Jane Maguire
BEES2041 - Practical Report 2
18/04/2025

Introduction:

The rocky shore of Maroubra Beach is a highly variable environment with different elevation levels and habitat types that are constantly changing due to the changing tide. The goal of this study is to investigate two different ecological aspects of the rocky shoreline. The first goal is to determine how gastropod abundance changes with the height of the shoreline. The second objective is to determine if species composition differs between rockpool and emergent habitats.

Methods:

Abundance of Gastropods:

Data on the abundance of gastropods was collected from the rocky shore platform at Maroubra Beach. The platform was divided into three elevation zones, low, mid, and high, based on distance from the water's edge. Each zone was measured out to be 4-6m high and 10m long. Within each zone, quadrats were placed at randomly chosen locations by throwing them blindly to minimize any placement bias. We then conducted separate sampling for each of three gastropod species (Nerita, Cellana, and Austrocochlea) to avoid interspecies interaction effects and ensure independence between samples. For each quadrat, we recorded the elevation zone and the number of individuals of the selected species.

Prior to analyzing this data, we first eliminated two high outliers as they were skewing our data making it difficult to read. After examining the data, there were species that were labelled differently that are actually the same. Nerita can also be known as the black snail, Cellana can be known as the limpet, and Austrocochlea the zebra snail. All of these different labels were under the species column so it appears as 6 different species even though it is only three. We dealt with this by mutating the dataset so that only the labels Nerita, Cellana, and Austrocochlea were shown for the different species.

Then we begin the analysis of this dataset. We will analyze this data by first creating a multiple predictors linear model using elevation and species. We tested the assumptions of this model using the check_model function in order to make sure that assumptions of linearity, independence of points, homoscedasticity, normality of residuals, absence of collinearity, and no outliers were met. Testing the assumptions of the model including its interaction term shows us that the model without the interaction term has much lower collinearity, making the parameter estimates much more trustworthy. The model with the interaction effect might be overfitting without much gain in how much variation the model accounts for. The R^2 value only increases from 0.247 to 0.2775. This is only a 3% change in the amount of variation accounted for which is not super significant. Because of this we will continue with the model with just the elevation and species effects now that we know assumptions are met. We then run an ANOVA to test the significance of main effects and estimate contrasts to find where the differences lie within groups and which groups are significantly different from each other.

Species Composition between Habitats:

To investigate whether species composition differs between rockpool and emergent habitats, we again collected data from the rocky shore platform on Maroubra Beach. The same as the first part of this experiment, quadrats were placed at randomly chosen locations by throwing them blindly to minimize any placement bias. However, we did ensure that an equal number of quadrats were placed in each habitat type, either within rock pools or on emergent rocky surfaces. Each quadrat served as a sampling unit in which all visible species were recorded. For each quadrat, we noted the habitat type and counted the number of individuals for each species listed on the data sheet. To maintain independence of samples, we ensured that no rockpool was sampled more than once. This method provided us with a multivariate dataset suitable for analyzing patterns in species composition between the two habitat types.

Prior to analyzing this data, we first eliminated a high outlier in the red_algae species. The algae abundance was measured based on the percentage of the quadrat covered in algae. In our dataset, there was one red_algae value of 300 which is not possible because 100 is the highest possible percentage value. Therefore, this datapoint is an error as it was likely misentered or misrecorded. After removing this datapoint we can continue the analysis of this dataset. We will analyze this data by first running a permanova to tell us if the difference in species composition between habitats is statistically significant. We then will do a PCA analysis in order to visualize these differences. When running the PCA, we will use a correlation matrix because we are dealing with variables that are measured on different scales (algae is measured in %, all others in count). Performing this analysis should then help us determine if species composition differs between rockpool and emergent habitats.

Results:

<u>Abundance of Gastropods:</u>

To explore the relationship between elevation and gastropod abundance while accounting for species differences, we ran a linear model with two main predictors, Elevation and Species. This allowed us to examine whether species abundance varies across elevation zones and among the three different gastropod species. To visualize the results, we created a faceted violin plot of abundance across elevation levels, split by species. This graph (Figure 1 below) includes jittered raw data points and error bars representing the estimated means and 95% confidence intervals for each group.

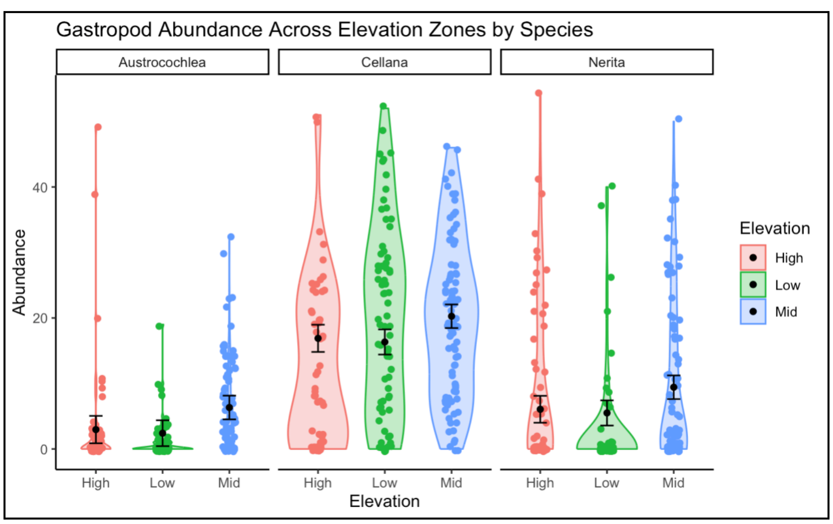

Figure 1: *Violin plots showing the distribution of gastropod abundance across elevation zones (Low, Mid, High) for each species (Nerita, Cellana, Austrocochlea). Raw data points are jittered, and black dots represent estimated group means with 95% confidence intervals derived from a multiple predictors linear model.*

The results of our ANOVA tell us that both the main effects of elevation and species are statistically significant with p values less than 0.001. There was a significant main effect of Elevation on Abundance_per_quadrat, ($df = 2$, $F = 8.3475$, $p = 0.0002648$), indicating that Low, Mid and High platforms differ in species Abundance counts. There was also a significant main effect of Species on Abundance_per_quadrat, ($df = 2$, $F = 93.0297$, $p < 0.001$), indicating that Nerita, Austrocochlea, and Cellana differ in Abundance as well. The ANOVA tells us that there are significant differences between groups overall, but it doesn't tell us which groups differ from each other.

Using the estimate contrasts function, we can find where the differences lie within groups and which groups are significantly different from each other. Running the estimate contrast function 2 different times we can look at the 2 different effects. We can see overall differences in elevation while averaged across species and we can see overall differences in species while averaged across elevations.

When looking at overall differences in elevation while averaged across species (Table 1 below), we see that generally mid elevation abundance is greater than high elevation abundance and mid elevation abundance is also greater than low elevation abundance. Both of these are statistically significant with p-values of <0.05. We can also see high elevation has slightly higher abundance than low elevation, however, the difference is not significant with a p-value of 0.628 and a confidence interval that includes 0.

Table 1:
Estimate Contrasts Results by Elevation

	Mean Difference	95% CI	P-value
Mid | High	3.37	1.22, 5.51	0.002
Mid | Low	3.93	1.91, 5.94	< 0.001
Low | High	-0.56	-2.82, 1.70	0.628

When looking at the overall differences in species while averaged across elevations (Table 2 below), we see that generally Cellana have greater abundance than Austrocochlea. We see that Nerita also has greater abundance than Austrocochlea. Lastly we see that Cellana generally has greater abundance than Nerita. All of these are significant results as they each have p-values of < 0.05.

Table 2:
Estimate Contrasts Results by Species

	Mean Difference	95% CI	P-value
Cellana | Austrocochlea	13.93	11.81, 16.05	< 0.001
Nerita | Austrocochlea	3.10	0.98, 5.21	0.004
Nerita | Cellana	-10.84	-12.92, -8.75	< 0.001

As a result of these statistical tests, we know that elevation and species both significantly influence gastropod abundance. To answer our original question, mid tidal zones appear to support higher abundances across species as there is a significant difference between mid/low and mid/high zones, and there is not a significant difference between the abundance of gastropods in low and high tidal zones.

Species Composition between Habitats:

To explore the relationship between species composition and habitat type, we used a permanova (permutational multivariate analysis of variance) to test for statistical differences in species composition between emergent surfaces and rockpools. The results of our permanova, (df = 1, F = 22.022, p = 0.001), tell us that the difference in species composition between habitats is statistically significant. It also gives us the R^2 value (0.15081) which tells us that this model accounts for 15.081% of the variation in the data. The overall takeaway from the permanova is that species composition does differ between the two habitat types, emergent and rockpool.

To visualize these differences, we can use Principal Components Analysis (PCA). The PCA ordination plot (Figure 2 below) shows samples plotted along the first two principal components, with each point representing a single sample. When color-coded by habitat, the plot shows us that samples from emergent zones and rock pools have some overlap but there still are distinct clusters seen between the two groups. This means species composition has some similarities across emergent and rock pool habitats, though there are definitely still differences between the two groups. There is overlap between groups so habitat type likely isn't the only driver of variation, though it definitely still contributes as we proved with the permanova above.

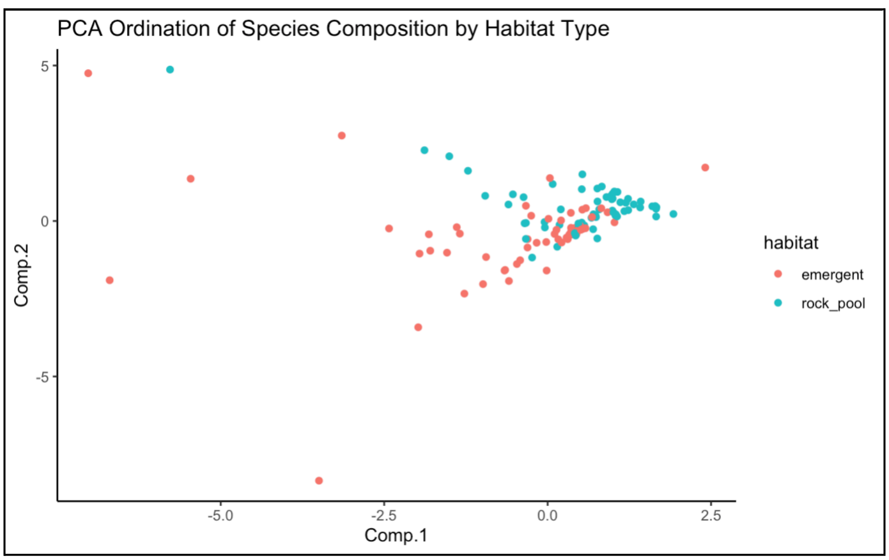

Figure 2: *PCA ordination plot showing variation in species composition between rockpool and emergent habitats. Each point represents a quadrat, positioned based on its species composition.*

The biplot (Figure 3 below) shows us how different species contribute to the variation along the PCA axes. Arrows point in the direction of increasing abundance for each species. Species like brown_algae, green_algae, and red_algae point strongly to the right, suggesting they are positively associated with PC1 and possibly with rock pools based on the location of the rockpool cluster from Figure 2. This makes sense as algae is most commonly found in water. Species like Actinia, Austrocochlea, and Patiriella point left/up, contributing more strongly to PC2, possibly more linked to emergent areas based on the location of the emergent cluster from Figure 2.

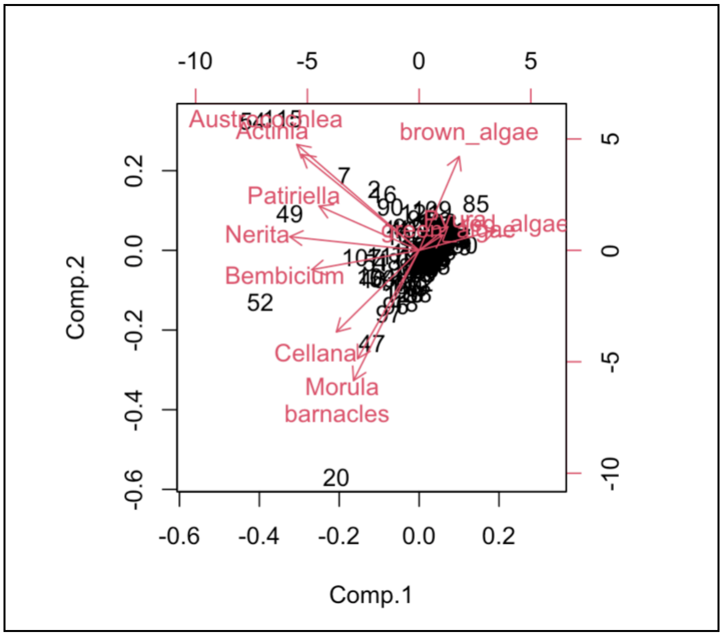

Figure 3: *Biplot of PCA based on species composition data using a correlation matrix. Points represent individual quadrats, and arrows represent species loadings, indicating how strongly each species contributes to the principal components.*

Based on the results of the permanova and pca, we can say species composition does differ between habitats. There is some overlap of species across habitats, with a few species driving differences. The PCA suggests that algae species and a few other species (those with longer arrows) are key contributors to differences in species composition between rock pools and emergent areas. Overall, the results highlight how both elevation and habitat type play significant roles in shaping gastropod abundance and intertidal species composition on the rocky shore at Maroubra Beach.