1. General introduction
   1. Background
      1. Background – research topic (nearshore fish communities)
         1. Overall goal paragraph
            1. The overall goal is to understand natural distribution of Alaskan nearshore fish communities
         2. Realm of the nearshore: see marine bio slides –
            1. define nearshore generally and why it is distinct realm

defined by habitat diversity

includes marine and estuarine settings

connectivity of habitats via biological migration and interacting physical forces

coastal processes are often tied to both terrestrial and oceanic forces

* + - * 1. why it’s important

in the ecosystem

clarify specifically the zone within the nearshore this work will focus on, the ‘inshore’ (intertidal to near-subtidal zone) – typically less than 10m depth

relatively shallow depths lead to high production and resource use

littoral zone

nursery habitats

Connectivity of habitats leads to many taxa using this zone as a migration corridor among functions uses

Eg, diadromy, rearing behaviors

and human access aspect – for impact and study

coastal communities among the most populous for humanity

anthropological uses

development = habitat alteration/degradation

restoration projects

resource extraction

varying levels of management and conservation practices

Also equates to some of the most well-studied systems glopbally

* + - 1. Fish communities
         1. For our purposes, defining a community as:

Ecological definition of a theoretical community:

Working under the assumption that fishes within a small area (i.e., sampling area/sample unit) all individuals can potentially interact with one another

Area we’re interested in: **theoretical** community defined by ‘nearshore’ - (e.g., 200m depth or less) OR inshore (as defined above)

In practice: defined by a gear type – such as beach seine, trawl, etc. and the specific net dimensions (or whatever) used, which limits zone and community sampled (subsets of community are sampled) that vary with depth, SAV, etc.

* + - * 1. Not including ichthyofauna (for this study) because….
        2. Use of nearshore habitat
      1. Alaska
         1. Arctic / subarctic / high latitude systems
         2. Prevalence of anadromous fishes (culturally and otherwise)
         3. Some well-studied regions/locations, such as Kachemak Bay, Lynn Canal, PWS

Studies often are either long term, but lack seasonal resolution, or have high temporal resolution (data throughout the year), but are generally shorter and for single locations

Some studies covered large spatial distribution, but more limited in temporal span (only sampled a few summers over hundreds of KM)

* + - 1. Literature dive / state of the science (nearshore fishes globally)
         1. Inshore fishes (specifically)
         2. Alaskan studies
         3. Recent studies
    1. Background - ecological themes
       1. Defining regional species pool for the inshore (nearshore)
          1. Ie, Alaska coastline and its subregions
       2. Community structuring processes
          1. Variability of scale
          2. Spatially and temporally
          3. Within community processes

Predator-prey interactions

Competition of resources

Inter and intra species avoidance/grouping behaviors

One of the harder to study aspects is the fine-scale timing of interactions and habitat uses by populations

* + - 1. Idea of gradients – true vs false gradients (Legendre)
         1. True (induced spatial dependence): spatial structure tied to external variables, like coastal formations and physiochemical conditions
         2. False (spatial autocorrelation): spatial structure defined by within community processes (see 2c)
      2. Spatial scale
         1. Grain size (size of sampling unit)
         2. Interval / lag (distance between units)
         3. Extent / range (total study area)
    1. Justification of research topic
       1. Community data have not been holistically studied across the state – there are many small scale, local, and semi-regional studies that we base understanding on
          1. Ie, PWS or SEAK wide, Kbay vs Lynn Canal
       2. Communities are important ecologically, home to primarily juvenile and/or YOY lifestage fishes, as well as migrating adults (including migration to spawning grounds)
          1. Many function guilds depend on coastal marine/esutuarine habitats
          2. Additionally, opportunistic feeders and estuarine residents / shallow subtidal residents are found together in these communities
       3. Practically, a lot of research is focused on species of interest (those of commercial/economic/etc value) – those with detailed management plans
          1. At the same time, many within the nearshore community do not have species-specific regulations or management plans due to limited knowledge of life history and seasonal functions
          2. Juvenile, larval, and YOY life stages for managed species are understudied relative to adult life stages (harvestable life stages)

Since juveniles do not typically have commercial value, there are few structures in place to gather data (outside of academically oriented research)

* + - 1. Aggregating many smaller-scaled projects (eg, species-specific or location based) should lead to insights on nearshore fish communities in a broad sense (inference of global population)
  1. General research questions
     1. Is pulling all these seines worth it?
        1. With what we know about our regional species pool and our surrounding environment, can we accurately describe / predict a fish community assemblage for a given time and place?
           1. i.e., can we model it?
        2. In terms of management/regulations, are there practices that can be applied to the community as a whole or indirectly via habitat or seasonal considerations
        3. Can we model it using a more specific question – is one of these more appropriate?
           1. Spatially?
           2. Temporally?
           3. By species / genus / family?
           4. By habitat?
           5. Regionally (SE vs SC, etc)?
           6. Local vs. larger processes?
        4. Does it scale up? (aggregate)
           1. All (?) majority of Alaskan studies report that a combination of local scale (define) and seasonal structures exist in these communities (often with an interaction effect); However, we cannot statistically differentiate between true and false gradients using studies on hand.
           2. Does the local (i.e., site specific) differences due to a larger physical gradient
           3. If not, what is an appropriate sampling level to look at?
     2. Can we discern biogeographic breaks in overall fish community structure throughout Alaska? (Chapter II)
        1. Are these caused by true or false gradients (unknowable mixture of the two)?
        2. Is this possible to answer with the currently available data?
           1. If not, what would be required?
     3. Does community structure align with known regime shifts throughout Alaska? (Chapter III)
        1. Shrimp to cod
        2. See Piatt research and Robards et al. 1999 references
        3. Eg, GOA ‘blob’ – see Arimitsu et al. 2023 and others
  2. Approach
     1. Applying common themes in ecology to data set
        1. Community data (sites, times, species, spatiotemporal variability, complex structuring, etc)
     2. The data - nearshore fish
     3. Spatial analysis – discerning what spatial structures can be teased out

1. Chapter 1
   1. See published paper!
2. Chapter 2
   1. Introduction
      1. Understanding spatial distributions in nearshore fish communities of AK using the nearshore fish database
      2. Question : is there a discernable biogeographic break along GOA coast? - where we have good data (high(est) resolution spatially and temporally of data)
         1. If so, where? (And, what drives the break?)
         2. If not, why? (because no break, or because data inadequate?)
   2. Method
      1. Overall : nearshore fish data -> wrangled to site x species (community comp) -> decided transformation / standardization of data -> community data matrix (distance or similarity – resemblance matrix) -> cluster analysis OR ordination OR spatial analysis (dbMEM)
      2. Analysis of community compositional data
         1. Sites (rows) x Species (cols)
      3. Data from NOAA’s Nearshore Fish Atlas Database
         1. Subset for collections by beach seines
         2. Lots of different projects producing similar data
            1. But not quite the same (eg, research questions, sampling design, gear type, etc)
         3. Able to be linked to ShoreZone habitat data
      4. Complications
         1. Lots of zeros in data
            1. Leads to double zero paradox
            2. due to sampling inefficiency or random species occurrences across long env gradients
         2. Varied sampling intensities
            1. Samples range in number of seines sets from 1 to 8
            2. Gear differences in mesh size
            3. Local conditions (habitat) targeted differently depending on project goal
         3. Likelihood of missing species
            1. Each replicate is unlikely to contain all possible taxa
            2. More sets will be closer to theoretical richness

But abundances will favor species most susceptible to capture method

* + 1. Solutions:
       1. Deal with excessive zeros!
          1. Beals smoothing

see vegan::beals() for explanation of uses

Use vegan::stepaccross() – similar to Swan’s version of Beal’s smoothing

For use when all else fails – when beta diversity is very high and a large proportion of dissimilarities are at upper limit

Not sure how to implement with abundance data – see ‘type’ explanation in beals()

* + - * 1. Should only consider asymmetrical resemblance indices
      1. Use of transformation formulas on raw data
         1. Allows for input of data into linear methods, eg, PCA RDA, k-means
         2. See figs 7.7 and 7.8 in numerical ecology
      2. Also use of distance matrix (distance coefficients) to compare sites
         1. Hellinger distance, Chi square distance, robust Aitchison distance, Bray-Curtis distance
    1. Distance based morans eigenvector maps (dbmem 7.4)
    2. Try quickMEM() page 327
       1. Need to understand ‘orthagonal’ and ‘detrended’ response variables

1. Chapter 3
   1. Introduction
      1. Temporal distributions over interannual (decadal) time periods
      2. Kachemak Bay has local long term monitoring, environmental data going back to 1950s, ROMS models that can hindcast data gaps
      3. Comparative studies from 1970s, 1990s, and recent – ask similar questions and build on them.
         1. Has there been a shift in community composition from 1990s to present? How does the shift compare to the shift from 1970s to 1990s?
         2. What were general oceanic environmental trends? (regime shifts)
            1. Shellfish->finfish
            2. ‘the blob’
   2. Method
      1. Borrow from 1990s paper (Robarts)
      2. Literature review for regime shifts
      3. Similar options for analyzing community shifts to chapter 2
         1. Cluster
         2. Ordination
         3. Modelling? Mixed model (glm or gam) where each period is modelled and then models are combined (70s, 90s, present)