**Optimal Bird-Watching in Massachusetts: Bird Frequency Forecasts and**

**Time-Location Recommendations**

**Authors:** Michael Aldoroty, Neil Ghosh, Nhat Pham, Hajera Siddiqui

**Github:** <https://github.com/nhathpham/Bird-Watching-Recommendation-System>

1. **Summary**

Bird watching is a valuable activity for enthusiasts, conservationists, and researchers, linking the joys of observation with the needs of environmental conservation. However, planning effective bird-watching excursions, especially for beginners, can be challenging. While recent technology advancements have improved bird identification through image and audio recognition (Green, 2020), there is still a need for better bird sighting recommendations. Current platforms, like eBird.com, while invaluable for their extensive historical data and explorative features, can be challenging to navigate and lack predictive capabilities. Our project fills this gap by offering an interactive dashboard with predictive models forecasting weekly detection rates of different bird species across Massachusetts counties. This tool also provides suggestions on the best hotspots and times for bird-watching based on past data. Our approach aims to simplify the bird-watching planning process and has potential applications in avian research and conservation.

Our project uses the eBird dataset from the Cornell Lab of Ornithology, which consists of global bird sightings recorded by professional and amateur bird watchers (Cornell, 2023). To understand local preferences, we conducted a survey among birders in Massachusetts (MA), receiving 102 responses. Notably, 83% indicated a tendency to visit regular or nearby locations (Appendix Figure A1). This insight, combined with computational limitations, guided our decision to focus our analysis exclusively on MA. We extracted 8.5GB of bird sighting data for MA from eBird.org, spanning from 2013 to 2022. Due to time constraint, our final forecast model focuses on 50 most frequently observed bird species.

The dataset is organized into two primary components: observation data and checklist data. In the observation data, each row represents an individual bird species sighting, providing detailed information about the species observed, the specific location, date, time, and any additional notes. Checklist data, on the other hand, compiles these individual observations into structured records centered around specific bird-watching outings or events. Each checklist summarizes the observations made during these events, including aggregated details such as the total count of each species and contextual information like the location, date, time, and the number of participants.

To predict weekly bird detection rates, we tested four statistical and machine learning time series forecasting models. After evaluating their performance and scalability, we narrowed our selection to Prophet and Greykite, with Greykite outperforming in overall efficacy. Prophet is a specialized time series forecasting model that decomposes time series data into trend, seasonality, and other components and combines them in an additive model. Greykite uses a hybrid forecasting model that can incorporate a wide range of forecasting techniques to handle diverse time series data.

Variability in model accuracy was noted across species and counties, reflecting diverse bird behaviors and regional differences. Nevertheless, all models maintained satisfactory predictive accuracy. Additionally, we analyzed historical data to identify optimal bird-watching times and hotspots within each county. All predictive and descriptive insights were integrated into an intuitive, interactive Tableau dashboard.

1. **Methods** 
   1. **Data processing**

The initial data cleaning and filtering process utilized the ‘auk' package in R, a tool specifically designed by Cornell Ornithology Lab for processing eBird data. This involved standardizing variable names, assigning appropriate data types to each variable, and collapsing group checklists to remove duplicates from shared eBird entries by multiple users. Given the semi-structured nature of the eBird dataset, we further refined the data to ensure accuracy. We included only complete checklists where users reported all the birds they observed, and prioritized only stationary and traveling protocols. To mitigate variability, we followed Cornell’s suggestions on effort variables. This included limiting our data to checklists of up to 6 hours in duration, covering distances no longer than 10km, recording at speeds under 100 km/h, and compiling groups of 10 or fewer observers.

Following the data cleaning in R, each row represented a bird sighting from a specific birding trip. The data was then grouped by county and aggregated into weekly intervals. This allows us to compute two metrics: the number of checklists in each county that recorded the bird of interest, and the total number of completed checklists, both with and without sightings of the bird of interest, for the same timeframe and county. Using these figures, 'detection\_rate' was calculated as the proportion of checklists that noted the species of bird to all checklists. For reliability, we included only rows with at least 5 checklists and focused on the 50 most common bird species. For each species in each county and year, we assigned a null detection rate to weeks without any checklists to maintain the continuity of the time series feature, accommodating models that required uninterrupted data sequences.

To prepare the data for the descriptive analyses on time-location trends, we performed additional processing on location data. The dataset contained over 3500 predefined hotspots, towns, and postal codes, along with over 110,000 user-input locations, the latter comprising a third of the observations. Our goal was to identify and retain locations within the referenced hotspots, even when entered under different names by users. To achieve this, we used the KDTree algorithm to efficiently pair each user-input address with a ‘nearest’ predefined location based on their longitudes and latitudes. The actual geographical distances between potential matches were calculated using Python's geodesic package. To ensure accuracy in matching, we also calculated the Levenshtein ratio to compare the similarity of names between user-input locations and their nearest referenced counterparts. We established matching criteria based on empirical testing: a maximum distance of 0.8km and a Levenshtein ratio of 0.65 or higher. This approach matched 10% of unique user-input locations. In our final dataset for descriptive analysis, unmatched user-input locations were excluded.

* 1. **Time Series Forecasting**

To forecast weekly detection rates for 50 bird species across 14 counties, we evaluated four algorithms on a few bird species with varied behaviors. Our assessment included traditional time series models, ARIMA/SARIMAX, as well as modern approaches like Gaussian Processes, Prophet, and Greykite. We found Prophet and Greykite outperformed the others in both performance and scalability, leading to their application across all bird species.

To assess Prophet and Greykite algorithms, we applied different model configurations: a baseline model, a model with finely-tuned parameters, and a model applied to smoothed data. Data smoothing techniques, such as rolling averages or exponential smoothing, were essential in handling the data for various bird species and counties. This necessity arose from the inherent volatility and sporadic patterns in the data. Specifically, since citizen science data can exhibit fluctuations in detection rates caused by factors like the timing of birdwatching, weather conditions, and other variables, which do not align with the actual migratory patterns of birds. These smoothing methods are crucial for mitigating noise in the data and emphasizing the underlying trends by ironing out these fluctuations. This variability arose from inherent bird behaviors and disparate data volume and quality across different counties. For a comprehensive evaluation, each model underwent a time series train-test split, employing data from 2013-2021 for training purposes and 2022 for testing. We used RMSE (Root Mean Square Error) and MAE (Mean Absolute Error) as our primary performance metrics, while also considering model run-time as a key indicator of scalability.

* + 1. **Greykite**

Greykite is an open-source Python package developed by Linkedin, featuring its main forecasting algorithm, Silverkite. Silverkite operates through a two-phase approach, initially focusing on creating a conditional mean model by extracting and transforming features from time series data, followed by fitting these features using machine learning algorithms. The second phase involves fitting a volatility model to the residuals. Silverkite also enables the inclusion of future-related data via regressors with past and future values, directly used in both model fitting and forecasting.

We deployed Silverkite to forecast weekly detection rates for each bird species across Massachusetts counties. For each model, we started with establishing a baseline using Silverkite's default parameters. Then, we fine-tuned this model by adjusting specific parameters. Observing varied growth patterns across species and counties, we employed linear growth with potential changepoints optimized via grid search. To mirror the birds' availability cycles, we integrated seasonality on both yearly and quarterly bases. The autoregressive parameter, leveraging past values and aggregated lags, was applied to capture temporal correlations. Ridge regression was chosen for fitting due to its effectiveness in diverse scenarios. Additionally, we implemented hyperparameter overrides for pre-processing steps, including setting thresholds for outliers and interpolating missing data. The performance of both the baseline and the tuned models was evaluated using the aforementioned train-test split, further supported by a 2-fold rolling window Cross-Validation with a forecast horizon of 52 weeks. We compared the RMSE and MAE of the baseline and tuned models, and selected the better-performing model for each bird-county combination.

* + 1. **Prophet**

Prophet, an open-source forecasting tool developed by Facebook, uses a Bayesian framework to forecast tasks with time series data. The basic principle behind Prophet is to represent time series data as a consolidation of trend, seasonality, and noise elements. The trend element determines the polarity of the direction, as in a positive or negative progression. The seasonality element examines patterns, whether it be weekly or monthly trends, and employs the Fourier series for its modeling. Unpredictable variations that are independent of trend or seasonality can be examined by the noise element. Prophet is well-suited for time series data that is characterized by strong seasonality and ample historical data, making it a reliable choice for applications like bird detection rate prediction. Additionally, Prophet excels at handling missing data and adapting to shifts in trends (Khare, 2023).

To prepare the data for Prophet, the data went through several data preprocessing steps such as removing unneeded rows and columns, changing the date and target column names, and modifying the datatypes of the data values to ensure they met the algorithm's specific input requirements. The default parameters, including the addition of yearly seasonality, number of periods, and type of frequency, were utilized, resulting in a quick and efficient setup. The model exhibited the capability to automatically capture underlying seasonality trends with minimal manual tuning. But there are many options to further tune the model to enhance results.

Although the model was able to generate good values for the evaluation metrics, some of the plots were found to not perform as well. The plots displayed large ranges of lower and upper bounds of the forecasts and some of the forecast lines did not exactly match the data points. To combat this, smoothing capabilities were explored and after conducting research, rolling average and exponential smoothing were determined to be good for smoothing time series data. Both performed really well on the Prophet model, resulting in better RMSE and MAE values than before smoothing. It was determined that since exponential smoothing performed slightly better than rolling average, this would be chosen to smoothe Prophet. This evaluation process, with Prophet being smoothed by exponential smoothing, took approximately one minute and 26 seconds to test on 50 bird species and 14 counties.

* 1. **Time-Location Descriptive Analysis and Tableau Dashboard**

To complement the forecasting models' predictions of weekly bird species detection rates, we conducted two descriptive analyses to provide users with recommendations for the best times and specific hotspots within each county for birdwatching, corresponding to the same forecasted period. These analyses used recent data from 2019-2022 to ensure relevance.

Our first analysis focused on location suggestions. After refining the dataset to 3,592 unique hotspots in Massachusetts, we calculated the weekly detection rate of each bird species at every hotspot over the four years. This was achieved by dividing the number of checklists mentioning a particular species by the total number of checklists at each hotspot. Our second analysis aimed to determine the best times of the day for bird sightings. Checklists were segmented into two-hour intervals throughout the day, and for each species, we computed the ratio of checklists featuring that species to the total checklists across Massachusetts for each time interval.

Lastly, we developed a Tableau dashboard to display both predictive and descriptive analyses’ results. Through its interactive interface, the tool aids users in making better-informed decisions for their bird watching trips based on specific bird preferences and desired timings.

1. **Results** 
   1. **Weekly Detection Rate Forecast**
      1. **Greykite and Prophet Comparison**

We conducted 700 models for 50 bird species using both Prophet and Greykite. Across both algorithms, we observed that data smoothing significantly improved model performance compared to using raw data. Specifically, exponential smoothing was preferred over rolling average due to its superior performance and relevance for our context, as we believe that more recent trends are more indicative of future patterns.

Greykite mostly outperformed Prophet in terms of performance metrics, demonstrating higher predictive accuracy. However, Prophet models were notably faster, with processing time of 5.36 seconds per bird, compared to Greykite's 2 minutes per bird. Despite this, for this phase of the project, we prioritized prediction accuracy and chose Greykite for its reliability in forecasting detection rates. This decision aligns with our goal of providing precise information to bird watchers. In our current setup, the model's processing speed is less critical as we're not running the models in real-time but rather storing prediction results for static use in the dashboard. Future optimizations are planned to either improve Greykite's runtime or enhance Prophet's accuracy.

Figure 1 compares the RMSE and MAE results of the fine-tuned Greykite and Prophet for three distinct bird species in selected counties. These birds were chosen for their varying behaviors: the migratory Yellow Warbler, present in Massachusetts only in summer; the resident Black-capped Chickadee, common throughout the year; and the Eastern Bluebird, whose winter presence in Massachusetts has been increasing in recent years. Table 1 contrasts Prophet and Greykite's performance by showcasing the median RMSEs and MAEs of 14 counties’ models for each species.

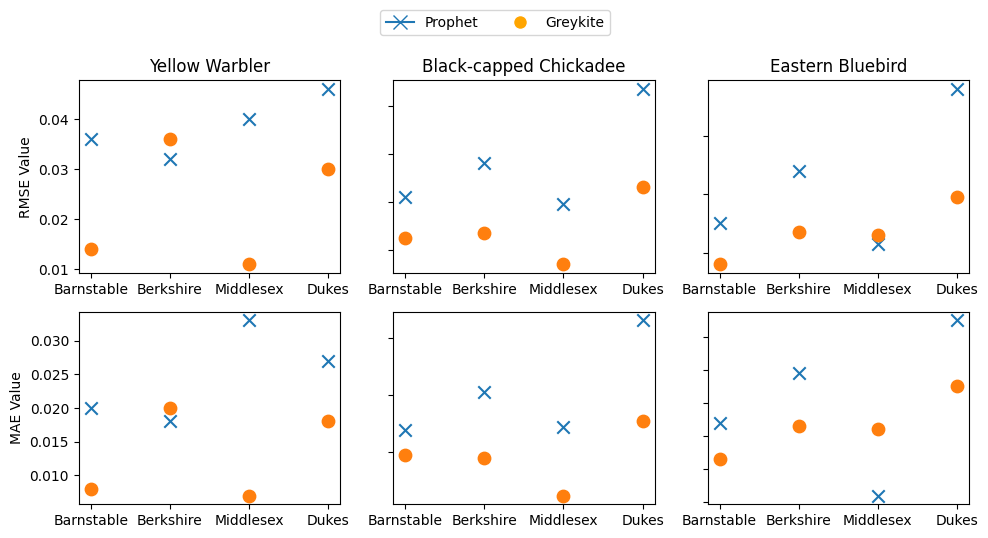


Figure 1. RMSE and MAE for Greykite and Prophet in 4 Counties

| Species | Prophet | | Greykite | |
| --- | --- | --- | --- | --- |
| Median RMSE | Median MAE | Median RMSE | Median MAE |
| Yellow Warbler | 0.043 | 0.029 | 0.021 | 0.013 |
| Black-capped Chickadee | 0.083 | 0.067 | 0.053 | 0.04 |
| Eastern Bluebird | 0.030 | 0.036 | 0.027 | 0.023 |

Table 1. Median RMSE and MAE for Greykite and Prophet across 14 Models/ Counties

* + 1. **Final Model Results**

Overall, the performance of the fine-tuned Greykite models was satisfactory, with a median RMSE of 0.034 and median MAE of 0.026. As anticipated, model efficacy varied across species and counties due to differing behavioral patterns and habitat preferences.

Models for migratory birds such as Bufflehead, Baltimore Oriole, Yellow Warbler, and Pine Warbler, demonstrated the highest predictive accuracy. This is attributed to their pronounced seasonal detection rates, yielding lower RMSEs, MAEs, and narrower confidence intervals. In contrast, models for resident species like American Robin, House Finch, Blue Jay, and Black-capped Chickadee exhibited higher RMSEs and MAEs along with wider confidence intervals. Factors potentially influencing this include their widespread occurrence leading to more variable detection patterns, potential underreporting by birdwatchers, and their complex behaviors responsive to environmental and human factors.

Model performance varied across different counties in Massachusetts. Hampshire, Berkshire, Barnstable, and Middlesex counties exhibited notably superior results in terms of RMSEs and MAEs. This heightened performance in Western Massachusetts counties like Hampshire and Berkshire can be attributed to their natural landscapes, lower population density, and diverse habitats. Barnstable's coastal and marshland environments likely contribute to consistently stable bird populations, while Middlesex, despite its urban density, maintains various conservation areas and consistently boasts the highest volume of annual birdwatching checklists. On the other hand, the islands of Nantucket and Dukes, which have limited data and unique bird populations, along with Hampden County's mixed urban-rural setting, introduced a greater degree of variability in bird detection rates. Appendix Figures A2-3 compare the RMSE and MAE of selected migratory and resident birds, and across counties.

Figures 2-7 shows the observed versus predicted detection rates over time for three selected species in Hampshire and Dukes counties. The data indicate that the models yield more accurate forecasts for migratory species compared to resident species, and also reveal variations in model performance across different counties.

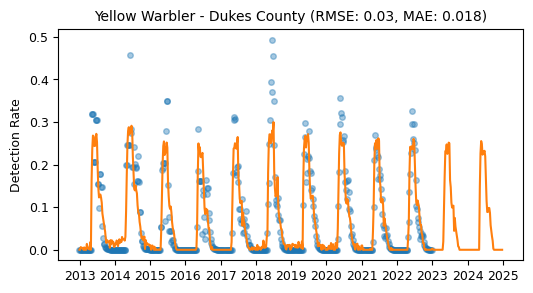
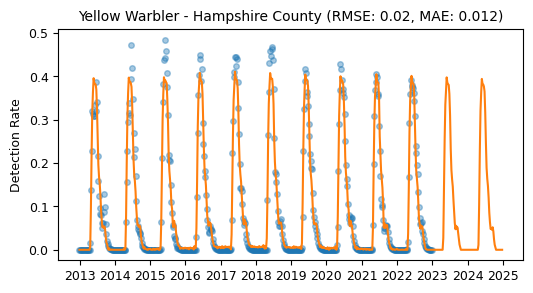
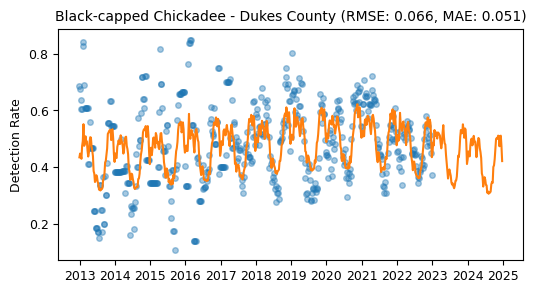
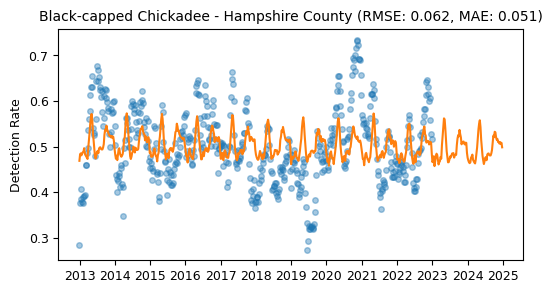


Figure 2. Yellow Warbler in Hampshire County Figure 3. Yellow Warbler in Dukes County

 Figure 4. Black-capped Chickadee in Hampshire County Figure 5. Black-capped Chickadee in Dukes County

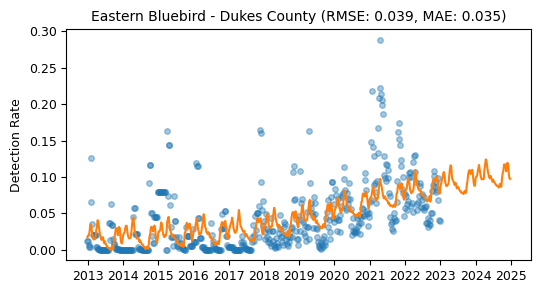
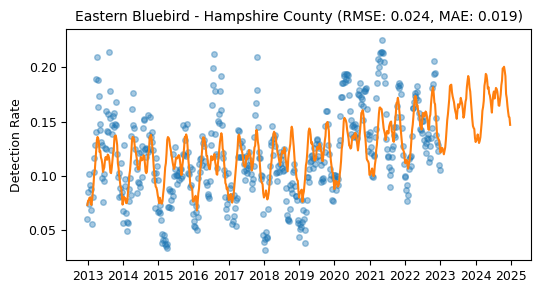


Figure 6. Eastern Bluebird in Hampshire County Figure 7. Eastern Bluebird in Dukes County

* 1. **Tableau Dashboard and Additional Time-Location Descriptive Analysis**

Figure 8 presents the layout of our interactive Tableau dashboard, which can be accessed publicly here: <https://bit.ly/3SqWssj> . Users can select a specific bird species, week, and forecast year to view the detection rate forecasts in all counties, ordered from the highest to the lowest. Adjacently, the upper right graph displays actual versus forecasted detection rates over time, offering insights into detection patterns and the accuracy of the forecast model. Subsequently, based on the user’s chosen reference year (spanning 2019-2022), the bottom-left graph shows the species' detection rates at times throughout the day during the same week as the forecasted one. The bottom-right bar chart and accompanying map highlight the locations with the highest detection rates within a county and week. The dashboard features interactive elements such as mutual filtering, where selecting a county, bird species, and year automatically filters related data. Additionally, selecting a location from the bar chart refines the map's focus to that particular site and selecting a location on the map directs users to the corresponding Google Maps site.

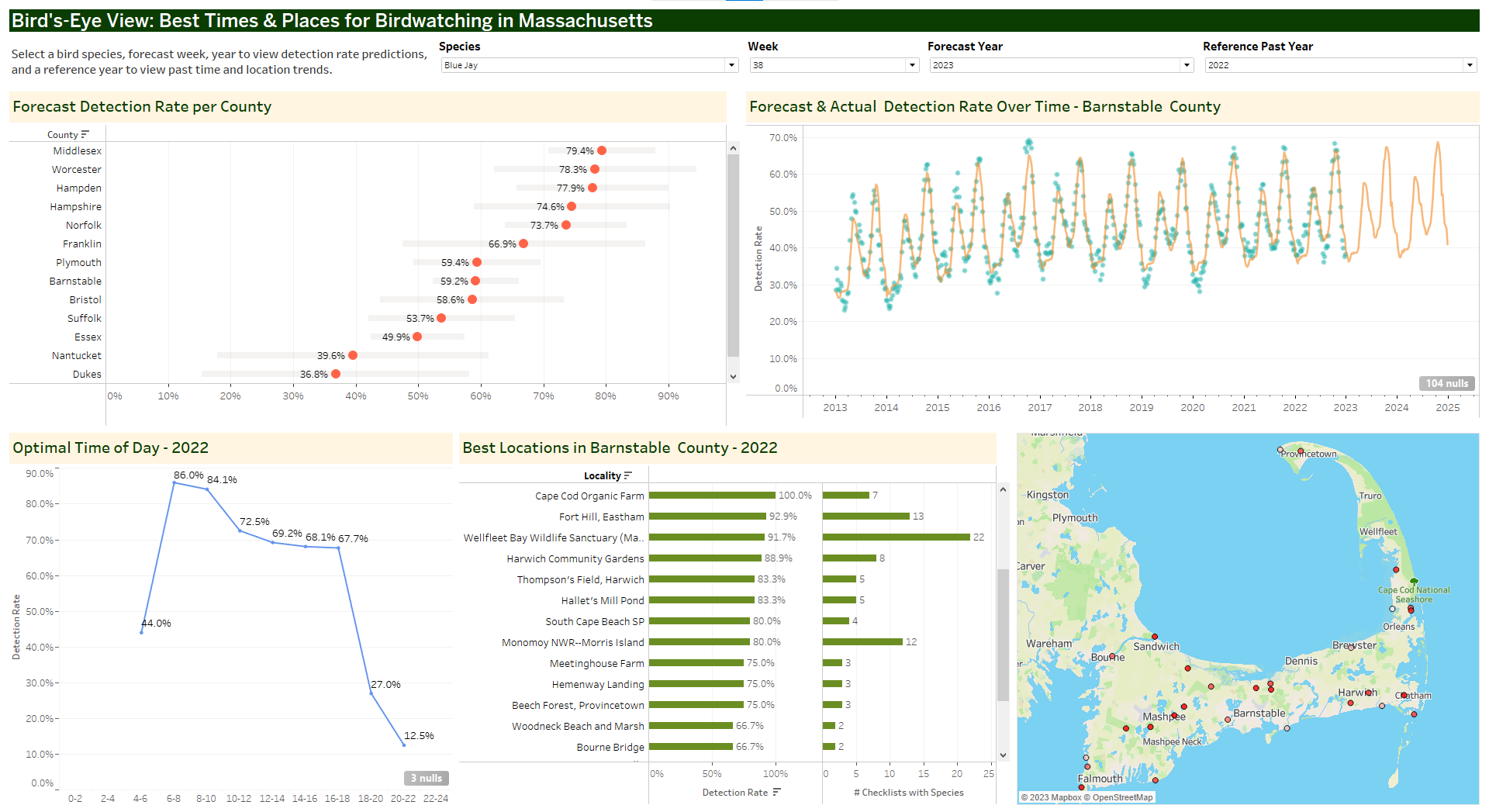
****

Figure 8. Tableau Dashboard

We conducted additional analysis using the time and hotspot dataset used to create time-location suggestions in the dashboard. The time-of-day detection rate data shows pronounced early morning activity (4-8am) for most bird species. Notably, coastal species such as Bufflehead and Gulls demonstrate comparable activity levels during afternoons, likely aligning with tidal cycles which influence food accessibility. Appendix Figure A4 shows the diurnal variations in bird activity for selected species, showcasing diverse temporal patterns. As we expand the project to include more species, we anticipate further variability, emphasizing the significance of this feature in understanding avian behavior.

In response to local birders' interest in discovering new, lesser-visited hotspots, we performed a preliminary analysis to explore overbirded and underbirded areas. Overbirded areas are characterized by a high volume of checklists yet lower detection rates, whereas underbirded areas see fewer checklists but boast higher detection rates. The scatter plot in Appendix Figure A5 illustrates the median weekly checklist count against the median detection rates for hotspots in certain counties in the spring and summer, when birdwatching is most popular. The graph highlights potential overbirded spots like Shady Knoll Campground (Brewster, MA). We plan to delve deeper into these findings in subsequent research.

1. **Discussion**

The results of this project provide numerous valuable insights into bird populations and their detection across various Massachusetts counties, providing a robust foundation for enhancing birdwatching activities and conservation efforts. By successfully meeting its objectives, the project has notably benefited the birdwatching community. The forecasting models currently deliver robust predictions, and the interactive Tableau dashboard enables birders to identify hotspots and ascertain the optimal times for bird observation. Additionally, the dashboard suggests new, promising locations for birdwatching, promoting the discovery of lesser-known areas with high detection rates. Overall, the predictive models and analyses conducted help users better understand avian behavior patterns, aiding both immediate birdwatching endeavors and long-term conservation planning.

However, a noteworthy challenge surfaced in the data collection process, particularly in Dukes and Nantucket counties. Although most counties contained plentiful data, the sparsity and volatility of the two counties made generating accurate forecasts challenging. The potential causes may include a limited number of citizens reporting bird sightings, potential mismatches between the birds in the dataset and the specific habitats on these islands.

Regarding the evaluated models, Greykite demonstrated more tuning flexibility and achieved better RMSE and MAE scores than Prophet. However, its increased tunability led to greater complexity and reduced scalability. Notably, Greykite also ran slower, a potential issue for larger datasets. Although currently tested on only 50 birds across 14 counties, and considering that Cornell's Lab of Ornithology updates the data monthly, the balance between accuracy and processing speed may not be immediately critical. However, as the data grows, training times could become a concern. Thus, further comparative tuning with Prophet is needed to determine its scalability relative to Greykite.

To further improve the project's impact, it can be scaled up and transitioned into a production-ready pipeline and model. Specifically, the pipeline of the project could take in new data as input, clean the data, run the models, and update the dashboard without someone individually controlling all of the steps. This would ensure that the information is always current, benefiting and enhancing the experience for both birdwatchers and conservationists.

Furthermore, incorporating spatial data such as habitat quality, land use changes, and migratory pathways into the model would offer a more holistic understanding of bird populations and their movements. Building a recommendation system for birdwatchers, suggesting prime locations and species to observe based on their historical data would make the project even more user-friendly and engaging for this community. By embracing these enhancements, the project can continue to meet its objectives effectively and, in the process, foster a stronger bond between birdwatching and conservation interests.

1. **Statement of contributions**

All members contributed equally to the project’s design and exploratory data analysis. Each member initially explored different forecasting algorithms, Hajera investigated Prophet, Michael analyzed Gaussian Processes, Neil focused on Greykite, and Nhat examined SARIMA. When the selection narrowed to Prophet and Greykite, Hajera and Michael spearheaded the design, development, training, and assessment of the Prophet model. In parallel, Neil Ghosh and Nhat Pham managed similar tasks for the Greykite model. Nhat Pham was in charge of data processing and developing the Tableau dashboard. Collective efforts were made in testing the dashboard and compiling the final report.

1. **References**

Cornell. (2023). *Use eBird data and tools*. eBird Science. Retrieved November 3, 2023, from https://science.ebird.org/en/use-ebird-data

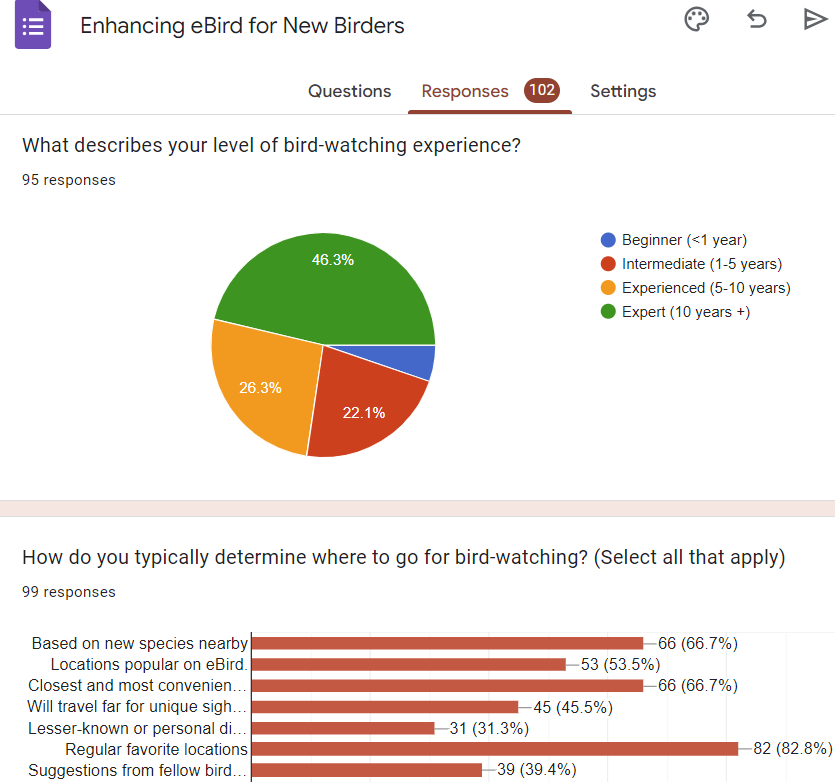
Greene, V. (2020, January 9). *Birding with Technology in the Year 2025: Our Predictions*. All About Birds. Retrieved November 3, 2023, from https://www.allaboutbirds.org/news/birding-with-technology-in-the-year-2025-our-predictions/

*Greykite - A flexible, intuitive and fast forecasting library*. (n.d.). LinkedIn Open Source. Retrieved October 16, 2023, from https://linkedin.github.io/greykite/

Khare, P. (2023, May 13). *Understanding FB Prophet: A Time Series Forecasting Algorithm*. Medium. Retrieved November 3, 2023, from https://medium.com/illumination/understanding-fb-prophet-a-time-series-forecasting-algorithm-c998bc52ca10

*12.2 Prophet model | Forecasting: Principles and Practice (3rd ed)*. (n.d.). OTexts. Retrieved October 16, 2023, from https://otexts.com/fpp3/prophet.html

1. **Appendix**

****

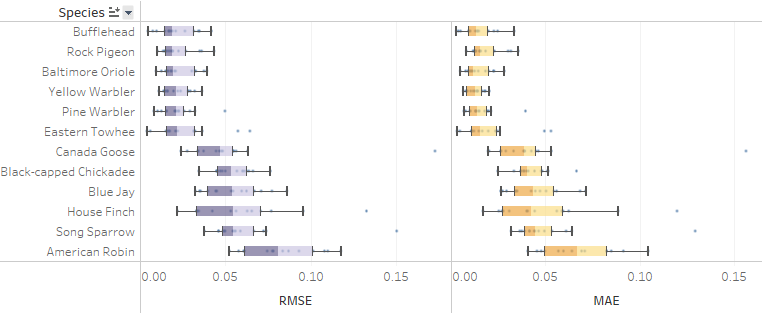
****

Figure A1. Comparison of RMSE, MAE among selected migratory and resident bird species

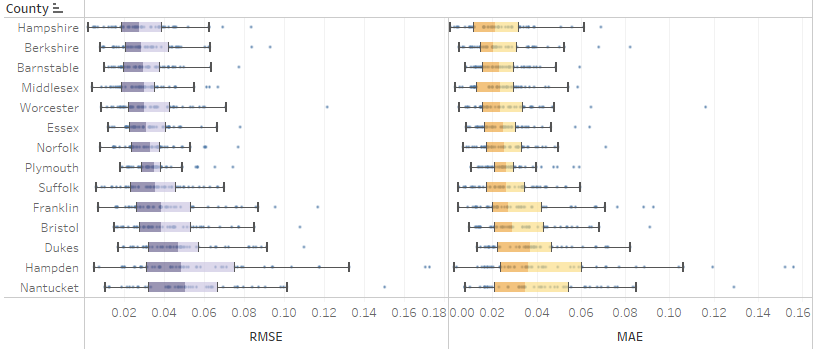
****

Figure A2. Comparison of RMSE, MAE across all counties

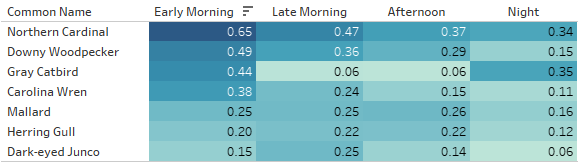


Figure A3. Diurnal Variations in Bird Activity

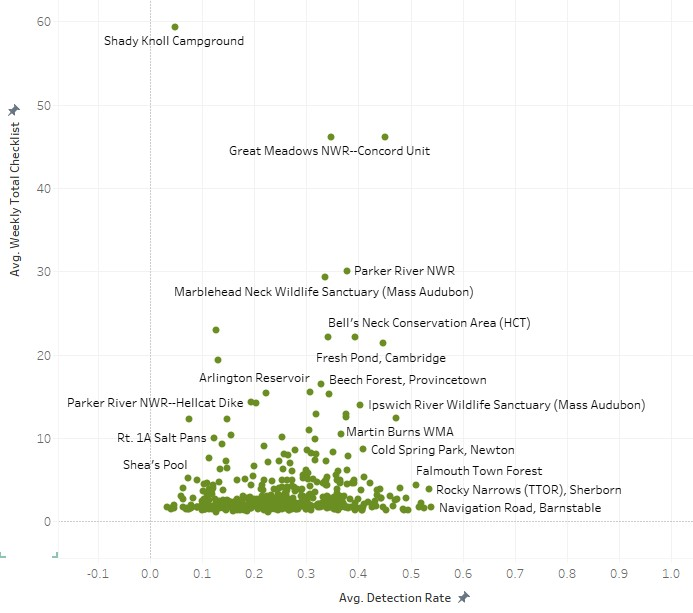
****

Figure A4. Average Weekly Number of Checklists and Average Detection Rate of Hotspots in Middlesex, Essex, and Hampshire in Spring and Summer (Week 10-39)