

A ROUGH ESTIMATE OF GHOST SHRIMP
(NEOTRYPAEA CALIFORNIENSIS)
POPULATION CHARACTERISTICS IN FALSE
BAY, SAN JUAN ISLAND, WASHINGTON

James Gavin Fitzpatrick

BIODIVERSITY AND MONITORING OF ESTUARINE ECOSYSTEMS AUTUMN
2021

Abstract

Neotrypaea californiensis are a burrowing decapod species that acts as an important ecosystem engineer and bioturbator in intertidal sandflats. Using a novel sampling method, the abundance of *N. californiensis* individuals was estimated for two sites in False Bay, San Juan Island, Washington. An equation for the biomass of an individual given the carapace length was derived from a regression of logarithmic carapace length and logarithmic mass and used to estimate the biomass of *N. californiensis* across both sample sites. Sediment samples were taken and subjected to a volume ratio analysis to quantify the percentage of sediment grain size that fell into >2mm, 2mm-0.5mm, and <0.5mm categories for each sample point. The influence of surface sediment grain size, beneath-surface sediment grain size categories, percent *Ulva* cover, and number of surface species on the spatial distribution of *N. californiensis* was examined using linear regressions and correlation matrices, but none were found to be significantly correlated or explanatory for the spatial distribution of individuals.

Keywords

False Bay, San Juan, *Neotrypaea californiensis*, Bay Ghost Shrimp, *Callinassa californiensis*, Population Estimate, Abundance, Biomass, *Ulva* spp

Acknowledgments

Special thanks to Ryan Pittsinger, who helped to design the sampling plan and who assisted with sample collection and recording data. Thanks also to Megan, James R., and Ryan Fitzpatrick, who provided emotional support and assisted in collecting samples for length/weight analysis.

This study would not be possible without the facilities and instruction given by Friday Harbor Laboratories at the University of Washington. This study was conducted as part of the Biodiversity and Monitoring of Estuarine Ecosystems course in autumn quarter, 2021. Thanks to the instructors of this course, who provided advice and instruction throughout the study period: Dr. Alan Trimble, Dr. Brooke Sullivan, Dr. Megan Dethier, and Dr. Wendel Raymond.

1. Theoretical Background

Neotrypaea californiensis, commonly known as bay ghost shrimp and formerly known as *Callinassa californiensis* (Dana 1854), are a decapod callinassid that lives in soft-sediment habitats throughout the U.S. West Coast. Ghost shrimp construct elaborate burrows in the sand with many surface entrances. This behavior provides the shrimp with oxygen in otherwise anoxic sediment, protection from surface predators, and food in the form of diatoms and other microorganisms eaten from the sediment that the shrimp digs through and detritus that the shrimp filters from the water flowing through the burrow (Feldman et al. 2000). The constant digging of the ghost shrimp means it acts as a bioturbator, facilitating faster flow of oxygen and nutrients through sediment that might otherwise stagnate and settle (Armstrong et al. 1995). Ghost shrimp are of interest in the habitats where they are found for a number of reasons. The first is that they are considered a nuisance in aquaculture farms focused on oysters, because they disturb sediment and suffocate young oysters shortly after they settle (Feldman et al. 2000). Another is that ghost shrimp can often be very important in habitats where they live due to their high influence over nutrient transport and oxygen content in soft sediment (Armstrong et al. 1995). Ghost shrimp are a major component of nitrogen cycling in habitats where they are present, since they reduce loss of nitrogen to sediment through their consumption of detritus in

the soil and subsequent deposition of fecal pellets on the surface outside their burrow, thereby returning nitrogen to the surface (Dewitt et al. 2004).

Ghost shrimp are highly prevalent in Washington tidal flats such as Willapa Bay and Gray's Harbor. This study aims to quantify ghost shrimp abundance within a tide flat on San Juan island. False Bay is a tide flat consisting largely of sandy and muddy sediments which opens directly into the Strait of Juan de Fuca. It is fed via this ocean water and the input of groundwater and creeks in the False Bay watershed, the largest in San Juan county. Surrounding farms and irrigation efforts have affected the watershed in ways that are still being quantified, and the input of nutrients and water into False Bay proper is not known.

Large blooms of *Ulva spp.* also occur within False Bay perennially with an unknown cause. These blooms result in large piles of *Ulva* biomass which settle and decay in the high intertidal. The impact of *Ulva* cover on ghost shrimp burrows in False Bay is not known, but decaying *Ulva* biomass is known to cause high levels of hydrogen sulfide (Nedergaard et al. 2003) in the sediment underneath it, and anoxic conditions due to suffocation and uptake of oxygen by decomposers. Ghost shrimp are resistant to anoxic conditions even over extended periods of time (Thompson & Pritchard 1969), but their distribution may still be negatively affected if they cannot tunnel through the *Ulva* to the surface to bring more oxygen into the burrow, or if high levels of hydrogen sulfide are harmful to them. Alternatively, *Ulva* may serve as a supplementary food source for ghost shrimp, which feed on any organic matter they can filter from the water. *Ulva* cover may also attract microorganisms and small or microscopic grazers or decomposers that could also provide food for *N. californiensis* even if the *Ulva* itself is not eaten. This relationship between ghost shrimp and *Ulva* is unknown and should be

quantified, because *Ulva* cover and *N. californiensis* abundance are likely both very important biotic factors in the distribution of other organisms.

Very little research of any kind has taken place in False Bay, and the need for baseline data on the presence of organisms and species in False Bay is apparent in order to better understand the ecosystem of the watershed. Ghost shrimp in particular are highly important as creators of habitat for commensals (Hoffman 1981) (Tamaki et al. 2018), as a limiting factor for juvenile oysters and seagrass seeds (Feldman et al. 2000) (Dumbauld & Wyllie-Echeverra 2003), as prey for many larger organisms (Armstrong et al. 1995), and as an essential component of nitrogen cycling (Dewitt et al. 2004). This study aims to answer the research question: “What is the abundance of *Neotrypaea californiensis* individuals in False Bay, and what biotic and abiotic factors might drive their distribution?” The first hypothesis of this study is that ghost shrimp will be more abundant where the sediment grain sizes are smaller on average. The second is that ghost shrimp will be more abundant where there is less *Ulva* cover. Additionally, data was collected on the presence of other invertebrate species near ghost shrimp burrows and the percent cover of *Ulva spp.* within the same area to determine if either of these biotic factors have an effect on ghost shrimp distribution.

2. Methodologies

2.1. Sample Collection

Data was collected from 48 points across two sampling areas in False Bay, San Juan Island, Washington (Fig. 1). The two selected areas were the mouth of False Bay Creek and the mouth of an unnamed stream which will hereafter be referred to as the Forest Stream. A point in the center of the mouth of each stream was selected and five transect lines were drawn outwards from it at equal angles of 25 degrees to form a fan. Each transect line was 160 meters long at the False Bay Creek site, while at the Forest Stream site each transect line was 80 meters long. From these transect lines, intersecting lines were drawn to approximate an arc, leaving four sectors of roughly equal area to each other at each sample location. At the False Bay Creek site, the sectors were drawn from 120 meters along each transect as the inner arc and from 160 meters as the outer arc. At the Forest Stream site, the sectors were drawn from 40 meters along each transect as the inner arc and from 80 meters as the outer arc. This results in the Forest Stream sectors being smaller in terms of surface area overall, but I believe that it captures a more diverse set of environmental factors (varied sediment grain sizes, varied *Ulva* cover) and visually appeared to contain shrimp burrow holes during site selection, unlike the proposed 120-160 meter sectors which appeared uniform in sediment grain size, lacked *Ulva* cover, and did not visually contain shrimp burrow holes.

Within each sector, a transect line was drawn from the farthest outside corner to the nearest inside corner and then reversed in the next sector so that each sample site had a “zig-zagged” line between all four sectors. This was chosen to maximize the surveyed area of each sector within the limited time of this study’s duration. Sample points were chosen at random along this line for both sites between a range of 5-15 meters using a random number generator, in

order to ensure that samples were taken about every 10 meters on average. The actual fieldwork used paces to record meter-distances along the quadrat, with our average paces being measured beforehand and found to approximate 1 meter closely enough for these methods. With this method, the Forest Stream (FT) site contained 18 sample points over a transect length of 176 meters and the False Bay Creek (FBC) site contained 30 sample points over a transect length of 295 meters.

At each sample point, two types of data were recorded- quadrat data at the precise GPS point on the sample transect and burrow data from the nearest *N. californiensis* burrow to the quadrat. Sampling for quadrat and burrow data was performed during the lowest low tide across 4 nights in November 2021, from the 10th to the 14th. Approximately 3 hours of sampling occurred on each of those nights with times ranging from 23:00 to 03:00, centered on the timing of low tide that night. The quadrats used were 0.5 meters by 0.5 meters and were placed directly at our feet as soon as we reached each randomly assigned meter distance along the transect. At each quadrat point the GPS latitude and longitude were recorded, as well as the number of *N. californiensis* burrow holes that were visible on the surface within the quadrat, the percent cover of *Ulva* spp. within the quadrat, the number of other animal species on the surface within the quadrat, and the most abundant grain size of sediment on the surface was approximated visually and by touch and categorized according to the Wentworth sediment size classification (Wentworth 1922). These were recorded as text categories (i.e. “fine sand”) and later converted to Phi values using the upper limit for each category. In addition to these observations, a sediment core of approximately 15 centimeters depth and 10 centimeters diameter was taken from every third quadrat point using a handheld trowel and returned to the lab for sediment grain size classification.

Sediment grain size categories were assessed for each core using a volume ratio method. The total wet volume of each sample was recorded in a graduated cylinder, with sediment allowed to settle in the container and ignoring any water remaining in the sediment sample. Then the sediment core was manually filtered through a stack of sieves consisting of a 2 millimeter sieve and a 500 micrometer sieve. The sieve stack containing the sediment was held partially submerged in water and shaken vigorously for approximately 1 minute per sample. Then the volume remaining in each sieve was recorded in the same graduated cylinder that recorded the total volume using the same method. The measurement from the 2 millimeter sieve is the volume of sediment in the sample with a grain size of larger than 2 millimeters. The measurement from the 500 micrometer sieve is the volume of sediment in the sample with a grain size smaller than 2 millimeters but more than 500 micrometers. The final sediment size category- the volume of sediment in the sample with a grain size of less than 500 micrometers- was obtained by subtracting the sum of the previous two measurements from the total volume of the sample. This allowed for rapid measurement of 16 total samples across both sample sites, and it also allowed for a time-efficient examination of the percentage of each sediment grain size category in each sample and each sector by extension.

The data collected at each quadrat was supplemented by data gathered from the nearest burrow to it. For the purposes of this study a burrow was arbitrarily defined as any collection of 4 or more connected *N. californiensis* burrow holes. After the quadrat data at each sample point was collected, the closest burrow to the quadrat was identified and the distance to it from the edge of the quadrat was recorded in paces of approximately 1 meter. The number of holes on the surface that were connected to the targeted burrow was recorded. If no burrows were found within 10 meters of the quadrat then no burrow data was recorded for that quadrat. Some

burrows were also not able to be located at certain quadrats due to conditions which made it impossible to see surface holes, such as the surrounding 10 meters being flooded with water or covered with *Ulva* spp. For each targeted burrow, a volume of sediment 50 centimeters deep with a diameter of 25 centimeters was excavated using either a clam gun or a shovel. All *N. californiensis* individuals within either the excavated sediment volume or the volume of pore water within the excavated hole were considered to be within the burrow and were recorded. The sex of each individual was recorded, identified by the relative sizes of its chelipeds. Males were assumed to be individuals where one cheliped was visibly drastically larger than the other, and females were assumed to be individuals where one cheliped was only slightly larger than the other. Classification of sex by secondary sex characteristics is less precise but necessary for this study- the loss in precision was outweighed by the need for fast visual identification. The carapace length (CL) for each individual was also recorded to the nearest millimeter, defined as the length from the anterior tip of the rostrum to the posterior end of the first segment. The presence or absence of egg masses on female individuals was also recorded, however ovigerous females were unlikely to be present in large numbers during the sample period (Dumbauld et al. 1996).

After sampling for quadrat and burrow data was completed, I conducted a targeted sampling of 6 burrows in Sector 1 of the False Bay Creek site in order to obtain *N. californiensis* individuals for a length/weight regression. This targeted sampling occurred from 23:00 to 02:00 on November 24th-25th, 2021. All burrows for this targeted sampling were located within approximately 10 meters of each other and were within the accuracy of the existing GPS point for FBC_30, the point at the end of the False Bay Creek sampling transect. The *N. californiensis* individuals obtained from this sampling were not included in any statistical analyses other than

the length/weight analysis. Individuals were obtained through excavation of 25 centimeter diameter by 50 centimeter depth holes using identical methods to the semi-random sampling in the quadrat-burrow data. These individuals were transferred into a small holding container filled with water for between 15 minutes and 2 hours for collection and transport before being transferred into a larger constant-flow holding tank at Friday Harbor Laboratories for 48 hours. All individuals were stored together at all times and were therefore considered to be exposed to the same biotic and abiotic conditions. 5 individuals out of 20 collected were deceased at the end of the 48 hour period. After 48 hours had passed the carapace length of each individual was measured to the nearest millimeter, and their weight was measured by a scale in grams to two decimal places. All individuals besides 5 were measured while alive; excess water was dried off with a paper towel prior to weighing. 2 of the deceased individuals were still intact with no visual signs of damage or decomposition. These individuals were weighed as well and included in the length/weight analysis. The other 3 deceased individuals were visibly damaged, decomposing, or missing body parts and were therefore not weighed or included in length/weight analysis.

2.2. Data Analysis

Data analysis was performed in Rstudio (RStudio Team 2020) using multiple packages, but the majority of processes were conducted using base R code (R Core Team 2021).

Correlations between the number of shrimp in a burrow as a response and other factors as explanatory- percent cover of *Ulva*, the primary surface sediment, the volume ratio of each grain size category, and the number of other species on the surface- were assessed using a correlation matrix generated using the `cor()` function from the `corrplot` R package (Wei & Simko 2021), and

a stepwise AIC (Akaike 1974) function from the MASS R package (Venables & Ripley 2002) also assisted in constructing a best-fit model for the factors influencing *N. californiensis* abundance. The stepwise AIC function starts from a regression model using number of ghost shrimp as the response variable against all of the measured abiotic and biotic factors and progressively simplified the model by removing factors. Per my hypothesis, a linear regression for the number of ghost shrimp versus the percent *Ulva* cover was plotted and assessed for fit using the summary() function in R to compare multiple R² and the p-value. A plot to test my second hypothesis was also constructed and a linear regression performed using the same method as for the first hypothesis, except with the variables being the number of ghost shrimp versus the percent volume of sediment grain size less than 500 micrometers. This plot was also assessed for fit using the summary() function and a comparison of the multiple R² and p-value.

Length/mass analysis was performed by assuming a $W = aL^b$ growth relationship which is common in other crustacean species (Winberg 1971). This relationship describes mass (W) in terms of length (L) multiplied by a constant (a) raised to the power (b). “b” is an index of the exponential relationship between length and weight; when it is higher than 3, this indicates that organisms are becoming heavier more quickly than they are becoming longer as their length increases (Winberg 1971). The $W = aL^b$ relationship was determined using the slope of the linear regression of the logarithmic carapace lengths and logarithmic weights, which is equal to “b” (Reeve & Huxley 1945), and a calculation which then solves the equation in terms of “a” and averages all the resultant values of “a” to find the model “a”. A second model was also constructed by performing a linear regression on the raw length and weight values (not the log values). These two models were compared using base R’s summary() function to compare the multiple R² values and p-values to determine which was the best-fitting model. The average

biomass of a single ghost shrimp was calculated according to the $W = aL^b$ relationship using the values for “a” and “b” identified during the analysis and the average carapace length of all the shrimp that were sampled.

Population estimates were performed using averages of measured data. Data were split into the sector that their GPS coordinates were contained in. An average was taken for the number of holes within each quadrat for each sector, which gave a hole density per $0.25m^2$. This density was multiplied by 4 to get the density per square meter of holes within each sector. This value was multiplied by the area for each sector to estimate a total number of holes within each sector. The number of holes in the burrow sampled at each quadrat was not included in this hole density in order to avoid introducing biases into the sampling, however this does result in some sectors having an estimated biomass and abundance of 0 despite containing measured values of shrimp. The total number of holes per sector was multiplied by another value: the per-sector number of shrimp per hole, which was obtained by dividing the average measured number of shrimp by the average number of burrow holes in the volume where they were sampled on a per-sector basis. This multiplication results in an estimate for total number of shrimp individuals for each sector- the *N. californiensis* abundance for each sector. Combined areas such as the Forest Stream Site (which includes all FT sectors) were estimated using averages of all of the relevant sector values together rather than by adding the averages of each sector together; this means that the sector values do not always add up to the value of the entire site. The estimated biomass of each sector was then calculated by multiplying the estimated abundance for each sector by the average biomass of an individual calculated during the length/weight analysis. Standard deviations for each mean and multiplication of means were calculated and displayed in the constructed figure as well.

All graphs and bar plots were constructed using the ggplot2 R package (Wickham 2016) or base R, while the table was constructed in R but exported as a .csv to Microsoft Excel (Microsoft Corporation 2021) for formatting and presentation.

3. Results

3.1. Sediment

Using the sediment volume percentages collected from every third sample point across both sites, bar plots were constructed for comparison (Fig. 2). All values for sediment grain size categories are given in terms of a percentage of the total sample volume rather than by exact values- exact values varied per sample due to human error and the inability of the trowel to dig a 100% accurate volume every time.

3.2. Individuals

Individual characteristics of the sampled shrimp across all sites were assessed to estimate the average carapace length in False Bay. The histogram and Q-Q plot generated in R (Fig. 3) showed that the frequency of carapace lengths across False Bay was reasonably normally distributed. Therefore the mean carapace length was taken to approximate the breadth of the data for the later biomass estimation using a length/weight relationship. This mean carapace length was determined to be 12.6 ± 3.8 millimeters, where 3.8 represents 1 standard deviation.

The ratio of males to females was also identified for each site. At the Forest Stream site the ratio of males to females was 2.33, while at the False Bay Creek site the ratio of males to females was 3.67. A one-proportion binomial two-tailed t-test for each site was performed using the `binom.test()` R function to see if the sex ratio was significantly different from 1.0. The Forest

Stream site had a p-value of 0.34375, while the False Bay Creek site had a p-value of 0.2101135. Therefore, neither site had a sex ratio significantly different from 1:1 and a roughly equal amount of males and females can be assumed for the population of False Bay.

3.3. Length/Mass

The length/mass relationship for *Neotrypaea californiensis* in False Bay was identified:

$$W = (7.267 * 10^{-4})L^{3.1024}$$

Mass (W) is outputted in grams and is derived from a wet weight rather than a dry weight measurement. Length (L) refers in this case to carapace length from the anterior tip of the rostrum to the posterior end of the first segment, measured in millimeters. “b” for this relationship is 3.1024, which is slightly greater than 3. This means that the growth in mass of the ghost shrimp will outpace the growth in length as the length increases. The regression for determining these values is displayed in Figure 4, which shows an R² value of 0.9153 and a p-value of less than 0.001, indicating that this regression (and the subsequent “b” value) are a good fit for these data.

The average biomass of a single shrimp in False Bay was found to be 1.88 ± 0.05 grams. This was determined using the length-weight relationship given above and the mean carapace length of 12.6 ± 2.8 mm. This per-shrimp biomass datum was later used to build an estimate of the total shrimp biomass across all sample sites.

3.4. Correlations

Very few of the explanatory variables compared in the correlation matrix (Fig. 5) were highly correlated with each other. The only variable-pairs that had a correlation value higher than

0.5 were the Phi value of surface sediment versus the percent volume of sediment grain size less than 500 micrometers, and the percent volume of sediment grain size greater than 2 millimeters versus the percent volume of sediment grain size less than 500 micrometers. The correlation value for the Phi value of surface sediment versus the percent volume of sediment grain size less than 500 micrometers was +0.52, which indicates that the variables are weakly positively correlated. The correlation value for the percent volume of sediment grain size greater than 2 millimeters versus the percent volume of sediment grain size less than 500 micrometers was -0.82, which indicates a strong negative correlation.

The strongest correlations between the number of ghost shrimp and an explanatory variable were the number of ghost shrimp versus the percent *Ulva* cover (-0.12), the number of ghost shrimp versus the number of other species on the surface (+0.12), and the number of ghost shrimp versus the percent volume of sediment grain size greater than 2 millimeters (-0.12). All of these correlation values indicate very weak correlations, most likely statistically insignificant.

The relationships represented in my hypotheses were plotted (Fig. 6 and 7) to determine if the relationship was statistically significant or explanatory of the variation in the data. These plots were the percent cover of *Ulva* versus the number of shrimp (Fig. 6) and the percentage of sediment sample volume smaller than 500 micrometers versus the number of shrimp (Fig. 7). The multiple R² value from both plots was less than 0.2, and the p-value for both plots was greater than 0.5, indicating that neither relationship was explanatory or statistically significant.

A stepwise AIC function from the MASS R package (Venables & Ripley 2002) was used to identify the best-fitting model for ghost shrimp distribution. The result of this function was a model identical to Figure 8., which has an AIC value of 119.6 compared to the starting model, which has an AIC of 134.3.

3.5. Estimates

Table 1 shows the final estimates calculated for shrimp density, abundance, and biomass for each site and sector in False Bay. The total abundance of shrimp across all sample sites is estimated at 8280.3 ± 22803.7 individuals, where 22803.7 is 1 standard deviation. The estimated total biomass of shrimp across all sites is a similarly large range, with a mean of 15529.3 ± 42824.0 grams, where 42824.0 is 1 standard deviation. Estimates per sector and per site do not always add together accurately, due to each estimate using an independent calculation based on averages rather than calculated the average of a site as the sum of averages of the sector. Some sectors, such as Forest Stream Sectors 2 and 3 and False Bay Creek Sector 3 show estimated values of 0 ± 0 for all categories despite having an N, or actual measured number of shrimp individuals, of >0 . This is due to the design of the methods for estimating shrimp density, which use an average of the number of holes in each quadrat to calculate an estimated number of holes per area. If none of the quadrats in a sector had a number of surface holes greater than 0, the estimated number of holes in the sector will therefore be 0 even if shrimp burrows do exist in the sector. The number of holes per burrow was not included in this holes per area calculation in order to avoid bias, as burrows were intentionally targeted and the number of holes in the burrow area may not be representative of the number of holes in the rest of the sector.

4. Discussion

The first part of my experimental question- “What is the abundance of *N. californiensis* in False Bay”- has begun to be answered by my data. Table 1 shows estimates of ghost shrimp abundance across my sample sites as well as several additional variables which were added later in the study plan and provide good “bonus” data that can act as a baseline for future research. Standard deviations of the means represented in Table 1 are high- often more than double the value of the mean they are associated with, such as in the case of the total abundance across all sectors and sites, which is 8280.3 ± 22803.7 . This wide spread is possibly indicative of the low number of samples taken per sector, in which case more data is needed. More relevantly, the high values of standard deviation indicate that there is a high amount of variability in the distribution of ghost shrimp throughout False Bay. Whether this variability is linked to any abiotic or biotic factors remains unknown; see the weak correlations between examined variables in Figure 5.

The biomass-carapace length relationship quantified in this study is the first analysis of its type in False Bay, and while more data would be needed to confirm its accuracy, the ability to apply a biomass number to the abundance of shrimp could be very useful for future ecological research in False Bay. Standard deviations for these mean values are similarly high, though this may simply be the result of the standard deviation of abundance being factored into the propagation of error through the biomass. The high variability of biomass across my sample sites is therefore linked to the high variability of abundance, which makes intuitive sense and means that improving the abundance estimate in the future will also increase the certainty of any biomass prediction as well.

The most surprising aspect of this study was the lack of any significant correlations between abiotic or biotic factors versus the number of ghost shrimp. I hypothesized that both the percent *Ulva* cover and the percentage of sediment that was less than 500 micrometers would

have a significant correlation with the number of ghost shrimp. In the case of percent *Ulva* cover, I thought that a negative correlation would occur, and with the percentage of small sediment grain size I believed a positive correlation would occur. As shown in figures 6 and 7, both of these hypotheses were incorrect- there was no significant correlation between either variable and the number of ghost shrimp. The stepwise AIC analysis also failed to yield any significant relationships between any of the environmental factors examined and the number of ghost shrimp- the best-fitting model identified was the aforementioned figure 7, which is neither explanatory nor significant. In more plain terms, this means that no possible combination of surface sediment grain size, percent *Ulva* cover, sediment size category, and number of surface species can be shown to significantly effect the spatial distribution of ghost shrimp across the sample sites I examined. It is possible that with more data points stronger correlations could be determined, as many of the points examined were unable to have a paired burrow assigned to them.

During the sample period, a number of environmental factors and methodological limitations made obtaining the full intended number of samples impossible and results in somewhat patchy data. Some quadrat points in the stream channel of False Bay Creek were entirely flooded with water, which makes identifying shrimp burrow holes impossible and therefore no burrows were able to be paired with the quadrat data. The quadrat-burrow pairs affected by this phenomenon were excluded from analysis to avoid skewing data too low, but this does result in FBC sector 2 containing only one valid data point, which means the potential for error is high in that sector. A similar issue occurred in FT sector 1, where the *Ulva* cover was dense enough and deep enough to make identifying burrow holes underneath it unfeasible within

the time limitations of the study. Thus FT sector 1 contains an artificially lower number of samples compared to the other FT sectors.

Other limitations on this study design include the narrow temporal window- sites were examined and sampled within a period of only 4 days, which means there is no assurance that temporal anomalies are not present in the data. This is an unfortunate and unavoidable issue imposed by the strict time limit of the Biodiversity and Monitoring of Estuarine Ecosystems course and this study's nature as a graded project for that course. The area examined by this study is also small compared to the total area of False Bay, and it is possible that the relationships uncovered in my data are not representative of the entirety of the sandflat.

Despite these limitations and issues, I believe the data set that I have gathered over the course of this project is an important first step to understanding the place and importance of *Neotrypaea californiensis* in False Bay. Over the course of this study I have established a solid baseline data set and methodology that is repeatable and has room for further expansion in the future. With the establishment of this data, future research in False Bay can focus on quantifying the relationships between ghost shrimp and other organisms in more certain terms.

5. Future Directions

Research on *Neotrypaea californiensis* in False Bay is only beginning. The baseline data and methodologies established in this study are new and statistically imprecise thus far, but very promising. With a larger number of samples in each area of this study, the abundance and biomass data obtained here in 2021 could be refined over future years of data collection, decreasing the standard deviation and standard error of my measurements and approaching mean values closer to the true population abundance and biomass of *N. californiensis* in False Bay.

The methods in this paper were designed to be fast and easy to apply compared to the more precise methods used in papers such as Dumbauld et al. 1996. My budget in terms of both the time and equipment that was available for this project was not high, but I was still able to obtain relevant results despite the small amount of time available to me and lack of sophisticated coring equipment for shrimp sampling. For this reason I believe that my sampling method and locations were an overall success and can be built upon in the future.

The lack of large correlations between *Ulva* cover and the number of shrimp or the percent volume of less than 500 micrometer sediment and the number of shrimp was surprising, and this lack of large correlations between any of my explanatory variables and the number of shrimp means that the main driving biotic and abiotic factors of *N. californiensis* spatial distribution remain unknown. The spatial distribution of ghost shrimp is not explained well by any category of sediment grain size, the number of other species on the surface, or the percent cover of *Ulva* on the surface. Future studies could focus on other potential explanatory variables, such as the characteristics of pore water beneath the surface or the abundance of other infauna.

6. Figures

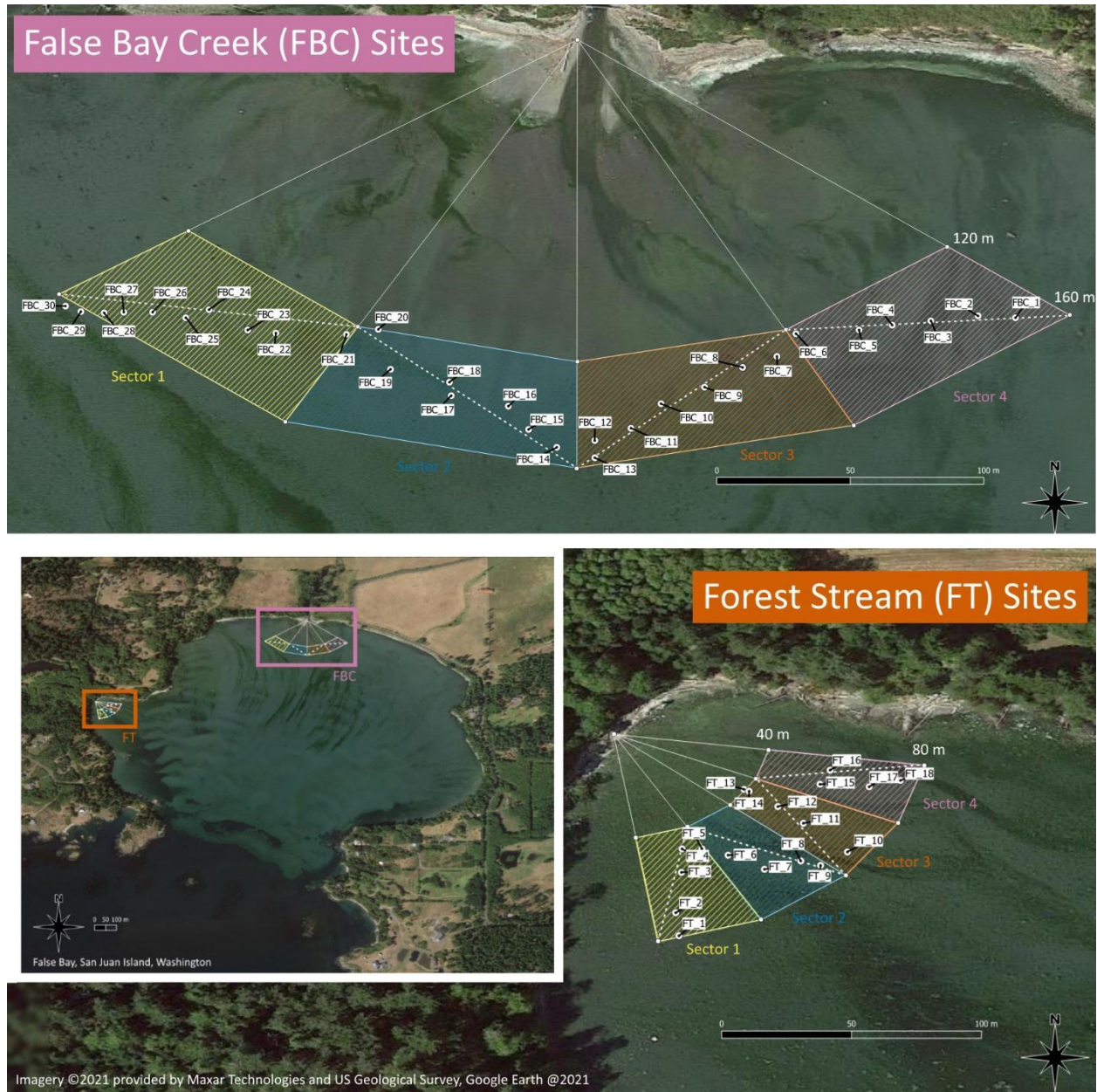


Figure 1. Map of Sample Sites in False Bay. Map shows an inset of the entirety of False Bay on San Juan Island, Washington, USA with both sample sites labeled and color-coded. The top part of the image shows a detailed view of the False Bay Creek sample sites, with each sector labeled and color-coded. The bottom right part of the image shows the Forest Stream sample sites, with each sector labeled and color-coded in the same manner as at False Bay Creek. Both images show the meter length of the point along the transect that the sectors were constructed from. Additionally, the planned sampling transect is shown as a dotted line while the actual GPS coordinates of each sample are shown and labeled according to the unique label of the sample in the data. Maps were constructed using Google Earth satellite imagery (Google 2021) and QGIS (QGIS Development Team 2021) software.

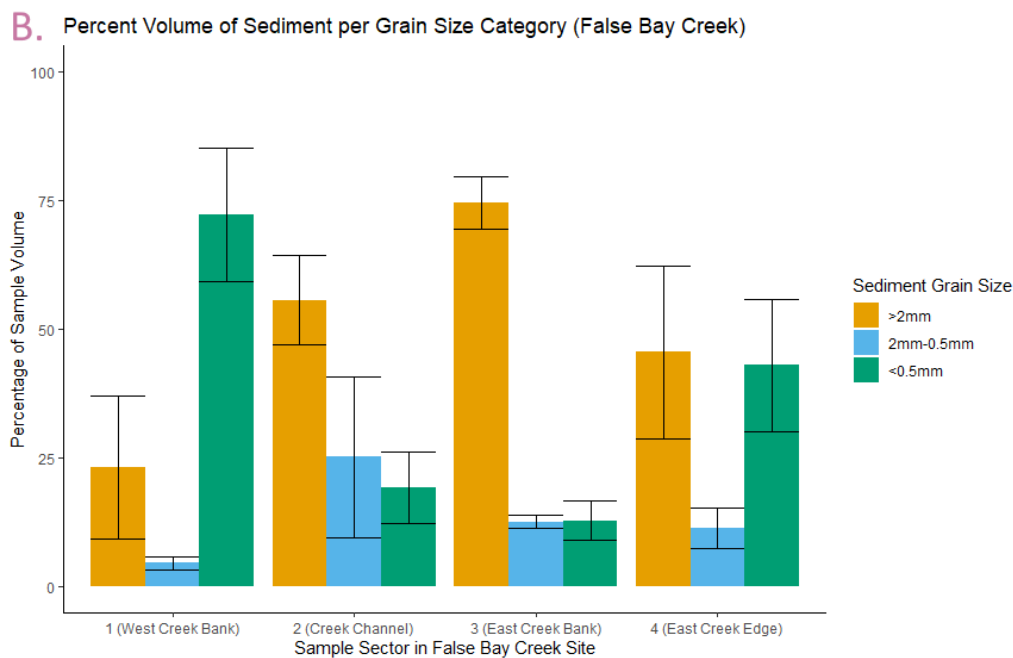
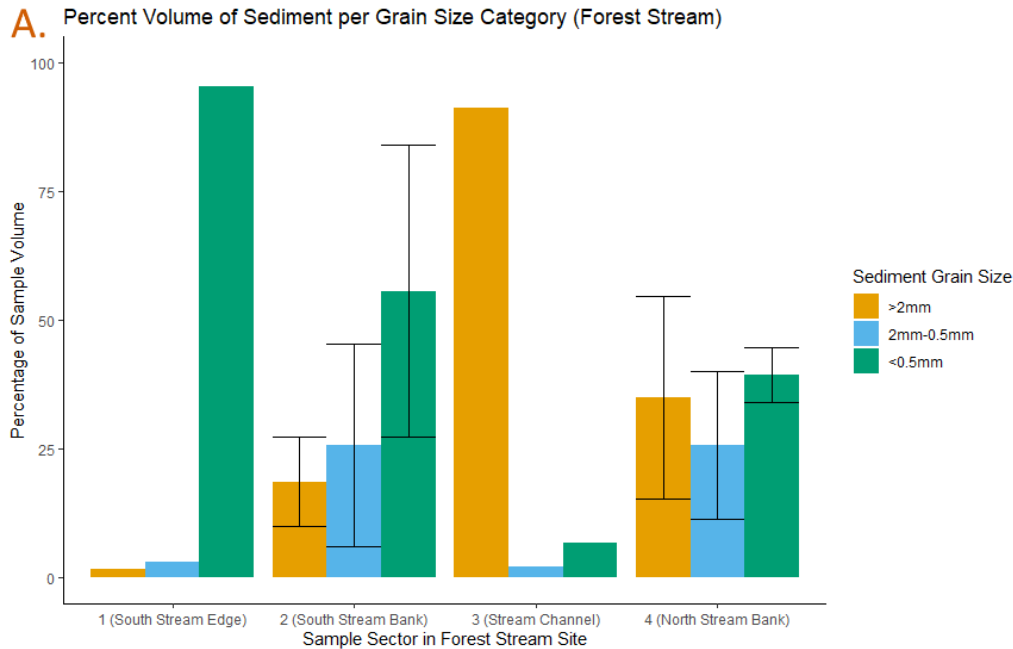
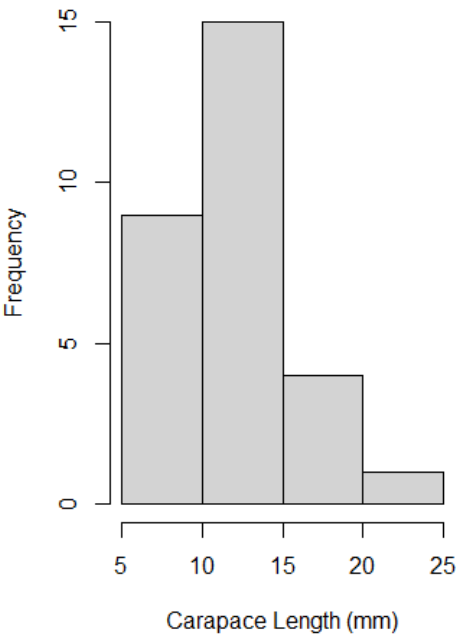


Figure 2. Comparison of Percent Volume of Sediment Grain Sizes Across Sites. Figure shows the percent volume of the total sample made up by each sediment grain size category. Fig 2A is the percent volume of sediment data for sites in the Forest Stream, while Fig 2B is the percent volume of sediment data for sites in the False Bay Creek sectors. Error bars represent the standard error of the mean values. Forest Stream sectors 1 and 3 contained only one sediment sample, so there is no standard error available for those measurements. Percent volumes were calculated as the mean of all of the percent volumes of each sample point that lies within the given sector. Data were analyzed in RStudio (RStudio Team 2021) using ggplot2 (Wickham 2016) and base R functions (R Core Team 2021).

Histogram of Carapace Lengths over All Sites



Normal Q-Q Plot of Carapace Length

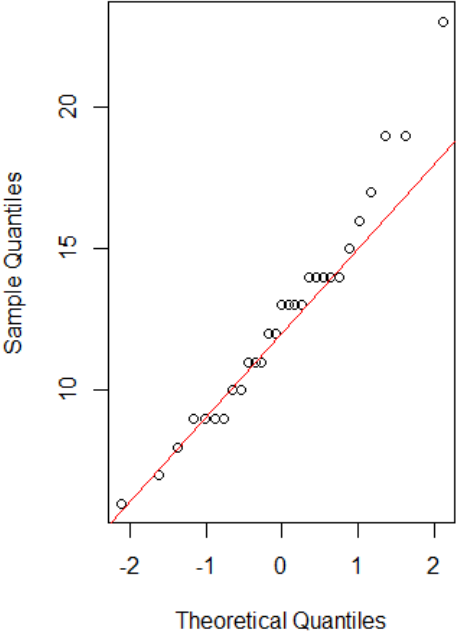


Figure 3. Histogram and Q-Q Plot of Carapace Length Across All Sites. Figure shows the frequency of carapace length values in millimeters measured from all shrimp across all sample sites on the left and a normal Q-Q plot of carapace length data on the right to assess the normality of the frequency distribution. Data were analyzed in RStudio (RStudio Team 2021) using base R functions (R Core Team 2021).

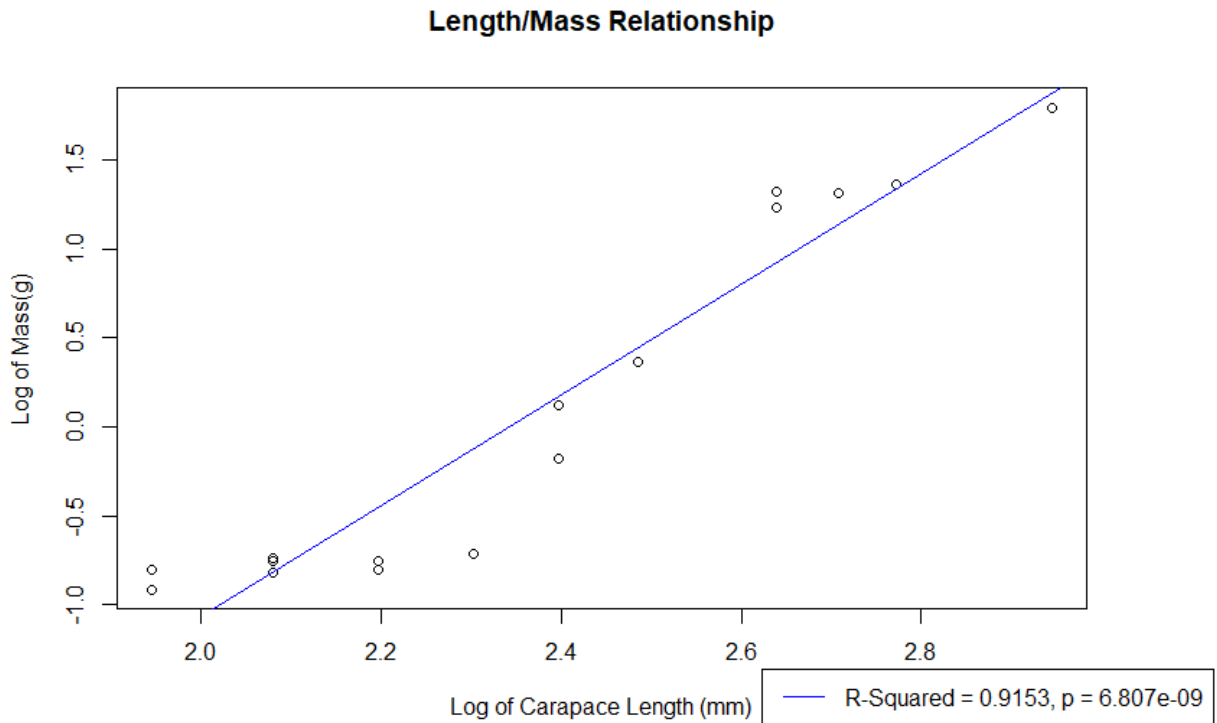


Figure 4. Linear Regression of Log Length/Mass Values. Figure shows the linear regression between the logarithm (base 10) of carapace length in millimeters versus the logarithm (base 10) of wet mass in grams for *Neotrypaea californiensis*. The multiple R^2 and p-value for the regression are displayed in the lower left corner. All values were obtained from 17 *N. californiensis* individual gathered specifically for this length/mass relationship from burrows within 10 meters of FBC_30. Data were analyzed in RStudio (RStudio Team 2021) using base R functions (R Core Team 2021).

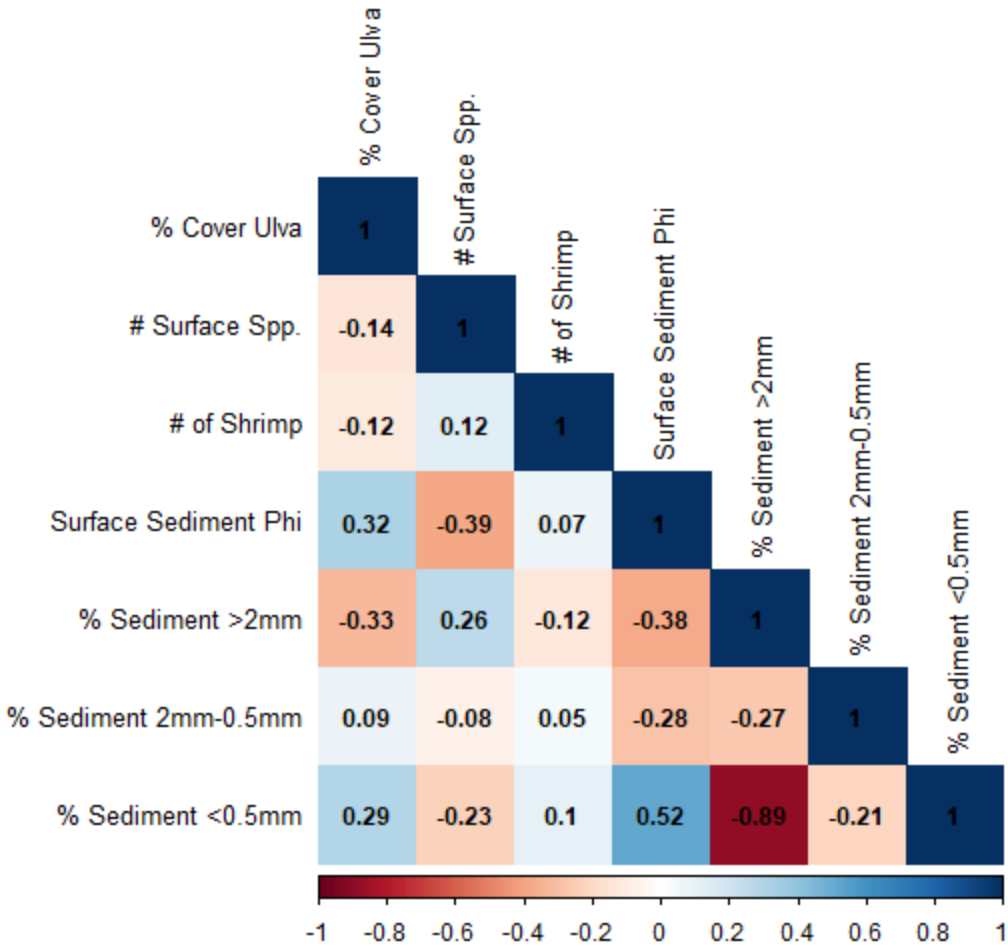


Figure 5. Correlation Matrix of Biotic and Abiotic Factors. Figure shows a matrix comparing the correlation values of biotic and abiotic factors measured at each sample site, generated using the `cor()` function in the `corrplot` R package (Wei and Simko, 2021). Number of shrimp refers to the count of *N. californiensis* individuals in the burrow paired with the sample point. Higher absolute values indicate a stronger correlation, and the positive or negative state of the value is the direction of the correlation. Correlation values for biotic and abiotic factors versus the number of *N. californiensis* individuals are shown to be relatively weak. Data were analyzed in RStudio (RStudio Team 2021).

Relationship between % Ulva cover and # of *N. californiensis*

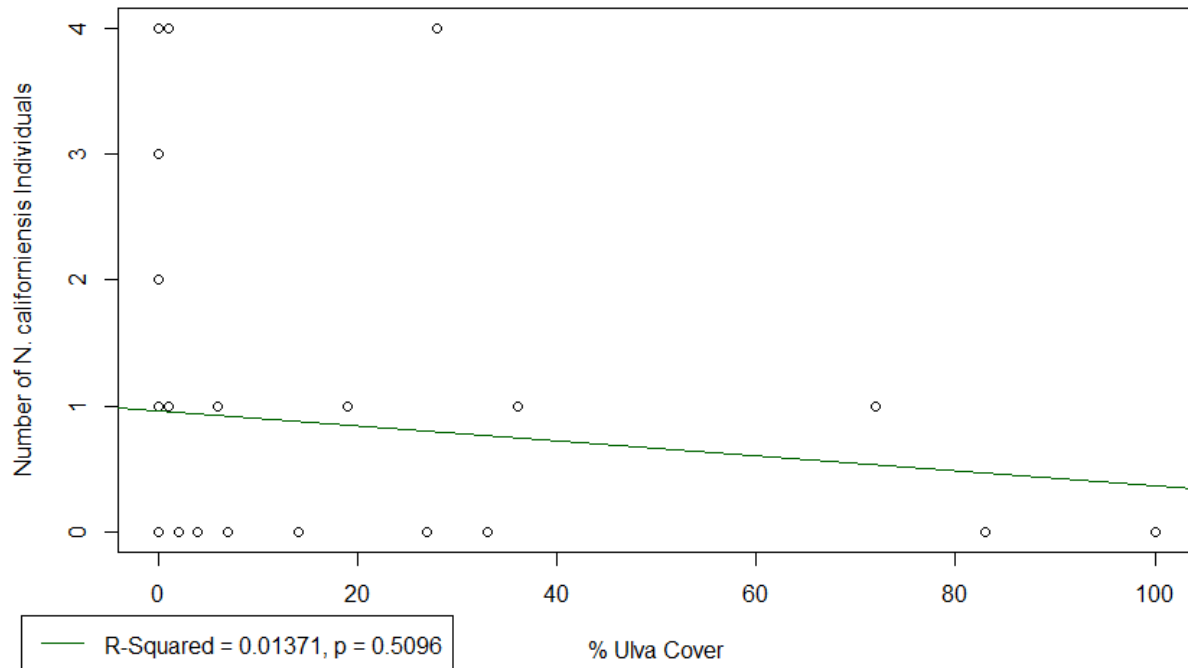


Figure 6. Linear Regression Model of Percent Ulva Cover vs. Number of *N. californiensis* Individuals. Figure shows a linear regression between the per-sample values of percent *Ulva* spp. cover and the number of *N. californiensis* individuals found in the burrow paired with the sample point. The multiple R^2 value and p-value displayed for the regression in the lower left suggests that this relationship is not explanatory or statistically significant between these two factors. Data were analyzed in RStudio (RStudio Team 2021) using base R functions (R Core Team 2021).

Relationship between Small Grain Sizes and # of *N. californiensis*

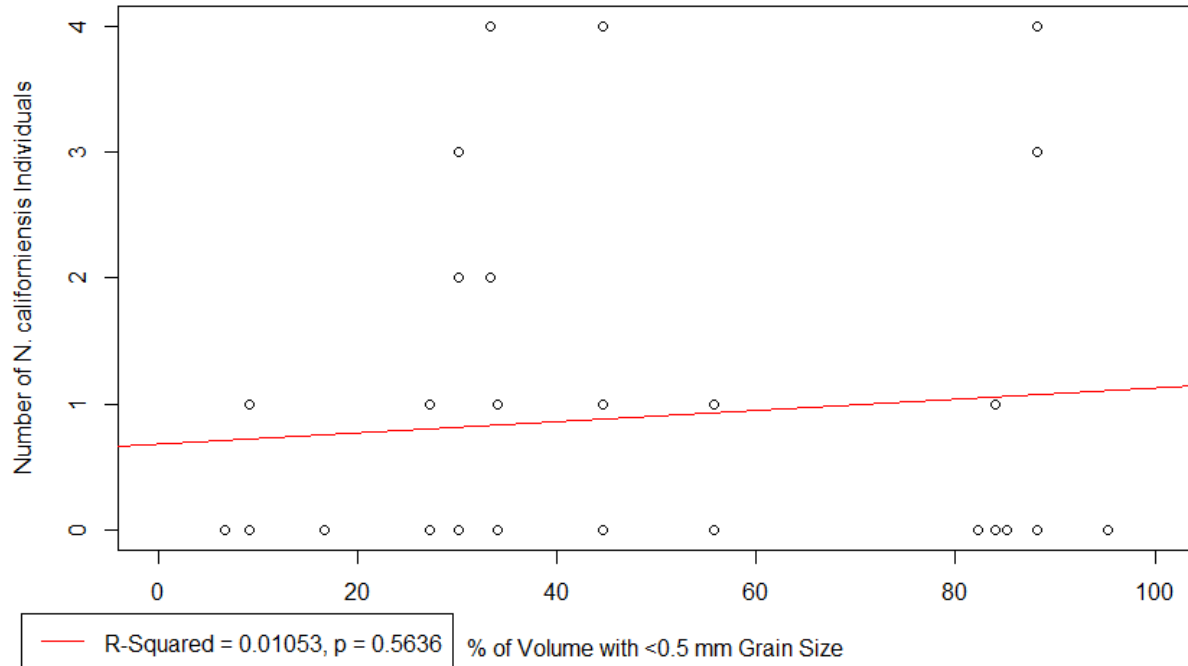


Figure 7. Linear Regression Model of Percent Volume of Sediment Grain Size Smaller Than 500 micrometers vs. Number of *N. californiensis* Individuals. Figure shows a linear regression between the per-sample values of the number of *N. californiensis* individuals found in the burrow paired with the sample point versus the percentage of the total volume of the sediment sample that had a grain size smaller than 500 micrometers. The multiple R² value and p-value is displayed for this regression in the lower left and suggests that this relationship is not explanatory or statistically significant between these two factors. This relationship was also identified by the stepwise AIC function in the MASS R package (Venables & Ripley 2002) as the best-fitting model for *N. californiensis* distribution after stepwise reduction in model complexity starting from a regression model taking all biotic and abiotic factors as explanatory to the number of *N. californiensis* individuals, with an AIC of 119.6 compared to the highest AIC of 134.3 (corresponding to the all-inclusive model that the stepwise function started with). Data were analyzed in RStudio (RStudio Team 2021) using base R functions (R Core Team 2021).

Table 1. *Neotrypaea californiensis* Population Abundance and Biomass Estimates. Table shows population estimates in terms of density, abundance, and biomass, as well as measured values of area and sampled shrimp for each location. The area of each sector as well as the combined area of each site and of all sites is shown in square meters. The “Shrimp Density” column, given in individuals per square meter, represents the shrimp per square meter density of the sector, calculated using the estimated abundance of the sector divided by the area of the sector. Abundance, given as a number of individuals, was estimated using the average of the measured number of surface holes per quadrat generalized to an average surface holes per unit area, multiplied by the area of the sector to get a total number of burrow holes in the sector. This was multiplied by the average of the measured shrimp per burrow hole value to obtain a number of individuals. Biomass, given in grams, was estimated using the formula $W = (7.267 \times 10^{-4})L^{3.1024}$, which was derived from a linear regression of log values. N is the measured number of shrimp that were sampled across all burrows in the given location. All values after the “±” symbol represent one standard deviation from the mean value. Values marked with an asterisk (*) do not have a standard deviation because only one data point is available for that sector. Data were analyzed in RStudio (RStudio Team 2021) and formatted for display in Microsoft Excel (Microsoft Corporation 2021).

Location	Area (sq m)	Shrimp Density (#/sq m)	Abundance (#)	Biomass (g)	N (#)
All Sites	12761	0.6 ± 10.9	8280.3 ± 138553	15549.3 ± 138553	29
All Forest (FT) Sectors	3069	0.8 ± 15.4	2385 ± 47264.2	4478.6 ± 47264.2	9
FT 1 South Stream Edge	826	0.3 ± 8.9	275.3 ± 7388	517 ± 7388	1
FT 2 South Stream Bank	828	0 ± 0	0 ± 0.1	0 ± 0.1	2
FT 3 Stream Channel	721	0 ± 14.3	0 ± 10318.1	0 ± 10318.1	0
FT 4 North Stream Bank	694	4.8 ± 13.6	3321.8 ± 9447.9	6237.8 ± 9447.9	6
All False Bay Creek (FBC) Sectors	9692	0.5 ± 9.4	5330.2 ± 91288.8	10009.4 ± 91288.8	20
FBC 1 West Creek Bank	2497	1.5 ± 13.6	3868.4 ± 34017.9	7264.3 ± 34017.9	11
FBC 2 Creek Channel	2515	1.1 *	2874.3 *	5397.5 *	2
FBC 3 East Creek Bank	2388	0 ± 0	0 ± 0.1	0 ± 0.1	1
FBC 4 East Creek Edge	2292	0.1 ± 1.6	318.3 ± 3742.8	597.8 ± 3742.8	6

References

- Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Armstrong, J. L., Armstrong, D. A., & Mathews, S. B. (1995). Food habits of estuarine staghorn sculpin, *Leptocottus armatus*, with focus on consumption of juvenile Dungeness crab, *Cancer magister*. *Fishery Bulletin*, 93(3), 456–470.
- Berkenbusch, K. & Rowden, A. A. (2007). An examination of the spatial and temporal generality of the influence of ecosystem engineers on the composition of associated assemblages. *Aquatic Ecology* 41, 129-147.
- Dewitt, T. H., Andrea, A. F. D., Brown, C. A., Griffen, B. D., & Eldridge, P. M. (2004). Impact of burrowing shrimp populations on nitrogen cycling and water quality in western North American temperate estuaries . In : A . Tamaki (ed .), Proceedings of the Symposium on Ecology of Large Bioturbators in Tidal Flats and Shallow Sublittoral Sedi. Quality, January, 107–118.
- Dumbauld, B. R., & Wyllie-Echeverria, S. (2003). The influence of burrowing thalassinid shrimps on the distribution of intertidal seagrasses in Willapa Bay, Washington, USA. *Aquatic Botany*, 77(1), 27–42. [https://doi.org/10.1016/S0304-3770\(03\)00077-9](https://doi.org/10.1016/S0304-3770(03)00077-9)
- Dumbauld, B. R., Armstrong, D. A., & Feldman, K. L. (1996). Life-history characteristics of two sympatric Thalassinidean shrimps, *Neotrypaea californiensis* and *Upogebia pugettensis*, with implications for oyster culture. *Journal of Crustacean Biology*, 16(4), 689-708.
- Feldman, K. L., Dumbauld, B. R., Box, P. O., Park, O., Dewitt, T. H., Doty, D. C., & North, C. W. (2000). Oysters , Crabs , and Burrowing Shrimp : Review of an Environmental Conflict Over Aquatic Resources and Pesticide Use in Washington State ' s (USA)

- Coastal Estuaries School of Fisheries 2111 Southeast Marine Science Drive. *Estuaries*, 23(2), 141–176.
- Google. (2019). Google Earth Imagery. 48.485963 N, -123.069218 W. San Juan Island, WA, USA.
- Hoffman, C. (1981). Associations between the Arrow Goby *Clevelandia ios* (Jordan and Gilbert) and the Ghost Shrimp *Callinassa californiensis* Dana in Natural and Artificial Burrows. 35(3), 211–216.
- Microsoft Corporation. (2021). Microsoft Excel. Retrieved from <https://office.microsoft.com/excel>
- Nedergaard, R. I., Risgaard-Petersen, N., & Finster, K. (2003). Erratum: The importance of sulfate reduction associated with *Ulva lactuca* thalli during decomposition: A mesocosm experiment (Journal of Experimental Marine Biology and Ecology (2002) 275 (15-29) PII: S0022098102002113). Journal of Experimental Marine Biology and Ecology, 283(1–2), 149. [https://doi.org/10.1016/S0022-0981\(02\)00564-6](https://doi.org/10.1016/S0022-0981(02)00564-6)
- QGIS Development Team (2021). QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>
- Reeve, C. & Huxley J. (1945) Some problems in the study of allometric growth. Essays on growth and form, 111-156.
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA <http://www.rstudio.com/>.

Tamaki, A., Kagesawa, T., Takeuchi, S., Ohashi, H., Yang, S., & Sassa, S. (2018). Facultative commensalism of a free-burrowing urothoid amphipod with a deep burrow-dwelling callianassid shrimp in intertidal sand. *Marine Biology*, 165(2), 1–15.

<https://doi.org/10.1007/s00227-018-3296-0>

Wei, T. & Simko, V. (2021). R package 'corrplot': Visualization of a Correlation Matrix (Version 0.92). Available from <https://github.com/taiyun/corrplot>

Wentworth, C. K. (1922). A Scale of Grade and Class Terms for Clastic Sediments

The Journal of Geology, 30(5), 377–392. <http://www.jstor.org/stable/30063207>

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.

Winberg, G. G. (1971). Growth, rate of development and fecundity in relation to environmental conditions. *Methods for the estimation of production of aquatic animals*, 33-64.

Venables, W. N. & Ripley, B. D. (2002). *Modern Applied Statistics with S*. Fourth Edition. Springer, New York. ISBN 0-387-95457-0