# preliminary_analyses_blending_groups

Debora

2025-06-16

## Data Grouping and Exclusion Rationale

The following groups were combined because there were no significant differences between them:

- 10°C and 15°C temperature treatments
- Lemna minor strains from a US lab and field collection
- Preparation methods with and without repeated first-daughter separation

```
library(lme4)
library(performance)
library(sjPlot)

## Testing for differences in preparation methods, average temperatures,
## and strains to justify merging of groups performed during data analysis

### This dataset contains replicates for which a preparation technique was performed (repeated first bo
original_dataset_2 <- read.csv("https://raw.githubusercontent.com/Cuddington-Lab/thermal-experiments/ma
original_dataset_2$prep <- rep("yes",times=length(original_dataset_2$Experiment_Number))

### This dataset contains replicates of experiments performed without a preparation technique
original_dataset_1 <- read.csv("https://raw.githubusercontent.com/Cuddington-Lab/thermal-experiments/ma
original_dataset_1$prep <- rep("no",times=length(original_dataset_1$Experiment_Number))

datin <- rbind(original_dataset_1,original_dataset_2)

datin <- datin[!(datin$Treatment == 0 & (datin$Obs_sd <= 2.1 | datin$Obs_sd >= 2.9))
               &!(datin$Treatment == 0.95 & (datin$Obs_sd <= 2.1 | datin$Obs_sd >= 2.9)),]

datin <- datin[!(datin$Treatment == 0 & (datin$Obs_ac <= -0.2 | datin$Obs_ac >= 0.2))
               &!(datin$Treatment == 0.95 & (datin$Obs_ac <= 0.92 | datin$Obs_ac >= 0.98)),]

datin <- subset(datin, !Errors == "y"|is.na(Errors))
datin <- subset(datin, !Treatment == "constant")

colnames(datin)[6] <- "Label"
table(datin$Label, datin$cat_1)
```

```
##
##                N N/A  P
##   0        66  0  16  0
##   0.95      0 78   0 73
##   constant  0  0   0  0
```

```r
levels(datin$cat_1) = c("no autocorrelation","hot-cold","no autocorrelation","cold-hot")
datin$Treatment<-paste0(datin$cat_1)

### Relabel groups to simplify visualization of results
library(stringr)
datin$Label <- str_replace(datin$Label, "0.95", "strong autocorrelation")
datin$Label <- str_replace(datin$Label, "0", "no autocorrelation")

datin$Exp_run <- str_sub(datin$Profile_name,-2,-1)
datin$Exp_run <- gsub('_','',datin$Exp_run)
datin$Exp_run <- as.numeric(datin$Exp_run)

# Subset data to keep only those Exp_run that have at least 3 occurrences within each Mean_Temperature
datin <- datin[
  ave(seq_along(datin$Exp_run),
      interaction(datin$Exp_run, datin$Mean_temperature, datin$Species),
      FUN = length) >= 3,
]

datin$total_living_fronds <- (datin$Frond_count_1 + datin$Frond_count_2 + datin$Frond_count_3)

# Datasets and variables to iterate over
datin1015 <- subset(datin, (Species == "Lab_LM" | Species == "Field_LM") &
                      (Mean_temperature == 10 | Mean_temperature == 15))

datasets <- list(datin1015 = datin1015)

predictors_list <- list(
  datin1015 = c("Mean_temperature*Treatment", "prep*Treatment")
)

# Initialize lists to save results
best_models <- list()
model_diagnostics <- list()

# Loop over datasets and predictors
for (data_name in names(datasets)) {
  data <- datasets[[data_name]]
  predictors <- predictors_list[[data_name]]

  for (pred in predictors) {
    # Fit models
    simple <- glm(as.formula(paste("total_living_fronds ~", pred)), data = data, family="poisson")
    exp_number <- glmer(as.formula(paste("total_living_fronds ~", pred, "+ (1|Exp_run)")), data = data,

    # Store candidate models
    Cand.modsF <- list("no random effects" = simple,
                       "experiment number" = exp_number)

    # Compare models
    compareF <- compare_performance(Cand.modsF)

    # Select best model and save it
```

```
    best_model_index <- which.min(compareF$AIC)
    best_model <- Cand.modsF[[best_model_index]]

    # Save best model and its diagnostics
    best_models[[paste(data_name, pred)]] <- best_model
}}

#at 10-15C, no differences between average temperature and preparation method
summary(best_models[[1]])
```

```
##
## Call:
## glm(formula = as.formula(paste("total_living_fronds ~", pred)),
##     family = "poisson", data = data)
##
## Coefficients:
##                                             Estimate Std. Error z value
## (Intercept)                                 2.987619   0.309239   9.661
## Mean_temperature                           -0.008150   0.023784  -0.343
## Treatmenthot-cold                           0.153296   0.434862   0.353
## Treatmentno autocorrelation                -0.179723   0.445021  -0.404
## Mean_temperature:Treatmenthot-cold         -0.011498   0.033533  -0.343
## Mean_temperature:Treatmentno autocorrelation  0.009838   0.034148   0.288
##                                            Pr(>|z|)
## (Intercept)                                  <2e-16 ***
## Mean_temperature                              0.732
## Treatmenthot-cold                             0.724
## Treatmentno autocorrelation                   0.686
## Mean_temperature:Treatmenthot-cold            0.732
## Mean_temperature:Treatmentno autocorrelation    0.773
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 15.024  on 47  degrees of freedom
## Residual deviance: 13.597  on 42  degrees of freedom
## AIC: 251.52
##
## Number of Fisher Scoring iterations: 4
```

```
summary(best_models[[2]])
```

```
##
## Call:
## glm(formula = as.formula(paste("total_living_fronds ~", pred)),
##     family = "poisson", data = data)
##
## Coefficients:
##                                    Estimate Std. Error z value Pr(>|z|)
## (Intercept)                         2.86479    0.07198  39.799   <2e-16 ***
## prepyes                             0.05837    0.12623   0.462    0.644
## Treatmenthot-cold                  -0.02094    0.10233  -0.205    0.838
## Treatmentno autocorrelation        -0.05868    0.10332  -0.568    0.570
```

```
## prepyes:Treatmenthot-cold            0.08346    0.17699   0.472    0.637
## prepyes:Treatmentno autocorrelation  0.01472    0.18073   0.081    0.935
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 15.024  on 47  degrees of freedom
## Residual deviance: 12.592  on 42  degrees of freedom
## AIC: 250.52
##
## Number of Fisher Scoring iterations: 4
```

```
#at 10-15C, no differences between duckweed strains (lab and field LM)
exp_number <- glmer(total_living_fronds ~ Species * Treatment + (1|Exp_run), data = datin1015, family="
simple <- glm(total_living_fronds ~ Species * Treatment, data = datin1015, family="poisson")
tab_model(simple)
```

total_living_fronds

Predictors

Incidence Rate Ratios

CI

p

(Intercept)

17.55

15.19 – 20.14

<0.001

Species [Field_LM]

1.06

0.82 – 1.35

0.644

Treatment [hot-cold]

0.98

0.80 – 1.20

0.838

Treatment [noautocorrelation]

0.94

0.77 – 1.15

0.570

Species [Field_LM] ×Treatment [hot-cold]

1.09

0.77 – 1.54

0.637

Species [Field_LM] ×Treatment [noautocorrelation]

1.01

0.71 – 1.45

0.935

Observations

48

R2 Nagelkerke

0.184