Clean_luca

2024-12-07

```r
#loading in all necessary packages

library(here)
```

```
## here() starts at C:/Users/Fiona/OneDrive/Documents/bamfield 2024/DS/Nudireactors Coding and Data/nud
```

```r
library(ggplot2)
library(readxl)
library(sf)
```

```
## Warning: package 'sf' was built under R version 4.4.2
```

```
## Linking to GEOS 3.12.2, GDAL 3.9.3, PROJ 9.4.1; sf_use_s2() is TRUE
```

```r
library(cowplot)
library(rnaturalearth)
```

```
## Warning: package 'rnaturalearth' was built under R version 4.4.2
```

```r
library(ggspatial)
```

```
## Warning: package 'ggspatial' was built under R version 4.4.2
```

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

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

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

```r
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v forcats   1.0.0      v stringr   1.5.1
## v lubridate 1.9.3      v tibble    3.2.1
## v purrr     1.0.2      v tidyr     1.3.1
## v readr     2.1.5


## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter()    masks stats::filter()
## x dplyr::lag()       masks stats::lag()
## x lubridate::stamp() masks cowplot::stamp()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
library(circular)
```

```
## Warning: package 'circular' was built under R version 4.4.2


##
## Attaching package: 'circular'
##
## The following objects are masked from 'package:stats':
##
##     sd, var
```

```r
library(knitr)
library(glmmTMB)
library(DHARMa)
```

```
## This is DHARMa 0.4.6. For overview type '?DHARMa'. For recent changes, type news(package = 'DHARMa')
```

```r
library(ggeffects)
```

```
##
## Attaching package: 'ggeffects'
##
## The following object is masked from 'package:cowplot':
##
##     get_title
```

```r
library(performance)
library(car)
```

```
## Loading required package: carData
##
## Attaching package: 'car'
##
## The following object is masked from 'package:purrr':
##
##     some
##
## The following object is masked from 'package:dplyr':
##
##     recode
```

```r
library(mgcv)
```

```
## Loading required package: nlme
## 
## Attaching package: 'nlme'
## 
## The following object is masked from 'package:dplyr':
## 
##     collapse
## 
## This is mgcv 1.9-1. For overview type 'help("mgcv-package")'.
## 
## Attaching package: 'mgcv'
## 
## The following object is masked from 'package:circular':
## 
##     dpnorm
```

```r
library(fitdistrplus)
```

```
## Loading required package: MASS
## 
## Attaching package: 'MASS'
## 
## The following object is masked from 'package:dplyr':
## 
##     select
## 
## Loading required package: survival
```

```r
library(goft)
```

```
## Loading required package: sn
## Loading required package: stats4
## 
## Attaching package: 'sn'
## 
## The following objects are masked from 'package:circular':
## 
##     sd, sd.default
## 
## The following object is masked from 'package:lubridate':
## 
##     dst
## 
## The following object is masked from 'package:stats':
## 
##     sd
```
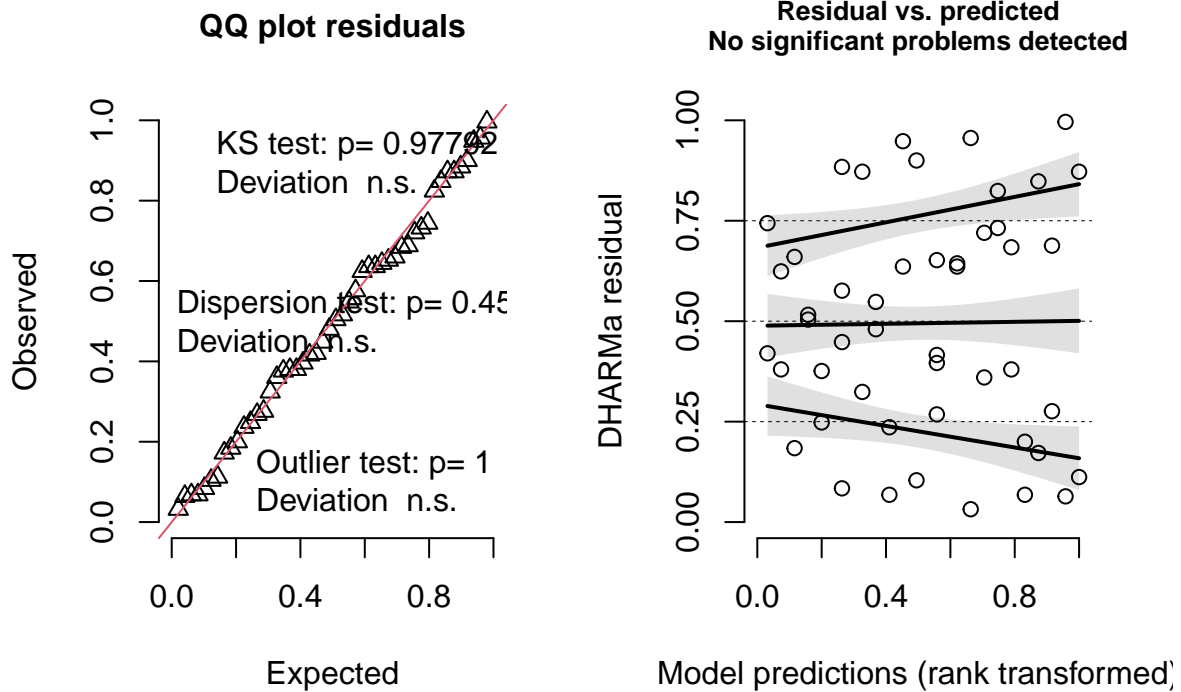
```r
library(gamlss)
```

```
## Loading required package: splines
## Loading required package: gamlss.data
##
## Attaching package: 'gamlss.data'
##
## The following object is masked from 'package:datasets':
##
##     sleep
##
## Loading required package: gamlss.dist
## Loading required package: parallel
##  **********   GAMLSS Version 5.4-22   **********
## For more on GAMLSS look at https://www.gamlss.com/
## Type gamlssNews() to see new features/changes/bug fixes.
##
##
## Attaching package: 'gamlss'
##
## The following object is masked from 'package:DHARMa':
##
##     getQuantile
##
## The following object is masked from 'package:glmmTMB':
##
##     refit
```

```r
#load in data, change temperature to factor, remove extra space in "conspecific"
nudi_data_factor <- read.csv("./clean_data/Meyknecht_Over_Parker_MRNE475_2024_combinedraw.csv")%>%
  mutate(temp_grp = as.factor(temp_grp))
nudi_data_factor$trial_type[nudi_data_factor$trial_type=="conspecific "] <- "conspecific"
```

```r
#make linear model with nudi data with interaction between temp and length
mod_nudi <- lm(mean_vel ~ temp_grp * avg_length_m,
            data = nudi_data_factor)

#plot simulated residuals to see whether it fits
simulateResiduals(mod_nudi)%>%
  plot()
```

DHARMa residual

QQ plot residuals

KS test: p= 0.97742
Deviation n.s.

Dispersion test: p= 0.45
Deviation n.s.

Outlier test: p= 1
Deviation n.s.

Observed / Expected

Residual vs. predicted
No significant problems detected

DHARMa residual

Model predictions (rank transformed)

```
#summary of the linear model
summary(mod_nudi)
```

```
##
## Call:
## lm(formula = mean_vel ~ temp_grp * avg_length_m, data = nudi_data_factor)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -6.033e-04 -1.810e-04 -1.209e-05  1.570e-04  6.814e-04
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)            0.0004563  0.0002229   2.047   0.0473 *
## temp_grp14            -0.0008480  0.0005674  -1.495   0.1429
## temp_grp16            -0.0006816  0.0006691  -1.019   0.3144
## temp_grp18            -0.0008107  0.0003933  -2.061   0.0458 *
## avg_length_m           0.0181207  0.0090403   2.004   0.0518 .
## temp_grp14:avg_length_m  0.0204947  0.0169694   1.208   0.2342
## temp_grp16:avg_length_m  0.0230516  0.0212045   1.087   0.2835
## temp_grp18:avg_length_m  0.0211581  0.0137253   1.542   0.1311
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000298 on 40 degrees of freedom
## Multiple R-squared:  0.4806, Adjusted R-squared:  0.3897
```

```
## F-statistic: 5.287 on 7 and 40 DF,  p-value: 0.0002471
```

```r
#fit model into gamma as interaction (temp*length)
nudi_gamma_factor_times <- glmmTMB(mean_vel ~ temp_grp * avg_length_m, family = Gamma(link = "log"), da

#fit simulated residuals in plot to see whether it fits (red means it doesn't, no red means it does)
simulateResiduals(nudi_gamma_factor_times)%>%
  plot()
```

DHARMa residual
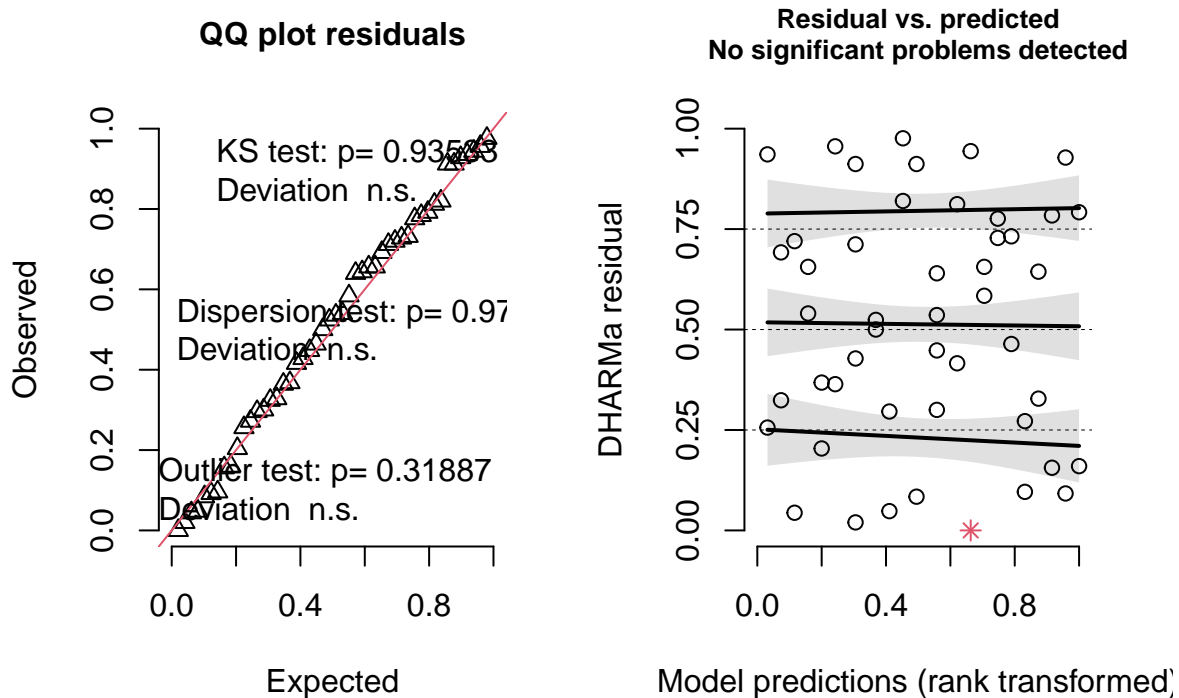

```r
#summarize the model
summary(nudi_gamma_factor_times)
```

```
##  Family: Gamma  ( log )
## Formula:          mean_vel ~ temp_grp * avg_length_m
## Data: nudi_data_factor
##
##      AIC      BIC   logLik deviance df.resid
##   -640.3   -623.5    329.1   -658.3       39
##
##
## Dispersion estimate for Gamma family (sigma^2): 0.082
##
## Conditional model:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -7.5489     0.2133  -35.40   <2e-16 ***
```

```
## temp_grp14                -1.0157    0.5475   -1.86   0.0636 .
## temp_grp16                -0.5313    0.6407   -0.83   0.4069
## temp_grp18                -0.9431    0.3790   -2.49   0.0128 *
## avg_length_m              21.1029    8.6412    2.44   0.0146 *
## temp_grp14:avg_length_m   23.9903   16.3507    1.47   0.1423
## temp_grp16:avg_length_m   17.8010   20.2979    0.88   0.3805
## temp_grp18:avg_length_m   22.9334   13.2102    1.74   0.0826 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Extract the coefficients
coef_table <- as.data.frame(summary(nudi_gamma_factor_times)$coefficients$cond)

# Save the table as a CSV file
write.csv(coef_table, "./luca_scripts/Meyknecht_Over_Parker_MRNE475_2024_nudi_gamma_factor_times_coeffic

#creating my own color palette for the different temp groups
my_cols <- c("12" = "red", "14" = "darkblue","16" = "green", "18"= "yellow" )

#boxplot mean velocity for different temperature groups without jitter
boxplot_nudi <- ggplot(nudi_data_factor, aes(x=temp_grp, y=mean_vel, fill=temp_grp)) +
    geom_boxplot(alpha=0.3) +
  scale_fill_manual(values = my_cols) +
  labs(y = "Mean velocity (m/s)\n", x = "\nTemperature (°C)",
       color = "Treatment group (°C)",
       fill = "Treatment group (°C)") +
  theme_classic()

#boxplot mean velocity for diffrent temperature groups with jitter
boxplot_jitter_nudi <- ggplot(nudi_data_factor, aes(x=temp_grp, y=mean_vel, fill=temp_grp)) +
    geom_point(aes(colour = temp_grp), position = position_jitterdodge()) +
    geom_boxplot(alpha=0.3) +
  scale_color_manual(values = my_cols) +
  scale_fill_manual(values = my_cols) +
  labs(y = "Mean velocity (m/s)\n", x = "\nTemperature (°C)",
       color = "Treatment group(°C)",
       fill = "Treatment group(°C)") +
  theme_classic()

#same boxplot, different theme
boxplot_nudi_goodtheme <- ggplot(nudi_data_factor, aes(x=temp_grp, y=mean_vel, fill=temp_grp)) +
    geom_boxplot(alpha=0.3) +
  scale_fill_manual(values = my_cols) +
  labs(y = "Mean velocity (m/s)\n", x = "\nTemperature (°C)",
       color = "Treatment group (°C)",
       fill = "Treatment group (°C)") +
theme_minimal() +
  theme(
    panel.grid.major = element_blank(),  # Remove major gridlines
    panel.grid.minor = element_blank(),  # Remove minor gridlines
    axis.line = element_line(color = "black", size = 0.5),  # Add axis lines for both x and y axes
    axis.ticks = element_line(color = "black", size = 0.5),  # Add ticks to both x and y axes
    axis.title.x = element_text(margin = margin(t = 10)),  # Add space between x-axis labels and numbers
    axis.title.y = element_text(margin = margin(r = 10)),  # Add space between y-axis labels and numbers
```
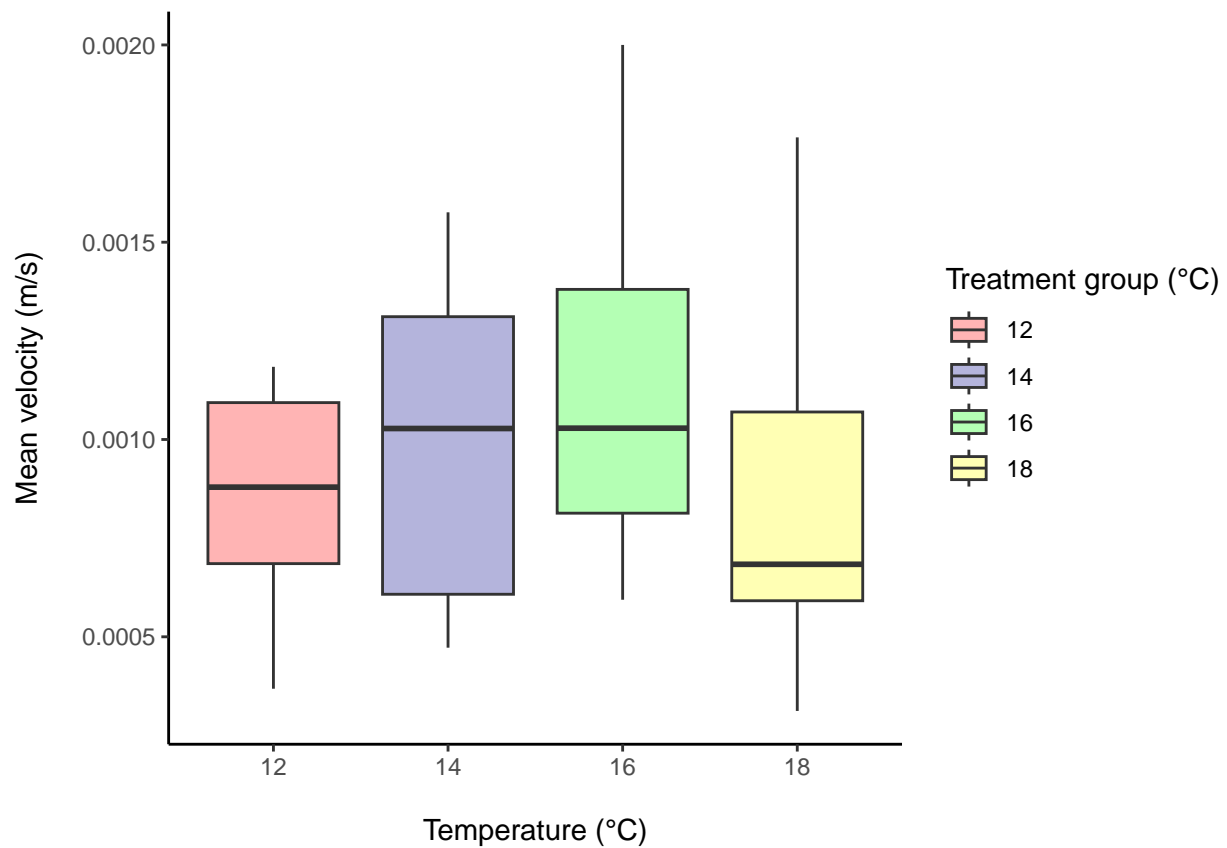
```
    strip.text = element_text(size = 12, face = "bold", margin = margin(b = 20))  # Add space between f
  )
```

```
## Warning: The 'size' argument of 'element_line()' is deprecated as of ggplot2 3.4.0.
## i Please use the 'linewidth' argument instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```
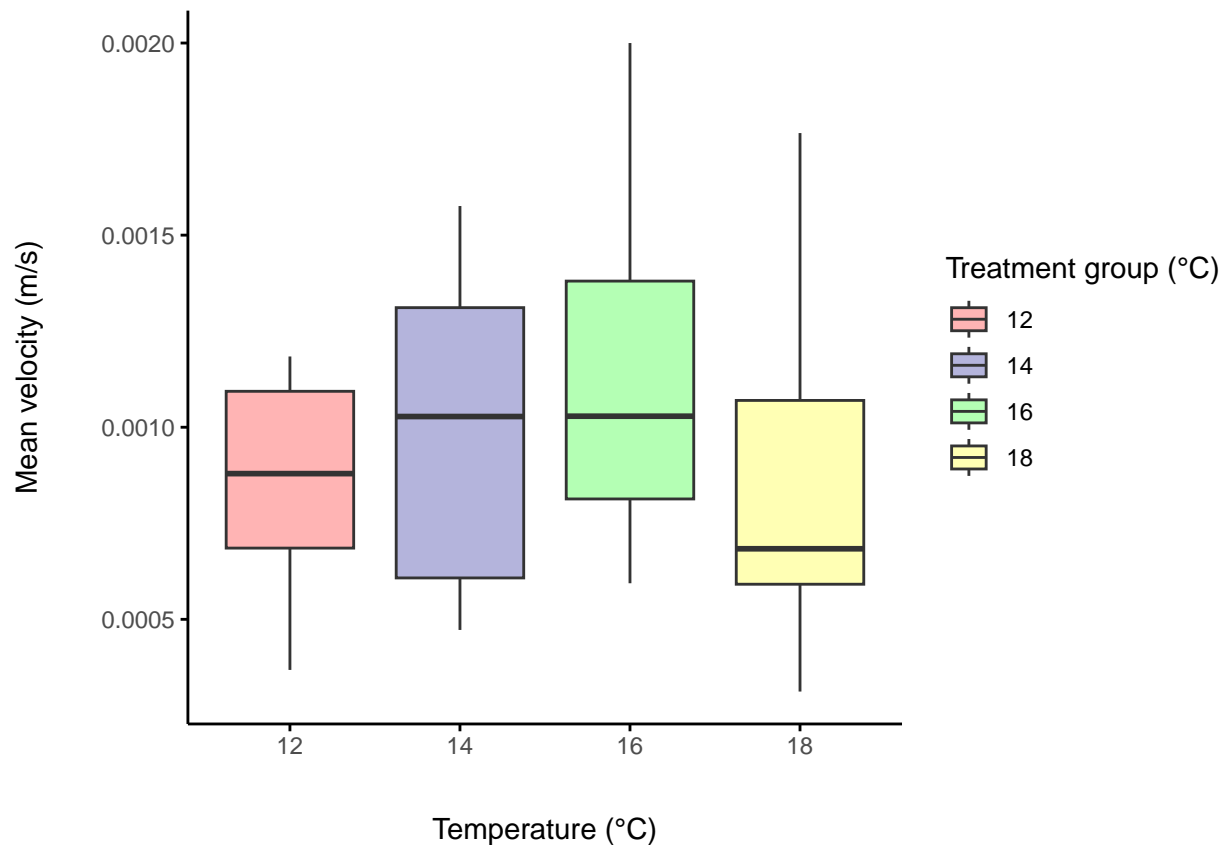
```
#show plots
boxplot_nudi
```



```
boxplot_jitter_nudi
```
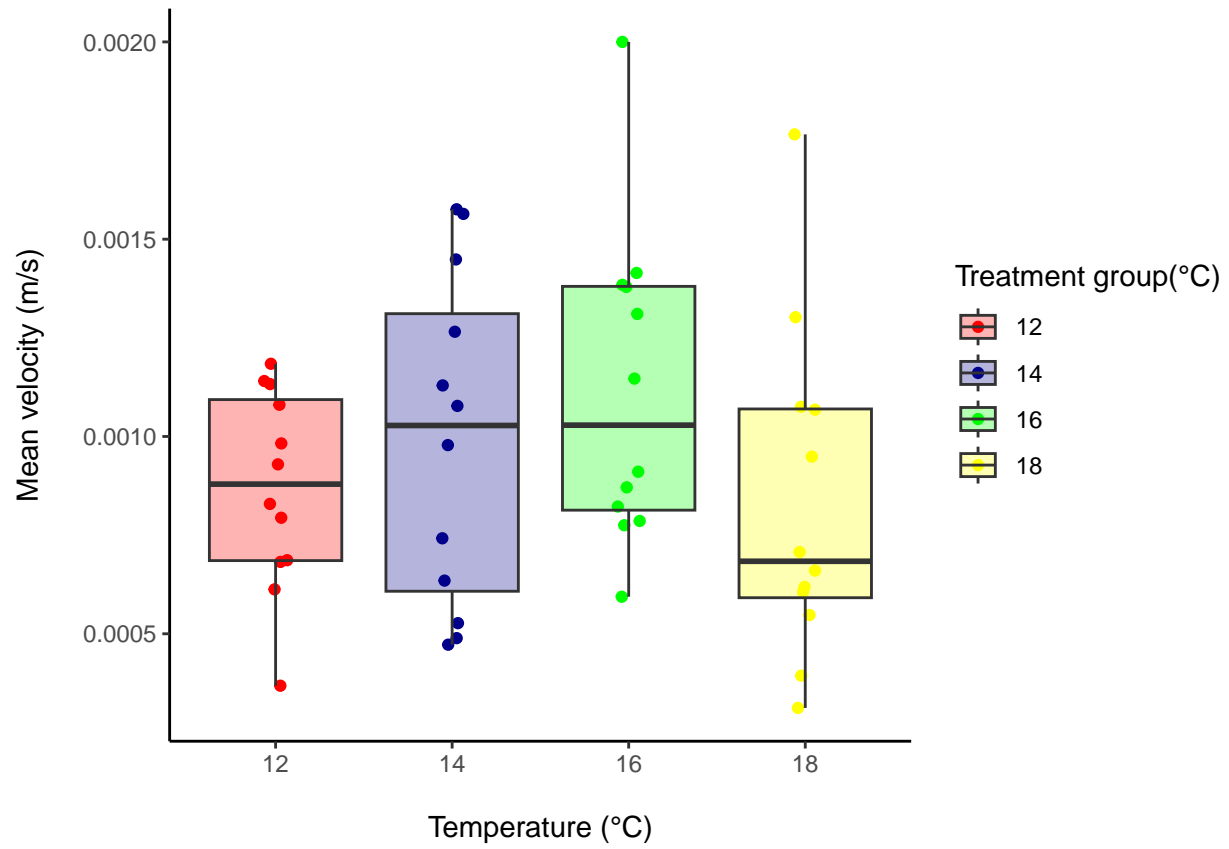
boxplot_nudi_goodtheme

```r
#save plots as picture
ggsave(plot = boxplot_nudi, filename = here("luca_scripts","Meyknecht_Over_Parker_MRNE475_2024_boxplot_
ggsave(plot = boxplot_jitter_nudi, filename = here("luca_scripts","Meyknecht_Over_Parker_MRNE475_2024_bo
ggsave(plot =boxplot_nudi_goodtheme, filename = here("luca_scripts", "Meyknecht_Over_Parker_MRNE475_202
```

```r
#print citations fro packages
print(citation("DHARMa"), style = "text")
```

```
## Hartig F (2022). _DHARMa: Residual Diagnostics for Hierarchical
## (Multi-Level / Mixed) Regression Models_. R package version 0.4.6,
## <https://CRAN.R-project.org/package=DHARMa>.
```

```r
print(citation("glmmTMB"), style = "text")
```

```
## Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen
## A, Skaug HJ, Maechler M, Bolker BM (2017). "glmmTMB Balances Speed and
## Flexibility Among Packages for Zero-inflated Generalized Linear Mixed
## Modeling." _The R Journal_, *9*(2), 378-400. doi:10.32614/RJ-2017-066
## <https://doi.org/10.32614/RJ-2017-066>.
```

```r
citation()
```

```
## To cite R in publications use:
```

```
## 
##   R Core Team (2024). _R: A Language and Environment for Statistical
##   Computing_. R Foundation for Statistical Computing, Vienna, Austria.
##   <https://www.R-project.org/>.
## 
## A BibTeX entry for LaTeX users is
## 
##   @Manual{,
##     title = {R: A Language and Environment for Statistical Computing},
##     author = {{R Core Team}},
##     organization = {R Foundation for Statistical Computing},
##     address = {Vienna, Austria},
##     year = {2024},
##     url = {https://www.R-project.org/},
##   }
## 
## We have invested a lot of time and effort in creating R, please cite it
## when using it for data analysis. See also 'citation("pkgname")' for
## citing R packages.
```

```r
#summarize the mean velocity for each temp group to get all the boxplot values + mean
summary(nudi_data_factor$mean_vel[nudi_data_factor$temp_grp=="12"])
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## 0.0003684 0.0006854 0.0008793 0.0008686 0.0010935 0.0011842
```

```r
summary(nudi_data_factor$mean_vel[nudi_data_factor$temp_grp=="14"])
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## 0.0004724 0.0006078 0.0010279 0.0009920 0.0013112 0.0015758
```

```r
summary(nudi_data_factor$mean_vel[nudi_data_factor$temp_grp=="16"])
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## 0.0005939 0.0008133 0.0010287 0.0011162 0.0013804 0.0020000
```

```r
summary(nudi_data_factor$mean_vel[nudi_data_factor$temp_grp=="18"])
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## 0.0003121 0.0005914 0.0006837 0.0008338 0.0010698 0.0017657
```

```r
#average and se of mean velocity
nudi_data_factor %>%
group_by(temp_grp) %>%
summarize(mean = mean(mean_vel)*100,
se = (sd(mean_vel)/sqrt(length(mean_vel)))*100)
```

```
## # A tibble: 4 x 3
##   temp_grp  mean      se
##   <fct>    <dbl>   <dbl>
```

```
## 1 12        0.0869 0.00726
## 2 14        0.0992 0.0120
## 3 16        0.112  0.0115
## 4 18        0.0834 0.0120
```

```r
# Calculate the average mean velocity for each temp_grp
mean_values <- nudi_data_factor %>%
  group_by(temp_grp) %>%
  summarise(mean_mean_vel = mean(mean_vel, na.rm = TRUE))

# create boxplot with average mean velocity per temp group in there aswell
boxplot_nudi <- ggplot(nudi_data_factor, aes(x = temp_grp, y = mean_vel, fill = temp_grp)) +
  geom_boxplot(alpha = 0.3)  +
  scale_fill_manual(values = my_cols) +
  geom_point(data = mean_values, aes(x = temp_grp, y = mean_mean_vel),
             color = "black", size = 3, shape = 21, fill="black") +
  labs(
    y = "Mean velocity (m/s)\n",
    x = "\nTemperature (°C)",
    color = "Treatment group (°C)",
    fill = "Treatment group (°C)"
  ) +
 theme_minimal() +
  theme(
    panel.grid.major = element_blank(),  # Remove major gridlines
    panel.grid.minor = element_blank(),  # Remove minor gridlines
    axis.line = element_line(color = "black", size = 0.5),  # Add axis lines for both x and y axes
    axis.ticks = element_line(color = "black", size = 0.5),  # Add ticks to both x and y axes
    axis.title.x = element_text(margin = margin(t = 10)),  # Add space between x-axis labels and number
    axis.title.y = element_text(margin = margin(r = 10)),  # Add space between y-axis labels and number
    strip.text = element_text(size = 12, face = "bold", margin = margin(b = 20))  # Add space between f
  )

#show plot
boxplot_nudi
```
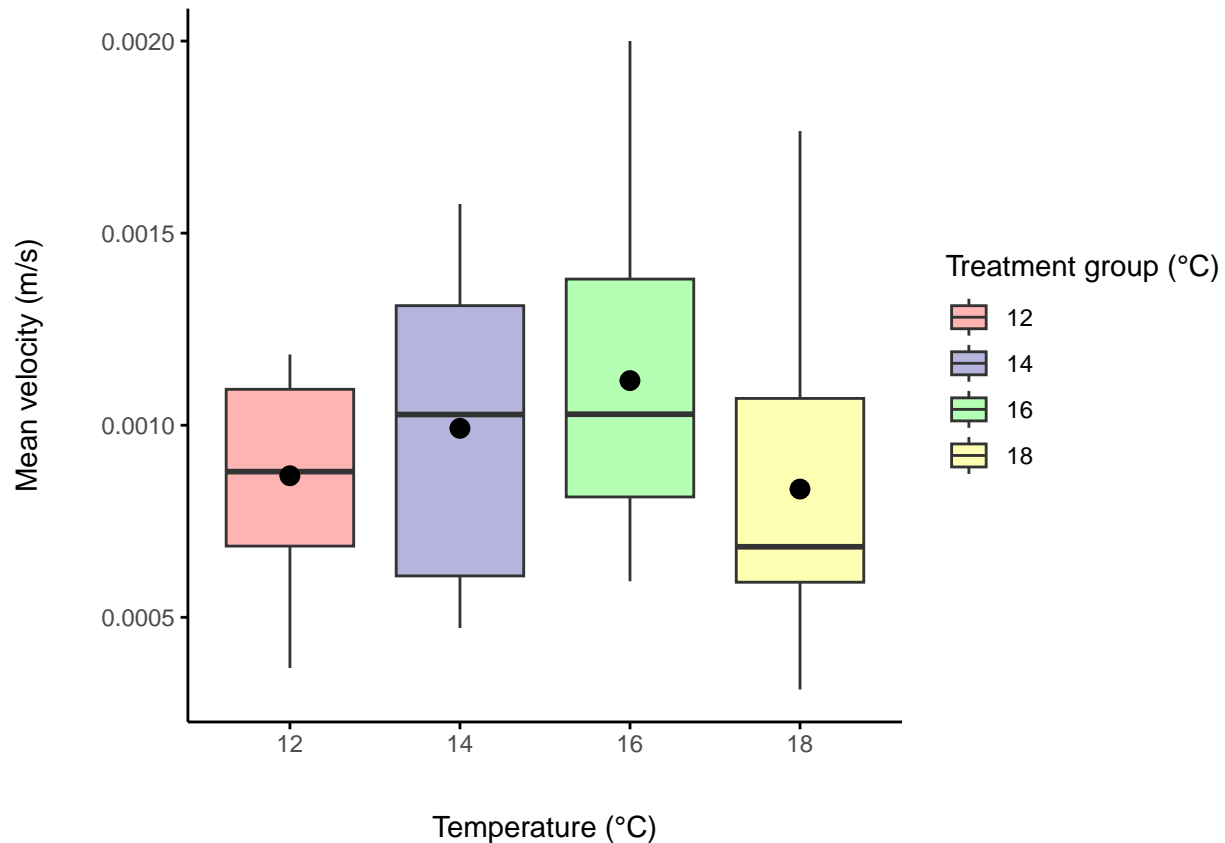
```
#save plot as picture
ggsave(plot =boxplot_nudi, filename = here("luca_scripts", "Meyknecht_Over_Parker_MRNE475_2024_perfectp]

#plot mean velocity and average length with points and glm, to see the relationship between the two
nudi_vel_length <- ggplot(nudi_data_factor, aes(x= avg_length_m, y = mean_vel)) +
  geom_point(size = 2) +
  labs(
    y = "Mean velocity (m/s)",
    x = "Average length (m)") +
  theme_minimal() +
  theme(
    panel.grid.major = element_blank(),  # Remove major gridlines
    panel.grid.minor = element_blank(),  # Remove minor gridlines
    axis.line = element_line(color = "black", size = 0.5),  # Add axis lines for both x and y axes
    axis.ticks = element_line(color = "black", size = 0.5),  # Add ticks to both x and y axes
    axis.title.x = element_text(margin = margin(t = 10)),  # Add space between x-axis labels and number:
    axis.title.y = element_text(margin = margin(r = 10)),  # Add space between y-axis labels and number:
    strip.text = element_text(size = 12, face = "bold", margin = margin(b = 20))  # Add space between f(
  )+
  geom_smooth(method = "glm", color = "darkgrey", fill = "lightgrey")

#save as picture
ggsave(plot = nudi_vel_length, filename = here("luca_scripts","Meyknecht_Over_Parker_MRNE475_2024_lm_le1

## 'geom_smooth()' using formula = 'y ~ x'
```

```
nudi_vel_length
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```