Rock Pool Temperature Model

Andy

12/7/2021

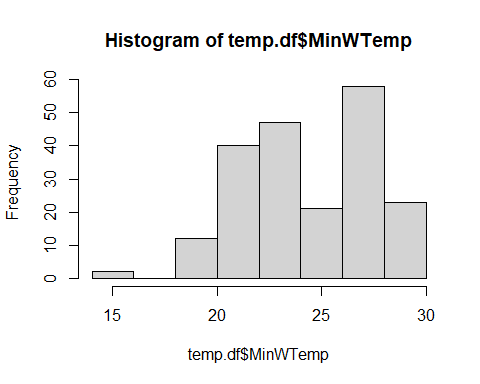
temp.df <- read.csv("temperature\_data.csv")

First, let’s take a look at some of the data and explore it for correlations.

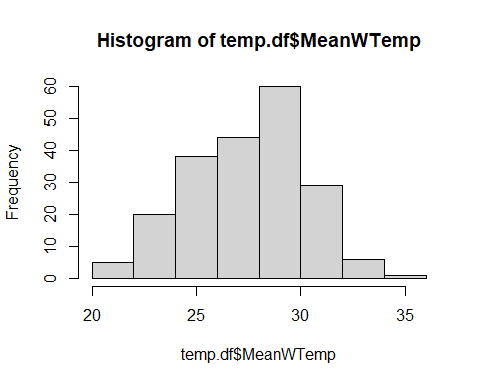
# Histograms

## Temperature Metrics

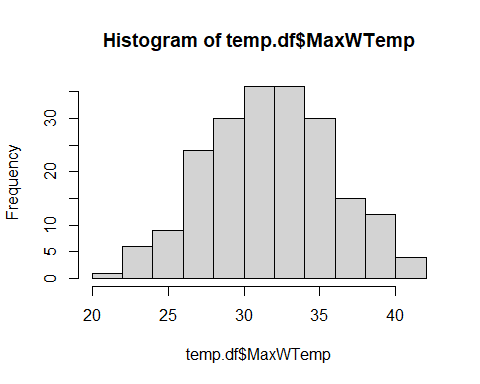
hist(temp.df$MinWTemp)



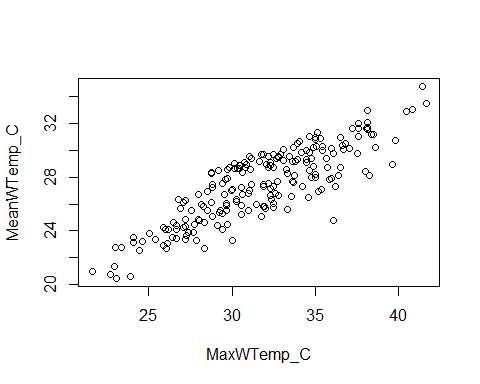
hist(temp.df$MeanWTemp)



hist(temp.df$MaxWTemp)



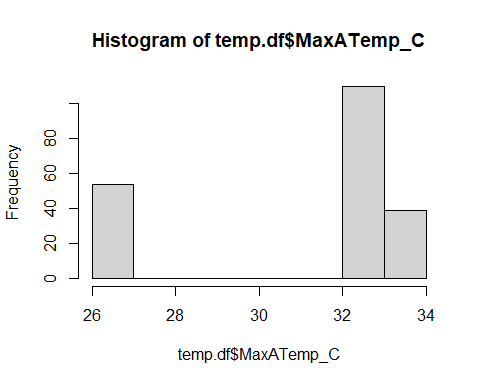
plot(MeanWTemp\_C ~ MaxWTemp\_C, data = temp.df)



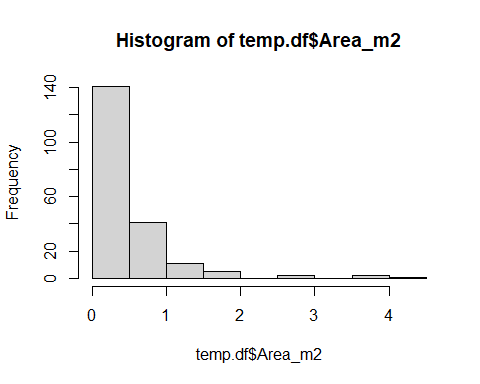
For this analysis, I’m choosing to focus in on maximum water temperature as our dependent variable, because it has a nice and pretty distribution and correlates well with mean water temperature (which is sort of what we’re more interested in here). We could easily use mean temperature as well, but I am hesitant to use it because it does not represent a true daily “mean”, because our data were collected over 8 hour periods and not a full 24 hours.

## Predictor Variables

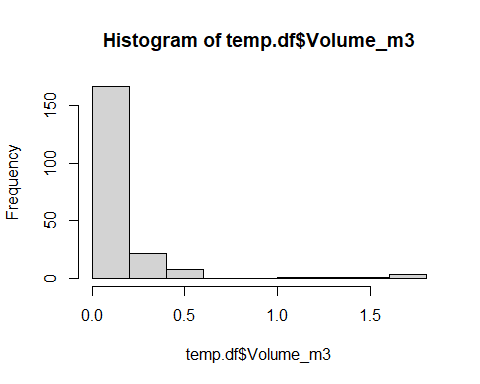
hist(temp.df$MaxATemp\_C)



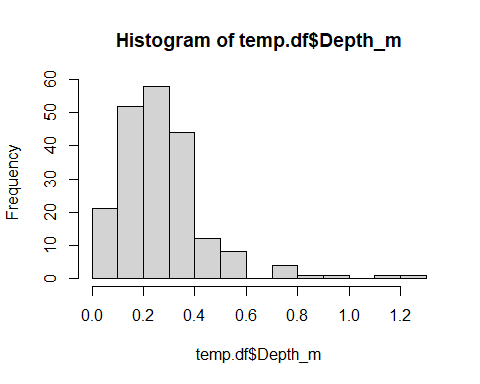
hist(temp.df$Area\_m2)



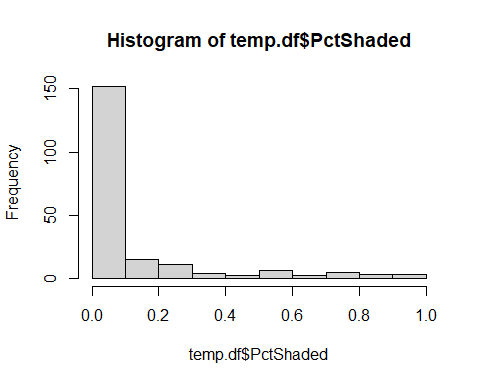
hist(temp.df$Volume\_m3)



hist(temp.df$Depth\_m)



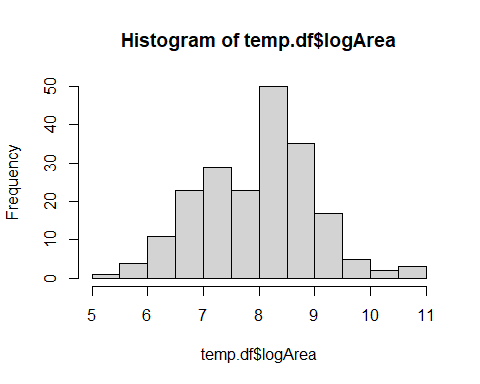
hist(temp.df$PctShaded)



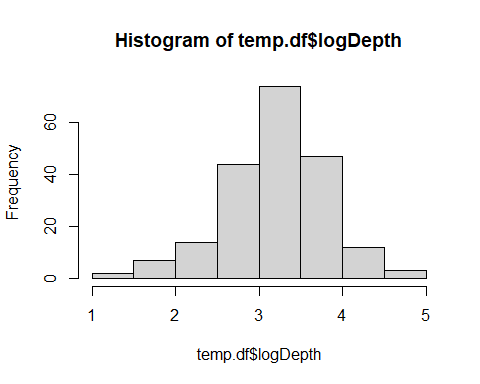
## Applying Transformations

Lots of right-skewing going on, so let’s transform these variables. We’ll use log +1 for Area and Depth, and either logit or arcsine for PctShaded since it’s a proportion.

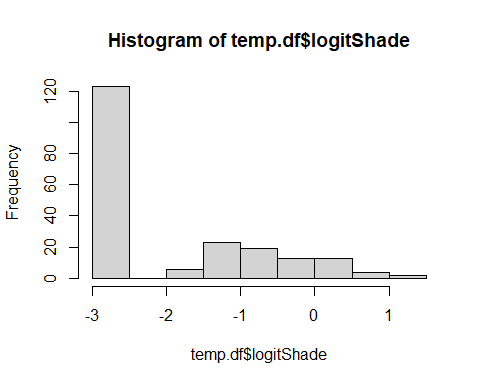
temp.df$Area\_cm2 <- temp.df$Area\_m2 \* 10000  
temp.df$logArea <- log(temp.df$Area\_cm2 + 1)  
temp.df$logDepth <- log(temp.df$Depth\_cm + 1)  
temp.df$logitShade <- log10(temp.df$PctShaded/(1-temp.df$PctShaded))  
temp.df$arcShade <- asin(sqrt(temp.df$PctShaded))  
temp.df$logShade <- log(temp.df$PctShaded + 1)  
temp.df$logVolume <- log( temp.df$Volume\_m3 + 1 )  
  
hist(temp.df$logArea)



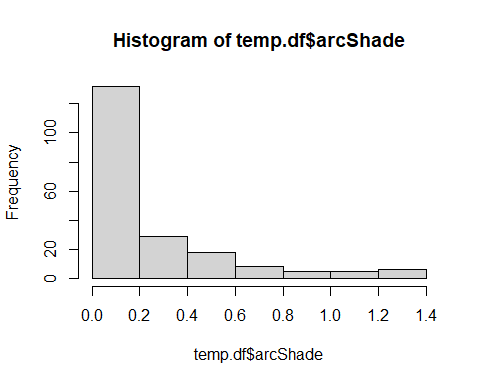
hist(temp.df$logDepth)



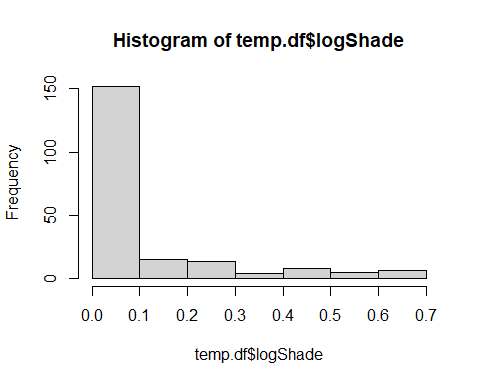
hist(temp.df$logitShade)



hist(temp.df$arcShade)



hist(temp.df$logShade)



temp.df$Date <- as.factor(temp.df$Date)

Arcsine seems to do a better job than logit, in this case, but it is still pretty zero inflated.

Are these distributions normal?

shapiro.test(temp.df$logArea)

##   
## Shapiro-Wilk normality test  
##   
## data: temp.df$logArea  
## W = 0.99165, p-value = 0.2963

shapiro.test(temp.df$logDepth)

##   
## Shapiro-Wilk normality test  
##   
## data: temp.df$logDepth  
## W = 0.98281, p-value = 0.01396

shapiro.test(temp.df$arcShade)

##   
## Shapiro-Wilk normality test  
##   
## data: temp.df$arcShade  
## W = 0.66518, p-value < 2.2e-16

Depth and shade are not. This presents a problem. Maybe it *is* time to bust out a Bayesian model.

# Correlation Plots

require(GGally)

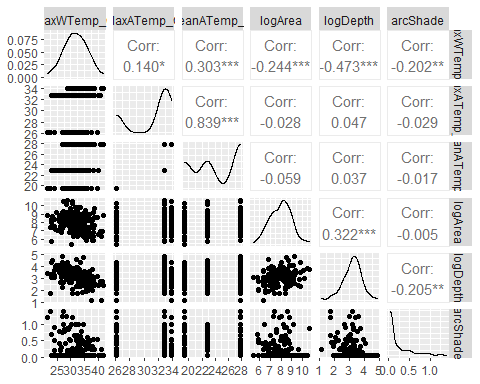
## Loading required package: GGally

## Warning: package 'GGally' was built under R version 4.1.2

## Loading required package: ggplot2

## Registered S3 method overwritten by 'GGally':  
## method from   
## +.gg ggplot2

temp2.df <- subset(temp.df, select = c("MaxWTemp\_C", "MaxATemp\_C", "MeanATemp\_C", "logArea", "logDepth", "arcShade"))  
ggpairs(data = temp2.df)



Area and depth look like good predictors, as does maximum air temperature and shade. Oddly, there is also some colinearity going on with some of our predictor variables, too (e.g., area x depth, shade x depth).

# Preliminary Modeling

Let’s try making a basic GLMM using these variables, with area and depth as fixed effects and maximum air temperature as a random effect. We’ll take a stepwise approach to model selection.

require(glmmTMB)

## Loading required package: glmmTMB

## Warning: package 'glmmTMB' was built under R version 4.1.2

## Warning in checkMatrixPackageVersion(): Package version inconsistency detected.  
## TMB was built with Matrix version 1.3.4  
## Current Matrix version is 1.3.3  
## Please re-install 'TMB' from source using install.packages('TMB', type = 'source') or ask CRAN for a binary version of 'TMB' matching CRAN's 'Matrix' package

require(performance)

## Loading required package: performance

## Warning: package 'performance' was built under R version 4.1.2

# Depth, Area, Shade, and Air temp   
mod1 <- glmmTMB(MaxWTemp\_C ~ logDepth \* logArea \* arcShade + (1 | MaxATemp\_C), data = temp.df )  
summary(mod1)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logDepth \* logArea \* arcShade + (1 | MaxATemp\_C)  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1051.4 1084.5 -515.7 1031.4 193   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## MaxATemp\_C (Intercept) 2.477 1.574   
## Residual 9.028 3.005   
## Number of obs: 203, groups: MaxATemp\_C, 3  
##   
## Dispersion estimate for gaussian family (sigma^2): 9.03   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 49.02214 10.92790 4.486 7.26e-06 \*\*\*  
## logDepth -3.85959 3.39196 -1.138 0.255   
## logArea -0.26667 1.39492 -0.191 0.848   
## arcShade -32.07827 33.66881 -0.953 0.341   
## logDepth:logArea -0.05983 0.42812 -0.140 0.889   
## logDepth:arcShade 7.70537 10.75391 0.717 0.474   
## logArea:arcShade 1.93673 4.27262 0.453 0.650   
## logDepth:logArea:arcShade -0.43604 1.35073 -0.323 0.747   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod1)

## # R2 for Mixed Models  
##   
## Conditional R2: 0.479  
## Marginal R2: 0.336

# Depth, Shade, and Air Temp  
mod2 <- glmmTMB(MaxWTemp\_C ~ logDepth \* arcShade + (1 | MaxATemp\_C), data = temp.df )  
summary(mod2)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logDepth \* arcShade + (1 | MaxATemp\_C)  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1046.4 1066.3 -517.2 1034.4 197   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## MaxATemp\_C (Intercept) 2.419 1.555   
## Residual 9.168 3.028   
## Number of obs: 203, groups: MaxATemp\_C, 3  
##   
## Dispersion estimate for gaussian family (sigma^2): 9.17   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 47.568 1.717 27.709 < 2e-16 \*\*\*  
## logDepth -4.558 0.445 -10.243 < 2e-16 \*\*\*  
## arcShade -17.641 3.323 -5.309 1.11e-07 \*\*\*  
## logDepth:arcShade 4.550 1.082 4.206 2.60e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod2)

## # R2 for Mixed Models  
##   
## Conditional R2: 0.470  
## Marginal R2: 0.330

# Area, Shade, and Air Temp  
mod3 <- glmmTMB(MaxWTemp\_C ~ logArea \* arcShade + (1 | MaxATemp\_C), data = temp.df )  
summary(mod3)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logArea \* arcShade + (1 | MaxATemp\_C)  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1117.6 1137.5 -552.8 1105.6 197   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## MaxATemp\_C (Intercept) 2.871 1.694   
## Residual 13.049 3.612   
## Number of obs: 203, groups: MaxATemp\_C, 3  
##   
## Dispersion estimate for gaussian family (sigma^2): 13   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 42.8266 2.7253 15.714 < 2e-16 \*\*\*  
## logArea -1.2484 0.3137 -3.979 6.91e-05 \*\*\*  
## arcShade -12.6812 7.2344 -1.753 0.0796 .   
## logArea:arcShade 1.2425 0.8875 1.400 0.1615   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod3)

## # R2 for Mixed Models  
##   
## Conditional R2: 0.263  
## Marginal R2: 0.101

# Depth, Shade, and Area  
mod4 <- glmmTMB(MaxWTemp\_C ~ logDepth \* logArea \* arcShade, data = temp.df )  
summary(mod4)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logDepth \* logArea \* arcShade  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1075.9 1105.7 -528.9 1057.9 194   
##   
##   
## Dispersion estimate for gaussian family (sigma^2): 10.7   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 52.55422 11.83072 4.442 8.91e-06 \*\*\*  
## logDepth -5.25888 3.68394 -1.428 0.153   
## logArea -0.71318 1.51507 -0.471 0.638   
## arcShade -34.47491 36.59332 -0.942 0.346   
## logDepth:logArea 0.09924 0.46494 0.213 0.831   
## logDepth:arcShade 8.86451 11.68336 0.759 0.448   
## logArea:arcShade 2.30442 4.64428 0.496 0.620   
## logDepth:logArea:arcShade -0.60723 1.46758 -0.414 0.679   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod4)

## [1] NA

# Depth, Area, and Air Temp  
mod5 <- glmmTMB(MaxWTemp\_C ~ logDepth \* logArea + (1 | MaxATemp\_C), data = temp.df )  
summary(mod5)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logDepth \* logArea + (1 | MaxATemp\_C)  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1088.3 1108.2 -538.1 1076.3 197   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## MaxATemp\_C (Intercept) 2.415 1.554   
## Residual 11.300 3.362   
## Number of obs: 203, groups: MaxATemp\_C, 3  
##   
## Dispersion estimate for gaussian family (sigma^2): 11.3   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 41.73632 9.06152 4.606 4.11e-06 \*\*\*  
## logDepth -1.89277 2.87819 -0.658 0.511   
## logArea -0.07118 1.16422 -0.061 0.951   
## logDepth:logArea -0.11681 0.36507 -0.320 0.749   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod5)

## # R2 for Mixed Models  
##   
## Conditional R2: 0.347  
## Marginal R2: 0.207

# Depth and Shade  
mod6 <- glmmTMB(MaxWTemp\_C ~ logDepth \* arcShade, data = temp.df )  
summary(mod6)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logDepth \* arcShade  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1069.9 1086.4 -529.9 1059.9 198   
##   
##   
## Dispersion estimate for gaussian family (sigma^2): 10.8   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 47.5660 1.5905 29.906 < 2e-16 \*\*\*  
## logDepth -4.6822 0.4823 -9.708 < 2e-16 \*\*\*  
## arcShade -16.9616 3.6101 -4.698 2.62e-06 \*\*\*  
## logDepth:arcShade 4.2781 1.1748 3.642 0.000271 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod6)

## [1] NA

# Depth and Air Temp  
mod7 <- glmmTMB(MaxWTemp\_C ~ logDepth + (1 | MaxATemp\_C), data = temp.df )  
summary(mod7)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logDepth + (1 | MaxATemp\_C)  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1087.4 1100.7 -539.7 1079.4 199   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## MaxATemp\_C (Intercept) 2.38 1.543   
## Residual 11.48 3.388   
## Number of obs: 203, groups: MaxATemp\_C, 3  
##   
## Dispersion estimate for gaussian family (sigma^2): 11.5   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 41.8249 1.5543 26.909 < 2e-16 \*\*\*  
## logDepth -3.0406 0.3953 -7.692 1.45e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod7)

## # R2 for Mixed Models  
##   
## Conditional R2: 0.336  
## Marginal R2: 0.198

# Depth and Area  
mod8 <- glmmTMB(MaxWTemp\_C ~ logDepth \* logArea, data = temp.df )  
summary(mod8)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logDepth \* logArea  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1106.3 1122.9 -548.2 1096.3 198   
##   
##   
## Dispersion estimate for gaussian family (sigma^2): 13   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 45.01405 9.63332 4.673 2.97e-06 \*\*\*  
## logDepth -3.12752 3.07211 -1.018 0.309   
## logArea -0.47561 1.24388 -0.382 0.702   
## logDepth:logArea 0.01814 0.39001 0.046 0.963   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod8)

## [1] NA

# Area and Shade  
mod9 <- glmmTMB(MaxWTemp\_C ~ logArea \* arcShade, data = temp.df )  
summary(mod9)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logArea \* arcShade  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1137.6 1154.2 -563.8 1127.6 198   
##   
##   
## Dispersion estimate for gaussian family (sigma^2): 15.1   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 41.9968 2.7322 15.371 < 2e-16 \*\*\*  
## logArea -1.2098 0.3377 -3.582 0.000341 \*\*\*  
## arcShade -10.5641 7.7783 -1.358 0.174415   
## logArea:arcShade 0.9744 0.9541 1.021 0.307118   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod9)

## [1] NA

# Area and Air Temp  
mod10 <- glmmTMB(MaxWTemp\_C ~ logArea + (1 | MaxATemp\_C), data = temp.df )  
summary(mod10)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logArea + (1 | MaxATemp\_C)  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1125.6 1138.9 -558.8 1117.6 199   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## MaxATemp\_C (Intercept) 2.823 1.680   
## Residual 13.860 3.723   
## Number of obs: 203, groups: MaxATemp\_C, 3  
##   
## Dispersion estimate for gaussian family (sigma^2): 13.9   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 40.0759 2.3119 17.335 < 2e-16 \*\*\*  
## logArea -0.9841 0.2607 -3.776 0.00016 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod10)

## # R2 for Mixed Models  
##   
## Conditional R2: 0.215  
## Marginal R2: 0.055

# Shade and Air Temp  
mod11 <- glmmTMB(MaxWTemp\_C ~ arcShade + (1 | MaxATemp\_C), data = temp.df )  
summary(mod11)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ arcShade + (1 | MaxATemp\_C)  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1130.1 1143.4 -561.1 1122.1 199   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## MaxATemp\_C (Intercept) 2.828 1.682   
## Residual 14.176 3.765   
## Number of obs: 203, groups: MaxATemp\_C, 3  
##   
## Dispersion estimate for gaussian family (sigma^2): 14.2   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 32.8183 1.0316 31.81 < 2e-16 \*\*\*  
## arcShade -2.5998 0.8442 -3.08 0.00207 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod11)

## # R2 for Mixed Models  
##   
## Conditional R2: 0.198  
## Marginal R2: 0.038

# Area alone  
mod12 <- glmmTMB(MaxWTemp\_C ~ logArea, data = temp.df )  
summary(mod12)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logArea  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1143.8 1153.7 -568.9 1137.8 200   
##   
##   
## Dispersion estimate for gaussian family (sigma^2): 15.9   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 39.7002 2.2438 17.693 < 2e-16 \*\*\*  
## logArea -1.0017 0.2792 -3.588 0.000333 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod12)

## [1] NA

# Shade alone  
mod13 <- glmmTMB(MaxWTemp\_C ~ arcShade, data = temp.df )  
summary(mod13)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ arcShade  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1147.8 1157.8 -570.9 1141.8 200   
##   
##   
## Dispersion estimate for gaussian family (sigma^2): 16.2   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 32.3165 0.3495 92.46 < 2e-16 \*\*\*  
## arcShade -2.6548 0.9030 -2.94 0.00328 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod13)

## [1] NA

# Depth alone  
mod14 <- glmmTMB(MaxWTemp\_C ~ logDepth, data = temp.df )  
summary(mod14)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logDepth  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1104.8 1114.7 -549.4 1098.8 200   
##   
##   
## Dispersion estimate for gaussian family (sigma^2): 13.1   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 41.9476 1.3608 30.827 < 2e-16 \*\*\*  
## logDepth -3.2095 0.4192 -7.657 1.91e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod14)

## [1] NA

# Air Temp alone  
mod15 <- glmmTMB(MaxWTemp\_C ~ (1 | MaxATemp\_C), data = temp.df )  
summary(mod15)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ (1 | MaxATemp\_C)  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1137.4 1147.3 -565.7 1131.4 200   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## MaxATemp\_C (Intercept) 2.871 1.694   
## Residual 14.844 3.853   
## Number of obs: 203, groups: MaxATemp\_C, 3  
##   
## Dispersion estimate for gaussian family (sigma^2): 14.8   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 32.224 1.022 31.52 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod15)

## # R2 for Mixed Models  
##   
## Conditional R2: 0.162  
## Marginal R2: 0.000

# There are four dates worth of temperature data but only three air temps for those dates. Perhaps date should be a random effect too?  
mod16 <- glmmTMB(MaxWTemp\_C ~ logDepth \* arcShade + ( 1 | Date ), data = temp.df )  
summary(mod16)

## Family: gaussian ( identity )  
## Formula: MaxWTemp\_C ~ logDepth \* arcShade + (1 | Date)  
## Data: temp.df  
##   
## AIC BIC logLik deviance df.resid   
## 1025.1 1044.9 -506.5 1013.1 197   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## Date (Intercept) 2.948 1.717   
## Residual 8.120 2.850   
## Number of obs: 203, groups: Date, 4  
##   
## Dispersion estimate for gaussian family (sigma^2): 8.12   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 47.2336 1.6251 29.066 < 2e-16 \*\*\*  
## logDepth -4.5175 0.4189 -10.785 < 2e-16 \*\*\*  
## arcShade -16.3940 3.1370 -5.226 1.73e-07 \*\*\*  
## logDepth:arcShade 4.1131 1.0217 4.026 5.68e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

r2(mod16)

## # R2 for Mixed Models  
##   
## Conditional R2: 0.517  
## Marginal R2: 0.341

Model 2, which excludes area but includes depth, has the lowest AIC. It indicates that depth and shade both negatively impact maximum water temperatures, and there is an interaction - deep, shady pools are probably especially cool. Additionally, maximum air temperature does pull out as a significant random effect. It has the second-highest r2 value, after model 1.

Curiously, the max air temperature for two of the dates is the same, so I tried it with date as a random effect instead of air temperature, which returned the highest AIC and r2 value. It’s likely that date accounts for more variation in other atmospheric conditions (i.e., temperature variation and mean) than maximum air temperature alone does.

It should be noted that there is some debate as to the validity and how to calculate r2 values for mixed models. See Nakagawa and Schielzeth 2012 in Methods in Ecology and Evolution.

## Exploring Other Model Types

Does the marginal R2 values being lower than the conditional R2 values suggest that air temperature is not a “useful” random effect? Some studies (e.g., Caissie et al. 2001) suggest that water temperature should be modeled as a logistic function of air temperature, instead. Others have done it as a linear function.

There’s also the possibility of doing it as a Bayesian hierarchical model, which could be fun, but perhaps not any more (or less) valid.

Ryland also suggested making a composite variable out of area and depth. Mike may know how.

Water temperature data for the river is unavailable.

## Some Quick “Occupancy” Modeling

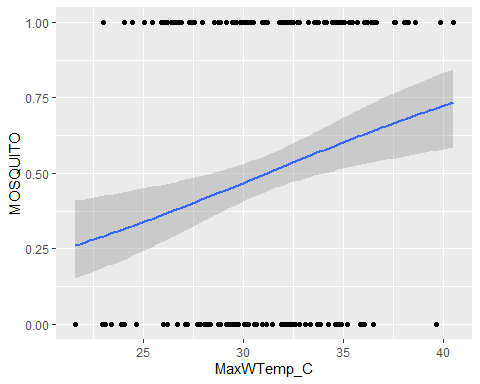
This is really just a few quick logistic fits of temperature vs. presence/absence of mosquitoes and dragonflies, for the dates we have overlap on. For the JASM abstract.

# Data manipulation  
occ.df <- read.csv("occupancy\_trimmed.csv")  
  
occ.df$PoolID <- occ.df$ï..POOL\_ID  
occ.df$ï..POOL\_ID <- NULL  
occ.df$PoolID <- as.factor(occ.df$PoolID)  
occ.df$Date <- occ.df$DATE  
occ.df$DATE <- NULL  
occ.df$Date <- as.factor(occ.df$Date)  
  
temp.df$PoolID <- as.factor(temp.df$PoolID)  
temp.df$Date <- as.factor(temp.df$Date)  
  
occ\_analysis.df <- merge(occ.df, temp.df, by = c("PoolID", "Date"))  
  
# Quick and dirty logistic models  
require(ggplot2)  
logmod1 <- glm(MOSQUITO ~ MeanWTemp\_C, family = binomial, data = occ\_analysis.df)  
summary(logmod1)

##   
## Call:  
## glm(formula = MOSQUITO ~ MeanWTemp\_C, family = binomial, data = occ\_analysis.df)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.6903 -1.1085 -0.5977 1.1323 1.6687   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -7.08366 1.63574 -4.331 1.49e-05 \*\*\*  
## MeanWTemp\_C 0.26542 0.06118 4.338 1.44e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 379.83 on 273 degrees of freedom  
## Residual deviance: 358.76 on 272 degrees of freedom  
## AIC: 362.76  
##   
## Number of Fisher Scoring iterations: 4

ggplot(occ\_analysis.df, aes(y = MOSQUITO, x = MaxWTemp\_C)) +  
 geom\_point() +   
 stat\_smooth(method = "glm", se = TRUE, method.args = list(family = binomial))

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

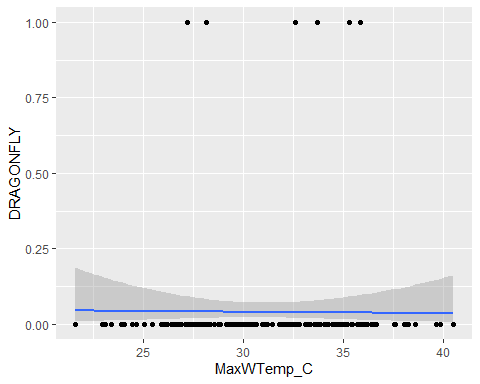


logmod2 <- glm(DRAGONFLY ~ MaxWTemp\_C, family = binomial, data = occ\_analysis.df)  
summary(logmod2)

##   
## Call:  
## glm(formula = DRAGONFLY ~ MaxWTemp\_C, family = binomial, data = occ\_analysis.df)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.3027 -0.2900 -0.2861 -0.2803 2.5583   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) -2.79851 2.52454 -1.109 0.268  
## MaxWTemp\_C -0.01214 0.08115 -0.150 0.881  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 92.287 on 273 degrees of freedom  
## Residual deviance: 92.265 on 272 degrees of freedom  
## AIC: 96.265  
##   
## Number of Fisher Scoring iterations: 6

ggplot(occ\_analysis.df, aes(y = DRAGONFLY, x = MaxWTemp\_C)) +  
 geom\_point() +   
 stat\_smooth(method = "glm", se = TRUE, method.args = list(family = binomial))

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



So… Logistic fit for the dragonflies doesn’t work at all, because there are very few pools in this set with dragons (like five?). The mosquito model suggests preference for warmer pools, but the r2 value is very low for both max and mean water temps (0.039 and 0.074 respectively) despite significant parameter estimates.