Aquaculture, Eelgrass, and Fish

Chris Comiskey

October 19, 2016

Project Information

Project No.: C1617-3
Client: M.S. student
Department: CEOAS - Marine Resource Management
Statistical Background: ST 511, ST 512
Major Professor: Flaxen Conway
Consultants: Chris Comiskey
Approximate Timeline: Defend June '17

1 Background and Objectives

Business interests drive shellfish aquaculture developments such as oyster farming, in estuaries like Tillamook Bay in Oregon, and Humboldt Bay in Northern California. Ecological concerns, on the other hand, focus on the importance of eelgrass habitats to underwater species abundance and diversity in these same habitats. Government regulatory bodies would like to understand the impact of oyster farming on eelgrass habitat and on fish abundance and diversity. This research explores the relationship, in two estuaries, between shellfish aquaculture developments; several measures of eelgrass habitat complexity; and fish species abundance, diversity, and activity.

2 Project Description

Estuaries exist where freshwater meets saltwater—where rivers transition to the sea. This research involves two estuaries—Tillamook Bay in Oregon and Humboldt Bay in California—and three eelgrass sites within each estuary. These estuary-site combinations were chosen based on similarities across site, within estuary, of tidal elevations; eelgrass habitat characteristics; oyster farm technique and layout; as well as availability of the sites.

Each site consists of three habitat types: eelgrass bed, shellfish aquaculture and edge, which is a transition between eelgrass bed and shellfish aquaculture. Essentially, we can think of these habitat types as areas of high (eelgrass bed), medium (edge), and low (shellfish aquaculture) eelgrass concentrations.

At each site, your research team placed a 50 meter transect across the middle of each habitat type, denoted by the dotted lines in Figure 1. The transect plays a key role, as described below in the study structure and data collection methods. Figure 1 represents one site, of which there are three in each bay.

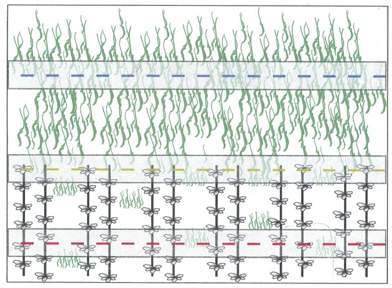

Figure 1: Three eelgrass habitat types in one site. From top to bottom, in generally descending order of eelgrass density: eelgrass beds, edge, shellfish aquaculture. The dotted lines represent transects that facilitate measuring eelgrass habitat complexity; species abundance, diversity, and activity; and predation pressure. The black lines in the shellfish aquaculture represent oyster farm seed lines.

3 Sampling Methods and Data Collection

Data for each site required a full day of effort, and therefore, all sampling was completed over the course of six days. In what follows, we categorize the types of data you collected into three groups—eelgrass habitat complexity; species abundance and activity; and predation pressure.

3.1 Eelgrass Habitat Complexity

At low tide on a given sampling day, your research team placed 20 quadrats along the transect within each habitat type at the site being sampled that day. The locations of these transects were randomly selected ahead of time. Inside each numbered quadrat your team visually assessed and recorded percent of eelgrass cover, shoot density, percent of epiphyte cover, and percent of microalgae cover. In addition, one eelgrass shoot was removed from the middle (approximately) of the quadrat—or haphazardly from nearby if no shoots were present in the quadrat—and it was stored for later laboratory processing and analysis.

3.2 Species Abundance and Activity

Also at low tide on the given sampling day, your research team placed three cinder block mounts at 5m, 25m, and 45m along each 50 meter transect, 4m perpendicularly from the transect, alternating sides. One hour before high tide on the same day, snorkelers deployed two GoPros at each cinder block location, one facing down, one facing up. The units were then retrieved one hour after high tide. The plan for the video data is to measure the abundance and activity of roughly seven fish species and one crab species.

At this stage, you have not translated the video recordings (in all 108 hours of video) into quantitative data.

3.3 Predation Pressure

In addition to placing the quadrats at 20 locations along each transect, your team placed a PTU at each of the 20 quadrat locations, on the opposite side of the transect from the quadrat. The PTU is essentially a stick with squid cubes tethered to it. One squid cube sits on the sediment, one hangs approximately 30 cm above the sediment. Once the PTU was submerged as the tide came in, your research team checked the PTU at regular intervals and recorded squid presence/absence. With these data, you hope to determine a predation rate for each PTU.

4 Recommendations

4.1 Inferential Scope

You have a limited number of estuaries and sites within estuaries. Furthermore, due to the non-random selection of those estuaries and sites within them, the statistical inferences you will be able to make from your study are rather narrow—they will be limited to these six sites. Any generalization of your results beyond these six sites within these two estuaries (three sites per estuary) will have to be made on the basis of scientific, subject-area arguments, not statistical arguments. In a statistical sense, this study is best characterized as a case study, looking at these estuaries and sites (it's important to note that the non-random selection of sites and estuaries limits the scope of inference). Case study analysis techniques, such as those described in Yin [2013] may help you to glean useful information from your study that could add to the research in this area, and/or suggest areas for future work. As an hypothetical example, if you were to see high correlation between algae cover and crab counts, it could motivate future research focused on this facet of eelgrass community dynamics.

4.2 Data Summary and (Possible) Modeling

There are a number of things you might consider for understanding the data you do have. There are some statistical models you can build since the quadrats and PTU's were randomly located within habitat types within sites (again, it's that you won't be able to generalize beyond these sites).

1. We always recommend that you look at plots as a starting point—there's an element of data snooping in doing this, but as this is a case study, you can look upon this as hypothesis generation. As an example, Figure 2 shows how shoot density varies across habitat type, at each site, in each bay.

Figure 2: Shoot density at each site, in each bay, by habitat type. Mercer shows an increase in shoot density across habitats, whereas Mad River shows very little. Tillamook sites show greater shoot density increase trends across habitats than do Humboldt sites.

2. If you are willing to summarize the PTU measurements into a univariate measurement, say minutes until the squid cubes are gone (though these are right censored in the sense that the maximum time over which you observed was one hour), or simply whether or not the squid cubes are gone after one hour, you could look at the relationship between that measurement and one or more of the eelgrass habitat complexity measurements. You should be aware, however, that you will have to consider the non-independence of the squid predation measurements; if a PTU has a predation event, this likely changes the probability that a nearby PTU will also be predated. Depending on what response you use for the predation, we could recommend a linear mixed or a generalized linear mixed model. At a minimum, site should be included as a blocking factor, and we recommend anaylzing data for Tillamook and Humboldt separately.

3. Shoot density seems to be the most objective measure of habitat complexity that you obtained. The potential for observer bias in visually assessed eelgrass habitat complexity measures should be considered. This is especially an issue with so many researchers making the assessments. Another study could explore this variability, by having a collection of researchers assess the same quadrats. This source of uncertainty is worth keeping in mind when you summarize your findings.

4. Confounding is a concern in any observational study because it makes it impossible to determine the source (cause) of an effect conclusively using statistical arguments. In your study, day and site are confounded because each site corresponds with exactly one day. Therefore, any effects

you do detect across sites could also be explained by differences in the days the sites were studied. Weather, for example, could increase turbidity and thus reduce fish counts. There is another issue at play here with respect to confounding.

Your overall objective is to determine the relationship between habitat type (shellfish aquaculture, edge, eelgrass), several different response variables having to do with fish/crab abundance, and predation. It's important to recognize that the eelgrass habitat complexity measurements are also likely to be related to habitat type. Therefore, a model for any of the response variables that includes both habitat type and habitat complexity variables may show no statistically significant evidence of a relationship with habitat type *simply because the habitat complexity variables are already accounting for that relationship*. In this regard, you should be careful about stating precisely the scientific hypotheses in which you are interested, and be careful not to undermine evidence against them by including variables that may be highly related.

5. The reliability of GoPro footage based species abundance and activity measurements presents several concerns, not least of which is how to convert the video footage to useable data.

    (a) Fish detection probability may vary across habitats, if fish are obscured by eelgrass, and/or across days if some days lead to greater water turbidity. Detection probability could vary across days if weather affects underwater visibility in any way. There is a vast literature on detection as an issue in wildlife studies, and fisheries are no different. While we do not have specific recommendations as to how you turn the video footage into data you can analyze (we'd probably have to look at some of the video with you and have more extensive conversations), we do recommend that you consider a method that accounts for differential detection probabilities.

    (b) You have no way of knowing whether the same fish (or crab) is repeatedly visiting the area near a GoPro, but this can artificially inflate your counts. One way around this is to use "sightings" as a response rather than counts—really, the numbers of sightings might equal the counts, but unless you can distinguish individuals, referring to the observations as sightings is more true to the information you have.

    (c) Because there are only three GoPro stations on each transect, it's not clear how you will match up whatever data you take from those GoPros to the data at the 20 quadrats and PTUs within each transect. One approach would be to match up GoPro data from each station with the nearest quadrat/PTU pair. Given your knowledge of this system, there may be better approaches.

    (d) We are concerned that the presence of the PTU may affect fish presence in some way, and artificially elevate the number of sightings. Data gleaned from GoPro footage can still be used as an index, or metric, *to compare sites and examine relative relationships*.

5 Summary

The limitations of this particular study design mean that the scope of inference is narrow in a statistical sense. Case study techniques can still give meaningful insight that add to the body of knowledge and motivate future research. Statistical models may have some value for looking at relationships among the variables you collected. We believe that well-considered, informative plots can go a long way towards explaining important relationships in this system.

References

Robert K Yin. Case study research: Design and methods. Sage publications, 2013.