Trends in Sea Star Abundance on the US West Coast

Paige Duffin Data Analysis Project

2019-11-29

# Summary/Abstract

*Write a summary of your project.*

# Introduction

## General Background Information

### Broad Scope: Characterizing Sea Star Wasting Disease

Outbreaks of infectious disease are an important, yet poorly understood, driving force in population biology and, while prevalent in both terrestrial and marine systems, the manner by which outbreaks influence biotic relationships, age demographics, community structure and function, and trophic interactions in the aquatic realm consistently lags behind that of terrestrial disease ecology. The growing body of evidence in this field, however, supports the observation that marine disease epidemics are increasing in frequency and severity (Orth et al., 2006) (Waycott et al., 2009). Elucidating the relationship between aquatic species and their pathogenic diseases is pertinent in the field of marine ecology because infection outbreaks have the potential to drastically alter ecosystem functionality (Burge et al., 2014). Many of the organisms which have faced chronic and/or severe outbreaks of disease (such as sea urchins, scleractinian corals and seagrasses) are also considered keystone species and/or ecosystem engineers, meaning that they contribute highly valuable services or functions to their surrounding community and ecosystem. Thus, it comes as no surprise that disease-driven mass mortalities of these species often generate waves of ecological permutation, ranging from temporary local disruptions to permanent phase shifts (Burge et al., 2014), such as the transformation from a coral to macroalgal dominated Caribbean reef structure that accompanied the massive *Diadema antillarum* (black sea urchin) die-off of the early 1980s (Lessios, 1988). This event, likely caused by a biological pathogen (Schultz, Cloutier, & Côté, 2016), reduced Diadema populations to less than 1% of their original size (Lessios, 1988). Mass mortality events impacting critical foundation species, ecosystem engineers, or keystone species such as the black sea urchin have been coined ‘marine disease emergencies’ due to the detrimental cascade of events that often succeeds them [miner\_large-scale\_2018].

Sea Star Wasting Disease (SSWD) is another epizootic crisis facing modern day coastal ecosystems. Outbreaks of asteroid wasting have been documented periodically since the late 1970s and SSWD describes a suite of symptoms observed across a broad range of sea star species, most of which play integral parts in shaping their community structure (Eisenlord et al., 2016) (Bucci et al., 2017) (Miner et al., 2018). Generally speaking, the wasting disease events which punctuated the past four decades were relatively brief, in localized areas, and largely failed to capture the attention of the scientific community. Beginning in summer 2013, however, mass mortalities of sea stars due to wasting disease have caused unprecedented damage, owing to the geographical and temporal extent of impact (Eisenlord et al., 2016) (Bucci et al., 2017) (Miner et al., 2018). This epizootic event, which has killed millions of asteroids across over 20 taxa, is widely referred to as the largest disease event sweeping through a wildlife marine species in documented history (Hewson et al., 2014) (Gudenkauf & Hewson, 2015) (Eisenlord et al., 2016). There is some discrepancy in the literature as to when the modern epidemic ended, with a handful of literature citing the duration as a one-year span, from 2013 to 2014 [@], while others assert it occurred until 2015 (e.g. [@]); a final set of articles document the outbreak as ongoing or remain vague about specific time frames (e.g. [@]). For the purpose of this investigation, I will be operating under the assumption that the timeframe of SSWD spanned from mid-to-late 2013 to early 2014, and will refer to this event as the “SSWD event” or the “disease event” hereafter.

The consensus of reports which emerged in the years after the disease event indicates that the ochre sea star, *Pisaster ochraceus*, suffered very high mortality [@]. Ochre stars occupy the lower to mid-intertidal on rocky shores, spanning an impressive range from Alaska to Baja California (**???**), and as keystone predators, the health of *P. ochraceus* populations heavily influence the surrounding community structure and species richness (**???**). There has been a special emphasis on the characterization of SSWD in this species, and researchers are highly motivated to address questions like: what can we learn from the individuals who survived and/or from the populations which have rebounded? How has community structure been altered as a result of this epidemic? Are the results at different locations similar? If not, what drives the differences? Though devastating, disease outbreak events such as this provide unique opportunities in the coming years to better understand how a suite of ecologically important organisms respond to unprecedented levels of ecological permutation.

## Description of data and data source

### The Source

The vast majority of sea star abundance surveys and wasting disease documentation cover the west coast of the United States. While it is almost certainly true that SSWD is most prevalent in this region, there is a strong bias towards the discovery of diseased organisms owing to the emphasis of long-term survey networks in the area, such as the extensive [Multi-Agency Rocky Intertidal Network](https://marine.ucsc.edu/index.html) (MARINe for short).

The purpose of this investigation is to make use of this data monitoring system to investigate spatial and temoporal changes in sea star abundance along the west coast of the United States, with particular emphasis on the years before, during, and after the SSWD outbreak.

### The Data

Below I explore the raw data I downloaded from a link I recieved after submitting a request for the data on the MARINe webpage.

Notably, it contains the following information (some parameters removed from this overview):

* groupcode: The unique code for each monitoring group
* marine\_site\_name: The name of the site where the survey was conducted.
* latitude and longitude: (self explanatory)
* marine\_season\_code: A four-character code to identify the Sampling season. The first two characters indicate the season and the last 2 characters indicate the year.
* method\_code: The method used for sampling the plot.
* size\_bin: The size of the species being counted, binned to the nearest 5 or 10 millimeter.
* target\_assemblage: The specific organism that this transect was created to monitor.
* total: Total number of individuals counted in a given size\_bin
* mpa\_designation: Describes whether the referenced site is located within a Marine Protected Area (MPA) or is a reference site. If this field is blank, the referenced site is not located within an MPA, or is not a reference site for an existing MPA site.
* georegion/ bioregion: geographic/biogeographic region in which site is located
* state\_province: The State or Province and Country where the referenced site is located.
* island: The name of the island where the referenced site is located. Sites not on islands are designated as mainland.

## Questions/Hypotheses to be addressed

How does species abundance change across time and space?

Are there discernable changes in spp. abundance before & after the focal SSWD event?

How did populations of the focal sp., P. ochraceus, respond & recover from the SSWD event?

Which variables predict species abundance?

Can we predict bioregion using the data and tree-fitting model?

**1. Are sea star populations changing over time?** I anticipate that there will be a decline in sea star populations over time, but that the change will not occur uniformly across geographical regions and species. Species that are reported to have been affected by SSWD will experience the biggest declines.

**2. Are there definable geographical regions throught space based on predictor variables?** I expect to be able to define rough species ranges.

**3. Is there evidence of sampling bias?** Is there a relationship between important survey parameters (mainly, abundance data) and sampling method (ex. group conducting the survey, plot sampling method, the target assemblage). At the very least, I expect there will be bias based on the targeted organism of the survey; if a given survey aims to characterize the presence of one species of sea stars, it is likely that they will seek out that species and record higher counts.

# Methods and Results

## Data import and cleaning

*Write code that reads in the file and cleans it so it’s ready for analysis. Since this will be fairly long code for most datasets, it might be a good idea to have it in one or several R scripts. If that is the case, explain here briefly what each file does. The files themselves should be commented well so everyone can follow along.*

The original data was downloaded from a link I recieved after submitting a request for the data on the Multi-Agency Rocky Intertidal Network, or MARINe, [webpage](https://marine.ucsc.edu/index.html). The raw data is found here: ./data/raw\_data/seastarkat\_size\_count\_totals\_download.csv.

This data was cleaned extensively prior to use in this analysis, and the code which performs this cleaning is found here: ./code/processing\_code/processing\_script. Briefly, this involved examining each variable by plotting factors as bar plots and integers/numerics as histograms. This allowed me to observe the structure of each variable and make necessary changes to clean the data.

## Exploratory analyses

The code which performed my exploratory analyses can be found here: ./code/analysis\_code/Exploratory\_Analysis.Rmd

### Mapping Survey Sites

First, I created geographical maps to visualize the sample sites. Sites were sampled in four states:

### How does species abundance change across time and space?

### Are there discernable changes in spp. abundance before & after the focal SSWD event?

### How did populations of the focal sp., P. ochraceus, respond & recover from the SSWD event?

## Full analysis

*Use one or several suitable statistical/machine learning methods to analyze your data and to produce meaningful figures, tables, etc. This might again be code that is best placed in one or several separate R scripts that need to be well documented. You can then load the results produced by this code*

### Which variables predict species abundance?

The first used abundance counts (or the variable total) as the outcome predictor on a reduced dataset that only retained the species of interest, *Pisaster ochraceus*. This code is found here: ./code/analysis\_code/Continuous\_Outcome\_Modeling.Rmd.

### Can we predict bioregion using variable selection and tree-fitting models?

*For this analysis, our main outcome of interest is if an outbreak was caused by the G2.4 strain of norovirus or not, and how other factors might be correlated with that strain*

*cross validation. We could do data splitting again as we did in the previous exercise, to have a final test set. But since you saw how it works in the previous exercise, we skip it here. We use the full data to fit to our models. We’ll still use cross-validation to get a more honest estimate of model performance. For a real data analysis, the choice to keep some for a final test or not is based on your goals. If your focus is predictive performance, you should consider this split. If your focus is inference or exploratory analysis, you might want to skip this.*

The second focused on the categorical variable bioregion and used a machine learning tree fitting method. This code is found here: ./code/analysis\_code/Continuous\_Outcome\_Modeling.Rmd.

## Discussion

### Summary and Interpretation

*Summarize what you did, what you found and what it means.*

### Strengths and Limitations

*Discuss what you perceive as strengths and limitations of your analysis.*

### Conclusions

*What are the main take-home messages?*

*Include citations in your Rmd file using bibtex, the list of references will automatically be placed at the end*

reference (**???**) here reference (**???**) here

### References

Bucci, C., Francoeur, M., McGreal, J., Smolowitz, R., Zazueta-Novoa, V., Wessel, G. M., & Gomez-Chiarri, M. (2017). Sea Star Wasting Disease in Asterias forbesi along the Atlantic Coast of North America. *PLOS ONE*, *12*(12), 1–20. <https://doi.org/10.1371/journal.pone.0188523>

Burge, C. A., Mark Eakin, C., Friedman, C. S., Froelich, B., Hershberger, P. K., Hofmann, E. E., … Harvell, C. D. (2014). Climate Change Influences on Marine Infectious Diseases: Implications for Management and Society. *Annual Review of Marine Science*, *6*(1), 249–277. <https://doi.org/10.1146/annurev-marine-010213-135029>

Eisenlord, M., Groner, M., Yoshioka, R. M., Elliott, J., Maynard, J., Fradkin, S., … Harvell, C. D. (2016). Ochre star mortality during the 2014 wasting disease epizootic: Role of population size structure and temperature . *Philosophical Transactions of the Royal Society B*, 201502212. <https://doi.org/https://doi.org/10.1098/rstb.2015.0212>

Gudenkauf, B. M., & Hewson, I. (2015). Metatranscriptomic Analysis of Pycnopodia helianthoides (Asteroidea) Affected by Sea Star Wasting Disease. *PLOS ONE*, *10*(5), 1–10. <https://doi.org/10.1371/journal.pone.0128150>

Hewson, I., Button, J. B., Gudenkauf, B. M., Miner, B., Newton, A. L., Gaydos, J. K., … Harvell, C. D. (2014). Densovirus associated with sea-star wasting disease and mass mortality. *Proceedings of the National Academy of Sciences*, *111*(48), 17278–17283. <https://doi.org/10.1073/pnas.1416625111>

Lessios, H. A. (1988). Mass Mortality of Diadema Antillarum in the Caribbean: What Have We Learned? *Annual Review of Ecology and Systematics*, *19*, 371–393. Retrieved from <http://www.jstor.org/stable/2097159>

Miner, C. M., Burnaford, J. L., Ambrose, R. F., Antrim, L., Bohlmann, H., Blanchette, C. A., … Raimondi, P. T. (2018). Large-scale impacts of sea star wasting disease (SSWD) on intertidal sea stars and implications for recovery. *PLOS ONE*, *13*(3), 1–21. <https://doi.org/10.1371/journal.pone.0192870>

Orth, R., Carruthers, T., Dennison, W., Duarte, C., Fourqurean, J., JR, K., … Williams, S. (2006). A Global Crisis for Seagrass Ecosystems. *BioScience*, *56*, 987–996. <https://doi.org/10.1641/0006-3568(2006)56[987:AGCFSE]2.0.CO;2>

Schultz, J. A., Cloutier, R., & Côté, I. M. (2016). Evidence for a trophic cascade on rocky reefs following sea star mass mortality in British Columbia. *PeerJ*.

Waycott, M., Duarte, C. M., Carruthers, T. J. B., Orth, R. J., Dennison, W. C., Olyarnik, S., … Williams, S. L. (2009). Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proceedings of the National Academy of Sciences*, *106*(30), 12377–12381. <https://doi.org/10.1073/pnas.0905620106>