α_i. We accounted for spatial autocorrelation by adding an iCAR structure to the α_i parameters describing the mean smooth coefficient per cell. This model structure makes it so that each α_i inherits information from the set N_i of its neighbors, here defined as all the adjacent cells to cell i (using the queen neighborhood rule, where each cell has at most eight neighbors). This is formalized by drawing α_i from the following distribution:

$$\alpha_i \sim \text{normal}\left(\frac{\sum_{n \in N_i} \alpha_n}{N_i}, \frac{\sigma_\alpha}{N_i}\right) \tag{3.3}$$

We fit our model in Stan (Stan Development Team, 2023) via the R programming language (R Core Team, 2018) using the package `rstan` (Stan Development Team, 2023). We estimated I parameters ψ_i, two parameters η and δ, I × K coefficients $\beta_{i,k}$ and I coefficients α_i for the smooth terms, one parameter a determining the shape of the gamma likelihood, and standard deviations σ_β, σ_α, and σ_ψ. We used a truncated Student-t prior on all the standard deviations and on the parameters η and δ, and specified it with three degrees of freedom, location equal to zero, and a scale parameter of 2.5. This choice of prior provides minimal information about possible parameter range to improve convergence, but does not make strong assumptions about

its real value (Gelman et al., 2008). The prior for the shape of the gamma distribution, a, is a flat gamma distribution with rate and scale equal to 0.01, which did not cause issues in our case because the variance was not close to zero (Gelman, 2006). In Table 3.1, we provide a list of model parameters, their interpretation, and the priors assigned to them.

Symbol	Prior	Interpretation
η	zero-truncated Student-t(3, 0, 2.5)	Coefficient of the correction for unobserved days within a season
δ	zero-truncated Student-t(3, 0, 2.5)	Coefficient of the correction for tracks lost to noise contamination within a season
α_i	normal$\left(\frac{\sum_{n \in N_i} \alpha_n}{N_i}, \frac{\sigma_\alpha}{N_i}\right)$	Cell-specific mean smooth coefficient (with iCAR structure)
ψ_i	zero-truncated Student-t(3, 0, 2.5)	Cell-level intercepts
$\beta_{i,k}$	normal$(\alpha_i, \sigma_\alpha)$	Smooth coefficient per cell per knot
σ_α	zero-truncated Student-t(3, 0, 2.5)	Standard deviation of cell-specific smooth coefficients
σ_β	zero-truncated Student-t(3, 0, 2.5)	Standard deviation from mean smooth across cells
σ_ψ	zero-truncated Student-t(3, 0, 2.5)	Standard deviation of the cell-level intercepts
a	gamma(0.01, 0.01)	Shape parameter of the gamma likelihood

Table 3.1: List of parameters with their verbal interpretation and the prior we assigned to them. Cells are indexed by i = 1...I, years are represented by t = 1...T, and knots are indexed by k = 1...K

We fit our model using 6500 Markov chain Monte Carlo (MCMC) iterations and 1000 warm-up iterations, evaluating convergence and parameter estimates in a thinned subset of 2400 samples. We used the fitted model to describe the posterior distributions of the population impact predictions from 2012 to 2022. These distributions were then used to obtain posterior estimates of the geometric mean trend for that period, which are samples of the geometric mean trend and describe the variation of estimates around the mean.

3.2.4 BBS and eBird Trends

We downloaded ensembles of eBird estimates of percent per year change from 2012 to 2022 for our target species — Purple Martin, Tree Swallow, Barn Swallow, Bank Swallow, and Cliff Swallow — using the package `ebirdst` (Strimas-Mackey et al., 2023). These ensembles are samples of trend estimates (n = 100 in this case) obtained from using different subsets of the data to make predictions with fitted double machine learning model, and are used to describe the mean estimate and the credible intervals for each species and cell (Fink et al., 2023). To make these estimates comparable to the radar and BBS estimates, we aggregated the estimates within each ensemble across species and between groups of 27-km cells to obtain a single ensemble of estimates for all our target species within each 100-km cell of our sampling grid. We do so as follows: we start with one ensemble of estimates of percent per year change per species per grid cell of 27 km size. These are aggregated in space using an average weighted by the relative abundance of that species (average from 2012 to 2022 provided in the data) to obtain one ensemble per species per cell within our grid of 100 km. Finally, using summed relative abundances across species as weights, we grouped the ensembles of the five species into a single ensemble of percent per year change per cell.

We obtained relative abundance estimates from BBS data using the generalized additive mixed model with iCAR structure described in Smith et al. (2023). Instead of adopting the standard grid of 1 degree cells typically used for analysis of BBS data, we fit the model with data partitioned within our grid of cells of the same size, but in slightly different locations. The `bbsBayes2` package provides an straightforward workflow to accomplish this using Stan (Edwards et al., 2023; Sauer et al., 2021). We fit the model for the five target species with data from 2000 to 2022, summing the relative abundances across species to get a posterior estimate of the total relative abundance. Lastly, we subset estimates for 2012 and 2022 to calculate the posterior distribution of the geometric mean trend between those years.

3.2.5 Comparison Strategy

Using the estimates of relative abundance or population impact percent per year change from 2012 to 2022 for each data source, we estimated the distribution of the correlation coefficients between all possible pairs of sources (BBS-eBird, BBS-radar, radar-eBird) across cells. To examine whether breeding population trends of a single species were driving observed population impact trends, we calculated distributions of estimated correlation coefficients between radar-derived trends and trends from BBS and eBird for each species. We then used the percent overlap between the empirical distributions of correlations to describe how similar/distinct the estimated coefficients were, following the procedure described in Pastore and Calcagnì (2019).

Additionally, we estimated the empirical cumulative distribution of the geometric mean estimates of change of each data source and obtained the probability of the trend within a given cell being positive or negative. We then used these probabilities to calculate the joint probabilities of every combination of trend direction between the sources, which provides estimates of the probabilities of directional agreements and disagreements between them. When examining directional probabilities for a single data source, we created four tiers corresponding to confidence levels that the trend has a given direction. If the observed probability is lower than 0.5 (the flip of a fair coin), we consider that trend direction to be unlikely ("Very Low" probability tier). We then divide the remaining values from 0.5 to 1 into three equally spaced tiers from "Low" to "High" probabilities of a given trend direction. We follow a similar rationale when examining the probabilities of the three-way comparison. If the observed probability of a combination of directions is lower than the probability of flipping three fair coins and getting that same outcome (1/8), we consider that combination improbable. The remaining values between 1/8 and 1 are divided into three equally spaced tiers from "Low" to "High", representing how much certainty we have that that combination is occurring. With the exception of the maps per species for BBS and eBird, the figures in this paper represent the tiered probabilities, but the reader will find the maps with probabilities in continuous scale for each source in Appendix D.1.

3.3 Results

The BBS estimates per species generally suggest relatively high probabilities of declines in the northern part of the Great Lakes region, particularly for Tree Swallows, Purple Martins, and Cliff Swallows (see Figure 3.1). These three species also face relatively high probabilities of a positive trend to the South of the Great Lakes region. The predicted probabilities for Bank and Barn swallows are less extreme (values closer to 0.5 than to 0 or 1) in both directions, indicating more uncertainty in the estimates for these two species. The eBird data, on the other hand, points at relatively high probabilities of widespread declines across the entire study region for all species but Purple Martins, for which it predicts relatively high probabilities of increases (see Figure 3.2). The aggregated estimates from BBS and eBird across all species are driven by these species-specific patterns. The aggregated estimates from BBS predict a diagonal line crossing southern Lake Michigan and separating a southern region of widespread positive trends from an area of spatially consistent negative trends in the northern Great Lakes region (see Figure 3.3a). The aggregated eBird estimates, however, predict a smooth range-wide decline (see Figure 3.3b). Our results from the radar data showed more spatial heterogeneity in population impact trends, when compared to the breeding population trends (see Figure 3.3c). In general, we observed medium to high probabilities of declines in roosting activity in the eastern Great Lakes, particularly in the region to the east of Lake Ontario and to the west of Lake Erie. Conversely, we found medium to high probabilities of increases in roosting activity in the west/southwestern Great Lakes region.

We estimated the correlations between the geometric mean trends of each pair of sources across cells. Our results show little evidence of correlation (see Figure 3.4), even though our estimates are biased towards predicting positive or negative association due to the significant spatial autocorrelation present within each variable. The average correlation between BBS and radar estimates was -0.01 (SD: 0.12), and between eBird and radar estimates was 0.2 (SD: 0.06). The average correlation between eBird and BBS was slightly higher, with an average of 0.25 (SD: 0.11). We further explored these results by calculating correlations between radar trends and

trends for each species across cells, for both BBS and eBird (see Figure 3.5). On both sources, no correlation estimate per species was significantly different from zero (all credible intervals overlap 0). Additionally, within each source, the posterior distributions of correlations had high overlapping area between species (> 0.6 average similarity index, see Figure 3.5), indicating that there isn't a significant difference between species in their estimated correlation distributions. The same is generally true when comparing BBS trends per species with radar-derived trends. We found, however, that BBS trends for Purple Martins were slightly more correlated to radar-derived trends (average correlation coefficient 0.14 SD: 0.11, versus average coefficients between -0.08 and 0.02 for the other four species) and that their distribution is less similar to other species (average similarity index: 0.5), however results are inconclusive since the correlation estimates are not significant.

Because eBird estimates within the Great Lakes region were consistently negative across the entire study region, any three-way comparison that incorporated the probability of a positive eBird trend resulted in a very low joint probability. We found that all sources likely agree on a negative trend on the eight cells surrounding Lake Ontario and on seven cells to the west of Lake Michigan, where the joint probability of a negative trend in all sources was greater than 0.42 and classified as medium or high (see Figure 3.6a). We found mostly very low or low probabilities of BBS predicting a positive trend when radar predicts a negative trend (Figure 3.6b), whereas there are 21 cells in which it is likely that BBS predicts a negative trend in disagreement with the trend predicted by the radar data, particularly on the four cells to the east of Lake Michigan and on the four northernmost cells to the east (Figure 3.6c). Lastly, even though BBS and radar disagree with eBird, they both agree on a positive trend in the southwest of our study region (south of Lake Michigan, see Figure 3.6d).

3.4 Discussion

Our study suggests that roosting activity is increasing towards the west/southwest Great Lakes region and decreasing in the opposite direction (particularly in Lake Ontario). Increases

in population impact could be a result of two biological processes: increase in roosting season duration and/or increases in daily relative abundances. Conversely, decreases in population impact could be caused by corresponding decreases in relative abundance and/or a shortening of the roosting season. Disentangling these two processes can be challenging. We have some evidence of breeding season advancements for all our study species except for Bank Swallows (for Cliff Swallows in Nebraska, US, see Brown & Brown, 2014; for Tree Swallows across the United States, see Dunn & Winkler, 1999; for Purple Martins across the United States, see Shave et al., 2019; lastly, Imlay et al., 2018, found no evidence of Bank Swallow breeding initiation advancements in Nova Scotia and New Brunswick, in Canada). Additionally, previous studies of roosting phenology in the Great Lakes region found that the second half of the roosting season (50%, 75% and 95% passage dates), has advanced on average between 2.23 and 4.09 dates over two decades (from 2001 to 2020, Deng et al., 2023). It is unclear, however, if those advancements are different enough to cause changes in roosting season length.

Regarldless of the processes causing the observed changes in population impact, trends in this metric could reflect changes in the impact these large aggregations have on the landscape that supports them and/or changes in the ecosystem services they provide. Studies indicate that seabird colonies concentrate nutrients around the islands that support them and increase the biomass and species composition of seagrasses (Powell et al., 1991). The same effects on local food webs could occur in areas around swallow and martin roosts, and our population impact metric could be used as a proxy for the intensity of these effects. In addition, because the presence of waterfowl has been associated with changes in water quality (Manny et al., 1994), and Cliff Swallow colonies have caused seasonal surges of higher concentration of *Escherichia coli* on the underlying creek (Sejkora et al., 2011), the population impact of roosts, which is proportional to how much fecal matter is deposited in an ecosystem, could be a state variable of interest for water quality monitoring schemes.

Our comparison between BBS, eBird, and radar data suggests that Lake Ontario is observing both a decrease in breeding populations and a decrease in roosting activity, indicating that this

region should be prioritized in future conservation efforts. This region could also be the focus of studies aimed at understanding if declines in local breeding populations are directly causing declines in roosting activity in the same region. Data from light-level geolocators suggests that Tree Swallows are using the Great Lakes region both as a breeding site and a stopover location (see Knight et al., 2018). It is possible, nonetheless, that individuals of other species breeding in Canada, for example, might use roosts in the Great Lakes region as stopover or staging sites, which would cause trends in seasonal impact in this region to reflect changes in northern breeding populations.

Regions that observe increases in roosting activity without a corresponding increase in breeding populations — that is, regions with a high probability of eBird (-), BBS (-), and radar (+) — give rise to many questions and hypotheses that can be tested in studies at smaller scales. This pattern could be caused, for example, by increases in relative abundance in northern breeding populations, or it could be a result of local birds staying in their pre-migratory roosts for much longer. The latter hypothesis is a cause of concern since it could suggest that birds might be facing challenges to refuel in time for their fall migration. This type of delay has been observed in Barn Swallows that needed to stay for a longer period in their nonbreeding grounds in order to prepare for spring migration (Saino et al., 2003). Conversely, regions in which we observe decreases in roosting activity paired with increases in breeding populations prompt us to wonder where and why are breeding birds moving to roost. These regions would show a high probability of a negative trend in population impact, combined with a high probability of a positive trend in BBS data, even though the latter does not agree with the negative eBird estimate.

We developed a workflow to combine multiple lines of evidence to provide robust inference on changes in population-level state variables. This technique can be generalized to other species or even include sources of trend estimates of other variables of interest, such as demographic rates obtained by the Monitoring Avian Productivity and Survivorship (MAPS) program (Saracco et al., 2008). This approach can be used with modeling frameworks that output realizations of the estimated trend, such as Bayesian hierarchical models or resampling meth-

ods. These realizations can be used to estimate regional directional probabilities of each trend, which can be spatially combined to obtain joint probabilities of directional agreements and disagreements between sources within each cell on a spatial grid that fits the purpose of the study while accommodating the sampling limitations of all the data sources.

Our proposed metrics for decision-making – the probabilities of agreement and disagreement – maximize the use of available information while describing both the uncertainty associated with each estimate and the uncertainty associated with our lack of knowledge about the system (epistemic uncertainty, Van Der Bles et al., 2019). Thus, for example, when implementing the structured decision-making framework proposed in Fournier et al. (2023), in addition to relying on expert opinion and working group consensus to establish the degree of uncertainty associated with the estimated trend, one could define thresholds on the probabilities of agreement dividing regions on a low to high scale of overall uncertainty levels across sources. This allows a distribution of conservation efforts proportional to how certain we are, what types of errors we are willing to make, and how much evidence we have that those resources are needed in a given region. Lastly, communicating directional probabilities instead of average estimates could effectively describe to the general public the uncertainty associated with the estimated trends.

Trend estimates from all three of our data sources have some degree of confounding between local changes in abundance and changes in annual cycle timing, though the radar data does so more explicitly. As we gather further evidence of advances in the breeding season in North America and across the globe (Horton et al., 2020; Parmesan & Yohe, 2003; Youngflesh et al., 2021), it becomes clearer that the fixed sampling dates of BBS and eBird could gradually be observing a different moment of the breeding period, which could lead to changes in the species detectability (e.g. availability). These changes would produce trends that might not correspond to actual population declines or increases as shown in Massimino et al. (2020). However, because the seasonal impact metric used for the radar data incorporates both daily relative abundances and season duration, it is expected that changes in timing alone will produce trends in

population impact. More research is needed to understand the impact of phenology changes on trend estimates from BBS and eBird data, and we suggest that the latter could be used to examine this issue in future studies. This concern, however, should not diminish the relevance of trend estimates of breeding populations, but it can help us improve our ability to interpret them.

In this study, we proposed a novel technique that can be used to spatially combine a wide range of data sources that monitor the state of a given population into a single metric describing the probability of agreement or disagreement across sources on the direction of the trend. This maximizes the use of available information while also describing the uncertainty associated with our incomplete knowledge of the limitations associated with the observation process and of the drivers and constraints of the real population dynamics. It also allows decisions to be made based on the amount of evidence that supports the conclusions about trend direction. We demonstrate the use of this method by combining eBird, BBS, and radar data to conclude that consensus around the declines observed in Lake Ontario require further investigation.

3.5 Figures

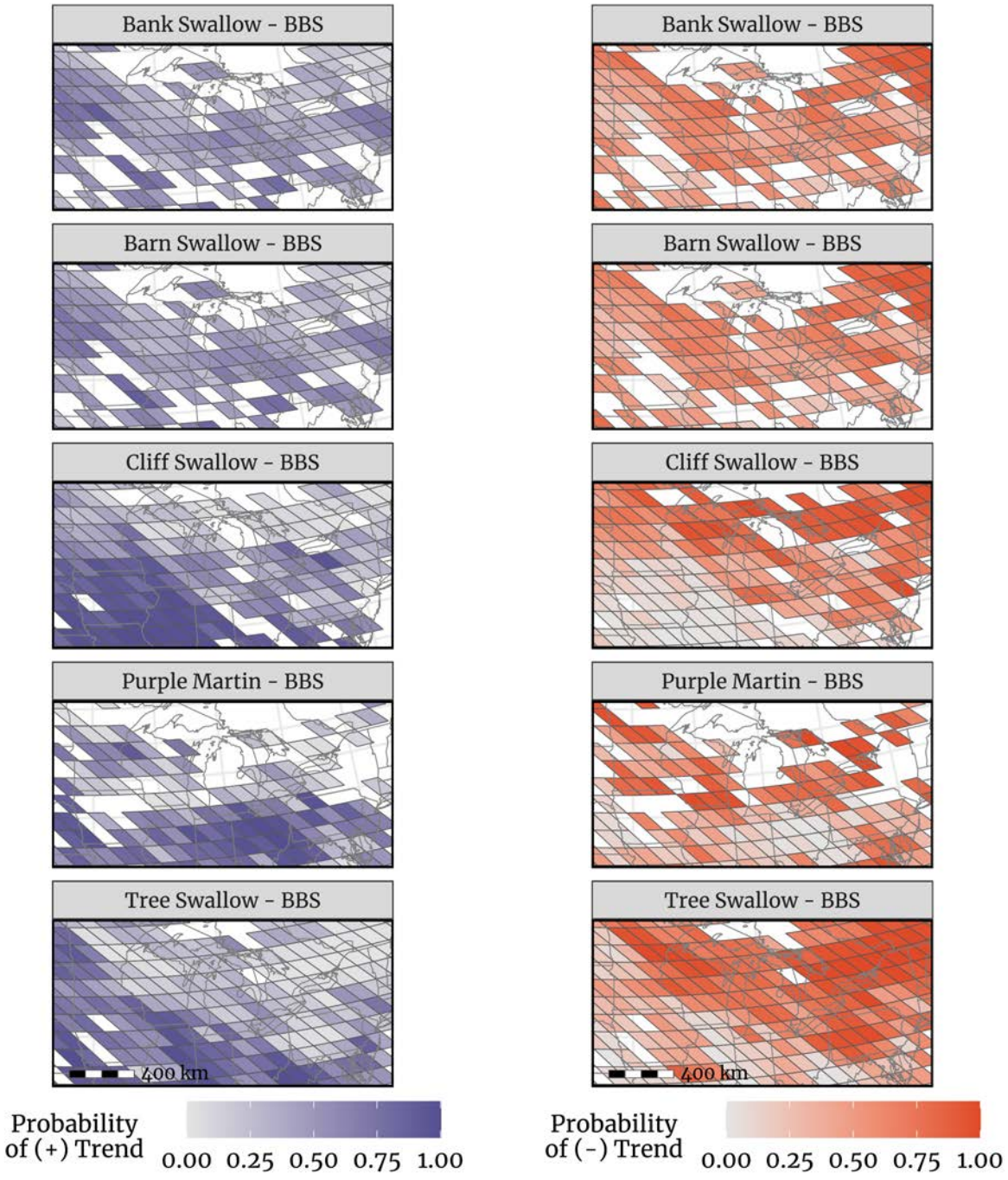


Figure 3.1: Directional probabilities of relative abundance trends for each species, obtained using BBS data. If the probability of a trend direction is lower than the probability of the flip of a fair coin, we consider that trend direction to be very unlikely. Shades of red from light to dark reflect increasing certainty on the direction of the trend in a cell.

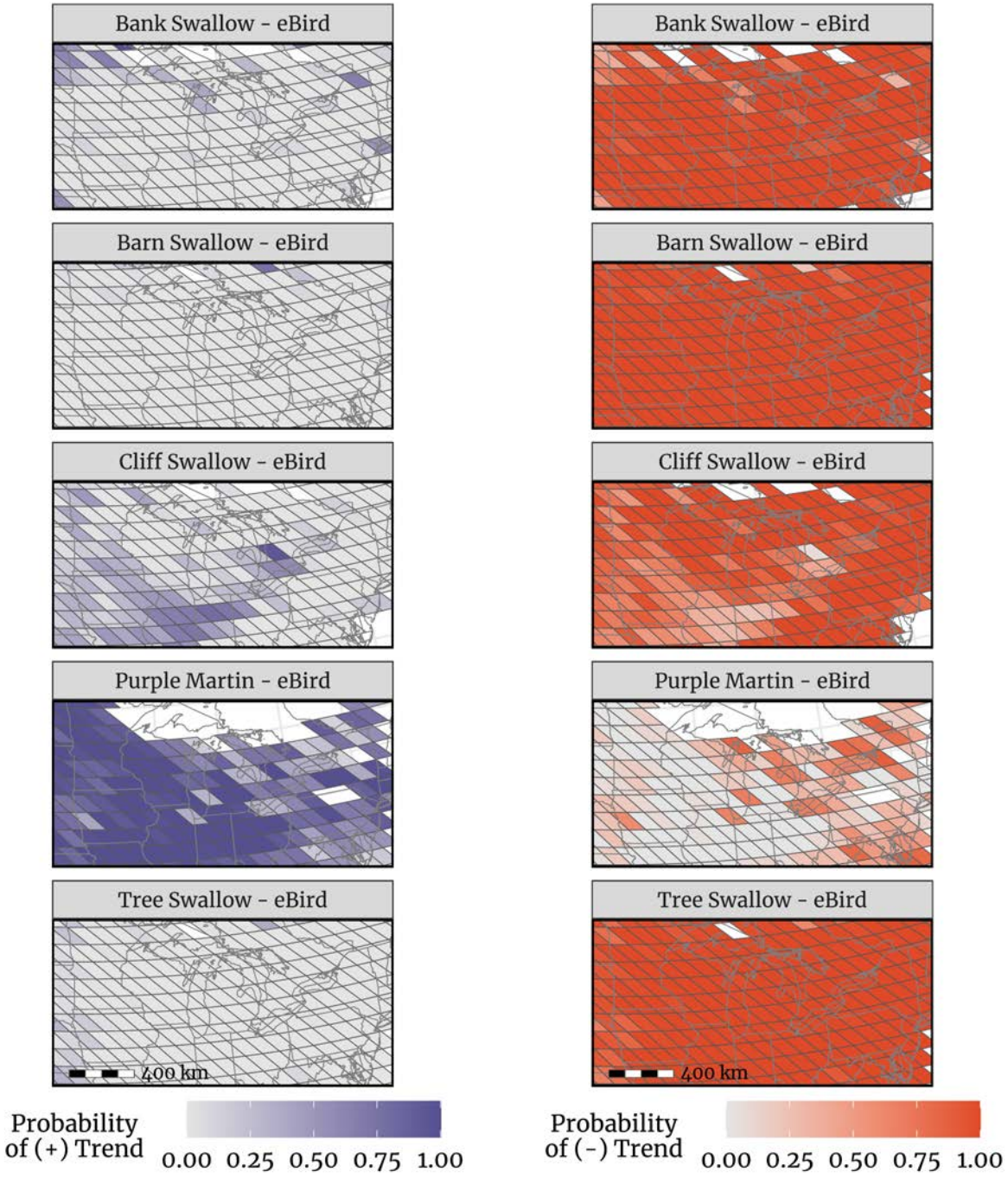
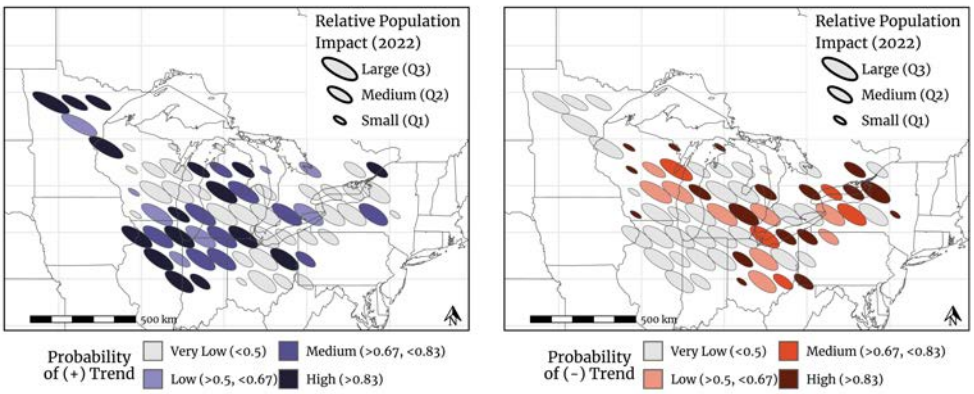
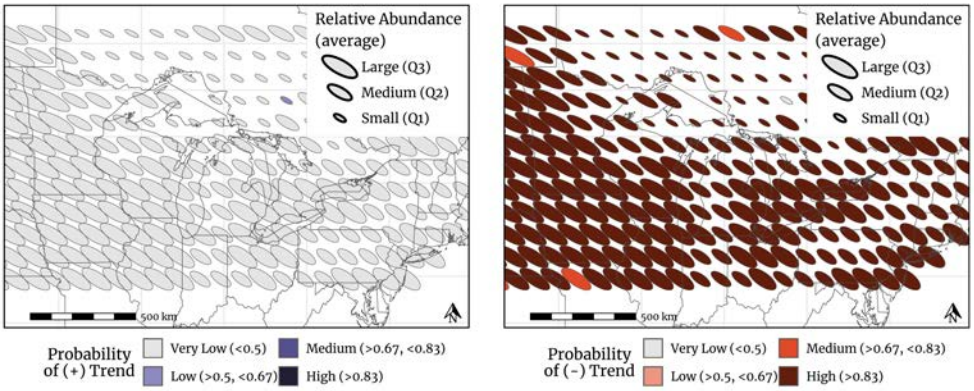


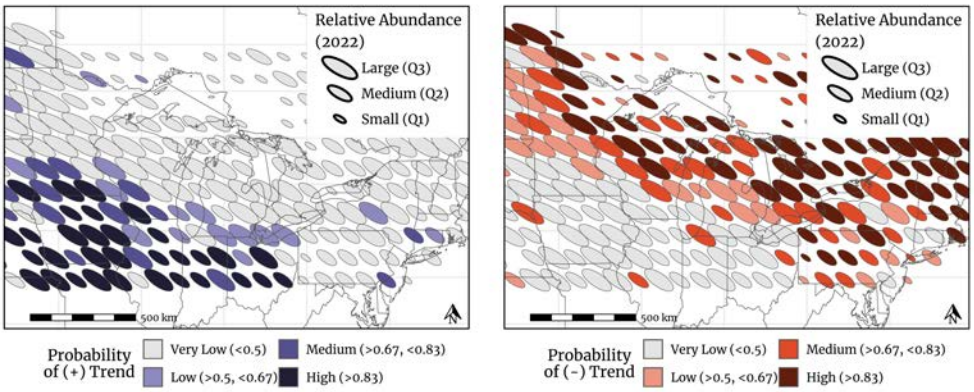
Figure 3.2: Directional probabilities of relative abundance trends for each species, obtained using eBird data. If the probability of a trend direction is lower than the probability of the flip of a fair coin, we consider that trend direction to be very unlikely. Shades of red from light to dark reflect increasing certainty on the direction of the trend in a cell.



(a) Directional probabilities of population impact trends obtained from radar data. Cell areas are proportional to estimated population impact in 2022.



(b) Directional probabilities of relative abundance trends obtained from eBird aggregated across all species. Cell areas are proportional to average relative abundance from 2012 to 2022.



(c) Directional probabilities of relative abundance trends obtained from Breeding Bird Survey data aggregated across all species. Cell areas are proportional to estimated relative abundances in 2022.

Figure 3.3: Maps representing the modeled probability of a negative or a positive trend for each of our data sources. If the probability of a trend direction is lower than the probability of the flip of a fair coin, we consider that trend direction to be very unlikely. Shades of red from light to dark reflect increasing certainty on the direction of the trend in a cell.

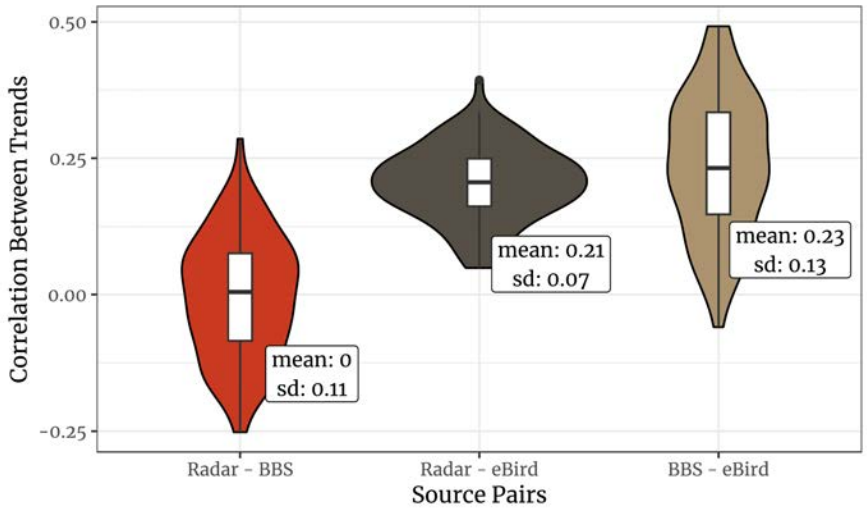
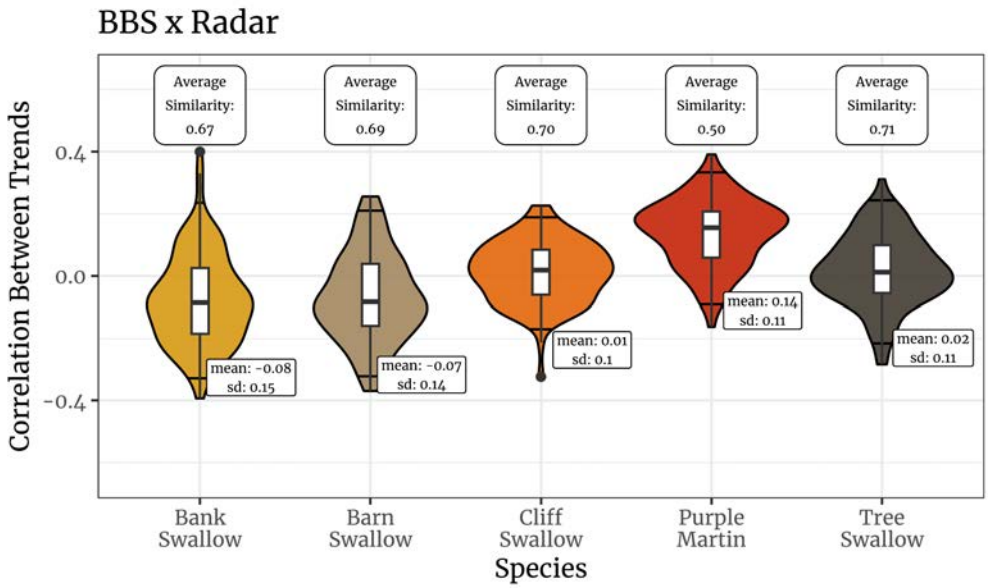
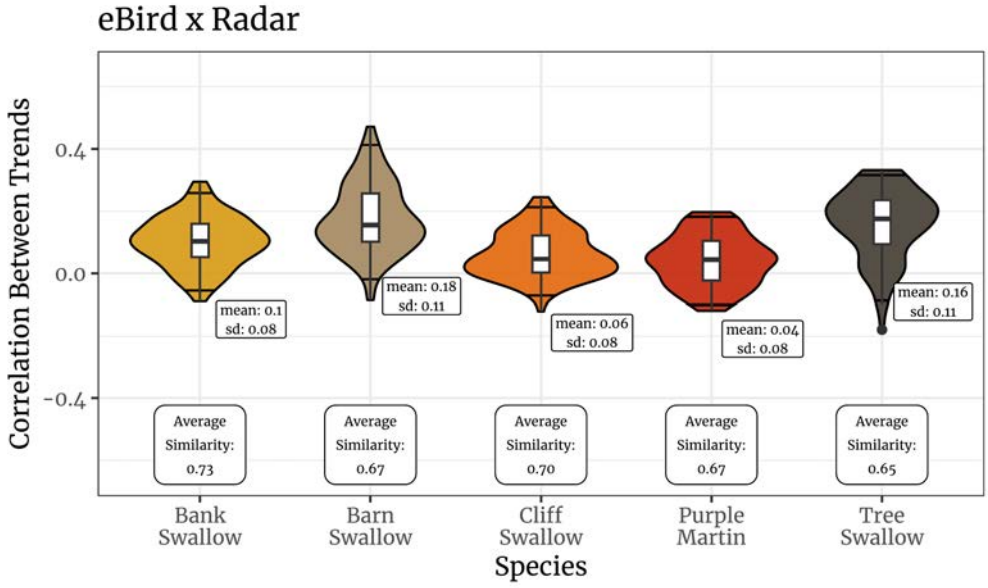


Figure 3.4: Density distribution of correlation coefficient estimates between geometric mean trends of each pair of sources, with respective means and standard deviations of each distribution.

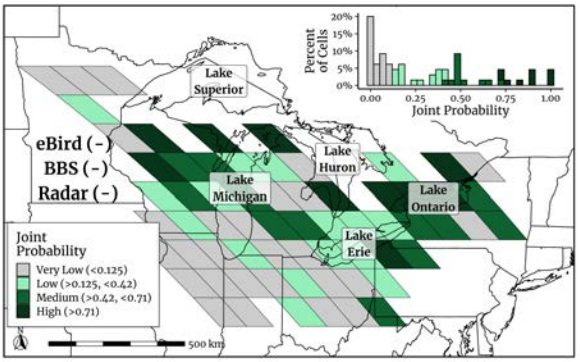


(a) Species-specific correlations between BBS trends and radar-derived trends.

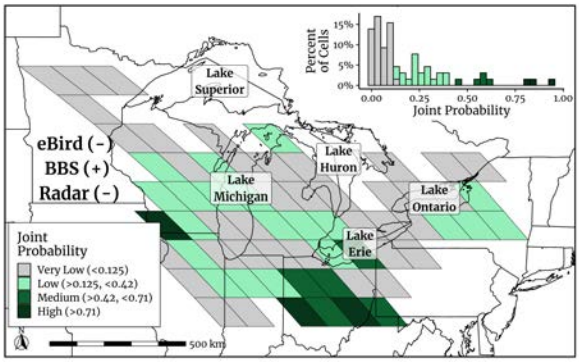


(b) Species-specific correlations between eBird trends and radar-derived trends.

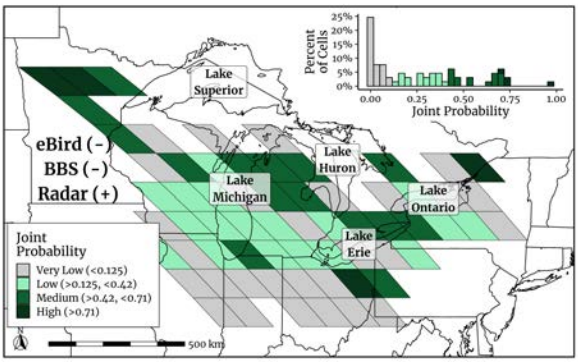
Figure 3.5: Density distributions of correlation coefficient estimates between radar-derived population impact trends and breeding population trends for each species, for each source (BBS or eBird). We calculated the percentage of overlapping area between empirical distributions of correlations of each pair of species. The mean similarity between a single species and all others is the average similarity described in the labels.



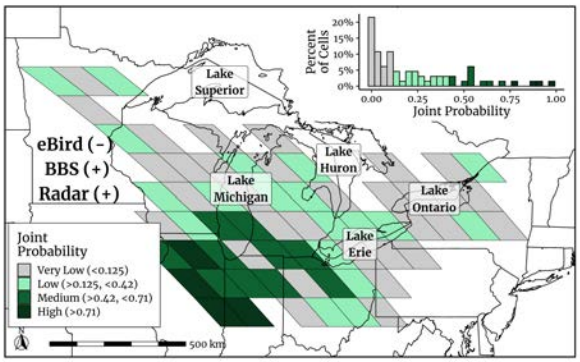
(a) Probability negative trend agreement across all sources.



(b) Probability of eBird (-), BBS (+) and radar (-) trends.



(c) Probability of eBird (-), BBS (-) and radar (+) trends.



(d) Probability of eBird (-), BBS (+) and radar (+) trends.

Figure 3.6: Maps representing joint directional probabilities of all possible combinations of trends across sources, except for eBird (+), since that is approximately equal to zero regardless of the estimates for the other two trends. Gray cells represent regions where the probability of a given combination is lower than the probability of obtaining that trend combination by flipping three fair coins. The different shades of green, from light to dark, reflect an increase in certainty about a given scenario. Histograms are a representation of the same joint probabilities, but better highlight the frequency of each value without the spatial information.

Bibliography

Belotti, M. C. T. D., Deng, Y., Zhao, W., Simons, V. F., Cheng, Z., Perez, G., Tielens, E., Maji, S., Sheldon, D., Kelly, J. F., & Horton, K. G. (2023). Long-term analysis of persistence and size of swallow and martin roosts in the US Great Lakes (V. Lecours & M. Van Den Broeke, Eds.). *Remote Sensing in Ecology and Conservation*, *9*(4), 469–482. https://doi.org/10.1002/rse2.323

Bled, F., Sauer, J., Pardieck, K., Doherty, P., & Royle, J. A. (2013). Modeling Trends from North American Breeding Bird Survey Data: A Spatially Explicit Approach. *Plos One*, *8*(12), e81867. https://doi.org/10.1371/journal.pone.0081867

Bridge, E. S., Pletschet, S. M., Fagin, T., Chilson, P. B., Horton, K. G., Broadfoot, K. R., & Kelly, J. F. (2015). Persistence and habitat associations of Purple Martin roosts quantified via weather surveillance radar. *Landscape Ecology*, *31*(1), 43–53. https://doi.org/10.1007/s10980-015-0279-0

Brown, C. R., & Brown, M. B. (2014). Breeding time in a migratory songbird is predicted by drought severity and group size. *Ecology*, *95*(10), 2736–2744. https://doi.org/10.1890/14-0425.1

Cheng, Z., Gabriel, S., Bhambhani, P., Sheldon, D., Maji, S., Laughlin, A., & Winkler, D. (2020). Detecting and Tracking Communal Bird Roosts in Weather Radar Data. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(01), 378–385. https://doi.org/10.1609/aaai.v34i01.5373

Chilson, P. B., Frick, W. F., Stepanian, P. M., Shipley, J. R., Kunz, T. H., & Kelly, J. F. (2012). Estimating animal densities in the aerosphere using weather radar: ToZor not toZ? *Ecosphere*, *3*(8), art72. https://doi.org/10.1890/es12-00027.1

Cox, A. R., Robertson, R. J., Lendvai, Á. Z., Everitt, K., & Bonier, F. (2019). Rainy springs linked to poor nestling growth in a declining avian aerial insectivore (*Tachycineta bicolor*). *Pro-*

ceedings of the Royal Society B: Biological Sciences, *286*(1898), 20190018. https://doi.org/10.1098/rspb.2019.0018

Deng, Y., Belotti, M. C. T. D., Zhao, W., Cheng, Z., Perez, G., Tielens, E., Simons, V. F., Sheldon, D. R., Maji, S., Kelly, J. F., & Horton, K. G. (2023). Quantifying long-term phenological patterns of aerial insectivores roosting in the Great Lakes region using weather surveillance radar. *Global Change Biology*, *29*(5), 1407–1419. https://doi.org/10.1111/gcb.16509

Droege, S. (1990). The North American Breeding Bird Survey. In *Survey Designs and Statistical Methods for the Estimation of Avian Population Trends* (p. 166, Vol. (1)). U.S. Fish; Wildlife Services. https://www.govinfo.gov/content/pkg/GOVPUB-I49-PURL-gpo59886/pdf/GOVPUB-I49-PURL-gpo59886.pdf

Dunn, P. O., & Winkler, D. W. (1999). Climate change has affected the breeding date of tree swallows throughout North America. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1437), 2487–2490. https://doi.org/10.1098/rspb.1999.0950

Dunning, J. (2008). CRC Handbook of Avian body masses. *CRC handbook of avian body masses, 2nd ed*, *384*. https://doi.org/10.1201/9781420064452

Edwards, M, B. P., Smith, A. C., & LaZerte, S. (2023). bbsBayes2: Hierarchical Bayesian Analysis of North American BBS Data. https://github.com/bbsBayes/bbsBayes2

Fink, D., Johnston, A., Strimas-Mackey, M., Auer, T., Hochachka, W. M., Ligocki, S., Oldham Jaromczyk, L., Robinson, O., Wood, C., Kelling, S., & Rodewald, A. D. (2023). A Double machine learning trend model for citizen science data. *Methods in Ecology and Evolution*, *14*(9), 2435–2448. https://doi.org/10.1111/2041-210x.14186

Fournier, A. M. V., Wilson, R. R., Gleason, J. S., Adams, E. M., Brush, J. M., Cooper, R. J., DeMaso, S. J., Driscoll, M. J. L., Frederick, P. C., Jodice, P. G. R., Ottinger, M. A., Reeves, D. B., Seymour, M. A., Sharuga, S. M., Tirpak, J. M., Vermillion, W. G., Zenzal, T. J., Lyons, J. E., & Woodrey, M. S. (2023). Structured Decision Making to Prioritize Regional Bird Monitoring Needs. *INFORMS Journal on Applied Analytics*, *53*(3), 207–217. https://doi.org/10.1287/inte.2022.1154

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, *1*(3), 515–534. https://doi.org/10.1214/06-ba117a

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, *2*(4). https://doi.org/10.1214/08-aoas191

Gregory, R. D., Vořišek, P., Noble, D. G., Van Strien, A., Klvaňová, A., Eaton, M., Meyling, A. W. G., Joys, A., Foppen, R. P., & Burfield, I. J. (2008). The generation and use of bird population indicators in Europe. *Bird Conservation International*, *18*(S1), S223–s244.

Gremillion, P. T., & Malone, R. F. (1986). Waterfowl Waste as a Source of Nutrient Enrichment in Two Urban Hypereutrophic Lakes. *Lake and Reservoir Management*, *2*(1), 319–322. https://doi.org/10.1080/07438148609354650

Hallmann, C. A., Foppen, R. P. B., Turnhout, C. A. M. v., Kroon, H. d., & Jongejans, E. (2014). Declines in insectivorous birds are associated with high neonicotinoid concentrations. *Nature*, *511*(7509), 341–343. https://doi.org/10.1038/nature13531

Horns, J. J., Adler, F. R., & Şekercioğlu, u. H. (2018). Using opportunistic citizen science data to estimate avian population trends. *Biological Conservation*, *221*, 151–159. https://doi.org/10.1016/j.biocon.2018.02.027

Horton, K. G., Doren, B. M. V., Sorte, F. A. L., Cohen, E. B., Clipp, H. L., Buler, J. J., Fink, D., Kelly, J. F., & Farnsworth, A. (2019). Holding steady: Little change in intensity or timing of bird migration over the Gulf of Mexico. *Global Change Biology*, *25*(3), 1106–1118. https://doi.org/10.1111/gcb.14540

Horton, K. G., La Sorte, F. A., Sheldon, D., Lin, T.-Y., Winner, K., Bernstein, G., Maji, S., Hochachka, W. M., & Farnsworth, A. (2020). Phenology of nocturnal avian migration has shifted at the continental scale. *Nature Climate Change*, *10*(1), 63–68. https://doi.org/10.1038/s41558-019-0648-9

Hudson, M.-A. R., Francis, C. M., Campbell, K. J., Downes, C. M., Smith, A. C., & Pardieck, K. L. (2017). The role of the North American Breeding Bird Survey in conservation. *The Condor*, *119*(3), 526–545. https://doi.org/10.1650/condor-17-62.1

Imlay, T. L., Mills Flemming, J., Saldanha, S., Wheelwright, N. T., & Leonard, M. L. (2018). Breeding phenology and performance for four swallows over 57 years: relationships with temperature and precipitation. *Ecosphere*, *9*(4), e02166. https://doi.org/10.1002/ecs2.2166

Johnson, A. E., Mitchell, J. S., & Brown, M. B. (2017). Convergent evolution in social swallows (Aves: Hirundinidae). *Ecology and Evolution*, *7*(2), 550–560. https://doi.org/10.1002/ece3.2641

Kardynal, K. J., Jardine, T. D., Génier, C. S. V., Bumelis, K. H., Mitchell, G. W., Evans, M., & Hobson, K. A. (2020). Mercury exposure to swallows breeding in Canada inferred from feathers grown on breeding and non-breeding grounds. *Ecotoxicology*, *29*(7), 876–891. https://doi.org/10.1007/s10646-020-02249-6

Keller, C. M. E., & Scallan, J. T. (1999). Potential Roadside Biases Due to Habitat Changes along Breeding Bird Survey Routes. *The Condor*, *101*(1), 50–57. https://doi.org/10.2307/1370445

Kelly, J. F., & Pletschet, S. M. (2017). Accuracy of swallow roost locations assigned using weather surveillance radar (M. Rowcliffe & D. Boyd, Eds.). *Remote Sensing in Ecology and Conservation*, *4*(2), 166–172. https://doi.org/10.1002/rse2.66

Kelly, J. F., Shipley, J. R., Chilson, P. B., Howard, K. W., Frick, W. F., & Kunz, T. H. (2012). Quantifying animal phenology in the aerosphere at a continental scale using NEXRAD weather radars. *Ecosphere*, *3*(2), art16. https://doi.org/10.1890/es11-00257.1

Knape, J. (2016). Decomposing trends in Swedish bird populations using generalized additive mixed models (G. Siriwardena, Ed.). *Journal of Applied Ecology*, *53*(6), 1852–1861. https://doi.org/10.1111/1365-2664.12720

Knight, S. M., Bradley, D. W., Clark, R. G., Gow, E. A., Bélisle, M., Berzins, L. L., Blake, T., Bridge, E. S., Burke, L., Dawson, R. D., Dunn, P. O., Garant, D., Holroyd, G. L., Hussell, D. J. T.,

Lansdorp, O., Laughlin, A. J., Leonard, M. L., Pelletier, F., Shutler, D., … Norris, D. R. (2018). Constructing and evaluating a continent-wide migratory songbird network across the annual cycle. *Ecological Monographs*, *88*(3), 445–460. https://doi.org/10.1002/ecm.1298

Lawler, J. J., & O'Connor, R. J. (2004). How Well do Consistently Monitored Breeding Bird Survey Routes Represent the Environments of the Conterminous United States? *The Condor*, *106*(4), 801–814. https://doi.org/10.1093/condor/106.4.801

Link, W. A., & Sauer, J. R. (2002). A hierarchical analysis of population change with application to the Cerulean Warbler. *Ecology*, *83*(10), 2832–2840.

Link, W. A., Sauer, J. R., & Niven, D. K. (2020). Model selection for the North American Breeding Bird Survey. *Ecological Applications*, *30*(6). https://doi.org/10.1002/eap.2137

Loh, J., Green, R. E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V., & Randers, J. (2005). The Living Planet Index: using species population time series to track trends in biodiversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1454), 289–295. https://doi.org/10.1098/rstb.2004.1584

Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. M., Elston, D. A., Scott, E. M., Smith, R. I., Somerfield, P. J., & Watt, A. D. (2010). Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in Ecology & Evolution*, *25*(10), 574–582. https://doi.org/10.1016/j.tree.2010.06.016

Manny, B. A., Johnson, W. C., & Wetzel, R. G. (1994). Nutrient additions by waterfowl to lakes and reservoirs: predicting their effects on productivity and water quality. *Hydrobiologia*, *279*(1), 121–132. https://doi.org/10.1007/bf00027847

Massimino, D., Harris, S. J., & Gillings, S. (2020). Phenological mismatch between breeding birds and their surveyors and implications for estimating population trends. *Journal of Ornithology*, *162*(1), 143–154. https://doi.org/10.1007/s10336-020-01821-5

Morris, M., Wheeler-Martin, K., Simpson, D., Mooney, S. J., Gelman, A., & DiMaggio, C. (2019). Bayesian hierarchical spatial models: Implementing the Besag York Mollié model in stan.

Spatial and Spatio-temporal Epidemiology, *31*, 100301. https://doi.org/10.1016/j.sste.2019.100301

Nebel, S., Mills, A., McCracken, J. D., & Taylor, P. D. (2010). Declines of Aerial Insectivores in North America Follow a Geographic Gradient. *Avian Conservation and Ecology*, *5*(2). https://doi.org/10.5751/ace-00391-050201

Parmesan, C., Gaines, S., Gonzalez, L., Kaufman, D. M., Kingsolver, J., Townsend Peterson, A., & Sagarin, R. (2005). Empirical perspectives on species borders: from traditional biogeography to global change. *Oikos*, *108*(1), 58–75. https://doi.org/10.1111/j.0030-1299.2005.13150.x

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, *421*(6918), 37–42. https://doi.org/10.1038/nature01286

Pastore, M., & Calcagnì, A. (2019). Measuring Distribution Similarities Between Samples: A Distribution-Free Overlapping Index. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01089

Powell, G. V. N., Fourqurean, J. W., Kenworthy, W. J., & Zieman, J. C. (1991). Bird colonies cause seagrass enrichment in a subtropical estuary: Observational and experimental evidence. *Estuarine, Coastal and Shelf Science*, *32*(6), 567–579. https://doi.org/10.1016/0272-7714(91)90075-m

R Core Team. (2018). *R: A Language and Environment for Statistical Computing* (v3.4.1). R Foundation for Statistical Computing. https://www.R-project.org/

Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., Stanton, J. C., Panjabi, A., Helft, L., Parr, M., & Marra, P. P. (2019). Decline of the North American avifauna. *Science*, *366*(6461), 120–124. https://doi.org/10.1126/science.aaw1313

Russell, K. R., & Gauthreaux, S. A. (1998). Use of Weather Radar to Characterize Movements of Roosting Purple Martins. *Wildlife Society Bulletin (1973-2006)*, *26*(1), 5–16. http://www.jstor.org/stable/3783794

Russell, K. R. (1996). *Spatial and temporal patterns of a Purple Martin (Progne subis) roost: a radar and direct visual study* [Master's thesis, Clemson University].

Saino, N., Szép, T., Romano, M., Rubolini, D., Spina, F., & Møller, A. P. (2003). Ecological conditions during winter predict arrival date at the breeding quarters in a trans-Saharan migratory bird. *Ecology Letters*, *7*(1), 21–25. https://doi.org/10.1046/j.1461-0248.2003.00553.x

Saracco, J. F., Desante, D. F., & Kaschube, D. R. (2008). Assessing Landbird Monitoring Programs and Demographic Causes of Population Trends. *The Journal of Wildlife Management*, *72*(8), 1665–1673. https://doi.org/10.2193/2008-129

Sauer, J. R., Niven, D. K., Hines, J. E., Ziolkowski, D. J., Jr, Pardieck, K. L., Fallon, J. E., & Link, W. A. (2021). *The North American Breeding Bird Survey, Results and Analysis 1966 - 2021* (tech. rep. No. Version 2.07.2021). USGS Patuxent Widlife Research Center. Laurel, Maryland.

Sejkora, P., Kirisits, M. J., & Barrett, M. (2011). Colonies of Cliff Swallows on Highway Bridges: A Source of Escherichia coli in Surface Waters1. *JAWRA Journal of the American Water Resources Association*, *47*(6), 1275–1284. https://doi.org/10.1111/j.1752-1688.2011.00566.x

Shave, A., Garroway, C. J., Siegrist, J., & Fraser, K. C. (2019). Timing to temperature: Egg-laying dates respond to temperature and are under stronger selection at northern latitudes. *Ecosphere*, *10*(12). https://doi.org/10.1002/ecs2.2974

Smith, A. C., D. Binley, A., Daly, L., Edwards, B. P. M., Ethier, D., Frei, B., Iles, D., Meehan, T. D., Michel, N. L., & Smith, P. A. (2023). Spatially explicit Bayesian hierarchical models improve estimates of avian population status and trends. *Ornithological Applications*, duad056. https://doi.org/10.1093/ornithapp/duad056

Smith, A. C., & Edwards, B. P. M. (2020). North American Breeding Bird Survey status and trend estimates to inform a wide range of conservation needs, using a flexible Bayesian hierarchical generalized additive model. *The Condor*, *123*(1). https://doi.org/10.1093/ornithapp/duaa065

Smith, A. C., Hudson, M.-A. R., Downes, C. M., & Francis, C. M. (2015). Change Points in the Population Trends of Aerial-Insectivorous Birds in North America: Synchronized in Time across Species and Regions (R. M. Brigham, Ed.). *Plos One*, *10*(7), e0130768. https://doi.org/10.1371/journal.pone.0130768

Somershoe, S. G., Twedt, D. J., & Reid, B. (2006). Combining Breeding Bird Survey and Distance Sampling to Estimate Density of Migrant and Breeding Birds. *The Condor*, *108*(3), 691–699. https://doi.org/10.1093/condor/108.3.691

Spiller, K. J., & Dettmers, R. (2019). Evidence for multiple drivers of aerial insectivore declines in North America. *The Condor*, *121*(2), duz010. https://doi.org/10.1093/condor/duz010

Stan Development Team. (2023). RStan: the R interface to Stan. https://mc-stan.org/

Strimas-Mackey, M., Ligocki, S., Auer, T., & Fink, D. (2023). *ebirdst: Access and Analyze eBird Status and Trends Data Products*. https://ebird.github.io/ebirdst/

Van Der Bles, A. M., Van Der Linden, S., Freeman, A. L., Mitchell, J., Galvao, A. B., Zaval, L., & Spiegelhalter, D. J. (2019). Communicating uncertainty about facts, numbers and science. *Royal Society open science*, *6*(5), 181870.

van Strien, A. J., Soldaat, L. L., & Gregory, R. D. (2012). Desirable mathematical properties of indicators for biodiversity change. *Ecological Indicators*, *14*(1), 202–208. https://doi.org/10.1016/j.ecolind.2011.07.007

Veech, J. A., Pardieck, K. L., & Ziolkowski, D. J., Jr. (2017). How well do route survey areas represent landscapes at larger spatial extents? An analysis of land cover composition along Breeding Bird Survey routes. *The Condor*, *119*(3), 607–615. https://doi.org/10.1650/condor-17-15.1

Walker, J., & Taylor, P. (2017). Using eBird data to model population change of migratory bird species. *Avian Conservation and Ecology*, *12*(1). https://doi.org/10.5751/ace-00960-120104

Wolkovich, E. M., Cook, B. I., McLauchlan, K. K., & Davies, T. J. (2014). Temporal ecology in the Anthropocene. *Ecology Letters*, *17*(11), 1365–1379. https://doi.org/10.1111/ele.12353

Wood, S. N. (2017, May). *Generalized Additive Models: An Introduction with R*. Chapman; Hall/CRC. https://doi.org/10.1201/9781315370279

Youngflesh, C., Socolar, J., Amaral, B. R., Arab, A., Guralnick, R. P., Hurlbert, A. H., LaFrance, R., Mayor, S. J., Miller, D. A. W., & Tingley, M. W. (2021). Migratory strategy drives species-level variation in bird sensitivity to vegetation green-up. *Nature Ecology & Evolution*, *5*(7), 987–994. https://doi.org/10.1038/s41559-021-01442-y

Chapter 4

Phenology and spatial distribution of swallow and martin roosts in the Low Negro-Solimões region of the Amazon Rainforest

4.1 Introduction

Every year, billions of migratory birds travel thousands of kilometers between their breeding and nonbreeding grounds. The fact that these two phases of these species' annual cycle occur in different places means that their population sizes could be influenced by conditions they encounter during migration or in entirely different regions of the world (Newton, 2004; Sherry & Holmes, 1996). Moreover, conditions experienced in one season could affect survival and fecundity in the following season, resulting in complex seasonal interactions that ultimately have important population ramifications (Harrison et al., 2011). Additionally, since individuals that breed together could spend their nonbreeding period in very distant regions, observing the effects of these seasonal interactions can be challenging (Cristol et al., 1999; Webster et al., 2002). These complexities create challenges for researching and conserving migratory species, especially when basic knowledge about the time they spend in the tropics is lacking. Within the Americas, for example, knowledge about the nonbreeding geographic range limits and habitat associations of some Nearctic-Neotropical migrants such as the Connecticut Warbler (*Geothlypis agilis*), or the Chimney Swift (*Cheatura pelagica*) is incomplete or based on very few records (Gomes, 2023; Steeves et al., 2020; Wolfe et al., 2012). This lack of knowledge is even more dramatic in the case of Austral migrants, for which data is scarce throughout the annual cycle (Chesser, 1994; Faaborg et al., 2010a; Jahn et al., 2004).

The literature suggests many reasons for these knowledge gaps, such as challenges associated with asymmetric collaborations between north and south scientific institutions (Malhado

et al., 2014), a tight link between national identity and science funding, particularly during the Cold War (Wagner et al., 2015), a perception amongst ornithologists that the tropical nonbreeding period of migrants was less important (Faaborg et al., 2010b), and the lack of an integrative conceptual framework that would allows us to develop new hypotheses and research questions about the full annual cycle of migratory species (Bowlin et al., 2010; Faaborg et al., 2010a; Marra et al., 2015). These cultural and structural challenges are being overcome (Malhado et al., 2014; Perez & Hogan, 2018; Wagner et al., 2015), however one barrier remains unaltered: one of the most important nonbreeding regions of the Americas, the Amazon Rainforest, is mostly inaccessible for systematic long-term field studies which would be needed if we were to fill some of these gaps (Lees et al., 2020).

Swallows and martins epitomize many of these problems. Species that reside in North America during their breeding period, such as Tree Swallows (*Tachycineta bicolor*), Purple Martins (*Progne subis*), and Barn Swallows (*Hirundo rustica*), are relatively well-studied and, like other species of aerial insectivores, have been facing important breeding population declines (Rosenberg et al., 2019; Tautin et al., 2017). However, knowledge about the period they spend in the Neotropics is still deficient, particularly for species that spend their nonbreeding period in the Amazon Rainforest, such as Purple Martins. Additionally, since many of the species of the *Progne* genus are virtually indistinguishable in the field, new locations are still being added every couple of years to the nonbreeding ranges of the intertropical migrants of this genus, such as the Caribbean Martin (*Progne domicensis*) and the Cuban Martin (*Progne cryptoleuca*, Cooper, 2024; García-Lau et al., 2021; Perlut et al., 2017). Lastly, Austral migrants that breed in southern South America and spend their nonbreeding period in the Amazon Rainforest, like the Southern Martin (*Progne elegans*) and the Brown-chested Martin subspecies fusca (*Progne tapera fusca*), have basic life history, phenology, and distribution knowledge gaps. Community science data (Santos et al., 2021), stable isotopes (Imlay et al., 2018), and light level geolocators (Fraser et al., 2012) have allowed us to advance towards closing some of these gaps, but data is still insufficient to understand how these species are responding to recent extreme droughts

and widespread wildfires in the Amazon and what can be done to reverse the declining trends in breeding populations in the Northern Hemisphere.

Weather radars can circumvent some of the challenges associated with the remoteness of the Amazon Rainforest, allowing us to collect data on the location, timing, altitude, and individual density of mass aerial movements for a broad range of taxa. In Europe and the United States, for example, data collected by operational weather radars has been used to observe, monitor, and predict nocturnal bird migration at a continental scale (see Bauer et al., 2018, for a comprehensive review), allowing us to reframe our conceptual understanding of these seasonal movements (Heffernan et al., 2014; Kelly & Horton, 2016). Additionally, radars have also been used to study emergences of aerial insects, which are notoriously difficult to observe (Bauer et al., 2024; Tielens et al., 2021).In these examples, radars can offer us information about biological events, but they lack the taxonomic precision of on the ground studies. However, in some unique instances, when use of the airspace at a given time and space is taxon-specific, we can refer to ancillary life history knowledge to infer the family or species observed by the radars. This is the case, for example, of Mexican free-tailed bats (*Tadarida brasiliensis*), whose enormous dispersals from well-known caves and bridges in Texas are systematically captured by local radars (Frick et al., 2012; Horn & Kunz, 2008; Stepanian & Wainwright, 2018).

In addition to select bat movements, swallows and martins are also identified during phases of their distinctive communal roosting behavior that can be systematically detected by nearby (150km) meteorological radars. During non-breeding phases of their annual cycle, some species of this family gather daily in communal roosts that can concentrate hundreds of thousands of individuals. These large aggregations offer protection against predators and could act as information-sharing centers, where birds learn about optimal foraging locations (Ward & Zahavi, 1973; Weatherhead, 1983; Winkler, 2006). The early morning dispersal of birds from the roosts creates a distinctive doughnut-shaped high reflectivity signal on radar reflectivity renderings, which can be used to quantify roost size, location, date, and timing (Belotti et al., 2023; Russell, 1996). With data from the NEXRAD network in the United States, comprehensive maps

of roost locations across most of the eastern range of these species in North America are coming into focus (Bridge et al., 2015; Kelly & Pletschet, 2017; Russell et al., 1998). Additionally, this network has been used to quantify roosting phenology and observe how temporal patterns have changed over the past 21 years, possibly due to climate change (Deng et al., 2023; Kelly et al., 2012). Studies have also estimated roost size distribution and persistence, based on 21 years of data collected from 12 radar stations in the Great Lakes region (Belotti et al., 2023). Moreover, this network has been instrumental in analyzing patterns of roost size changes during the non-breeding season of Tree Swallows in Louisiana and Florida (Laughlin et al., 2016). However, as of yet, no systematic analysis via radar has been conducted documenting roosting dynamics in the heart of non-breeding biomes, namely the Amazon Rainforest.

To fill this gap, we give the first description of swallow and martin roosts in Brazil based on data collected by a radar in Manaus, in the state of Amazonas, Brazil, from 2014 to 2015. These two years of data are used to describe the yearly phenology of swallow and martin roosting behavior in the Amazon Rainforest. We also describe the spatial distribution of these aggregations, and how they vary within and across years. We compare the roosting behavior observed in the region around Manaus with observations derived from North America's Great Lakes region during the same period. Because the Amazon Rainforest hosts swallows and martins migrating both from North America (Purple Martins and Barn Swallows) and from southern South America (Southern Martins and Brown-chested martins subspecies fusca), we hypothesize that, contrary to the dynamics observed within the Great Lakes, roosting activity in Manaus will occur throughout the year. Additionally, since a well-documented roost composed predominantly of Purple Martins (but also gathering significant numbers of Brown-Chested Martins) occurs close to Manaus, in the Negro River, from January to May, we expected to observe a peak in roosting activity during that time (Arruda, 2023; Oliveira, 2021).

4.2 Materials and Methods

4.2.1 Study Region

The region surrounding the urban area of Manaus is a mosaic of terra firme moist forests (forests that are not seasonally waterlogged), secondary forests, sparse pasture areas associated with roads, and two major rivers and their affluents, the Negro river and the Amazon/Solimões river, with their associated floodplains (so called river-floodplain systems, Junk, 1997; Olson et al., 2001; Project, 2021). These river-floodplain systems can be further divided into white water systems with high productivity due to high loads of suspended sediment and nutrients, such as the Amazon/Solimões river, and nutrient-poor black water systems such as the Negro river (Junk et al., 2012, 2015). Floodplains associated with white water rivers are called *várzeas* (singular: *várzea*), whereas those associated with black water rivers are referred to as *igapós* (singular: *igapó*, Junk et al., 2011). Wetlands correspond to approximately 34% of the total area of the Amazon basin (Junk et al., 2011; Wittmann et al., 2022).

Floodplains in the Amazon basin are drained by a combination of overbank diffusive flow and channelized flows through numerous floodplain channels (Trigg et al., 2012). These floodplain channels play a key role in determining the hydrodynamic characteristics of river-floodplain systems, which in turn regulate biological and biogeochemical processes in the Amazon basin. Due to its mostly flat landscape, wetlands in the Amazon lowlands are subject to a flood pulse of predictable magnitude and timing, with one high-water and one low–water period during the year (Junk et al., 1989; Wittmann et al., 2022, see Figure E.2 for plots characterizing this behavior). This monomodal flood pulse characterizes the flow regime in most of the Amazonian lowlands, that is primarily driven by precipitation in the upper catchment area, which explains the time lag in flood peaks that exists between upstream and downstream reaches (Correa et al., 2022). The flow of water from the main river to its network of floodplain channels creates a complex pattern of erosion and sedimentation that is driven by the differences between white and black water rivers and results in spatial heterogeneity in vegetation community composition. This heterogeneity, paired with the yearly disturbance of the flood pulse, gives rise to

diverse seasonally available habitats and influences vertebrate species distributions and landscape connectivity within the alluvial landscape (Junk et al., 1989; Peterman, 1997; Salo et al., 1986; Wittmann et al., 2022).

4.2.2 Data Extraction and Processing

For this study, we used a set of two years of volumetric data collected by the S-band Doppler weather radar located in the city of Manaus, Amazonas (3°8′56.0″ S, 59°59′29.1″ W) and made available to us by the Amazon Protection System (SIPAM, Sistema de Proteção da Amazônia). The technical specifications of the radar are summarized in Table 4.1. We downloaded and rendered all scans made between January 1st, 2014 and December 31st, 2015 using pyArt version 1.18.1 (Helmus & Collis, 2016). We narrowed our search to an hour before and two hours after local sunrise (approximately 5:00 to 8:00 in local time) which is usually when roost departures occur (Belotti et al., 2023; Kelly & Pletschet, 2017; Russell, 1996). We then used Label Studio to manually annotate each scan, drawing bounding boxes around each roost and classifying them as either clear roosts, noise-contaminated roosts, or roosts caught in the radar ground clutter region (Tkachenko et al., 2020). We also made notes for each day about the presence of precipitation or anomalous propagation that could have affected our ability to detect roosts if they were present (for an example of a roost caught in clutter and its vertical cross section, see E.1).

After manual annotation, we used the coordinates of the bounding boxes to extract and filter reflectivity measurements within each roost. Because one of our roosts was frequently within the radar's ground clutter region, we designed clutter masks to remove some of the ground clutter from our scans. We did so by selecting scans within the lower 2% quantile of summed reflectivity within our study period. We used these scans to calculate two average masks of ground clutter reflectivity (in cm^6/km^3) for every elevation angle, for each azimuthal configuration (294 or 360 azimuths). We limited our masks to a region within up to 27.5 km from the radar, where we usually observed ground clutter during our manual screening. We then subtracted this mask from reflectivity measurements in every scan before estimating the number of birds within each

Instrument Parameter	Value
Location	3°8′56.0″ S, 59°59′29.1″ W
Wavelength	0.10 m
Frequency	2.997925×10^9 Hz
Horizontal Beamwidth	1.8°
Vertical Beamwidth	1.8°
Range gate	500 m
Volume coverage pattern	fixed
Elevation angles	0.9°, 1.5°, 2°, 3°, 4°, 5°, 6°, 7°, 8°, 9°, 10.5°, 12°, 13.5°, 15°, 16.5°, 18°, 19°
Pulse width	10^5 μs
Pulse repetition time	0.00161 s
Unambiguous range	240000 m
Nyquist velocity	15.525 $\mathrm{m\,s^{-1}}$
Azimuthal strategy	294 or 360 azimuths per sweep

Table 4.1: Technical specifications of the S-band Doppler weather radar located in Manaus, Amazonas.

bounding box. We refer the reader to Figure E.3, where we plot the number of birds considered to be clutter for each sweep within each detection. We also removed from our analysis pixels that had very high (greater than 30 dBZ) or very low (lower than 5 dBZ) reflectivity since those are unlikely to correspond to swallows and martins.

Calculating the volume sampled by the radar at each azimuth and range presents a challenge because the beamwidth of the Manaus radar is 1.8°, while its azimuthal strategy consists of performing a rotation of 1° (360 azimuths) or 1.22° (294 azimuths) between each azimuth, for each sweep. Because of the rotation angle of both strategies is smaller than the radar beamwidth, a region between each pair of consecutive azimuths is sample twice. We accounted for this double sampling by using Monte Carlo integration to estimate the volume of the intersection between two radar sampling volumes, assuming that the radar sampling volume has the shape of a conical frustum of beamwidth 1.8°. We could then use the estimated intersection volume to calculate the number of birds within the intersection and the number of birds outside of it. Within the intersection, since we have two measurements of reflectivity, we use their average to procede with the calculation of the number of birds.

We estimated the number of birds within each roost following the method described in Chilson et al. (2012): we converted reflectivity factor (Z_e, measured in dBZ) within each radar sampling volume to linear scale (mm^6/m^3), and then transformed it into reflectivity (η), which can be interpreted as the density of scatterers of a given radar cross section (RCS) in the atmosphere (units are cm^6/km^3). We can obtain the radar cross section of a typical Hirundinidae from its mass by using the relationship proposed in Horton et al. (2019): $\log(\text{RCS}) = 0.699x\log(\text{mass})$. Since Purple Martins are the largest Hirundinidae in the Americas, we used their mass of 51 g (Dunning, 2008), which corresponds to an RCS of 15.2 cm^2, to obtain conservative estimates of the number of birds. In dividing η by this RCS, we are left with the density in units of number of birds per km^3. The last step, thus, is to obtain the volume sampled by the radar at each azimuth (antenna rotation angle) and range.

Calculating this volume presents a challenge because the beamwidth of the Manaus radar is 1.8°, while its azimuthal strategy consists of performing a rotation of 1° (360 azimuths) or 1.22° (294 azimuths) between each azimuth, for each sweep. Because the rotation angle of both strategies is smaller than the radar beamwidth, there is a region between each pair of consecutive azimuths that is sampled twice. We accounted for this double sampling by using Monte Carlo integration to estimate the volume of the intersection between two radar sampling volumes in each azimuthal strategy, assuming that the radar sampling volume has the shape of a conical frustum of beamwidth 1.8°. We could then use the estimated intersection volume to calculate the number of birds within the two intersections (N_{left} and N_{right}) and the number of birds within the single-sampled region (N_{core}). The number of birds caught within each intersection was divided between the two corresponding azimuths so that the total number of birds per azimuth and range was equal to $\frac{N_{left}}{2} + N_{core} + \frac{N_{right}}{2}$. In doing this for all the azimuth and range combinations within each bounding box and summing the results, we get estimates of the number of birds within the roost dispersal at the time of each sweep. We restricted our counts to the first sweep (at an elevation angle of 0.9°), since sweeps at higher elevation angles either overlap the first one (due to the high beamwidth) or overshoot the roost dispersal.

The bounding boxes represent a single scan or frame of the dispersal event. A full dispersal typically happens across multiple scans, and thus has multiple detections. The groups of detections that represent a single dispersal event are called tracks. We grouped detections into tracks by using a heuristic procedure that matched each detection with the one in the previous scan with which it overlapped the most. If a given detection overlapped more than one detection in the previous scan, we paired it with the nearest one. The results of this procedure were manually inspected for accuracy. We could then further summarize counts per track by extracting the peak count within each track. We summed counts across roosts on each day to obtain estimates of daily relative abundance. Because roosts often occur in the same place day after day, some regions accumulate roost occurrences throughout a season. In order to understand the temporal patterns of use of each roosting region within a year and between our two study years, we used mean shift clustering to find regions that were consistently used by swallows and martins to roost throughout our study period (this approach was first employed in Belotti et al. (2023)).

In our analyses, we considered a day to have been sampled if it had at least 6 scans (approximately 1/3 of the expected 15 scans within our daily sampling period), and those scans were not radar failures and had no precipitation or anomalous propagation. If a scan had precipitation or anomalous propagation, but a roost was visible, we flagged it as a presence, but did not use it for our quantitative analyses of roost size.

We compared daily relative abundances and roost size distribution from Manaus with those observed in the Great Lake region. For this purpose, we used a dataset from the same period collected from 12 operational S-band radar stations in the Great Lakes region of the United States (Belotti et al., 2024). We obtained this dataset by running a machine learning pipeline that is capable of detecting roost signatures and tracking them as they disperse in the airspace. The output of this pipeline are bounding boxes, which were labeled following a protocol similar to what we just described for the dataset from Manaus. Reflectivity measurements were extracted from each bounding box using the same approach, except no clutter filtering was implemented

(roosts were not typically found within the clutter region of the radars). To assert that the two datasets were comparable, we extracted the minimum and maximum detected reflectivity for each range gate for all scans made within our study period in the Manaus dataset, and for a sample of 10% of the scans from the Great Lakes region. We averaged measurements across scans to obtain one estimate of the minimum and one of the maximum detectable signals for each range gate (see Figures E.4 and E.5). We found that the NEXRAD radars were capable of detecting much lower reflectivity measurements than the radar from Manaus, indicating that the latter is more likely to underestimate the number of birds since it's not as sensitive to lower reflectivities.

4.3 Results

We found 308 days in 2014 and 205 days in 2015 with at least six scans available in the archive during our sampling window. In 2014, 12 days in November had to be removed due to radar failure. Precipitation or anomalous propagation led us to discard 118 days in 2014 and 160 days in 2015 in our estimation of roost size, even though roosts were observed on 78 of the contaminated days in 2014 and on 131 of the contaminated days in 2015. After removing days in which sampling did not satisfy our criteria due to radar failure or lack of scans, we were left with 178 sampled days in 2014 and 145 sampled days in 2015 for the purposes of estimating roost size. In our analysis of presence/absence of roosts, which combines clear and contaminated scans, we were left with 256 days (178 clear, 78 weather contaminated) in 2014 and 276 days (145 clear, 131 weather contaminated) in 2015.

Roosting activity was observed in our study region throughout the year in both 2014 and 2015. In 2014, of the 256 days sampled for the presence/absence of roosts, only 13 were true zeros, which are days when we could observe the entire radar field but still found no evidence of roosts (see Figure 4.1). Similarly, in 2015, of the 276 days sampled for presence/absence of roosts, only seven were true zeros (see Figure 4.1). The activity pattern in both years, however, was distinct. In 2014, from January to May, when the presence of Purple Martins likely drives

roosting activity, we found daily counts ranging from 1,487 to 31,649 birds (see Figure 4.1). Activity steadily decreases between May and June and picks up again in July. Starting in August 2014, however, when Purple Martins began arriving in the Amazon, we observed daily counts twice as large as those from the beginning of 2014, with a steady increase until the end of that year (see Figure 4.1). In 2015, the daily counts continued to rise, reaching a peak of 422,538 birds on March 26. The numbers decrease between May and June and start increasing again in late July, similar to what we observed in 2014 (see Figure 1).

We found eight distinct spatial clusters of roosting activity and, of those, three clusters were systematically used in both years (Figure 4.2). The remaining clusters, which we will refer to as satellite roosts, were used transiently, especially when we observed the dramatic increase in the daily number of birds in late 2014 and 2015 (see Figure 4.3). From East to West, the first stable cluster (CL0000) was located in the eastern region of the Anavilhanas archipelago, mainly within the Anavilhanas National Park (federally protected land with IUCN Category II). The second stable cluster, CL0001, is the previously documented Negro River roost, described in Arruda (2023) and Oliveira (2021). This roost occurs on Comaru Island, an ephemeral partially inundated island within the limits of a state-managed protected area reserved for sustainable use (IUCN Category V) between Manaus and Iranduba (see Figure 4.2). The third stable roost (CL0002) was in the unprotected region surrounding Lago do Rei, in the municipality of Careiro da Várzea (see Figure 4.2). The largest satellite roost, CL0003, was located on demarcated indigenous land (Terra Indígena Murutinga/Tracajá, see Funai (2024)), while satellite roost CL0004 is on land claimed by indigenous communities of ethnicity Mura (Ricardo et al., 2023). In both years, we observed that the main roost moved from the Anavilhanas archipelago (CL0000) to Comaru Island (CL0001) just before the peak of the Purple Martin season (February 15, 2014, and February 14, 2015, see Figure 4.3). It moved back to Anavilhanas on July 1, 2014, much later than it did on 2015 (May 7).

We subsetted four years of data from the Great Lakes (2013-2016), in North America, for our comparison. The area sampled in the Great Lakes corresponds to the union of 12 partially

overlapping radar ranges, covering approximately 787,229 km^2. This area is 8.7 times larger than the area sampled in Manaus (90,000 km^2). However, we found that the average number of birds per day using the region around Manaus in 2014 — 71,606 (sd: 106,553) individuals per day — was 1.26 larger than the largest average number of birds per day we found on the Great Lakes region — 56,956 (sd: 59,695) individuals per day in 2015 (see Figure 4.1). Furthermore, in 2015, we observed between 4.62 and 7.57 times more birds per day in the region around Manaus than in the Great Lakes (see Figure 4.1). Lastly, with regards to the distribution of roost sizes, we found that Manaus had a much larger average roost size than the Great Lakes region. In 2014, the average roost size in Manaus was approximately 52,578 birds (SD: 76,331) whereas in 2015 that average was 172,416 (SD: 152,508, see Figure 4.4b). In the Great Lakes region, on the other hand, the average roost size was 5,440 individuals (SD: 5,364) in 2014, and 6,938 individuals (SD: 7,686) in 2015 (see Figure 4.4a).

4.4 Discussion

We used two years of data collected by a weather radar in Manaus to characterize for the first time the phenology of swallow and martin roosting activity in the Amazon Rainforest. We found that swallows and martins roost in the region around Manaus throughout the year, with a peak in activity between January and May. This peak corresponds to the annual formation of the previously documented Negro River roost, which still occurs in the same ephemeral island. The daily relative abundance pattern observed in both years, however, was distinct: in 2015, our study region hosted substantially more birds than in 2014. We found that, on average, roosts gather more birds in Manaus than they do in the Great Lakes region, and that the region around Manaus hosted more birds per day than the entire Great Lakes region, which is more than eight times larger. Lastly, we observed that the roost changed location throughout the year, moving from the Anavilhanas archipelago to Comaru Island just before the peak of the Purple Martin season and returning to Anavilhanas in May, in 2015, and in July, in 2014.

Our data confirms previous reports that communal roosting activity occurs in the Amazon throughout the year. In Figure E.6, we summarize the number of records per month and per species known to join roosts in the Amazon, based on data available in the Global Biodiversity Network Facility and in the Brazilian community science platform Wikiaves. Current accounts of roosting activity in Manaus suggest that resident Brown-chested Martins and Barn Swallows join Purple Martins in the Negro River roost from January to May, though in much smaller numbers (Arruda, 2023). Furthermore, historical reports of swallow and martin roosts in the region around Manaus indicate that resident Grey-breasted Martin (*Progne chalybea*) may also join these aggregations early in the year (Sick, 2001). Since the Amazon receives Austral migrants such as the Southern Martin and the Brown-chested Martin fusca subspecies between July and September, it is possible that these two species are driving roost activity during that period. Lastly, we note that Purple Martins are seen in low numbers in the Amazon even between June and July (see Figure E.6), and on one occasion this species was photographed sharing a roost with Southern Martins in early September (Miranda Júnior & Gomes, 2023). Thus, we suggest that communal roosts could act as a link between two migration systems or possibly three, if we find that intertropical migrants such as the Caribbean Martin also join these aggregations when they spend time in the Amazon.

The surge of activity during the Purple Martin season in 2015 was unexpected. More data is needed to establish if the activity pattern observed in 2015 is still observed today, if it happened just once, or if it happens every few years. We highlight that 2015 observed a record-breaking El Niño Southern Oscillation event, which is associated with droughts and warmer temperatures in the Amazon (Jiménez-Muñoz et al., 2016). However, the mechanisms by which El Niño events could influence roosting activity are unclear. These events typically create optimal conditions for the establishment of fires in dryland forests (Holmgren et al., 2001), nevertheless the peak of the wildfire season would be in August and September, after the observed surge of roosting activity. Interestingly, we remark that this specific occurrence of El Niño was marked by temperature anomalies larger than 1 °C as early as September of 2014, approximately when our counts

start to increase (Jiménez-Muñoz et al., 2016). Regardless of whether the extreme El Niño event was the cause of the observed pattern, the considerable number of Purple Martins using the Manaus region to roost from January to May in 2015 should be taken into account when planning human interventions in this area. Nonetheless, our study indicates that the areas used by swallows and martins to roost are well-protected, except for the region around Lago do Rei and the areas hosting satellite roosts.

The observed difference between roost size and daily number of birds in the Great Lakes region and in the Amazon Rainforest prompts further investigation. Roosting activity in these two regions occur at different moments of the annual cycle of swallows and martins and thus could be driven by different social and environmental factors. In the Great Lakes region, roosting occurs for a much shorter period, after swallows and martins finish breeding and before they start their southwards migration. In the Amazon, in contrast, roosting activity seems to occur throughout the nonbreeding season, when birds are finishing molting (Niles, 1972) and preparing to travel north as soon as possible to compete for better breeding opportunities (Stutchbury et al., 2016). Furthermore, since the Great Lakes region lies in the northern breeding range of most swallows and martins, roosts in that region likely congregate only the portion of the total population that breeds in the northern United States and Canada, whereas the Amazon is estimated to concentrate approximately 80% of the eastern subspecies of Purple Martins (Fraser et al., 2012).

Explaining the change in roost location within a season from the Anavilhanas National Park to the Comaru Island requires that we answer two questions: (1) why did the roost move away from Anavilhanas in February, and (2) why does it move back between May and July? Both locations are within the same black water river, the Negro River, where the water level can vary from approximately 3 m in the dry season to 15 m in the wet season (see river level and precipitation plots in the supplements). The area of Comaru Island during the wet season is greatly reduced, which could explain why the birds abandon it, but the difference of almost two months between the dates when birds moved up the river in both years can't be explained with current data (see

Figures E.7b and E.7a for true color renderings of Sentinel data demonstrating the change in surface area of Comaru island). In addition, studies are needed to understand the movement from Anavilhanas to Comaru Island. This movement occurred in mid-February, three months after the river levels started rising (see water level and precipitation plots in the supplements). It is possible that the influx of Purple Martins, paired with roosting substrate reduction due to the rising water levels, could prompt birds to move as a group to the new roosting location, but more data is needed to test this hypothesis.

Lastly, we remark that the Amazon River experiences interannual cyclical flooding and drought events, which are influenced by factors such as El Niño and La Niña climatic phenomena, as well as moisture transport from the tropical Atlantic Ocean, causing variations in the region's water cycle (Marengo & Espinoza, 2015; Sikora de Souza et al., 2020). Recent studies indicate that the intensity and frequency of extreme hydrological events in the Amazon River Basin are increasing, with both droughts and floods becoming more prolonged over time (Correa et al., 2022; Espinoza et al., 2022). These extended periods of extreme conditions not only disrupt the region's water cycle but also intensify ecological and socio-economic impacts on local communities and ecosystems. Severe droughts can cause significant environmental disruptions, hindering efforts to accurately assess floodplain and river channel topography, which is a key information for understanding habitats shaped by the flood pulse and their ecological dynamics (Fassoni-Andrade et al., 2021). Extreme climatic events, such as prolonged droughts, can lead to a disruption in the ecosystem balance within the Amazon forest, potentially undermining the long-term stability and functioning of these vital ecosystems (De Faria et al, 2020). In this region, there are nine operational S-band weather radars that could offer us a cost-effective way to monitor swallow and martin roosting activity — and bird migration more generally — at unprecedented spatial scales. Access to these data could help us understand how migratory birds are responding to these severe threats and what can be done to mitigate this damage.

4.5 Figures

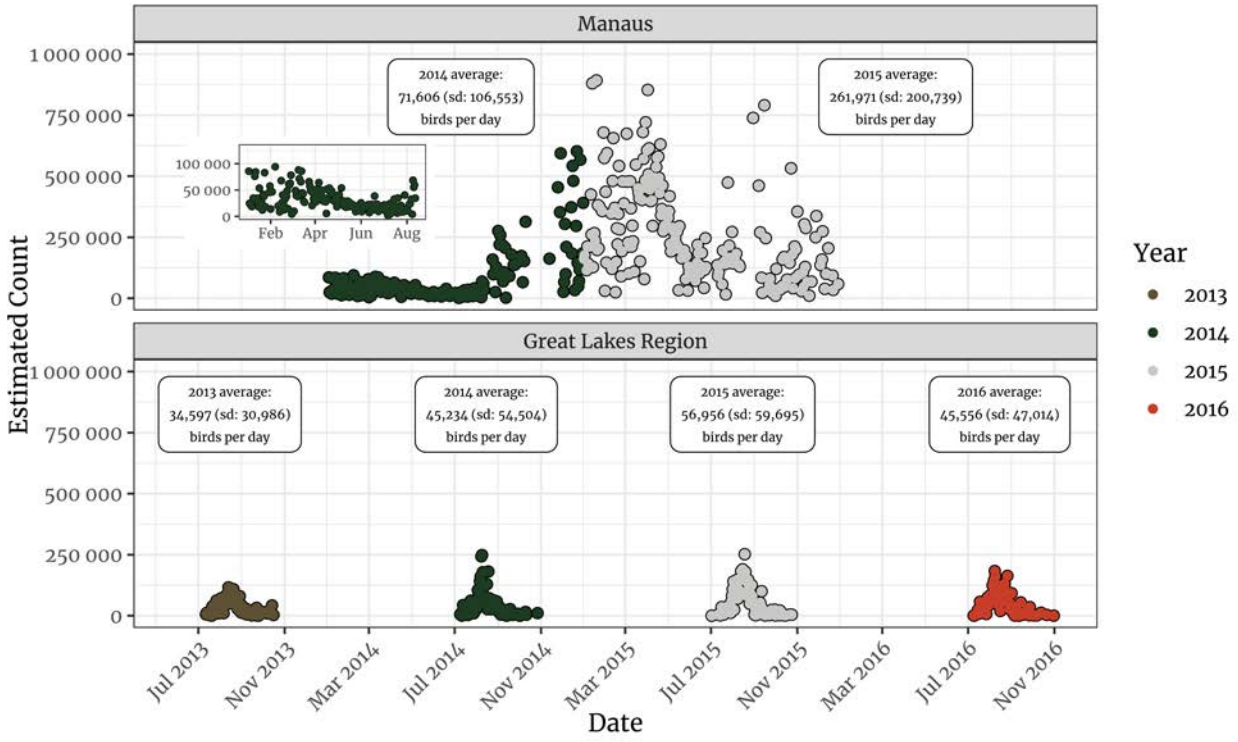


Figure 4.1: Plots of the daily counts observed in Manaus (upper) and in the Great Lakes region (lower). Colors represent different years. Sampling in the Great Lakes occurred from 2013 to 2016, during the period between June 1st to October 31st, whereas in Manaus we examined the entire two years of data.

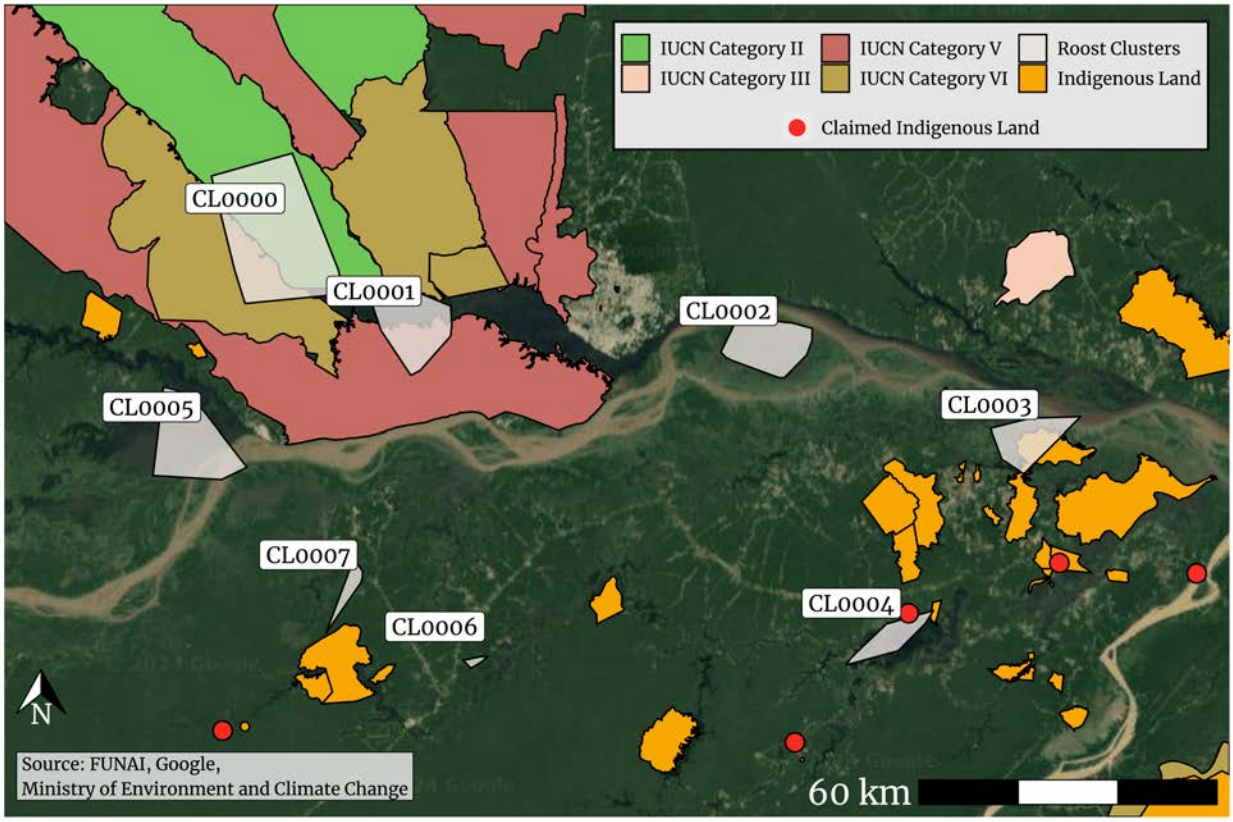
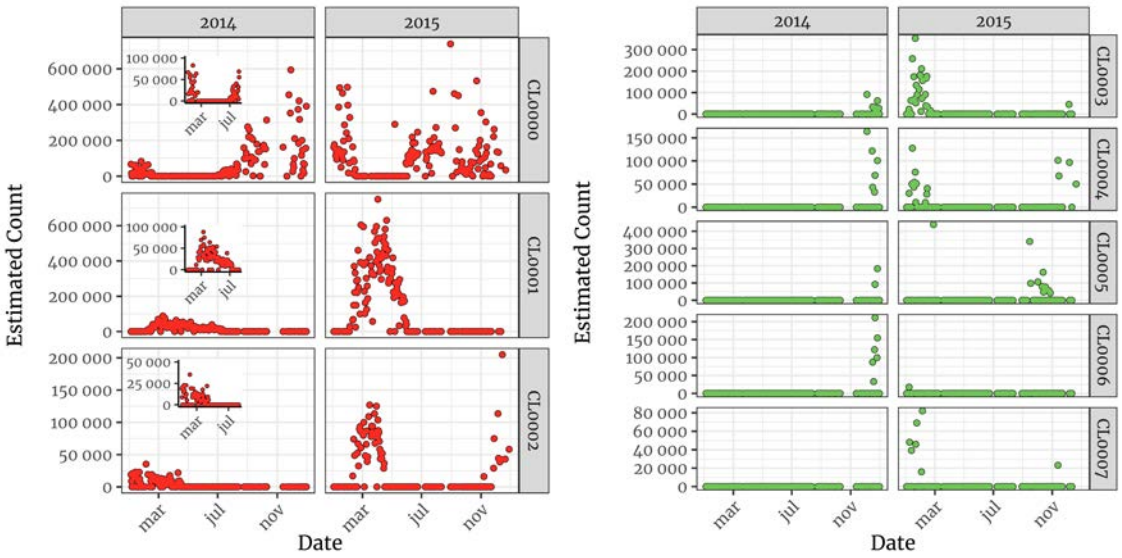


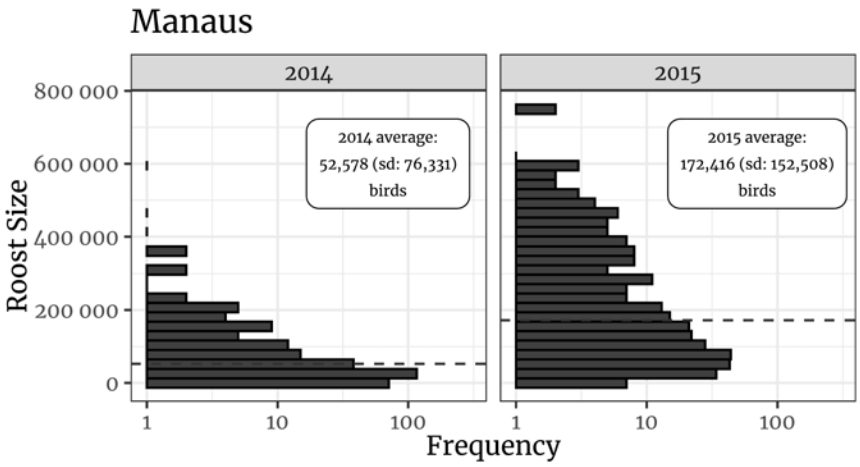
Figure 4.2: Map of the roost clusters found in Manaus (light grey). Polygons represent either protected land (colored based on IUCN Category) orP indigenous land (yellow). Red points represent regions claimed by the indigenous communities of ethnicity Mura since 1997. IUCN Categories: II - national park, III - natural monument or feature, V - protected landscape, and VI - protected area with sustainable use of natural resources.



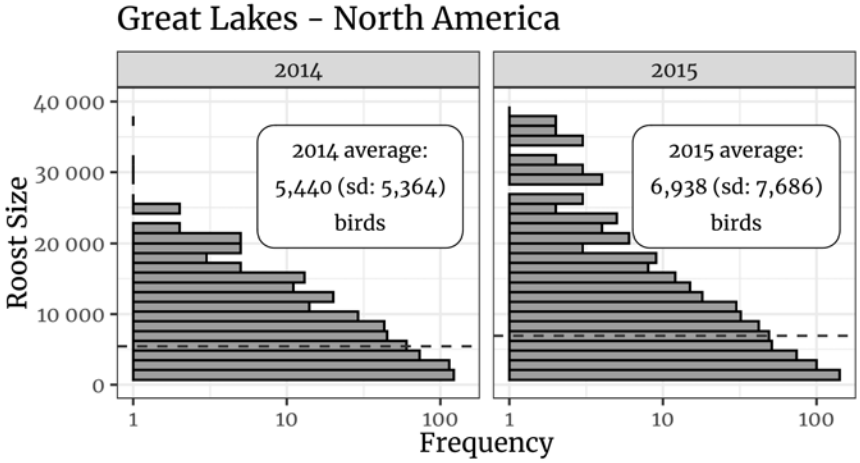
(a) Stable roosts.

(b) Satellite roosts.

Figure 4.3: Daily counts of birds in stable (a) and satellite (b) roosts in the region around Manaus. Insets represent counts from January 1st to August 10.



(a) Roost size distribution in Manaus in 2014 and 2015.



(b) Roost size distribution in the Great Lakes region in 2014 and 2015

Figure 4.4: Plots representing the roost size distribution in the Great Lakes (a) and in Manaus (b). The distribution of roost sizes is heavy-tailed. The x axis on both plots is in logarithmic scale.

Bibliography

Arruda, E. F. (2023). *Dinâmica Temporal e Espacial em um Megadormitório de Andorinha-Azul (Progne subis), uma Ave Migratória Invernando na Amazônia* [Master's thesis, Instituto Nacional de Pesquisas da Amazônia]. https://repositorio.inpa.gov.br/bitstream/1/39147/1/DISSERTA%c3%87%c3%83O%20-%20ELIENE%20FONTES%20ARRUDA.pdf

Bauer, S., Shamoun-Baranes, J., Nilsson, C., Farnsworth, A., Kelly, J. F., Reynolds, D. R., Dokter, A. M., Krauel, J. F., Petterson, L. B., Horton, K. G., & Chapman, J. W. (2018). The grand challenges of migration ecology that radar aeroecology can help answer. *Ecography*, *42*(5), 861–875. https://doi.org/10.1111/ecog.04083

Bauer, S., Tielens, E. K., & Haest, B. (2024). Monitoring aerial insect biodiversity: a radar perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1904). https://doi.org/10.1098/rstb.2023.0113

Belotti, M., Zhao, W., Deng, Y., Cheng, Z., Perez, G., Simons, V., tielens elske, e., Maji, S., Sheldon, D., Kelly, J., & Horton, K. (2024). Swallow and martin roosts detected on WSR in the Great Lakes region from 2000 to 2020. https://doi.org/10.6084/m9.figshare.20137961

Belotti, M. C. T. D., Deng, Y., Zhao, W., Simons, V. F., Cheng, Z., Perez, G., Tielens, E., Maji, S., Sheldon, D., Kelly, J. F., & Horton, K. G. (2023). Long-term analysis of persistence and size of swallow and martin roosts in the US Great Lakes (V. Lecours & M. Van Den Broeke, Eds.). *Remote Sensing in Ecology and Conservation*, *9*(4), 469–482. https://doi.org/10.1002/rse2.323

Bowlin, M. S., Bisson, I.-A., Shamoun-Baranes, J., Reichard, J. D., Sapir, N., Marra, P. P., Kunz, T. H., Wilcove, D. S., Hedenstrom, A., Guglielmo, C. G., Akesson, S., Ramenofsky, M., & Wikelski, M. (2010). Grand Challenges in Migration Biology. *Integrative and Comparative Biology*, *50*(3), 261–279. https://doi.org/10.1093/icb/icq013

Bridge, E. S., Pletschet, S. M., Fagin, T., Chilson, P. B., Horton, K. G., Broadfoot, K. R., & Kelly, J. F. (2015). Persistence and habitat associations of Purple Martin roosts quantified via

weather surveillance radar. *Landscape Ecology*, *31*(1), 43–53. https://doi.org/10.1007/s10980-015-0279-0

Chesser, R. T. (1994). Migration in South America: an overview of the austral system. *Bird Conservation International*, *4*(2-3), 91–107. https://doi.org/10.1017/s0959270900002690

Chilson, P. B., Frick, W. F., Stepanian, P. M., Shipley, J. R., Kunz, T. H., & Kelly, J. F. (2012). Estimating animal densities in the aerosphere using weather radar: ToZor not toZ? *Ecosphere*, *3*(8), art72. https://doi.org/10.1890/es12-00027.1

Cooper, J. C. (2024). Ecological niche divergence or ecological niche partitioning in a widespread Neotropical bird lineage. *PeerJ*, *12*, e17345. https://doi.org/10.7717/peerj.17345

Correa, S. B., van der Sleen, P., Siddiqui, S. F., Bogotá-Gregory, J. D., Arantes, C. C., Barnett, A. A., Couto, T. B. A., Goulding, M., & Anderson, E. P. (2022). Biotic Indicators for Ecological State Change in Amazonian Floodplains. *BioScience*, *72*(8), 753–768. https://doi.org/10.1093/biosci/biac038

Cristol, D., Baker, M., & Carbone, C. (1999). Differential migration revisted: Latitudunal segergation by age and sex class. *Current Ornithology*, *15*, 33–88.

Deng, Y., Belotti, M. C. T. D., Zhao, W., Cheng, Z., Perez, G., Tielens, E., Simons, V. F., Sheldon, D. R., Maji, S., Kelly, J. F., & Horton, K. G. (2023). Quantifying long-term phenological patterns of aerial insectivores roosting in the Great Lakes region using weather surveillance radar. *Global Change Biology*, *29*(5), 1407–1419. https://doi.org/10.1111/gcb.16509

Dunning, J. (2008). CRC Handbook of Avian body masses. *CRC handbook of avian body masses, 2nd ed*, *384*. https://doi.org/10.1201/9781420064452

Espinoza, J.-C., Marengo, J. A., Schongart, J., & Jimenez, J. C. (2022). The new historical flood of 2021 in the Amazon River compared to major floods of the 21st century: Atmospheric features in the context of the intensification of floods. *Weather and Climate Extremes*, *35*, 100406. https://doi.org/10.1016/j.wace.2021.100406

Faaborg, J., Holmes, R. T., Anders, A. D., Bildstein, K. L., Dugger, K. M., Gauthreaux, S. A., Heglund, P., Hobson, K. A., Jahn, A. E., Johnson, D. H., Latta, S. C., Levey, D. J., Marra, P. P., Merkord,

C. L., Nol, E., Rothstein, S. I., Sherry, T. W., Sillett, T. S., Thompson, F. R., & Warnock, N. (2010a). Conserving migratory land birds in the New World: Do we know enough? *Ecological Applications*, *20*(2), 398–418. Retrieved February 19, 2022, from http://www.jstor.org/stable/27797816

Faaborg, J., Holmes, R. T., Anders, A. D., Bildstein, K. L., Dugger, K. M., Gauthreaux, S. A., Heglund, P., Hobson, K. A., Jahn, A. E., Johnson, D. H., Latta, S. C., Levey, D. J., Marra, P. P., Merkord, C. L., Nol, E., Rothstein, S. I., Sherry, T. W., Sillett, T. S., Thompson, F. R., & Warnock, N. (2010b). Recent advances in understanding migration systems of New World land birds. *Ecological Monographs*, *80*(1), 3–48. https://doi.org/10.1890/09-0395.1

Fassoni-Andrade, A. C., Fleischmann, A. S., Papa, F., Paiva, R. C. D. d., Wongchuig, S., Melack, J. M., Moreira, A. A., Paris, A., Ruhoff, A., Barbosa, C., Maciel, D. A., Novo, E., Durand, F., Frappart, F., Aires, F., Abrahão, G. M., Ferreira-Ferreira, J., Espinoza, J. C., Laipelt, L., … Pellet, V. (2021). Amazon Hydrology From Space: Scientific Advances and Future Challenges. *Reviews of Geophysics*, *59*(4). https://doi.org/10.1029/2020rg000728

Fraser, K. C., Stutchbury, B. J. M., Silverio, C., Kramer, P. M., Barrow, J., Newstead, D., Mickle, N., Cousens, B. F., Lee, J. C., Morrison, D. M., Shaheen, T., Mammenga, P., Applegate, K., & Tautin, J. (2012). Continent-wide tracking to determine migratory connectivity and tropical habitat associations of a declining aerial insectivore. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1749), 4901–4906. https://doi.org/10.1098/rspb.2012.2207

Frick, W. F., Stepanian, P. M., Kelly, J. F., Howard, K. W., Kuster, C. M., Kunz, T. H., & Chilson, P. B. (2012). Climate and Weather Impact Timing of Emergence of Bats (B. Fenton, Ed.). *PLoS ONE*, *7*(8), e42737. https://doi.org/10.1371/journal.pone.0042737

Funai. (2024). Terra Indígenas no Brasil. Retrieved August 29, 2024, from https://www.gov.br/funai/pt-br/atuacao/terras-indigenas/geoprocessamento-e-mapas

García-Lau, I., Bani Assadi, S., Kent, G., González, A., Rodríguez-Ochoa, A., Jiménez, A., Acosta, M., Mugica, L., & Meyer, K. (2021). Tracking Cuban Martin (Progne cryptoleuca) migra-

tion to wintering location and back using geolocators: solving a mystery. *Ornithology Research*, *29*(2), 106–112. https://doi.org/10.1007/s43388-021-00057-y

Gomes, A. M. (2023). *Aves do Estado do Amazonas, Brasil: Diversidade e Padrões de Distribuição* [Master's thesis, Instituto Nacional de Pesquisas da Amazônia]. https://repositorio.inpa.gov.br/bitstream/1/39591/1/Disserta%c3%a7%c3%a3o%20Completa%20-%20Aves%20do%20Amazonas_vers%c3%a3o%20final.pdf

Harrison, X. A., Blount, J. D., Inger, R., Norris, D. R., & Bearhop, S. (2011). Carry-over effects as drivers of fitness differences in animals. *Journal of Animal Ecology*, *80*(1), 4–18. https://doi.org/10.1111/j.1365-2656.2010.01740.x

Heffernan, J. B., Soranno, P. A., Angilletta Jr, M. J., Buckley, L. B., Gruner, D. S., Keitt, T. H., Kellner, J. R., Kominoski, J. S., Rocha, A. V., Xiao, J., Harms, T. K., Goring, S. J., Koenig, L. E., McDowell, W. H., Powell, H., Richardson, A. D., Stow, C. A., Vargas, R., & Weathers, K. C. (2014). Macrosystems ecology: understanding ecological patterns and processes at continental scales. *Frontiers in Ecology and the Environment*, *12*(1), 5–14. https://doi.org/10.1890/130017

Helmus, J. J., & Collis, S. M. (2016). The Python ARM Radar Toolkit (Py-ART), a Library for Working with Weather Radar Data in the Python Programming Language. *Journal of Open Research Software*, *4*(1), 25. https://doi.org/10.5334/jors.119

Holmgren, M., Scheffer, M., Ezcurra, E., Gutiérrez, J. R., & Mohren, G. M. (2001). El Niño effects on the dynamics of terrestrial ecosystems. *Trends in Ecology & Evolution*, *16*(2), 89–94.

Horn, J. W., & Kunz, T. H. (2008). Analyzing NEXRAD Doppler Radar Images to Assess Nightly Dispersal Patterns and Population Trends in Brazilian Free-Tailed Bats (Tadarida brasiliensis). *Integrative and Comparative Biology*, *48*(1), 24–39. Retrieved March 2, 2022, from http://www.jstor.org/stable/25160117

Horton, K. G., Doren, B. M. V., Sorte, F. A. L., Cohen, E. B., Clipp, H. L., Buler, J. J., Fink, D., Kelly, J. F., & Farnsworth, A. (2019). Holding steady: Little change in intensity or timing of

bird migration over the Gulf of Mexico. *Global Change Biology*, *25*(3), 1106–1118. https://doi.org/10.1111/gcb.14540

Imlay, T. L., Mills Flemming, J., Saldanha, S., Wheelwright, N. T., & Leonard, M. L. (2018). Breeding phenology and performance for four swallows over 57 years: relationships with temperature and precipitation. *Ecosphere*, *9*(4), e02166. https://doi.org/10.1002/ecs2.2166

Jahn, A. E., Levey, D. J., & Smith, K. G. (2004). Reflections Across Hemispheres: A System-Wide Approach to New World Bird Migration. *The Auk*, *121*(4), 1005–1013. https://doi.org/10.1093/auk/121.4.1005

Jiménez-Muñoz, J. C., Mattar, C., Barichivich, J., Santamaría-Artigas, A., Takahashi, K., Malhi, Y., Sobrino, J. A., & Schrier, G. v. d. (2016). Record-breaking warming and extreme drought in the Amazon rainforest during the course of El Niño 2015–2016. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep33130

Junk, W., Bayley, P., & Sparks, R. (1989). The Flood Pulse Concept in River-Floodplain Systems. *Can. Spec. Public Fish. Aquat. Sci.*, *106*.

Junk, W. J. (1997). General Aspects of Floodplain Ecology with Special Reference to Amazonian Floodplains. In *The Central Amazon Floodplain: Ecology of a Pulsing System* (1st ed., pp. 3–22, Vol. 126). Springer.

Junk, W. J., Piedade, M. T. F., Schöngart, J., Cohn-Haft, M., Adeney, J. M., & Wittmann, F. (2011). A Classification of Major Naturally-Occurring Amazonian Lowland Wetlands. *Wetlands*, *31*(4), 623–640. https://doi.org/10.1007/s13157-011-0190-7

Junk, W. J., Piedade, M. T. F., Schöngart, J., & Wittmann, F. (2012). A classification of major natural habitats of Amazonian white-water river floodplains (várzeas). *Wetlands Ecology and Management*, *20*(6), 461–475. https://doi.org/10.1007/s11273-012-9268-0

Junk, W. J., Wittmann, F., Schöngart, J., & Piedade, M. T. F. (2015). A classification of the major habitats of Amazonian black-water river floodplains and a comparison with their white-water counterparts. *Wetlands Ecology and Management*, *23*(4), 677–693. https://doi.org/10.1007/s11273-015-9412-8

Kelly, J. F., & Horton, K. G. (2016). Toward a predictive macrosystems framework for migration ecology. *Global Ecology and Biogeography*, *25*(10), 1159–1165. https://doi.org/10.1111/geb.12473

Kelly, J. F., & Pletschet, S. M. (2017). Accuracy of swallow roost locations assigned using weather surveillance radar (M. Rowcliffe & D. Boyd, Eds.). *Remote Sensing in Ecology and Conservation*, *4*(2), 166–172. https://doi.org/10.1002/rse2.66

Kelly, J. F., Shipley, J. R., Chilson, P. B., Howard, K. W., Frick, W. F., & Kunz, T. H. (2012). Quantifying animal phenology in the aerosphere at a continental scale using NEXRAD weather radars. *Ecosphere*, *3*(2), art16. https://doi.org/10.1890/es11-00257.1

Laughlin, A. J., Sheldon, D. R., Winkler, D. W., & Taylor, C. M. (2016). Quantifying non-breeding season occupancy patterns and the timing and drivers of autumn migration for a migratory songbird using Doppler radar. *Ecography*, *39*(10), 1017–1024. https://doi.org/10.1111/ecog.01988

Lees, A. C., Rosenberg, K. V., Ruiz-Gutierrez, V., Marsden, S., Schulenberg, T. S., & Rodewald, A. D. (2020). A roadmap to identifying and filling shortfalls in Neotropical ornithology. *The Auk*, *137*(4). https://doi.org/10.1093/auk/ukaa048

Malhado, A. C. M., de Azevedo, R. S. D., Todd, P. A., Santos, A. M. C., Fabré, N. N., Batista, V. S., Aguiar, L. J. G., & Ladle, R. J. (2014). Geographic and Temporal Trends in Amazonian Knowledge Production. *Biotropica*, *46*(1), 6–13. https://doi.org/https://doi.org/10.1111/btp.12079

Marengo, J. A., & Espinoza, J. C. (2015). Extreme seasonal droughts and floods in Amazonia: causes, trends and impacts. *International Journal of Climatology*, *36*(3), 1033–1050. https://doi.org/10.1002/joc.4420

Marra, P. P., Cohen, E. B., Loss, S. R., Rutter, J. E., & Tonra, C. M. (2015). A call for full annual cycle research in animal ecology. *Biology Letters*, *11*(8), 20150552. https://doi.org/10.1098/rsbl.2015.0552

Miranda Júnior, I., & Gomes, A. (2023). eBird Checklist: S149256060 - eBird: An online database of bird distribution and abundance [web application]. https://ebird.org/ebird/view/checklist/S36954356

Newton, I. (2004). Population limitation in migrants. *Ibis*, *146*(2), 197–226. https://doi.org/10.1111/j.1474-919X.2004.00293.x

Niles, D. M. (1972). Molt Cycles of Purple Martins (Progne subis). *The Condor*, *74*(1), 61. https://doi.org/10.2307/1366450

Oliveira, J. A. d. (2021). *Contribuição de Nutrientes e Implicações Tróficas da Deposição de Sazonal de Fezes de Andorinha-Azul (Progne subis) em um Megadormitório no Rio Negro, Amazônia Central* [Master's thesis, Instituto Nacional de Pesquisas da Amazônia].

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., & Kassem, K. R. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, *51*(11), 933. https://doi.org/10.1641/0006-3568(2001)051[0933:teotwa]2.0.co;2

Perez, T. M., & Hogan, J. A. (2018). The changing nature of collaboration in tropical ecology and conservation. *Biotropica*, *50*(4), 563–567. https://doi.org/https://doi.org/10.1111/btp.12573

Perlut, N. G., Klak, T. C., & Rakhimberdiev, E. (2017). Geolocator Data Reveal the Migration Route and Wintering Location of a Caribbean Martin (Progne dominicensis). *The Wilson Journal of Ornithology, 129*(3), 605–610. https://doi.org/10.1676/16-142.1

Peterman, P. (1997). Birds. In W. J. Junk (Ed.), *The Central Amazonian Floodplain: Ecology of a Pulsing System* (First Edition, pp. 419–451). Springer.

Project, M. (2021). Mapbiomas Project: Annual Series of Land Use and Land Cover Maps of Brazil (V6.0.0). Retrieved October 1, 2022, from https://plataforma.brasil.mapbiomas.org/

Ricardo, F., Klein, T., & dos Santos, T. M. (2023). *Povos Indígenas do Brasil 2017/2022* (2ed). Instituto Socioambiental.

Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., Stanton, J. C., Panjabi, A., Helft, L., Parr, M., & Marra, P. P. (2019). Decline of the North American avifauna. *Science*, *366*(6461), 120–124. https://doi.org/10.1126/science.aaw1313

Russell, K. R., Mizrahi, D. S., & Gauthreaux, S. A. (1998). Large-Scale Mapping of Purple Martin Pre-Migratory Roosts Using WSR-88D Weather Surveillance Radar. *Journal of Field Ornithology*, *69*(2), 316–325. http://www.jstor.org/stable/4514321

Russell, K. R. (1996). *Spatial and temporal patterns of a Purple Martin (Progne subis) roost: a radar and direct visual study* [Master's thesis, Clemson University].

Salo, J., Kalliola, R., Häkkinen, I., Mäkinen, Y., Niemelä, P., Puhakka, M., & Coley, P. D. (1986). River dynamics and the diversity of Amazon lowland forest. *Nature*, *322*(6076), 254–258. https://doi.org/10.1038/322254a0

Santos, C. O., Branco, J. M., Belotti, M. C. T. D., Abilleira, P., Siegrist, J., Fischer, J., Lima, L. M., Cohn-Haft, M., & Hingst-Zaher, E. (2021). Distribution and migration phenology of Purple Martins (Progne subis) in Brazil. *Ornithology Research*, *29*(4), 213–222. https://doi.org/10.1007/s43388-021-00071-0

Sherry, T. W., & Holmes, R. T. (1996). Winter Habitat Quality, Population Limitation, and Conservation of Neotropical-Nearctic Migrant Birds. *Ecology*, *77*(1), 36–48. Retrieved May 23, 2022, from https://esajournals-onlinelibrary-wiley-com.ezproxy2.library.colostate.edu/doi/abs/10.2307/2265652

Sick, H. (2001). *Ornitologia Brasileira* (3ª Reimpressão). Editora Nova Fronteira.

Sikora de Souza, V. A., Moreira, D. M., Rotunno Filho, O. C., & Rudke, A. P. (2020). Extreme rainfall events in Amazonia: The Madeira river basin. *Remote Sensing Applications: Society and Environment*, *18*, 100316. https://doi.org/10.1016/j.rsase.2020.100316

Steeves, T. K., Kearney-McGee, S. B., Rubega, M. A., Cink, C. L., & Collins, C. T. (2020, March). Chimney Swift (Chaetura pelagica) (A. F. Poole, Ed.). https://doi.org/10.2173/bow.chiswi.01

Stepanian, P. M., & Wainwright, C. E. (2018). Ongoing changes in migration phenology and winter residency at Bracken Bat Cave. *Global Change Biology*, *24*(7), 3266–3275. https://doi.org/10.1111/gcb.14051

Stutchbury, B. J., Fraser, K. C., Silverio, C., Kramer, P., Aeppli, B., Mickle, N., Pearman, M., Savage, A., & Mejeur, J. (2016). Tracking mated pairs in a long-distance migratory songbird: migration schedules are not synchronized within pairs. *Animal Behaviour*, *114*, 63–68. https://doi.org/10.1016/j.anbehav.2016.01.016

Tautin, J., Cousens, B. F., & Kostka, K. (2017). Addressing regional declines in Purple Martin populations. *Proceedings of the Fourth International Partners in Flight Conference: Tundra to Tropics*.

Tielens, E. K., Cimprich, P. M., Clark, B. A., DiPilla, A. M., Kelly, J. F., Mirkovic, D., Strand, A. I., Zhai, M., & Stepanian, P. M. (2021). Nocturnal city lighting elicits a macroscale response from an insect outbreak population. *Biology Letters*, *17*(3). https://doi.org/10.1098/rsbl.2020.0808

Tkachenko, M., Malyuk, M., Holmanyuk, A., & Liubimov, N. (2020). Label Studio: Data labeling software. https://github.com/heartexlabs/label-studio

Trigg, M. A., Bates, P. D., Wilson, M. D., Schumann, G., & Baugh, C. (2012). Floodplain channel morphology and networks of the middle Amazon River. *Water Resources Research*, *48*(10). https://doi.org/10.1029/2012wr011888

Wagner, C. S., Park, H. W., & Leydesdorff, L. (2015). The Continuing Growth of Global Cooperation Networks in Research: A Conundrum for National Governments (W. Glanzel, Ed.). *Plos One*, *10*(7), e0131816. https://doi.org/10.1371/journal.pone.0131816

Ward, P., & Zahavi, A. (1973). The important of certain assemblages of birds as "information centers" for food-finding. *Ibis*, *115*(4), 517–534. https://doi.org/https://doi.org/10.1111/j.1474-919X.1973.tb01990.x

Weatherhead, P. J. (1983). Two Principal Strategies in Avian Communal Roosts. *The American Naturalist*, *121*(2), 237–243. https://doi.org/10.1086/284053

Webster, M. S., Marra, P. P., Haig, S. M., Bensch, S., & Holmes, R. T. (2002). Links between worlds: unraveling migratory connectivity. *Trends in Ecology & Evolution*, *17*(2), 76–83. https://doi.org/https://doi.org/10.1016/S0169-5347(01)02380-1

Winkler, D. (2006). Roosts and migration of swallows. *El Hornero*, *21*(2), 85–97.

Wittmann, F., Householder, J. E., Piedade, M. T. F., Schöngart, J., Demarchi, L. O., Quaresma, A. C., & Junk, W. J. (2022). A Review of the Ecological and Biogeographic Differences of Amazonian Floodplain Forests. *Water*, *14*(21), 3360. https://doi.org/10.3390/w14213360

Wolfe, J., Carvalho Diniz, F., Anciaes, M., & Stouffer, P. (2012). New record of Connecticut Warbler (Oporornis agilis) in central Amazonian Brazil. *Cotinga*, 127–128.

Appendix A

Screening Protocol

In this appendix, we describe the screening protocol used to find the true positives amongst the tracks suggested by the machine learning pipeline. We examined renderings of reflectivity factor (dBZ) and radial velocity (m/s) of each scan at the lowest elevation angle available (between 0.3° and 0.5°). Bounding boxes surrounding the roosts were highlighted in each scan. Figures S1 and S2 define the general decision-making procedures followed by the screeners. The text that follows delineates each decision rule in more detail.

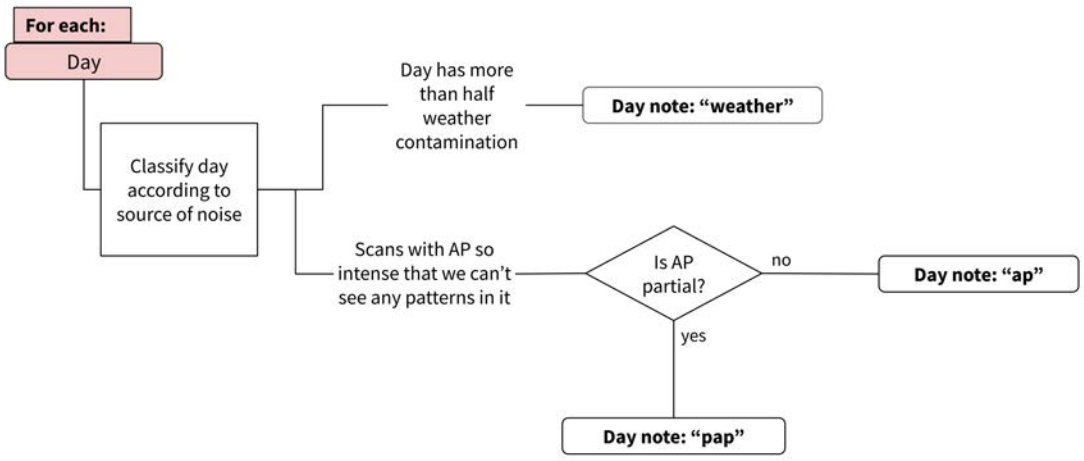


Figure A.1: Rules employed to classify days that needed to be discarded from the analysis due to partial or total non-biological contamination (weather or anomalous propagation).

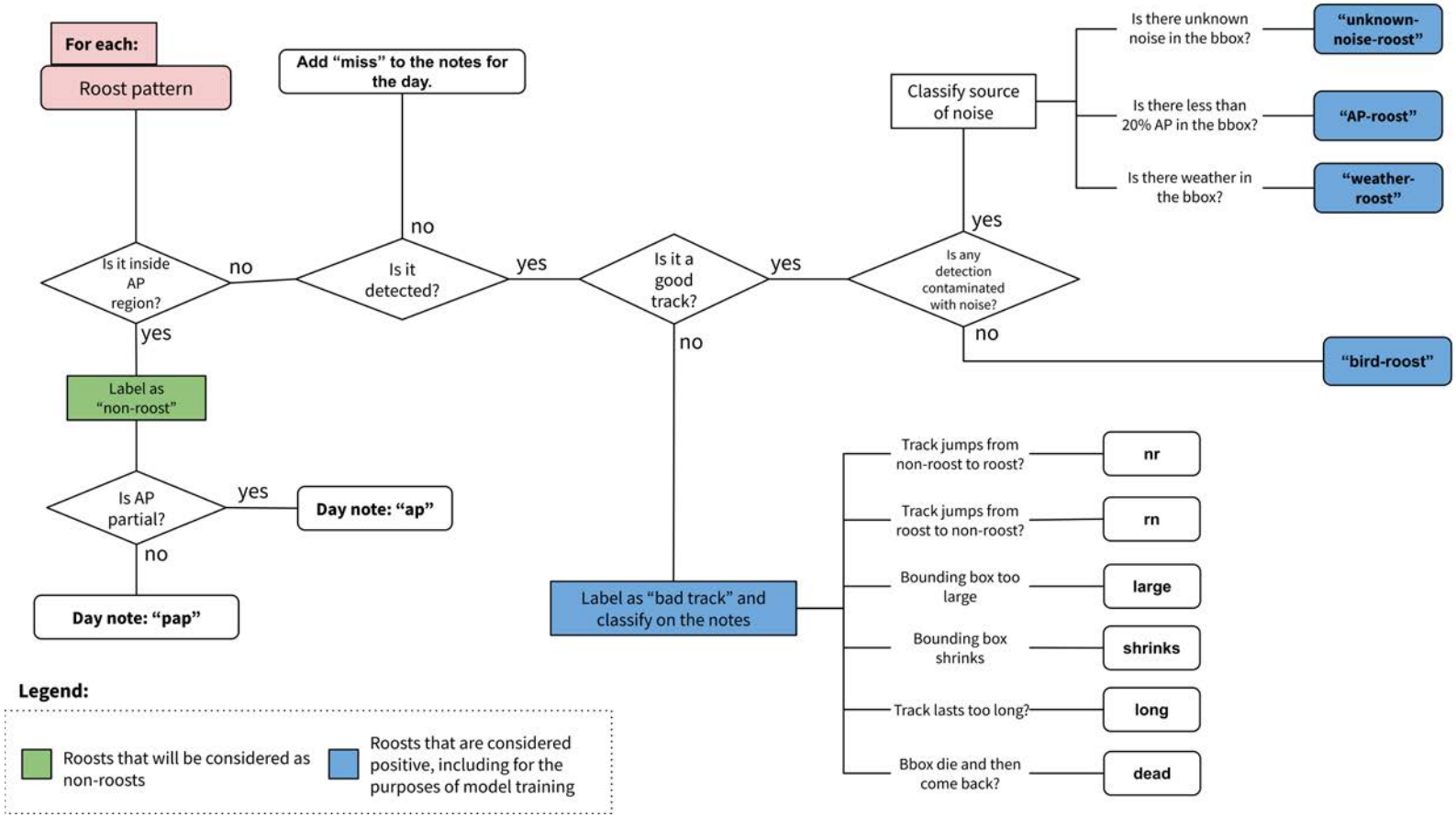


Figure A.2: Rules adopted to classify each roost pattern (high-reflectivity, donut-shaped dispersal event) based on the amount of contamination found inside the bounding box proposed by the machine learning pipeline.

A.1 General Remarks

General approach to decision-making: if a bounding box appears to have more noise (yellow and red areas that don't follow the spatiotemporal pattern typical of roosts) than birds, mark it using its specific noise label, or as "non-roost".

Rule 1.1. If the same roost is detected twice, keep the best track and mark the other one as "duplicate":

- The track where bounding boxes follow the roost more tightly and/or
- The track with more well-captured frames

Rule 1.2. Some roosts get missed by the machine learning pipeline. When this happens, add "miss" to the notes for that day.

Rule 1.3. Good tracks that are cut by the border of the image should be labeled as "bird-roost".

Rule 1.4. Overlapping roosts should be labeled as usual and considered individually for their quality.

A.2 Weather Contamination

This happens when weather phenomena (rain, hail, thunderstorms etc) get detected and will result in high bird counts that are not actually birds, OR when most of the image is contaminated by weather in a way such that if there were roosts, we would not see them.

Rule 2.1. If more than half of the image is contaminated by intense, homogenous areas of $Z > 20$ dBZ, and there are no clear roosts outside of the high-reflectivity region, change all detected roosts to "non-roosts" and add a note "weather" for this day.

Rule 2.2. If a good track has any bounding box that is partially contaminated by weather, we mark it as "weather-roost".

A.3 Anomalous Propagation

In "normal" air temperature conditions, we can roughly predict the path that an electromagnetic wave will make through different layers of the atmosphere, and with this we can measure the sampling volume, the height, and the location of the beam emitted by the radar. Sometimes, conditions are not normal (specifically, air temperature inversions or sharp decreases in moisture with height), and the propagation of the electromagnetic wave through the air is unexpected or unpredictable. In these situations, we see phenomena called ducting and super refraction. When either of those occur, we see an increase in ground clutter ranges, and sometimes the entire scan can be taken by high reflectivity patches.

If anomalous propagation occurs and no roosts are present, we can be confident that we are looking at a true negative. On the other hand, if roosts are inside the anomalous propagation region, we will need to discard these roosts because we don't know how to accurately measure its spatial position. In the latter case, we can't consider this a true negative, since we know the birds to be present.

Rule 3.1. If there is anomalous propagation (however much of it) AND we see a roost inside it:

Note: Intense anomalous propagation is characterized by patchy areas where Z > 20 dBZ. Sometimes it occurs in the entire scan, but it can also happen in the regions at closer range from the radar station, while the area farther away is clear. The radial velocity will be messy, with intercalating positive, zero and negative velocities.

1. Mark all roosts inside the AP region as non-roosts.

2. Roosts outside the AP region should be labeled normally, that is, following all the other rules (don't label them as "AP-roost").

3. If a small part (<20%) of a good roost is caught in the AP region, label it as "AP-roost".

4. If the anomalous propagation occurs in all 360 degrees of the scan, add "ap" to the notes for that day. If it occurs in less than 360 degrees, add "pap" to notes for that day.

Rule 3.2. If there is granular anomalous propagation, where we see scattered pixels of less than 5 dBZ without saturation of the region closer to the radar, we can label the roosts normally following the other rules. Before doing so, take a good look at the region close to the radar. It should not have more ground clutter than normal.

A.4 Unknown Noise

This happens when roosts are caught in the middle of areas that have high reflectivity (greater than 25 dBZ) for undetermined reasons (most commonly it is ground clutter, sun interference or other biological scatterers).

Rule 4.1. If a good track has any bounding box that is partially contaminated by unknown sources of noise, mark them as “unknown-noise-roost”.

Rule 4.2. We found a specific case where we see clusters of small to medium roosts that get poorly and inconsistently detected, most commonly close to the radar center, and usually embedded in unknown noise. If we see this, we can label the roosts in this region as “non-roost” and add “cluster” to the notes for that day.

A.5 Bad Tracks

Bad tracks result from issues with the detector. We need to label these so we can measure how much these influence the overall results. All the following cases should be labeled as “bad-track”.

Rule 5.1. Large bad tracks: if a bounding box becomes much larger (at least two times larger than the actual roost), label it as “bad-track” and add a note to the track with the word “large”.

Rule 5.2. Shrinking bad tracks: if the bounding box is going fine in a track and suddenly shrinks, becoming much smaller than the actual roost, label it as “bad-track” and add a note to the track with the word “shrinks”.

Rule 5.3. Long bad tracks: if a track has more frames than it should, meaning that the roost disappeared before the track actually ends, label it as “bad-track” and add a note to the track with the word “long”.

Rule 5.4. Jumping bad tracks. Beware: jumping tracks can easily be mistaken for new tracks.

Rule 5.4.1. Jumping bad tracks from roost to non-roost: if a track is following a roost just fine and suddenly starts tracking something else (usually noise or another roost), label it as “bad-track” and add a note to the track with the word “nr” (non-roost to roost abbreviation).

Rule 5.4.2. Jumping bad tracks from non-roost to roost: if a track is following some random scatter, and suddenly starts following an actual roost, label it as “bad-track” and add a note to the track with the word “rn” (roost to non-roost abbreviation).

Rule 5.5. Dead bad tracks: sometimes tracks reappear after they have already finished tracking a good roost. Once they reappear, they start tracking noise. Label them as “bad-track” and add a note to the track with the word “dead”.

Appendix B

Great Lakes data processing: handling overlaps and mean shift clustering

This appendix describes details of the data processing procedures used in the second chapter of this dissertation (Chapter 2). We divide it in four parts:

1) Grouping overlapping tracks

2) Mean shift clustering

3) Data Completeness Plots

4) Detection Biases Analyses

Our initial dataset contains all the roost detections that were screened and labeled as “swallow-roost”, “weather-roost”, “unknown-noise-roost”, “AP-roost”, “bad-track”, or “duplicate”. These detections represent valid roost locations, even though they can’t be used to estimate abundance due to noise contamination. Except for duplicates, all tracks were included in the two data processing steps described in this appendix.

B.1 Finding overlapping tracks

We used the heuristic procedure described below to find overlapping tracks.

for *day* ***in*** *dataset* **do**
    Filter detections on current day;
    Get bounding boxes around each detection;
    for *track_id* ***in*** *current day* **do**
        Get the union of bounding boxes of all the detections in the current track (*track buffers*)
    end
    Create empty graph where each node is a track id on the current day **for** *track_id* ***in*** *current day* **do**
        for *other track_id* ***in*** *current day* **do**
            Get area of the intersection between the two track buffers;
            Get area of the union of the two track buffers;
            Get ratio between intersection and union;
            if *ratio* > 0.2 **then**
                Create edge between two track_id nodes;
            end
        end
    end
    Create unique identifier to all the connected components of the final graph for this day
end

B.1.1 Analysis of Overlaps

There is a trade off between tolerating a certain overlap percentage and using a no-tolerance policy. It frequently happens that tracks are very close together but are still two different departure events. Because we wanted to capture these cases as two individual roosts, we used a threshold of 20% overlap tolerance. In the plots below, we analyze the percent overlap (area of the intersection between two tracks divided by the total area of their union) for overlapping tracks captured by the same radar station and for overlapping tracks found on different stations. Our goal was to minimize the number of same-station overlaps, while maximizing the number of overlaps captured from two different stations.

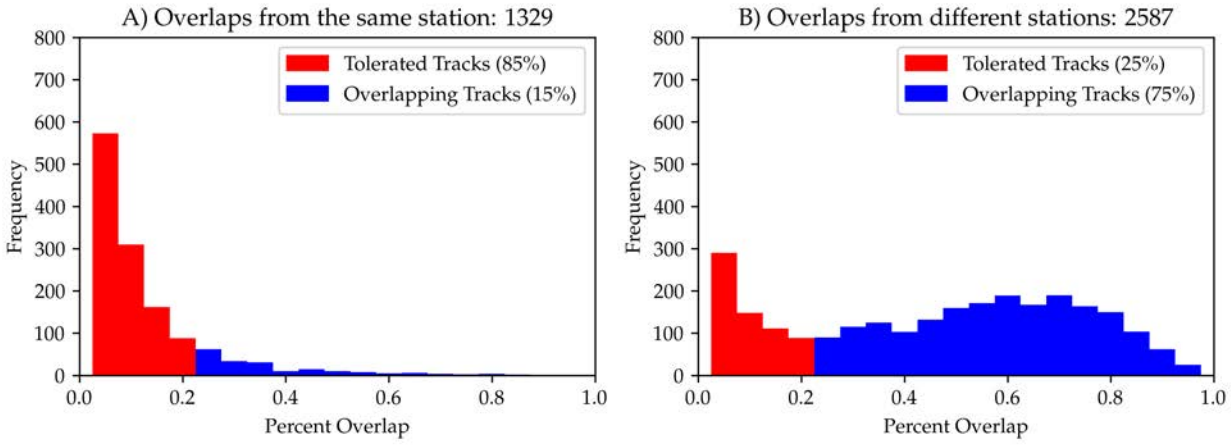


Figure B.1: Histograms of the percent overlaps found in our dataset for tracks occurring in the same station (A) and for tracks occurring in different stations (B). In red, we represent the tracks that were grouped because they overlapped by more than 20%. In blue, tracks that remained as individual occurrences.

B.2 Mean Shift Clustering

We used the mean shift clustering implementation available in the Python package `scikit-learn` (Pedregosa et al., 2011). Because the first detection of a roost is the most precise estimate of its location, we subset our dataset to the first detection of each group of tracks (tracks grouped to account for overlaps as described above). The goal of a clustering procedure is to find groups of observations that minimize intra-cluster variance while maximizing inter-cluster variance. This quality can be evaluated with multiple metrics, but we selected two implemented in `scikit-learn`: the Calinski-Harabasz index and the silhouette coefficient. We ran the mean shift clustering algorithm with varying bandwidth parameters to observe which value would maximize the silhouette coefficient and while maintaing a low Calinski-Harabasz index. The plots below show the results of this test.

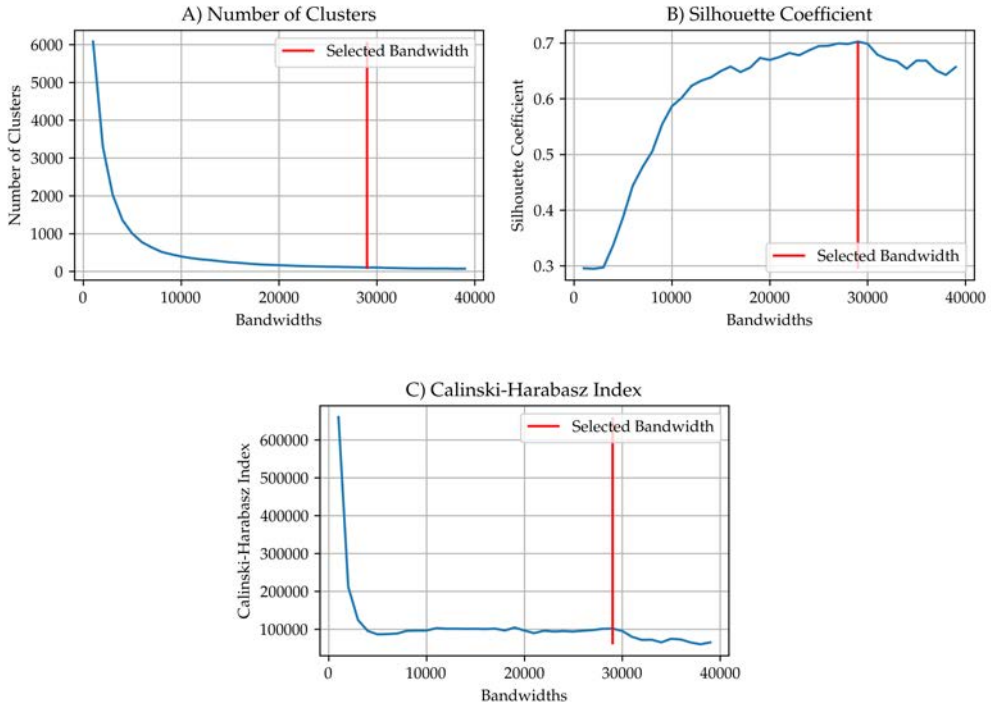


Figure B.2: Plots show the number of clusters, the silhouette coefficient and the Calinski-Harabasz index for the mean shift clustering algorithm with varying bandwidths. We selected the bandwidth that maximized the silhouette coefficient. This value is somewhat similar to the bandwidth that maximized the Calinski-Harabasz index after the initial drop expected for this index.

B.3 Data Completeness Plots

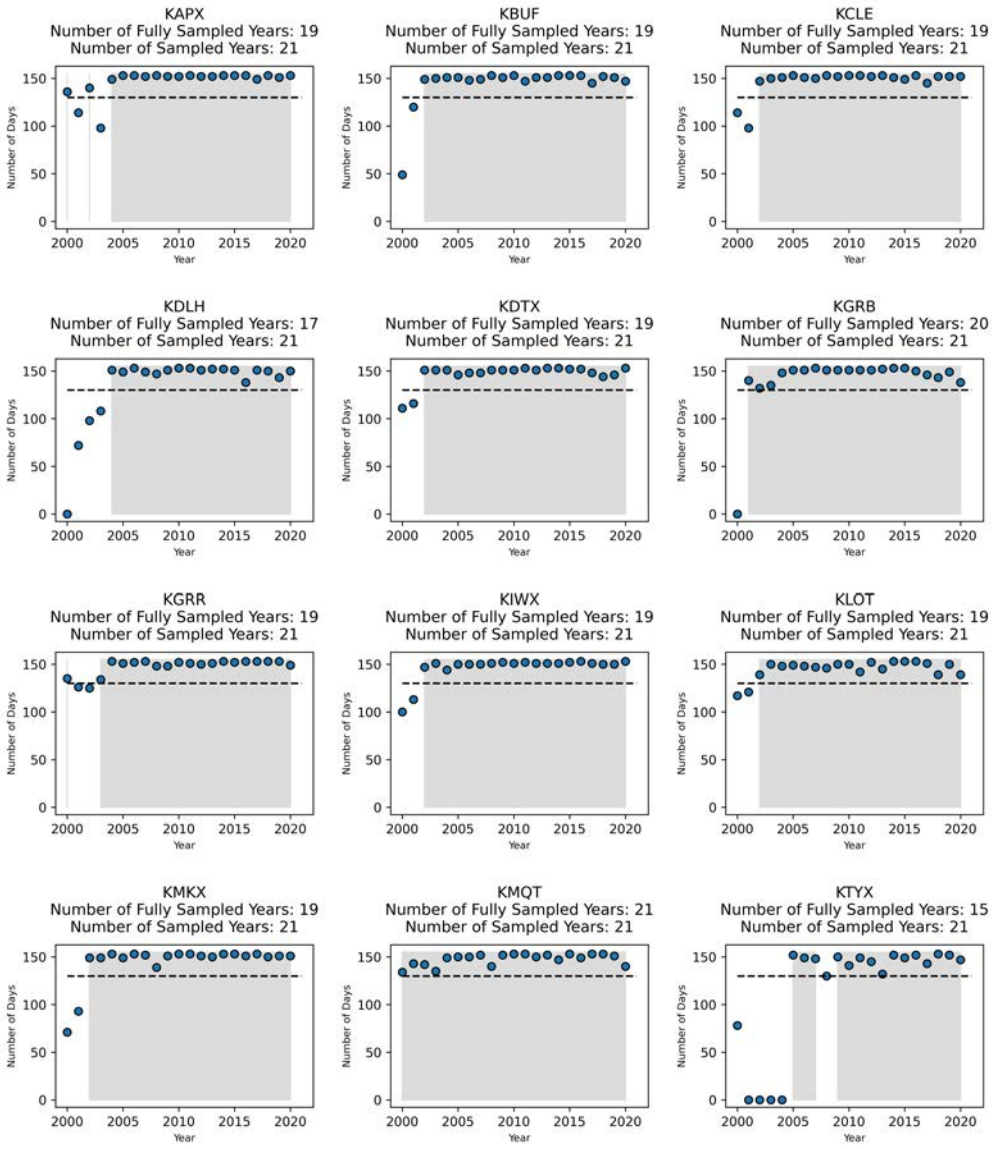


Figure B.3: Plots of the number of days with available scans per station. Shaded regions represent years with more than 130 days between June and October. Number of sampled years refers to the number of years when there was at least one month in the archive, whereas the number of fully sampled years is the number of years with more than 130 days.

B.4 Detection Biases Analyses

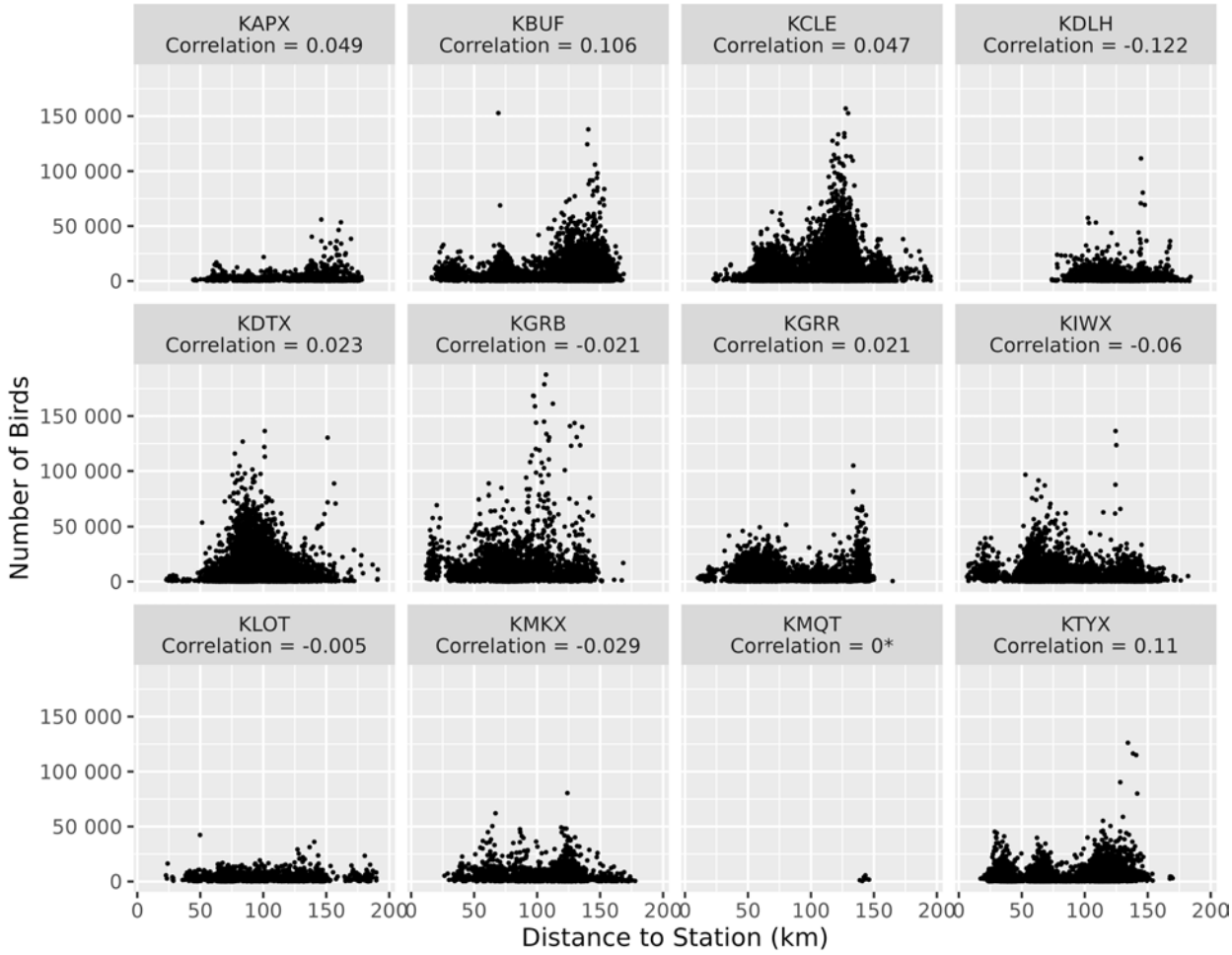


Figure B.4: Scatter plots of the distance to the radar station against the number of birds counted in each detection of our dataset. Correlation coefficients obtained using the Kendall rank correlation test. All results were statistically significant, except for KMQT (likely due to small sample size).

Bibliography

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

Appendix C

Spatial sampling schemes for radar, eBird, and BBS data

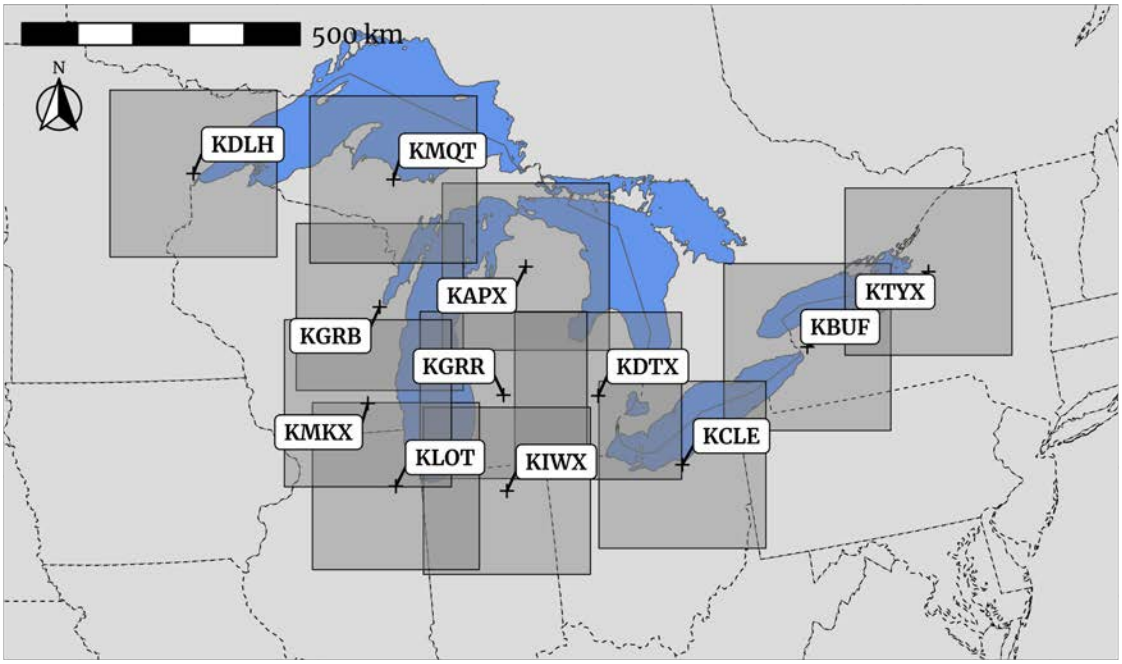


Figure C.1: Map of the sampled regions around each of the twelve radar stations in the Great Lakes region

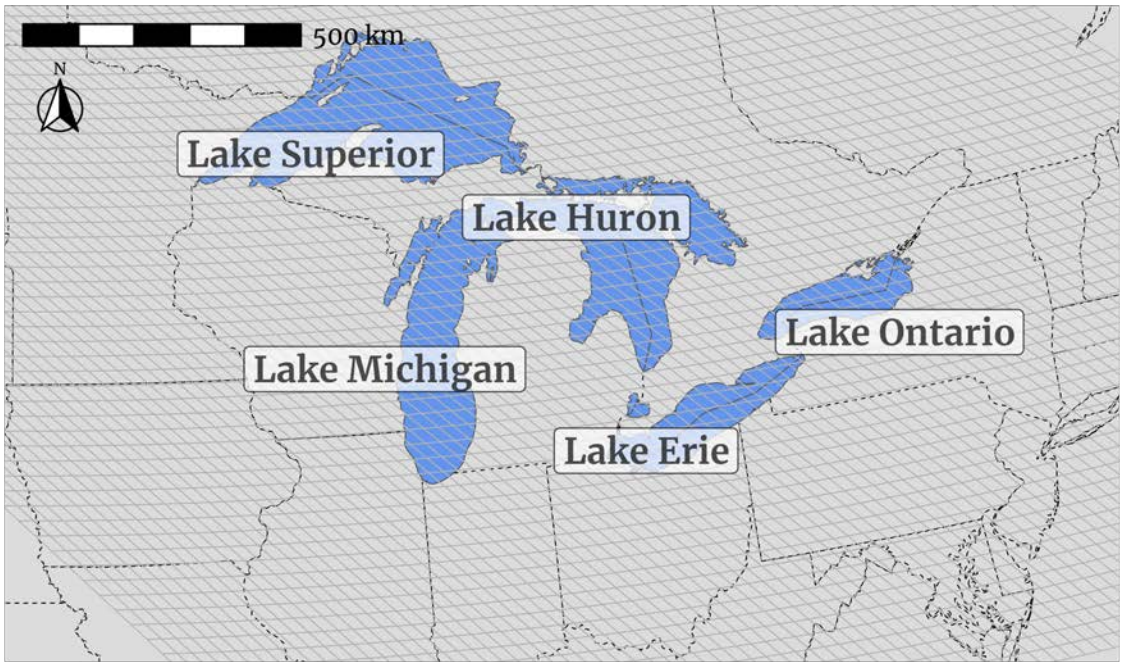


Figure C.2: Map of the sampling grid used by eBird. Each cell has 27 km of width and height.

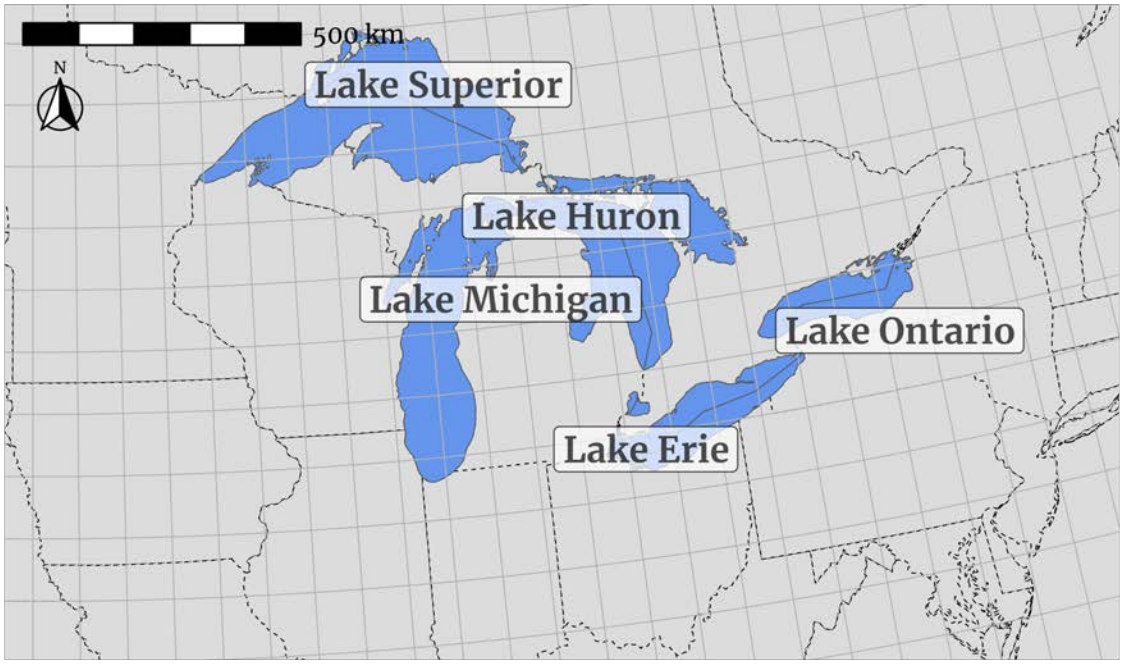
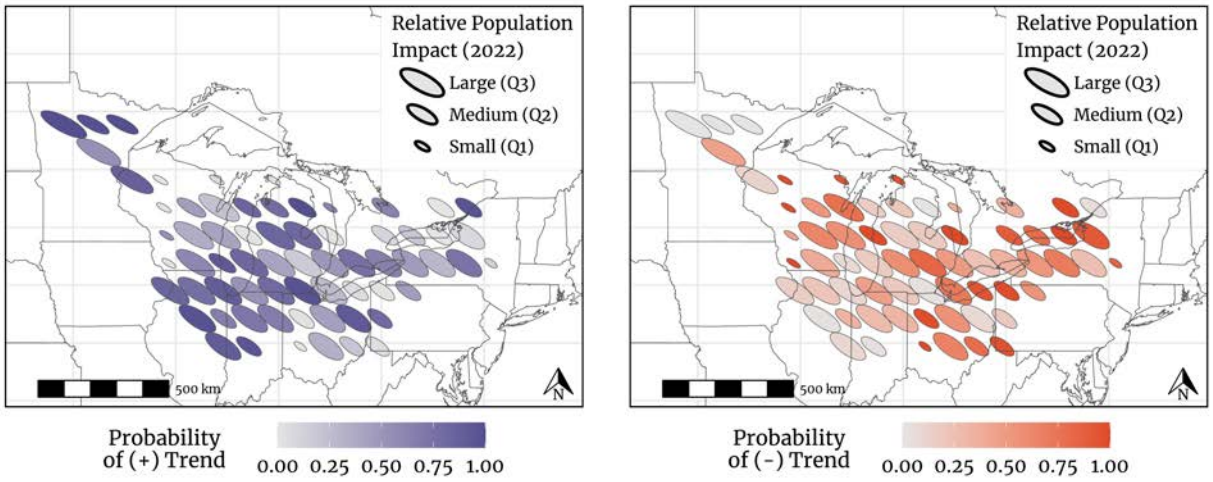


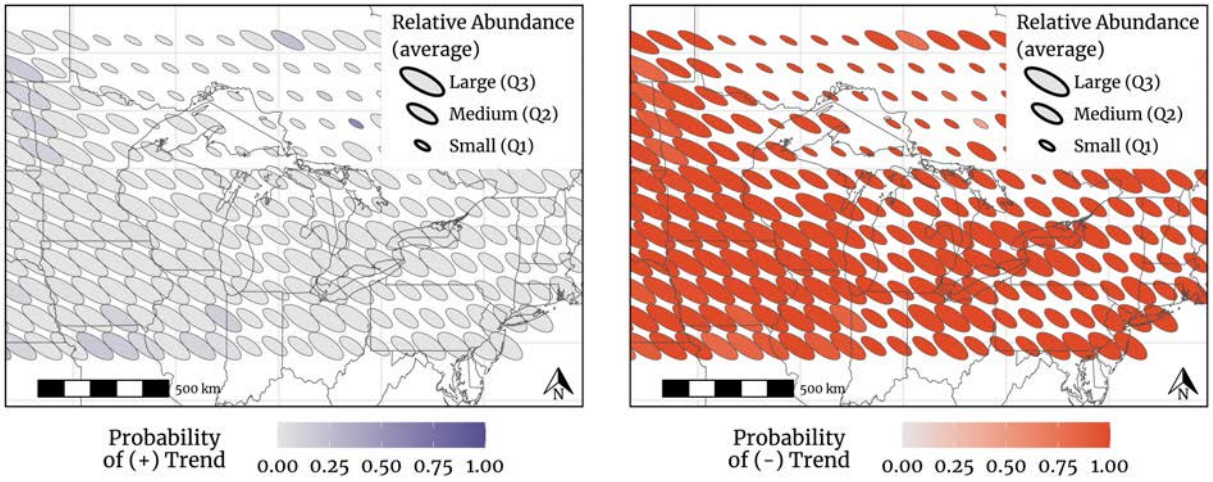
Figure C.3: Map of the sampling grid used by the Breeding Bird Survey. Each cell has 1° of width and height, which corresponds to approximately 100 km within our study region.

Appendix D

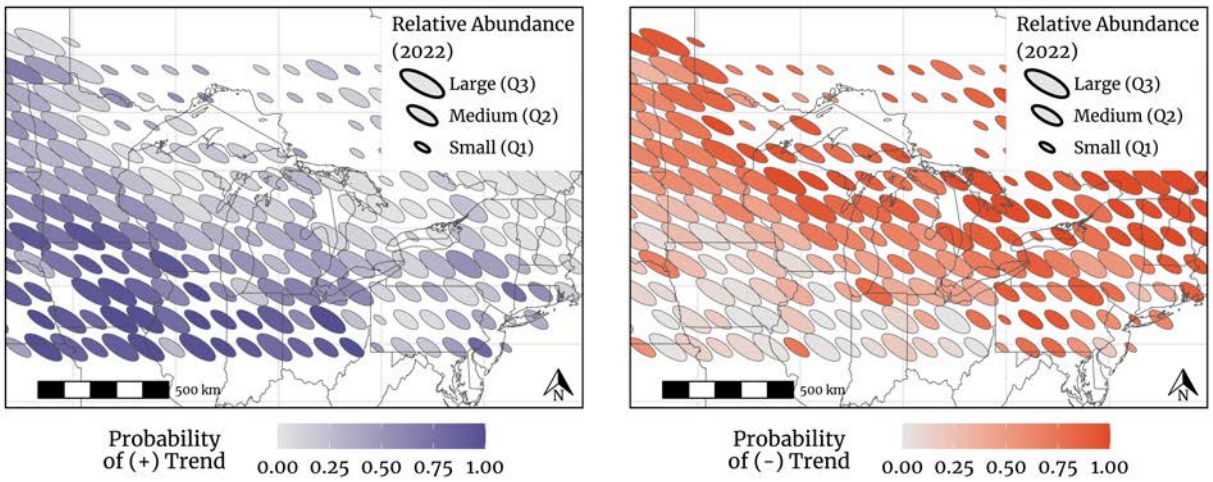
Maps of directional probabilities of trends for each source in continuous scale



(a) Directional probabilities of population impact trends obtained from radar data. Cell areas are proportional to estimated population impact in 2022.



(b) Directional probabilities of relative abundance trends obtained from eBird aggregated across all species. Cell areas are proportional to average relative abundance from 2012 to 2022.



(c) Directional probabilities of relative abundance trends obtained from Breeding Bird Survey data aggregated across all species. Cell areas are proportional to estimated relative abundances in 2022.

Figure D.1: Maps representing the modeled probability of a negative or a positive trend for each of our

Appendix E

Supplemental figures supporting the analysis of the data from Manaus

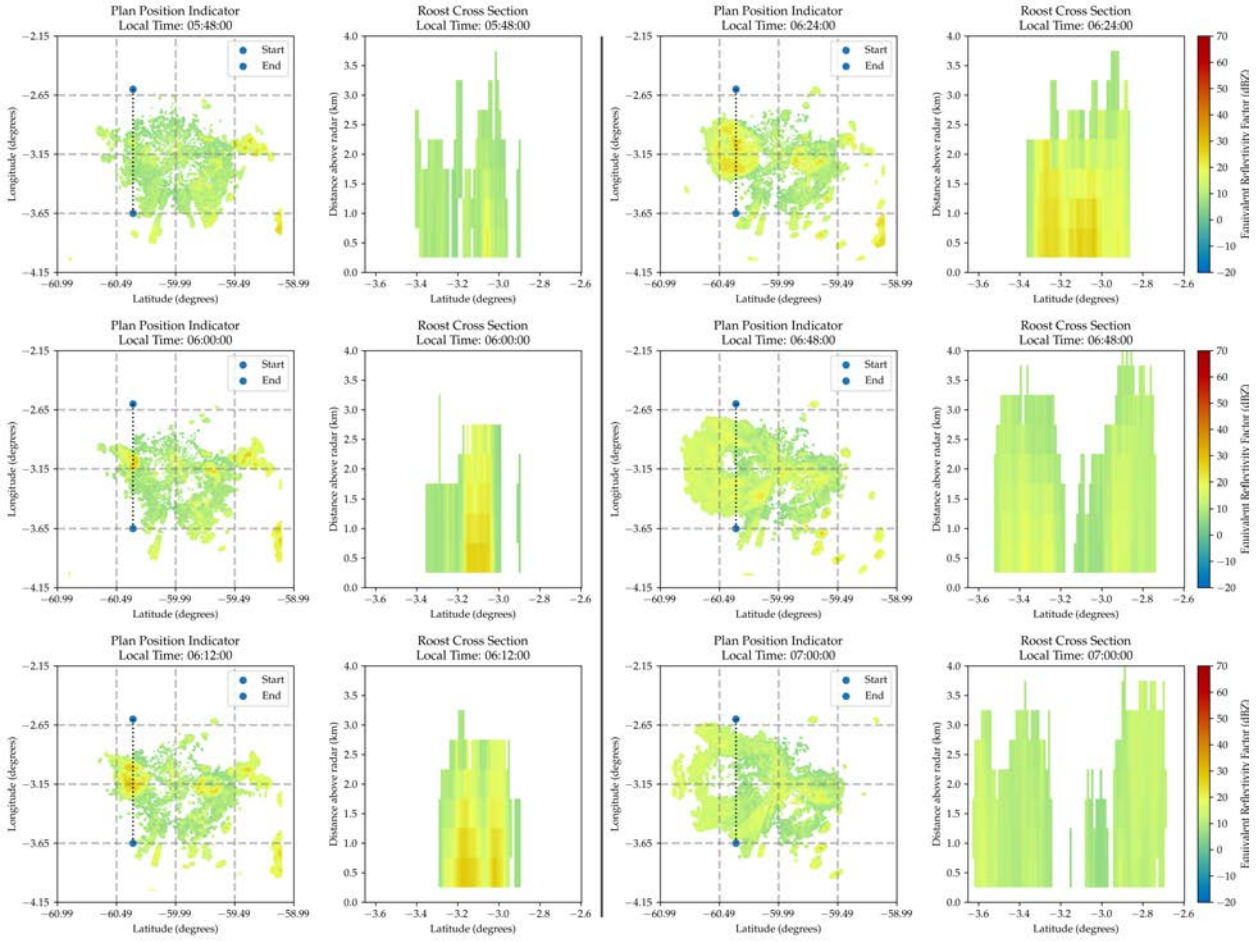


Figure E.1: This panel shows subsequent scans made on March 3, 2015. For each pair of plots, figures on the left are plan position indicators with the radar at the center, while figures on the right are cross sections of the main roost dispersing from Comaru Island. The cross section was made at the line represented on the plan position indicator.

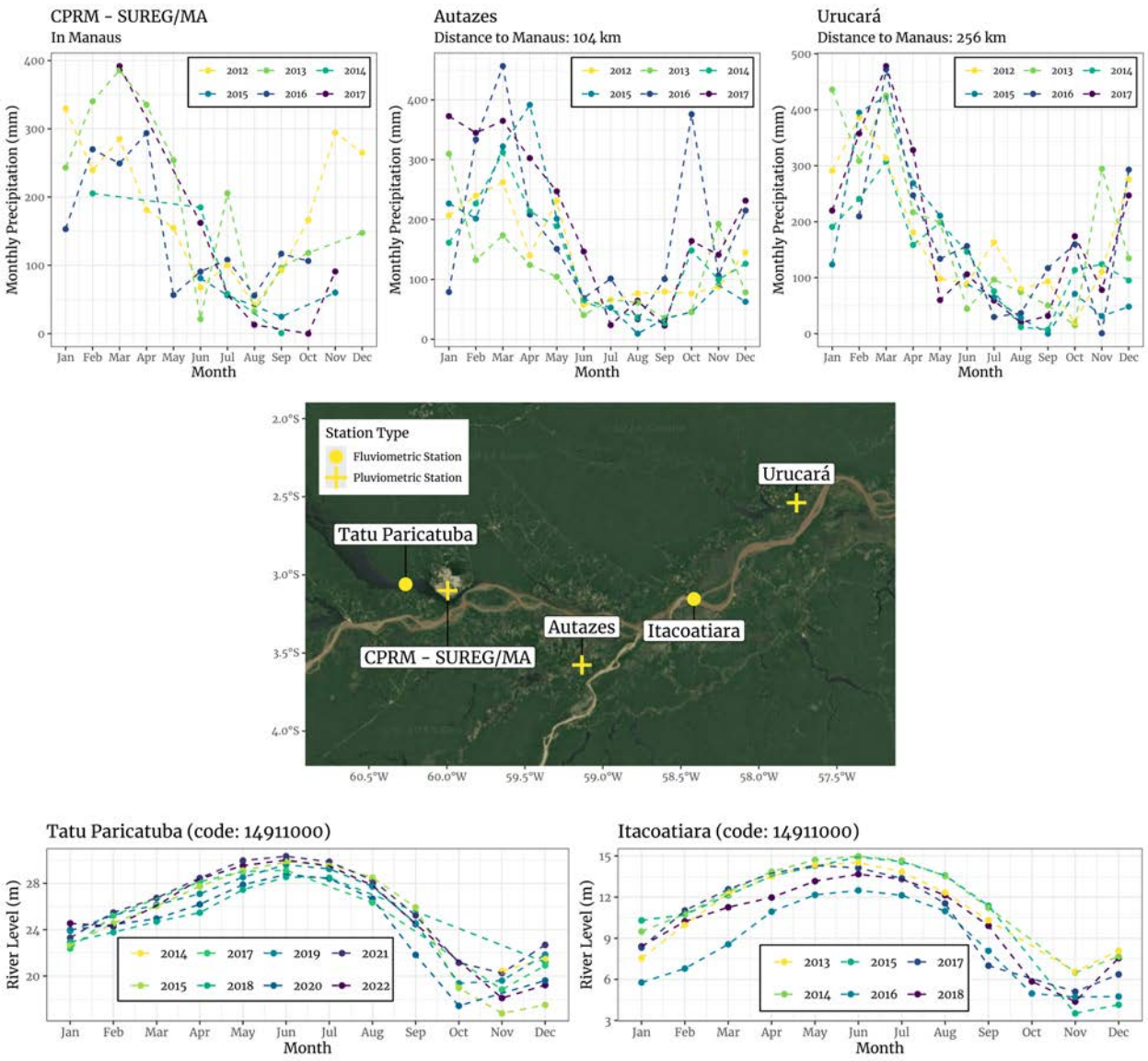


Figure E.2: This panel characterizes the so-called monomodal flood pulse that regulates biogeochemical cycles in the Amazon basin. The upper plots contain historical measurements of monthly precipitation (measured in milimeters) from three pluviometric stations up to 256 km from Manaus. The bottom plots represent historic water level measurements collected by monitoring instruments close to our study area. The map depicts instrument locations. The data is collected, processed, and maintained by the Agência Nacional de Águas of the Brazilian government.

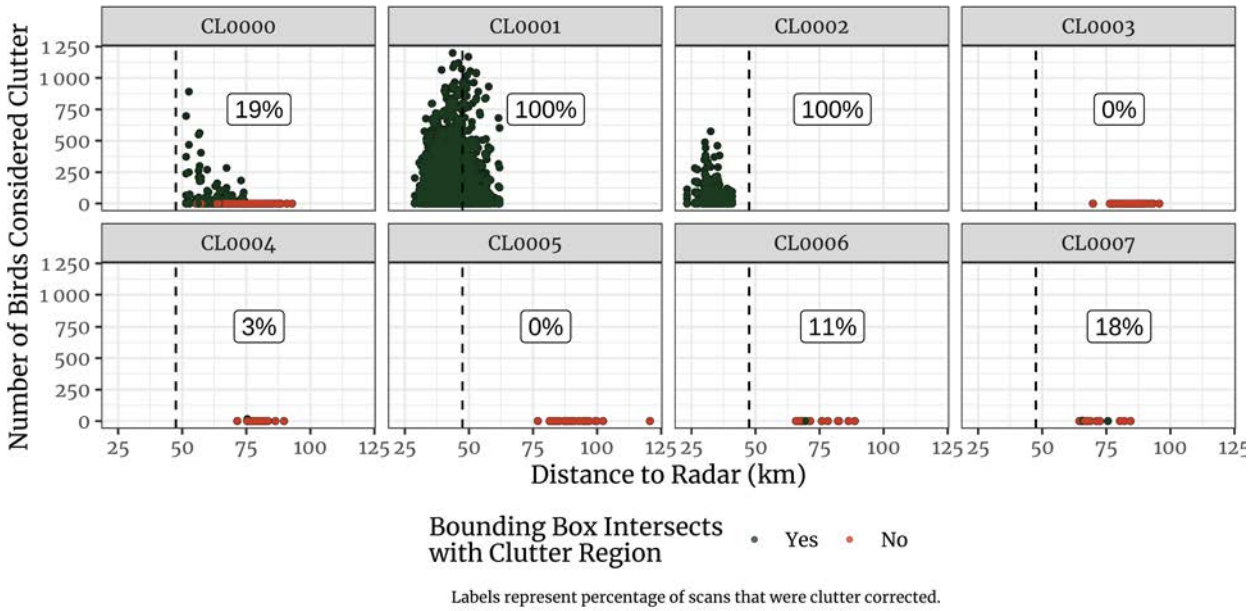


Figure E.3: The plots represent, for each cluster of roosts, the number of birds that were removed by our clutter filtering procedure as a function of distance between the radar and the centroid of the bounding box (measured in kilometers). The labels contain percentages of the total number of sweeps that were effectively filtered. The points are colored in dark green if the bounding box intersects the clutter mask, and in red if they do not. Notice that even though the centroid of a bounding box might be outside of the mask range of 27.5 km, the bounding box itself can still intersect the clutter mask.

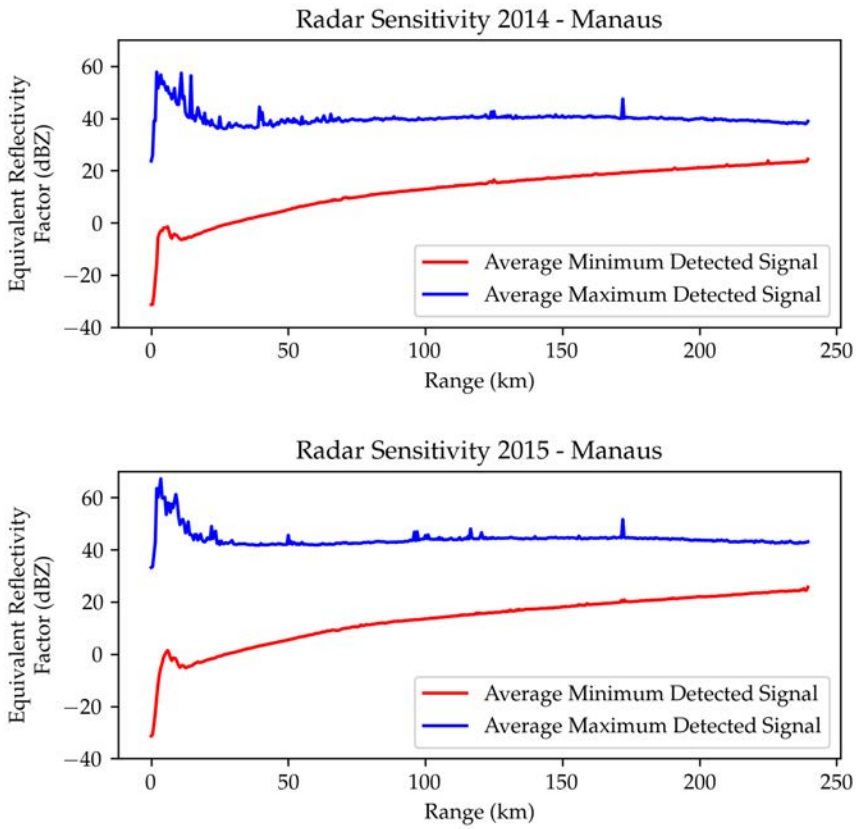


Figure E.4: Plots representing the minimum (red) and maximum (blue) detected equivalent reflectivity factor in dBZ for each range gate, averaged across all scans made by the Manaus radar within our study period for each year. We note that the curves for both years are similar, indicating that there were no instrument changes between 2014 and 2015.

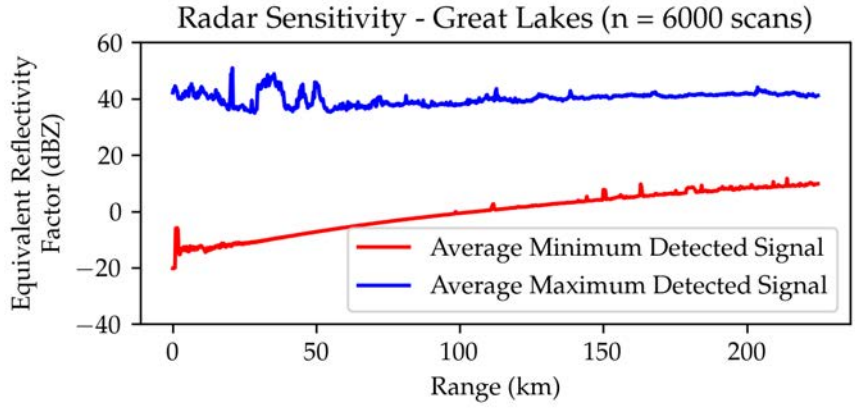


Figure E.5: Plots representing the minimum (red) and maximum (blue) detected equivalent reflectivity factor in dBZ for each range gate, averaged across a sample of 6000 scans made at dawn by twelve radar stations in the Great Lakes region from 2001 to 2022.

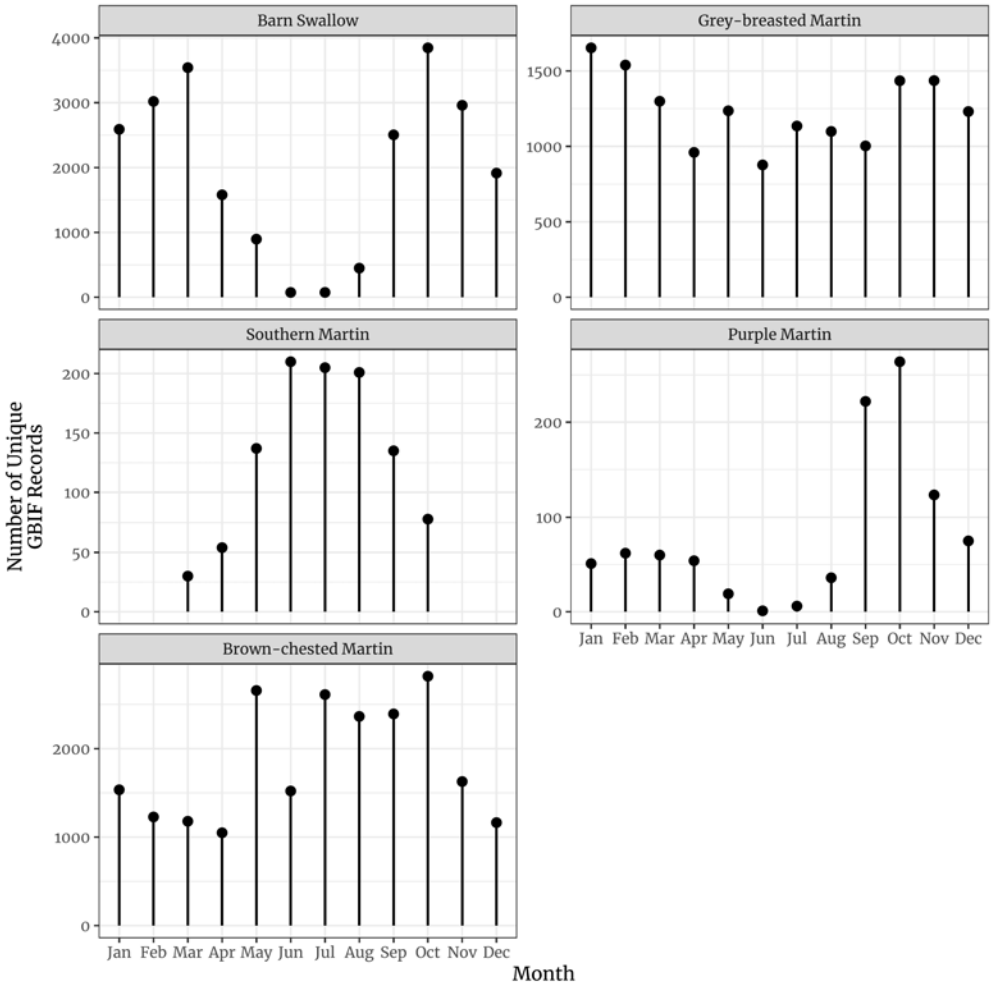
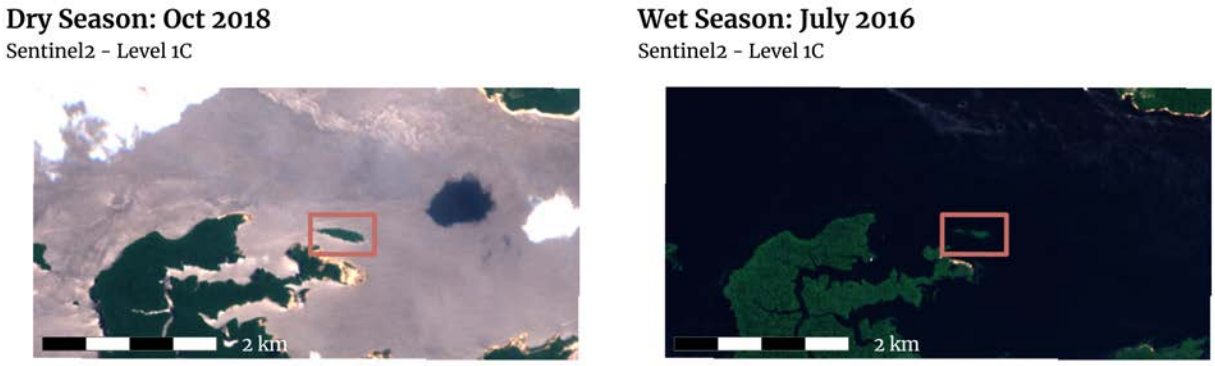
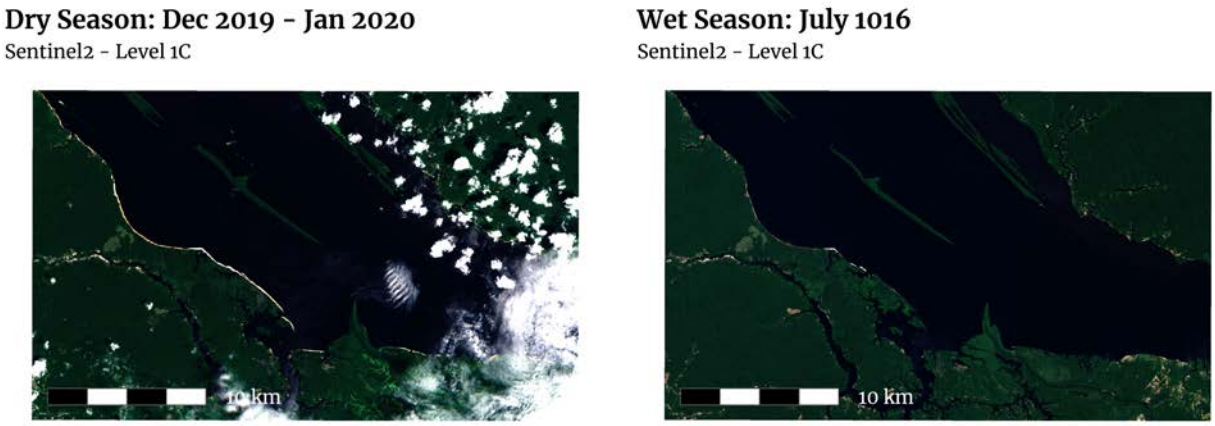


Figure E.6: Plots representing the number of records available in the Global Biodiversity Information Facility and in the Brazilian community science platform called Wikiaves for each of the species known to join roosts in larger numbers in the Amazon Rainforest. Purple Martins are Nearctic-Neotropical migrants (Gbif, 2024d). Southern Martins and Brown-chested Martins subspecies fusca are Austral migrants (Gbif, 2024c, 2024e). Gray-breasted Martins and Barn Swallows are partial migrants, and Brown-chested Martins subspecies tapera are residents in the Amazon (Gbif, 2024a, 2024b). Data is collected from museums and community science databases such as iNaturalist and eBird, and represent occurrence-only observations from 1819 to 2024. Individual records can represent hundreds of thousands of individuals.



(a) Comaru Island (highlighted) in the Negro River.



(b) Anavilhanas National Park in the Negro River.

Figure E.7: True color renderings of Sentinel reflectance data (10 m resolution) (European Space Agency, 2022). We created mosaics from month/year combinations with minimal cloud blockage. The peak of the dry season is typically between October and November, whereas the peak of the wet season is between June and July.

Bibliography

European Space Agency, E. (2022). Sentinel-2 MSI Level-1C TOA Reflectance. http://dx.doi.org/10.5270/S2_-742ikth

Gbif. (2024a). Hirundo rustica - Occurrence Download. https://doi.org/10.15468/dl.syk7gh

Gbif. (2024b). Progne chalybea - Occurrence Download. https://doi.org/10.15468/dl.zzr2t3

Gbif. (2024c). Progne elegans - Occurrence Download. https://doi.org/10.15468/dl.qgrem4

Gbif. (2024d). Progne subis - Occurrence Download. https://doi.org/10.15468/dl.j683da

Gbif. (2024e). Progne tapera - Occurrence Download. https://doi.org/10.15468/dl.pjvagv