Thermal and warming tolerance analysis

Andrew Villeneuve

08/07/2020

### The Study

The purpose of the study was to evaluate the thermal tolerances (or ) of *Urosalpinx cinerea* from a range of populations across latitude in both the Atlantic and Pacific Oceans. Further, these populations of snails were incubated, hatched, acclimated at two different temperatures (20° and 24° C), to test for plasticity in thermal tolerances across population. We expect populations from northern “cold” sites to show lower then populations from southern “warm” sites.

Egg cases were reared in aerated tanks at either 20C or 24C until hatching. After hatching, snails were held and fed oyster spat for 8-16 days at their acclimation temperature before they entered the heat bar experiment.

The heatbar was controlled on one end by a PID controller to control heat ramping, while a cooling water bath was hooked up to the other end. For the published manuscript, we include data from one protocol of heatbar ramping. However, we did use one other ramping protocol in an attempt to create finer divisions between heatbar positions (see the section below addressing alternate protocols).

The heatbar was designed to test three populations at once, with 30 snails from each population. Prior to the start of the heat bar experiment, each snail was weighed as a proxy for age as well as size to include in models. We did this to ensure that size and age did not affect our results. An individual, weighed snail was placed in a 5ml tube of seawater from the proper acclimation (20 or 24C) temperature. Nitex mesh of 200microns was secured with a plastic collar so that the mesh forms a plug below the waterline. This allows for free exchange of air. Once all snails were placed in tubes and their ID recorded so we could match survivorship with weight, population, and specific snail, we began the heatbar trial. See the separate r file heatbar\_ramping.r for the specifics of how the heatbar was ramped over 5 hours. After hours, tubes were immersed in either 20C or 24C water as appropriate overnight to allow for recovery. The next day, we sorted live and dead snails by looking for foot attachment on the sides of each tube or by poking the foot with a probe if unclear.

There were two main steps in analyzing thermal tolerance data. 1) We used binomial models to extract the LT50 of each run (see metadata for explanation). Essentially, for every 30 snails from a single population and acclimation temp, we got a single LT50 value. 2) We constructed models to test the relationship between population and acclimation LT50s across different environmental parameters. Through model selection, we were able to select which environmental predictor best explained patterns of LT50 with environment.

## Protocol

We created one other heat bar ramping protocol during the experimental phase of research. This was formulated to decrease the temperature difference between heatbar positions and potentially get a more precise LT50 value. Unfortunately, we did not have enough juvenile snails to achieve the sample sizes necessary to enable a proper analysis. The heatbar protocol for this other ramp is shown in the heatbar\_ramping.r file, but the analysis of the LT50 data is not shown in this Rmarkdown. Future research into protocol differences should be undertaken, as initial results were promising.

## Metadata

# What is a run?

A “run” is a sequence of 15-30 snails for a given population and acclimation. Each “run” has no column replicates, so if we did not randomly distribute them by row they would be in a single row. A heatbar run is when up to three different “runs” are run simultaneously.

# Date

The date is the date of the heatbar “run” (as defined below) for a given population and acclimation.

# Site

Two-letter abbreviation of sites used. Correspond to the following sites:

site.meta<-data.table(siteID = c("GB","WH","OY","BF","FB","HM","TO"),site=c("Great Bay","Woods Hole","Oyster","Beaufort","Folly Beach","Humboldt","TO"),state=c("NH","MA","VA","NC","SC","CA","CA"))  
site.meta

## siteID site state  
## 1: GB Great Bay NH  
## 2: WH Woods Hole MA  
## 3: OY Oyster VA  
## 4: BF Beaufort NC  
## 5: FB Folly Beach SC  
## 6: HM Humboldt CA  
## 7: TO TO CA

# Acc

Acc is the acclimation temperature. Each population was reared in either 20°C or 24°C seawater (30PSU). We are interested in knowing if acclimation temperature will expose signs of plasticity in thermal tolerance for each population.

# col

Col is the column in which the individual from that particular run/population/acc was placed. For both protocols 1 and 2, all snails were placed in positions 9-38. No snails were in 1-8 due to the low temperature in these positions closeest to the cooling element.

# rand

a random number generated to assign column positions to snails

# row

The row (A,B,C) randomly assigned for each run’s snail

# weight

The dry weight of each snail before being used in the heatbar experiment.Snails were dabbed dry to remove excess water on and inside the shell.

# T0-T5

These are the timepoints at which temperatures were taken. Each has a step in between of 1 hr. The table below shows the ramping rate on the PID controller (setpoints) in relation to time and the timepoints at which temperature was measured.

# Survival

Survival of the snails was established between 15-18 hrs after the experiment end. Tubes of the snails were floated in their acclimation temperature seawater immediately after the experiment for the 15-18 hours. Afterwards, each snail was examined under a microscope and prodded with a probe. If the foot retracted, they were scored as “alive” (or 1), and if not, they were scored as “dead” (or 0). In at least one run, some snails were discovered to have been drilled (cannibalized) before the experiment. These were classed as “d” for drilled, and were not included in analysis. Future runs had each snail examined for signs of drilling before experiment ran. In the run with drilled snails, I counted snails that had tissue in the shell but clearly dead as “0”s.

# Protocol

Three total protocols were tested in the experiment, with only one included in data analysis. Most runs were completed under protocol 1, wherein temperature was ramped from 25C to 60C and ice added every hour for the first two hours of the five hour experiment. Protocol 2 had the same ramping rate but had no ice added. Protocol 3 had no ice added and had a lower end setpoint of 50C. P1 had a very wide range in temperature, and so we developed P2 to lessen the spread between temperature positions. P3 was developed for the same reason, but the projected CTmax point was too close to the heating element (not enough “0”s for the model).

ramp

## time\_hrs timepoint setpoint\_P1P2 P1\_ice P2P3\_ice setpoint\_P3  
## 1: 0 T0 20/24 yes no 20/24  
## 2: 0.5 25 no no 25  
## 3: 1 T1 30 yes no 28  
## 4: 1.5 35 no no 31  
## 5: 2 T2 40 yes no 34  
## 6: 2.5 45 no no 38  
## 7: 3 T3 50 no no 42  
## 8: 3.5 55 no no 46  
## 9: 4 T4 60 no no 50  
## 10: 4.5 60 no no 50  
## 11: 5 T5 60 no no 50

In the first heatbar run, I included cannibalized snails in the heatbar. I was able to differentiate at the end of the heatbar run which snails had died because of cannibalisim by looking for drill holes on the shells. These snails were denoted with a “d”. Snails which had no drill holes and still had the body of the snail within the shell were counted as mortalities in the heatbar.

I also need to get rid of protocol 2 runs, as the ramping rate was too dissimilar to protocol 1.

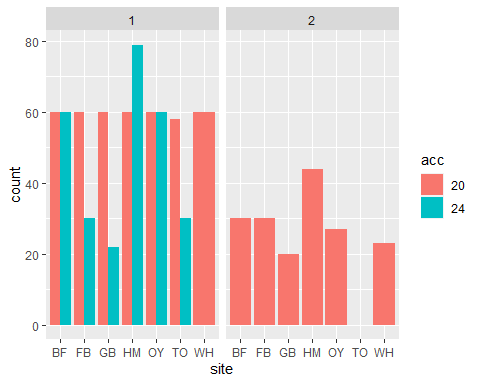
Finally, not every column in the heatbar was used, and so these are blank rows we can get rid of.

## , , = 20  
##   
##   
## BF FB GB HM OY TO WH  
## 0 38 37 31 47 37 20 29  
## 1 52 53 49 57 50 38 54  
##   
## , , = 24  
##   
##   
## BF FB GB HM OY TO WH  
## 0 25 10 13 28 26 11 0  
## 1 35 20 9 51 34 19 0

## BF FB GB HM OY TO WH   
## 150 120 102 183 147 88 83

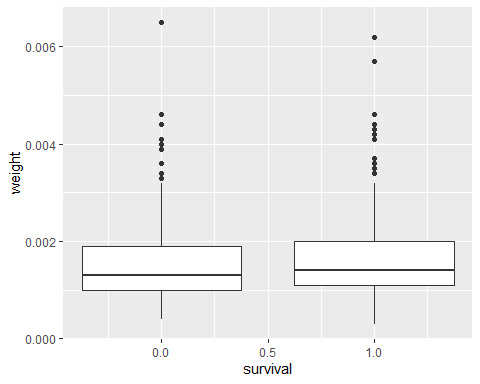
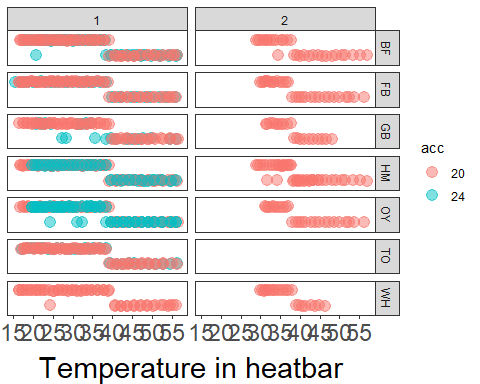
Data is now fully cleaned! Time to do some exploration.How well did we sample all the sites across acclimation temperatures?

ggplot(data,aes(x=site, fill=acc))+geom\_bar(position="dodge")+facet\_wrap(protocol~.)

 Because coverage was not complete across sites for protocol 2, and ramping rate was not a primary research objective, we did not include protocol 2 data in our analysis. We present the data above for transparency.

## Visualization of all data

Let’s visualize survival data in the heatbar across all sites and protocols. Further, did age of the snails drive survivorship?



## Df Sum Sq Mean Sq F value Pr(>F)  
## weight 1 0.62 0.6222 2.587 0.108  
## Residuals 871 209.45 0.2405

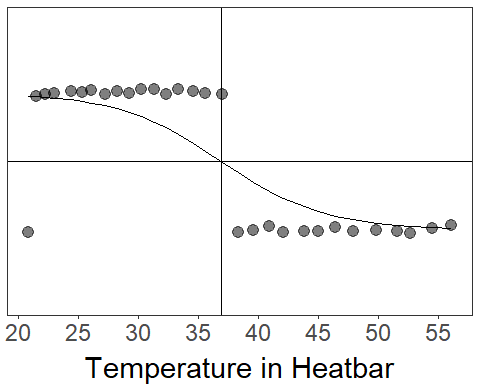
Survival does not determined by weight, which is our proxy for age.

### Individual Analysis

Now that I have a workflow, I will start off by producing binary models and extracting from each run. This is so I can make a for-loop later to perform this over all runs.

##   
## Call:  
## brglm(formula = survival ~ t5, family = "binomial", data = run1)  
##   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 8.63084 2.91917 2.957 0.00311 \*\*  
## t5 -0.23392 0.07946 -2.944 0.00324 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 35.397 on 29 degrees of freedom  
## Residual deviance: 18.879 on 28 degrees of freedom  
## Penalized deviance: 12.58527   
## AIC: 22.879

## [1] 36.89719



Here we can see how I will analyze each “run.” Note that there is an outlier death.

Below is a loop I used to extact LT50s by run and store in a dataframe.

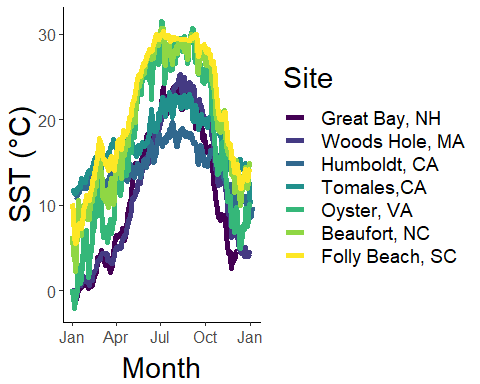
##forloop  
  
runs<-unique(data$run\_ID)  
df<-list()  
for (i in 1:length(runs)){  
 tmp<-filter(data,run\_ID==runs[i])  
 df[[i]]<-tmp  
   
}  
  
modlist<-list()  
for (i in 1:length(df)){  
 mod<-brglm(data=df[[i]],survival~t5,family="binomial")  
 modlist[[i]]<-mod  
}  
names(modlist)<-runs  
  
LT50<-list()  
for(i in 1:length(modlist)){  
 lt<-dose.p(modlist[[i]],p=0.5)[1]  
 LT50[[i]]<-lt  
}  
  
dflt<-data.frame(LT50,ncol=length(LT50))  
dflt<-melt(as.data.table(dflt))  
dflt<-dflt[-c(31),]#32 if we keep run\_ID ==3  
dfdlt2<-data.frame(dflt,runs)  
lt50s<-subset(dfdlt2,select=-c(variable))  
names(lt50s)<-c("LT50","run\_ID")  
#View(lt50s)  
  
run.count <- read.csv(here::here("data/run.count.csv"))  
lt50s<-merge(lt50s,run.count,by="run\_ID")

I’ve now extracted all of my values. I just need to merge this with other metadata for each run so that I can analyze by environmental predictors

### Environmental Data Exploration

Latitude is not a great proxy for temperature. I expect the slope of the LT50~site plot to be different then when latitude is on the x axis. There are a lot of different metrics I can use from the site temperature data I have. I chose four to use: mean sst, summer sst (Jun 01 - September 31), upper 25% quartile of summer sst, and the 90th percentile of summer sst.

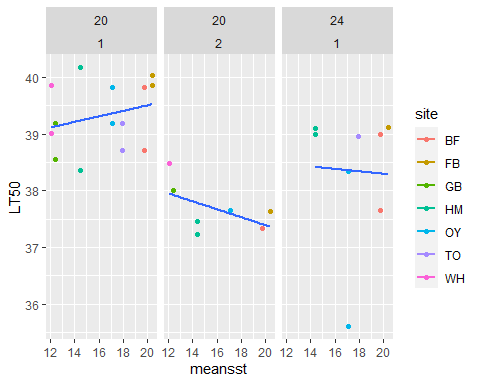
## `summarise()` regrouping output by 'day', 'month', 'year' (override with `.groups` argument)



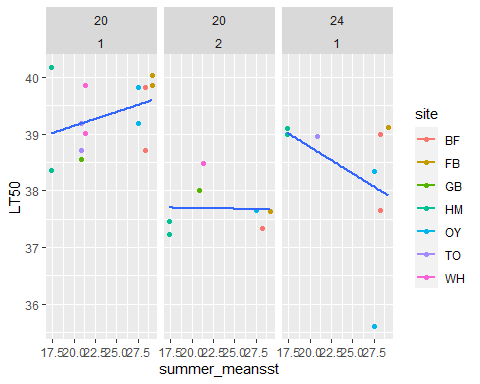
## Comparison of LT50 by four target environmental metrics.

## run\_ID LT50 date site   
## Min. : 1 Min. :35.60 Length:29 Length:29   
## 1st Qu.:15 1st Qu.:38.01 Class :character Class :character   
## Median :22 Median :38.96 Mode :character Mode :character   
## Mean :22 Mean :38.66   
## 3rd Qu.:29 3rd Qu.:39.19   
## Max. :40 Max. :40.18   
## acc run\_count P mean.sst s.mean.sst   
## Min. :20.0 Min. :14.00 Min. :1.000 Min. :12.10 Min. :17.36   
## 1st Qu.:20.0 1st Qu.:30.00 1st Qu.:1.000 1st Qu.:14.43 1st Qu.:20.89   
## Median :20.0 Median :30.00 Median :1.000 Median :17.08 Median :21.35   
## Mean :21.1 Mean :28.52 Mean :1.241 Mean :16.53 Mean :23.77   
## 3rd Qu.:24.0 3rd Qu.:30.00 3rd Qu.:1.000 3rd Qu.:19.74 3rd Qu.:28.28   
## Max. :24.0 Max. :30.00 Max. :2.000 Max. :20.35 Max. :29.12   
## q.mean.sst t.mean.sst oce max   
## Min. :18.77 Min. :19.78 Length:29 Min. :21.80   
## 1st Qu.:22.15 1st Qu.:22.71 Class :character 1st Qu.:24.12   
## Median :23.60 Median :24.30 Mode :character Median :26.00   
## Mean :25.10 Mean :25.87 Mean :27.83   
## 3rd Qu.:29.10 3rd Qu.:30.00 3rd Qu.:31.30   
## Max. :29.80 Max. :30.07 Max. :33.57   
## lat   
## Min. :32.66   
## 1st Qu.:34.82   
## Median :38.13   
## Mean :38.09   
## 3rd Qu.:40.85   
## Max. :43.09

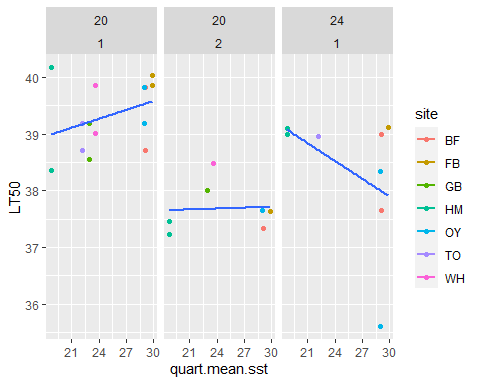
## `geom\_smooth()` using formula 'y ~ x'



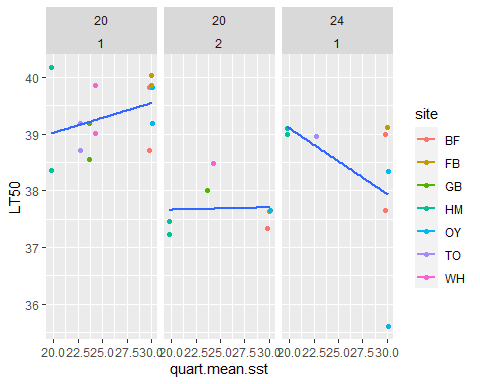
## `geom\_smooth()` using formula 'y ~ x'



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



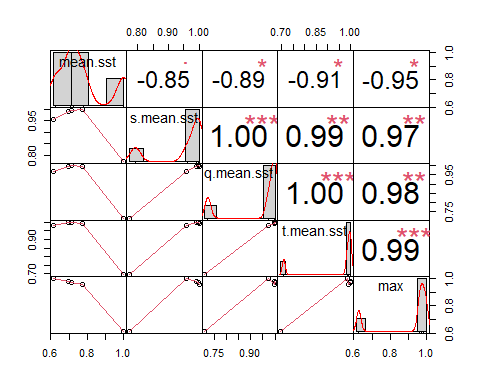
## `geom\_smooth()` using formula 'y ~ x'



We can see that the variance about the lines of best fit and the slope of the lines do differ amongst the different environmental metrics. Therefore, we will have to include them in our model selection framework to determin which best drives patterns in LT50.

## Competing environmental data/ models

# Comparing lt50s among acclimations



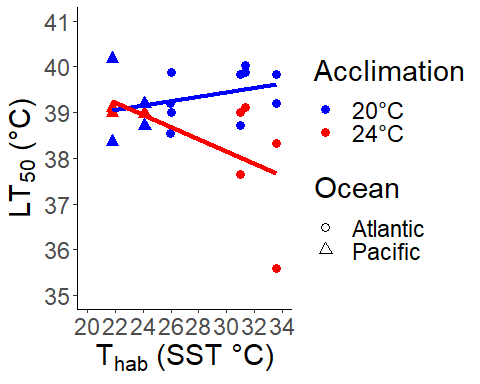
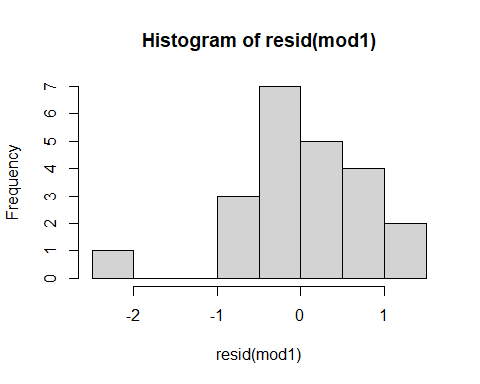
##   
## Model selection based on AICc:  
##   
## K AICc Delta\_AICc AICcWt Cum.Wt LL  
## acc 3 61.57 0.00 0.18 0.18 -27.12  
## max\*acc 5 61.79 0.22 0.16 0.34 -24.02  
## t.mean\*acc 5 63.51 1.94 0.07 0.41 -24.88  
## maxacc 4 63.96 2.39 0.05 0.47 -26.80  
## q.mean\*acc 5 64.07 2.50 0.05 0.52 -25.16  
## lat.acc 4 64.09 2.52 0.05 0.57 -26.87  
## oacc 4 64.16 2.59 0.05 0.62 -26.90  
## mean.acc 4 64.30 2.73 0.05 0.67 -26.97  
## t.mean.acc 4 64.33 2.76 0.05 0.71 -26.99  
## q.mean.acc 4 64.44 2.87 0.04 0.76 -27.04  
## s.mean\*acc 5 64.47 2.90 0.04 0.80 -25.36  
## null 2 65.00 3.43 0.03 0.83 -30.19  
## olat.acc 5 66.41 4.84 0.02 0.85 -26.33  
## lat\*acc 5 66.85 5.28 0.01 0.86 -26.55  
## max 3 66.88 5.31 0.01 0.87 -29.77  
## omean.acc 5 67.05 5.48 0.01 0.88 -26.65  
## t.mean 3 67.30 5.73 0.01 0.89 -29.98  
## maxoacc 5 67.36 5.79 0.01 0.90 -26.80  
## os.mean.acc 5 67.39 5.82 0.01 0.91 -26.82  
## q.mean 3 67.44 5.87 0.01 0.92 -30.05  
## oq.mean.acc 5 67.48 5.91 0.01 0.93 -26.86  
## mean\*acc 5 67.48 5.92 0.01 0.94 -26.87  
## s.mean 3 67.49 5.92 0.01 0.95 -30.08  
## s.mean.acc 3 67.49 5.92 0.01 0.96 -30.08  
## ot.mean.acc 5 67.55 5.98 0.01 0.97 -26.90  
## lat 3 67.67 6.10 0.01 0.98 -30.17  
## mean 3 67.70 6.13 0.01 0.99 -30.18  
## ot.mean 4 70.22 8.65 0.00 0.99 -29.93  
## oq.mean 4 70.45 8.88 0.00 0.99 -30.05  
## olat 4 70.49 8.92 0.00 0.99 -30.07  
## os.mean 4 70.51 8.94 0.00 1.00 -30.08  
## omean 4 70.60 9.03 0.00 1.00 -30.12  
## olat\*acc 9 74.25 12.68 0.00 1.00 -20.62  
## os.mean\*acc 9 74.84 13.27 0.00 1.00 -20.92  
## omean\*acc 9 75.04 13.47 0.00 1.00 -21.02  
## maxoacc\* 9 76.52 14.95 0.00 1.00 -21.76  
## oq.mean\*acc 9 76.87 15.30 0.00 1.00 -21.94  
## ot.mean\*acc 9 80.19 18.62 0.00 1.00 -23.59

##   
## Call:  
## lm(formula = LT50 ~ max \* acc, data = prot1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.0718 -0.4062 -0.0082 0.4908 1.1459   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 17.18529 12.38796 1.387 0.1823   
## max 0.95638 0.43607 2.193 0.0417 \*  
## acc 1.03966 0.56794 1.831 0.0838 .  
## max:acc -0.04539 0.01993 -2.277 0.0352 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.797 on 18 degrees of freedom  
## Multiple R-squared: 0.4291, Adjusted R-squared: 0.334   
## F-statistic: 4.511 on 3 and 18 DF, p-value: 0.0158

## # A tibble: 2 x 13  
## # Groups: acc [2]  
## acc r.squared adj.r.squared sigma statistic p.value df logLik AIC BIC  
## <int> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 20 0.0368 -0.0114 0.982 0.764 0.392 1 -29.8 65.5 68.8  
## 2 24 0.0368 -0.0114 0.982 0.764 0.392 1 -29.8 65.5 68.8  
## # ... with 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>

##   
## Call:  
## lm(formula = LT50 ~ max, data = a20)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.76732 -0.42952 -0.00582 0.36106 1.13846   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.97841 1.10844 34.263 2.43e-13 \*\*\*  
## max 0.04861 0.03964 1.226 0.244   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5927 on 12 degrees of freedom  
## Multiple R-squared: 0.1114, Adjusted R-squared: 0.03731   
## F-statistic: 1.504 on 1 and 12 DF, p-value: 0.2436

## [1] 0.6352615



## [1] 39.32378

## [1] NaN

## [1] 0.6040965

## [1] NA

## Warming Tolerance

Deutsch et al. (2008 PNAS) discuss “Warming Tolerance,” which they define as WT = -, or the difference between thermal tolerance and temperature of the habitat. is usually measured as the mean temperature. The smaller the warming tolerance is, they more likely that population is to be vulnerable to climate change. I will calculate a mean WT, an upper 90th percentile wt, and Wt with habitat maximum.

##   
## Model selection based on AICc:  
##   
## K AICc Delta\_AICc AICcWt Cum.Wt LL  
## Mod35 5 61.79 0.00 0.68 0.68 -24.02  
## Mod34 4 63.96 2.17 0.23 0.90 -26.80  
## Mod33 3 66.88 5.09 0.05 0.96 -29.77  
## Mod36 5 67.36 5.57 0.04 1.00 -26.80  
## Mod37 9 76.52 14.73 0.00 1.00 -21.76  
## Mod5 3 85.20 23.42 0.00 1.00 -38.94  
## Mod10 4 85.27 23.48 0.00 1.00 -37.46  
## Mod16 5 86.89 25.10 0.00 1.00 -36.57  
## Mod21 4 88.22 26.43 0.00 1.00 -38.93  
## Mod26 5 88.26 26.47 0.00 1.00 -37.26  
## Mod4 3 91.77 29.98 0.00 1.00 -42.22  
## Mod9 4 92.14 30.35 0.00 1.00 -40.89  
## Mod15 5 94.30 32.51 0.00 1.00 -40.27  
## Mod25 5 94.47 32.68 0.00 1.00 -40.36  
## Mod20 4 94.54 32.75 0.00 1.00 -42.09  
## Mod3 3 96.31 34.53 0.00 1.00 -44.49  
## Mod19 4 97.18 35.40 0.00 1.00 -43.42  
## Mod24 5 97.41 35.62 0.00 1.00 -41.83  
## Mod8 4 97.87 36.08 0.00 1.00 -43.76  
## Mod32 9 98.66 36.87 0.00 1.00 -32.83  
## Mod14 5 99.91 38.12 0.00 1.00 -43.08  
## Mod30 9 101.36 39.58 0.00 1.00 -34.18  
## Mod31 9 104.24 42.45 0.00 1.00 -35.62  
## Mod18 4 106.10 44.32 0.00 1.00 -47.87  
## Mod23 5 106.26 44.48 0.00 1.00 -46.26  
## Mod29 9 107.44 45.65 0.00 1.00 -37.22  
## Mod17 4 108.47 46.68 0.00 1.00 -49.06  
## Mod22 5 108.59 46.81 0.00 1.00 -47.42  
## Mod28 9 112.00 50.21 0.00 1.00 -39.50  
## Mod27 4 113.88 52.09 0.00 1.00 -51.76  
## Mod1 3 125.28 63.49 0.00 1.00 -58.97  
## Mod6 4 128.09 66.30 0.00 1.00 -58.87  
## Mod2 3 128.78 67.00 0.00 1.00 -60.73  
## Mod12 5 130.10 68.31 0.00 1.00 -58.17  
## Mod7 4 131.64 69.85 0.00 1.00 -60.64  
## Mod13 5 133.75 71.96 0.00 1.00 -60.00  
## Mod11 3 135.96 74.17 0.00 1.00 -64.31

##   
## Call:  
## glm(formula = wt.max ~ max \* acc, family = "gaussian", data = prot1)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.0718 -0.4062 -0.0082 0.4908 1.1459   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.97841 1.49053 25.480 1.42e-15 \*\*\*  
## max -0.95139 0.05330 -17.851 6.79e-13 \*\*\*  
## acc24 4.15862 2.27177 1.831 0.0838 .   
## max:acc24 -0.18156 0.07973 -2.277 0.0352 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 0.6352615)  
##   
## Null deviance: 462.627 on 21 degrees of freedom  
## Residual deviance: 11.435 on 18 degrees of freedom  
## AIC: 58.037  
##   
## Number of Fisher Scoring iterations: 2

##   
## Call:  
## lm(formula = wt.max ~ max \* acc, data = prot1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.0718 -0.4062 -0.0082 0.4908 1.1459   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.97841 1.49053 25.480 1.42e-15 \*\*\*  
## max -0.95139 0.05330 -17.851 6.79e-13 \*\*\*  
## acc24 4.15862 2.27177 1.831 0.0838 .   
## max:acc24 -0.18156 0.07973 -2.277 0.0352 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.797 on 18 degrees of freedom  
## Multiple R-squared: 0.9753, Adjusted R-squared: 0.9712   
## F-statistic: 236.7 on 3 and 18 DF, p-value: 1.2e-14

