Linear Models Describing Female Dungeness Crab Carapace Size Before and After Molting

By Alana R. McGraw

Abstract

Dungeness crabs are one of many marine creatures that have been threatened by overfishing. Size restrictions are in place in order to allow male crabs to mate before they are captured. If the law changes, allowing females to be fished, then it will be important to understand the growth patterns of adult female Dungeness crabs in order to create new size restrictions. In this report we will study the growth patterns of female Dungeness crabs in order to help generate size restriction recommendations for the fishing of female Dungeness crabs. Using Mathematica, we numerically and graphically summarized the pre molt and post molt carapace sizes of Dungeness crabs from capture-recapture sampling, laboratory sampling, and the mixture of both. We generated three linear models of the post molt size on the pre molt size from the capture-recapture, laboratory, and mixture data. All three linear models have very high R-squared values, at 0.9808, 0.9810, and 0.9328 for the mixed, laboratory, and capture-recapture models respectively, but none of the models followed the assumptions of homoskedasticity or normally distributed residuals. The high R-squared values indicate that the data strongly follows these models, but the inability to follow the assumptions of the linear regression suggests that some caution be used when using these models to predict pre molt sizes of Dungeness crabs as some of these predictions may be unreliable.

Introduction

Overfishing has been a threat to marine creatures for decades. Dungeness crabs (*Cancer magister*) are one of the largest and most abundant crabs on Pacific coast; however, every year in US waters, nearly all adult male Dungeness crabs are fished between the months of December and June. There are some restrictions in place to protect the viability of the Dungeness crab population in the US: female crabs are not fished, and male crabs have size restrictions so they can have at least one chance to mate before they are caught. However, it has been argued that the imbalanced sex ratio has a role in the large fluctuations of yearly catches and the declining crab population. In Canada, they catch male and female Dungeness crabs and do not experience the same yearly fluctuations, but more needs to be known about the female crabs growth in order to create protective size restrictions if the fishing laws are changed in the US.

Dungeness crabs have a hard exterior shell that covers their backs, and provides them with protection; this is called a carapace. The crabs periodically molt their shells in order to accommodate their growth. They do this by shedding their old shells and taking in lots of water as they generate new shells. Once the new shell is hardened they release the water and shrink back to their original size, with more room to grow in their new shells. Right after the molting season, shells are clean, indicating that the crab inside is likely smaller than its outer shell indicates.

More information about adult female Dungeness crab size before and after molting is needed to determine the female crabs growth patterns. Information about the increase in shell width based only on the size of the shell after molting would be particularly helpful because it would allow predictions for how much the shell size changed. In this report we will study the growth patterns of female Dungeness crabs in order to help generate size restriction recommendations for the fishing of female crabs.

Methods

The data used in this study was found at the UC Berkeley Stats Lab website. The two data sets used in this report are part of a study of female Dungeness crabs. The first set has pre molt and post molt widths of carapaces (shells) from 472 female Dungeness crabs collected from laboratory and capture-recapture data over three seasons (1981,1982, and 1992). Capture-recapture crabs were tagged and measured before molting and returned for measurements after molting. Crabs from the laboratory data were first collected and measured at the beginning of the molting season, and then measured again a few days after molting. The second data set was collected from 362 crabs in 1983 after the molting season. The measurements for each data set were made on the widest part of the external carapace including spines.

The first data set entitled “Crab molting” and second entitled “Sample of crabs” were downloaded and imported into Mathematica, the program used for this data analysis. All entries without data were removed from both data sets. The order to the first two columns, “presz” and “postsz” was reversed. Several subsets of data were made for the “Crab molting” data set after reversing the order of the first two columns: (i) both the pre and post molt data of all sample (PostPre), (ii) pre molt of all samples (Pre All), (iii) post molt of all samples (Post All), (iv) both the pre and post molt data of laboratory sample (Lab), (v) pre molt of laboratory samples (Pre Lab), (vi) post molt of laboratory samples (Post Lab), (vii) both the pre and post molt data of capture-recapture samples (Ocean), (viii) pre molt of capture-recapture samples (Pre Ocean), and (ix) post molt of capture-recapture samples (Post Ocean).

Descriptive statistics were calculated for subsets (ii), (iii), (v), (vi), (viii), and (ix) to describe the data center, spread, and distribution. These calculations include the mean, standard deviation, skewness, kurtosis, and five-point summary (Table 1). Histograms for each of these subsets was created as a graphical representation of the carapace sizes (Figure 1). Smooth histograms were generated for each of these subsets; (ii) and (iii) were displayed together and (v), (vi), (viii), and (ix) were displayed together to compare the distributions between groups (Figure 2) . A side by side box and whisker plot was constructed for these data sets to graphically represent data spread, dispersion, and outliers (Figure 3). Quantile plots were generated for these data sets to graphically compare the data distribution to the normal distribution (Figure 4).

Linear models were generated from the PostPre, Lab, and Ocean data subsets, generating regression lines of pre molt on post molt. Graphic representations were made for each linear model (Figure 5).The R-squared was calculated for each model to determine how close the data is to the fitted regression line. The mean, standard deviation, skewness, kurtosis, and five-point summary were calculated for the residuals of each model. List plots and smooth histograms overlain with the corresponding normal distribution were generated in order to graphically represent the residuals from each of these linear models. Quantile plots were created for the residuals to graphically compare their distribution to the normal distribution.

The Lab linear model was used to predict the premolt sizes of the crabs from the "Sample of crabs" data set. The predicted pre molt carapace sizes were graphically represented with a histogram.

Results

The mean, standard deviation, skewness, kurtosis, and five-point summary are described for subsets (ii), (iii), (v), (vi), (viii), and (ix) . The histograms and a side-by-side box and whisker plot for these data subsets visually describe the distribution of the data (Figure 1 and Figure 3). The maximum, minimum, median, upper quartile, and lower quartile numerically summarize the data distribution (Table 1). The skewness of all the data sets is negative. All kurtosis values are above 3 but the values for the pre (4.76) and post (5.24) molt ocean subsets are the closest to 3. The quantile plots for (ii), (iii), (v), (vi), and (viii) curve downward at the top and bottom but the middles follow the theoretical quantile. The quantile plot for (ix) curves downward at the bottom but the remainder follows the theoretical quantile.

Table 1. Description of crab carapace sizes in mm before (Pre) and after (Post) molting including mean, standard deviation, skewness, kurtosis and a five-point summary. Data was analyzed from laboratory samples (Lab), ocean samples (Ocean), and both groups together (All).

	Pre All	Post All	Pre Lab	Post Lab	Pre Ocean	Post Ocean
Mean	129.2	143.9	126.2	141.1	139.0	153.0
Standard Deviation	15.9	14.6	16.6	15.3	7.3	6.7
Skewness	−2.004	−2.347	−1.889	−2.288	−1.111	−1.1190
Kurtosis	9.766	13.12	9.024	12.44	4.761	5.241
Five−Point Summary						
Minimum	31.1	38.8	31.1	38.8	113.6	127.7
2nd Quartile	121.6	137.9	119.3	135.2	136.1	150.0
Median	132.8	147.4	128.9	143.7	140.1	154.0
3rd Quartile	140.0	153.4	137.4	150.8	143.9	157.0
Maximum	155.1	166.8	155.1	166.8	153.9	166.5

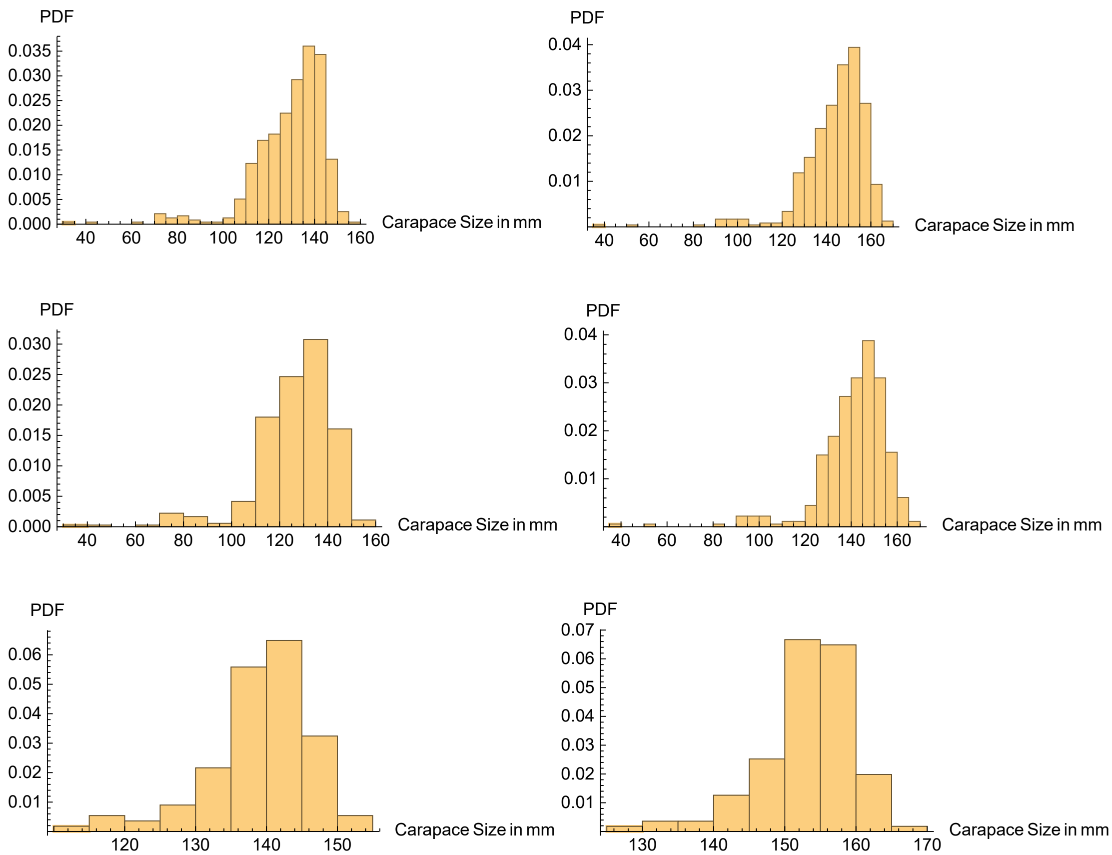


Figure 1. Carapace size data for Dungeness crabs before (Left) and after (Right) molting. The top describes the carapace sizes for the mixture of Ocean and Lab samples, the middle describes the Lab sample, and the bottom describes the Ocean sample.

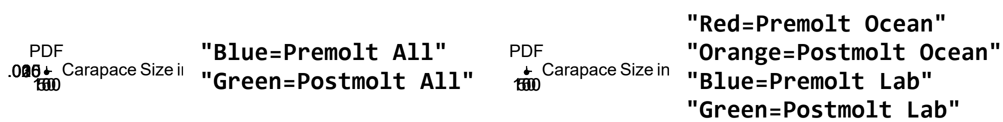


Figure 2. Carapace sizes for Dungeness crabs before and after molting. (Top) The pre and post molt sizes for the mixture of Lab and Ocean samples (All). (Bottom) The pre and post molt sizes for the Lab samples and the Ocean samples.

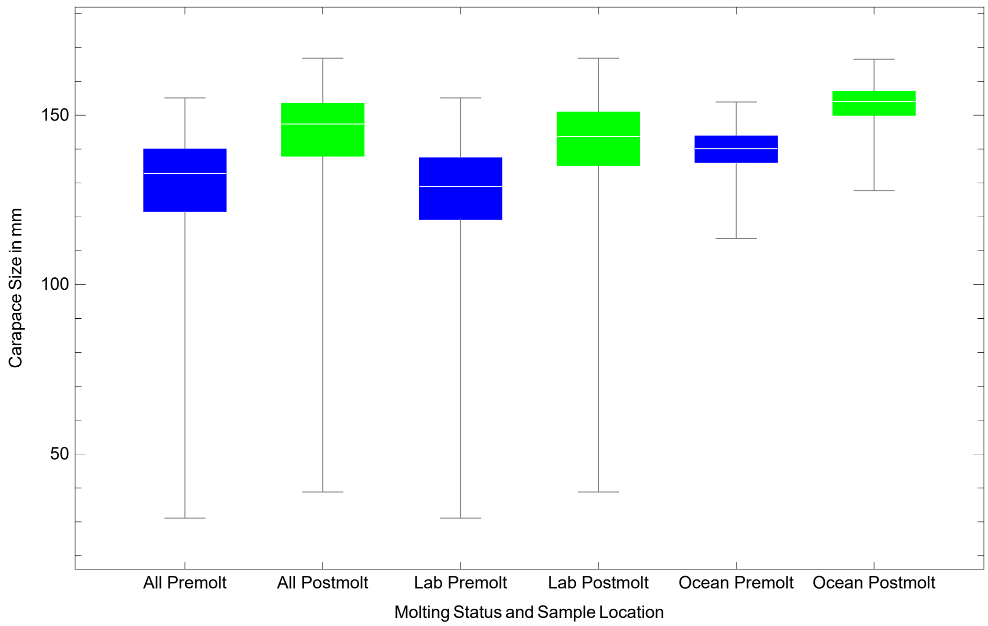


Figure 3. Pre and Post molt carapace sizes of Dungeness crabs measured from the Lab, Ocean, and a mixture of Lab and Ocean samples (All).

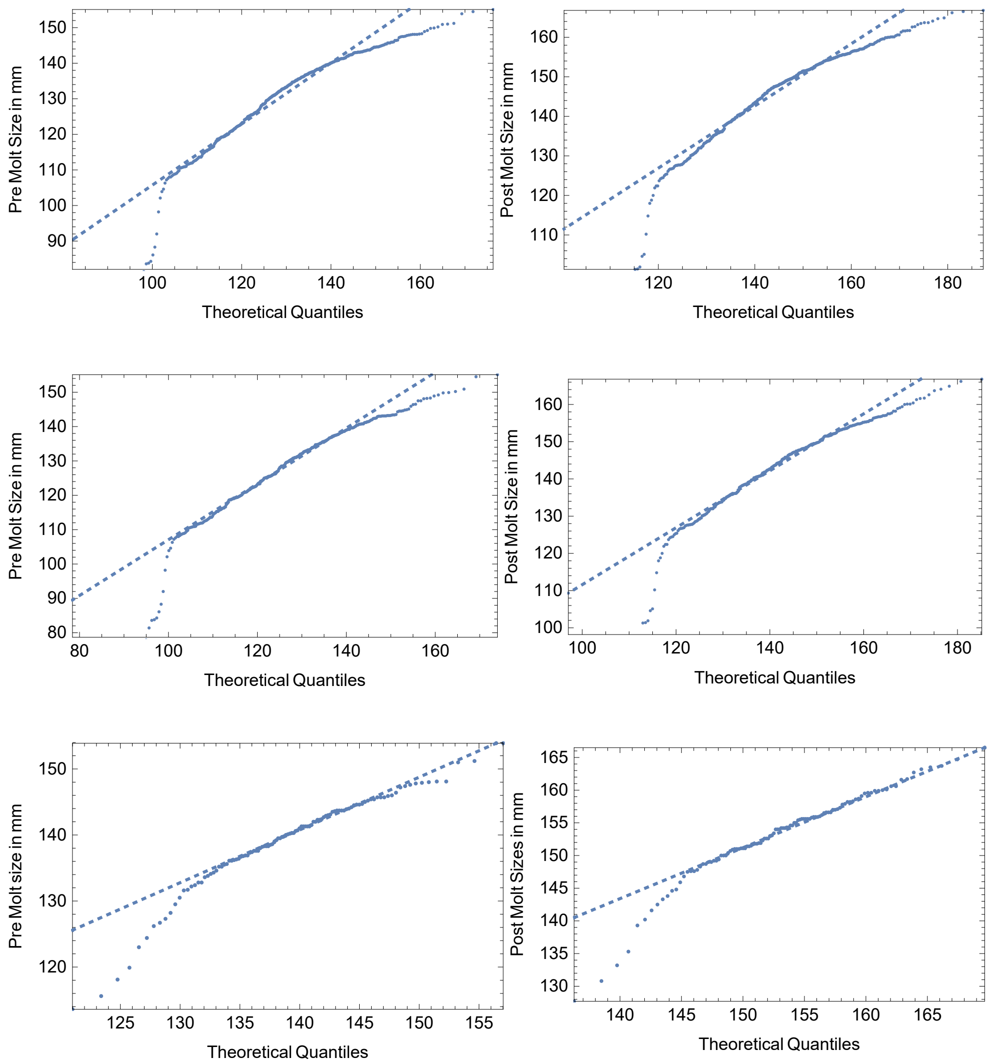
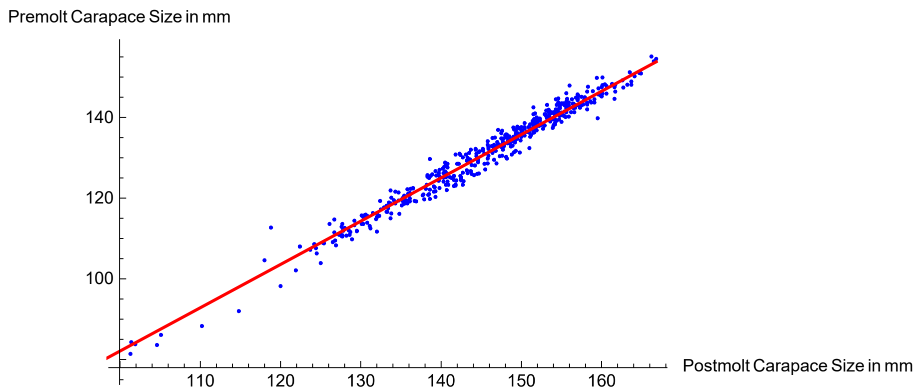
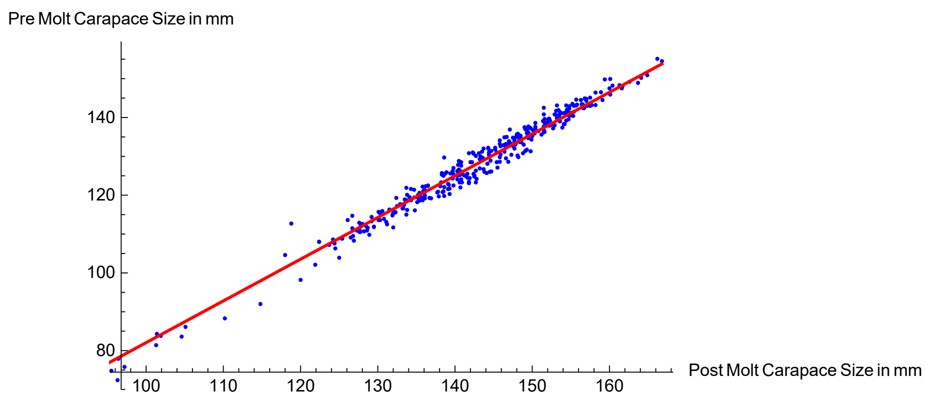


Figure 4. The quantile plots for pre (Left) and post (Right) molt sizes of Dungeness crab. (Top) Data that is a mixture of Lab and Ocean crabs. (Middle) Crabs measured and grown in the Lab. (Bottom) Crabs measured from Ocean sampling.

The post molt vs pre molt linear models generated from the mixture of Lab and Ocean measurements, just Lab measurements, and just Ocean measurements of the Dungeness crabs are graphically displayed in Figure 5. The R-squared value of the mixed data linear model is 0.9808. The R-squared from the laboratory linear model is 0.9810. The R-squared for the capture-recapture linear model is 0.9328

Table 2 summarizes the mean, standard deviation, skewness, kurtosis, and five-point summary of the residuals for the mixed, Lab, and Ocean linear models. None of the kurtosis values are close to 3, but the kurtosis value of the Ocean residuals is the closest at 3.913. The residuals do not have a constant variance from their mean (Figure 6). The quantile plots for the residuals of all three models curve downward on the left side and upward on the right side but they maintains a linear relationship with the theoretical quantiles in the middle of the plot (Figure 7).





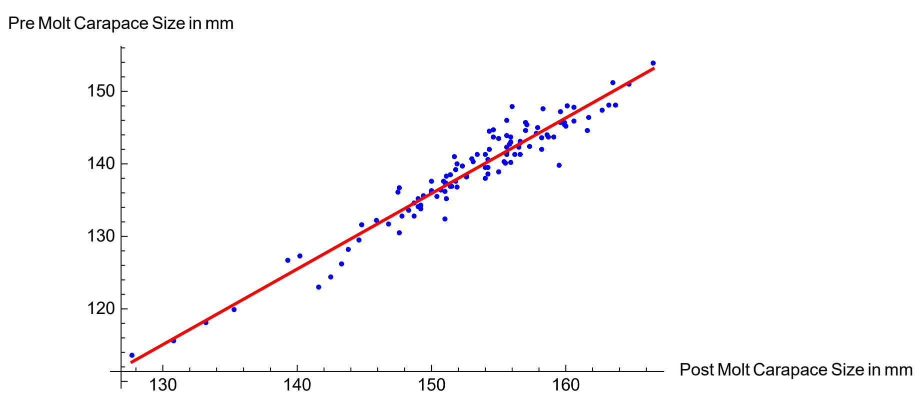


Figure 5. Linear models of post molt carapace sizes on pre molt carapace sizes of Dungeness crabs. (Top) The model generated for a mixture of laboratory and capture-recapture measurements. (Middle) The model generated for the laboratory measurements. (Bottom) The model generated for the capture-recapture measurements.

Table 2. Description of the residuals from the post molt vs. pre molt linear models including mean, standard deviation, skewness, kurtosis and a five-point summary. Linear models were generated using data collected from laboratory grown crabs (Lab), capture-recapture crabs (Ocean), and both groups combined (All).

	All	Lab	Ocean
Mean	−8.451∗10^−14	−3.720∗10^−14	−2.394∗10^−14
Standard Deviation	2.196	2.284	1.880
Skewness	0.8454	1.012	0.03556
Kurtosis	8.379	9.001	3.913
Five−Point Summary			
1st Quartile	−6.156	−5.944	−6.021
2nd Quartile	−1.307	−1.404	−1.100
Median	0.05639	0.08352	−0.2496
3rd Quartile	1.315	1.325	1.211
Maximum	14.68	14.78	5.727

Figure 5.

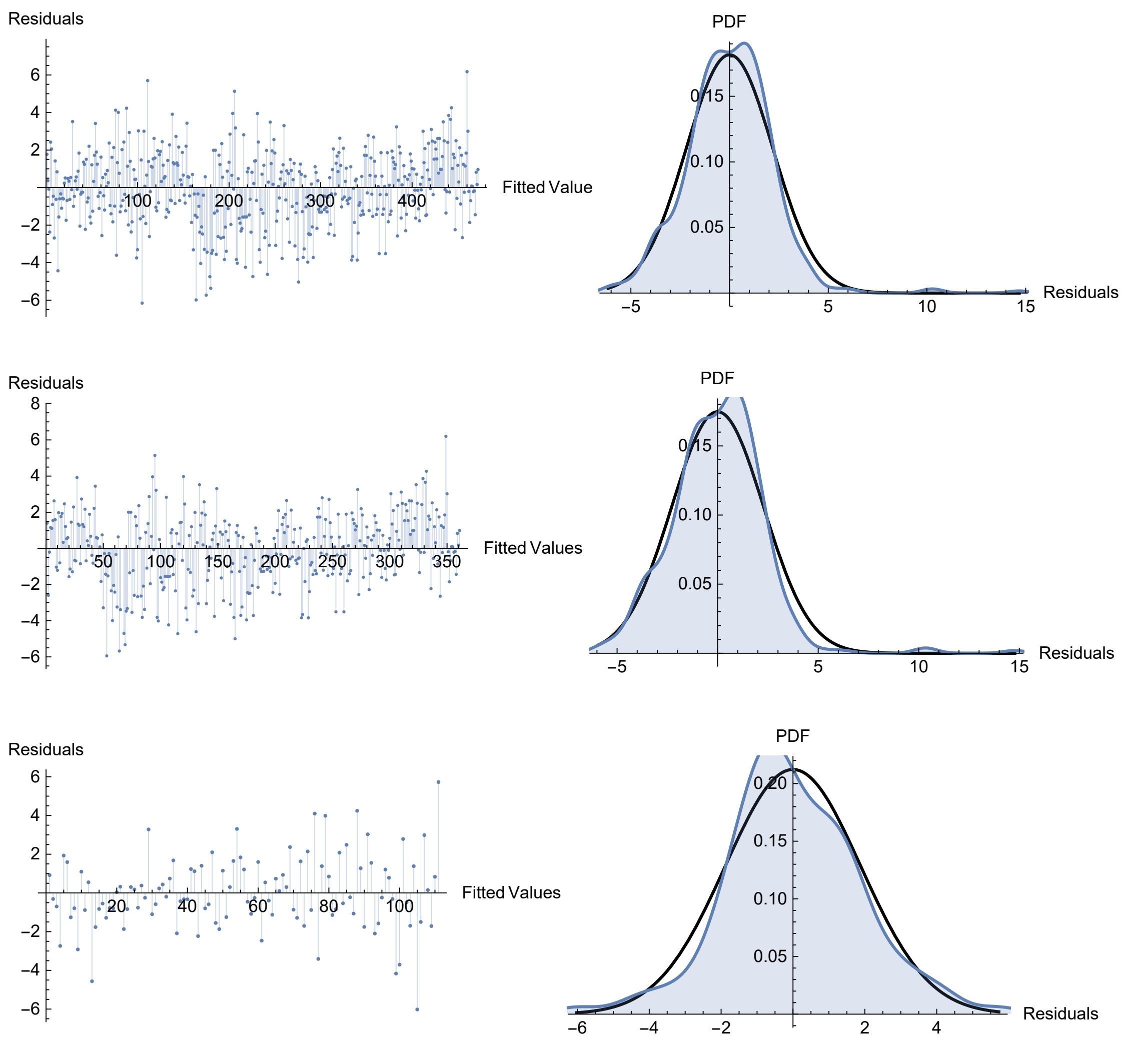


Figure 6. Plots of the residuals from the the mixed (Top), laboratory (Middle), and capture-recapture (Bottom) linear models. (Right) The plots of the residuals from the the mixed, laboratory, and capture-

recapture linear models with an overlay of the corresponding normal distribution.

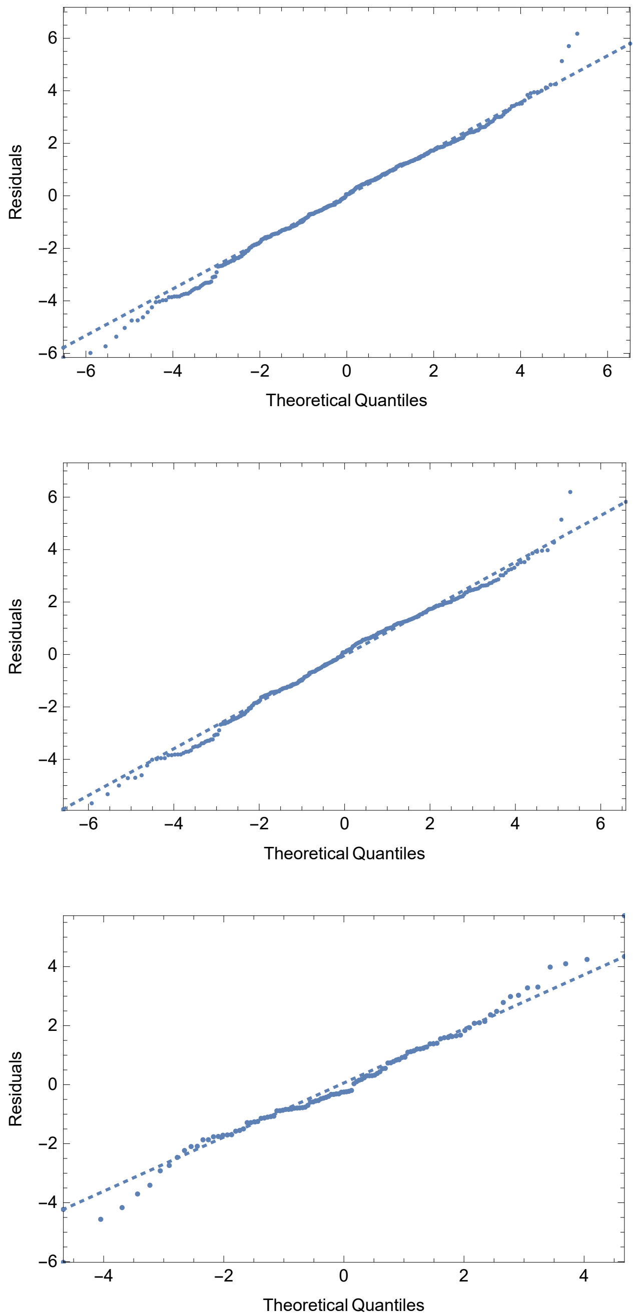


Figure 7. The quantile plots of residuals from the mixed (Top), laboratory (Middle), and capture-recapture (Bottom) linear models.

The laboratory sample's linear model predicts the pre molt carapace sizes based on the post molt sizes of crabs caught from the "Sample of crabs" data set (Figure 8).

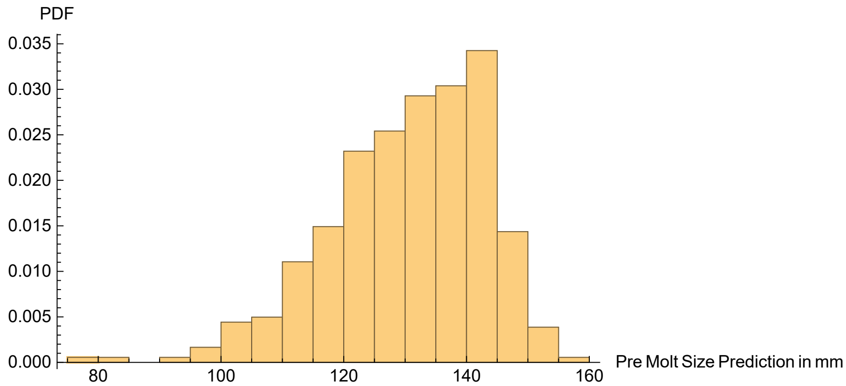


Figure 8. The predicted pre molt sizes of Dungeness crabs based on the laboratory linear model.

Discussion

The objective of this report is to provide information about the relationship between post molt and pre molt carapace sizes in female Dungeness crabs that can inform size restrictions for fishing. In the summary of carapace sizes, we found that the mean for the capture-recapture Dungeness crab sizes is higher than the laboratory data for both the pre molt and post molt sizes. The dispersion of the capture-recapture data is also much smaller based on the standard deviation and 5 point summary (Figure 2 and Figure 3). This is likely a reflection of the traps used to capture the female Dungeness crabs by the commercial fishers in this study, because these traps are designed to catch larger male Dungeness crabs. The pre and post molt measurements from the mixed, laboratory, or capture-recapture subsets do not follow the normal distribution based on the negative skewness, kurtosis values greater than 3, and quantile plots; the data is too peaky. Of the three, the capture-recapture data has the kurtosis values the closest to 3 at 4.761 and 5.241 for pre and post molt size data subsets respectively, but it is still too peaky to follow the normal distribution.

All three linear models have very high R-squared values, at 0.9808, 0.9810, and 0.9328 for the mixed, laboratory, and capture-recapture models respectively. This demonstrates a strong linear relationship between post molt and pre molt carapace sizes. The variance of the residuals from the mean for all three linear models are not constant, indicating that the residuals do not follow the assumption of homoskedasticity; they are heteroscedastic for all three linear regressions. The mean of the residuals is 0, which is expected with the definition of residuals. The skewness, kurtosis, and quantile plots for all of the linear models demonstrate that the residuals are not normally distributed for any of the linear models; however, the capture-recapture linear model has the skewness closest to 0 and a kurtosis value close to 3 (3.913), indicating that it has the residuals with the closest to normal distribution.

The laboratory linear model was used to predict the pre molt sizes for the “Sample of crabs” data because it has the highest R-squared value; however, the laboratory linear model has the kurtosis value farthest from 3. This indicates that the linear model’s residuals are too peaky to follow a normal distribution, so the model does not follow the linear regression assumptions of normally distributed residuals or homoskedasticity. Therefore some of the predictions may be misleading.

While the high R-squared values show that the data fits all three models well, the linear models generated in this report do not follow the assumptions of linear regressions. The assumptions of normally distributed residuals and homoskedasticity fail to be met for each of the generated models. This indicates that their ability to predict the pre molt size from post molt size may be unreliable. These models may also be limited by the gap in time between when the data was collected and now; we might not be able to apply these models to current Dungeness crab populations. The commercial traps that caught the crabs in the capture recapture data were designed to catch large male crabs so the females caught were typically larger than 155mm, and not necessarily a representation of the entire population of adult female Dungeness crabs. The second set of data also did not control for whether the crabs had molted in the most recent seasons or not. In order to make this report more applicable to current Dungeness crab populations, a random sample taken from the current population should be analyzed to account for any changes in nutrient availability and ocean conditions over the past 30-40 years.

In this report, we generated three linear models with strong relationships between the post molt and pre molt sizes of female Dungeness crabs. Models like these, that can predict pre molt size from post molt size can inform the size restrictions in place for fishing Dungeness crabs, and help inform more sustainable fishing practices.

References

1. Casella, George, et al., editors. "Dungeness Crab Growth" Stat Labs: Mathematical Statistics Through Application, by Deborah Nolan and Terry Speed, Springer, 2000, pp. 139-152.