NOAA nearshore fish atlas project proposal

Wendel Raymond

September 3, 2018

Data

Data was obtained from the NOAA nearshore fish atlas (Johnson et al 2012). This data set compiles 555 beach seines across Alaska in 4 common nearshore habitats; eelgrass/seagrass, kelp, bedrock, and sand.

Structure

The data set is structured as every row being a unique fish with measured fork length in mm ('ForkLength'). Date, site, gear type, location, habitat, temperature, salinity, and other information is also reported.

```
##  [1] "Data_Events_SiteID" "EventID"            "Date"
##  [4] "Season"             "Mon"               "SeasonNoYear"
##  [7] "Year"               "Gear"              "Temp"
## [10] "Salinity"           "SpCode"            "LifeStage"
## [13] "ForkLength"         "Unmeasured"        "AtlasID"
## [16] "Data_Sites_SiteID"  "Region"            "Locale"
## [19] "Location"           "SubLocale"         "Nickname"
## [22] "Habitat"            "Habitat_LC"        "HabitatVeg"
## [25] "Lat1"               "Long1"
```

Current analytical status

I used a subset of this data for my class project in FISH 631 and so have some existing code and understanding of the data. In that project I investigated the effect of habitat type on fish community composition in beach seines from southeast Alaska. That project found that dispersion among the four habitat groups was significant indicating that variation in fish community differs among habitat types (F = 48.01, p-value < 0.0001). PERMANOVA results indicated that fish community composition was significantly different among habitats (psudo-F = 17.04, p-value < 0.001).

Since that project I have also played around this with data set a bit looking as salmonid use of eelgrass/seagrass habitat. I feel fairly confident that my familiarity with the data and with the code I have already written that I can work up other analyses fairly easily. One major hurdle if I/we decide on using fork length data is that fish that were not directly measured (i.e. just counted because they already measured 50 from that seine) need to be assigned lengths. I have done this with my dissertation data and in theory should not be difficult to adapt the code to this data set (they are structured very similarly)

Case Study Goal

I would like to build on what I did for FISH 631 to build a more robust model for fish abundance/community composition that accounts for season, latitude, and other factors. Perhaps PDO, NPGO SST, etc. once approach would be to focus effort on certain species, however, I am partial to the community approach. So essentially the question remains the same; "how does fish community composition differ across habitat types while accounting for the above?". Also, while the full data set contains data from all over Alaska, I plan on restricting analysis to southeast Alaska.

Where I could benefit from help

1. Data wrangling and model building that accounts for space and time.
2. Obtaining data on SST, PDO, NPGO etc.
3. Group think on global model creation. Whats important to consider? Whats not? What am I not thinking about?
4. What do people know about ecological niche modeling? Could this data be used in that context?