Urosalpinx growth rate (length)

Andrew Villeneuve

5/26/2020

# The Study - read first

For this part of the study, we measured growth rates of juvenile Urosalpinx snails. We measured snail length using ImageJ within two days of hatching, and grew them in tea strainers separated by population for 24 days in a common garden experiment. Nine replicates per population were distributed in six temperatures, with three groups of three subreplicates in each temperature/population treatment. Snails were given food ad libitum. The end length was recorded using digital calipers after 24 days, and we substracted the starting length from the beginning length to get the growth rate (over 24 days; see data exploration for reasoning). From this growth data, we can create thermal performance curves (TPCs) for each population to show the growth response of Urosaplinx populations to a range of laboratory temperatures.

The analysis of these growth curves requires two steps:

1. the creation of models that describe the curved TPC shape using segmented and quadratic models and
2. the extraction of breakpoints (thermal optima = x breakpoint, maximal trait performance = y optima) for each population’s TPC and modeling which environmental factors best describe any patterns in these optima across populations. I do this across two methods - segmented and quadratic regression. We want to see if both methods give approximately the same results, and based on model outputs decide which method we should use.

These two steps are presented, separately, in the Data Analysis section. The organization of these analyses thusly:

* Broken stick regression - I create, analyze, and plot segmented regressions
  + Breakpoint analysis, broken stick - I extract the thermal optima (x brkpt) and maximal trait performance (y brkpt) of each population’s segmented regression
* Quadratic model - I create, analyze, and plot quadratic regressions.
  + Breakpoint analysis, quadratic - I extract the thermal optima (x brkpt) and maximal trait performance (y brkpt) of each population’s quadratic regression

# Metadata

* code
  + Unique code for each indiviudal snail, corresponding to population, temperature treatment. First digit = temperature (1=16,2=20,3=24,4=26, 5=28, 6=30), second digit = site (1=Willapa, 2=Humboldt, 3=Great Bay, 4=Woods Hole, 5=Oyster, 6=Beaufort, 7=Folly Beach, 8=Skidaway), third digit = tupperware bin number (1-3), fourth digit = snail replicate (1-3)
* pop
  + Source population of each snail. See data table below for list of site abbreivations with site.
* temp
  + Common garden temperature the snails were raised in for 24 days. Degrees C
* hatch
  + hatch date of each snail from it’s egg case (m/dd/yyyy)
* exp.date
  + Date on which hatchling snails were placed in the common garden experiment. Not more then 2 days from the hatch date. (m/dd/yyyy)
* grow.date
  + End date where growth measurements were taken. 24 days after exp.date, therefore no more then 26 days post hatch (m/dd/yyyy)
* alive
  + Tracks if snails survived the experient. m marks missing, n marks no, y marks yes
* rem.oysters
  + Was there a surplus of food at the end of the experiment? n marks no, y marks yes
* cal.length.start
  + caliper length of hatchlings upon entering the experment. We took photos of snails before entering snails into the experiment, and then used ImageJ to extract snail sizes. Size in mm
* cal.length.end
  + caliper length of hatchling at the end of the experiemnt. We took caliper measurements of the snails, as well as verifying the measurements using a subset of photographs in ImageJ. Size in mm
* wt
  + End weight of snail. Note that no initial starting weight was recorded. Weight in g.
* ran.out
  + Did the snail ever run out of food during the consumption experiment? 1 for yes, 0 for no
* bin
  + Bin number, controls subreplication. The third digit of the code.
* oce
  + Ocean (Atlantic or Pacific)

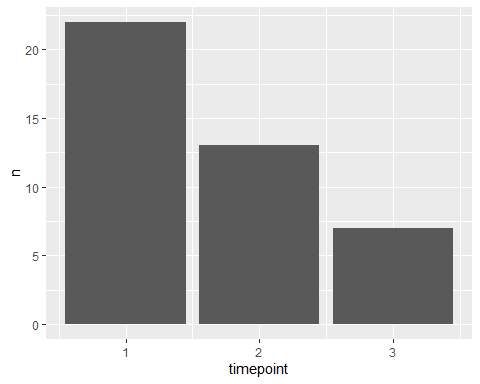
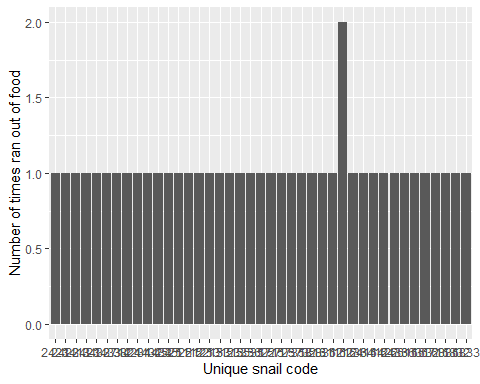
## site.abbreviation site  
## 1 gbj Great Bay, NH  
## 2 wh Woods Hole, MA  
## 3 oy Oyster, VA  
## 4 bf Beaufort, NC  
## 5 fb Folly Beach, SC  
## 6 gcsk Skidaway, GA  
## 7 nah1516 Willapa, WA  
## 8 hmi2 Humboldt, CA

# Data Setup

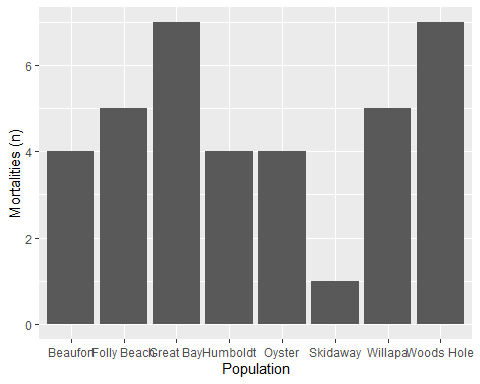
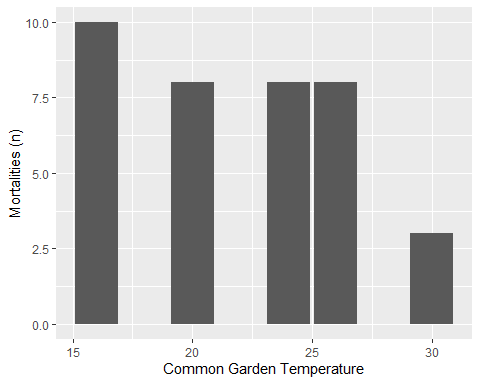
We had to clean parts of the data to prepare it for analysis. The first issue we had to resolve was the extreme weight and caliper length outliers, which was clearly due to a misplaced decimal or missing 0 during data entry (these data points were off by a factor of 10 from “correct” data). In the silenced part of the markdown, this had to be corrected 5/432 times.

The second issue was that some snails ran out of food during the growth experiment, which could jeopardize our assumption of unlimited growth while in the common garden experiment. We checked snail consumption three times over the course of the experiment, both to ensure snails had food but also to record which ones ran out of food. The vast majority (391/432) of snails never ran out of food. 40/432 ran out of food once, while 1/432 ran out of food twice. The plot below shows this breakdown, with snails that never ran out food removed. For the purposes of this experiment, we decided to include snails that missed a single meal during the entirety of the experiment, but removed the single case in which a snail ran out of food twice.

Further, the second plot below shows that most snails consumed all their food at timepoint 1, followed by timepoint 2 and finally timepoint 3 (seven snails ran out at t3)



The third issue was that some snails died, or went missing, over the course of the experiment. 37 snails died, and 1 went missing. We removed these snails from consideration, as we could not get a final growth measurement. The two plots below show the distribution of dead snails across site and temperature. Intriguingly, snails tended to died at lower temperatures, especially at 16C. We will see later on this was the temperature of lowest growth as well.

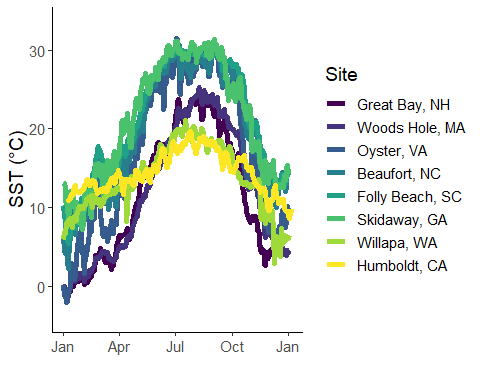


In, between removing snails that ran out of food twice (n=1), died (n=37), or went missing (n=1), there were 393 snails whos growth rates we kept.

## Environmental Data

We extracted temperature data from each site. With this data, we calculated different environmental predictors that might explain patterns in growth.

* “Quantile” (°C)  
  \*The average SST of the upper 75th percentile of summer months (06/01/2018 - 09/30/2018)
* “Decile” (°C) \*The average SST of the upper 90th percentile of summer months (06/01/2018 - 09/30/2018)
* “max” \*The maximum SST value recorded during summer months (06/01/2018 - 09/30/2018)
* “mean” (°C) \*The mean SST of the site, calculated across the entire year
* “summer mean” (°C) \*The mean SST of the site, calculated during the summer months (06/01/2018 - 09/30/2018)
* “seasonlength10/16” (days) \*The number of days where the average daily temperature exceeded a threshold. Here, we use 16C, since this was our lowest temperature in experimentation and resulted in very slow growth (accelerates at 20C) - from Cheng et al. 2017. However, things become signicant when performed at 10C. We can also use 12.5C and get somewhat significant results, after the breakpoint for Oxygen consumption found in Schick’s dissertation (Schick 1971)

We also used latitude as a predictor (not shown in table below) 

## site quantile decile max mean summer.mean seasonlength10  
## 1 gbj 22.91100 23.69700 25.96300 12.40097 20.88969 179  
## 2 wh 23.60000 24.30000 26.00000 12.10044 21.35360 197  
## 3 oy 28.92496 30.07281 33.56934 17.08424 27.54020 250  
## 4 bf 29.10000 29.80000 31.00000 19.73769 28.28414 330  
## 5 fb 29.80000 30.00000 31.30000 20.34944 29.12126 310  
## 6 nah1516 19.80700 20.43960 26.49120 13.91728 18.73544 285  
## 7 hmi2 18.77000 19.78000 21.80000 14.43312 17.36318 348  
## 8 to3 22.14730 22.71100 24.12180 17.90621 20.91789 366  
## 9 gcsk 30.26200 30.97800 33.47800 21.45739 29.36123 348  
## seasonlength16  
## 1 137  
## 2 132  
## 3 186  
## 4 224  
## 5 214  
## 6 123  
## 7 113  
## 8 224  
## 9 237

# Data exploration

## Do sites differ?

Do populations differ in their growth rate across the common garden experiemnt? Here, ANOVA tells us that growth betwween sites are significantly different. We are justified in pursuing population and temperature level differences.

anovasites<-(aov(cal.length~pop\*temp,data=growth.alive))  
summary(anovasites)

## Df Sum Sq Mean Sq F value Pr(>F)   
## pop 7 23.2 3.3 7.274 3.47e-08 \*\*\*  
## temp 1 352.0 352.0 771.549 < 2e-16 \*\*\*  
## pop:temp 7 10.8 1.5 3.385 0.00162 \*\*   
## Residuals 374 170.6 0.5   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

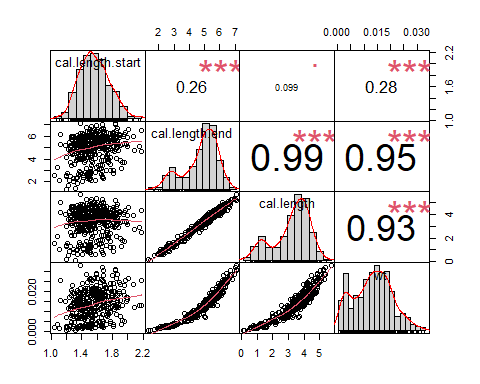
growth.alive$pop<-as.factor(growth.alive$pop)

## Shell size by population

Should we use end caliper lengths, or do we need to subtract initial caliper length from end caliper length? In other words, do initial caliper lengths differ, requiring us to standardize our growth rate? Here, we find that populations do differ in initial growth. Tukey post-hoc comparisons (silenced, in code) further support this. Thus, we must standardize growth by creating a growth rate of Final size - initial size. The univariate boxplots below also show that while outliers do appear to be present, they can be attributed to population or temperature level differences.

## Df Sum Sq Mean Sq F value Pr(>F)   
## pop 7 2.16 0.30856 8.539 9.92e-10 \*\*\*  
## Residuals 382 13.80 0.03614   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Correlations

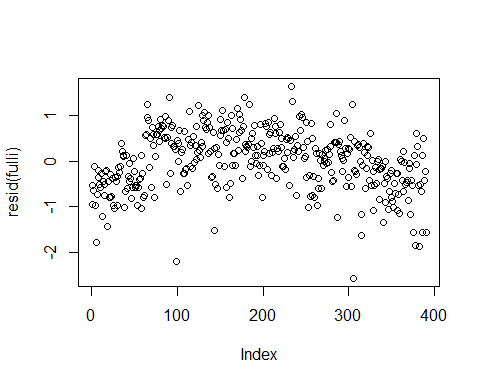


Here, we see that cal length and weight are both highly correlated. Thus, we will not include weight in any of our models.

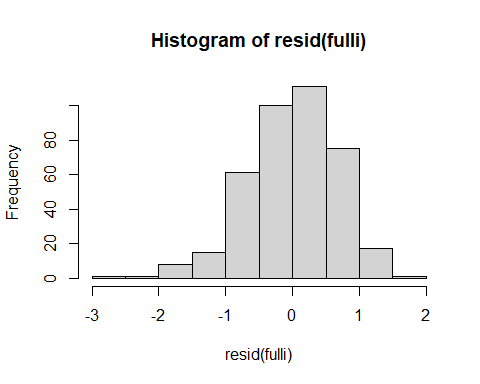
## Residuals of full model

Here, we investigate the residuals of the full model (lm(cal.length~pop*temp*oce\*bin,growth.alive)). This is to check if we need to transform our data, and if our assumptions of normality are warranted.

growth.alive<-na.omit(growth.alive)  
fulli<-(lm(cal.length~pop\*temp\*oce\*bin,growth.alive))  
  
plot(resid(fulli))



hist(resid(fulli))



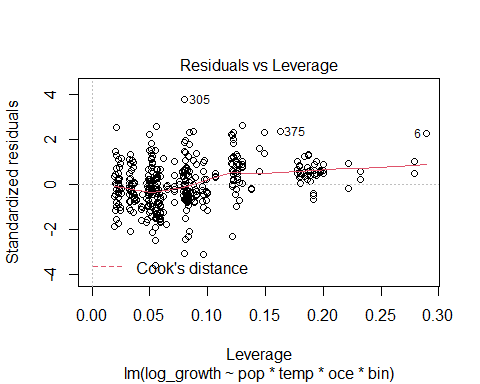
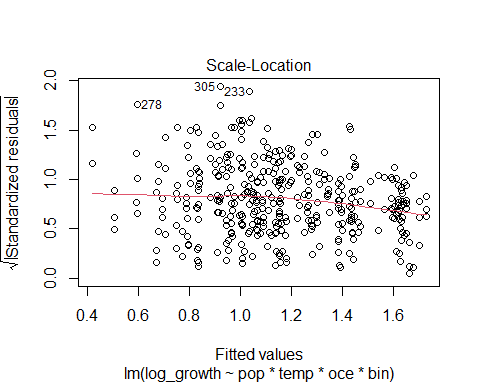
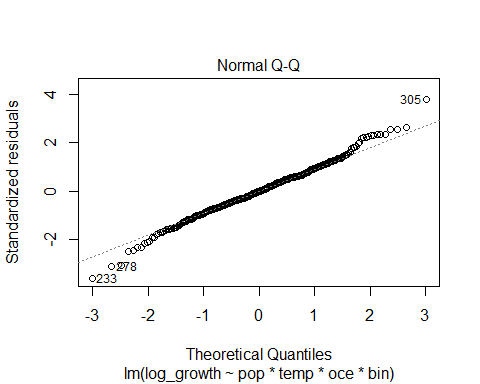
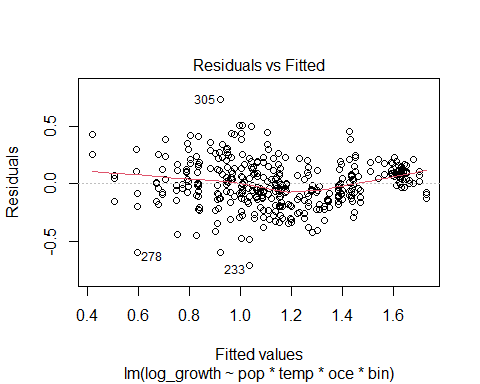
skewness(growth.alive$cal.length)

## [1] -0.7552799

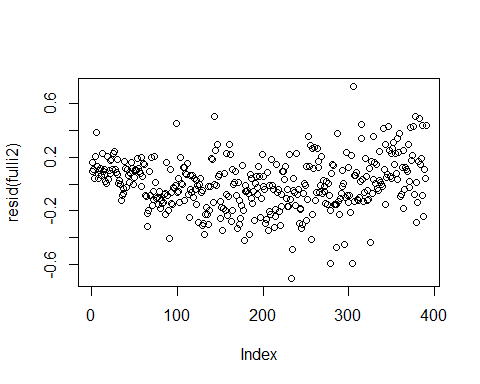
#negative skew  
growth.alive<-tidyr::drop\_na(growth.alive)  
growth.alive<-na.omit(growth.alive)

Things look ok, but maybe log-transforming growth will improve our skew and residual plots.

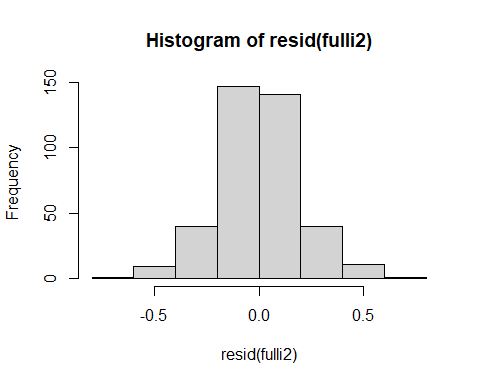
growth.alive$log\_growth<-log(max(growth.alive$cal.length+1)-growth.alive$cal.length)  
fulli2<-(lm(log\_growth~pop\*temp\*oce\*bin,growth.alive))  
  
plot(fulli2)



plot(resid(fulli2))



hist(resid(fulli2))



skewness(growth.alive$log\_growth)

## [1] 0.04950179

While the log transformation improves the skewness slightly, we will assume the distribution of residuals are normal and proceed with untransformed growth rate.

## Coplots

Other than temperature and population, which we control for, do we see different reactions depending on tupperware bin (our subreplication)? Here, reactions appaer to be the same no matter what bin we used. We do not need to include bin in our models (further supported by AIC below in data analysis). Note: Margins were too big on this figure to include in the markdown - viewable in R console.

# Data Analysis

## Model predictors

We are going to create TPCs for each population by temperature, using both piecewise (segmented) regression and quadratic regression. What predictors should be used in these models?

##   
## Model selection based on AICc:  
##   
## K AICc Delta\_AICc AICcWt Cum.Wt LL  
## pop\*temp 17 820.01 0.00 0.3 0.30 -392.18  
## pop\*temp\*oce 17 820.01 0.00 0.3 0.59 -392.18  
## pop\*temp\*out 32 820.81 0.81 0.2 0.79 -375.45  
## pop\*temp\*oce\*out 32 820.81 0.81 0.2 0.98 -375.45  
## pop.temp 10 828.90 8.89 0.0 0.99 -404.16  
## pop.temp.oce 10 828.90 8.89 0.0 0.99 -404.16  
## pop.temp.bin 11 829.65 9.64 0.0 0.99 -403.47  
## pop.temp.oce.bin 11 829.65 9.64 0.0 1.00 -403.47  
## pop.temp.out 11 831.01 11.01 0.0 1.00 -404.16  
## pop.temp.oce.out 11 831.01 11.01 0.0 1.00 -404.16  
## fulla 12 831.77 11.76 0.0 1.00 -403.47  
## temp\*out 5 834.30 14.30 0.0 1.00 -412.07  
## bin\*out\*temp 9 838.27 18.27 0.0 1.00 -409.90  
## oce\*temp\*out 9 841.00 20.99 0.0 1.00 -411.26  
## pop\*temp\*bin 33 844.44 24.44 0.0 1.00 -386.07  
## pop\*temp\*oce\*bin 33 844.44 24.44 0.0 1.00 -386.07  
## temp\*oce\*bin\*out 17 851.09 31.09 0.0 1.00 -407.72  
## temp 3 853.64 33.63 0.0 1.00 -423.79  
## oce 3 853.64 33.63 0.0 1.00 -423.79  
## temp.bin 4 854.64 34.63 0.0 1.00 -423.27  
## temp.out 4 855.56 35.56 0.0 1.00 -423.73  
## temp.oce 4 855.68 35.68 0.0 1.00 -423.79  
## bin.out.temp 5 856.53 36.52 0.0 1.00 -423.18  
## temp\*bin 5 856.53 36.52 0.0 1.00 -423.19  
## oce.temp.bin 5 856.69 36.69 0.0 1.00 -423.27  
## temp\*oce 5 857.47 37.46 0.0 1.00 -423.66  
## oce.temp.out 5 857.61 37.61 0.0 1.00 -423.73  
## temp.oce.bin.out 6 858.59 38.58 0.0 1.00 -423.18  
## fulli 56 860.55 40.54 0.0 1.00 -364.69  
## oce\*temp\*bin 9 862.78 42.77 0.0 1.00 -422.15  
## pop.out 10 1237.12 417.12 0.0 1.00 -608.27  
## out 3 1240.95 420.94 0.0 1.00 -617.44  
## bin.out 4 1242.62 422.61 0.0 1.00 -617.26  
## oce.out 4 1242.79 422.78 0.0 1.00 -617.34  
## bin\*out 5 1244.04 424.03 0.0 1.00 -616.94  
## oce.bin.out 5 1244.46 424.45 0.0 1.00 -617.15  
## oce\*out 5 1244.83 424.82 0.0 1.00 -617.33  
## pop 9 1247.38 427.37 0.0 1.00 -614.45  
## pop.oce 9 1247.38 427.37 0.0 1.00 -614.45  
## pop\*oce 9 1247.38 427.37 0.0 1.00 -614.45  
## pop.bin 10 1248.75 428.74 0.0 1.00 -614.08  
## null 2 1249.56 429.55 0.0 1.00 -622.76  
## oce\*bin\*out 9 1250.54 430.54 0.0 1.00 -616.03  
## pop\*out 17 1250.68 430.68 0.0 1.00 -607.52  
## bin 3 1250.96 430.96 0.0 1.00 -622.45  
## oce.bin 4 1252.95 432.95 0.0 1.00 -622.42  
## oce\*bin 5 1253.80 433.79 0.0 1.00 -621.82  
## pop\*bin 17 1260.64 440.64 0.0 1.00 -612.50

Here, we see that a few models are well supported. We choose the interactive pop\*temp model only because 1) the additive model only tells us if populations are different at each temperature, while the interactive model also tells us the populations slopes with temperature and gives us a TPC 2) oce adds nothing to the models, so is removed. 3) Looking at our coplots of bin, bin had no effect on growth.

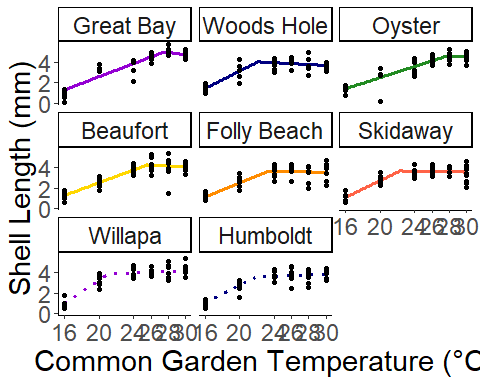
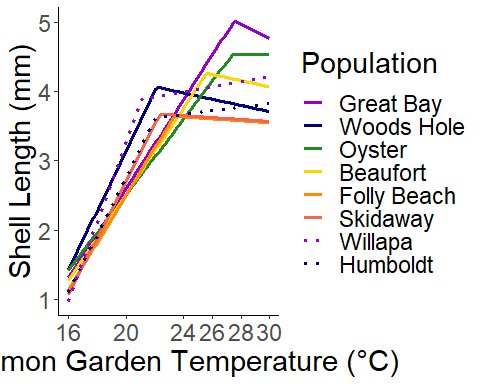
##   
## Call:  
## lm(formula = cal.length ~ pop \* temp, data = growth.alive)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.66991 -0.40576 0.02389 0.45756 1.57275   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.77328 0.49386 -3.591 0.000374 \*\*\*  
## popFolly Beach 0.53984 0.70613 0.765 0.445045   
## popGreat Bay -1.02905 0.71996 -1.429 0.153745   
## popHumboldt 0.39919 0.69843 0.572 0.567972   
## popOyster -0.45481 0.69034 -0.659 0.510417   
## popSkidaway 0.63587 0.68474 0.929 0.353679   
## popWillapa 0.48179 0.72715 0.663 0.508016   
## popWoods Hole 1.45221 0.72167 2.012 0.044905 \*   
## temp 0.21233 0.02002 10.605 < 2e-16 \*\*\*  
## popFolly Beach:temp -0.03559 0.02878 -1.237 0.216949   
## popGreat Bay:temp 0.05824 0.02928 1.989 0.047417 \*   
## popHumboldt:temp -0.02477 0.02840 -0.872 0.383738   
## popOyster:temp 0.02420 0.02802 0.864 0.388399   
## popSkidaway:temp -0.03873 0.02790 -1.388 0.165985   
## popWillapa:temp -0.01411 0.02929 -0.482 0.630423   
## popWoods Hole:temp -0.06146 0.02907 -2.114 0.035146 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.6754 on 374 degrees of freedom  
## Multiple R-squared: 0.6935, Adjusted R-squared: 0.6812   
## F-statistic: 56.41 on 15 and 374 DF, p-value: < 2.2e-16

Based on this supported model strucutre, we will construct segmented and quadratic models that follow this formulation: growth~pop\*temp.

## Broken Stick regression

It’s hard to compare TPC against one another. One method we’ve settled on is the use of broken stick regression to allow us to quantify the shape of the reaction as well as the thermal optima (x) and the maximal trait performance (y). Here, we used the segmented package to create single-optima broken stick regressions that also allow us to extract optimas.

We used a separate model for each population, as segmented does not allow for grouping. We used the following formulation: glm(shell length ~ temperature, population,family=gaussian), such that we modeled the response of shell lengths from a single population to temperature.



Here, I show the segmented model fits both on a single plot and facetted to show spread of points. Just by eyeballing the results, we see that more or less the northern sites have optima above those of southern sites. We will test for this later. First, how confident are we in these breakpoints? The tests below tell us.

* P-score: The P-score tests the null hypothesis for no difference in slopes, i.e. no breakpoint. If P is below 0.05, then there is a breakpoint. We see here that all p-scores are signficant, and thus breakpoints do exist in our data
* Davies test: we perform this analysis, but is less powerful for one breakpoint analyses. I am not certain what this means, since the null hypothesis is no breakpoint but we get radically different results this way.
* CI Low/High: Confidence interval of the breakpoint
* Breakpoint: Thermal optima breakpoint

## site P.Score Davies.Test CI.Low CI.High breakpoint  
## 1 Great Bay 1.513501e-04 0.0006151857 26.0537 29.0353 27.5445  
## 2 Woods Hole 1.355661e-10 0.4191475386 20.3966 23.9135 22.1550  
## 3 Oyster 3.418056e-02 0.8691763589 25.1749 29.7494 27.4621  
## 4 Beaufort 6.195021e-05 0.8964569173 23.7413 27.4960 25.6186  
## 5 Folly Beach 9.908718e-06 0.8001403394 20.6380 26.0520 23.3450  
## 6 Skidaway 2.705474e-10 0.1706811552 20.6209 24.1080 22.3644  
## 7 Willapa 1.695057e-10 0.5011343864 19.8379 22.3878 21.1129  
## 8 Humboldt 1.747966e-07 0.5606387987 20.2784 23.9175 22.0979

## Breakpoint analysis, broken stick

To be able to complete statistical analysis of the differences in TPC curves, I extracted the x and y componenets of each curve to give me the thermal optima and maximal trait performance, respectively. This extraction is silenced in code. Once we have extracted the thermal optima and maximal trait performance, we can move on to the relationship between environment and each breakpoint componenet.

### Maximum trait Performance (y axis), broken stick

The AIC table below tells us which environmental predictors best describe the relationship of maximal trait performance across populations.

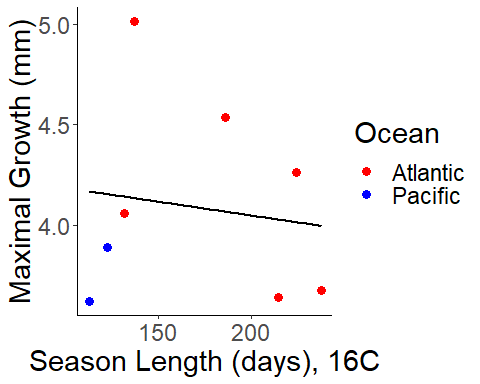
##   
## Model selection based on AICc:  
##   
## K AICc Delta\_AICc AICcWt Cum.Wt LL  
## seasonlength10 3 16.03 0.00 0.49 0.49 -2.01  
## nullo 2 16.82 0.80 0.33 0.82 -5.21  
## lat 3 21.63 5.60 0.03 0.85 -4.81  
## mean 3 22.20 6.17 0.02 0.87 -5.10  
## seasonlength16 3 22.27 6.24 0.02 0.89 -5.13  
## s.mean 3 22.41 6.38 0.02 0.91 -5.20  
## omax 3 22.41 6.38 0.02 0.93 -5.20  
## t.mean 3 22.41 6.39 0.02 0.95 -5.21  
## q.mean 3 22.42 6.40 0.02 0.97 -5.21  
## lato 4 23.25 7.23 0.01 0.98 -0.96  
## seasonlengtho10 4 24.77 8.75 0.01 0.99 -1.72  
## q.meano 4 26.79 10.76 0.00 0.99 -2.73  
## seasonlengtho16 4 26.81 10.79 0.00 1.00 -2.74  
## s.meano 4 26.84 10.82 0.00 1.00 -2.75  
## t.meano 4 27.36 11.34 0.00 1.00 -3.01  
## maxo 4 29.28 13.25 0.00 1.00 -3.97  
## meano 4 29.98 13.95 0.00 1.00 -4.32  
## o\*lat 5 41.61 25.59 0.00 1.00 -0.81  
## o\*seasonlength10 5 43.42 27.40 0.00 1.00 -1.71  
## o\*q.mean 5 44.85 28.82 0.00 1.00 -2.42  
## o\*s.mean 5 44.87 28.84 0.00 1.00 -2.43  
## seasonlength16 5 44.97 28.94 0.00 1.00 -2.49  
## o\*t.mean 5 45.58 29.56 0.00 1.00 -2.79  
## o\*max 5 46.98 30.95 0.00 1.00 -3.49  
## o\*mean 5 48.45 32.42 0.00 1.00 -4.22

Here, it appears that the season length as calculated at 10C without ocean is the best predictor. This is followed by latitude, mean sst, and season length calculated at 16C, all without ocean. Latitude doesn’t tell us much about the environment, so let’s take a look at how both season length metrics we calculated perform when calculating maximal trait performance.

#### Maximal trait performance, season length = 16 C

##   
## Call:  
## lm(formula = brkpty ~ seasonlength16, data = brkpts)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.5467 -0.3379 -0.1735 0.3034 0.8821   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.323378 0.708748 6.100 0.000884 \*\*\*  
## seasonlength16 -0.001384 0.004003 -0.346 0.741278   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5308 on 6 degrees of freedom  
## Multiple R-squared: 0.01954, Adjusted R-squared: -0.1439   
## F-statistic: 0.1196 on 1 and 6 DF, p-value: 0.7413

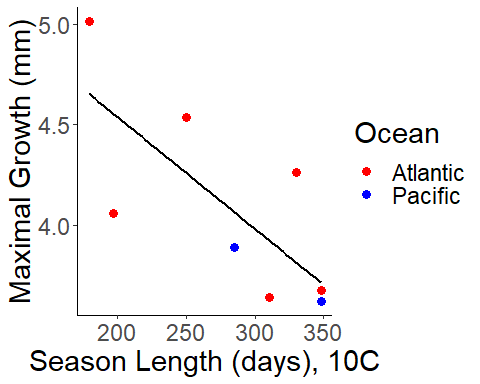
## `geom\_smooth()` using formula 'y ~ x'



#### Maximal trait performance, season length = 10 C

##   
## Call:  
## lm(formula = brkpty ~ seasonlength10, data = brkpts)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.49443 -0.20332 -0.06673 0.29972 0.44687   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.649935 0.590265 9.572 7.43e-05 \*\*\*  
## seasonlength10 -0.005564 0.002052 -2.711 0.035 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3593 on 6 degrees of freedom  
## Multiple R-squared: 0.5506, Adjusted R-squared: 0.4757   
## F-statistic: 7.351 on 1 and 6 DF, p-value: 0.03504

## `geom\_smooth()` using formula 'y ~ x'



When we look at season length when calculated at 10C, we see a significant relationship between maximal trait performance and season length. This trend disappears when we calculate it at 16C, mostly because our Pacific sites “ungroup” from the Atlantic sites. At 10C season length, we could be seeing possible countergradient variation! The one weird point in the Atlantic is Woods Hole. Anecdotal, I could attribute this to the very hot local conditions in the estuary, but can’t say for sure.

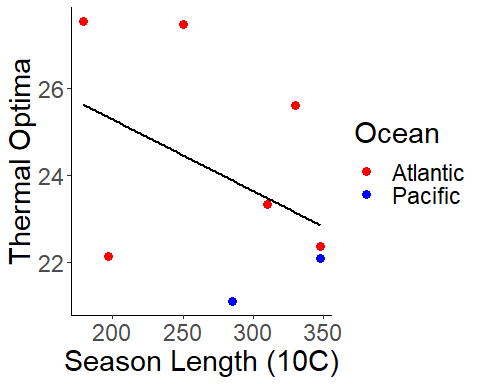
But what about the thermal optima?

### Thermal Optima (Breakpoint X axis), broken stick

##   
## Model selection based on AICc:  
##   
## K AICc Delta\_AICc AICcWt Cum.Wt LL  
## nullo 2 43.03 0.00 0.49 0.49 -18.32  
## mean 3 45.43 2.40 0.15 0.64 -16.72  
## seasonlength10 3 47.04 4.01 0.07 0.71 -17.52  
## t.mean 3 47.29 4.26 0.06 0.77 -17.65  
## q.mean 3 47.48 4.45 0.05 0.82 -17.74  
## omax 3 47.64 4.61 0.05 0.87 -17.82  
## s.mean 3 47.75 4.72 0.05 0.92 -17.88  
## seasonlength16 3 48.21 5.18 0.04 0.95 -18.11  
## lat 3 48.48 5.45 0.03 0.99 -18.24  
## lato 4 53.99 10.96 0.00 0.99 -16.33  
## seasonlengtho10 4 54.05 11.02 0.00 0.99 -16.36  
## meano 4 54.06 11.02 0.00 0.99 -16.36  
## seasonlengtho16 4 54.38 11.35 0.00 0.99 -16.52  
## q.meano 4 54.53 11.50 0.00 1.00 -16.60  
## s.meano 4 54.54 11.51 0.00 1.00 -16.60  
## t.meano 4 54.67 11.64 0.00 1.00 -16.67  
## maxo 4 54.79 11.76 0.00 1.00 -16.73  
## o\*lat 5 72.05 29.02 0.00 1.00 -16.03  
## o\*seasonlength10 5 72.30 29.27 0.00 1.00 -16.15  
## o\*mean 5 72.60 29.57 0.00 1.00 -16.30  
## seasonlength16 5 72.94 29.91 0.00 1.00 -16.47  
## o\*q.mean 5 73.10 30.07 0.00 1.00 -16.55  
## o\*s.mean 5 73.11 30.08 0.00 1.00 -16.55  
## o\*t.mean 5 73.23 30.20 0.00 1.00 -16.61  
## o\*max 5 73.36 30.33 0.00 1.00 -16.68

##   
## Call:  
## lm(formula = brkptx ~ seasonlength10, data = brkpts)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.1814 -1.2694 -0.3195 2.0500 2.9938   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 28.56345 4.10084 6.965 0.000435 \*\*\*  
## seasonlength10 -0.01638 0.01426 -1.149 0.294342   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.496 on 6 degrees of freedom  
## Multiple R-squared: 0.1803, Adjusted R-squared: 0.0437   
## F-statistic: 1.32 on 1 and 6 DF, p-value: 0.2943

## `geom\_smooth()` using formula 'y ~ x'



Here, mean tempeature is the best predictor. However, season length 10 is also well-supported, so for consistency sake I present season length 10. There is no signficant pattern between season length and thermal optima (none with mean temp either). I can attribute this insignicance mostly to the presence of the Woods Hole population. If we remove this, the trend becomes slightly more signficant. It is worth pointing out this point and that because of our observations, we could attribute this outlier to other environmental effects, although I don’t think we can actually eliminate it from analysis.

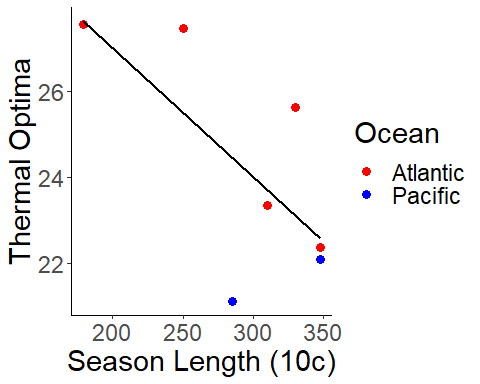
#### Sensitivity analysis

Woods Hole seems to be an outlier for this pattern of decreasing thermal optima with season length. This is possibly due to very local-level effects of where we collected the woods hole population. what would happen if we remove it?

##   
## Model selection based on AICc:  
##   
## K AICc Delta\_AICc AICcWt Cum.Wt LL  
## nullo 2 39.39 0.00 0.60 0.60 -16.19  
## seasonlength10 3 41.79 2.40 0.18 0.78 -13.89  
## mean 3 44.00 4.61 0.06 0.84 -15.00  
## t.mean 3 45.38 5.99 0.03 0.87 -15.69  
## q.mean 3 45.53 6.14 0.03 0.90 -15.77  
## omax 3 45.78 6.39 0.02 0.92 -15.89  
## s.mean 3 45.83 6.44 0.02 0.95 -15.91  
## seasonlength16 3 46.23 6.84 0.02 0.97 -16.12  
## lat 3 46.35 6.96 0.02 0.99 -16.18  
## seasonlengtho16 4 48.88 9.49 0.01 0.99 -10.44  
## seasonlengtho10 4 50.87 11.48 0.00 0.99 -11.43  
## lato 4 50.93 11.55 0.00 1.00 -11.47  
## s.meano 4 51.27 11.88 0.00 1.00 -11.64  
## q.meano 4 51.72 12.33 0.00 1.00 -11.86  
## t.meano 4 52.74 13.36 0.00 1.00 -12.37  
## maxo 4 54.31 14.93 0.00 1.00 -13.16  
## meano 4 56.11 16.72 0.00 1.00 -14.06  
## o\*lat 5 88.28 48.89 0.00 1.00 -9.14  
## o\*seasonlength10 5 90.19 50.80 0.00 1.00 -10.09  
## seasonlength16 5 90.77 51.38 0.00 1.00 -10.39  
## o\*s.mean 5 93.24 53.85 0.00 1.00 -11.62  
## o\*q.mean 5 93.66 54.28 0.00 1.00 -11.83  
## o\*t.mean 5 94.64 55.25 0.00 1.00 -12.32  
## o\*max 5 96.24 56.85 0.00 1.00 -13.12  
## o\*mean 5 97.96 58.57 0.00 1.00 -13.98

##   
## Call:  
## lm(formula = brkptx ~ seasonlength10, data = brkpts\_wh)  
##   
## Residuals:  
## 1 2 3 4 5 6 7   
## -3.34256 -0.47605 -0.07659 1.96145 2.50710 -0.36382 -0.20953   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 32.96679 4.13387 7.975 0.0005 \*\*\*  
## seasonlength10 -0.02986 0.01386 -2.155 0.0837 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.084 on 5 degrees of freedom  
## Multiple R-squared: 0.4816, Adjusted R-squared: 0.3779   
## F-statistic: 4.645 on 1 and 5 DF, p-value: 0.0837

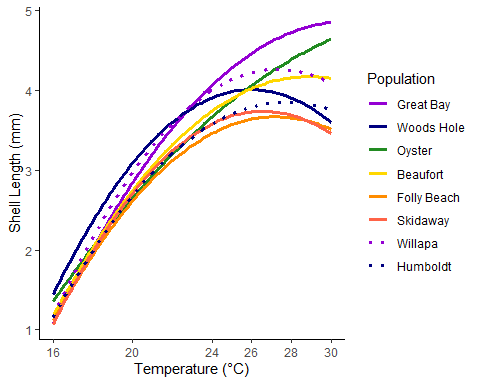
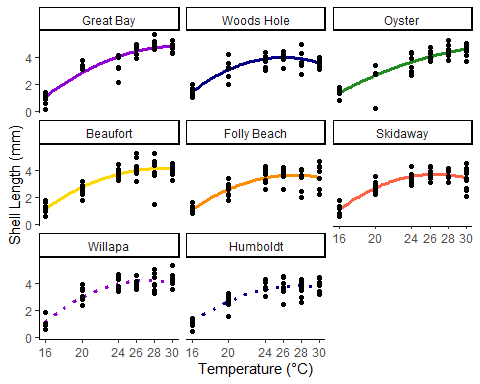
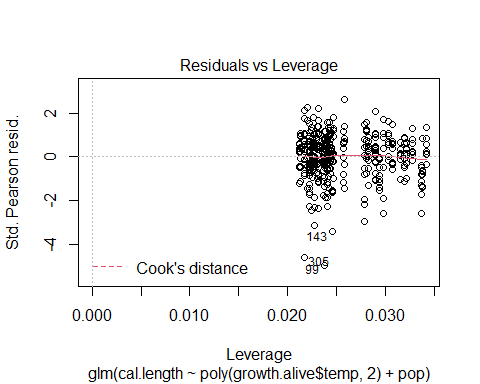
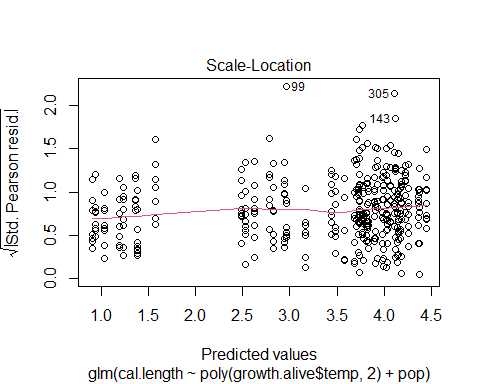
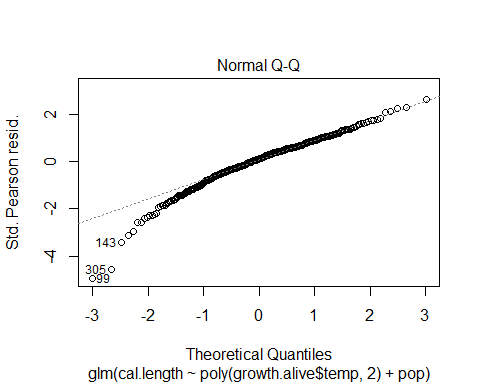
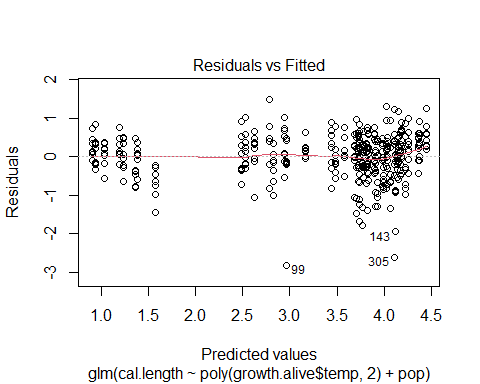
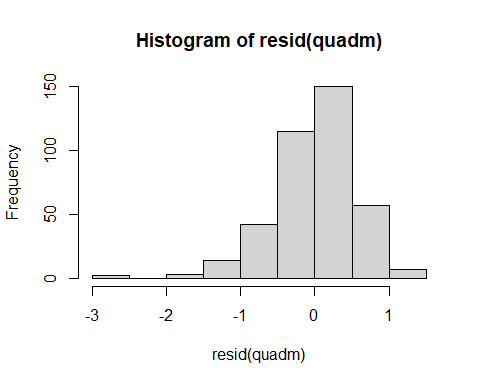
## `geom\_smooth()` using formula 'y ~ x'



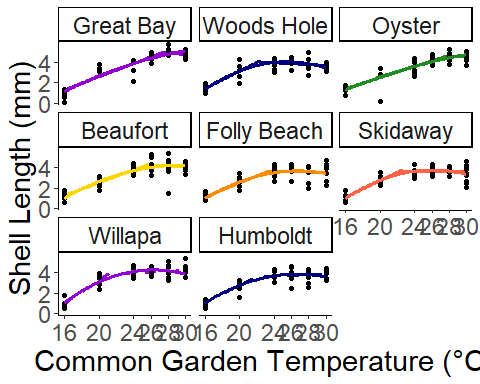
With Woods Hole dropped, we see a better indication of a negative relationship between season length and optima. We probably can’t justify droppping Woods Hole, but this is good to know.

## Quadratic

One reason we are checking quadratic is that sometimes the optima isn’t the optima! If you closely examine the segmented regressions, you see that in some cases the second segment’s slope does not appear to be negative. Instead, it plateaus. This raises the question whether we can really define our breakpoint as optima. Potentially, we could interpret the breakpoints as the lowest temperature of maximum growth. At any rate, one idea is to redo all the analysis we just did for segmented regression but with a quadratic model, and seeing if we get a similar result in both the stacking but also the trends of the maximal growth and optima as we did with the segmented regression.

 Here, we see more or less similar patterns of cold pops stacked on warm pops. How do the breakpoints stack up?

Before we extract quadratic breakpoints, let’s line up segmented regression and quadratic regression and see if they look close. The curves do match up nicely, but the breakpoints will be different.

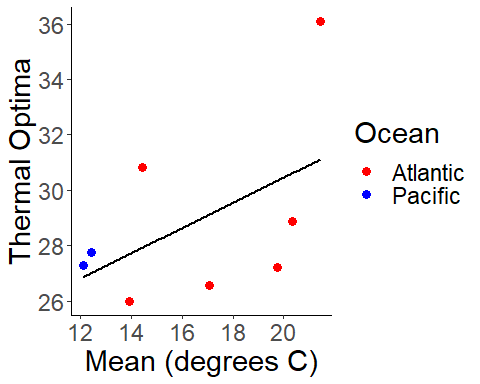


### Thermal optima (X Breakpoint), Quadratic

##   
## Model selection based on AICc:  
##   
## K AICc Delta\_AICc AICcWt Cum.Wt LL  
## nullo 2 47.14 0.00 0.57 0.57 -20.37  
## mean 3 50.30 3.16 0.12 0.69 -19.15  
## omax 3 51.73 4.60 0.06 0.75 -19.87  
## seasonlength10 3 51.96 4.82 0.05 0.80 -19.98  
## t.mean 3 52.18 5.04 0.05 0.84 -20.09  
## q.mean 3 52.35 5.21 0.04 0.89 -20.17  
## s.mean 3 52.44 5.30 0.04 0.93 -20.22  
## seasonlength16 3 52.71 5.57 0.04 0.96 -20.35  
## lat 3 52.74 5.60 0.03 1.00 -20.37  
## meano 4 59.38 12.24 0.00 1.00 -19.02  
## maxo 4 61.06 13.92 0.00 1.00 -19.86  
## seasonlengtho10 4 61.09 13.95 0.00 1.00 -19.88  
## lato 4 61.26 14.12 0.00 1.00 -19.96  
## seasonlengtho16 4 61.44 14.30 0.00 1.00 -20.05  
## t.meano 4 61.48 14.34 0.00 1.00 -20.07  
## q.meano 4 61.56 14.43 0.00 1.00 -20.12  
## s.meano 4 61.57 14.44 0.00 1.00 -20.12  
## o\*mean 5 78.04 30.90 0.00 1.00 -19.02  
## o\*max 5 79.54 32.40 0.00 1.00 -19.77  
## o\*seasonlength10 5 79.65 32.52 0.00 1.00 -19.83  
## o\*lat 5 79.81 32.67 0.00 1.00 -19.91  
## seasonlength16 5 80.10 32.96 0.00 1.00 -20.05  
## o\*t.mean 5 80.13 32.99 0.00 1.00 -20.07  
## o\*q.mean 5 80.22 33.08 0.00 1.00 -20.11  
## o\*s.mean 5 80.23 33.09 0.00 1.00 -20.11

##   
## Call:  
## glm(formula = brkptyq ~ mean, family = "gaussian", data = brkpts)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -3.1107 -1.9277 -0.6359 1.2881 4.9942   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 21.3715 5.2134 4.099 0.00636 \*\*  
## mean 0.4539 0.3103 1.463 0.19386   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 9.367229)  
##   
## Null deviance: 76.246 on 7 degrees of freedom  
## Residual deviance: 56.203 on 6 degrees of freedom  
## AIC: 44.299  
##   
## Number of Fisher Scoring iterations: 2

## `geom\_smooth()` using formula 'y ~ x'

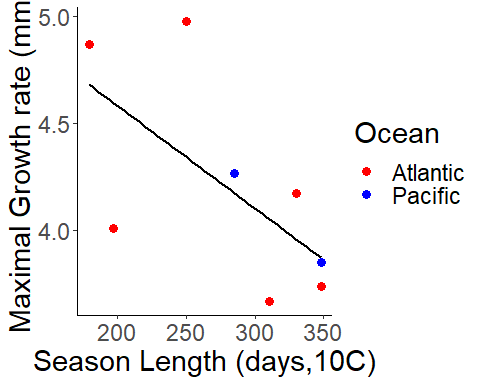
 This is not signficant in the slightest. part of the issue is that the quadratic optima for Oyster site is near 34C. Very unlikely! This is likely due to the very small breakpoint change in Oyster that is interpreted by the quadratic as a single slope. Basically, nothing predicts it well

### Maximal trait performance (Y Breakpoint), Quadratic

##   
## Model selection based on AICc:  
##   
## K AICc Delta\_AICc AICcWt Cum.Wt LL  
## nullo 2 16.72 0.00 0.49 0.49 -5.16  
## seasonlength10 3 18.07 1.35 0.25 0.74 -3.04  
## lat 3 20.83 4.11 0.06 0.80 -4.41  
## seasonlength16 3 21.85 5.13 0.04 0.84 -4.93  
## mean 3 22.12 5.40 0.03 0.87 -5.06  
## s.mean 3 22.19 5.47 0.03 0.90 -5.10  
## q.mean 3 22.27 5.55 0.03 0.93 -5.13  
## omax 3 22.28 5.56 0.03 0.96 -5.14  
## t.mean 3 22.30 5.58 0.03 0.99 -5.15  
## lato 4 27.02 10.30 0.00 0.99 -2.84  
## seasonlengtho10 4 27.38 10.66 0.00 1.00 -3.02  
## seasonlengtho16 4 29.47 12.76 0.00 1.00 -4.07  
## s.meano 4 30.15 13.43 0.00 1.00 -4.41  
## q.meano 4 30.24 13.52 0.00 1.00 -4.45  
## t.meano 4 30.59 13.87 0.00 1.00 -4.63  
## meano 4 31.40 14.68 0.00 1.00 -5.03  
## maxo 4 31.40 14.68 0.00 1.00 -5.03  
## o\*lat 5 45.68 28.96 0.00 1.00 -2.84  
## o\*seasonlength10 5 46.00 29.28 0.00 1.00 -3.00  
## seasonlength16 5 47.42 30.70 0.00 1.00 -3.71  
## o\*s.mean 5 48.10 31.38 0.00 1.00 -4.05  
## o\*q.mean 5 48.23 31.51 0.00 1.00 -4.11  
## o\*t.mean 5 48.70 31.98 0.00 1.00 -4.35  
## o\*max 5 49.47 32.75 0.00 1.00 -4.73  
## o\*mean 5 49.64 32.92 0.00 1.00 -4.82

##   
## Call:  
## glm(formula = brkptxq ~ seasonlength10, family = "gaussian",   
## data = brkpts)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.58531 -0.19727 0.03542 0.19236 0.63475   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.536229 0.670849 8.253 0.000171 \*\*\*  
## seasonlength10 -0.004781 0.002332 -2.050 0.086290 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 0.1667831)  
##   
## Null deviance: 1.7013 on 7 degrees of freedom  
## Residual deviance: 1.0007 on 6 degrees of freedom  
## AIC: 12.073  
##   
## Number of Fisher Scoring iterations: 2

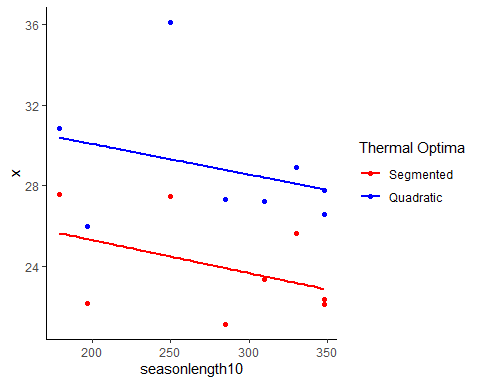
## `geom\_smooth()` using formula 'y ~ x'



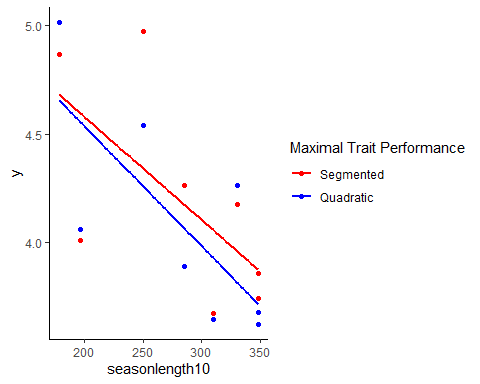
Here, we see that the maxima growth rate pattern is more or less preserved. While we don’t trust the quadratic estimates for the optima, I do trust that there is a pattern of increasing growth rate with higher latitude. This is evidence for countergradient variation in growth. We can confirm this by looking further at growth in weight as well as in consumption rate.

Finally, let us compare the breakpoints we have extracted from segmented and quadratic regression. We want to know if we see signficantly different results in methods, because if we do then we have to be careful about which once we select (seg v. quad). We see here that while thermal optima estimates differ signficantly between quad and seg, they do not when estimating maximal trait performance. Importantly, the sign of both relationships is the same. Therefore, we can be confident that our conclusion of countergradient variation in maximal trait peroformance is correct.

## `geom\_smooth()` using formula 'y ~ x'



## `geom\_smooth()` using formula 'y ~ x'



##   
## Call:  
## lm(formula = x ~ modeltypex, data = brkptnewx)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.850 -1.822 -1.311 1.743 7.274   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 23.963 1.043 22.97 1.63e-12 \*\*\*  
## modeltypexbrkptyq 4.869 1.475 3.30 0.00526 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.95 on 14 degrees of freedom  
## Multiple R-squared: 0.4376, Adjusted R-squared: 0.3974   
## F-statistic: 10.89 on 1 and 14 DF, p-value: 0.005258

##   
## Call:  
## lm(formula = y ~ modeltypey, data = brkptnewy)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.5243 -0.4216 -0.1060 0.2428 0.9288   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.1935 0.1749 23.98 9.07e-13 \*\*\*  
## modeltypeybrkpty -0.1065 0.2473 -0.43 0.673   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4946 on 14 degrees of freedom  
## Multiple R-squared: 0.01306, Adjusted R-squared: -0.05743   
## F-statistic: 0.1853 on 1 and 14 DF, p-value: 0.6734