```
# Neurobiological and behavioural responses of cleaning mutualisms to ocean warming and acidification
#
# José Ricardo Paula1*, Tiago Repolho1, Maria Rita Pegado1, Per-Ove Thörnqvist2, Regina Bispo3,
# Svante Winberg2, Philip L. Munday4, Rui Rosa1
#
# 1 MARE – Marine and Environmental Sciences Centre, Laboratório Marítimo da Guia,
# Faculdade de Ciências da Universidade de Lisboa, Av. Nossa Senhora do Cabo, 939, 2750-374 Cascais, Portugal
#
# 2 Department of Neuroscience, Physiology Unit, Biomedical Centre (BMC),
# Uppsala University, Box 593, Uppsala SE-75124, Sweden
#
# 3 Departamento de Matemática, Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia,
Universidade Nova de Lisboa, Campus da Caparica, 2829-516 Caparica, Portugal
#
# 4 ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, QLD 4811, Australia
#
# * Corresponding author: José Ricardo Paula; Telephone: +351 214869211; E-mail: jrpaula@fc.ul.pt


##set the working directory & read the data
setwd("~/Dropbox/R Projects/Neurobiological")

#Housekeeping####
#Load packages from R and support functions
library(MASS)
library(lattice)
library(mgcv)
library(ggplot2)
library(plyr)
library(effects)
source("HighstatLibV10.R") # Functions from: Mixed effects models and extensions in ecology with R. (2009).
#Zuur, AF, Ieno, EN, Walker, N, Saveliev, AA, and Smith, GM. Springer. Copyright Highland Statistics LTD.

#Import the data####
#Import the data from a tab delimited ascii file (this file is a TXT file saved from the sheet all_data in the
xlsx file data.xlsx)

data <- read.table(file = "all_data.txt",
                   header = TRUE,
                   dec = ".")

#Inspect the file
names(data)
str(data)
data$temp<-factor(data$temp)
data$ID<-factor(data$ID)

#A Outliers
names(data)
MyVar1<-
c("ID","temp","pco2","n_interactions","int_jolt","started_cleaners","started_clients","posingratio","duration_sec",

MyVar2<-
c("FB_DA_LD","FB_5HT_LD","FB_5HIAA_LD","FB_DOPAC_LD","MB_DA_LD","MB_5HT_LD","MB_5HIAA_LD","MB_DOPAC_LD","HB_DA_LD",

MyVar3<-
c("FB_DA_NE","FB_5HT_NE","FB_5HIAA_NE","FB_DOPAC_NE","MB_DA_NE","MB_5HT_NE","MB_5HIAA_NE","MB_DOPAC_NE","HB_DA_NE",


Mydotplot(data[,MyVar1])
Mydotplot(data[,MyVar2])
Mydotplot(data[,MyVar3])
summary(data)

#B. Collinearity

pairs(data[, MyVar1],
      lower.panel = panel.cor)
pairs(data[, MyVar2],
      lower.panel = panel.cor)
pairs(data[, MyVar3],
      lower.panel = panel.cor)
#C. Zero inflation
data2<-na.omit(data)
sum(data2$duration == 0) / nrow(data2)
sum(data2$started_cleaners == 0) / nrow(data2)
sum(data2$jolts == 0) / nrow(data2)
sum(data2$ts == 0) / nrow(data2)
sum(data$n_interactions == 0) / nrow(data)
sum(data$pousingratio == 0) / nrow(data)

#D. Are categorical covariates balanced?
table(data$temp)
table(data$pco2)

##Behaviour variables ####
```

```
### Data analysis ####

##### Number of interactions ####
plot(density(data$n_interactions))
M <- glm(n_interactions ~ temp*pco2, data=data, family=poisson(link=log))
step(M)
summary(M)

E <- resid(M, type = "pearson")
N <- nrow(data)
p <- length(coef(M))
Dispersion <- sum(E^2) / (N - p)
Dispersion # overdispersion 2.014

M <- glm.nb(n_interactions ~ temp*pco2, data=data)
step(M)
E <- resid(M, type = "pearson")
N <- nrow(data)
p <- length(coef(M))
Dispersion <- sum(E^2) / (N - p)
Dispersion # OK 1.36
summary(M)


# Call:
# glm.nb(formula = n_interactions ~ temp * pco2, data = data_ld,
#        init.theta = 8.85338122, link = log)
#
# Deviance Residuals:
#    Min       1Q   Median       3Q      Max
# -2.6348  -0.9189  -0.1288   0.4265   2.3522
#
# Coefficients:
#    Estimate Std. Error z value Pr(>|z|)
# (Intercept)       2.7726     0.1481  18.722  < 2e-16 ***
#    temp32           -2.6391     0.4038  -6.535 6.35e-11 ***
#    pco2High         -1.9076     0.2978  -6.406 1.50e-10 ***
#    temp32:pco2High   3.2210     0.5228   6.161 7.24e-10 ***
#    ---
#    Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# (Dispersion parameter for Negative Binomial(8.8534) family taken to be 1)
#
# Null deviance: 132.067  on 30  degrees of freedom
# Residual deviance:  38.932  on 27  degrees of freedom
# AIC: 152.25
#
# Number of Fisher Scoring iterations: 1
#
#
# Theta:  8.85
# Std. Err.:  7.78
#
# 2 x log-likelihood:  -142.254
# Model validation
# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)


#Normality
hist(E, main = "Normality", breaks=10)
#Or qq-plot
qqnorm(E)
qqline(E)

#Dependence due to model misfit
#Plot residuals versus covariates
plot(x = data$temp,
     y = E)
abline(h = 0, lty = 2)
plot(x = data$pco2,
     y = E)
abline(h = 0, lty = 2)

#Look at influential observations
plot(cooks.distance(M), type = "h", ylim = c(0, 1))
abline(h = 1, col = 2,lwd = 3) #ok
```

```
p <- ggplot()
p <- p + geom_point(data = data,
                    aes(y = E, x = n_interactions),
                    shape = 1,
                    size = 1)
p <- p + xlab("int") +
  ylab("Residuals")
p <- p + theme(text = element_text(size = 15))
p <- p + facet_grid(temp ~pco2)
p

boxplot(data$n_interactions~data$temp*data$pco2,
        ylab="int", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)


### Proportion of interactions started by cleaners ####
data$propcleaners_start <- data$started_cleaners/data$n_interactions
plot(density(na.omit(data$propcleaners_start)))
M <- glm(cbind(started_cleaners,started_clients) ~ temp*pco2, data=data, family=binomial)
step(M)
M <- glm(cbind(started_cleaners,started_clients) ~ pco2, data=data, family=binomial)
summary(M)

E <- resid(M, type = "pearson")
N <- nrow(data)
p <- length(coef(M))
Dispersion <- sum(E^2) / (N - p)
Dispersion #0.71

# Call:
#   glm(formula = cbind(started_cleaners, started_clients) ~ pco2,
#       family = binomial, data = data)
#
# Deviance Residuals:
#   Min       1Q   Median       3Q      Max
# -1.8049  -0.6281  -0.3414   0.4217   1.4688
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)   0.6712     0.1812   3.703 0.000213 ***
#   pco2High     -3.4846     0.6214  -5.607 2.06e-08 ***
#   ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# (Dispersion parameter for binomial family taken to be 1)
#
# Null deviance: 84.234  on 26  degrees of freedom
# Residual deviance: 19.371  on 25  degrees of freedom
# AIC: 56.107
#
# Number of Fisher Scoring iterations: 5

# Model validation
# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)


#Normality
hist(E, main = "Normality", breaks=10)
#Or qq-plot
qqnorm(E)
qqline(E)


#Look at influential observations
plot(cooks.distance(M), type = "h", ylim = c(0, 1))
abline(h = 1, col = 2,lwd = 3) #ok



boxplot(data$propcleaners_start~data$temp*data$pco2,
        ylab="Prop. interactions started by cleaners", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)



##### Client posing ratio ####
plot(density(na.omit(data$posingratio)))
M<- glm(posingratio ~ temp*pco2, data=data, family=gaussian(link=identity))
```

```
step(M)
summary(M)


# Call:
#   glm(formula = posingratio ~ temp * pco2, family = gaussian(link = identity),
#       data = data)
#
# Deviance Residuals:
#   Min        1Q    Median        3Q       Max
# -0.89582  -0.09285   0.00224   0.04394   0.63819
#
# Coefficients:
#   Estimate Std. Error t value Pr(>|t|)
# (Intercept)      0.06105    0.10948   0.558    0.582
# temp32           0.15838    0.16026   0.988    0.332
# pco2High         0.83922    0.15482   5.420 9.86e-06 ***
#   temp32:pco2High  0.30540    0.22283   1.371    0.182
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# (Dispersion parameter for gaussian family taken to be 0.09588213)
#
# Null deviance: 11.2415  on 30  degrees of freedom
# Residual deviance:  2.5888  on 27  degrees of freedom
# AIC: 21.008
#
# Number of Fisher Scoring iterations: 2

# Model validation
# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)



#Normality
hist(E, main = "Normality", breaks=10) #Reference Quinn and Keough (2002) Normality is not important
#Or qq-plot
qqnorm(E)
qqline(E)

#Dependence due to model misfit #should be no patterns, if there are some maybe we miss a covariate etc..
#Plot residuals versus covariates
plot(x = data$temp,
     y = E)
abline(h = 0, lty = 2)
plot(x = data$pco2,
     y = E)
abline(h = 0, lty = 2)

#Look at influential observations
plot(cooks.distance(M), type = "h", ylim = c(0, 1))
abline(h = 1, col = 2,lwd = 3) #ok


p <- ggplot()
p <- p + geom_point(data = data,
                    aes(y = E, x = posingratio),
                    shape = 1,
                    size = 1)
p <- p + xlab("int") +
  ylab("Residuals")
p <- p + theme(text = element_text(size = 15))
p <- p + facet_grid(temp ~pco2)
p

boxplot(data$posingratio~data$temp*data$pco2,
        ylab="posing", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)



##### Proportion of interactions with Tactile Stimulation #####
plot(density(na.omit(data$ts)))
data$int_ts <- data$n_interactions*data$ts
data$int_wots <- data$n_interactions-data$int_ts

M <- glm(cbind(int_ts,int_wots) ~ temp*pco2, data=data, family=binomial)
step(M)
```

```
E <- resid(M, type = "pearson")
N <- nrow(data)
p <- length(coef(M))
Dispersion <- sum(E^2) / (N - p)
Dispersion #1.43 ok!

M <- glm(cbind(int_ts,int_wots) ~ pco2, data=data, family=binomial)
summary(M)

E <- resid(M, type = "pearson")
N <- nrow(data)
p <- length(coef(M))
Dispersion <- sum(E^2) / (N - p)
Dispersion #1.39 ok!


# Call:
#   glm(formula = cbind(int_ts, int_wots) ~ pco2, family = binomial,
#       data = data)
#
# Deviance Residuals:
#   Min       1Q   Median       3Q      Max
# -1.7204  -1.1013  -0.6592   0.1903   3.8187
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)  -2.1650     0.2822  -7.672 1.69e-14 ***
#   pco2High      1.2354     0.4155   2.974  0.00294 **
#   ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# (Dispersion parameter for binomial family taken to be 1)
#
# Null deviance: 54.718  on 26  degrees of freedom
# Residual deviance: 46.017  on 25  degrees of freedom
# AIC: 75.336
#
# Number of Fisher Scoring iterations: 4


# Model validation
# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)



## Predicted vs fitted values
P <- predict(M)
plot(x = Fx,
     y = P,
     xlab = "Fitted values",
     ylab = "Predicted")
abline(h = 0, v = 0, lty = 2) # OK
#Normality
hist(E, main = "Normality", breaks=10) #refenrence quinn nd keough (2002) normality is not important
#Or qq-plot
qqnorm(E)
qqline(E)


#Look at influential observations
plot(cooks.distance(M), type = "h", ylim = c(0, 1))



boxplot(data$ts~data$temp*data$pco2,
        ylab="TS", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)


### Proportion of interactions with client Jolts ####

data$propint_jolts <- data$int_jolt/data$n_interactions
data$int_wojolts <- data$n_interactions-data$int_jolt
plot(density(na.omit(data$propint_jolts)))

M <- glm(cbind(int_jolt,int_wojolts) ~ pco2*temp, data=data, family=binomial)

E <- resid(M, type = "pearson")
```

```
N <- nrow(data)
p <- length(coef(M))
Dispersion <- sum(E^2) / (N - p)
Dispersion #1.12 ok!

step(M)


summary(M)

# Call:
# glm(formula = cbind(int_jolt, int_wojolts) ~ pco2 * temp, family = binomial,
#     data = data)
#
# Deviance Residuals:
#   Min       1Q   Median       3Q      Max
# -1.5010  -0.8478  -0.3975   0.3304   2.2342
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)      -2.7081     0.3651  -7.416  1.2e-13 ***
#   pco2High          1.0341     0.7274   1.422    0.155
# temp32            1.6094     0.8944   1.799    0.072 .
# pco2High:temp32  -1.9504     1.2162  -1.604    0.109
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# (Dispersion parameter for binomial family taken to be 1)
#
# Null deviance: 31.578  on 26  degrees of freedom
# Residual deviance: 27.317  on 23  degrees of freedom
# (4 observations deleted due to missingness)
# AIC: 58.367
#
# Number of Fisher Scoring iterations: 4


# Model validation
# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)



#Normality
hist(E, main = "Normality", breaks=10) #reference quinn and keough (2002) normality is not important
#Or qq-plot
qqnorm(E)
qqline(E)


#Look at influential observations
plot(cooks.distance(M), type = "h", ylim = c(0, 1))
abline(h = 1, col = 2,lwd = 3) #ok

boxplot(data$propint_jolts~data$temp*data$pco2,
        ylab="Prop. interactions with jolts", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)


### Interaction duration ####
plot(density(na.omit(data$duration_sec),bw=0.5))
M <- glm(duration_sec ~ temp*pco2, data=data, family=gaussian(link=identity))
step(M)
summary (M)
# Call:
#   glm(formula = duration ~ temp * pco2, family = gaussian(link = identity),
#       data = data)
#
# Deviance Residuals:
#   Min       1Q   Median       3Q      Max
# -3.2229  -1.5976  -0.5729   1.6477   4.9271
#
# Coefficients:
#   Estimate Std. Error t value Pr(>|t|)
# (Intercept)       4.6229     0.8078   5.723 7.91e-06 ***
#   temp32           -2.3104     1.3991  -1.651    0.112
# pco2High         -1.0499     1.1424  -0.919    0.368
# temp32:pco2High   2.3351     1.8319   1.275    0.215
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
#
# (Dispersion parameter for gaussian family taken to be 5.22009)
#
# Null deviance: 134.78  on 26  degrees of freedom
# Residual deviance: 120.06  on 23  degrees of freedom
# AIC: 126.91
#
# Number of Fisher Scoring iterations: 2
# Model validation
# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)


#Normality
hist(E, main = "Normality", breaks=10) #refenrence quinn nd keough (2002) normality is not important
#Or qq-plot
qqnorm(E)
qqline(E)

#Look at influential observations
plot(cooks.distance(M), type = "h", ylim = c(0, 1))
abline(h = 1, col = 2,lwd = 3) #ok



boxplot(data$duration_sec~data$temp*data$pco2,
        ylab="Duration (sec)", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)


##### Cleaner fish brain####
##Housekeeping####
#Load packages from R
library(glmmADMB)
# Import the data from a tab delimited ascii file (this file is a TXT file saved from the sheets
# cleaner_brain and client_brain in the xlsx file data.xlsx)



cleaner_brain <- read.table(file = "cleaner_brain.txt",
                            header = TRUE,
                            dec = ".")

names(cleaner_brain)
str(cleaner_brain)
cleaner_brain$temp<-factor(cleaner_brain$temp)
cleaner_brain$ID<-factor(cleaner_brain$ID)

LDDA<-cleaner_brain[,c(1:5)]
LDDA<-na.omit(LDDA)
LDDA2 <- within(LDDA, region <- relevel(region, ref = 3))
LDSER<-cleaner_brain[,c(1:4,6)]
LDSER<-na.omit(LDSER)
LDSER2 <- within(LDSER, region <- relevel(region, ref = 3))
LDHIAA<-cleaner_brain[,c(1:4,7)]
LDHIAA<-na.omit(LDHIAA)
LDHIAA2 <- within(LDHIAA, region <- relevel(region, ref = 3))
LDDOPAC<-cleaner_brain[,c(1:4,8)]
LDDOPAC<-na.omit(LDDOPAC)
LDDOPAC2 <- within(LDDOPAC, region <- relevel(region, ref = 3))

cleaner_brain2 <- within(cleaner_brain, region <- relevel(region, ref = 3))



### Dopamine ####
plot(density(na.omit(LDDA$DA)))
M <- glmmadmb(DA ~ pco2*temp*region + (1|ID), data=LDDA, family="Gamma", link="log")

stats::step(M)

M <- glmmadmb(DA ~ pco2*temp+region + (1|ID), data=LDDA, family="Gamma", link="log")
summary(M)
# Call:
#   glmmadmb(formula = DA ~ pco2 * temp + region + (1 | ID), data = LDDA,
#            family = "Gamma")
#
# AIC: 359.9
#
```

```
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)        1.139      0.270    4.22  2.5e-05 ***
#   pco2High          -0.892      0.299   -2.99   0.0028 **
#   temp32            -0.562      0.303   -1.85   0.0637 .
# regionHB           0.998      0.240    4.16  3.2e-05 ***
#   regionMB           1.263      0.244    5.17  2.3e-07 ***
#   pco2High:temp32    1.126      0.415    2.71   0.0067 **
#   ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=71, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)  0.02918 0.1708

M <- glmmadmb(DA ~ pco2*temp+region + (1|ID), data=LDDA2, family="Gamma")
summary(M)

# Call:
#   glmmadmb(formula = DA ~ pco2 * temp + region + (1 | ID), data = LDDA2,
#            family = "Gamma")
#
# AIC: 359.9
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)        2.402      0.258    9.31  < 2e-16 ***
#   pco2High          -0.892      0.299   -2.99   0.0028 **
#   temp32            -0.562      0.303   -1.85   0.0637 .
# regionFB          -1.263      0.244   -5.17  2.3e-07 ***
#   regionHB          -0.265      0.235   -1.13   0.2602
# pco2High:temp32    1.126      0.415    2.71   0.0067 **
#   ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=71, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)  0.02917 0.1708
#
# Gamma shape parameter: 1.576 (std. err.: 0.27063)
#
# Log-likelihood: -171.95


# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)



#Normality
hist(E, main = "Normality", breaks=10)
#Or qq-plot
qqnorm(E)
qqline(E)

#Dependence due to model misfit #should be no patterns, if there are some maybe we miss a covariate etc..
#Plot residuals versus covariates
plot(x = LDDA$temp,
     y = E)
abline(h = 0, lty = 2)
plot(x = LDDA$pco2,
     y = E)
abline(h = 0, lty = 2)
plot(x = LDDA$region,
     y = E)
abline(h = 0, lty = 2)



p <- ggplot()
p <- p + geom_point(data = LDDA,
                    aes(y = E, x = DA),
                    shape = 1,
                    size = 1)
p <- p + xlab("Duration") +
```

```
  ylab("Residuals")
p <- p + theme(text = element_text(size = 15))
p <- p + facet_grid(temp ~pco2)
p

boxplot(LDDA$DA~LDDA$temp*LDDA$pco2*LDDA$region,
        ylab="Dopamine", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)



### DOPAC ####
plot(density(LDDOPAC$DOPAC))
M <- glmmadmb(DOPAC ~ pco2*temp*region + (1|ID), data=LDDOPAC, family="Gamma", link="inverse")
summary(M)
stats::step(M)

# Call:
#   glmmadmb(formula = DOPAC ~ pco2 * temp * region + (1 | ID), data = LDDOPAC,
#            family = "Gamma", link = "inverse")
#
# AIC: 50.4
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)                1.918      0.581    3.30  0.00096 ***
#   pco2High                   1.403      1.135    1.24  0.21619
# temp32                    -0.318      0.726   -0.44  0.66165
# regionHB                   0.057      0.382    0.15  0.88146
# regionMB                   1.689      1.059    1.60  0.11071
# pco2High:temp32           -1.587      1.320   -1.20  0.22925
# pco2High:regionHB          0.970      1.612    0.60  0.54722
# pco2High:regionMB         -0.360      1.956   -0.18  0.85384
# temp32:regionHB            0.296      0.538    0.55  0.58179
# temp32:regionMB            1.011      1.646    0.61  0.53916
# pco2High:temp32:regionHB  -1.385      1.682   -0.82  0.41043
# pco2High:temp32:regionMB   5.878      3.780    1.56  0.11993
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=72, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)   0.7667 0.8756
#
# Gamma shape parameter: 1.729 (std. err.: 0.30532)
#
# Log-likelihood: -11.1977

M <- glmmadmb(DOPAC ~ pco2*temp*region + (1|ID), data=LDDOPAC2, family="Gamma", link="inverse")
summary(M)


# Call:
#   glmmadmb(formula = DOPAC ~ pco2 * temp * region + (1 | ID), data = LDDOPAC2,
#            family = "Gamma", link = "inverse")
#
# AIC: 50.4
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)                3.607      1.073    3.36  0.00077 ***
#   pco2High                   1.043      1.772    0.59  0.55618
# temp32                     0.693      1.675    0.41  0.67908
# regionFB                  -1.689      1.059   -1.60  0.11071
# regionHB                  -1.632      1.065   -1.53  0.12540
# pco2High:temp32            4.291      3.704    1.16  0.24669
# pco2High:regionFB          0.360      1.956    0.18  0.85384
# pco2High:regionHB          1.331      2.144    0.62  0.53488
# temp32:regionFB           -1.011      1.646   -0.61  0.53916
# temp32:regionHB           -0.715      1.672   -0.43  0.66902
# pco2High:temp32:regionFB  -5.878      3.780   -1.56  0.11993
# pco2High:temp32:regionHB  -7.263      3.890   -1.87  0.06186 .
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=72, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)   0.7667 0.8756
#
# Gamma shape parameter: 1.729 (std. err.: 0.30532)


# Homogeneity
par(mfrow =c(1,1))
```

```
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)



#Normality
hist(E, main = "Normality", breaks=10)
#Or qq-plot
qqnorm(E)
qqline(E)

#Dependence due to model misfit #should be no patterns, if there are some maybe we miss a covariate etc..
#Plot residuals versus covariates
plot(x = LDDOPAC$temp,
     y = E)
abline(h = 0, lty = 2)
plot(x = LDDOPAC$pco2,
     y = E)
abline(h = 0, lty = 2)
plot(x = LDDOPAC$region,
     y = E)
abline(h = 0, lty = 2)



p <- ggplot()
p <- p + geom_point(data = LDDOPAC,
                    aes(y = E, x = DOPAC),
                    shape = 1,
                    size = 1)
p <- p + xlab("Ser") +
  ylab("Residuals")
p <- p + theme(text = element_text(size = 15))
p <- p + facet_grid(temp ~pco2)
p

boxplot(LDDOPAC$DOPAC~LDDOPAC$temp*LDDOPAC$pco2*LDDOPAC$region,
        ylab="DOPAC", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)


### Serotonin ####
plot(density(na.omit(LDSER$SER), bw=0.5))

M <- glmmadmb(SER ~ pco2*temp*region + (1|ID), data=LDSER, family="Gamma", link="inverse")

stats::step(M)
M <- glmmadmb(SER ~ pco2*region + (1|ID), data=LDSER, family="Gamma", link="inverse")
summary(M)


# glmmadmb(formula = SER ~ pco2 * region + (1 | ID), data = LDSER,
# family = "Gamma", link = "inverse")
#
# AIC: 266.1
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)         0.0896     0.0103    8.69  < 2e-16 ***
#   pco2High            0.0140     0.0162    0.86    0.387
# regionHB            0.3128     0.0495    6.32  2.5e-10 ***
#   regionMB            0.5427     0.0735    7.38  1.5e-13 ***
#   pco2High:regionHB  -0.1617     0.0588   -2.75    0.006 **
#   pco2High:regionMB  -0.1383     0.0947   -1.46    0.144
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=70, ID=24
# Random effect variance(s):
#   Group=ID
# Variance    StdDev
# (Intercept) 1.125e-07 0.0003353
#
# Gamma shape parameter: 6.2897 (std. err.: 1.0362)
#
# Log-likelihood: -125.064

M <- glmmadmb(SER ~ pco2*region + (1|ID), data=LDSER2, family="Gamma", link="inverse")
summary(M)

# Call:
# glmmadmb(formula = SER ~ pco2 * region + (1 | ID), data = LDSER2,
```

```
#          family = "Gamma", link = "inverse")
#
# AIC: 266.1
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)         0.6323     0.0728    8.69  < 2e-16 ***
#   pco2High           -0.1243     0.0934   -1.33   0.1830
# regionFB           -0.5427     0.0735   -7.38  1.5e-13 ***
#   regionHB           -0.2299     0.0874   -2.63   0.0085 **
#   pco2High:regionFB   0.1383     0.0947    1.46   0.1444
# pco2High:regionHB  -0.0234     0.1092   -0.21   0.8301
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=70, ID=24
# Random effect variance(s):
#   Group=ID
# Variance    StdDev
# (Intercept) 1.125e-07 0.0003353
#
# Gamma shape parameter: 6.2897 (std. err.: 1.0362)
#
# Log-likelihood: -125.064

# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)



#Normality
hist(E, main = "Normality", breaks=10)
#Or qq-plot
qqnorm(E)
qqline(E)

#Dependence due to model misfit #should be no patterns, if there are some maybe we miss a covariate etc..
#Plot residuals versus covariates
plot(x = LDSER$temp,
     y = E)
abline(h = 0, lty = 2)
plot(x = LDSER$pco2,
     y = E)
abline(h = 0, lty = 2)
plot(x = LDSER$region,
     y = E)
abline(h = 0, lty = 2)



p <- ggplot()
p <- p + geom_point(data = LDSER,
                    aes(y = E, x = SER),
                    shape = 1,
                    size = 1)
p <- p + xlab("Ser") +
  ylab("Residuals")
p <- p + theme(text = element_text(size = 15))
p <- p + facet_grid(temp ~pco2)
p

boxplot(LDSER$SER~LDSER$temp*LDSER$pco2*LDSER$region,
        ylab="Seratonin", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)



### HIAA ####
plot(density(LDHIAA$HIAA))
M <- glmmadmb(HIAA ~ pco2*temp*region + (1|ID), data=LDHIAA, family="Gamma", link="inverse")
summary(M)
stats::step(M)

# Call:
#   glmmadmb(formula = HIAA ~ pco2 * temp * region + (1 | ID), data = LDHIAA,
#            family = "Gamma", link = "inverse")
#
# AIC: 131.8
#
# Coefficients:
```

```
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)                 0.8056     0.2037    3.95  7.7e-05 ***
#   pco2High                  -0.2046     0.2648   -0.77   0.4398
# temp32                     -0.0925     0.2804   -0.33   0.7416
# regionHB                   -0.1313     0.1344   -0.98   0.3284
# regionMB                    2.5513     0.9066    2.81   0.0049 **
#   pco2High:temp32             0.2556     0.3913    0.65   0.5136
# pco2High:regionHB           2.7354     0.8818    3.10   0.0019 **
#   pco2High:regionMB          -1.1011     1.0652   -1.03   0.3013
# temp32:regionHB             2.2136     0.7777    2.85   0.0044 **
#   temp32:regionMB           -1.1937     1.0711   -1.11   0.2651
# pco2High:temp32:regionHB  -4.7573     1.1852   -4.01  6.0e-05 ***
#   pco2High:temp32:regionMB   0.6831     1.2949    0.53   0.5978
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=72, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)  0.07751 0.2784
#
# Gamma shape parameter: 2.2796 (std. err.: 0.38461)
#
# Log-likelihood: -51.8951

M <- glmmadmb(HIAA ~ pco2*temp*region + (1|ID), data=LDHIAA2, family="Gamma", link="inverse")
summary(M)

# Call:
#   glmmadmb(formula = HIAA ~ pco2 * temp * region + (1 | ID), data = LDHIAA2,
#            family = "Gamma", link = "inverse")
#
# AIC: 131.8
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)                 3.357      0.900    3.73  0.00019 ***
#   pco2High                  -1.306      1.058   -1.23  0.21733
# temp32                     -1.286      1.060   -1.21  0.22511
# regionFB                   -2.551      0.907   -2.81  0.00489 **
#   regionHB                  -2.683      0.904   -2.97  0.00299 **
#   pco2High:temp32            0.939      1.280    0.73  0.46343
# pco2High:regionFB           1.101      1.065    1.03  0.30128
# pco2High:regionHB           3.837      1.362    2.82  0.00484 **
#   temp32:regionFB            1.194      1.071    1.11  0.26507
# temp32:regionHB             3.407      1.295    2.63  0.00849 **
#   pco2High:temp32:regionFB   -0.683      1.295   -0.53  0.59783
# pco2High:temp32:regionHB   -5.440      1.714   -3.17  0.00151 **
#   ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=72, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)  0.07751 0.2784
#
# Gamma shape parameter: 2.2796 (std. err.: 0.38461)
#
# Log-likelihood: -51.8951

# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)



#Normality
hist(E, main = "Normality", breaks=10)
#Or qq-plot
qqnorm(E)
qqline(E)

#Dependence due to model misfit #should be no patterns, if there are some maybe we miss a covariate etc..
#Plot residuals versus covariates
plot(x = LDHIAA$temp,
     y = E)
abline(h = 0, lty = 2)
plot(x = LDHIAA$pco2,
```

```
     y = E)
abline(h = 0, lty = 2)
plot(x = LDHIAA$region,
     y = E)
abline(h = 0, lty = 2)



p <- ggplot()
p <- p + geom_point(data = LDHIAA,
                    aes(y = E, x = HIAA),
                    shape = 1,
                    size = 1)
p <- p + xlab("Ser") +
  ylab("Residuals")
p <- p + theme(text = element_text(size = 15))
p <- p + facet_grid(temp ~pco2)
p

boxplot(LDHIAA$HIAA~LDHIAA$temp*LDHIAA$pco2*LDHIAA$region,
        ylab="RAtio Serotonin", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)

##### Client fish brain####

#Import the data from a tab delimited ascii file (this file is a TXT file saved from the sheets client_brain and
client_brain in the xlsx file data.xlsx)



client_brain <- read.table(file = "client_brain.txt",
                           header = TRUE,
                           dec = ".")

names(client_brain)
str(client_brain)
client_brain$temp<-factor(client_brain$temp)
client_brain$ID<-factor(client_brain$ID)

NEDA<-client_brain[,c(1:5)]
NEDA<-na.omit(NEDA)
NEDA2 <- within(NEDA, region <- relevel(region, ref = 3))
NESER<-client_brain[,c(1:4,6)]
NESER<-na.omit(NESER)
NESER2 <- within(NESER, region <- relevel(region, ref = 3))
NEHIAA<-client_brain[,c(1:4,7)]
NEHIAA<-na.omit(NEHIAA)
NEHIAA2 <- within(NEHIAA, region <- relevel(region, ref = 3))
NEDOPAC<-client_brain[,c(1:4,8)]
NEDOPAC<-na.omit(NEDOPAC)
NEDOPAC2 <- within(NEDOPAC, region <- relevel(region, ref = 3))

client_brain2 <- within(client_brain, region <- relevel(region, ref = 3))



### Dopamine ####
plot(density(na.omit(NEDA$DA)))
M <- glmmadmb(DA ~ pco2*temp*region + (1|ID), data=NEDA, family="Gamma", link="inverse")

stats::step(M)
M <- glmmadmb(DA ~  pco2 + temp + region + pco2:temp  + (1|ID), data=NEDA, family="Gamma", link="inverse")

summary(M)
# Call:
# glmmadmb(formula = DA ~ pco2 + temp + region + pco2:temp + (1 |
#                                                              ID), data = NEDA, family = "Gamma", link =
"inverse")
#
# AIC: 35
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)       0.1800     0.0483    3.73  0.00019 ***
#   pco2High          0.1736     0.0865    2.01  0.04459 *
#   temp32            0.0599     0.0671    0.89  0.37243
# regionHB          1.8664     0.2213    8.43  < 2e-16 ***
#   regionMB         10.4973     1.1146    9.42  < 2e-16 ***
#   pco2High:temp32  -0.1585     0.1147   -1.38  0.16691
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=72, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)  0.00416 0.0645
#
# Gamma shape parameter: 3.8814 (std. err.: 0.68169)
```

```
#
# Log-likelihood: -9.49106

M <- glmmadmb(DA ~ pco2 + temp + region + pco2:temp  + (1|ID), data=NEDA2, family="Gamma", link="inverse")
summary(M)

# Call:
#   glmmadmb(formula = DA ~ pco2 + temp + region + pco2:temp + (1 |
#                                                                ID), data = NEDA2, family = "Gamma", link =
"inverse")
#
# AIC: 35
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)      10.6774     1.1150    9.58   <2e-16 ***
#   pco2High          0.1736     0.0865    2.01    0.045 *
#   temp32            0.0599     0.0671    0.89    0.372
# regionFB        -10.4973     1.1146   -9.42   <2e-16 ***
#   regionHB         -8.6309     1.1356   -7.60    3e-14 ***
#   pco2High:temp32  -0.1585     0.1147   -1.38    0.167
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=72, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)  0.00416 0.0645
#
# Gamma shape parameter: 3.8814 (std. err.: 0.68169)
#
# Log-likelihood: -9.49106

# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)



#Normality
hist(E, main = "Normality", breaks=10)
#Or qq-plot
qqnorm(E)
qqline(E)

#Dependence due to model misfit #shouNE be no patterns, if there are some maybe we miss a covariate etc..
#Plot residuals versus covariates
plot(x = NEDA$temp,
     y = E)
abline(h = 0, lty = 2)
plot(x = NEDA$pco2,
     y = E)
abline(h = 0, lty = 2)
plot(x = NEDA$region,
     y = E)
abline(h = 0, lty = 2)



p <- ggplot()
p <- p + geom_point(data = NEDA,
                    aes(y = E, x = DA),
                    shape = 1,
                    size = 1)
p <- p + xlab("Dopamine") +
  ylab("Residuals")
p <- p + theme(text = element_text(size = 15))
p <- p + facet_grid(temp ~pco2)
p

boxplot(NEDA$DA~NEDA$temp*NEDA$pco2*NEDA$region,
        ylab="Dopamine", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)



### DOPAC ####
plot(density(NEDOPAC$DOPAC))
M <- glmmadmb(DOPAC ~ pco2*temp*region + (1|ID), data=NEDOPAC, family="Gamma", link="log")
```

```
summary(M)
stats::step(M)

#Call:
#glmmadmb(formula = DOPAC ~ pco2 * temp * region + (1 | ID), data = NEDOPAC,
#        family = "Gamma", link = "log")
#
# AIC: -490.2
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)               -3.7709     0.4872   -7.74  9.9e-15 ***
#   pco2High                   0.6207     0.6596    0.94     0.35
# temp32                     0.2354     0.6596    0.36     0.72
# regionHB                  -0.0766     0.6596   -0.12     0.91
# regionMB                  -2.7278     0.6596   -4.14  3.5e-05 ***
#   pco2High:temp32           -0.8510     0.9114   -0.93     0.35
# pco2High:regionHB         -0.0585     0.9114   -0.06     0.95
# pco2High:regionMB         -0.3568     0.9114   -0.39     0.70
# temp32:regionHB           -0.3992     0.9114   -0.44     0.66
# temp32:regionMB           -0.7858     0.9114   -0.86     0.39
# pco2High:temp32:regionHB   1.0383     1.2734    0.82     0.41
# pco2High:temp32:regionMB   0.5194     1.2734    0.41     0.68
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=71, ID=24
# Random effect variance(s):
#   Group=ID
# Variance  StdDev
# (Intercept) 1.739e-05 0.00417
#
# Gamma shape parameter: 0.84279 (std. err.: 0.12263)
#
# Log-likelihood: 259.117

M <- glmmadmb(DOPAC ~ pco2*temp*region + (1|ID), data=NEDOPAC2, family="Gamma", link="log")
summary(M)


# Call:
# glmmadmb(formula = DOPAC ~ pco2 * temp * region + (1 | ID), data = NEDOPAC2,
#          family = "Gamma", link = "log")
#
# AIC: -490.2
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)               -6.499      0.445  -14.61  < 2e-16 ***
#   pco2High                   0.264      0.629    0.42     0.67
# temp32                    -0.550      0.629   -0.88     0.38
# regionFB                   2.728      0.660    4.14  3.5e-05 ***
#   regionHB                   2.651      0.629    4.22  2.5e-05 ***
#   pco2High:temp32           -0.332      0.889   -0.37     0.71
# pco2High:regionFB          0.357      0.911    0.39     0.70
# pco2High:regionHB          0.298      0.889    0.34     0.74
# temp32:regionFB            0.786      0.911    0.86     0.39
# temp32:regionHB            0.387      0.889    0.43     0.66
# pco2High:temp32:regionFB  -0.519      1.273   -0.41     0.68
# pco2High:temp32:regionHB   0.519      1.258    0.41     0.68
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=71, ID=24
# Random effect variance(s):
#   Group=ID
# Variance  StdDev
# (Intercept) 1.739e-05 0.00417
#
# Gamma shape parameter: 0.84279 (std. err.: 0.12263)
#
# Log-likelihood: 259.117


# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)
```

```
#Normality
hist(E, main = "Normality", breaks=10)
#Or qq-plot
qqnorm(E)
qqline(E)

#Dependence due to model misfit #shouNE be no patterns, if there are some maybe we miss a covariate etc..
#Plot residuals versus covariates
plot(x = NEDOPAC$temp,
     y = E)
abline(h = 0, lty = 2)
plot(x = NEDOPAC$pco2,
     y = E)
abline(h = 0, lty = 2)
plot(x = NEDOPAC$region,
     y = E)
abline(h = 0, lty = 2)



p <- ggplot()
p <- p + geom_point(data = NEDOPAC,
                    aes(y = E, x = DOPAC),
                    shape = 1,
                    size = 1)
p <- p + xlab("DOPAC") +
  ylab("Residuals")
p <- p + theme(text = element_text(size = 15))
p <- p + facet_grid(temp ~pco2)
p

boxplot(NEDOPAC$DOPAC~NEDOPAC$temp*NEDOPAC$pco2*NEDOPAC$region,
        ylab="DOPAC", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)


### Serotonin ####
plot(density(na.omit(NESER$SER), bw=0.10))


M <- glmmadmb(SER ~ pco2*temp*region + (1|ID), data=NESER, family="Gamma", link="log")
summary(M)
stats::step(M)

M <- glmmadmb(SER ~ pco2 + region + (1|ID), data=NESER, family="Gamma", link="log")
summary(M)

#
# Call:
#   glmmadmb(formula = SER ~ pco2 + region + (1 | ID), data = NESER,
#            family = "Gamma", link = "log")
#
# AIC: 19
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)   0.5071     0.0692    7.33  2.3e-13 ***
#   pco2High     -0.1215     0.0758   -1.60     0.11
# regionHB     -0.3266     0.0765   -4.27  2.0e-05 ***
#   regionMB     -1.2369     0.0761  -16.25  < 2e-16 ***
#   ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=72, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)  0.01135 0.1065
#
# Gamma shape parameter: 14.566 (std. err.: 2.9617)
#
# Log-likelihood: -3.47729

M <- glmmadmb(SER ~ pco2+region + (1|ID), data=NESER2, family="Gamma", link="log")
summary(M)

# Call:
# glmmadmb(formula = SER ~ pco2 + region + (1 | ID), data = NESER2,
#          family = "Gamma", link = "log")
#
# AIC: 19
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)  -0.7299     0.0701   -10.4   <2e-16 ***
#   pco2High     -0.1215     0.0758    -1.6     0.11
# regionFB      1.2369     0.0761    16.2   <2e-16 ***
#   regionHB      0.9104     0.0761    12.0   <2e-16 ***
#   ---
```

```
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=72, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)  0.01135 0.1065
#
# Gamma shape parameter: 14.566 (std. err.: 2.9617)
#
# Log-likelihood: -3.47729



# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)



#Normality
hist(E, main = "Normality", breaks=10)
#Or qq-plot
qqnorm(E)
qqline(E)

#Dependence due to model misfit #shouNE be no patterns, if there are some maybe we miss a covariate etc..
#Plot residuals versus covariates
plot(x = NESER$temp,
     y = E)
abline(h = 0, lty = 2)
plot(x = NESER$pco2,
     y = E)
abline(h = 0, lty = 2)
plot(x = NESER$region,
     y = E)
abline(h = 0, lty = 2)



p <- ggplot()
p <- p + geom_point(data = NESER,
                    aes(y = E, x = SER),
                    shape = 1,
                    size = 1)
p <- p + xlab("Ser") +
  ylab("Residuals")
p <- p + theme(text = element_text(size = 15))
p <- p + facet_grid(temp ~pco2)
p

boxplot(NESER$SER~NESER$temp*NESER$pco2*NESER$region,
        ylab="Seratonin", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)



### HIAA ####
plot(density(NEHIAA$HIAA))

M <- glmmadmb(HIAA ~ pco2 *temp * region (1|ID), data=NEHIAA, family="Gamma", link="inverse")
stats::step(M)

M <- glmmadmb(HIAA ~ pco2 + temp + region + pco2:region+ (1|ID), data=NEHIAA, family="Gamma", link="inverse")
summary(M)
stats::step(M)

# Call:
#   glmmadmb(formula = HIAA ~ pco2 + temp + region + pco2:region +
#              (1 | ID), data = NEHIAA, family = "Gamma", link = "inverse")
#
# AIC: -147.2
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)         2.9262     0.3965    7.38  1.6e-13 ***
#   pco2High           -0.0213     0.4330   -0.05    0.961
# temp32             -0.6125     0.4055   -1.51    0.131
# regionHB            2.6598     0.5784    4.60  4.3e-06 ***
#   regionMB            4.4418     0.7501    5.92  3.2e-09 ***
```

```
#   pco2High:regionHB   2.3119     0.9817    2.36    0.019 *
#   pco2High:regionMB  -0.1820     1.0408   -0.17    0.861
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=72, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)   0.3955 0.6289
#
# Gamma shape parameter: 8.0942 (std. err.: 1.6384)
#
# Log-likelihood: 82.577
M <- glmmadmb(HIAA ~ pco2 + temp + region + pco2:region+(1|ID), data=NEHIAA2, family="Gamma", link="inverse")
summary(M)
#
# Call:
#   glmmadmb(formula = HIAA ~ pco2 + temp + region + pco2:region +
#              (1 | ID), data = NEHIAA2, family = "Gamma", link = "inverse")
#
# AIC: -147.2
#
# Coefficients:
#   Estimate Std. Error z value Pr(>|z|)
# (Intercept)          7.368      0.762    9.67  < 2e-16 ***
#   pco2High            -0.203      1.017   -0.20    0.842
# temp32              -0.612      0.405   -1.51    0.131
# regionFB            -4.442      0.750   -5.92  3.2e-09 ***
#   regionHB            -1.782      0.880   -2.03    0.043 *
#   pco2High:regionFB    0.182      1.041    0.17    0.861
# pco2High:regionHB    2.494      1.347    1.85    0.064 .
# ---
#   Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Number of observations: total=72, ID=24
# Random effect variance(s):
#   Group=ID
# Variance StdDev
# (Intercept)   0.3955 0.6289
#
# Gamma shape parameter: 8.0942 (std. err.: 1.6384)
#
# Log-likelihood: 82.577

# Homogeneity
par(mfrow =c(1,1))
E <- resid(M)
Fx <- fitted(M)

plot(x = Fx,
     y = E,
     xlab = "Fitted values",
     ylab = "Residuals",
     main = "Homogeneity?")
abline(h = 0, v = 0, lty = 2)



#Normality
hist(E, main = "Normality", breaks=10)
#Or qq-plot
qqnorm(E)
qqline(E)

#Dependence due to model misfit #shouNE be no patterns, if there are some maybe we miss a covariate etc..
#Plot residuals versus covariates
plot(x = NEHIAA$temp,
     y = E)
abline(h = 0, lty = 2)
plot(x = NEHIAA$pco2,
     y = E)
abline(h = 0, lty = 2)
plot(x = NEHIAA$region,
     y = E)
abline(h = 0, lty = 2)



p <- ggplot()
p <- p + geom_point(data = NEHIAA,
                    aes(y = E, x = HIAA),
                    shape = 1,
                    size = 1)
p <- p + xlab("HIAA") +
  ylab("Residuals")
p <- p + theme(text = element_text(size = 15))
p <- p + facet_grid(temp ~pco2)
```

```
p

boxplot(NEHIAA$HIAA~NEHIAA$temp*NEHIAA$pco2*NEHIAA$region,
        ylab="RAtio Serotonin", cex.axis=0.8, cex.lab=0.8, cex.main=0.8)

######## CCA: Canonical Correlation Analysis #####

library(pgirmess)
library(car)

library(ggplot2)
library(knitr)
library(pander)
library(tm)
library(SnowballC)
library(RColorBrewer)
library(wordcloud)
library(biclust)
library(cluster)
library(igraph)
library(extrafont)
library(reshape2)
library(factoextra)
library(psych)
library(FactoMineR)
library(Hmisc)
library(plotrix)
library(gridExtra)
library(CCA)
library(scales)
fancy_scientific <- function(l) {
  # turn in to character string in scientific notation
  l <- format(l, scientific = TRUE)
  # quote the part before the exponent to keep all the digits
  l <- gsub("^(.*)e", "'\\1'e", l)
  # turn the 'e+' into plotmath format
  l <- gsub("e", "%*%10^", l)
  # return this as an expression
  parse(text=l)
}

crossloalplot<-function(res.rcc,dimx,dimy){

  par(mgp=c(2.5,1,0),mar=c(4,4,2,2))
  par(oma=c(0,0,0,0))
  par(mfrow=c(1,2))
  par(pty="s")
  x1<-res.rcc$scores$corr.Y.xscores[,dimx]
[abs(res.rcc$scores$corr.Y.xscores[,dimx])>=0.5|abs(res.rcc$scores$corr.Y.xscores[,dimy])>=0.5]
  y1<-res.rcc$scores$corr.Y.xscores[,dimy]
[abs(res.rcc$scores$corr.Y.xscores[,dimx])>=0.5|abs(res.rcc$scores$corr.Y.xscores[,dimy])>=0.5]
  x2<-res.rcc$scores$corr.Y.xscores[,dimx][-c(which(res.rcc$scores$corr.Y.xscores[,dimx] %in% x1))]
  y2<-res.rcc$scores$corr.Y.xscores[,dimy][-c(which(res.rcc$scores$corr.Y.xscores[,dimy] %in% y1))]
  plot(x1,y1,col=4,xlim=c(-1.2,1.2),ylim=c(-1.2,1.2),pch=16,ylab=paste("Canonical Variate",
dimy),xlab=paste("Canonical Variate", dimx),type="n",main=paste("Canonical Variate", dimx))
  rect(xleft = -0.5,xright = 0.5,ybottom = -1.5,ytop = 1.5,col="grey",density=20)
  points(x1,y1,col=4,xlim=c(-1.2,1.2),ylim=c(-1.2,1.2),pch=16,type="n")
  points(x2,y2,col="grey",xlim=c(-1.2,1.2),ylim=c(-1.2,1.2),pch=16,ylab=paste("Canonical Variate",
dimy),xlab=paste("Canonical Variate", dimx),type="n")
  text(x1,y1,labels=names(x1),cex=0.6,col=4)
  text(x2,y2,labels=names(x2),cex=0.6,col="grey")
  abline(h=0.5,v=0.5,lty=2)
  abline(h=-0.5,v=-0.5,lty=2)

points(res.rcc$scores$corr.X.yscores[,dimx],res.rcc$scores$corr.X.yscores[,dimy],col=2,xlim=c(-1,1),ylim=c(-1,1),p


text(res.rcc$scores$corr.X.yscores[,dimx],res.rcc$scores$corr.X.yscores[,dimy],labels=rownames(res.rcc$xcoef),cex=

  plot(x1,y1,col=4,xlim=c(-1.2,1.2),ylim=c(-1.2,1.2),pch=16,ylab=paste("Canonical Variate",
dimy),xlab=paste("Canonical Variate", dimx),type="n",main=paste("Canonical Variate", dimy))
  rect(xleft = -1.5,xright = 1.5,ybottom = -0.5,ytop = 0.5,col="grey",density=20)
  points(x1,y1,col=4,xlim=c(-1,1),ylim=c(-1,1),pch=16,type="n")
  points(x2,y2,col="grey",xlim=c(-1,1),ylim=c(-1,1),pch=16,ylab=paste("Canonical Variate",
dimy),xlab=paste("Canonical Variate", dimx),type="n")
  text(x1,y1,labels=names(x1),cex=0.6,col=4)
  text(x2,y2,labels=names(x2),cex=0.6,col="grey")
  abline(h=0.5,v=0.5,lty=2)
  abline(h=-0.5,v=-0.5,lty=2)

points(res.rcc$scores$corr.X.yscores[,dimx],res.rcc$scores$corr.X.yscores[,dimy],col=2,xlim=c(-1,1),ylim=c(-1,1),p


text(res.rcc$scores$corr.X.yscores[,dimx],res.rcc$scores$corr.X.yscores[,dimy],labels=rownames(res.rcc$xcoef),cex=


}
```

```
crossloalplot_i<-function(res.rcc,dimx,dimy){

  par(mgp=c(2.5,1,0),mar=c(4,4,2,2))
  par(oma=c(0,0,0,0))
  par(mfrow=c(1,2))
  par(pty="s")

  x1<-res.rcc$scores$xscores[,dimx]
  y1<-res.rcc$scores$xscores[,dimy]
  x2<-res.rcc$scores$yscores[,dimx]
  y2<-res.rcc$scores$yscores[,dimy]

  max.x=max(max(x1),max(x2))
  min.x=min(min(x1),min(x2))

  max.y=max(max(y1),max(y2))
  min.y=min(min(y1),min(y2))
  plot(x1,y1,ylab=paste("Canonical Variate", dimy),xlab=paste("Canonical Variate", dimx),main="Independent
Variates",type="n",ylim=c(min.y-0.2,max.y+0.2),xlim=c(min.x-0.2,max.x+0.2),cex.main=0.8)
  text(x1,y1,labels=paste0("Client",c(1:2,4:12)),cex=0.6,font = 2)
  abline(h=0,v=0,lty=2)
  plot(x2,y2,ylab=paste("Canonical Variate", dimy),xlab=paste("Canonical Variate", dimx),main="Dependent
Variates",type="n",ylim=c(min.y-0.2,max.y+0.2),xlim=c(min.x-0.2,max.x+0.2),cex.main=0.8)
  text(x2,y2,labels=paste0("Cleaner",c(1:2,4:12)),cex=0.6,font = 2)
  abline(h=0,v=0,lty=2)
}
##Cleaners####

cleaner1<-data[,c(2:4,9:22,35,38)]# Cleaner fish behavioural and neurobiological data
rownames(cleaner1)<-paste0("Cleaner",rownames(cleaner1))
cleaner1<-cleaner1[complete.cases(cleaner1),]

x<-as.matrix(scale(cleaner1[,c(3:5,18:19)]))#Behaviour variables
y<-as.matrix(scale(cleaner1[,-c(1:5,18:19)]))#Neurobiological Variables

res.reg<-CCA::estim.regul(x,y)

res.rcc<-CCA::rcc(x,y,lambda1 = res.reg$lambda1,lambda2 = res.reg$lambda2)



bp<-barplot(res.rcc$cor,xlab="Canonical functions",ylab="Canonical correlations",names.arg =
paste0("CF",1:min(ncol(x),ncol(y))),ylim=c(0,1))
text(bp,res.rcc$cor-0.05,labels = paste(round(res.rcc$cor,2)))

#enfc: number of canonical functions
nfc<-min(dim(x)[2],dim(y)[2])

cancor<-data.frame(`Canonical Function`=1:nfc,`Canonical Correlation`=res.rcc$cor[1:nfc],`Canonical
R2`=res.rcc$cor[1:nfc]^2)

logical<-cancor$Canonical.Correlation>0.7

result.r<-character()
for(i in 1:length(logical)){
  result.r[i]<-ifelse(logical[i]=="TRUE",print(paste0("CF",i,":
r=",round(cancor$Canonical.Correlation[i],2))),print(paste0("CF",i,": r<0.70")))
}

kable(cancor,digits = 2,caption = "Canonical correlations")
res.rcc$names$ind.names

#Redundancy index
# Dependent canonical variables

dcv<-seq(1,nfc*2,2)

for(i in 1:nfc){

assign(paste0("RedunIndex_r",dcv[i]),data.frame(Função=i,Variables="Dependent",Average.Loading=sum(res.rcc$scores$
(sum(res.rcc$scores$corr.Y.yscores[,i]^2)/length(res.rcc$scores$corr.Y.yscores[,i]))*(res.rcc$cor[i]^2)))
}

# Independent canonical variables

icv<-seq(2,nfc*2+1,2)

for(i in 1:nfc){

assign(paste0("RedunIndex_r",icv[i]),data.frame(Função=i,Variables="Independent",Average.Loading=sum(res.rcc$scores
(sum(res.rcc$scores$corr.X.xscores[,i]^2)/length(res.rcc$scores$corr.X.xscores[,i]))*(res.rcc$cor[i]^2)))}

RI_index<-c(dcv,icv)
RI<-data.frame()
for(i in RI_index){
  stepRI<-get(paste0("RedunIndex_r",i))
  RI<-rbind(RI,stepRI)
}
```

```
#independents
Var.indep<-data.frame(`Canonical Function`=1:nfc,Shared.Var=RI$Average.Loading[RI$Variables=="Independent"])


Var.indep$CumVar<-cumsum(Var.indep$Shared.Var)
Var.indep$CanR2<-res.rcc$cor[1:nfc]^2
Var.indep$RI<-Var.indep$Shared.Var*Var.indep$CanR2
Var.indep$CumRI<-cumsum(Var.indep$RI)

#dependents
Var.dep<-data.frame(`Canonical.Function`=1:nfc,Shared.Var=RI$Average.Loading[RI$Variables=="Dependent"])
Var.dep$CumVar<-cumsum(Var.dep$Shared.Var)
Var.dep$CanR2<-res.rcc$cor[1:nfc]^2
Var.dep$RI<-Var.dep$Shared.Var*Var.dep$CanR2
Var.dep$CumRI<-cumsum(Var.dep$RI)

kable(Var.indep[,-4],digits = 3,caption = "Variance of the behavioral variables explained by the themselves
(Shared.Var and CumVar) and neurobiological variables (RI and CumRI)\\label{tab:t30}")
kable(Var.dep[,-4],digits = 3,caption = "Variance of neurotransmitters explained by themselves (Shared.Var and
CumVar) and by behavioural variables (RI e CumRI)\\label{tab:t31}")
kable(res.rcc$xcoef[,1:nfc],digits = 3,caption = "Standarized canonical coefficients (_weights_) for independent
variables (_behaviour_)\\label{tab:t32}",col.names = paste("CV",1:nfc))

kable(res.rcc$ycoef[rev(order(abs(res.rcc$ycoef[,1]))),1:nfc],digits = 3,caption = "Standarized canonical
coefficients (_weights_) for dependent variables (Neurotransmitters)\\label{tab:t33}",col.names =
paste("DV",1:nfc))

cl_ind<-res.rcc$scores$corr.X.xscores[,1:nfc]
cl_dep<-res.rcc$scores$corr.Y.yscores[,1:nfc]

kable(cl_ind,digits = 3,caption = "Canonical Loadings for independent variables
(Behaviour)\\label{tab:t34}",col.names = paste("IV",1:nfc))

kable(cl_dep,digits = 3,caption = "Canonical Loadings for dependent variables
(Neurotransmitters)\\label{tab:t35}",col.names = paste("DV",1:nfc))

ccl_ind<-res.rcc$scores$corr.X.yscores[,1:nfc]
kable(ccl_ind,digits = 3,caption = "Canonical crossloadings for independent variables
(_Behaviour_)\\label{tab:t36}",col.names = paste("CF",1:nfc))
ccl_dep<-res.rcc$scores$corr.Y.xscores[,1:nfc]
kable(ccl_dep,digits = 3,caption = "Canonical crossloadings for dependent variables
(Neurotransmitters)\\label{tab:t37}",col.names = paste("CF",1:nfc))

crossloalplot(res.rcc,dimx = 1,dimy = 2)
varY.x<-ccl_ind^2*100
kable(varY.x,digits = 1,caption = "Variability of behaviour explained by the dependent canonical variables
(Neurotransmitters)\\label{tab:t38}",col.names = paste("CV",1:nfc))

varX.y<-ccl_dep^2*100
kable(varX.y,digits = 1,caption = "Variability of neurotransmitters explained by the independent canonical
variables (Behaviour)\\label{tab:t39}",col.names = paste("CV",1:nfc))


#FC1
##B
datavarYx<-data.frame(r2=varY.x[,1])
datavarYx$a<-rownames(datavarYx)
datavarYx$a<-factor(datavarYx$a,levels = rownames(datavarYx))
#
# a50.1<-datavarYx[c(which(datavarYx[,1]>=50)),1]
#
plot1<-ggplot(data = datavarYx,aes(x = a,y = r2,group=1))+
  geom_bar(stat = "identity",width = 0.5, fill= "burlywood") +
  scale_y_continuous("Var. explained by neurotransmitters (%)",limits = c(0,100))+
  scale_x_discrete(" ")+
  theme(text =
element_text(size=7),axis.text.x=element_text(size=6),axis.title.y=element_text(margin=margin(0,8,0,0),size=7))+
  geom_hline(yintercept = 50,colour=2,linetype=2)+
  ggtitle("Canonical Function 1")
#
# ##N
datavarXy1<-data.frame(r2=varX.y[,1])
datavarXy1$N<-rownames(datavarXy1)
datavarXy1$N<-factor(datavarXy1$N,levels = rownames(datavarXy1))
#
# N50.1<-datavarXy1[c(which(datavarXy1[,1]>=50)),1]
#
plot2<-ggplot(data = datavarXy1,aes(x = N,y = r2,group=1))+
  geom_bar(stat = "identity",width = 0.5, fill= "grey54") +
  scale_y_continuous("Var. explained by behaviour (%)",limits = c(0,100))+
  scale_x_discrete(" ")+
  theme(text =
element_text(size=7),axis.text.x=element_text(vjust=0.5,size=5),axis.title.y=element_text(margin=margin(0,8,0,0),s

  geom_hline(yintercept = 50,colour=2,linetype=2)
# #
```

```
grid.arrange(plot1,plot2)

#FC2

indiceFC<-2

##B
datavarYx<-data.frame(r2=varY.x[,indiceFC])
datavarYx$a<-rownames(datavarYx)
datavarYx$a<-factor(datavarYx$a,levels = rownames(datavarYx))
#
# a50.1<-datavarYx[c(which(datavarYx[,indiceFC]>=50)),1]
#
plot1<-ggplot(data = datavarYx,aes(x = a,y = r2,group=1))+
  geom_bar(stat = "identity",width = 0.5, fill= "burlywood") +
  scale_y_continuous("Var. explained by neurotransmitters (%)",limits = c(0,100))+
  scale_x_discrete(" ")+
  theme(text =
element_text(size=7),axis.text.x=element_text(size=6),axis.title.y=element_text(margin=margin(0,8,0,0),size=7))+
  geom_hline(yintercept = 50,colour=2,linetype=2)+
  ggtitle("Canonical Function 2")
#
# ##N
datavarXy1<-data.frame(r2=varX.y[,indiceFC])
datavarXy1$N<-rownames(datavarXy1)
datavarXy1$N<-factor(datavarXy1$N,levels = rownames(datavarXy1))
#
# N50.1<-datavarXy1[c(which(datavarXy1[,indiceFC]>=50)),1]
#
plot2<-ggplot(data = datavarXy1,aes(x = N,y = r2,group=1))+
  geom_bar(stat = "identity",width = 0.5, fill= "grey54") +
  scale_y_continuous("Var. explained by behaviour (%)",limits = c(0,100))+
  scale_x_discrete(" ")+
  theme(text =
element_text(size=7),axis.text.x=element_text(vjust=0.5,size=5),axis.title.y=element_text(margin=margin(0,8,0,0),s

  geom_hline(yintercept = 50,colour=2,linetype=2)
# #
# #

grid.arrange(plot1,plot2)



Brcc1<-res.rcc$scores$xscores[,1]
Nrcc1<-res.rcc$scores$yscores[,1]

Brcc2<-res.rcc$scores$xscores[,2]
Nrcc2<-res.rcc$scores$yscores[,2]

Brcc3<-res.rcc$scores$xscores[,3]
Nrcc3<-res.rcc$scores$yscores[,3]

#1st Canonical function

layout(matrix(c(1,2,3,3),2,2,byrow = T),widths = c(2,2),heights = c(2,2),respect = TRUE)


par(las=1)
par(mgp=c(2.5,1,0))
par(mar=c(4,5,2,2),oma=c(0,0,0,0))
x1<-res.rcc$scores$xscores[,1]
x1<-data.frame(x=x1,col=as.numeric(cleaner1$pco2))
x1<-x1[order(x1$x),]

b1<-barplot(x1$x,horiz = TRUE,xlim=c(-3,3), xlab = "Scores Canonical Function
1",main="Behaviour",cex.main=0.8,cex.lab=0.8,cex.names = 0.8,cex.axis = 0.8,names.arg = rownames(x1),space = 0.5)
abline(h=b1[which(x1$x>0)[1]]-1.2/2,lty=2)




y1<-res.rcc$scores$yscores[,1]
y1<-data.frame(y=y1,col=as.numeric(cleaner1$pco2))
y1<-y1[order(y1$y),]
barplot(y1$y,horiz = TRUE,xlim=c(-3,3), xlab = "Scores Canonical Function 1", main="Neurotransmitters",cex.names
= 0.8,cex.main=0.8,cex.lab=0.8,cex.axis = 0.8,names.arg = rownames(y1),space = 0.5)
abline(h=b1[which(y1$y>0)[1]]-1.2/2,lty=2)


par(pty="s")
plot(Brcc1,Nrcc1,pch=16,col=1,xlab="Scores CF1 Behaviour",ylab="Scores CF1
Neurotransmitters",xlim=c(-3,3),ylim=c(-3,3),cex.lab=0.8,cex.axis=0.8)
text(Brcc1,Nrcc1,labels=res.rcc$names$ind.names,pos=c(2,4),cex=0.8,col=1)
abline(a = 0,b = 1,col=1)

#2nd Canonical function
```

```
layout(matrix(c(1,2,3,3),2,2,byrow = T),widths = c(2,2),heights = c(2,2),respect = TRUE)


par(las=1)
par(mgp=c(2.5,1,0))
par(mar=c(4,5,2,2),oma=c(0,0,0,0))
x2<-res.rcc$scores$xscores[,2]
x2<-data.frame(x=x2,col=as.numeric(cleaner1$pco2))
x2<-x2[order(x2$x),]

b2<-barplot(x2$x,horiz = TRUE,xlim=c(-3,3), xlab = "Scores Canonical Function 2",names.arg =
rownames(x2),main="Behaviour",cex.names = 0.8,cex.main=0.8,cex.lab=0.8,cex.axis = 0.8,space = 0.5)
abline(h=b2[which(x2$x>0)[1]]-1.2/2,lty=2)

y2<-res.rcc$scores$yscores[,2]
y2<-data.frame(y=y2,col=as.numeric(cleaner1$pco2))
y2<-y2[order(y2$y),]

barplot(y2$y,horiz = TRUE,xlim=c(-3,3), xlab = "Scores Canonical Function 2",names.arg = rownames(y2),
main="Neurotransmitters",cex.names = 0.8,cex.main=0.8,cex.lab=0.8,cex.axis = 0.8,space = 0.5)
abline(h=b2[which(y2$y>0)[1]]-1.2/2,lty=2)



par(pty="s")
plot(Brcc2,Nrcc2,pch=16,col=1,xlab="Scores CF2 Behaviour",ylab="Scores CF2
neurotransmitters",xlim=c(-3,3),ylim=c(-3,3))
text(Brcc2,Nrcc2,labels=res.rcc$names$ind.names,pos=c(4,2),cex=0.8,col=1)
abline(a = 0,b = 1,col=1)

#References Jolliffe, I. T. (2002). Principal component analysis. Springer, New York


##Clients####
par(mfrow=c(1,1))
Client1<-data[,c(2:4,7:10,23:34,38)]# Client behavioural and neurobiological data
rownames(Client1)<-paste0("Client",rownames(Client1))
Client1<-Client1[complete.cases(Client1),]

x<-as.matrix(scale(Client1[,c(3:7,20)]))#Behaviour variables
y<-as.matrix(scale(Client1[,-c(1:7,20)]))#Neurobiological Variables

res.reg<-CCA::estim.regul(x,y)

res.rcc<-CCA::rcc(x,y,lambda1 = res.reg$lambda1,lambda2 = res.reg$lambda2)

bp<-barplot(res.rcc$cor,xlab="Canonical functions",ylab="Canonical correlations",names.arg =
paste0("CF",1:min(ncol(x),ncol(y))),ylim=c(0,1))
text(bp,res.rcc$cor-0.05,labels = paste(round(res.rcc$cor,2)))

#enfc: number of canonical functions
nfc<-min(dim(x)[2],dim(y)[2])

cancor<-data.frame(`Canonical Function`=1:nfc,`Canonical Correlation`=res.rcc$cor[1:nfc],`Canonical
R2`=res.rcc$cor[1:nfc]^2)

logical<-cancor$Canonical.Correlation>0.7

result.r<-character()
for(i in 1:length(logical)){
  result.r[i]<-ifelse(logical[i]=="TRUE",print(paste0("CF",i,":
r=",round(cancor$Canonical.Correlation[i],2))),print(paste0("CF",i,": r<0.70")))
}

kable(cancor,digits = 2,caption = "Canonical correlations")
res.rcc$names$ind.names

#Redundancy index
# Dependent canonical variables

dcv<-seq(1,nfc*2,2)

for(i in 1:nfc){

assign(paste0("RedunIndex_r",dcv[i]),data.frame(Função=i,Variables="Dependent",Average.Loading=sum(res.rcc$scores$
(sum(res.rcc$scores$corr.Y.yscores[,i]^2)/length(res.rcc$scores$corr.Y.yscores[,i]))*(res.rcc$cor[i]^2)))
}

# Independent canonical variables

icv<-seq(2,nfc*2+1,2)

for(i in 1:nfc){

assign(paste0("RedunIndex_r",icv[i]),data.frame(Função=i,Variables="Independent",Average.Loading=sum(res.rcc$scores
(sum(res.rcc$scores$corr.X.xscores[,i]^2)/length(res.rcc$scores$corr.X.xscores[,i]))*(res.rcc$cor[i]^2)))}

RI_index<-c(dcv,icv)
```

```
RI<-data.frame()
for(i in RI_index){
  stepRI<-get(paste0("RedunIndex_r",i))
  RI<-rbind(RI,stepRI)
}

#independents
Var.indep<-data.frame(`Canonical Function`=1:nfc,Shared.Var=RI$Average.Loading[RI$Variables=="Independent"])


Var.indep$CumVar<-cumsum(Var.indep$Shared.Var)
Var.indep$CanR2<-res.rcc$cor[1:nfc]^2
Var.indep$RI<-Var.indep$Shared.Var*Var.indep$CanR2
Var.indep$CumRI<-cumsum(Var.indep$RI)

#dependents
Var.dep<-data.frame(`Canonical.Function`=1:nfc,Shared.Var=RI$Average.Loading[RI$Variables=="Dependent"])
Var.dep$CumVar<-cumsum(Var.dep$Shared.Var)
Var.dep$CanR2<-res.rcc$cor[1:nfc]^2
Var.dep$RI<-Var.dep$Shared.Var*Var.dep$CanR2
Var.dep$CumRI<-cumsum(Var.dep$RI)

kable(Var.indep[,-4],digits = 3,caption = "Variance of the behavioral variables explained by the themselves
(Shared.Var and CumVar) and neurobiological variables (RI and CumRI)\\label{tab:t30}")
kable(Var.dep[,-4],digits = 3,caption = "Variance of neurotransmitters explained by themselves (Shared.Var and
CumVar) and by behavioural variables (RI e CumRI)\\label{tab:t31}")
kable(res.rcc$xcoef[,1:nfc],digits = 3,caption = "Standarized canonical coefficients (_weights_) for independent
variables (_behaviour_)\\label{tab:t32}",col.names = paste("CV",1:nfc))

kable(res.rcc$ycoef[rev(order(abs(res.rcc$ycoef[,1]))),1:nfc],digits = 3,caption = "Coeficientes canónicos
standartizados (_weights_) para as variáveis dependentes (Neurotransmitters)\\label{tab:t33}",col.names =
paste("DV",1:nfc))

cl_ind<-res.rcc$scores$corr.X.xscores[,1:nfc]
cl_dep<-res.rcc$scores$corr.Y.yscores[,1:nfc]

kable(cl_ind,digits = 3,caption = "Canonical Loadings for independent variables
(Behaviour)\\label{tab:t34}",col.names = paste("IV",1:nfc))

kable(cl_dep,digits = 3,caption = "Canonical Loadings for dependent variables
(Neurotransmitters)\\label{tab:t35}",col.names = paste("DV",1:nfc))

ccl_ind<-res.rcc$scores$corr.X.yscores[,1:nfc]
kable(ccl_ind,digits = 3,caption = "Canonical crossloadings for independent variables
(_Behaviour_)\\label{tab:t36}",col.names = paste("CF",1:nfc))
ccl_dep<-res.rcc$scores$corr.Y.xscores[,1:nfc]
kable(ccl_dep,digits = 3,caption = "Canonical crossloadings for dependent variables
(Neurotransmitters)\\label{tab:t37}",col.names = paste("CF",1:nfc))

crossloalplot(res.rcc,dimx = 1,dimy = 2)
varY.x<-ccl_ind^2*100
kable(varY.x,digits = 1,caption = "Variability of behaviour explained by the dependent canonical variables
(Neurotransmitters)\\label{tab:t38}",col.names = paste("CV",1:nfc))

varX.y<-ccl_dep^2*100
kable(varX.y,digits = 1,caption = "Variability of neurotransmitters explained by the independent canonical
variables (Behaviour)\\label{tab:t39}",col.names = paste("CV",1:nfc))


#FC1
##B
datavarYx<-data.frame(r2=varY.x[,1])
datavarYx$a<-rownames(datavarYx)
datavarYx$a<-factor(datavarYx$a,levels = rownames(datavarYx))
#
# a50.1<-datavarYx[c(which(datavarYx[,1]>=50)),1]
#
plot1<-ggplot(data = datavarYx,aes(x = a,y = r2,group=1))+
  geom_bar(stat = "identity",width = 0.5, fill= "burlywood") +
  scale_y_continuous("Var. explained by neurotransmitters (%)",limits = c(0,100))+
  scale_x_discrete(" ")+
  theme(text =
element_text(size=7),axis.text.x=element_text(size=6),axis.title.y=element_text(margin=margin(0,8,0,0),size=7))+
  geom_hline(yintercept = 50,colour=2,linetype=2)+
  ggtitle("Canonical Function 1")
#
# ##N
datavarXy1<-data.frame(r2=varX.y[,1])
datavarXy1$N<-rownames(datavarXy1)
datavarXy1$N<-factor(datavarXy1$N,levels = rownames(datavarXy1))
#
# N50.1<-datavarXy1[c(which(datavarXy1[,1]>=50)),1]
#
plot2<-ggplot(data = datavarXy1,aes(x = N,y = r2,group=1))+
  geom_bar(stat = "identity",width = 0.5, fill= "grey54") +
  scale_y_continuous("Var. explained by behaviour (%)",limits = c(0,100))+
  scale_x_discrete(" ")+
  theme(text =
```

```
element_text(size=7),axis.text.x=element_text(vjust=0.5,size=5),axis.title.y=element_text(margin=margin(0,8,0,0),s:

  geom_hline(yintercept = 50,colour=2,linetype=2)
# #

grid.arrange(plot1,plot2)

#FC2

indiceFC<-2

##B
datavarYx<-data.frame(r2=varY.x[,indiceFC])
datavarYx$a<-rownames(datavarYx)
datavarYx$a<-factor(datavarYx$a,levels = rownames(datavarYx))
#
# a50.1<-datavarYx[c(which(datavarYx[,indiceFC]>=50)),1]
#
plot1<-ggplot(data = datavarYx,aes(x = a,y = r2,group=1))+
  geom_bar(stat = "identity",width = 0.5, fill= "burlywood") +
  scale_y_continuous("Var. explained by neurotransmitters (%)",limits = c(0,100))+
  scale_x_discrete(" ")+
  theme(text =
element_text(size=7),axis.text.x=element_text(size=6),axis.title.y=element_text(margin=margin(0,8,0,0),size=7))+
  geom_hline(yintercept = 50,colour=2,linetype=2)+
  ggtitle("Canonical Function 2")
#
# ##N
datavarXy1<-data.frame(r2=varX.y[,indiceFC])
datavarXy1$N<-rownames(datavarXy1)
datavarXy1$N<-factor(datavarXy1$N,levels = rownames(datavarXy1))
#
# N50.1<-datavarXy1[c(which(datavarXy1[,indiceFC]>=50)),1]
#
plot2<-ggplot(data = datavarXy1,aes(x = N,y = r2,group=1))+
  geom_bar(stat = "identity",width = 0.5, fill= "grey54") +
  scale_y_continuous("Var. explained by behaviour (%)",limits = c(0,100))+
  scale_x_discrete(" ")+
  theme(text =
element_text(size=7),axis.text.x=element_text(vjust=0.5,size=5),axis.title.y=element_text(margin=margin(0,8,0,0),s:

  geom_hline(yintercept = 50,colour=2,linetype=2)
# #
# #

grid.arrange(plot1,plot2)



Brcc1<-res.rcc$scores$xscores[,1]
Nrcc1<-res.rcc$scores$yscores[,1]

Brcc2<-res.rcc$scores$xscores[,2]
Nrcc2<-res.rcc$scores$yscores[,2]

Brcc3<-res.rcc$scores$xscores[,3]
Nrcc3<-res.rcc$scores$yscores[,3]

#1st Canonical function

layout(matrix(c(1,2,3,3),2,2,byrow = T),widths = c(2,2),heights = c(2,2),respect = TRUE)


par(las=1)
par(mgp=c(2.5,1,0))
par(mar=c(4,5,2,2),oma=c(0,0,0,0))
x1<-res.rcc$scores$xscores[,1]
x1<-data.frame(x=x1,col=as.numeric(Client1$pco2))
x1<-x1[order(x1$x),]

b1<-barplot(x1$x,horiz = TRUE,xlim=c(-3,3), xlab = "Scores Canonical Function
1",main="Behaviour",cex.main=0.8,cex.lab=0.8,cex.names = 0.8,cex.axis = 0.8,names.arg = rownames(x1),space = 0.5)
abline(h=b1[which(x1$x>0)[1]]-1.2/2,lty=2)




y1<-res.rcc$scores$yscores[,1]
y1<-data.frame(y=y1,col=as.numeric(Client1$pco2))
y1<-y1[order(y1$y),]
barplot(y1$y,horiz = TRUE,xlim=c(-3,3), xlab = "Scores Canonical Function 1", main="Neurotransmitters",cex.names
= 0.8,cex.main=0.8,cex.lab=0.8,cex.axis = 0.8,names.arg = rownames(y1),space = 0.5)
abline(h=b1[which(y1$y>0)[1]]-1.2/2,lty=2)


par(pty="s")
plot(Brcc1,Nrcc1,pch=16,col=1,xlab="Scores CF1 Behaviour",ylab="Scores CF1
Neurotransmitters",xlim=c(-3,3),ylim=c(-3,3),cex.lab=0.8,cex.axis=0.8)
```

```
text(Brcc1,Nrcc1,labels=res.rcc$names$ind.names,pos=c(2,4),cex=0.8,col=1)
abline(a = 0,b = 1,col=1)

#2nd Canonical function

layout(matrix(c(1,2,3,3),2,2,byrow = T),widths = c(2,2),heights = c(2,2),respect = TRUE)


par(las=1)
par(mgp=c(2.5,1,0))
par(mar=c(4,5,2,2),oma=c(0,0,0,0))
x2<-res.rcc$scores$xscores[,2]
x2<-data.frame(x=x2,col=as.numeric(Client1$pco2))
x2<-x2[order(x2$x),]

b2<-barplot(x2$x,horiz = TRUE,xlim=c(-3,3), xlab = "Scores Canonical Function 2",names.arg =
rownames(x2),main="Behaviour",cex.names = 0.8,cex.main=0.8,cex.lab=0.8,cex.axis = 0.8,space = 0.5)
abline(h=b2[which(x2$x>0)[1]]-1.2/2,lty=2)

y2<-res.rcc$scores$yscores[,2]
y2<-data.frame(y=y2,col=as.numeric(Client1$pco2))
y2<-y2[order(y2$y),]

barplot(y2$y,horiz = TRUE,xlim=c(-3,3), xlab = "Scores Canonical Function 2",names.arg = rownames(y2),
main="Neurotransmitters",cex.names = 0.8,cex.main=0.8,cex.lab=0.8,cex.axis = 0.8,space = 0.5)
abline(h=b2[which(y2$y>0)[1]]-1.2/2,lty=2)



par(pty="s")
plot(Brcc2,Nrcc2,pch=16,col=1,xlab="Scores CF2 Behaviour",ylab="Scores CF2
neurotransmitters",xlim=c(-3,3),ylim=c(-3,3))
text(Brcc2,Nrcc2,labels=res.rcc$names$ind.names,pos=c(4,2),cex=0.8,col=1)
abline(a = 0,b = 1,col=1)
```