Predicting Community Assemblage in The Rocky Intertidal of Acadia National Park

Final Proposal

DSE6311 – SU2 2025
Nick Lagoni
7/22/2025

1 Introduction

1.1 Background

The Gulf of Maine is changing rapidly, warming at an accelerated rate compared to most of the globe (Friedland et al. 2020). Intertidal communities are experiencing a similarly rapid change, including changes in vertical zonation (Trott 2022) and introduction of invasive and range expanding species (Cohen et al. 1995, Yamada 2001, Epifanio 2013, Johnson 2015, Cheng et al. 2025). These changes may compound over time into significant changes to community structure. Because of this and the importance of these ecosystems ecologically and as a source of ecosystem services (including as interpretive spaces in National Parks), it is important that we have an understanding of what the future holds for the rocky intertidal in the Gulf of Maine, and which factors contribute most to that future.

1.2 Research Question and Hypothesis

The primary research question posed by this analysis is: can we accurately predict how the community assemblage of the rocky intertidal in Acadia National Park will change in the near future, and what variables (including substrate composition, invertebrate abundance, and temperature) are the strongest predictors of that. While long-term monitoring has been ongoing through the Northeast Temperate Network (NETN) in the Inventory and Monitoring Division of the National Park Service (NPS) since 2013, no significant analysis has been done on this subject since that protocol began.

I hypothesize that temperature and presence and abundance of motile invertebrates will significantly predict the substrate percent cover of *Fucus vesiculosus* (Linnaeus 1753) and *Ascophyllum nodosum* (Linnaeus 1753) in the Acadia rocky intertidal over time.

I predict that an increase in the abundance of motile invertebrates *Littorina obtusata* will significantly be associated with a significant reduction in the algal cover of *F. vesiculosus* and *A. nodosum*, as these periwinkles have been shown to feed on both of these species (Hadlock 1979, Watson and Norton 1987). I predict that *Littorina littorea* abundance will not significantly predict the algal cover of *F. vesiculosus* and *A. nodosum*, as this species has been shown to avoid feeding on these species (Watson and Norton 1985). I predict that increased temperatures will be correlated with a decrease in *F. vesiculosus and A. nodosum* in the higher plots, and an increase in both target species in the lower plots (i.e., the red algae plots); these species have been shown to expand their range as waters warm (Jueterbock et al. 2013, Marbà et al. 2017), which may extend to their local zonation range as well, expanding into deeper water.

2 Data and Analysis

2.1 Data Source

NETN has a dataset spanning 8-years (2013 – 2021) of long-term ecological monitoring data across six sites in Acadia, as well as three sites in the Boston Harbor Islands National Park (Northeast Temperate Network 2021). All nine sites are visited annually for data collection. While the Acadia sites are the focus of this analysis, as the Boston Harbor Islands are a very different ecosystem, both naturally, with an uncommon mixed coarse substrate intertidal habitat, and artificially, in part due to Boston's position as a global shipping hub introducing a significant amount of invasive species (Putnam et al. 2024). I am working directly with the Data Manager for NETN to see if I can procure additional years' data (2022 – 2025), depending on the NPS distribution policy. The dataset is broken down into different sub-protocols, including: motile invertebrate and substrate percent cover photoplots, line intercept vertical zonation transects, tide pool motile invertebrate band transects, barnacle recruitment photoplots, and water temperature logging.

2.1.1 Motile Invertebrate and Substrate Percent Cover Photoplots

Each site is divided into subsites based on target species (*A. nodosum, F. vesiculosus,* red algae, barnacles, and mussels), with each of those comprising five photoplots each. This subprotocol includes count data of all motile invertebrate individuals collected from each photoplot, separated by species, as well as measurements of a random subsample of ten individuals for each species and photoplot, or however many were collected if there were less than ten in a plot. This data also includes percent cover data for each photoplot, determined using a proprietary model with visual validation by a human.

2.1.2 Vertical Zonation Point Intercept Transects

Each site has three vertical transects meant to capture trends in vertical zonation. These transects run parallel to each other and are generally ~30m, though there is variation for sites with a very shallow grade. The data from this subprotocol is point intercept data containing the species or substrate found every 0.3m along the transect, as well as the distance of each bolt on the transect. This allows measurements between bolts to be grouped together into bolt ranges for more clear comparison between years, as the bolts are static, which accounts for human error in laying the transect and reading at consistent intervals.

2.1.3 Tidepool Motile Invertebrate Band Transects

Each site has three 10m transects that are set up to run along the length of a tide pool. For each of these transects, a band spanning 1m on either side of the transect is searched for all echinoderms, including *Strongylocentrotus droebachiensis* (Müller 1776), *Asterias forbesi (Desor 1848)*, *Asterias rubens* (Linné and Salvius 1758), and *Henricia sanguinolenta* (Müller 1776). A total count of each species for each pool is taken, and a random sample of 10 of each species are measured for length.

2.1.4 Barnacle Recruitment Photoplots

Each of the five barnacle recruitment plots per site are scraped clean annually in the winter and photographed in the summer to quantify barnacle recruitment in the plots. Plots are photographed and the number of barnacles present in the summer are counted. This represents the number of new barnacles that recruited to the plot.

2.1.5 Water Temperature Logging

There are three temperature loggers on each site. These record temperature readings once every hour and have their readings downloaded annually at each site. This data is then filtered to only include water temperature due to heterogeneity in the air temperature of intertidal habitats.

2.2 Predictor Variables

My target variable will be the percent cover of *F. vesiculosus* and *A. nodosum*, which are particularly important to the whole of the rocky intertidal as ecosystem engineers (Råberg and Kautsky 2007, Westerbom and Koivisto 2022). These are a focus of the NETN protocol and, from personal anecdotal experience working in the field on this protocol, have been undergoing dramatic changes in distribution along the intertidal in recent years, which is a trend I would like to verify using the data. The predictor variables will include temperature, percent cover of each algal species (or group when speciation is inappropriate or unfeasible) and sessile invertebrates like *Semibalanus balanoides* (Linné and Salvius 1758) and *Mytilus edulis* (Linné and Salvius 1758), and abundance of each motile invertebrate.

2.3 Analytical Plan

My analytical plan is to start with simple regression machine learning models, such as logistic regression, before transitioning to more robust and complex decision tree-based models. I will use cross-validation to tune the hyperparameters of both the simple regression models and the tree-based models and evaluate their performance against a validation dataset. I will evaluate these models based on overall accuracy and extract importance scores. Whichever model combines accuracy and interpretability of importance will be selected as the final model, as these are both important needs for answering the posed question. I anticipate random forest being a good fit for this data for determining importance scores and the large number of features contained in this dataset. Logistic regression (and the variations therein such as ridge regression and LASSO regression) may prove valuable for interpretability's sake. There are other variables contained within each dataset, such as the person recording the data, the month, and the site, that I would like to explore as potential confounding variables. I intend to explore the effects of these variables by including them as predictors: if I see significant prediction from something like the recorder, that will indicate a flaw in the data, and I may have to filter out the most impactful recorders. I may also include these as mixed effects in the regression equation to explore this as well. If the site or month turns out to be significant, that tells me that there is more work to be done on determining what about those specific values is so important to the abundance of the target variable.

2.4 Evaluating the Question and Hypotheses

I will know that my question has been answered whether or not the model performs. If the model performs well, then yes, we can predict the changes in community assemblage. If the model predicts poorly, then, at least with the current data available, we cannot.

The hypothesis relies first on the model performing significantly well and then relies on that model producing an interpretable and meaningful list of important variables. If the model is able to do both of these things, and our hypothesized predictors out to be strong predictors of whichever target variable I select, we can consider the hypothesis supported.

3 Technical Details

3.1 Language and GitHub Repo

I intend to code exclusively in R via RStudio, though I may dip into Python if there is a significantly demonstrated need for it in my project. I may do some data cleaning work in Excel as well. My GitHub repo can be found here.

3.2 Additional Resources

I may benefit from seeking out some more Gulf of Maine related datasets that include the timespan from 2013 – 2021, to use as additional covariates for the model. For now, I intend to use NETN's proprietary data as their protocol is designed specifically to inform on the overall trends in these sites. However, depending on what I find from the model results and the importance scores, supplemental datasets may further elucidate these trends. This may include things like mean-high high-water levels, meteorological data, and park visitation data.

References

Cheng, H., M. D. McMahan, S. B. Scyphers, L. McClenachan, and J. H. Grabowski. 2025. Observations, perceptions and concerns of the American lobster industry regarding the range-expansion of Black Sea Bass. Marine Policy 173:106517.

Cohen, A. N., J. T. Carlton, and M. C. Fountain. 1995. Introduction, dispersal and potential impacts of the green crab Carcinus maenas in San Francisco Bay, California. Marine Biology 122:225–237.

Desor, É. 1848. Proceedings of the Boston Society of Natural History. Boston Society of Natural History.

Epifanio, C. E. 2013. Invasion biology of the Asian shore crab Hemigrapsus sanguineus: A review. Journal of Experimental Marine Biology and Ecology 441:33–49.

Friedland, K. D., R. E. Morse, J. P. Manning, D. C. Melrose, T. Miles, A. G. Goode, D. C. Brady, J. T. Kohut, and E. N. Powell. 2020. Trends and change points in surface and bottom thermal environments of the US Northeast Continental Shelf Ecosystem. Fisheries Oceanography 29:396–414.

Hadlock, R. H. 1979. The distribution of Littorina obtusata (L.) in the rocky intertidal: effects of competition with Littorina littorea (L.). Master's thesis, Department of Zoology, University of Rhode Island, Kingston, RI.

Johnson, D. S. 2015. The Savory Swimmer Swims North: A Northern Range Extension of the Blue Crab Callinectes Sapidus? Journal of Crustacean Biology 35:105–110.

Jueterbock, A., L. Tyberghein, H. Verbruggen, J. A. Coyer, J. L. Olsen, and G. Hoarau. 2013. Climate change impact on seaweed meadow distribution in the North Atlantic rocky intertidal. Ecology and evolution 3:1356–1373.

Linnaeus, C. 1753. Species plantarum, exhibentes plantas rite cognitas ad genera relatas cum differentiis specificis, nominibus trivialibus, synonymis selectis, locis natalibus, secundum systema sexuale digestas. Vol 1.

Linné, C. von, and L. Salvius. 1758. Caroli Linnaei...Systema naturae per regna tria naturae :secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Impensis Direct. Laurentii Salvii, Holmiae.

Marbà, N., D. Krause-Jensen, B. Olesen, P. B. Christensen, A. Merzouk, J. Rodrigues, S. Wegeberg, and R. T. Wilce. 2017. Climate change stimulates the growth of the intertidal macroalgae Ascophyllum nodosum near the northern distribution limit. Ambio 46:119–131.

Müller, O. F. 1776. Zoologiae Danicae Prodromus, seu Animalium Daniae et Norvegiae Indigenarum characters, nomina, et synonyma imprimis popularium. Typis Hallageriis, Havniae 32:1–282.

Northeast Temperate Network. 2021, January 1. Long-term Rocky Intertidal - Database. https://irma.nps.gov/DataStore/Reference/Profile/2289832.

Putnam, A. B., S. C. Endyke, A. R. Jones, L. A. D. Lockwood, J. Taylor, M. Albert, and M. D. Staudinger. 2024. Historical insights, current challenges: tracking marine biodiversity in an urban harbor ecosystem in the face of climate change. Marine Biodiversity 54.

Råberg, S., and L. Kautsky. 2007. A comparative biodiversity study of the associated fauna of perennial fucoids and filamentous algae. Estuarine, Coastal and Shelf Science 73:249–258.

Trott, T. J. 2022. Mesoscale Spatial Patterns of Gulf of Maine Rocky Intertidal Communities. Diversity 14:557.

Watson, D. C., and T. A. Norton. 1985. Dietary preferences of the common periwinkle, Littorinalittorea (L.). Journal of Experimental Marine Biology and Ecology 88:193–211.

Watson, D. C., and T. A. Norton. 1987. The habitat and feeding preferences of Littorina obtusata (L.) and L. mariae sacchi et rastelli. Journal of Experimental Marine Biology and Ecology 112:61–72.

Westerbom, M., and M. Koivisto. 2022. Mussels and canopy-forming algae as ecosystem engineers: their contribution to community organization in the rocky sublittoral. Frontiers in Marine Science 9.

Yamada, S. B. 2001. Global invader: the European green crab.