Humpback Whales and Ship Noise

Fabian Blasch

04/30/2022

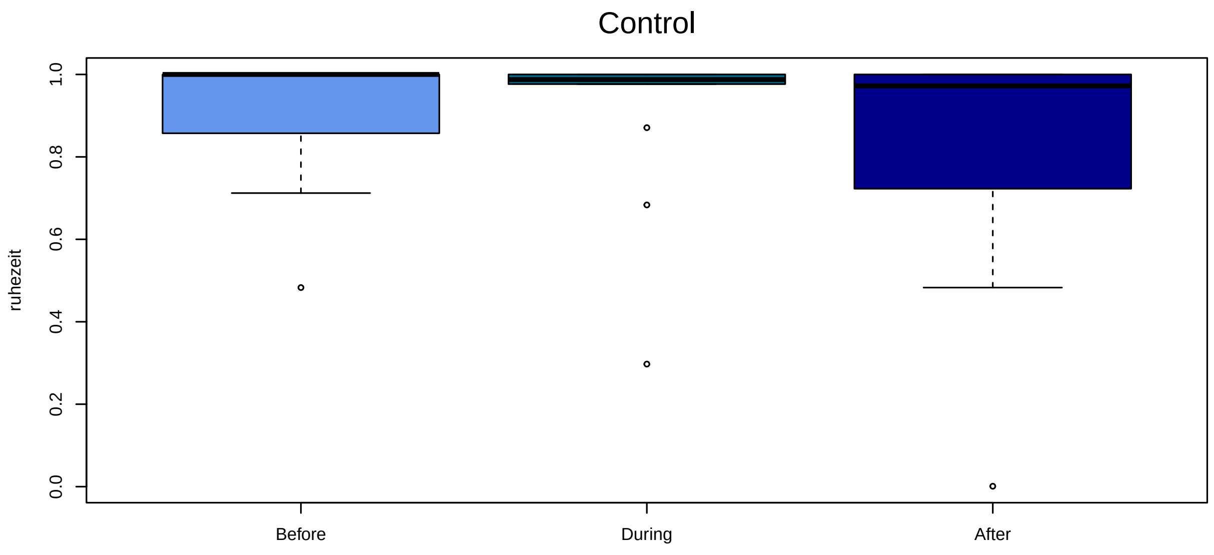
1 Data and Desciptive Statistics

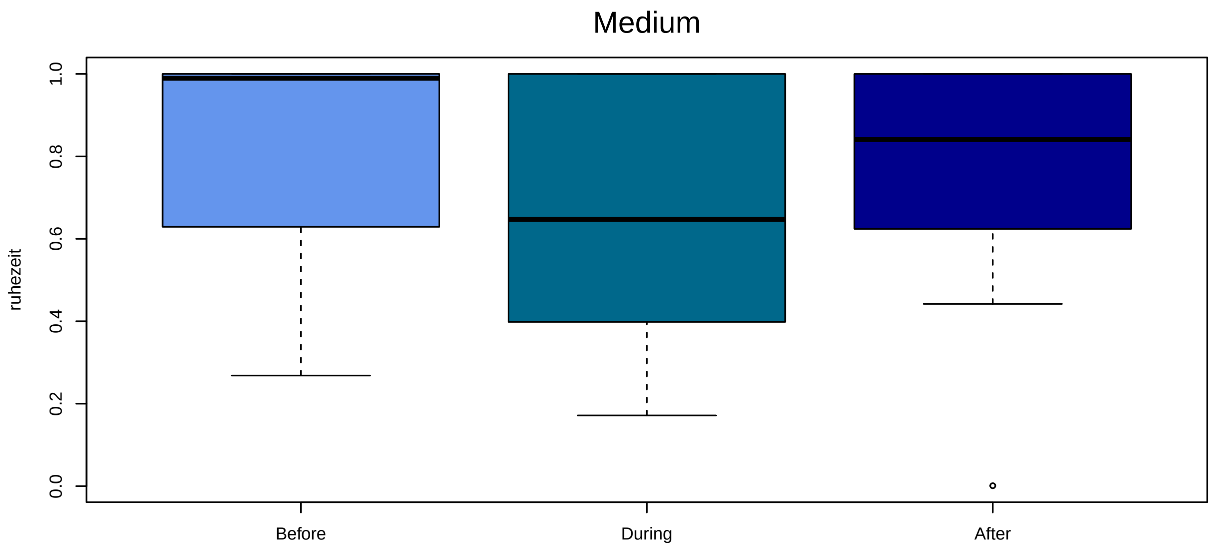
```r
# import data (this dataset is unfortunatelly not public)
openxlsx::read.xlsx("./../Data/Humpback_Whales_Data.xlsx") -> dat_whale

# fist a quick look at the missing values in the data
sapply(dat_whale,  \(x) sum(is.na(x))) |> knitr::kable(col.names = "NAs")
```

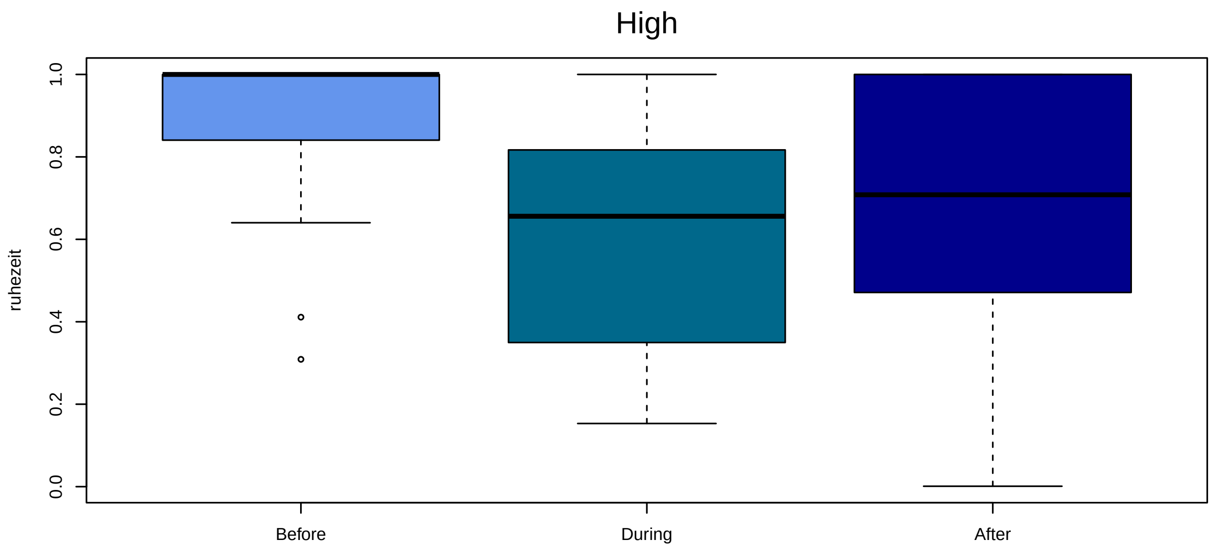
	NAs
Individuum	0
Treatment	0
Szenario	0
ruhezeit	0
speed	5
Atem	0

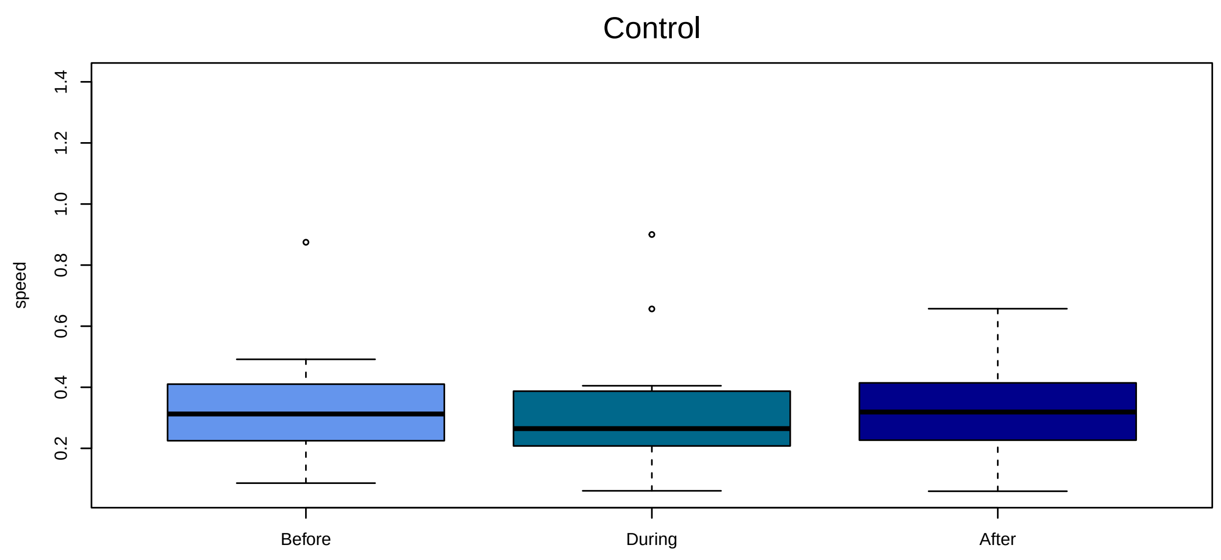
```r
# harmonize names
colnames(dat_whale) <- tolower(colnames(dat_whale))

# to numeric
lapply(dat_whale[, c("ruhezeit", "speed", "atem")], as.numeric) -> dat_whale[, c("ruhezeit", "speed", "

# to factor
lapply(dat_whale[ ,!(colnames(dat_whale) %in%  c("ruhezeit", "speed", "atem"))],
       as.factor) -> dat_whale[ ,!(colnames(dat_whale) %in%  c("ruhezeit", "speed", "atem"))]

# relevel
factor(dat_whale[, "szenario"],
       levels = c("Before", "During", "After")) -> dat_whale[, "szenario"]
factor(dat_whale[, "treatment"], c("Control", "Medium", "High")) -> dat_whale[, "treatment"]

# add log
within(dat_whale,{
   logspeed <- log(speed)
   atem[atem == 0] <- 0.001
   ruhezeit[ruhezeit == 0] <- 0.001
   logatem <- log(atem)
   logruhezeit <- log(ruhezeit)
   sqrtatem <- sqrt(atem)
   sqrtspeed <- sqrt(speed)
   sqrtruhezeit <- sqrt(ruhezeit)
}) -> dat_whale


# frist split into different intensities
dat_whale_intens <- split(dat_whale, dat_whale[, "treatment"])
```

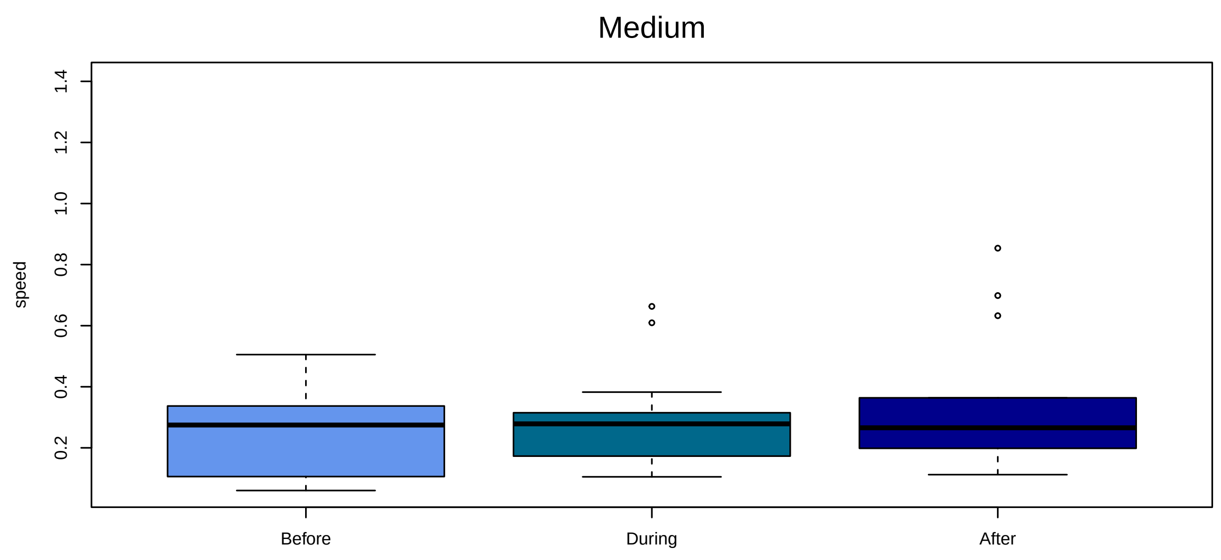
```r
# build formulas
formulae <- paste(c("ruhezeit", "speed", "atem"), "~", "szenario")

# max and min for plot y-axis
sapply(c(min, max), \(x){
   sapply(dat_whale[, c("ruhezeit", "speed", "atem")], \(y) x(y, na.rm = TRUE))


}) -> ylims

# over szenarios
invis.Map(\(y, nom, lims){

   # safe for presentation
   # pdf(paste0("./../Presentation/", nom, ".pdf"))

   # align
   par(mfrow = c(3, 1), mar = c(2, 4, 4, 2) + 0.1)

   # over treatment
   invis.Map(\(x, nom){

      # boxplots
      boxplot(as.formula(y), data = x,
              col = c("cornflowerblue", "deepskyblue4", "darkblue"),
              ylim = c(lims[1], lims[2]))

      # add label
      mtext(nom, side = 3, line = 1, cex = 1.2)

   }, dat_whale_intens, names(dat_whale_intens))

   # close graph. device
   # dev.off()

}, formulae, c("resting", "speed", "respatory"), ylims |> t() |> as.data.frame())
```

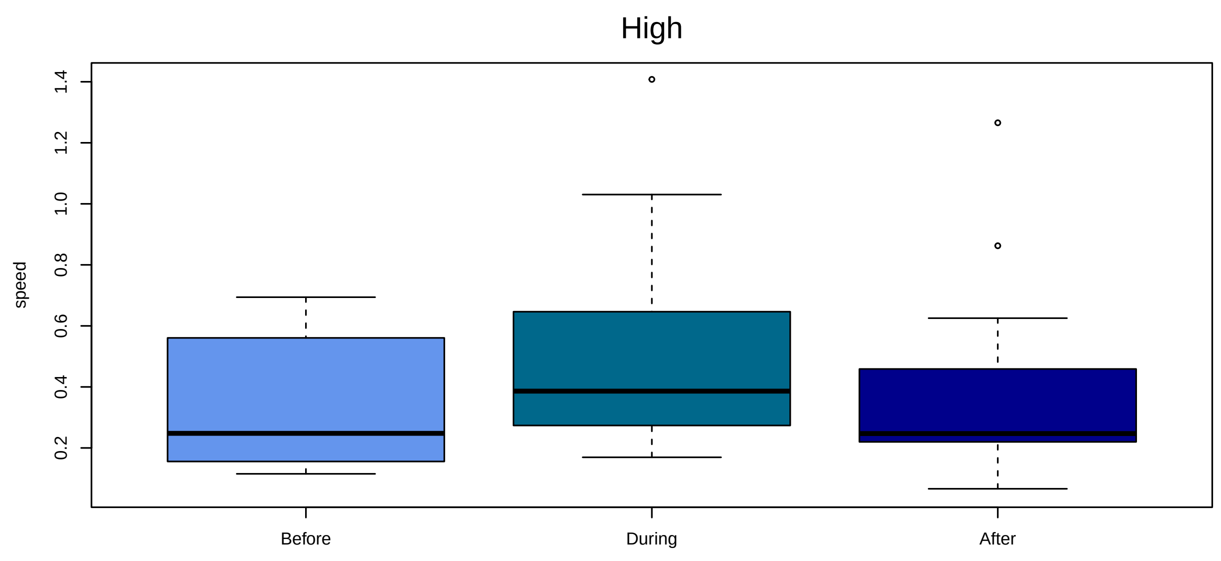

Control
ruhezeit
1.0
0.8
0.6
0.4
0.2
0.0
Before
During
After

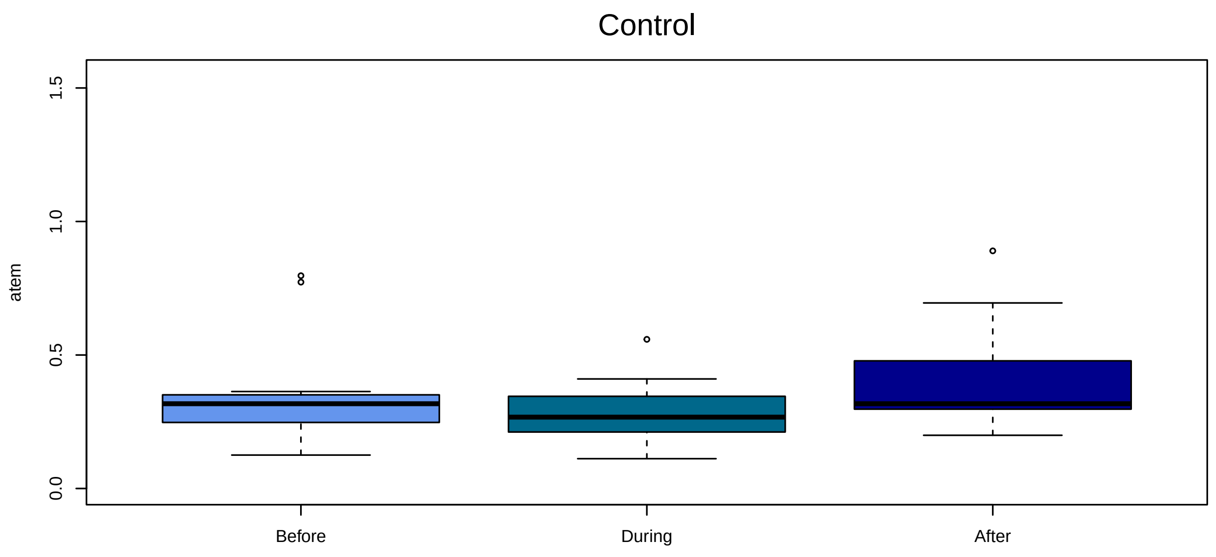

Medium
ruhezeit
1.0
0.8
0.6
0.4
0.2
0.0
Before
During
After

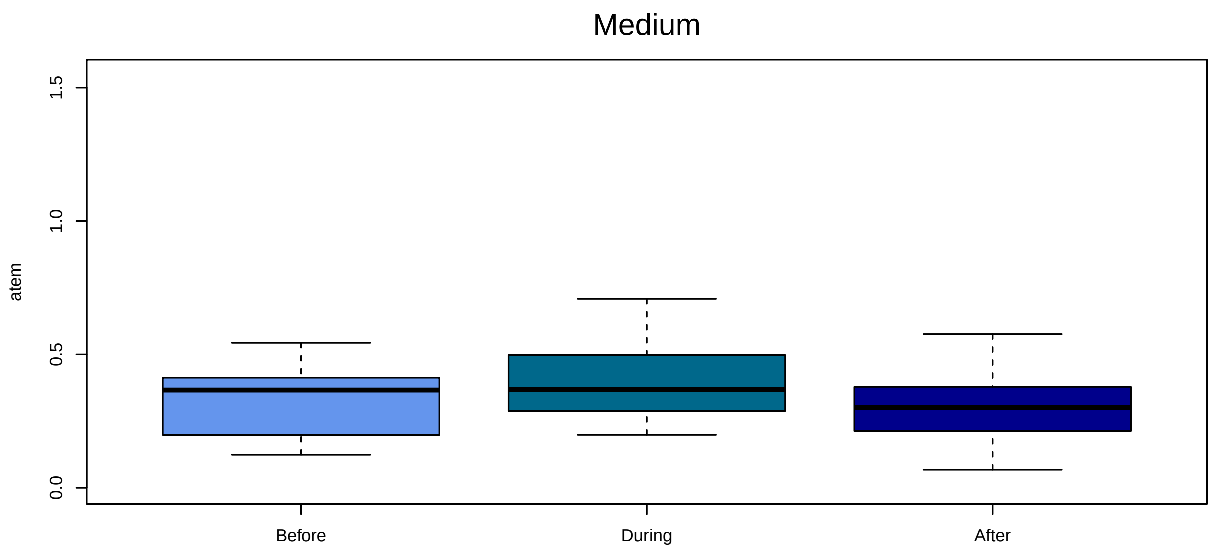

High
ruhezeit
1.0
0.8
0.6
0.4
0.2
0.0
Before
During
After

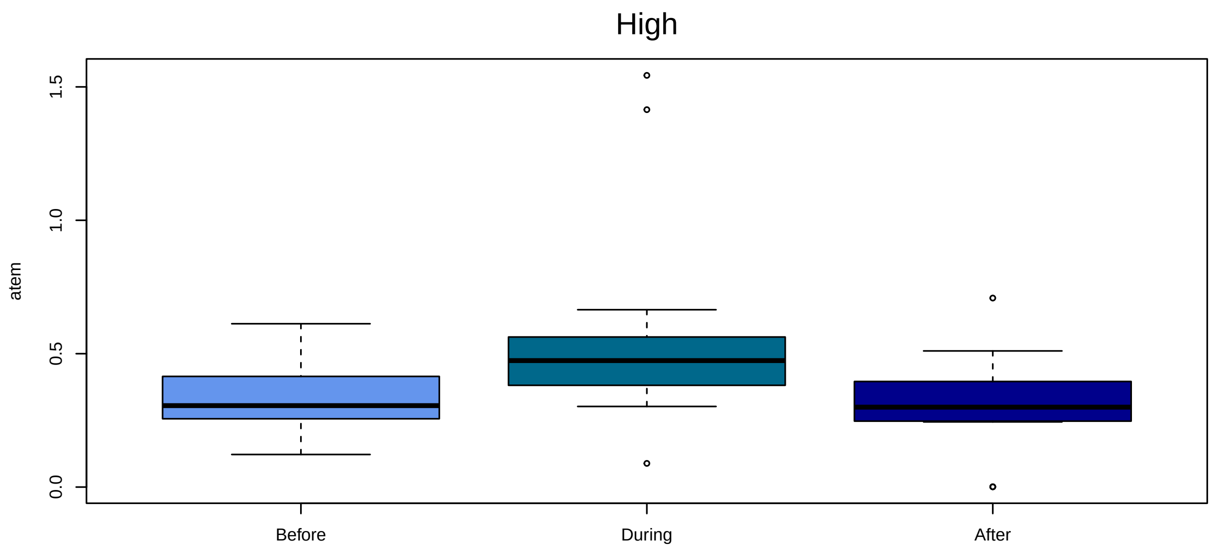

Control
speed
1.4
1.2
1.0
0.8
0.6
0.4
0.2
Before
During
After


Medium
speed
1.4
1.2
1.0
0.8
0.6
0.4
0.2
Before
During
After


High
speed
1.4
1.2
1.0
0.8
0.6
0.4
0.2
Before
During
After

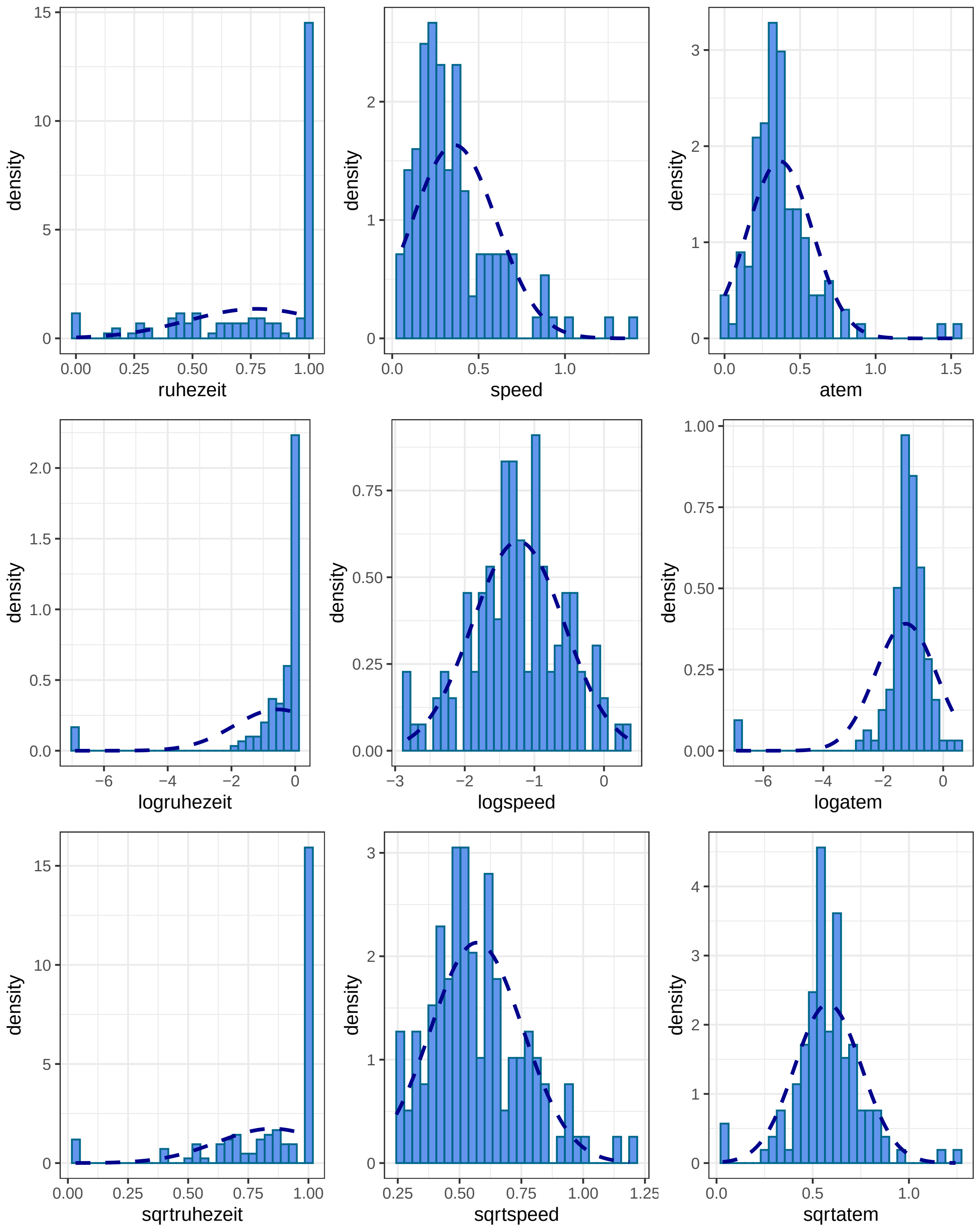

Control
atem
0.0
0.5
1.0
1.5
Before
During
After


Medium
atem
0.0
0.5
1.0
1.5
Before
During
After


High
atem
0.0
0.5
1.0
1.5
Before
During
After

2 Distribution of Covariates

```r
# vars
nom <- c("ruhezeit", "speed", "atem")
nom <- c(nom, paste0("log", nom), paste0("sqrt", nom))

# loop to generate plots
lapply(nom, \(x) Dens_norm_plot(y = x)) -> plots

# display (remove useless bin width messages)
print((plots[[1]] + plots[[2]] + plots[[3]]) /
      (plots[[4]] + plots[[5]] + plots[[6]]) /
      (plots[[7]] + plots[[8]] + plots[[9]])) |>
      suppressWarnings() |>
      suppressMessages()
```

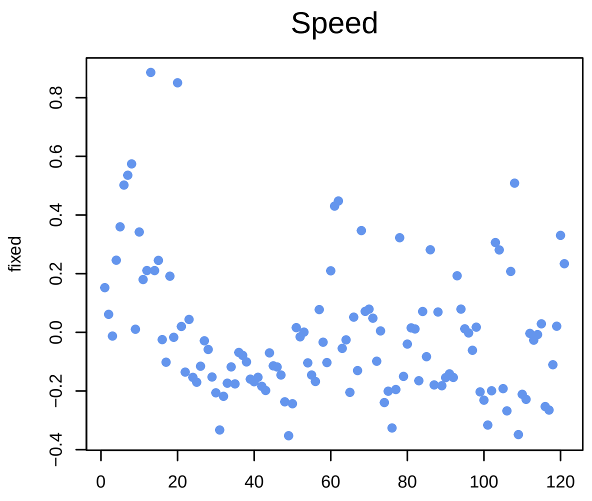

density
ruhezeit
density
speed
density
atem
density
logruhezeit
density
logspeed
density
logatem
density
sqrtruhezeit
density
sqrtspeed
density
sqrtatem

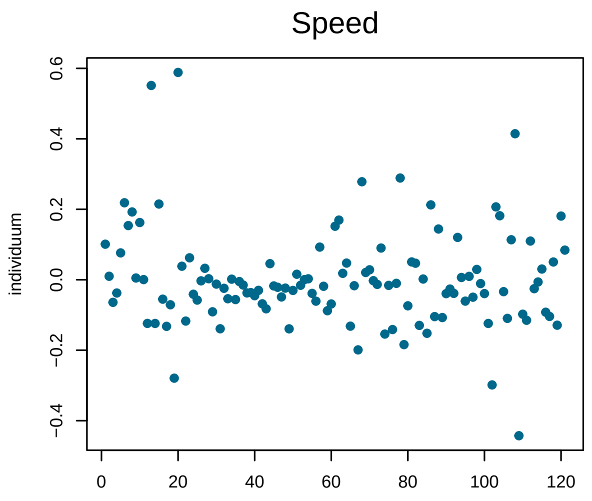
3 Model

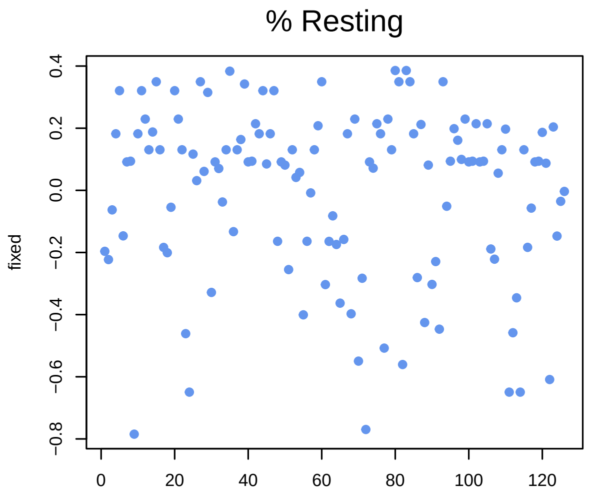
```r
# formulas
formulae_glmm <- c("speed ~ I(treatment) * I(szenario)",
                   "ruhezeit ~ I(treatment) * I(szenario)",
                   "atem ~ I(treatment) * I(szenario)")

# fit models
lapply(formulae_glmm, \(x){

   # fit
   MASS::glmmPQL(as.formula(x), random = ~ 1 | individuum,
                 family = quasi(link = "identity", variance = "constant"),
                 data = dat_whale) -> fit

   # summary
   list(fit,
        summary(fit))

}) |> setNames(c("speed", "ruhezeit", "atem")) |> suppressMessages() -> models

# fits and summaries
glmmPQL_fits <- lapply(models, "[[", 1)
glmmPQL_summaries <- lapply(models, "[[", 2)
```

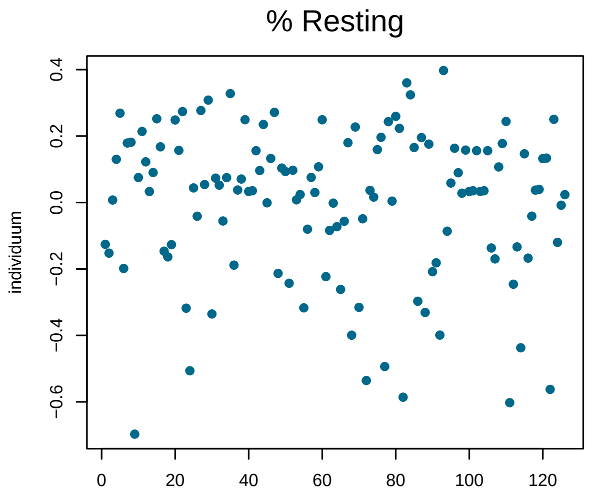
4 Residual Diagnostics

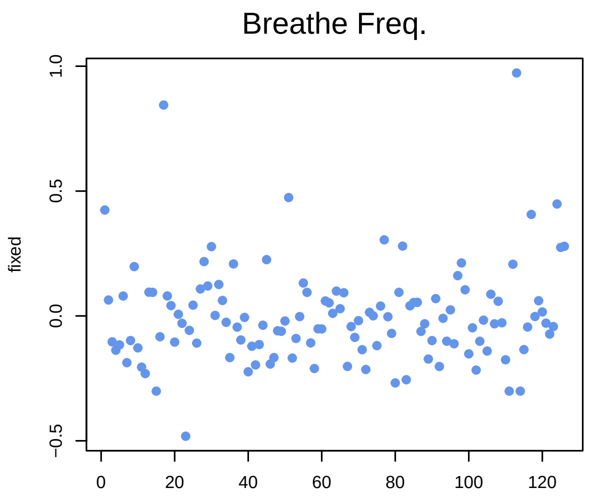
```r
# align
par(mfrow = c(3, 2), mar = c(2, 4, 4, 2) + 0.1)

# residual plots
invis.Map(\(x, nom){

   invis.Map(\(y, col){

      # plots
      plot(x[["residuals"]][, y], type = "p", ylab = y, xlab = "", pch = 19,
           col = col)

      # label
      mtext(nom, side = 3, line = 1, cex = 1.2)

   }, c("fixed", "individuum"), c("cornflowerblue", "deepskyblue4"))

}, glmmPQL_fits, c("Speed", "% Resting", "Breathe Freq."))
```

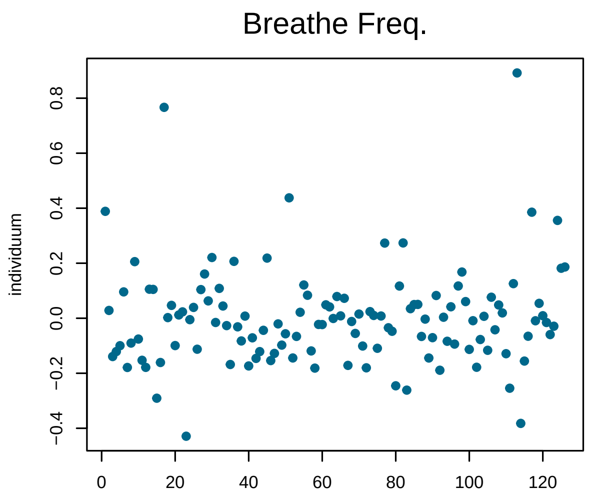

Speed
fixed
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
0
20
40
60
80
100
120


Speed
individuum
−0.4
−0.2
0.0
0.2
0.4
0.6
0
20
40
60
80
100
120


% Resting
fixed
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0
20
40
60
80
100
120


% Resting
individuum
−0.6
−0.4
−0.2
0.0
0.2
0.4
0
20
40
60
80
100
120


Breathe Freq.
fixed
−0.5
0.0
0.5
1.0
0
20
40
60
80
100
120


Breathe Freq.
individuum
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
0
20
40
60
80
100
120

5 Coefficients and t-Tests

```r
invis.lapply(c("speed", "ruhezeit", "atem"), \(x){

  # mods
  mod <- models[[x]][[2]][["tTable"]]

  # print table
  print(knitr::kable(mod))
})
```

```
##
##
## |                                     |      Value| Std.Error| DF|    t-value|   p-value|
## |:------------------------------------|----------:|---------:|--:|----------:|---------:|
## |(Intercept)                          |  0.3392015| 0.0675730| 73|  5.0197817| 0.0000035|
## |I(treatment)Medium                   | -0.0796651| 0.0950922| 39| -0.8377670| 0.4072666|
## |I(treatment)High                     |  0.0128286| 0.0923222| 39|  0.1389543| 0.8902013|
## |I(szenario)During                    | -0.0132775| 0.0659039| 73| -0.2014679| 0.8408931|
## |I(szenario)After                     |  0.0365651| 0.0677241| 73|  0.5399125| 0.5909006|
## |I(treatment)Medium:I(szenario)During |  0.0571340| 0.0932022| 73|  0.6130111| 0.5417742|
## |I(treatment)High:I(szenario)During   |  0.1831087| 0.0900419| 73|  2.0335943| 0.0456279|
## |I(treatment)Medium:I(szenario)After  |  0.0556378| 0.0941069| 73|  0.5912197| 0.5561995|
## |I(treatment)High:I(szenario)After    |  0.0262402| 0.0934876| 73|  0.2806813| 0.7797488|
##
##
## |                                     |      Value| Std.Error| DF|    t-value|   p-value|
## |:------------------------------------|----------:|---------:|--:|----------:|---------:|
## |(Intercept)                          |  0.9084175| 0.0787899| 78| 11.5296214| 0.0000000|
## |I(treatment)Medium                   | -0.0905042| 0.1094179| 39| -0.8271426| 0.4131887|
## |I(treatment)High                     | -0.0389885| 0.1076475| 39| -0.3621866| 0.7191678|
## |I(szenario)During                    | -0.0022044| 0.0992740| 78| -0.0222049| 0.9823412|
## |I(szenario)After                     | -0.1227239| 0.0992740| 78| -1.2362142| 0.2200887|
## |I(treatment)Medium:I(szenario)During | -0.1364697| 0.1378648| 78| -0.9898803| 0.3252935|
## |I(treatment)High:I(szenario)During   | -0.2528542| 0.1356341| 78| -1.8642375| 0.0660507|
## |I(treatment)Medium:I(szenario)After  |  0.0754663| 0.1378648| 78|  0.5473937| 0.5856712|
## |I(treatment)High:I(szenario)After    | -0.0962581| 0.1356341| 78| -0.7096892| 0.4800131|
##
##
## |                                     |      Value| Std.Error| DF|    t-value|   p-value|
## |:------------------------------------|----------:|---------:|--:|----------:|---------:|
## |(Intercept)                          |  0.3487213| 0.0568680| 78|  6.1321165| 0.0000000|
## |I(treatment)Medium                   | -0.0228728| 0.0789743| 39| -0.2896229| 0.7736389|
## |I(treatment)High                     | -0.0161951| 0.0776965| 39| -0.2084406| 0.8359698|
## |I(szenario)During                    | -0.0644547| 0.0757239| 78| -0.8511814| 0.3972742|
## |I(szenario)After                     |  0.0673726| 0.0757239| 78|  0.8897145| 0.3763551|
## |I(treatment)Medium:I(szenario)During |  0.1421530| 0.1051600| 78|  1.3517785| 0.1803532|
## |I(treatment)High:I(szenario)During   |  0.3022035| 0.1034585| 78|  2.9210116| 0.0045609|
## |I(treatment)Medium:I(szenario)After  | -0.0945670| 0.1051600| 78| -0.8992675| 0.3712780|
## |I(treatment)High:I(szenario)After    | -0.0978333| 0.1034585| 78| -0.9456288| 0.3472582|
```