What's the Catch? Recreational Fishing Trends in North Carolina (1990-2019)

https://github.com/ardathdixon/Data_FinalProject

Ardath Dixon, Annie Harshbarger, Eva May

Spring 2021

Contents

1 Rationale and Research Questions **5**

2 Dataset Information **6**

2.1 Data Retrieval: 6

2.2 Data Wrangling: 6

2.2.1 MORE ON WRANGLING HERE 6

3 Exploratory Analysis **7**

4 Analysis **9**

4.1 Question 1: Are there trends in the amount of these fish caught over time? How do they compare? 9

4.2 Question 2: What could these trends look like in the future? 12

5 Summary and Conclusions **14**

5.1 Strong seasonal trends 14

5.2 Overall positive trend 14

5.3 Limitations 15

5.4 Future recommendations 15

6 References **16**

List of Tables

1 General Information About the Data Used 6
2 Total Catch Summaries (Number of Fish) 6
3 Number of missing values from NOAA MRIP data 7
4 Seasonal Mann Kendall Tests 11

List of Figures

1 Total Catch Patterns over Time 8
2 Seasonal and Trend Decomposition for All Fish Total Catch 9
3 Seasonal and Trend Decomposition for Bluefish Total Catch 10
4 Seasonal and Trend Decomposition for Black Sea Bass Total Catch 11
5 Holt-Winters Catch Forecasts 12

1 Rationale and Research Questions

As public awareness of increasing strains on ocean resources and organisms grows (e.g. see the reach of films such as Seaspiracy and Sonic Sea), more attention is being placed on understanding fishing patterns and the impacts they have on the oceans. Most of this attention is placed on commercial and industrial fishing operations, which are studied and managed by the federal agency NOAA, or the National Oceanic and Atmospheric Administration. However, there are fewer studies on recreational fishing, for which NOAA also aids in overseeing and collects data. For this study, we chose to investigate recreational fishing trends in North Carolina over a thirty-year period. The data, whose origins are discussed more below, initially included 27 variables, detailing information such as mode of fishing and wave (time period) the data was caught in. For this analysis, we wanted to look specifically at total catch during each wave, as we were running time series analyses during the course of the project. Therefore, we focused on only one catch estimation variable: total catch. We chose North Carolina due to our familiarity with species here, and we chose two popular recreational fishing species to investigate alongside all species combined. Trends in recreational fishing data can give us information about human behavior, species populations, and species movement patterns, which is why we found this topic interesting and wanted to investigate it further. We chose the following questions to guide our work:

1. Are there trends in the amount of these fish caught over time? How do they compare?
2. What could these trends look like in the future?

2 Dataset Information

2.1 Data Retrieval:

For this analysis, we used data collected during Marine Recreational Information Program (MRIP) surveys conducted by NOAA (**NOAA n.d.**). NOAA works with state and local partners to collect information on the species and number of fish caught by fishers via in-person communication, telephone surveys, and mail-in surveys. We retrieved the data using the NOAA Recreational Fisheries Statistics Queries "download query" tool (found **here**). We created three separate queries to download data: one for all species, one for bluefish (*Pomatomus saltatrix*), and one for black sea bass (*Centropristis striata*). For each query, we used the date range 1990-2019 and requested catch estimate data by "wave", or two-month period, for all waves, fishing modes, and areas of fishing for the state of North Carolina. We downloaded the CSVs as ZIP files and added them to our project repository.

All data and code for this project can be retrieved from the **GitHub repository**.

2.2 Data Wrangling:

We began our analysis by converting waves to months, in order to process the six annual waves using time series analyses. For NOAA fishing records, wave 1 represents January and February, wave 2 represents March and April, and this continues through the year. Therefore, we assigned wave 1 catches to the date of January 1, wave 2 catches to March 1, and beyond.

2.2.1 MORE ON WRANGLING HERE

Table 1: General Information About the Data Used

Detail	Description
Data Source	NOAA MRIP
Retrieved from	https://www.fisheries.noaa.gov/data-tools/recreational-fisheries-statistics-queries
Variables Used	Year, Wave, Total Catch (Number of fish), Mode, Area
Date Range	January 1990 - December 2019

Table 2: Total Catch Summaries (Number of Fish)

Summary Statistics	All Fish	Bluefish	Black Sea Bass
Minimum	11869	26	1168
Mean	12402954	1342064	411196
Median	11292146	1064369	313437
Maximum	34932698	5254124	1746847

3 Exploratory Analysis

Following initial wrangling, we checked the number of waves without catch records for each dataset by joining the existing data to a list of all possible waves between Wave 1 of 1990 (represented by 1990-01-01) and Wave 6 of 2019 (2019-11-01). The results of this exploration, which informed our approach for interpolation, can be found in Table 3.

Table 3: Number of missing values from NOAA MRIP data

Dataset	Number of missing values
All fish	11
Bluefish	17
Black sea bass	13

To fill the gaps with no data, we interpolated the likely values of missing time periods. This linear interpolation incorporated the catch numbers on either side chronologically. If Wave 1 of 1990 or wave 6 of 2019 was missing, we did not interpolate the value, as there would not be a measurement available on both sides. We graphed the total catch trends over time (with the newly interpolated values for missing periods) as shown in Figure 1. With this visualization, we could compare the three categories' recreational fishing catch patterns: all fish, bluefish, and black sea bass.

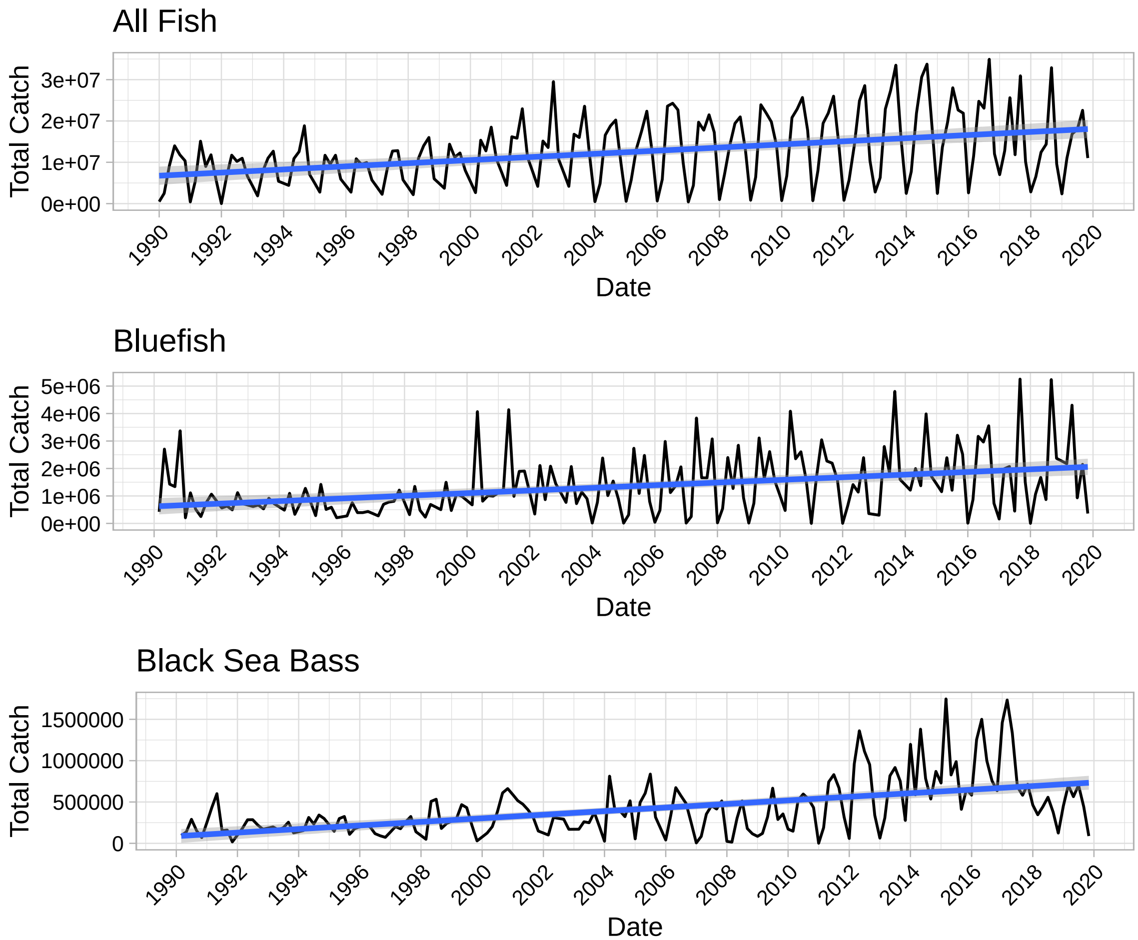

Figure 1: Total Catch Patterns over Time

4 Analysis

4.1 Question 1: Are there trends in the amount of these fish caught over time? How do they compare?

After our exploratory analysis and interpolation of the missing data points for each dataset, we created three time series for further analysis. We investigated the trends in total catch for all fish (Figure 2), bluefish (Figure 3), and black sea bass (Figure 4) by decomposing the time series into their seasonal, trend, and remainder components. For all three time series, we observed a strong seasonal trend with low catch totals in the winter and high catch totals in the summer. Furthermore, each trend component showed an apparent increase in total catch over time.

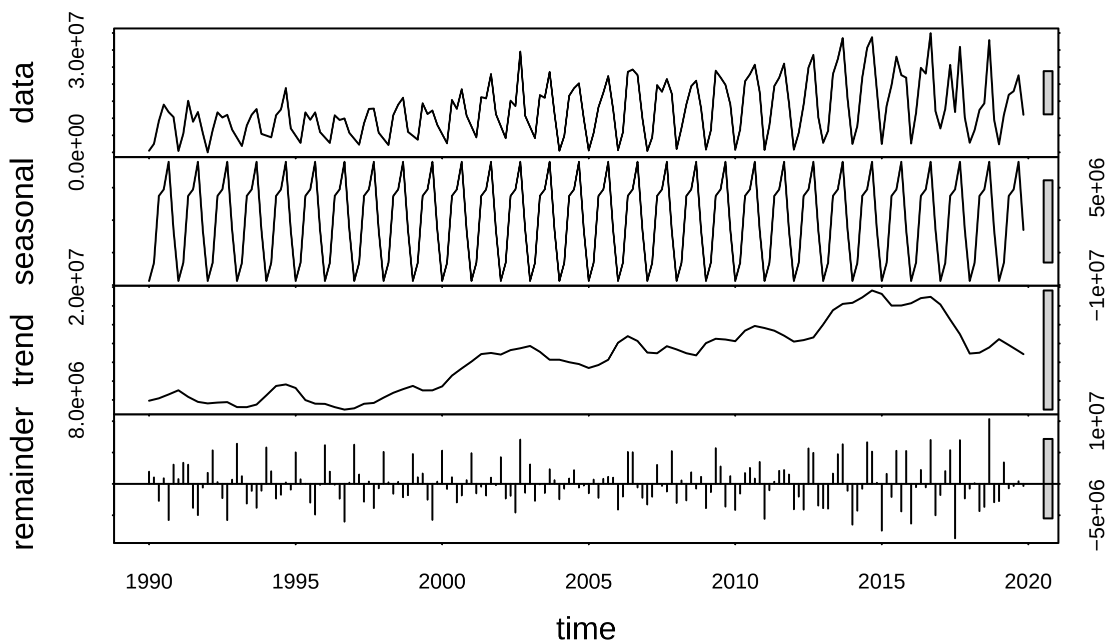

Figure 2: Seasonal and Trend Decomposition for All Fish Total Catch

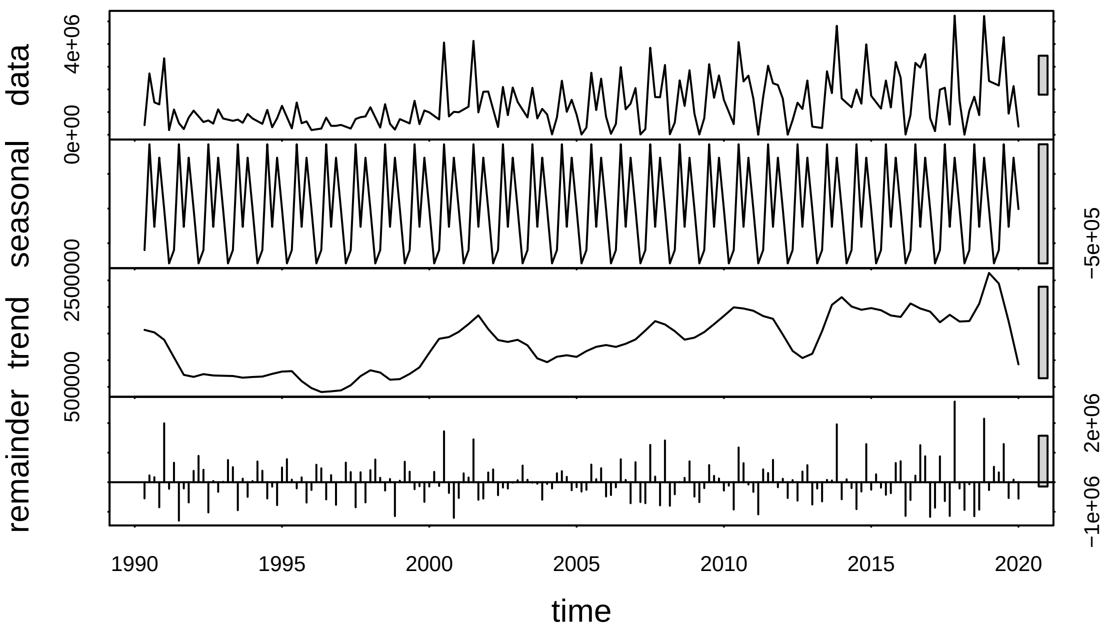

Figure 3: Seasonal and Trend Decomposition for Bluefish Total Catch

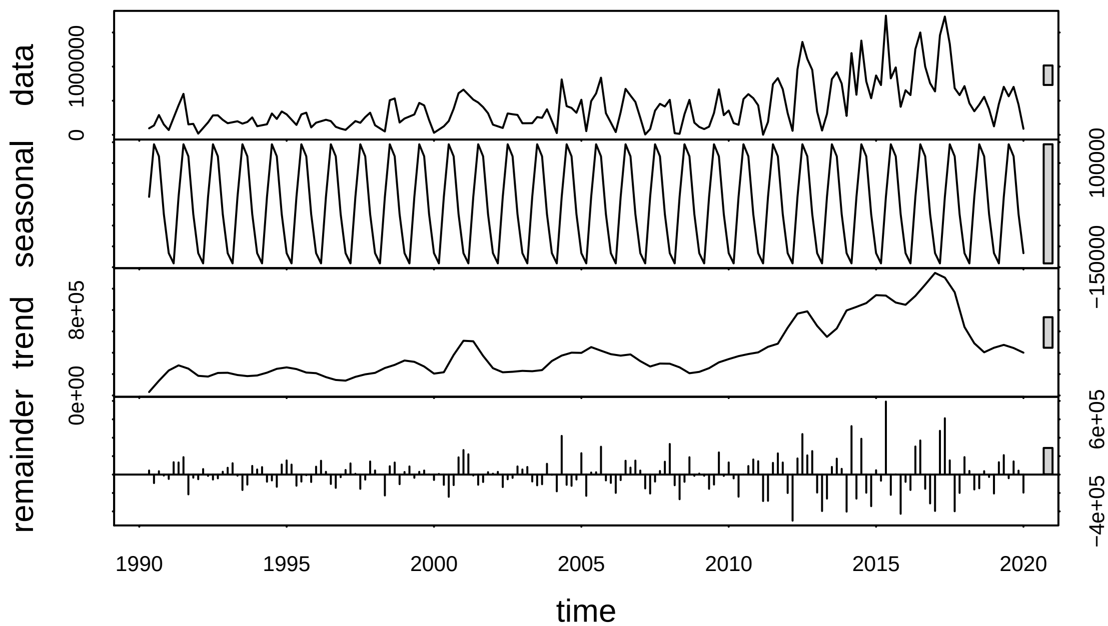

Figure 4: Seasonal and Trend Decomposition for Black Sea Bass Total Catch

Table 4: Seasonal Mann Kendall Tests

Fish Category	tau	2-Sided P-value
All Fish	0.4896552	0.000000e+00
Bluefish	0.3235180	8.748902e-10
Black Sea Bass	0.4095312	8.437695e-15

We ran a Seasonal Mann-Kendall test on each time series to test whether there was a monotonic trend in the total number of fish caught over time (Table 4). All three tests returned a statistically significant result (All fish: tau = 0.49, p < 2.22 x 10^-16; bluefish: tau = 0.32, p = 8.75 x 10^-10; black sea bass: tau = 0.41, p = 8.44 x 10^-15). Therefore, for all three time series, we reject the null hypothesis that there is no monotonic trend in favor of the alternative hypothesis that there is a trend in the data over time.

The strengths of these trends vary; catch for all fish has the strongest trend while bluefish has the weakest trend. Nonetheless, catch for both individual species and all species combined increases over the span of the data, between the beginning of 1990 and the end of 2019.

4.2 Question 2: What could these trends look like in the future?

To investigate future trends, we used forecasting in R to predict future data based on the existing past data we pulled from NOAA. There are several different methods through which to forecast, though only some of them account for seasonality, which was necessary here due to the seasonal components found in all three of our time series. We chose to use the Holt-Winters forecasting method for this data. Holt-Winters is more complex than simpler methods like naive but requires knowledge of fewer additional variables that other models, like SARIMA, need in order to run. Holt-Winters uses smooth exponentiating and varying weights of past data - with more recent data weighed more - to predict future data. Here, we predicted five years of data (where h = number of periods, and each period = 2 months). In the resulting plots, the dark purple area represents the 80% confidence interval level for predicted data, while the light purple area represents the 95% confidence interval level.

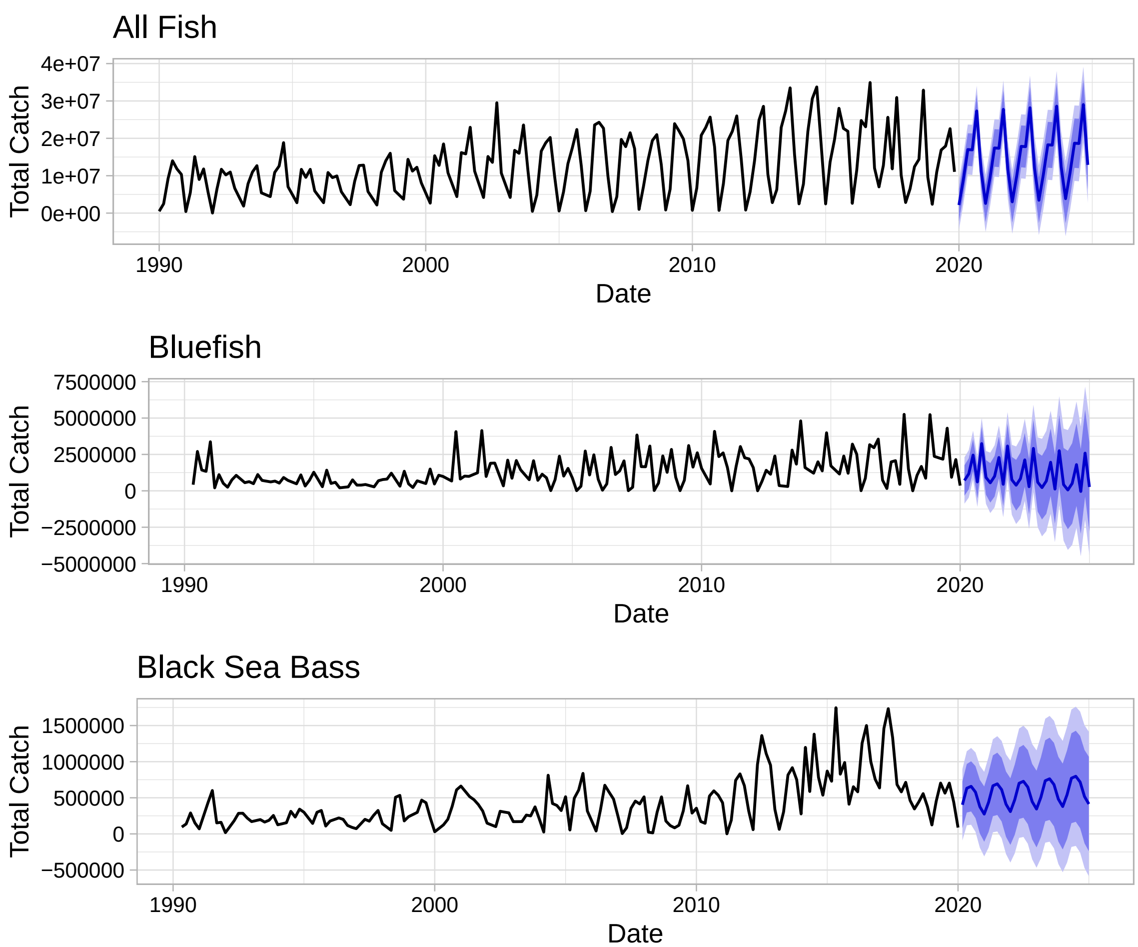

Figure 5: Holt-Winters Catch Forecasts

The Holt-Winters plots all show clear continued seasonality patterns and trends. For Black Sea Bass and all species combined, overall future trends are expected to be positive, like the past trends. For Bluefish, the forecasted trend is slightly negative – this is likely because the most recent years of bluefish data do show a slight negative slope, even though the trend is positive for the full twenty years in the dataset (though as noted above, bluefish has the smallest tau, and therefore the weakest trend). These forecasting plots are useful visualizations but should not be considered fully accurate because of these inherent complexities in fishing catch data that models like Holt-Winters cannot account for.

> *Insert visualizations and text describing your main analyses. Format your R chunks so that graphs are displayed but code and other output is not displayed. Instead, describe the results of any statistical tests in the main text (e.g., "Variable x was significantly different among y groups (ANOVA; df = 300, F = 5.55, p < 0.0001)"). Each paragraph, accompanied by one or more visualizations, should describe the major findings and how they relate to the question and hypotheses. Divide this section into subsections, one for each research question.*
>
> *Each figure should be accompanied by a caption, and each figure should be referenced within the text*

5 Summary and Conclusions

Summarize your major findings from your analyses in a few paragraphs. What conclusions do you draw from your findings? Relate your findings back to the original research questions and rationale.

5.1 Strong seasonal trends

NOAA marine recreational fishing catch totals for North Carolina show strong seasonal trends. Many more fish are caught in the summer, and much fewer fish are caught in the winter, as demonstrated above (Figure 1). The high seasonality for all three datasets analyzed was confirmed with the Seasonal Mann Kendall Tests, where all three P-values < 0.05 (Table 4).

This seasonality is likely influenced by recreational fishing patterns, where fishers are more likely to fish in the warm summer weather than the cool winter weather. Another potential cause for the seasonal trends is fish abundance and migration patterns, with higher populations of fish in North Carolina waters during the summer than during the winter. Total catch trends for all fish and Black Sea Bass showed unimodal peaks and valleys overall, while Bluefish showed bimodal trends (Figure 1). These bimodal Bluefish peaks could be due to their seasonal migration patterns (ASMFC 2021).

5.2 Overall positive trend

There was an increase in total catch of bluefish, black sea bass, and all species combined over time. The driver of this increase in recreational fishery landings is unknown, but it could be attributed to increased fishing effort, with more fishers participating in recreational fishing, or a change in methods or regulations that resulted in individual fishers catching more fish.

##I think it would be beneficial to add some references in this section – working on this! - Annie

Although the trend was generally positive for all three time series, it was not uniform; in other words, the rate of change varied over time, and there were brief periods where catch plateaued or decreased. This variation could be caused by changes in recreational fishing regulations over time, such as the increase or decrease of catch limits and size limits, or temporary area closures. Furthermore, recreational fishing is subject to all of the myriad factors that can influence behavior, including but not limited to climate variation and current events. In a recent example of how current events could impact recreational fishing that is beyond the scope of our analysis, which extended only to the end of 2019, the COVID-19 pandemic could have either increased recreational fishing in 2020, when people were spending more time outside, or decreased recreational fishing, as people were less able to travel to the coast.

5.3 Limitations

The MRIP system works very well for its intended purposes, but it is ultimately still a system based on estimation. Through MRIP, NOAA interviews only some fishers, then uses mathematical modeling to extrapolate on this collected survey data in order to create statewide (or area-wide, depending on the dataset) estimates. While MRIP is the best source of recreational fishing catch data, there is always room for some error in its estimations.

Within each of our datasets, there were some missing values. These values were not side by side or frequent, so we chose to linearly (*?*) interpolate our data in order to fill them in. This helped our figures and analyses (e.g. Mann-Kendall) to appear and run cleaner, but interpolating has some drawbacks. While interpolations seemed to follow the clear seasonal pattern in each dataset, most often the interpolated data was in wave periods that were minimums in other years of the datasets, so our interpolated values may have been a bit higher than actual catch rates during those times. Interpolation is an estimator for missing data, and it is important to acknowledge that our interpolated data may not be representative of true values.

Finally, our forecasted data are much cleaner and less noisy than our existing data. Noise in these datasets comes from external factors like changes in fish catch limits or weather patterns, which this forecasting method cannot take into account or predict. Our forecasting outputs are therefore limited by the relatively simple methodology we chose to employ. Holt-Winters remains a popular forecasting method and is reasonable to use for our purposes here, but the inherent uncertainty in this predicted data is important to acknowledge.

5.4 Future recommendations

- Comparisons of other species or other states
- Catch per unit effort
- Include earlier data

6 References

<add references here if relevant, otherwise delete this section>

National Oceanic and Atmospheric Administration. (n.d.). "About the Marine Recreational Information Program". Retrieved from: https://www.fisheries.noaa.gov/recreational-fishing-data/about-marine-recreational-information-program, accessed 25 Apr 2021.