help

2023-05-20

library(lme4)

## Loading required package: Matrix

library(multcomp)

## Loading required package: mvtnorm

## Loading required package: survival

## Loading required package: TH.data

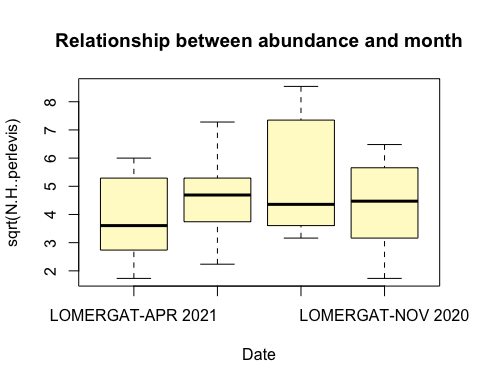
## Loading required package: MASS

##   
## Attaching package: 'TH.data'

## The following object is masked from 'package:MASS':  
##   
## geyser

Hypothesis 1: The relationship between abundance of H.perlevis and the months (from 2023 study) independent: Month of study, Dependent: abundance Analysis: Glmer model including a random effect, modelling the variability among different levels, chi-squared test, post-hoc tukey analysis

prey\_ab\_vol <- read.csv('Dataa/prey\_ab\_vol.csv')  
prey\_ab\_vol$row\_id <- factor(1:nrow(prey\_ab\_vol))  
  
boxplot(sqrt(N.H..perlevis)~ Date,prey\_ab\_vol,col="lemonchiffon", main = "Relationship between abundance and month")



m1 <- glm(N.H..perlevis ~ Date,prey\_ab\_vol,family='poisson')  
m1 <- glmer(N.H..perlevis ~ Date + (1|row\_id),prey\_ab\_vol,family='poisson')  
drop1(m1,test='Chisq')

## Single term deletions  
##   
## Model:  
## N.H..perlevis ~ Date + (1 | row\_id)  
## npar AIC LRT Pr(Chi)  
## <none> 574.87   
## Date 3 574.94 6.0749 0.108

summary(glht(m1, linfct = mcp(Date = "Tukey")))

##   
## Simultaneous Tests for General Linear Hypotheses  
##   
## Multiple Comparisons of Means: Tukey Contrasts  
##   
##   
## Fit: glmer(formula = N.H..perlevis ~ Date + (1 | row\_id), data = prey\_ab\_vol,   
## family = "poisson")  
##   
## Linear Hypotheses:  
## Estimate Std. Error z value  
## LOMERGAT-JUNE 2020 - LOMERGAT-APR 2021 == 0 0.4080 0.2295 1.778  
## LOMERGAT-JUNE 21 - LOMERGAT-APR 2021 == 0 0.5396 0.2312 2.333  
## LOMERGAT-NOV 2020 - LOMERGAT-APR 2021 == 0 0.1873 0.2610 0.718  
## LOMERGAT-JUNE 21 - LOMERGAT-JUNE 2020 == 0 0.1316 0.2051 0.642  
## LOMERGAT-NOV 2020 - LOMERGAT-JUNE 2020 == 0 -0.2207 0.2382 -0.927  
## LOMERGAT-NOV 2020 - LOMERGAT-JUNE 21 == 0 -0.3523 0.2399 -1.469  
## Pr(>|z|)   
## LOMERGAT-JUNE 2020 - LOMERGAT-APR 2021 == 0 0.2825   
## LOMERGAT-JUNE 21 - LOMERGAT-APR 2021 == 0 0.0895 .  
## LOMERGAT-NOV 2020 - LOMERGAT-APR 2021 == 0 0.8895   
## LOMERGAT-JUNE 21 - LOMERGAT-JUNE 2020 == 0 0.9179   
## LOMERGAT-NOV 2020 - LOMERGAT-JUNE 2020 == 0 0.7895   
## LOMERGAT-NOV 2020 - LOMERGAT-JUNE 21 == 0 0.4547   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
## (Adjusted p values reported -- single-step method)

Hypothesis 2: The relationship between volume of H.perlevis and the months (2023 study)

cor.test(log(prey\_ab\_vol$N.H..perlevis),log(prey\_ab\_vol$Volumen.Hym..mL.))

##   
## Pearson's product-moment correlation  
##   
## data: log(prey\_ab\_vol$N.H..perlevis) and log(prey\_ab\_vol$Volumen.Hym..mL.)  
## t = 4.9905, df = 69, p-value = 4.326e-06  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.3201512 0.6680298  
## sample estimates:  
## cor   
## 0.5149891

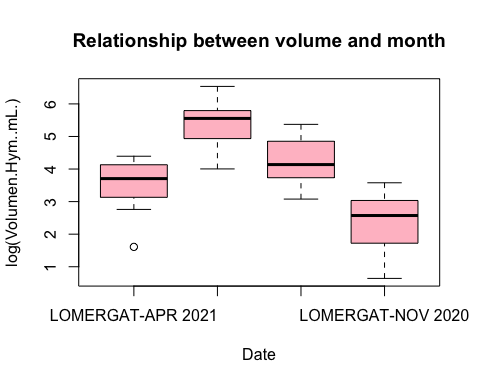
m1 <- aov(log(Volumen.Hym..mL.) ~ Date,prey\_ab\_vol)  
anova(m1)

## Analysis of Variance Table  
##   
## Response: log(Volumen.Hym..mL.)  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Date 3 82.274 27.4246 51.381 < 2.2e-16 \*\*\*  
## Residuals 67 35.762 0.5338   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

TukeyHSD(m1,ordered=TRUE)

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
## factor levels have been ordered  
##   
## Fit: aov(formula = log(Volumen.Hym..mL.) ~ Date, data = prey\_ab\_vol)  
##   
## $Date  
## diff lwr upr p adj  
## LOMERGAT-APR 2021-LOMERGAT-NOV 2020 1.1909930 0.46159661 1.920389 0.0003224  
## LOMERGAT-JUNE 21-LOMERGAT-NOV 2020 1.8540258 1.17472826 2.533323 0.0000000  
## LOMERGAT-JUNE 2020-LOMERGAT-NOV 2020 3.0564074 2.38303878 3.729776 0.0000000  
## LOMERGAT-JUNE 21-LOMERGAT-APR 2021 0.6630328 0.01230726 1.313758 0.0442034  
## LOMERGAT-JUNE 2020-LOMERGAT-APR 2021 1.8654144 1.22088054 2.509948 0.0000000  
## LOMERGAT-JUNE 2020-LOMERGAT-JUNE 21 1.2023817 0.61514235 1.789621 0.0000057

boxplot(log(Volumen.Hym..mL.)~Date,prey\_ab\_vol,main = "Relationship between volume and month",col= 'pink')



library(car)

## Loading required package: carData

leveneTest(m1)

## Warning in leveneTest.default(y = y, group = group, ...): group coerced to  
## factor.

## Levene's Test for Homogeneity of Variance (center = median)  
## Df F value Pr(>F)  
## group 3 0.4505 0.7178  
## 67

Hypothesis 3: The relationship between average volume of H.perlevis per month and pH, salinity and temperature averages for the months (2022)

TpHSal <- read.csv('Dataa/TempsalinitypH.csv')  
TpHSal$row\_id <- factor(1:nrow(TpHSal))  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following object is masked from 'package:car':  
##   
## recode

## The following object is masked from 'package:MASS':  
##   
## select

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

data <- data.frame(  
 ID\_SITE = c(5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5),  
 DATE = as.Date(c("10/6/2020", "18/6/2020", "26/6/2020", "5/11/2020", "12/11/2020", "20/11/2020", "27/11/2020", "3/4/2021", "8/4/2021", "16/4/2021", "7/6/2021", "15/6/2021", "21/6/2021", "29/6/2021"), format = "%d/%m/%Y"),  
 T = c(15.416, 15.892, 16.824, 13.417, 13.241, 13.048, 12.848, 11.072, 11.23, 10.971, 15.061, 16.295, 15.688, 15.399),  
 S = c(34.733, 34.632, 34.704, 34.552, 34.424, 34.56, 34.31, 33.62, 33.8, 34.156, 34.626, 34.762, 34.865, 35.014),  
 PH = c(8.138, 7.959, 7.953, 7.799, 8.055, 7.791, 7.794, 7.9, 7.93, 7.91, 7.94, 7.91, 7.93, 7.88)  
)  
data$DATE <- as.POSIXlt(data$DATE)  
data$Month <- data$DATE$mon + 1  
data$Year <- data$DATE$year + 1900  
averagestemp\_ph\_S <- aggregate(cbind(T, S, PH) ~ Month + Year, data, FUN = mean)  
june\_2020 <- subset(averagestemp\_ph\_S, Month == 6 & Year == 2020)  
june\_2021 <- subset(averagestemp\_ph\_S, Month == 6 & Year == 2021)  
print(averagestemp\_ph\_S)

## Month Year T S PH  
## 1 6 2020 16.04400 34.68967 8.016667  
## 2 11 2020 13.13850 34.46150 7.859750  
## 3 4 2021 11.09100 33.85867 7.913333  
## 4 6 2021 15.61075 34.81675 7.915000

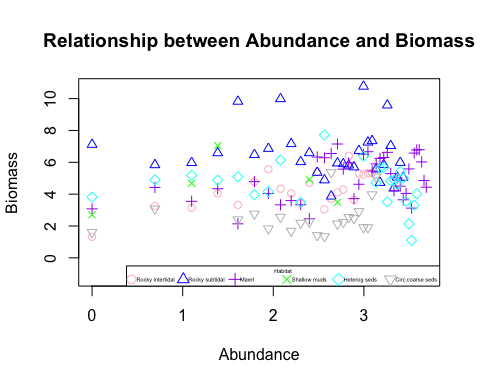
averages2 <- aggregate(Volumen.Hym..mL. / N.H..perlevis ~ Date, prey\_ab\_vol, FUN = mean)  
print(averages2)

## Date Volumen.Hym..mL./N.H..perlevis  
## 1 LOMERGAT-APR 2021 3.8781924  
## 2 LOMERGAT-JUNE 2020 12.6688840  
## 3 LOMERGAT-JUNE 21 3.0448443  
## 4 LOMERGAT-NOV 2020 0.8122513

averagesvolume <- data.frame(  
 Month = c(6, 11, 4, 6),  
 Year = c(2020, 2020, 2021, 2021),  
 Volumen.Hym..mL..N.H..perlevis = c(3.8781924, 12.6688840, 3.0448443, 0.8122513)  
)  
merged\_dataset <- merge(averagesvolume, averagestemp\_ph\_S, by = c('Year', 'Month'))  
correlation2 <- cor(merged\_dataset[, c("Volumen.Hym..mL..N.H..perlevis", "T", "S", "PH")])

Hypothesis 4: The relationship between abundance of H.perlevis and biomass per haibitat (2022 study)

data2022 <- read.csv('Dataa/data2022.csv')  
data2022$row\_id <- factor(1:nrow(data2022))  
data2022 <- data2022[data2022$Abundance>0,]  
data2022$log.n <- log(data2022$N)  
data2022$log.biomass <- log(data2022$biomass)  
  
plot(data2022$log.n, data2022$log.biomass, xlab = "Abundance", ylab = "Biomass", main = "Relationship between Abundance and Biomass", type = 'n')  
  
j <- 0  
for (i in unique(data2022$Habitat)) {  
 points(data2022$log.n[data2022$Habitat == i], data2022$log.biomass[data2022$Habitat == i],  
 pch = j + 1, col = c("pink", "blue", "purple","green", "cyan",'grey')[j + 1])  
 j <- j + 1  
}  
  
legend("bottomright", legend = unique(data2022$Habitat),  
 pch = 1:length(unique(data2022$Habitat)),  
 col = c("pink", "blue", "purple", "green", "cyan", 'grey'),  
 title = "Habitat", horiz = TRUE, cex = 0.35, pt.cex = 1.2, pt.bg = "white")



correlation3 <- cor(data2022$log.n, data2022$log.biomass)  
correlation3

## [1] 0.2006319

m10 <- lm(log.biomass ~ Habitat\*log.n, data = data2022)  
m10b <- lm(log.biomass ~ Habitat + log.n, data = data2022)  
anova(m10b,m10)

## Analysis of Variance Table  
##   
## Model 1: log.biomass ~ Habitat + log.n  
## Model 2: log.biomass ~ Habitat \* log.n  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 139 264.52   
## 2 134 237.25 5 27.265 3.0799 0.01154 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

summary(m10)

##   
## Call:  
## lm(formula = log.biomass ~ Habitat \* log.n, data = data2022)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.2712 -0.8922 -0.0812 0.6594 4.4258   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.0782 0.8740 1.234 0.219489   
## HabitatHeterog seds 4.0560 1.1477 3.534 0.000562 \*\*\*  
## HabitatMaerl 1.9553 1.1229 1.741 0.083925 .   
## HabitatRocky intertidal 1.1119 1.1975 0.928 0.354843   
## HabitatRocky subtidal 6.2112 1.1556 5.375 3.3e-07 \*\*\*  
## HabitatShallow muds 3.0138 1.4101 2.137 0.034393 \*   
## log.n 0.5668 0.3601 1.574 0.117799   
## HabitatHeterog seds:log.n -0.7834 0.4466 -1.754 0.081728 .   
## HabitatMaerl:log.n 0.1448 0.4349 0.333 0.739702   
## HabitatRocky intertidal:log.n 0.3753 0.4975 0.754 0.451927   
## HabitatRocky subtidal:log.n -0.8832 0.4591 -1.924 0.056472 .   
## HabitatShallow muds:log.n -0.2526 0.7123 -0.355 0.723416   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.331 on 134 degrees of freedom  
## Multiple R-squared: 0.525, Adjusted R-squared: 0.486   
## F-statistic: 13.47 on 11 and 134 DF, p-value: < 2.2e-16