Final Project - Proposal

Aaron Shavitz

Modeling drivers of *Zostera marina* (eelgrass) decline at Duck Harbor

For my project, I intend to model changes in eelgrass percent cover from Duck Harbor in Wellfleet, MA. This is an offshoot of my current work with the Massachusetts Bays National Estuary Partnership, in which I am examining potential drivers of eelgrass decline near Wellfleet. For this work, I have been collecting, organizing, and visualizing various environmental time series data sets. My final project will be informed by this analysis, but will be a distinct exercise, as my current work with MassBays does not involve modeling percent cover changes. If time allows, the final project for 607 can also be expanded to other areas along the coast of MA for which eelgrass percent cover and associated data is available.

Primary Research Question

In my current work with MassBays, the primary research question is: What are potential driving factors of eelgrass decline observed near Wellfleet, MA between 1995 and present? As the scope of this exercise only looks to survey environmental time series data, the analysis is lacking the establishment of a predictive model that relates potential driving factors (variables) of eelgrass decline to metrics associated with the health of an eelgrass meadow, or how the interaction of these variables may impact these health metrics.

One potential health metric for eelgrass is percent cover. At Duck Harbor, the National Park Service (NPS) has recorded percent cover and other metrics for a series of 3 transects during peak summer eelgrass biomass since 2003. At first glance of the time series, there appears to be a decline in average eelgrass percent coverage across the transects.

Thus my initial research question is: **Can eelgrass percent cover be effectively predicted by associated environmental variables?**

Of course, I do not intend to identify causality. However, if it can be shown that parameters such as sea surface temperature, salinity, wave/current direction & intensity, boat activity, or any other of a host of other variables can be used to effectively predict percent cover, then the

most influential of these factors can be prioritized for further examination in other areas with remaining eelgrass beds.

Data Sources:

The primary percent cover data I intend to use can be found from the Duck Harbor page on seagrass.net. This data only goes through 2017, however I have direct access to data through 2023 from NPS.

Various water quality data from a sampling station off the coast of Wellfleet is available from the Center for Coastal Studies - Cape Cod Water Quality Station 5S (station 2 in the data download). This data is available from 2006 - 2023.

Other data sets still need to be determined. The seagrass.net page fro Duck Harbor says that "Monitoring results reveal both spatial and temporal heterogeneity in the eelgrass population in response to physical wave disturbance" however they do not provide data on this physical wave disturbance. The Coastal Data Information Program (CDIP) has a buoy in Cape Cod Bay (Buoy CDIP 221), with wave and current (and other) data available via ERDDAP, but the time series is too short (only started collecting data in 2016, and some variables not until 2022). The NERACOOS A01 buoy in Massachusetts Bay has wave and current data available, however it's unclear if that data would be relevant given the distance to Duck Harbor.

I may be able to get some vessel traffic related data from Northeast Ocean Data Center.

Framework and Analyses

My intent is to put together summertime and/or annual summary statistics of variables shown to have changed over time along the coast of MA, based in part on my work with Mass Bays. As my response variable for any model(s) I intend to build is percent cover, for which I will only have annual values, I intend to build a data frame with annually measured percent cover values and all other variables I may consider for analysis. From there I will likely go through a series of visualizations examining potential correlations and interactions, prior to starting to build models and check assumptions.

Regarding analysis, I expect I am likely to use a general linear with beta distribution, as we discussed in class that beta distributions are useful for bounded data such as percent cover. I am not yet clear on best practice, but I imagine the building of this model will be iterative, either by including or removing variables and interactions in subsequent stages of creating an optimized and neat model. (I doubt it will ever really be neat).

Depending on the results of my model, and time required to complete this, I could consider collecting similar data for another NPS eelgrass monitoring station, and test the ability of my initial model to predict percent cover.

Concerns...

Once concern I have is that since I think I'll be working with just annual measurements, I won't have a large data set. I have an itching feeling that there is a different approach I should take, and that I'm just focused on time series annual data because it's what I've primarily been working with this semester...