

# THE EFFECTS OF ABIOTIC AND BIOTIC FACTORS ON PERIPHYTON AND FISH COMMUNITIES IN EVERGLADES NATIONAL PARK

Danielle Montocchio & Megan Ridgway

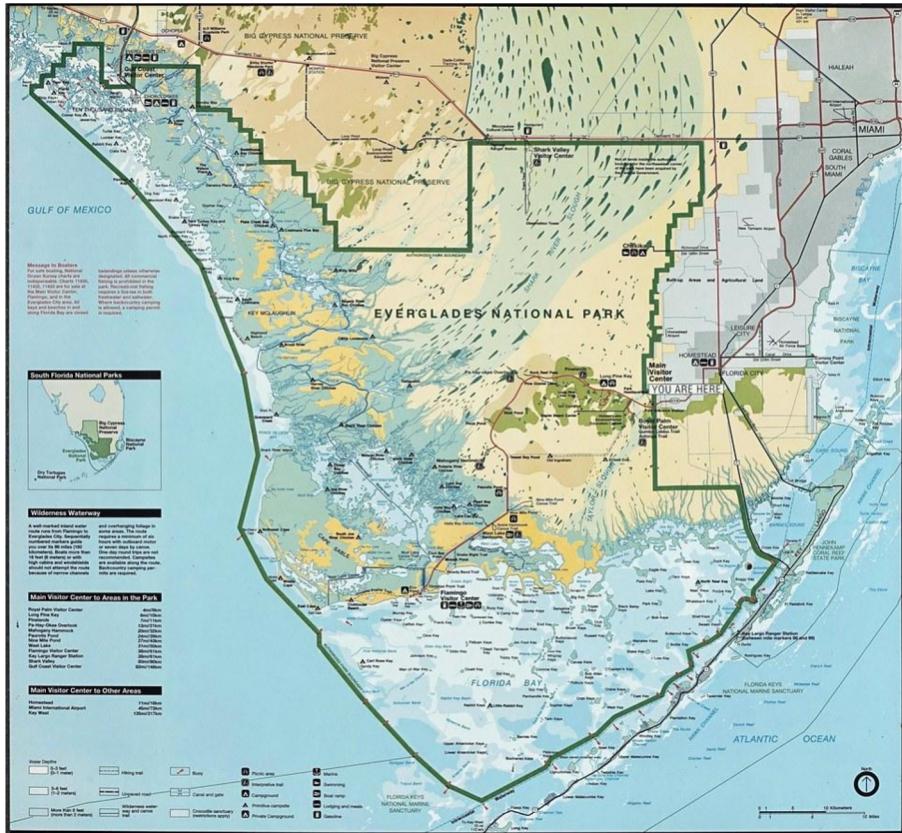
# OUTLINE

- Everglades National Park
- Data collection
- Question and hypotheses
- Data joining and cleaning
- Making the model
- Next steps



# EVERGLADES NATIONAL PARK

- Located in the Florida Everglades
- Third largest National park in the United States
- Two major biomes: temperate deciduous forest and flooded grasslands
- Includes a variety of ecosystems such as marine/estuarine, coastal lowland, mangroves, and freshwater sloughs



# AREAS OF INTEREST

- Shark River Slough (SRS)
  - dominant path for flow of water into ENP
  - Mixture of sawgrass marshes, tree islands and wet prairies
  - Begins in Water Conservation Area 3, flows through ENP and ends in Florida Bay
- Taylor Slough
  - Smaller wetland system to the east of SRS
  - Both sloughs together form the principle natural drainages for the freshwater Everglades



# DATA COLLECTION: WATER QUALITY

- Water samples collected at each wetland sites every 3–4 weeks
- Analysed the samples for total phosphorus, total nitrogen and salinity
- Grab samples were analysed for inorganic nutrients
- Water temperature and turbidity was also measured in the field



# DATA COLLECTION: PLANT COMMUNITY

- For every 1 m<sup>2</sup> throw trap, emergent plant species were identified
- All samples returned to the lab for further analysis including total species counts



# DATA COLLECTION: PERIPHYTON

- Artificial blades, which acts as seagrass would, collected periphyton over time (2 months)
- Blades are collected periphyton is isolated and periphyte volume is determined
- Periphytometers are placed on the water for 2 months to estimate periphyton cover



# DATA COLLECTION: FISH COMMUNITY

- Stratified random sampling design
- Each canal was divided into multiple sampling stations 200m apart which are randomly selected in advance of sampling
- Sampling repeated 4 times per year
- Electrofishing used to collect larger species, which were identified to and measured
- **Counted:** all centrarchids, fundulus and small, rare fish
- **Not counted:** mollies, mosquitofish, flagfish, bluefins, sheepsheads and silversides
- Seine-fishing used to collect smaller species, invertebrates and other organisms



# BIOLOGICAL QUESTIONS:

What biological impact does water temperature and dissolved oxygen have on fish biomass and fish diversity?

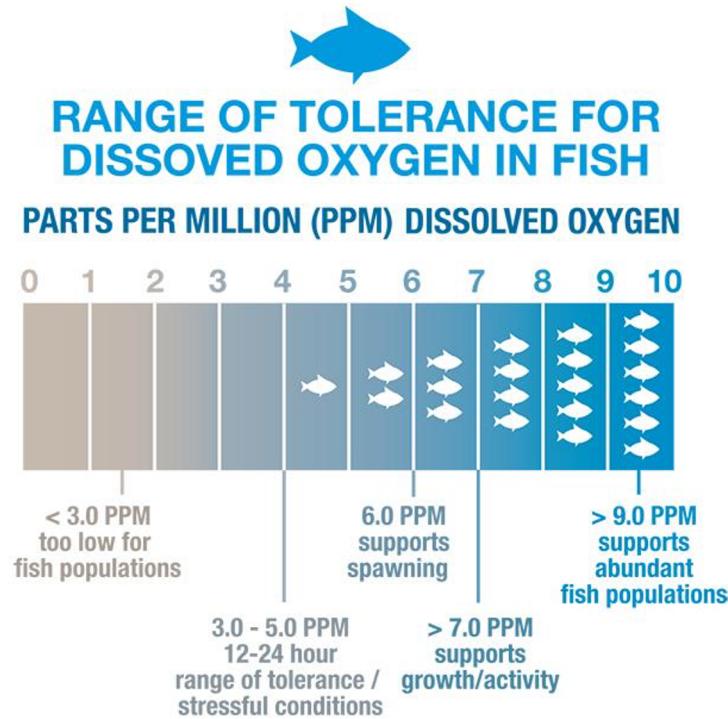
How does average plant cover and water quality impact periphyton density?

How does periphyton coverage impact fish biomass/composition of fish functional groups?

# HYPOTHESIS 1:

## WATER TEMPERATURE AND DO IMPACTS FISH BIOMASS AND DIVERSITY

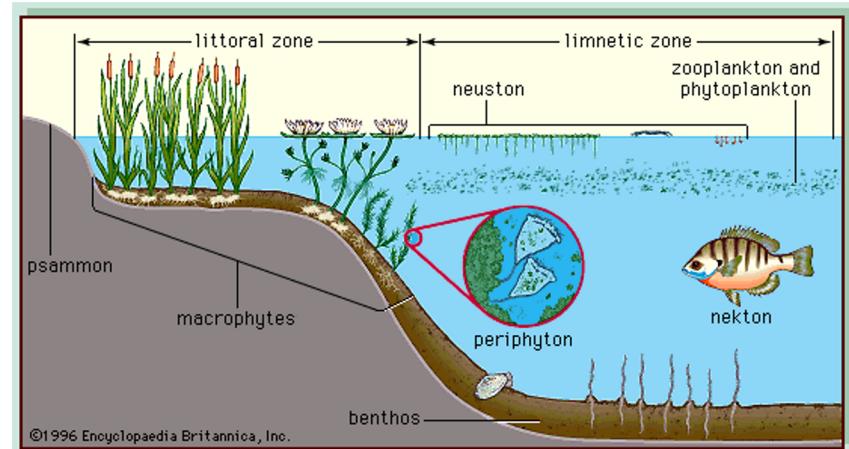
- As temperature increases, and DO decreases, fish biomass and diversity will decrease
- Elevated temperatures and lower DO will decrease the carrying capacity of an aquatic environment to support a larger fish community
- Diversity will decrease, as more sensitive species would no longer be able to inhabit the area, and only hypoxia-tolerant species would remain



## HYPOTHESIS 2:

# WATER QUALITY AND PLANT COVERAGE IMPACTS PERIPHYTON COVERAGE

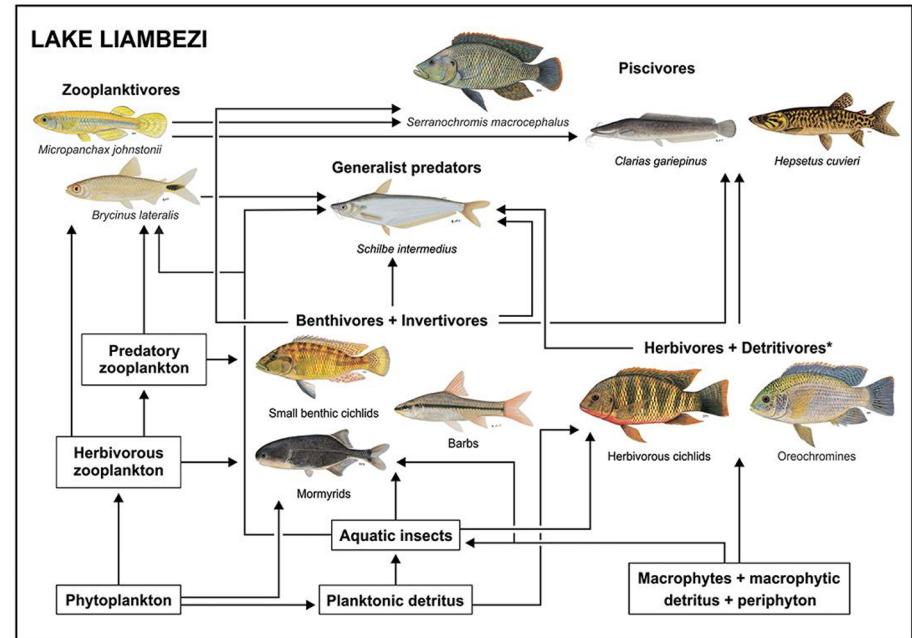
- Periphyton is a critical food source for many invertebrates and fish species, in particular, herbivorous and omnivorous species
- Periphyton is linked to nutrient loading, particularly total phosphorus (TP) and DO (Kannavillil and Kurissey, 2013)
- As TP increases, and DO decreases, periphyton coverage will increase
- Since periphyton requires substrate, periphyton coverage will increase as plant coverage/density increases



# HYPOTHESIS 3:

## PERiphyton COVERAGE IMPACTS FISH COMMUNITIES

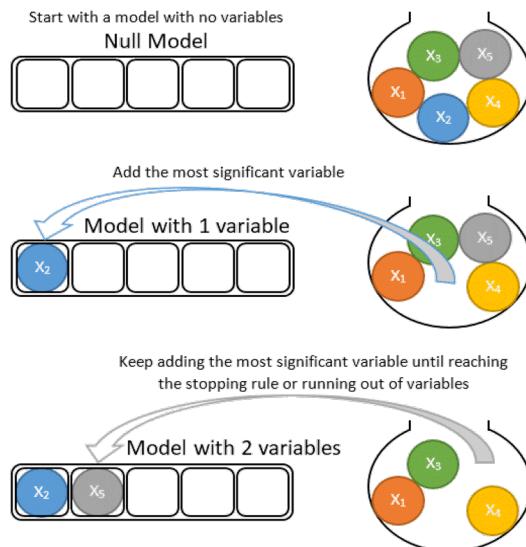
- As periphyton coverage increases, so too will fish biomass since periphyton can be a food source for herbivorous fish
- Periphyton coverage will have a stronger effect on herbivorous and omnivorous fish species more so than carnivorous species
- Periphyton will still have an effect on carnivorous species due to a bottom-up effect through the food web



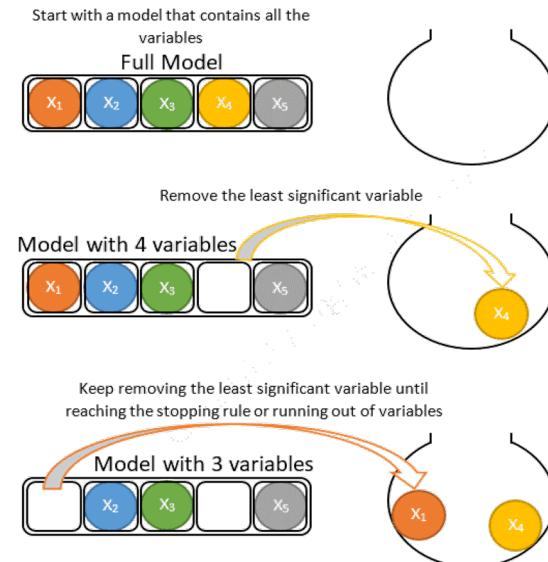
# HOW DID WE TEST THESE HYPOTHESES?

- Diagnostic and residuals analysis
- Correlation matrices with explanatory variables
- Stepwise Linear regressions (generalized linear regression models)

Forward stepwise selection example with 5 variables:

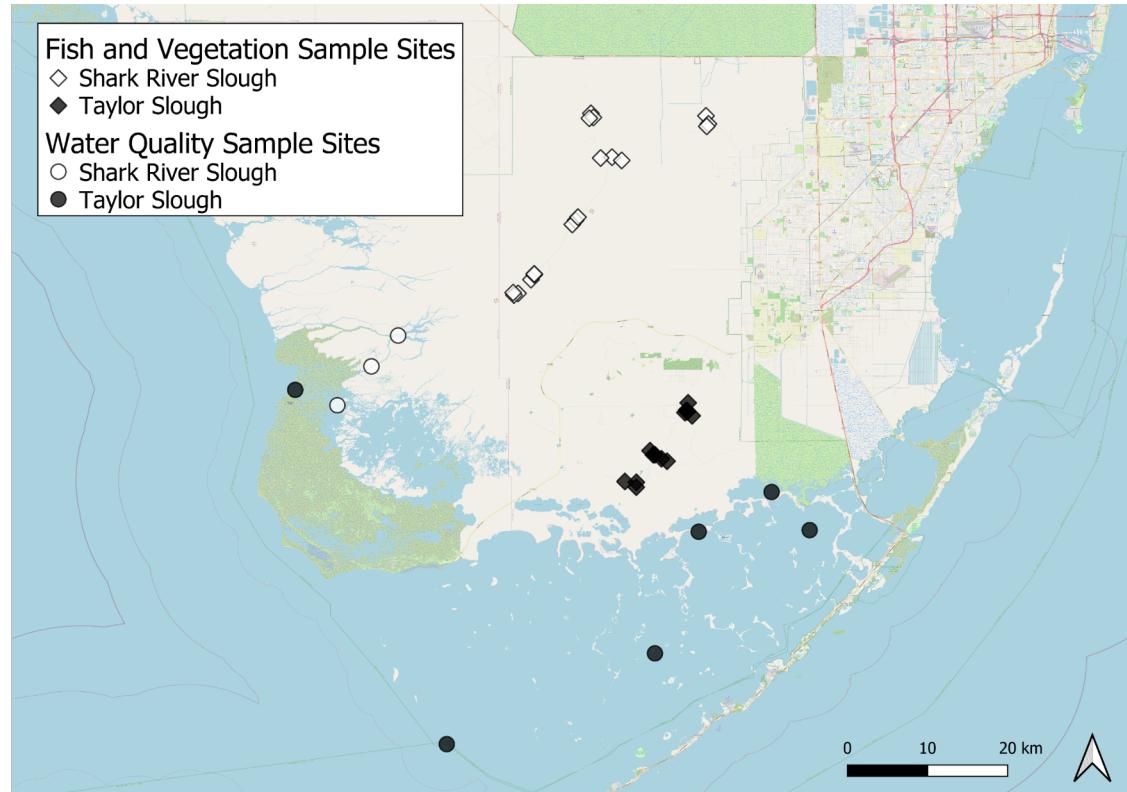


Backward stepwise selection example with 5 variables:



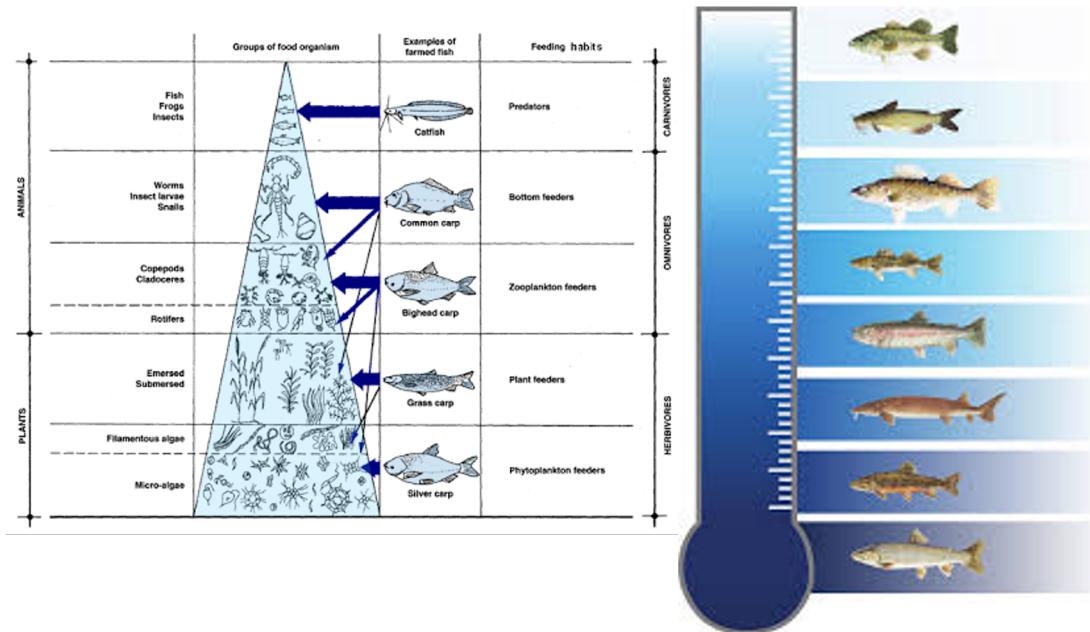
# DATA JOINING

- Data was collected by different researchers
  - Abiotic and biotic were separate
  - Missing linking site ID code
  - GIS analysis revealed most overlap occurred at the waterway level (Areas:Shark River and Taylor)



# DATA JOINING

- Fish data join with categorical variables
  - Functional Groups
    - invertivore, piscivore, planktivore, detritivore, algivore, omnivore
  - Thermal Guild
    - cold ( $\leq 19^{\circ}\text{C}$ ), cool ( $> 19^{\circ}\text{C}, < 22^{\circ}\text{C}$ ), warm ( $\geq 22^{\circ}\text{C}$ ), cool/warm ( $> 19^{\circ}\text{C}$ ; thermal-tolerant)

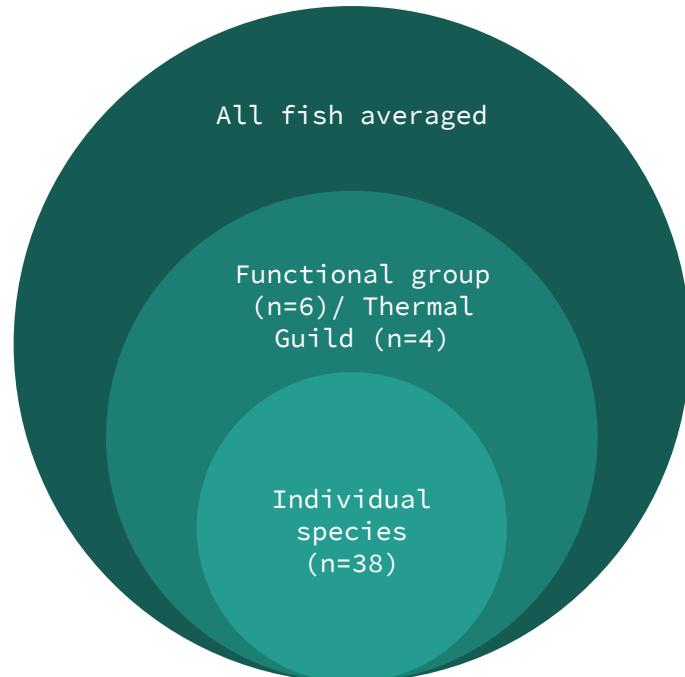


# DATA CLEANING AND ANALYSIS FLOW (HYPOTHESIS 1)

1. Check variable distributions for approximate normality
  - i. explanatory variables were roughly normal
  - ii. Biomass (one of the response variables) was not
    - a. this was log-transformed
2. Remove all NA's and 0's (same thing in this dataset) as these had no biological meaning for the variables (such as 0 depth and biomass = dry site/not sampled)
3. A full model with all explanatory variables was created for the species-level dataset
  - i. residuals of this model were evaluated with boxplots broken up into potentially influential categorical variables
    - a. i.e. Area, Month, Year --> all were found to have no differential effect on the residuals
      - March appeared to be an outlier, this was determined to be because there was only one observation in this category
  - ii. residuals with thermal guild-level dataset were evaluated
    - a. thermal guild was evaluated to have an effect

# MAKING THE MODEL

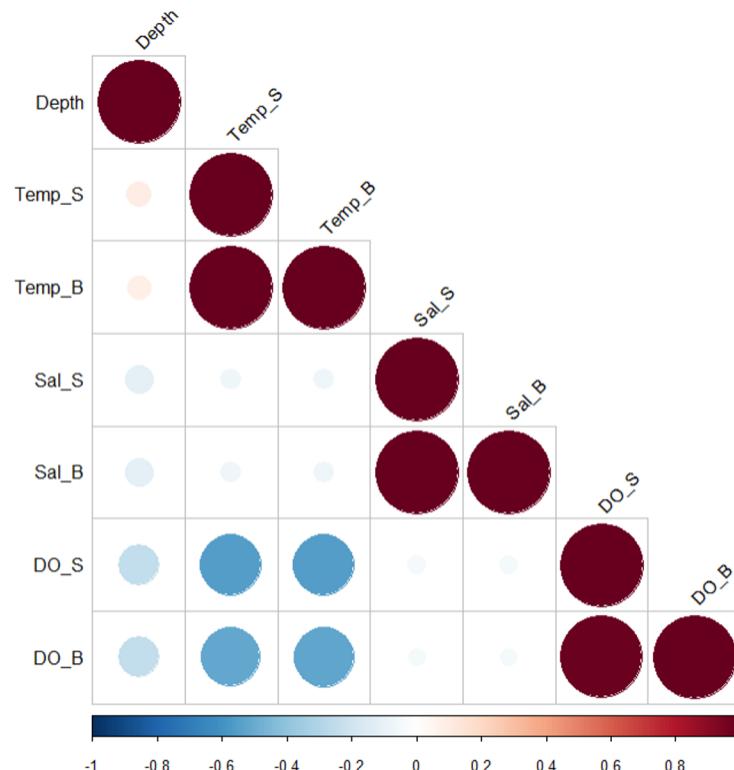
- Afterwards stepwise linear regressions (performed backwards and forwards) was performed
  - diagnostic plots were evaluated
    - species-level data results indicated model assumption violations
    - thermal-guild level → DO (bottom), depth, and guild was found to be significant
    - all fish → DO (bottom), temperature (bottom), salinity (bottom), and depth was the best fitted model for logBiomass
    - all fish → DO (bottom) and depth for 2 variable regression for logBiomass
    - all fish → salinity (bottom) was best fit for species richness
      - this was a generalized linear model with a Poisson error distribution and a logarithmic link function (count data)



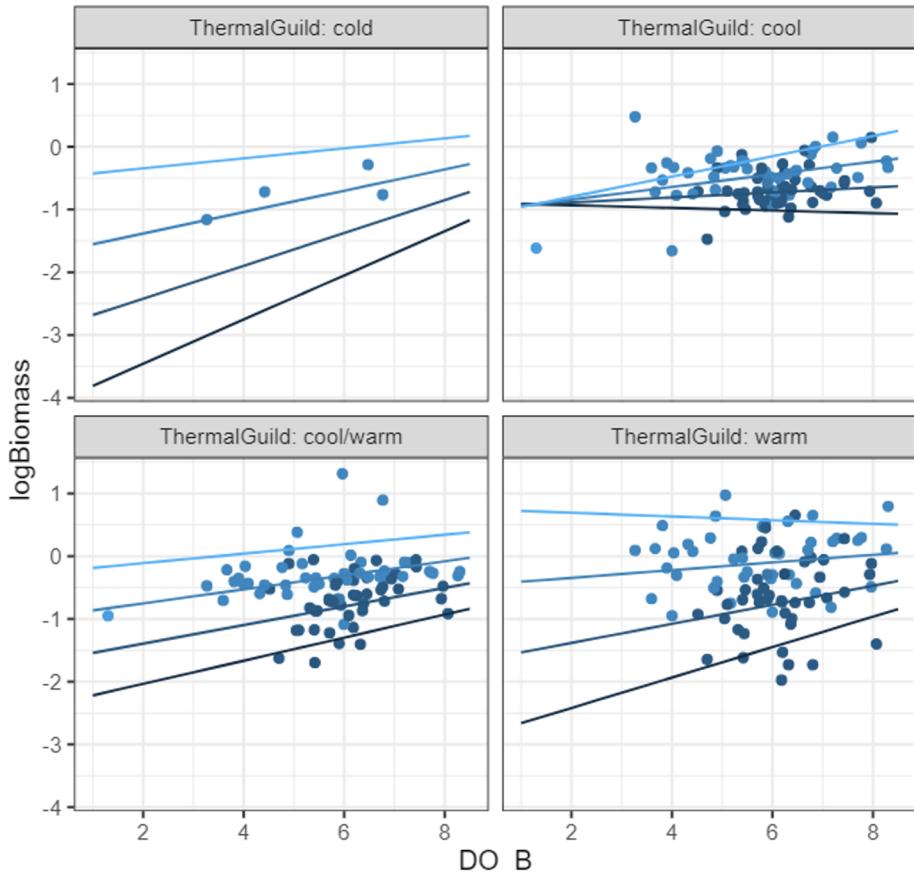
# DIAGNOSTIC PLOTS

## Hypothesis 1 – Explanatory variable correlation matrix

- Temperature and DO
  - Surface and bottom highly correlated
  - Pearson correlation coefficient  $\sim 1$ 
    - Only used bottom (where a majority of our fish species spend their time)
  - Highly correlated with each other



# MODEL 1



```

call:
lm(formula = logBiomass ~ DO_B * Depth * ThermalGuild, data = fish_therm_sub4)

Residuals:
    Min      1Q  Median      3Q     Max 
-1.33732 -0.21195 -0.01246  0.23313  1.55673 

Coefficients: (2 not defined because of singularities)
                                         Estimate Std. Error t value Pr(>|t|)    
(Intercept)                         -7.8165   3.0747  -2.542  0.0116 *  
DO_B                                 0.6238   0.5067  1.231  0.2194    
Depth                                2.4372   1.1562  2.108  0.0360 *  
ThermalGuildcool                      7.1552   3.8917  1.839  0.0672 .  
ThermalGuildcool/warm                 3.2727   3.8917  0.841  0.4012    
ThermalGuildwarm                      1.2560   0.8545  1.470  0.1428    
DO_B:Depth                            -0.1813   0.1888  -0.960  0.3377    
DO_B:ThermalGuildcool                -0.8270   0.6397  -1.293  0.1972    
DO_B:ThermalGuildcool/warm           -0.3310   0.6397  -0.517  0.6053    
DO_B:ThermalGuildwarm                -0.1092   0.1561  -0.700  0.4846    
Depth:ThermalGuildcool               -2.5901   1.4968  -1.730  0.0848 .  
Depth:ThermalGuildcool/warm          -1.0098   1.4968  -0.675  0.5005    
Depth:ThermalGuildwarm               NA        NA        NA        NA      
DO_B:Depth:ThermalGuildcool         0.3029   0.2453  1.235  0.2182    
DO_B:Depth:ThermalGuildcool/warm   0.1089   0.2453  0.444  0.6576    
DO_B:Depth:ThermalGuildwarm        NA        NA        NA        NA      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 0.4304 on 251 degrees of freedom
Multiple R-squared:  0.2963, Adjusted R-squared:  0.2599 
F-statistic:  8.13 on 13 and 251 DF,  p-value: 1.354e-13

```

## MODEL 2

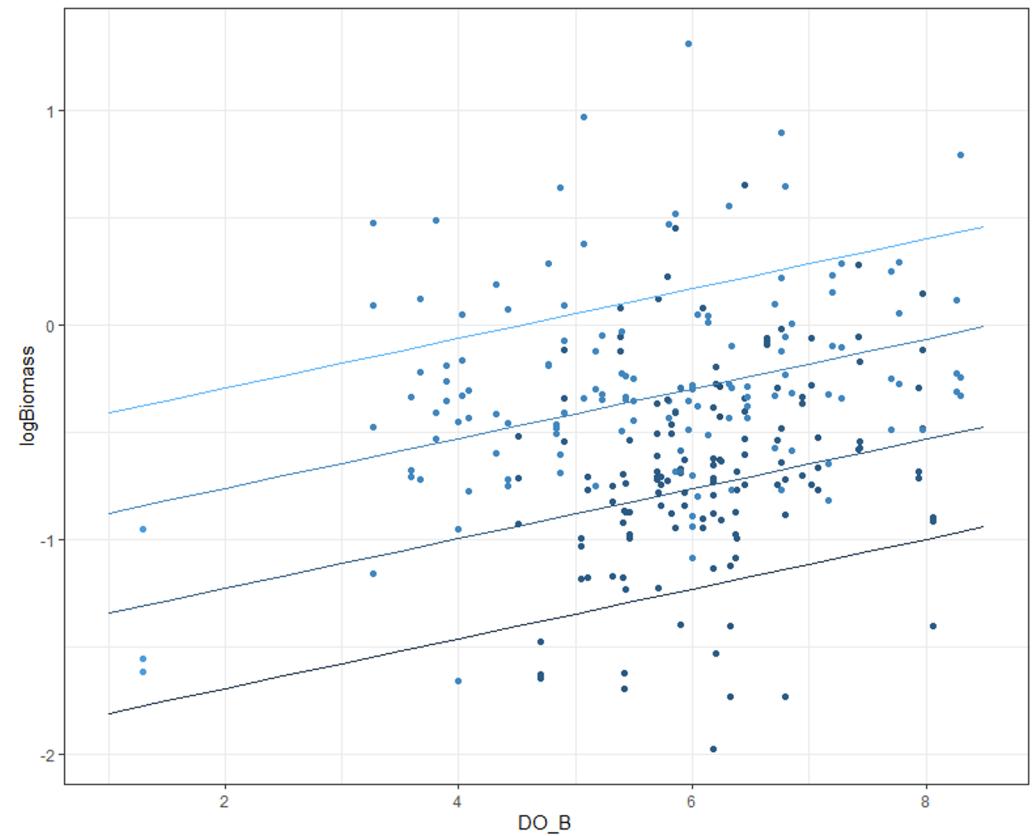
```
Call:
lm(formula = logBiomass ~ DO_B + Temp_B + Depth + Sal_B, data = all_fish_sub2)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.99273 -0.18445 -0.01944  0.21710  0.93336 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -1.783997   0.912053  -1.956   0.0538 .  
DO_B         0.067000   0.041519   1.614   0.1104  
Temp_B       -0.016151   0.014071  -1.148   0.2544  
Depth        0.854308   0.256645   3.329   0.0013 ** 
Sal_B        -0.003950   0.006735  -0.587   0.5591  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

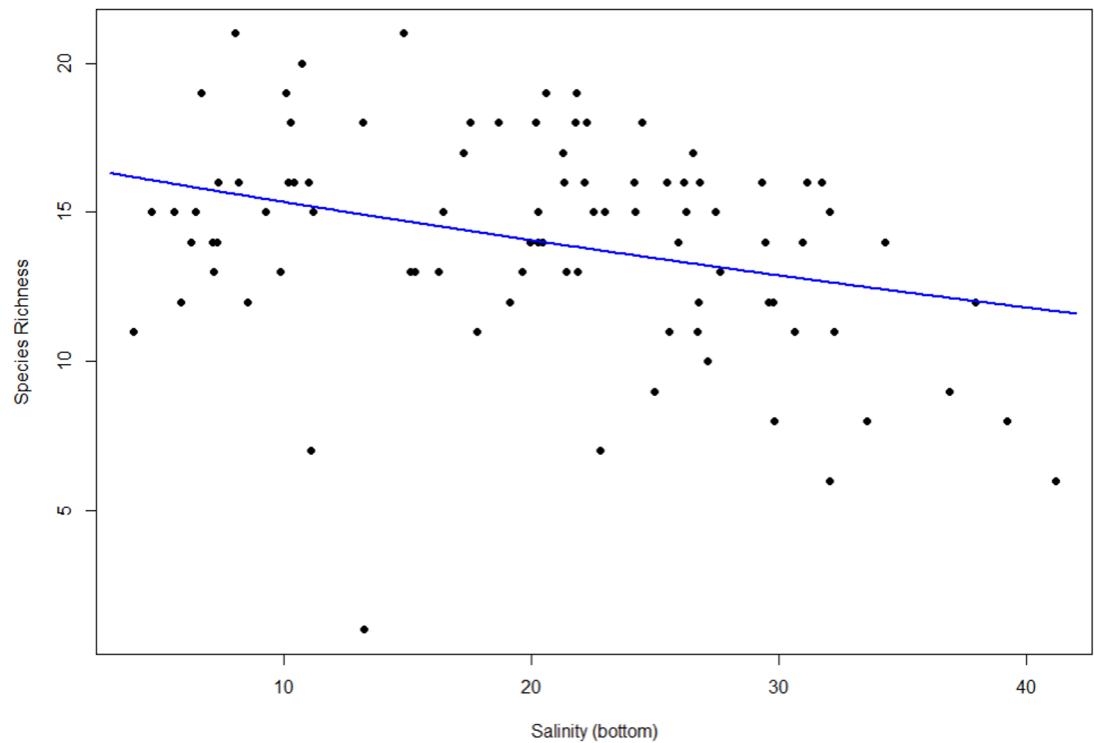
Residual standard error: 0.388 on 83 degrees of freedom
Multiple R-squared:  0.3043, Adjusted R-squared:  0.2707 
F-statistic: 9.074 on 4 and 83 DF,  p-value: 3.944e-06
```

# MODEL 3



```
Call:  
lm(formula = logBiomass ~ DO_B + Depth, data = fish_therm_sub3)  
  
Residuals:  
    Min      1Q  Median      3Q     Max  
-1.31023 -0.23812 -0.03378  0.24737  1.55233  
  
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) -3.32237   0.33080 -10.044 < 2e-16 ***  
DO_B         0.11556   0.02314   4.993 1.08e-06 ***  
Depth        0.93297   0.11430   8.162 1.36e-14 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.4464 on 264 degrees of freedom  
Multiple R-squared:  0.2208, Adjusted R-squared:  0.2149  
F-statistic: 37.41 on 2 and 264 DF, p-value: 4.977e-15
```

# MODEL 4



```
call:  
glm(formula = SpeciesRichness ~ Sal_B, family = poisson, data = all_fish_sub1)  
  
Deviance Residuals:  
    Min      1Q   Median      3Q     Max  
-4.7344 -0.4705  0.0373  0.6750  1.5413  
  
Coefficients:  
            Estimate Std. Error z value Pr(>|z|)  
(Intercept) 2.818110  0.066657 42.278 < 2e-16 ***  
Sal_B       -0.008755  0.003088 -2.835  0.00458 **  
---  
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
(Dispersion parameter for poisson family taken to be 1)  
  
Null deviance: 92.036 on 87 degrees of freedom  
Residual deviance: 83.982 on 86 degrees of freedom  
AIC: 479.15  
  
Number of Fisher scoring iterations: 4
```

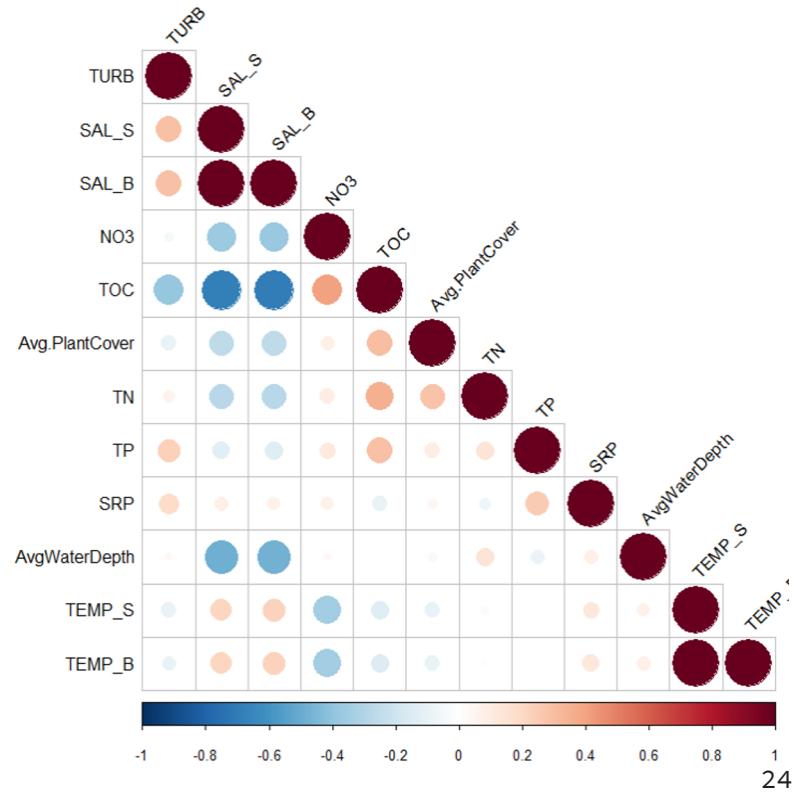
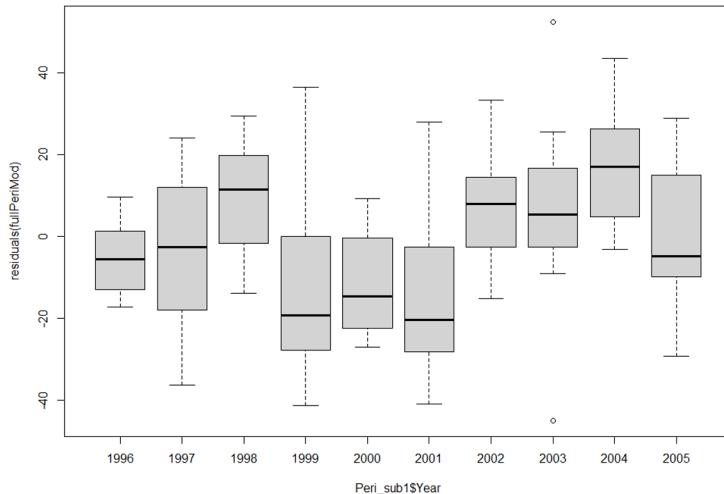
# DATA CLEANING AND ANALYSIS FLOW (HYPOTHESIS 2)

1. Check variable distributions for approximate normality
  - i. explanatory variables transformed to log or square root
    - a.  $\log(\text{average plant cover})$ ,  $\sqrt{NO_3}$ ,  $\sqrt{TN}$ ,  $\log(TP)$ ,  
 $\log(SRP)$ ,  $\log(TOC)$ ,  $\log(\text{turbidity})$
  - ii. Average periphyton cover was not transformed
2. Check for correlation between explanatory variables
3. Remove all NA's (= dry site and/or not sampled)
4. A full model with all explanatory variables was created
  - i. residuals of this model were evaluated with boxplots broken up into potentially influential categorical variables
    - a. i.e. Area, Month, Year → all were found to have no differential effect on the residuals
      - EXCEPT Year → included as categorical explanatory variable

# DIAGNOSTIC PLOTS

Hypothesis 2 – Explanatory variable correlation matrix and boxplot

- Surface and bottom measurements highly correlated
  - Pearson correlation coefficient  $\sim 1$ 
    - Only used bottom
- Salinity and TOC negatively correlated
  - Used TOC only



# MODEL 1

```
Call:
lm(formula = Avg.PeriphytonCover ~ Year + logPlantCover + AvgWaterDepth +
    sqrtTN + logTOC + logTURB + logSRP, data = Peri_sub1)

Residuals:
    Min      1Q  Median      3Q     Max 
 -54.703 -10.633 -0.338  9.970  41.335 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 184.6045   52.1520   3.540 0.000598 ***
Year1997      2.5222    8.1883   0.308 0.758676  
Year1998     20.5811   8.2414   2.497 0.014068 *  
Year1999     -3.3076   9.6466  -0.343 0.732381  
Year2000     -4.2842   9.5720  -0.448 0.655385  
Year2001     -8.4167   9.1839  -0.916 0.361528  
Year2002     20.3704   9.8539   2.067 0.041170 *  
Year2003     19.3561   9.2116   2.101 0.038011 *  
Year2004     28.3937   9.8115   2.894 0.004628 ** 
Year2005     12.3475   9.9378   1.242 0.216827  
logPlantCover -15.2958   7.5044  -2.038 0.044040 *  
AvgWaterDepth -0.2404   0.1120  -2.146 0.034155 *  
sqrtTN        5.3337   2.4417   2.184 0.031160 *  
logTOC        -45.4441  16.8397  -2.699 0.008115 ** 
logTURB        -20.7728  7.7007  -2.698 0.008140 ** 
logSRP         8.9529   6.0663   1.476 0.142979  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17.57 on 105 degrees of freedom
Multiple R-squared:  0.4131, Adjusted R-squared:  0.3292 
F-statistic: 4.927 on 15 and 105 DF,  p-value: 3.436e-07
```

# NEXT STEPS...

- Issues with previous models?
  - Species richness
- Repeat steps for Hypothesis 3
- Interpret results and link back to biological context
- Write report



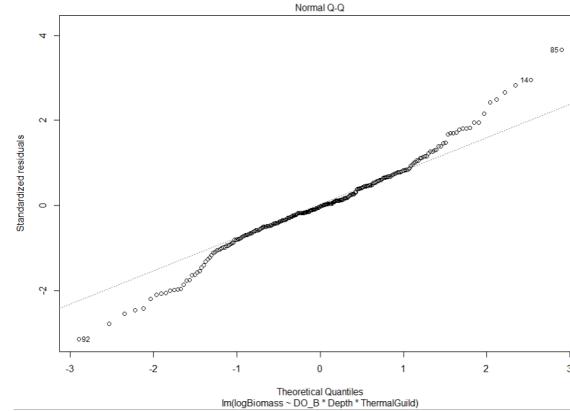
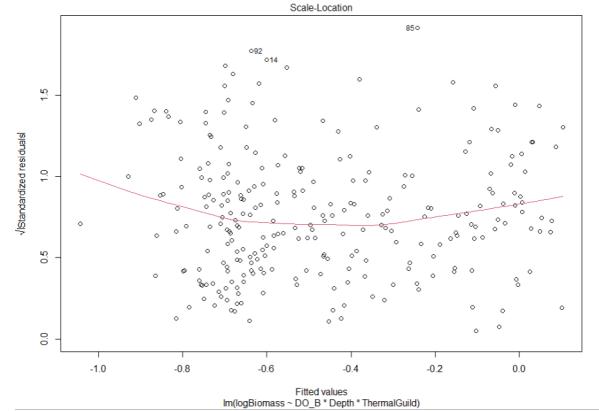
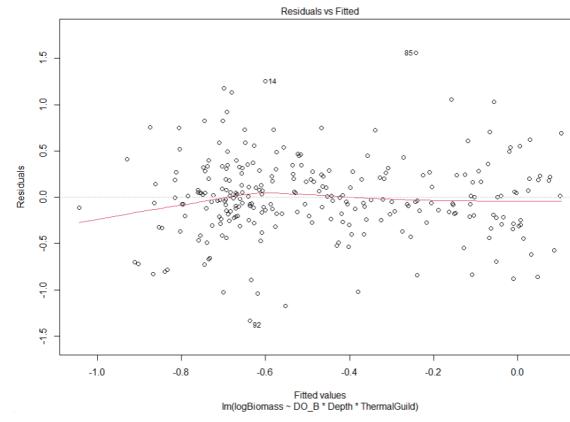
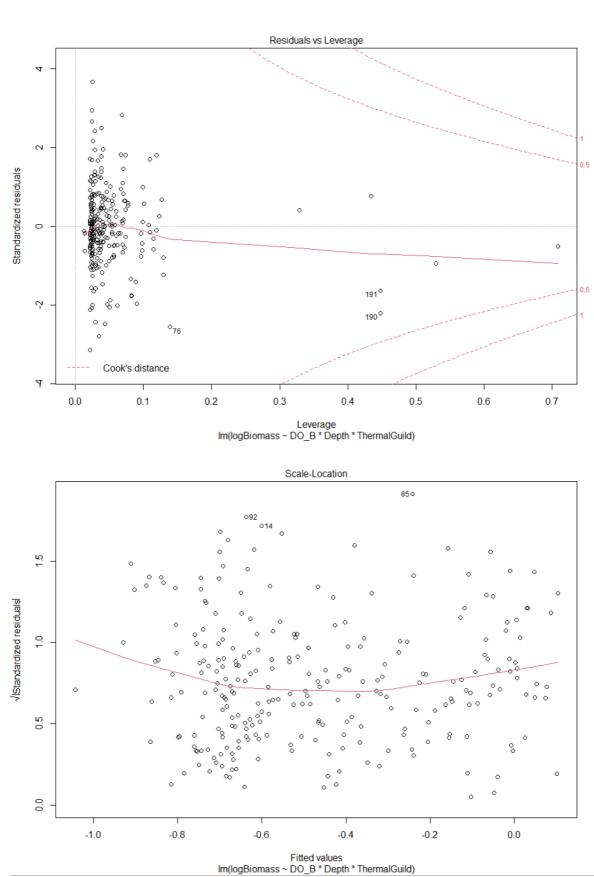


Thank You for Visiting  
Everglades National Park  
Please Drive Carefully

Questions!?

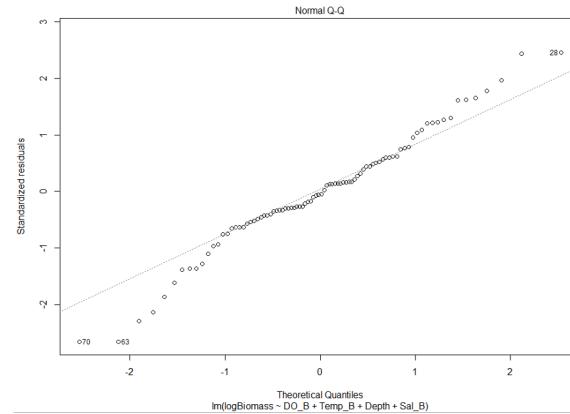
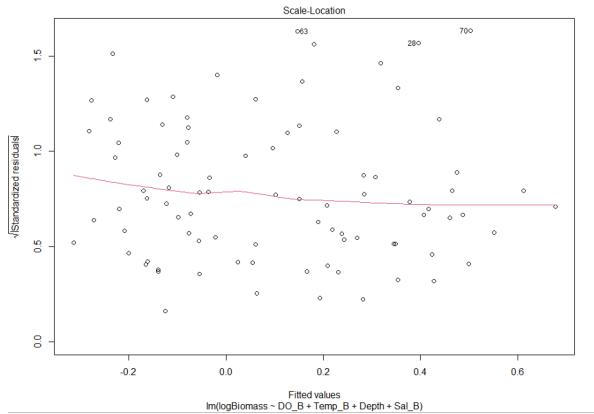
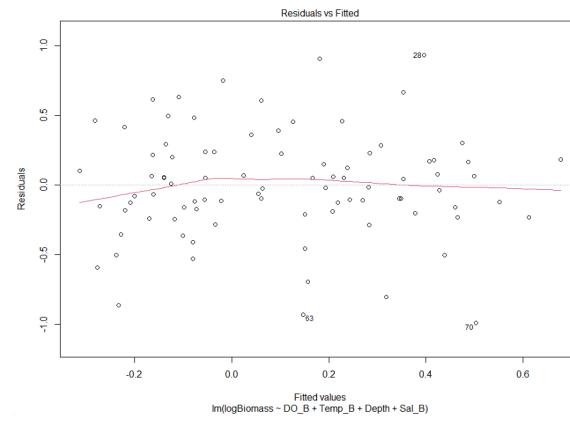
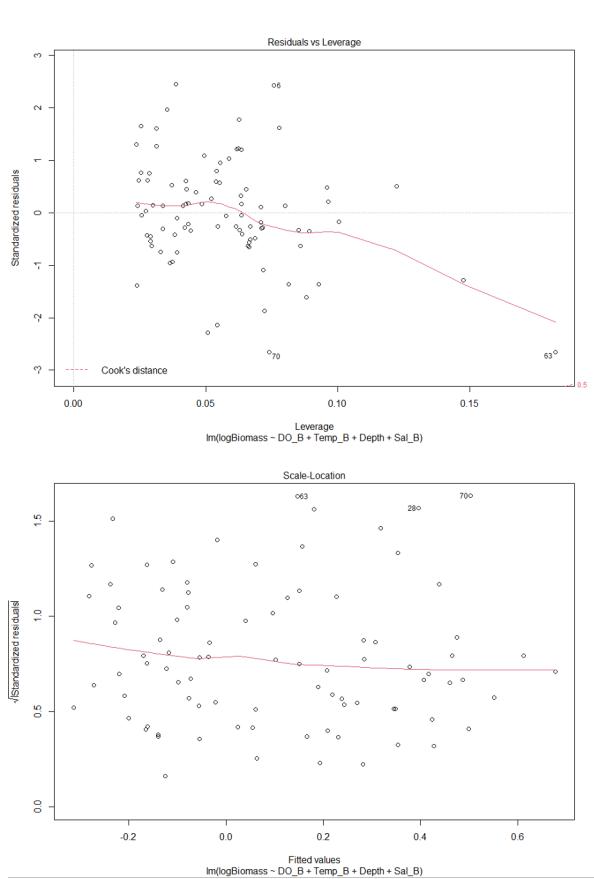
# DIAGNOSTIC PLOTS

Model 1 -  $\log\text{Biomass} \sim \text{DO}_B * \text{Depth} * \text{ThermalGuild}$



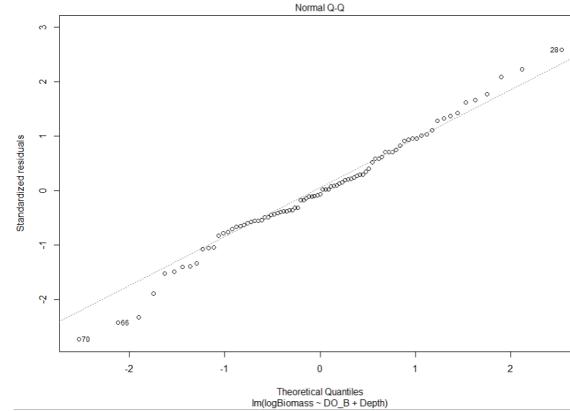
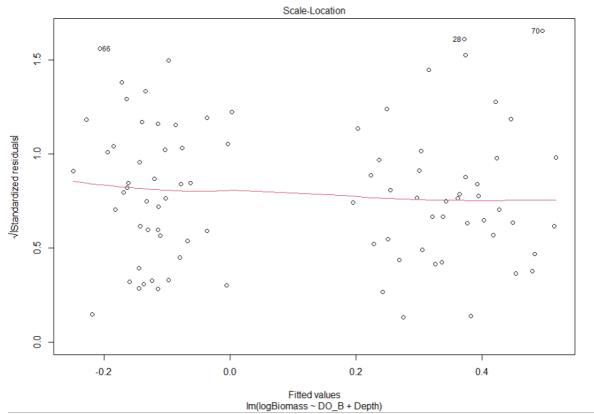
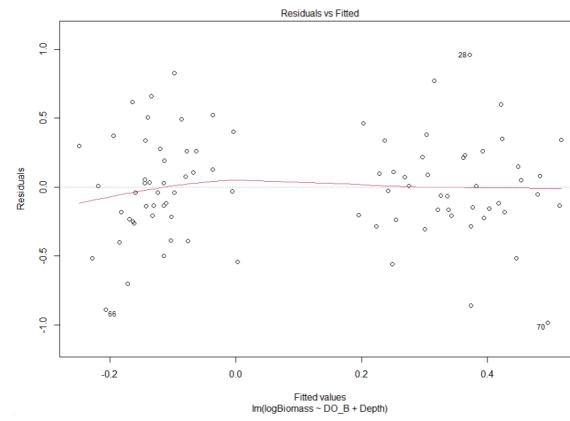
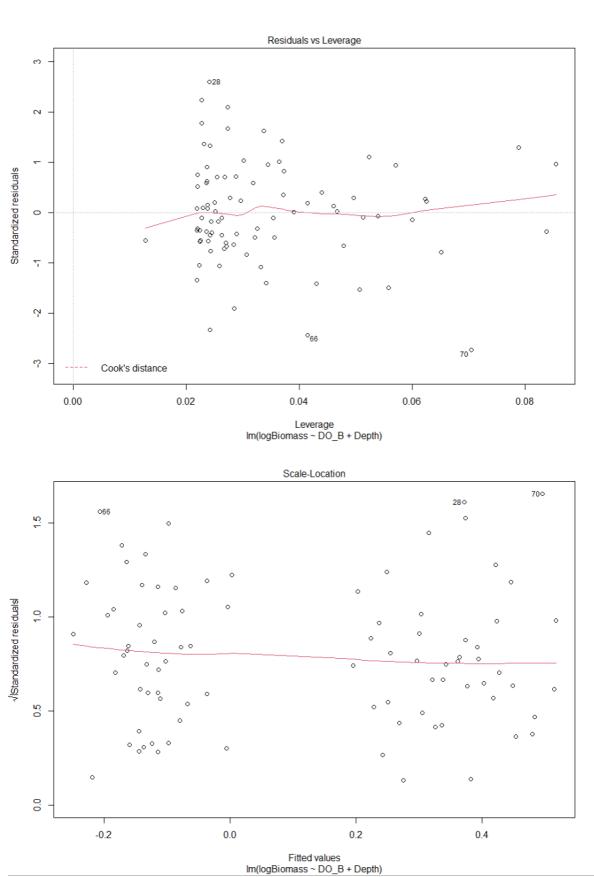
# DIAGNOSTIC PLOTS

Model 2 -  $\log\text{Biomass} \sim \text{DO}_B + \text{Depth} + \text{Sal}_B + \text{Temp}_B$  (stepwise selection)



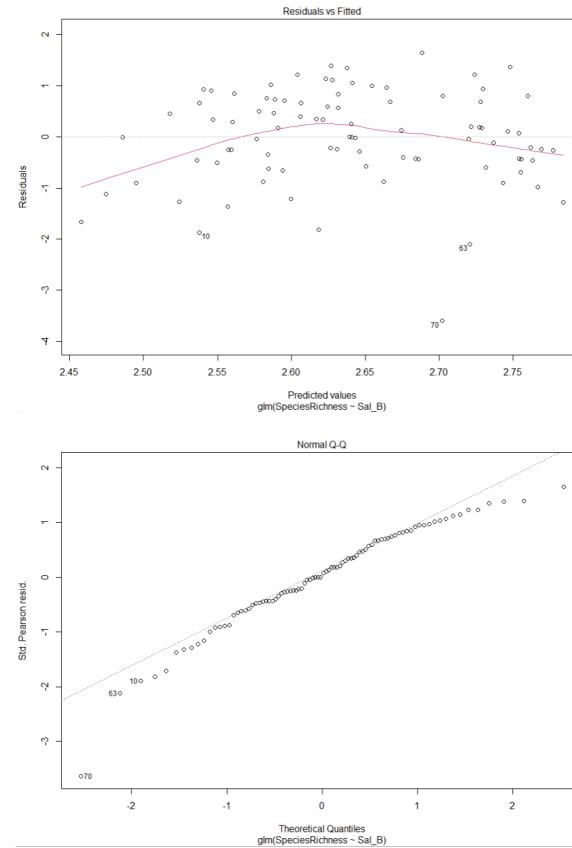
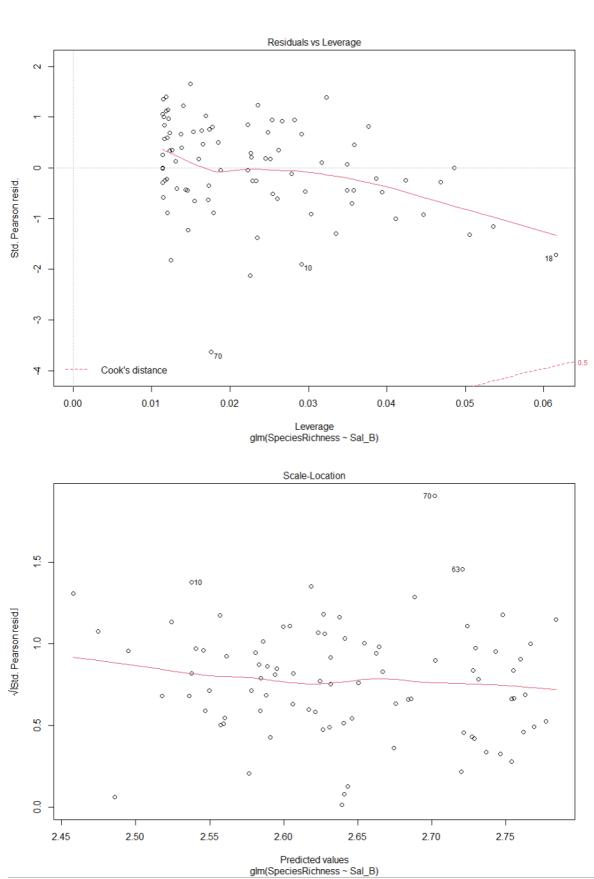
# DIAGNOSTIC PLOTS

Model 3 -  $\log\text{Biomass} \sim \text{DO\_B} + \text{Depth}$  (stepwise selection, nmax = 2)



# DIAGNOSTIC PLOTS

Model 4 - SpeciesRichness ~ Sal\_B (stepwise selection)



# DIAGNOSTIC PLOTS

Model 1 - Avg.PeriCover ~ logPlantCover + WaterDepth + sqrtTN + logSRP + logTOC + logTURB

