Spatial Hotspots and Seasonal Shifts in Juvenile Loggerhead Sea Turtle Telemetry Data (2002–2005)

GitHub: https://github.com/uilliam-c/Connolly_Liam_ENE872_FinalProject

Liam Connolly

Contents

List of Tables

List of Figures

1 Rationale

According to the online database, from which I collected my data, juvenile loggerhead turtles (Caretta caretta) move around frequently. Due to their juvenile phase of life, they are highly mobile yet, what isn't known is how or why they move to certain places in different seasons or ocean regions. From a spatial perspective, we don't have answers. Being able to identify where and when these turtles move/are found can help better understand their habitats. It would be interesting to discover "hotspots" of more concentrated turtle presence as that may guide additional ecological analyses in the future. With the data available, it should be possible to assemble a general sense of the spatial patterns these loggerheads exhibit.

The background of this dataset relies on using satellite tracking locations where each record present is indicative of a geo-referenced observation of a juvenile loggerhead. Each record comes with latitude/longitude, an individual identifier (series_id), and a timestamp (date_time). It is likely the dataset will include repeated observations from the same loggerhead over the course of the study (2002-2005), the data is still relevant for basic analyses. I will be focusing on examing seasonal distribution and spatial concentration for the loggerheads. Finally, the data contains a location "class" code (classitem) and some precision fields to help evaluate the patterns of data quality.

2 Research Questions

Research Topic 1) Seasonal distribution Shifts Research Question: Do juvenile loggerhead sea turtle locations (latitude) differ among seasons (Winter, Spring, Summer, Fall)? Hypothesis: Mean latitude differs among seasons. Rationale for hypothesis: By nature, seasonal changes in ocean conditions such as temperature, current direction, etc should influence where juvenile turtles are found. If this were true that juvenile loggerheads go to different places throughout the year, or if observation density varies by the different seasons in varying areas, then there should be differences in latitude among seasons.

Research Topic 2) Spatial Hotspots of Presence Research Question : Are loggerhead turtles locations concentrated in a few specific geographic areas ("Hotspots") or are they more evenly distributed? Hypothesis: The locations are spatially clustered, with a small amount of places accounting for a disproportionate amount of loggerhead observations. Rationale: It isn't uncommon for animals, when they find a favorable place to live with good food availability or water temperature, to appear repeatedly in the same area. Therefore, if I can discover repeated observations of loggerheads in the same types of areas, it should allow for the creation of measurable hotspots in a spatial setting.

Research Topic 3) Classitem Differences by Location Class Research Question: Does the distribution of classitem categories differ by season? Hypothesis: The variety of observations in each classitem category will differ among seasons (Winter, Spring, Summer, Fall), indicating that the mix of class item codes will also vary over time. Rationale: While examining the dataset, I came across the columns with a largely unexplained categorical code (I think its categorical) that wasn't mentioned anywhere in the public database. That code being classitem. It appears the classitem has some variablity between observations and that the recorded observations have no even distribution across the study. In theory, the data collected via satellite would be under different conditions including seasonal shifts in loggerhead behavior so, it is possible that the proportion of observations in each class item category differs among seasons.

3 Dataset Information

This dataset originates from OBIS-SEAMAP. Specifically, the dataset is called "SWFSC juvenile loggerhead sea turtle tracking 2002–2005" (Dataset ID 758; Version 1.0.1; internal DOI: 10.82144/4774e438). Actual URL is: https://seamap.env.duke.edu/dataset/758 The metadata explains in further detail that the source is satellite tracking data of the juvenile loggerhead turtle (Caretta caretta). The dataset has data from October 14, 2002 to July 7, 2005. Location coordinates are provided in decimal degrees (coordinate system: WGS84), and the spatial extent for the dataset is approximately 20.95–41.913° N latitude and 153.571–246.713° E longitude.

Each row in the points file is a single location observation for an individually separate turtle. Also included with an observation is the latitude and longitude (represented in decimal degrees) and observation date/time. There are some limitations as publically downloading the data limits me to core fields including date/time, latitude/longitude, and species information. It is unclear if the categories I can't download will influence the data analyses I conduct. Reported precision fields (lprecision and tprecision) were constant in this dataset, so precision-based comparisons were not informative.

Table 1: Dataset information.

Detail	Description
Data source	OBIS-SEAMAP
Retrieved from	https://seamap.env.duke.edu/dataset/758
Dataset title	SWFSC juvenile loggerhead sea turtle tracking 2002–2005
Dataset ID	758
DOI	https://doi.org/10.82144/4774e438
Date range	2002-10-14 to 2005-07-07
Coordinate system	WGS84; decimal degrees

Table 2: Data structure and Basic ranges for Key Variables Used.

Detail	Value
Number of observations	6861
Number of turtles	18
Start date-time	2002-10-14 18:26:57
End date-time	2005-07-07 18:35:28
Minimum latitude	20.95
Maximum latitude	41.913
Minimum longitude	-179.996
Maximum longitude	179.977
Median lprecision	4

4 Data Wrangling

I imported the points CSV into R-studio Online and converted the date_time column to date_time in UTC. I removed any records with missing latitude, longitude, or timestamp information. In this case, just one record was removed. Not knowing this in advance that only one record would be removed, I intended to ensure all analyses had a complete set of spatial observations recorded to remove outlier data. Additionally, I created a categorical season variable from the months included in date_time. Winter was defined as Dec–Feb, Spring = Mar–May, Summer = Jun–Aug, and Fall = Sep–Nov. Finally, I converted class_item to a categorical variable for later group comparison.

Table 3: Representing Missing Data with Key Variables.

Detail	Value
n_total	6861
missing_lat	0
missing_lon	0
missing_date_time	0
missing_lprecision	0
missing_classitem	0

5 Exploratory Analysis

This analysis of the data used gives a general summary of the spatial structure of the juvenile loggerhead turtle location dataset prior to testing my three hypotheses. First, I use maps of latitude and longitude to describe the geographic extent of observations and then I try to do a summary over time to get a sense of the patterns of observation across months and seasons. To inquire the spatial concentration for question 2, I tried to express a visualization of point density to identify whether sightings appear clustered into potential "hotspots." Additionally, I was curious to see if I could separate the observations between individual turtles (via series_id) to find repeated sightings. Finally, season-based visualizations provide descriptive context for seasonal latitude patterns (RQ1) and seasonal shifts in classitem composition (RQ3).

This table shows, after filtering, that there are no missing values within the key variables used for analysis (latitude, longitude, timestamp, precision, and class code). With this in mind, further plots and tests are being conducted off a complete dataset of 6861 location observations.

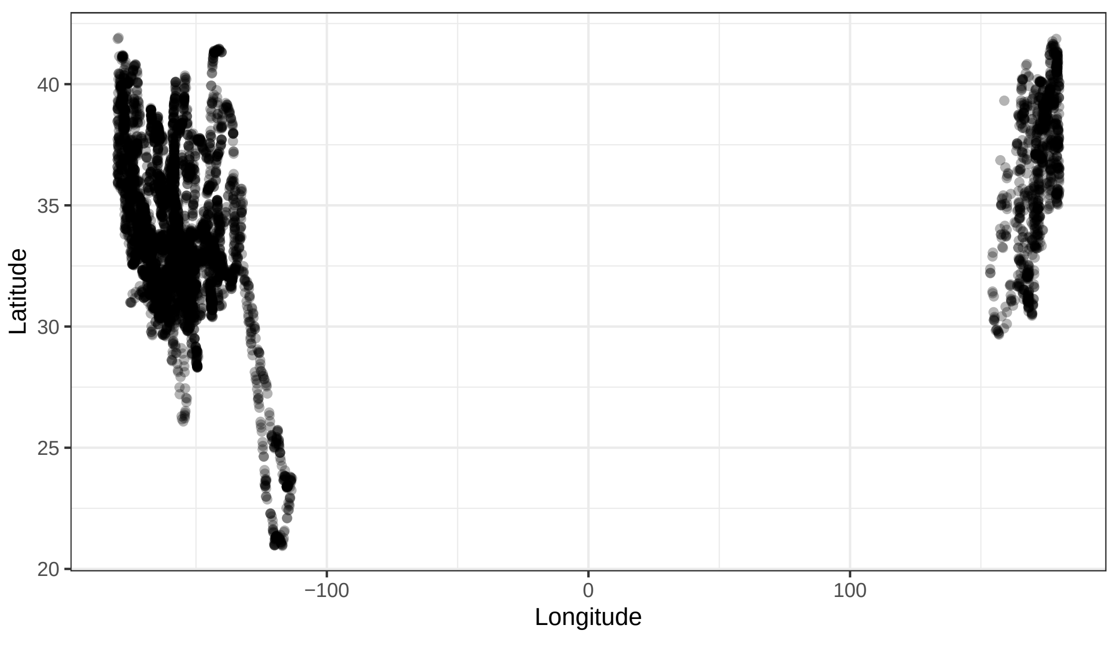

Figure 1: Spatial Patterns for Turtles.

Figure shows the geographic extent of juvenile loggerhead turtle locations within this dataset.

This Figure shows sampling intensity across the duration of the study period.

This figure shows latitude spread across Winter, Spring, Summer, and Fall.

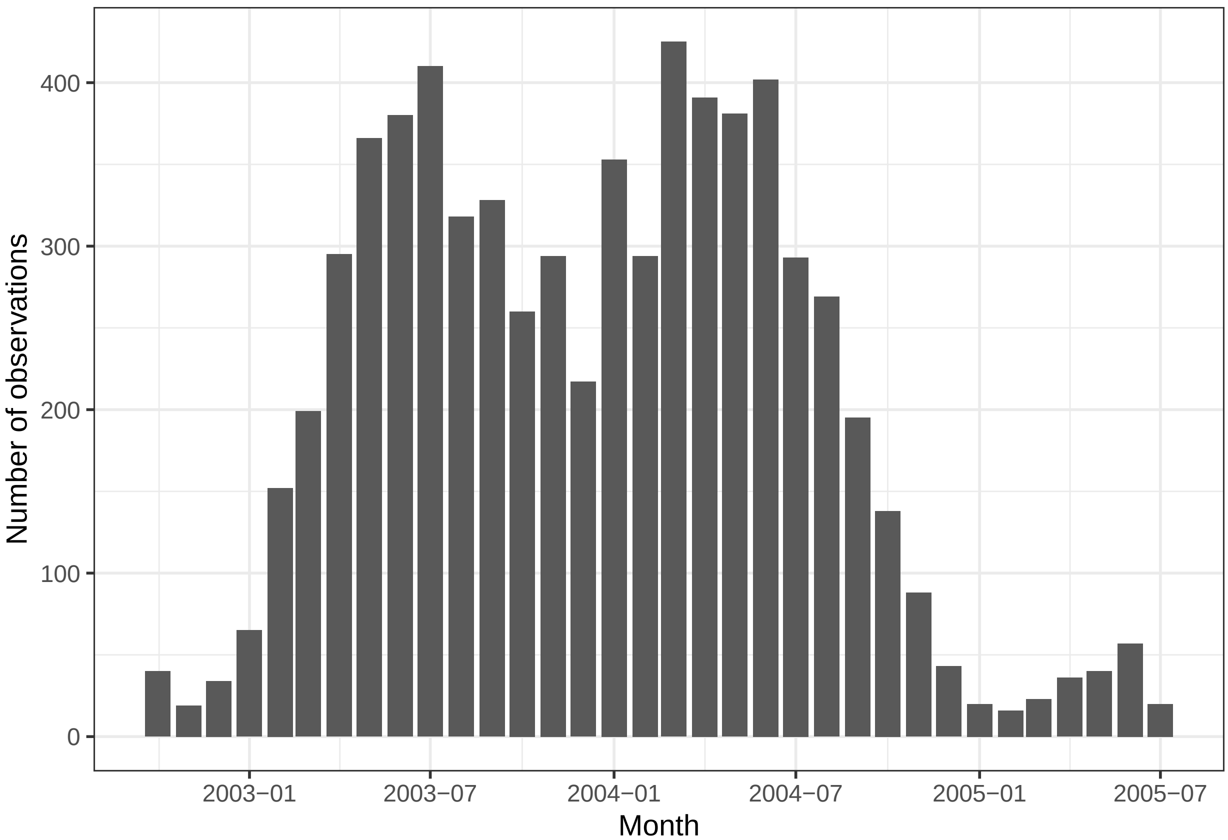

Figure 2: Number of location observations over time (by month).

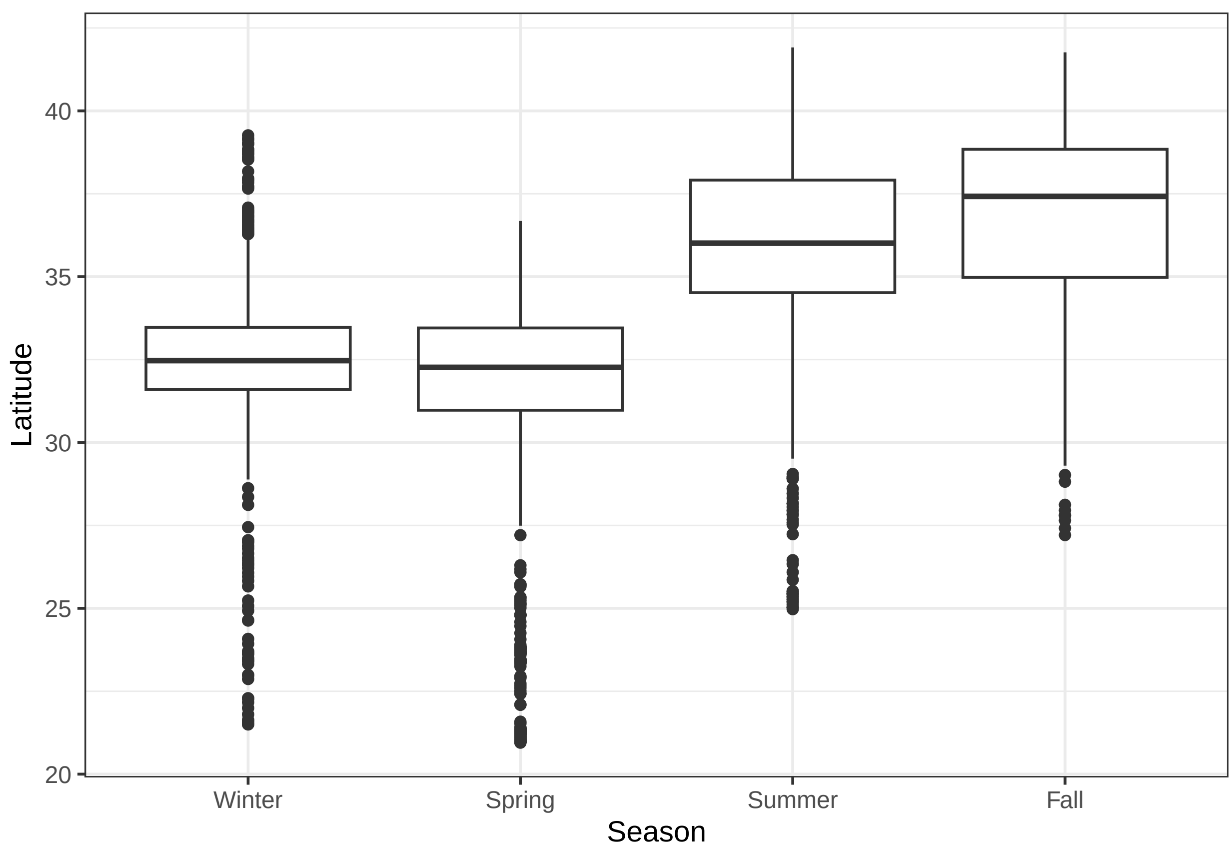

Figure 3: Latitude by season.

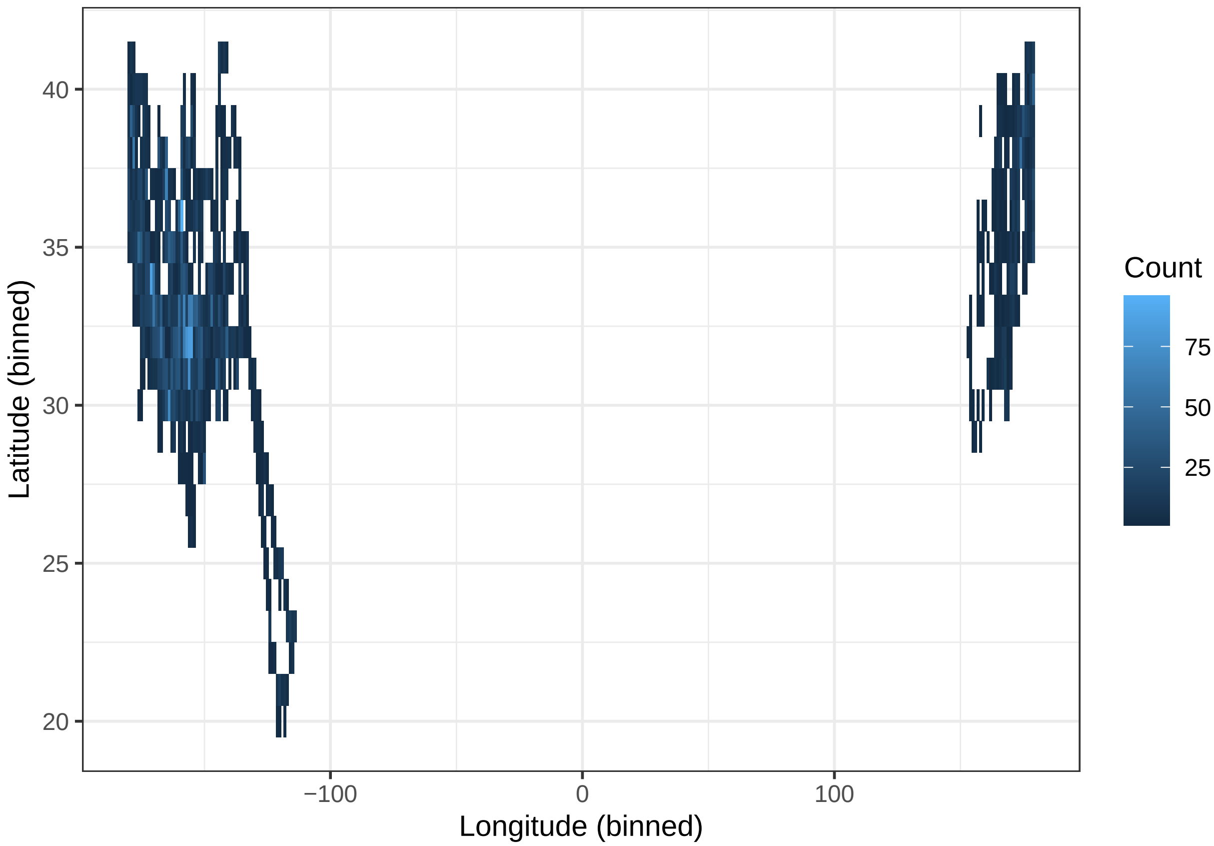

Figure 4: Exploratory point density (1-degree grid).

This figure inquires and suggests areas of concentrated observations that motivate the hotspot analysis for Q2.

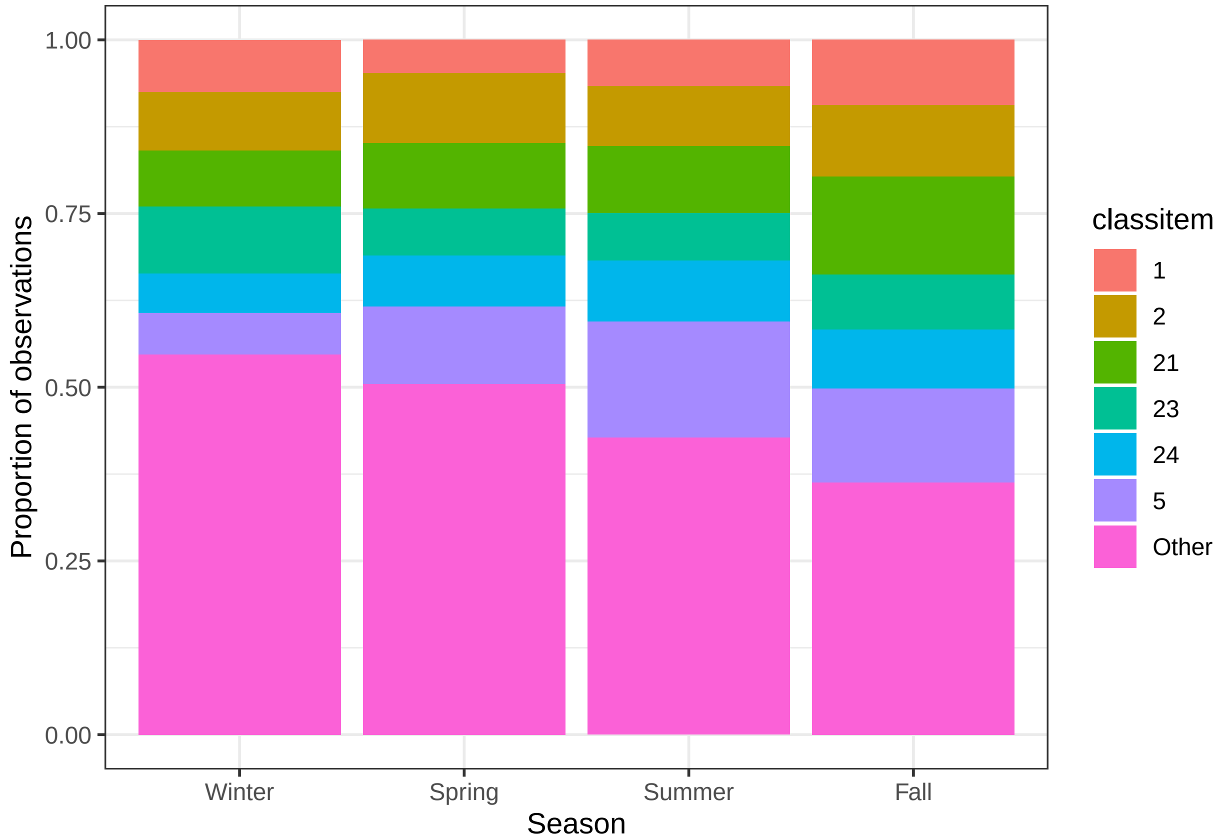

Figure 5: Proportion of observations in the most common classitem categories by season.

This figure shows the varying proportionate composition of classitem categories across different seasons. While many classitem categories are shown, a few make up most observations. The relative proportions appear similar between seasons.

6 Analysis

6.1 Question 1: Do juvenile loggerhead sea turtle locations (latitude) differ among seasons (Winter, Spring, Summer, Fall)?

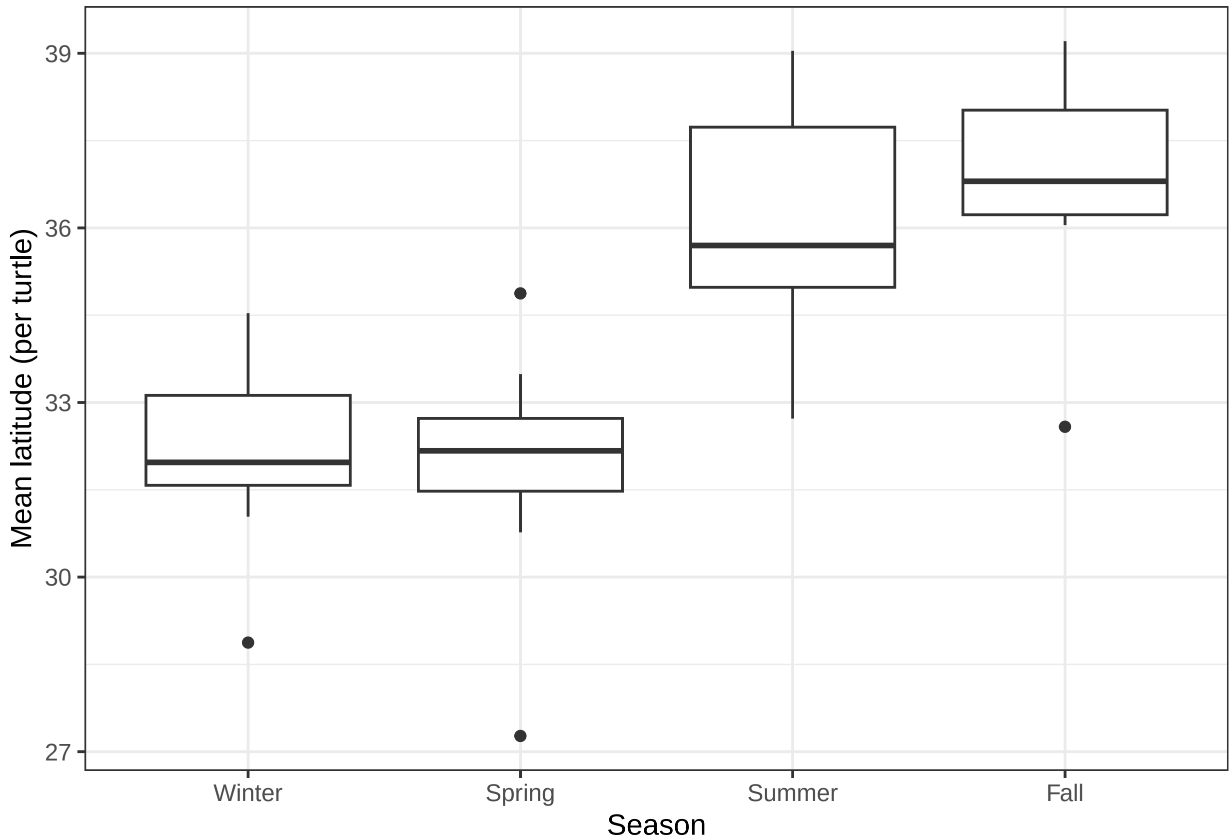

Figure 6: Mean latitude by season (summarized per turtle).

After being able to summarize latitude by individual turtle occurrences and season, mean latitude differed significantly among seasons (ANOVA; df = 3, 58; F = 38.3; p = 8.97 × 10^-14 or p<0.0001). This figure clearly shows seasonal separation in per-turtle mean latitude. With that said, these results support my hypothesis (seeking to solve Research Question 1) that juvenile loggerhead latitude varies by season in this dataset. Mean latitude was highest in Fall (36.89°N) and lowest in Spring (32.02°N). This is consistent with a seasonal shift in latitude.

6.2 Question 2: Are loggerhead turtles locations concentrated in a few specific geographic areas ("Hotspots") or are they more evenly distributed?

To evaluate RQ2 (Research Question 2), all loggerhead locations were summarized into 1-degree latitude–longitude grid cells and the number of observations per cell was counted. Under an even spatial distribution, counts would be similar across grid cells. However, hotspot clustering would be supported if a small subset of cells had a disproportionate share of the locations. In this case, the results show clear spatial concentration. The top 10 grid cells contained 11.2% of all locations. The highest-use cells had substantially higher counts of locations than the others. These findings support the RQ2 hypothesis I made that loggerhead turtle locations are clustered into a small number of hotspots rather than evenly distributed across the study area.

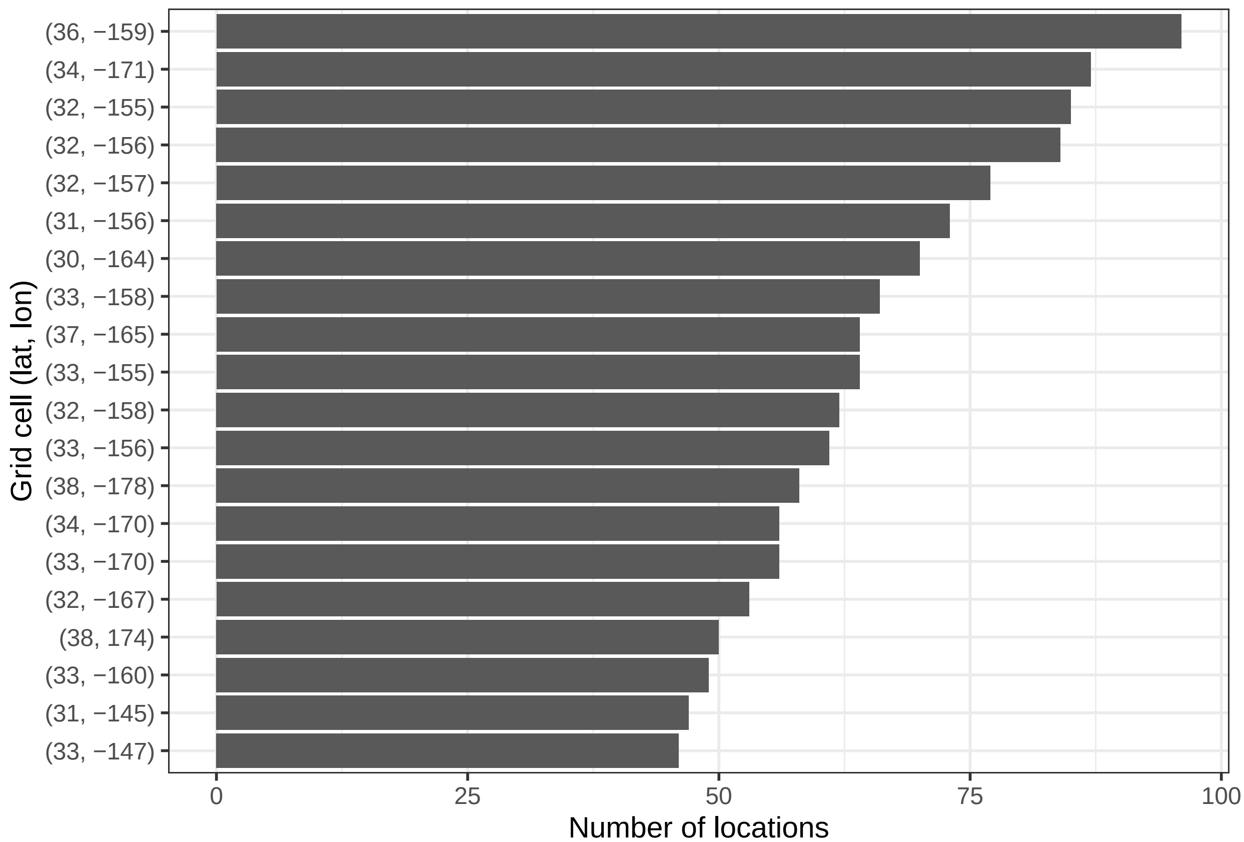

Figure 7: Top 20 most-used 1-degree grid cells (counts of turtle locations per grid cell).

Table 4: Top 10 Most-Used 1-Degree Grid cells.

lat_cell	lon_cell	n_cell
36	-159	96
34	-171	87
32	-155	85
32	-156	84
32	-157	77
31	-156	73
30	-164	70
33	-158	66
33	-155	64
37	-165	64

Given the possible data variability, as a data-quality check, longitudes were standardized to -180 to 180 where (min = -180.0, max = 180.0). Of those, 15.4% of locations observed occurred at positive longitudes.

6.3 Question 3: Does the distribution of classitem categories differ by season?

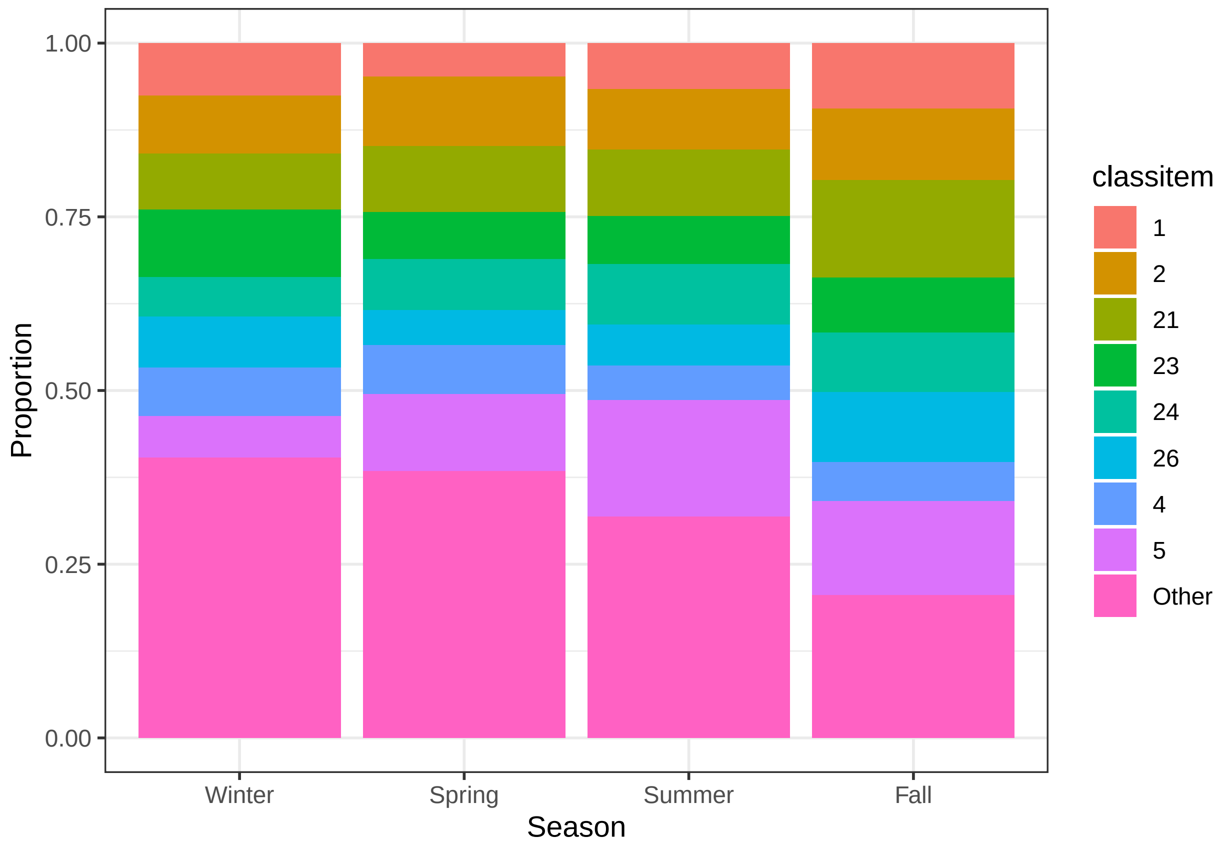

Figure 8: Classitem composition by season (top categories shown; remaining grouped as 'Other').

For RQ3 (Research Question 3), I compared the composition of classitem categories across seasons. Due to classitem containing many categories, I only displayed the most frequent categories and all remaining categories were noted as "Other". The proportion plot shows that classitem composition varied across seasons. For example, the "Other" category decreases from 0.404 in Winter to 0.206 in Fall. By contrast, while classitem 5 increases from 0.059 in Winter to 0.168 in Summer, Classitem 21 also increases toward Fall (0.080 in Winter to 0.141 in Fall). The tables also summarize the seasonal proportions for the categories and identify the categories with the largest seasonal changes (those being classitem 5 and 21). This analysis is descriptive and does not include a test of statistical significance.

In review, these results match my RQ3 hypothesis that classitem composition would differ by season. This is because multiple categories, as well as the non listed "Other" group, change noticeably across Winter, Spring, Summer, and Fall. However, it is important to note as I did above, that because this analysis is descriptive, the results support the hypothesis based on observed patterns. As I did not run a test of statistical significance, its possible for the results to be different.

Table 5: Proportion of observations in each classitem category by season (top categories + Other).

classitem2	Winter	Spring	Summer	Fall
1	0.075	0.048	0.067	0.094
2	0.084	0.100	0.087	0.103
21	0.080	0.095	0.096	0.141
23	0.097	0.068	0.068	0.079
24	0.057	0.073	0.087	0.085
26	0.074	0.051	0.059	0.101
4	0.070	0.071	0.050	0.057
5	0.059	0.111	0.168	0.135
Other	0.404	0.384	0.319	0.206

Table 6: Top 10 Classitem Categories with the Largest Seasonal Change in Proportion (top categories + Other).

classitem2	min_prop	max_prop	range_prop
Other	0.206	0.404	0.198
5	0.059	0.168	0.108
21	0.080	0.141	0.061
26	0.051	0.101	0.050
1	0.048	0.094	0.046
24	0.057	0.087	0.031
23	0.068	0.097	0.029
4	0.050	0.071	0.021
2	0.084	0.103	0.019

7 Summary and Conclusions

This dataset analyzed juvenile loggerhead sea turtle tracking locations between 2002–2005 to determine seasonal latitude patterns, spatial clustering into hotspots, and seasonal differences in classitem composition. Analyses performed in this report first cleaned the observations with longitude wrapping and season definitions (Winter, Spring, Summer, Fall).

RQ1: Does latitude differ among seasons?

Seasonal latitude differed strongly when analyzing the mean latitude per turtle per season. This was done to reduce overweighting of individuals with many observations to avoid unbalancing the data. An ANOVA of per-turtle seasonal mean latitude indicated that the season had a significant effect of ($F(3, 58) = 38.3$, $p = 8.97e\text{-}14$). The mean latitude was lowest in Spring (32.02) and Winter (32.16) and highest in Summer (36.12) and Fall (36.89). This is consistent with an observed seasonal movement north during Summer/Fall as relative to Winter/Spring.

RQ2: Are turtle locations clustered into geographic hotspots?

The spatial concentration was evaluated through aggregating locations into 1-degree latitude–longitude grid cells and counting observations per cell. The top 10 grid cells contained 11.2% of all locations and the highest-use cells had much larger counts than the rest of the cells. This assortment of concentrations supports the hypothesis that locations are clustered into a small number of hotspots rather than evenly distributed across the study area.

RQ3: Does classitem composition differ by season?

The seasonal differences in classitem were assessed descriptively using proportional composition by the different seasons. Due to classitem containing many categories and not being documented in the metadata, the most frequent categories were displayed and all remaining categories were labeled as "Other". The proportional composition varied across the different seasons. For example, the "Other" category decreased from 0.404 in Winter to 0.206 in Fall, while classitem 5 increased from 0.059 in Winter to 0.168 in Summer. Also, classitem 21 increased from 0.080 in Winter to 0.141 in Fall. These patterns are consistent with the hypothesis I made that classitem composition differs by season. However, this conclusion was drawn from proportional changes and not from a significance test.

7.1 Limitations and next steps

There were multiple limitations I encountered that definitely played a role in influencing my interpretation. First, sampling effort was not even between individuals and time. My first research question tried to reduce this bias through per-turtle summaries, but the hotspot counts can still reflect this variable sampling. Second, with classitem not being defined in the metadata, I had to interpret the results as differences in an undocumented categorical code rather than the specific behaviors they were meant to be. This is where only being able to download certain public data while other data was excluded created some issues. Third, hotspot identification depends on the grid's resolution (in this case, 1-degree cells); but using a finer resolution may have yielded better overall results. For the future, assuming there was access to the full dataset without variables such as classitem being excluded, it would be interesting to consider other spatial clustering analyses as well as more significance tests. For example, doing a chi square test might prove useful. Additionally, examining the "clearly defined behaviors" that were excluded from what I had access to could prove instrumental in learning more about the turtle's spatial location and movement throughout the seasons.

8 References

(paraphrased a few things in introduction) Howell, E.A., P.H. Dutton, J.J. Polovina, H. Bailey, D.M. Parker and G.H. Balazs. 2010. Oceanographic influences on the dive behavior of juvenile loggerhead turtles (Caretta caretta) in the North Pacific Ocean. Marine Biology. 157:1011-1026.