Host activity before and during infection influences resulting parasite intensities

David R. Clark

2024-09-04

```r
# Clear the working environment
rm(list = ls())
# Visualization
library(ggplot2)
library(visreg)
library(gridExtra)
source("http://highstat.com/Books/BGS/GAMM/RCodeP2/HighstatLibV6.R")
# (generalized) Linear mixed modeling
library(lme4)
library(glmmTMB)
library(lmodel2)
# Statistcal analysis reporting and model validation
library(performance)
library(car)
library(lmtest)
library(DHARMa)
# Data wrangling
library(dplyr)
library(plyr)
library(tidyverse)
library(tidylog)
library(splancs)
```

Visualize some patterns in the raw data

```r
# Load in dataset from above. This bit of code is meant to save time so people
# dont have to rerun the entire data parsing and calculating step above. Code
# will be saved for reproducibility.
IndBehav7 <- read_csv("IndividualBehaviors_20240501_WA.csv")
```

```
## Rows: 347 Columns: 27
## -- Column specification -------------------------------------------------------
## Delimiter: ","
## chr  (3): TrialTime, Sex, Treatment
## dbl (23): fishID, BehavGroup, AvgVel, BehavVig, VarvelBef, PreWeight, PreLen...
## lgl  (1): ContrPeriod
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
IndBehav7 <- IndBehav7 %>%
    rename(Wormaf = Wormbf, AUC = AUC2)
```

rename: renamed 2 variables (AUC, Wormaf)

```
# Subsetting down to Infected individuals
IndBehav7I <- IndBehav7 %>%
    filter(Infection == 1 & wormJump >= 0)
```

filter: removed 172 rows (50%), 175 rows remaining

```
# getting the mean and SD of worms jumped
summary(IndBehav7I$wormJump)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   2.000   2.000   1.909   2.000   3.000
```

```
sqrt(var(IndBehav7I$wormJump))
```

[1] 0.4189282

```
# Setting some of the factors back to factors
IndBehav7$fishID <- as.factor(IndBehav7$fishID)
IndBehav7$TrialTime <- as.factor(IndBehav7$TrialTime)
IndBehav7$Sex <- as.factor(IndBehav7$Sex)
IndBehav7$Infection <- as.factor(IndBehav7$Infection)
IndBehav7$ContInf <- as.factor(IndBehav7$ContInf)
IndBehav7$Died <- as.factor(IndBehav7$Died)
IndBehav7$Treatment <- as.factor(IndBehav7$Treatment)


# Calculating tissue tolerance for each individual.
IndBehav7 <- IndBehav7 %>%
    mutate(ChSMI = LateSMI - PreSMI) %>%
    mutate(TisTol = ChSMI/Totworm)
```

mutate: new variable 'ChSMI' (double) with 85 unique values and 0% NA
mutate: new variable 'TisTol' (double) with 47 unique values and 46% NA

```
# Subsetting down to female and males only to scale their SMI Males
IndBehavM <- IndBehav7 %>%
    filter(Sex == "M") %>%
    mutate(ScPSMI = scale(PreSMI), ScLSMI = scale(LateSMI))
```

filter: removed 180 rows (52%), 167 rows remaining
mutate: new variable 'ScPSMI' (double) with 41 unique values and 0% NA
new variable 'ScLSMI' (double) with 41 unique values and 0% NA

```r
# Females
IndBehavF <- IndBehav7 %>%
    filter(Sex == "F") %>%
    mutate(ScPSMI = scale(PreSMI), ScLSMI = scale(LateSMI))
```

```
## filter: removed 167 rows (48%), 180 rows remaining
## mutate: new variable 'ScPSMI' (double) with 45 unique values and 0% NA
##          new variable 'ScLSMI' (double) with 45 unique values and 0% NA
```

```r
# Combine the two separate dataframes together
IndBehav8 <- rbind(IndBehavF, IndBehavM)

# Scaling some variables to make them biologically comparable and better for
# model fitting
IndBehav8 <- IndBehav8 %>%
    mutate(ScVarvelBef = c(scale(VarvelBef)), ScNRatebf = c(scale(NRatebf)), ScBehavTol = c(scale(Behav
         ScTotworm = c(scale(Totworm)), ScAUC = c(scale(AUC)), ScBehavVig = c(scale(BehavVig)),
         ScChSMI = c(scale(ChSMI)), ScTisTol = c(scale(TisTol)))
```

```
## mutate: new variable 'ScVarvelBef' (double) with 86 unique values and 2% NA
##          new variable 'ScNRatebf' (double) with 85 unique values and 37% NA
##          new variable 'ScBehavTol' (double) with 49 unique values and 1% NA
##          new variable 'ScTotworm' (double) with 47 unique values and 46% NA
##          new variable 'ScAUC' (double) with 49 unique values and 43% NA
##          new variable 'ScBehavVig' (double) with 87 unique values and 1% NA
##          new variable 'ScChSMI' (double) with 85 unique values and 0% NA
##          new variable 'ScTisTol' (double) with 47 unique values and 46% NA
```

```r
# Filtering down to only infected and only individuals who have VIE injections
# for sample size questions
IndBehavind <- IndBehav8 %>%
    distinct(fishID, .keep_all = TRUE) %>%
    filter(Treatment == "VIE") %>%
    filter(Infection == "1")
```

```
## distinct: removed 260 rows (75%), 87 rows remaining
## filter: removed 52 rows (60%), 35 rows remaining
## filter: removed 18 rows (51%), 17 rows remaining
```

```r
IndBehavCh <- IndBehav8 %>%
    filter(TrialTime != "Before") %>%
    mutate(TrialTime = recode(TrialTime, Early = "1", Late = "2", Later = "3"))
```

```
## filter: removed 86 rows (25%), 261 rows remaining
## mutate: changed 261 values (100%) of 'TrialTime' (0 new NA)
```

```r
IndBehavCh$TrialTime <- as.factor(IndBehavCh$TrialTime)
IndBehavCh$Infection <- as.factor(IndBehavCh$Infection)


IndBehavChI <- IndBehavCh %>%
    filter(Infection == "1")
```

```
## filter: removed 117 rows (45%), 144 rows remaining
```

```r
IndBehavChU <- IndBehavCh %>%
    filter(Infection == "0")
```

```
## filter: removed 144 rows (55%), 117 rows remaining
```

What hypotheses we want to test with these data and what data we can use to test them?

- Does host behavioral vigor trade-off with their ability to resist parasite infection?
- Does higher host activity during infection lead to higher parasite infection?

What factors are important for host infection intensity over the course of infection and do is there any sexual variation in host infection intensity?

Visually inspection of the explanatory variables that will be used in the analyses

```r
pairs(~AUC + Sex + ScPSMI + ScVarvelBef + ScBehavTol + ScBehavVig + PreLength + Treatment,
    lower.panel = panel.smooth, diag.panel = panel.hist, upper.panel = panel.cor,
    data = IndBehav8)
```
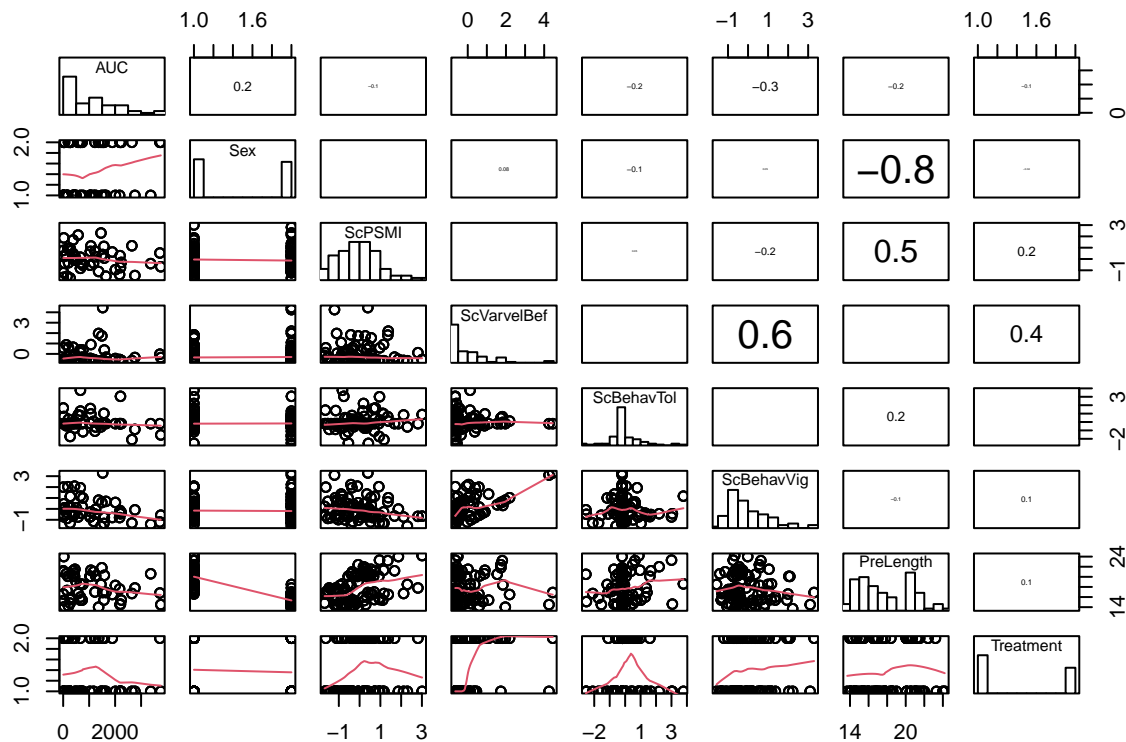

Description, development, and fitting of linear model for the analysis

We will use a linear model to analyze how infection intensity differs by sexual variation and other important host traits. Given each host only has one infection intensity measure, we do not need the fishID random effect used in previous models.
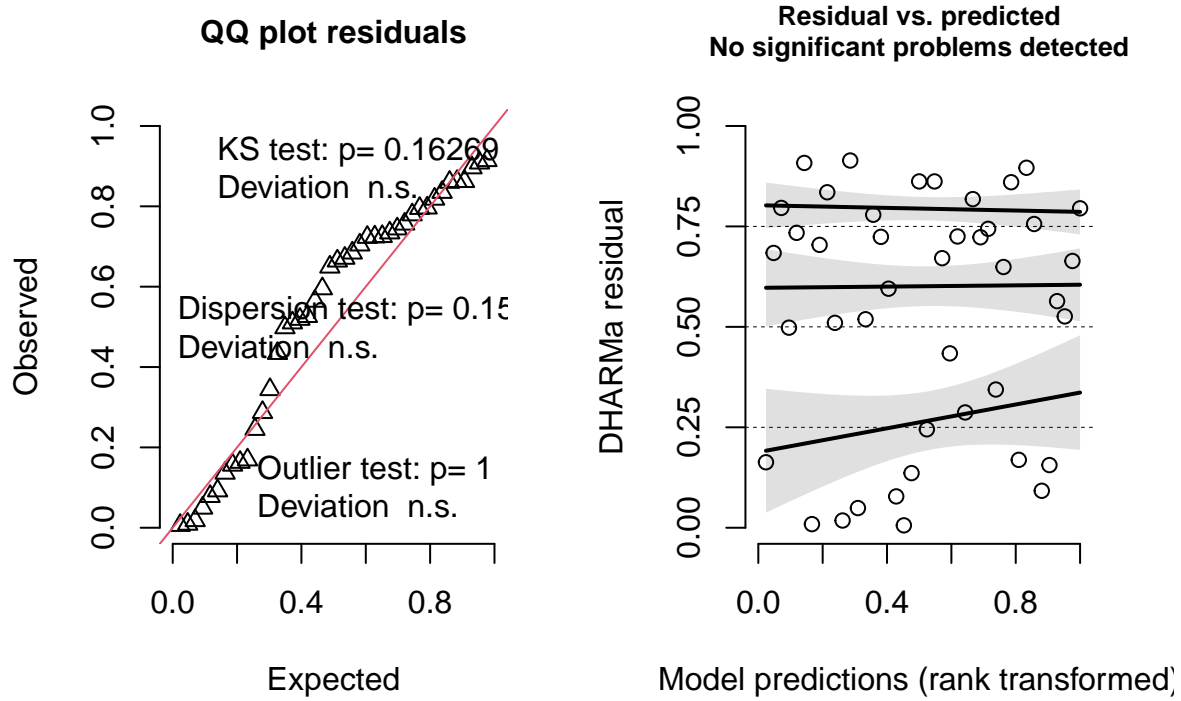
4

- Deterministic

- AUC_{det} = a + b_1Sex + b_2ScBehavVig + b_3ScVarVelBef + b_4ScResidPLSMI + b_5ScRPLength + b_6Sex:ScBehavVig + b_7Sex:ScResidPLSMI

- Stochastic

  - AUC ~ $N(AUC_{det}, \sigma^2)$

- Fixed

  - Sex
  - Scaled behavioral vigor
  - Scaled Pre-infection SMI
  - Scaled variance in velocity before infection
  - Scaled Residuals of Length and Sex
  - VIE treatment
  - Interaction between Sex and behavioral vigor
  - Interaction between sex and body condition

```r
# Scaling our pre-infection sclaed mass index
IndBehavI$ScPSMI <- as.numeric(IndBehavI$ScPSMI)
# Fit a linear model for infection integral Fitting a glm because we only have
# one measure of beahvioral vigor and infection integral
AUCLM <- glm(AUC ~ Sex + ScBehavVig + ScVarvelBef + ScRPLength + Treatment + wormJump +
    Sex:ScRPLength, family = Gamma(link = "log"), IndBehavI)

# Summary to see the relationship of the variables.
summary(AUCLM)
```

```
##
## Call:
## glm(formula = AUC ~ Sex + ScBehavVig + ScVarvelBef + ScRPLength +
##     Treatment + wormJump + Sex:ScRPLength, family = Gamma(link = "log"),
##     data = IndBehavI)
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      6.19754    0.61592  10.062 9.95e-12 ***
## SexM             0.53484    0.26311   2.033  0.04995 *
## ScBehavVig      -0.55887    0.15749  -3.549  0.00116 **
## ScVarvelBef      0.45869    0.18094   2.535  0.01601 *
## ScRPLength      -0.55197    0.16684  -3.308  0.00222 **
## TreatmentVIE    -0.05194    0.29171  -0.178  0.85975
## wormJump         0.20270    0.30905   0.656  0.51631
## SexM:ScRPLength  0.51726    0.29635   1.745  0.08994 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.689384)
##
##     Null deviance: 69.030  on 41  degrees of freedom
## Residual deviance: 55.847  on 34  degrees of freedom
##   (6 observations deleted due to missingness)
## AIC: 675.03
```

```
##
## Number of Fisher Scoring iterations: 9
```

Validate that the model fits well and there are no problems

```
# Using the check_model function from the perforamnce package to check the
# model validation
check_model(AUCLM)
```

Posterior Predictive Check
Model–predicted lines should resemble observed da

Homogeneity of Variance
Reference line should be flat and horizontal

Influential Observations
Points should be inside the contour lines

Collinearity
High collinearity (VIF) may inflate parameter uncer

Uniformity of Residuals
Dots should fall along the line


```
# Using the Dharma package to check quantile residuals First simulating the
# quantile residuals
sim_residuals_AUCLM <- simulateResiduals(AUCLM, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_AUCLM)
```

DHARMa residual

QQ plot residuals

KS test: p= 0.16269
Deviation n.s.

Dispersion test: p= 0.15
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)

```
# Testing for dispersion
testDispersion(sim_residuals_AUCLM)
```

DHARMa nonparametric dispersion test via sd of
residuals fitted vs. simulated
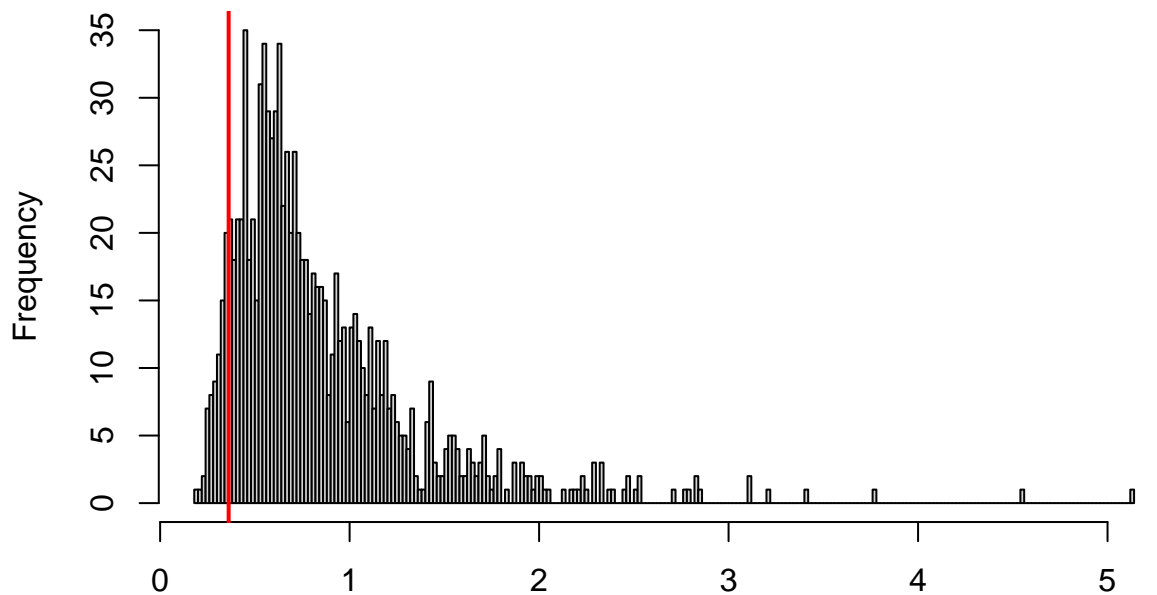
Frequency

Simulated values, red line = fitted model. p−value (two.sided) = 0.154

```
##
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
##
## data:  simulationOutput
## dispersion = 0.42402, p-value = 0.154
## alternative hypothesis: two.sided
```

Testing the significance of factors in our model

```
Anova(AUCLM, type = 3)
```

```
## Analysis of Deviance Table (Type III tests)
##
## Response: AUC
##                LR Chisq Df Pr(>Chisq)
## Sex              3.9769  1   0.046127 *
## ScBehavVig      10.4695  1   0.001214 **
## ScVarvelBef      7.2344  1   0.007152 **
## ScRPLength       6.8738  1   0.008747 **
## Treatment        0.0325  1   0.856908
## wormJump         0.5585  1   0.454867
## Sex:ScRPLength   2.5644  1   0.109294
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Visualize the important explanatory factors for infection intensity

```
# Behavioral Vigor graph
InfIntbyVig = visreg(AUCLM, scale = "response", "ScBehavVig", partial = T, gg = TRUE) +
    theme_classic() + theme(legend.position = "none") + ylab("Infection intensity") +
    xlab("Scaled Behavioral Vigor") + theme(text = element_text(size = 18))
```

```
## Conditions used in construction of plot
## Sex: F
## ScVarvelBef: -0.5119451
## ScRPLength: 0.04746859
## Treatment: UNTOUCHED
## wormJump: 2
```

```
# Print the graph
print(InfIntbyVig)
```
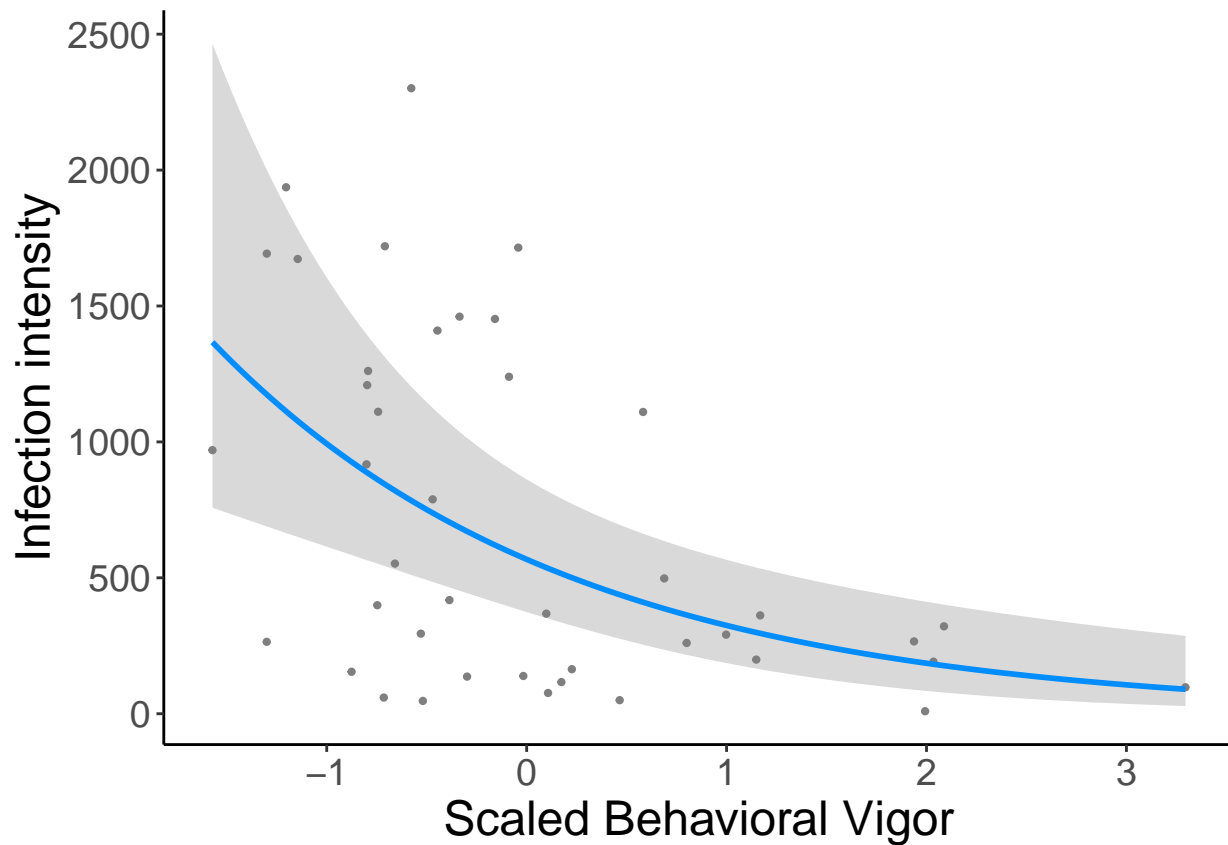
```
# Behavioral Vigor graph Extracting fit and residuals from model
InfIntbyLen = visreg(AUCLM, scale = "response", "ScRPLength", "Sex", partial = T,
    plot = FALSE, overlay = T)
# Extracting fit
InfIntbyLenfit <- InfIntbyLen$fit
# Extracting residuals
InfIntbyLenres <- InfIntbyLen$res
cpsex = c("darkseagreen4", "lightsalmon3")
InfIntbyLengraph <- ggplot() + geom_smooth(data = InfIntbyLenfit, aes(x = ScRPLength,
    y = visregFit, group = Sex, color = Sex)) + geom_ribbon(data = InfIntbyLenfit,
    aes(x = ScRPLength, y = visregFit, ymin = visregLwr, ymax = visregUpr, fill = Sex),
    alpha = 0.5) + geom_point(data = InfIntbyLenres, aes(x = ScRPLength, y = visregRes,
    group = Sex, color = Sex)) + theme_classic() + theme(legend.position = "none") +
    ylab("Infection intensity") + xlab("Scaled Residual body length") + theme(text = element_text(size =
    scale_color_manual(values = cpsex) + scale_fill_manual(values = cpsex) + ylim(0,
    7000)
# Print the graph
print(InfIntbyLengraph)
```

```
## 'geom_smooth()' using method = 'loess' and formula = 'y ~ x'
```
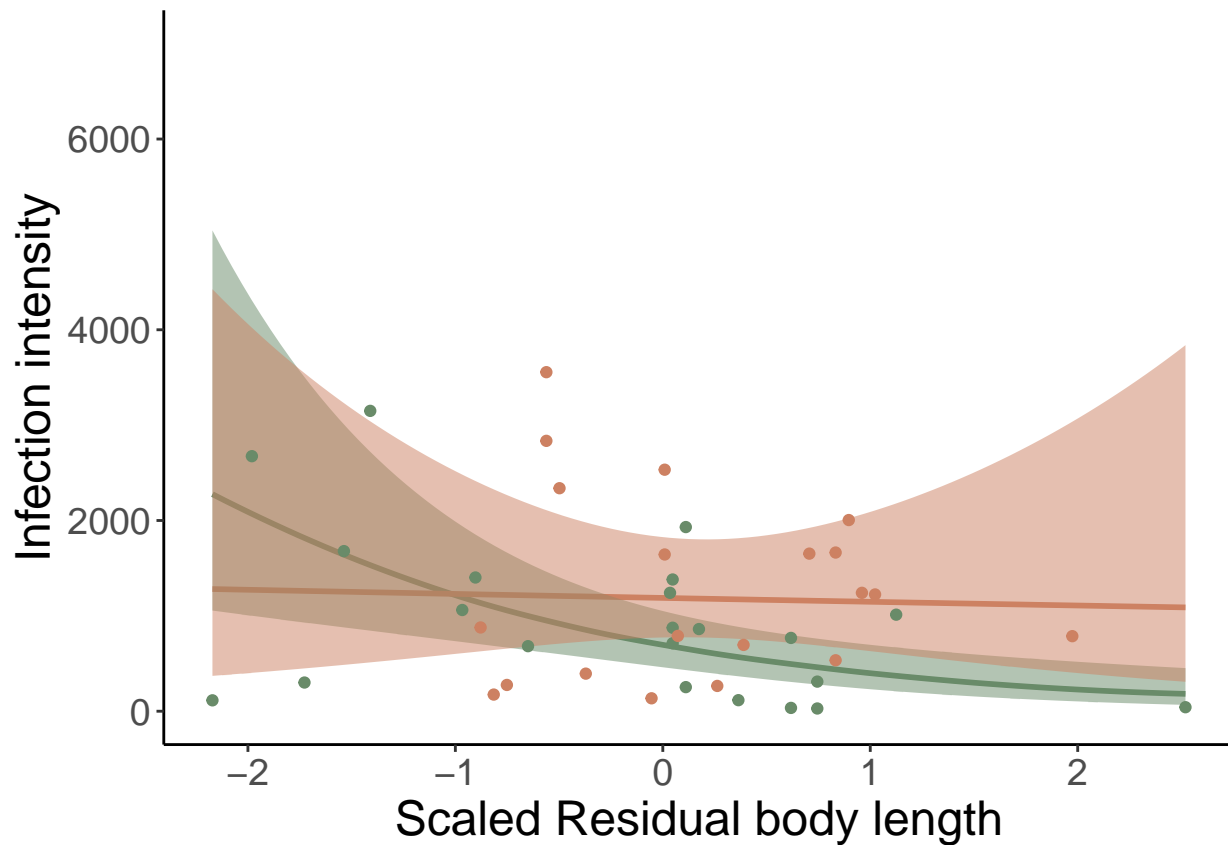
```r
# Sex by SMI graph Extracting the fit and residuals
InfIntbySex = visreg(AUCLM, scale = "response", "Sex", partial = T, plot = FALSE)
# Extract fit
InfIntbySexFit <- InfIntbySex$fit
# Extract residuals
InfIntbySexRes <- InfIntbySex$res
# Add colorscheme by sex
cpsex = c("darkseagreen4", "lightsalmon3")
# Plot the graph
InfSexgraph <- ggplot(data = InfIntbySexRes, aes(x = Sex, y = visregRes)) + geom_boxplot(aes(fill = Sex
    geom_jitter(alpha = 0.5) + theme_classic() + theme(legend.position = "none") +
    ylab("Infection intensity") + xlab(" ") + theme(text = element_text(size = 18)) +
    scale_fill_manual(values = cpsex) + scale_x_discrete(breaks = c("F", "M"), labels = c("Females",
    "Males"))
# Loading in gridExtra for multiple graphs in one image
library(gridExtra)

#
InfIntgraphs <- grid.arrange(InfIntbyLengraph, InfIntbyVig, nrow = 1)
```

```
## 'geom_smooth()' using method = 'loess' and formula = 'y ~ x'
```
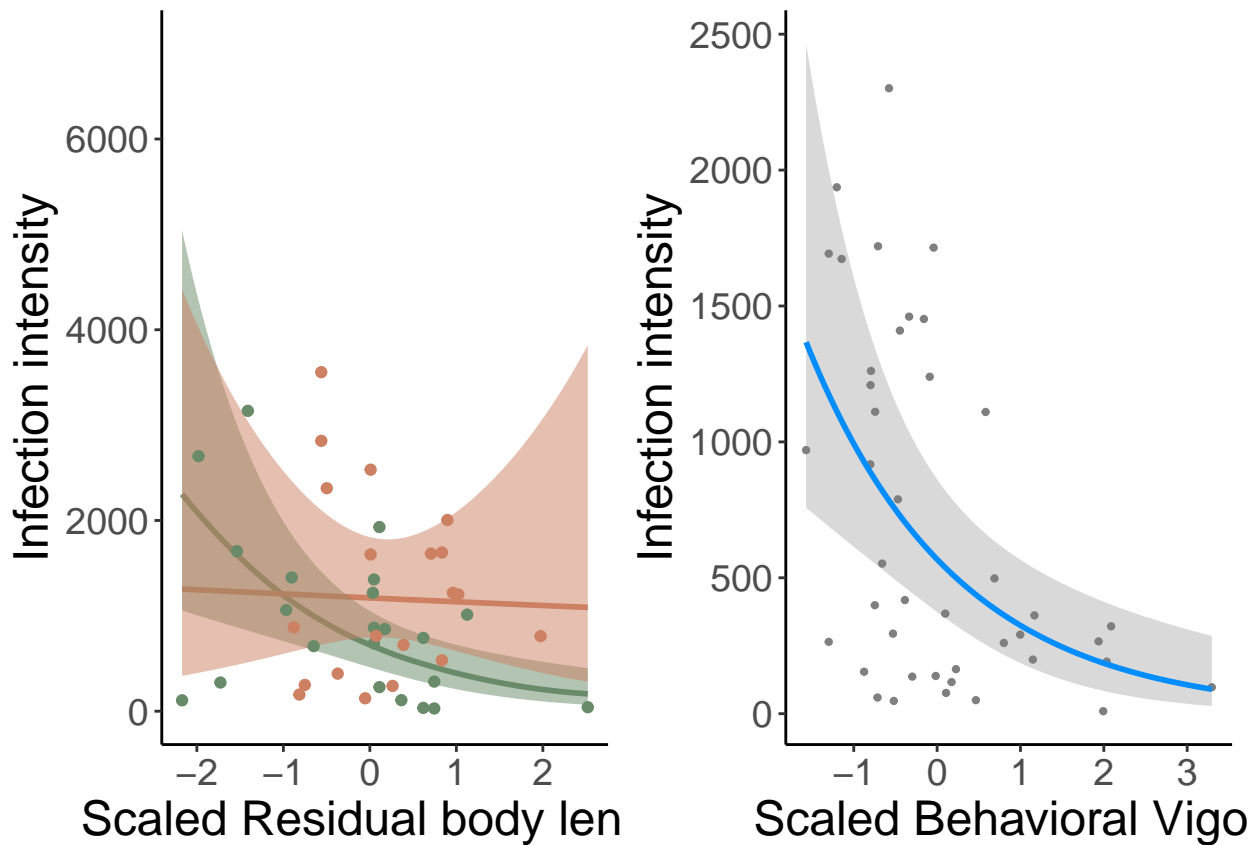
Fitting posthoc model to test for pattern within sex

```r
# Subsetting down to females only
IndBehavIF <- IndBehavI %>%
    filter(Sex == "F")
```

```
## filter: removed 23 rows (48%), 25 rows remaining
```

```r
# Subsetting down to males only
IndBehavIM <- IndBehavI %>%
    filter(Sex == "M")
```

```
## filter: removed 25 rows (52%), 23 rows remaining
```

```r
# Fit a linear model for infection intensity by sex Female model
AUCLMF <- glm(AUC ~ ScBehavVig + ScVarvelBef + ScResidPLSMI + ScRPLength + Treatment,
    family = Gamma(link = "log"), IndBehavIF)
# Male model
AUCLMM <- glm(AUC ~ ScBehavVig + ScVarvelBef + ScResidPLSMI + ScRPLength + Treatment,
    family = Gamma(link = "log"), IndBehavIM)

# Summary to see the relationship of the variables.  Females
summary(AUCLMF)
```

```
##
## Call:
## glm(formula = AUC ~ ScBehavVig + ScVarvelBef + ScResidPLSMI +
##     ScRPLength + Treatment, family = Gamma(link = "log"), data = IndBehavIF)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   6.3853     0.2365  27.004 5.12e-16 ***
## ScBehavVig   -0.2783     0.1928  -1.444 0.166036
## ScVarvelBef   0.3585     0.3033   1.182 0.252570
## ScResidPLSMI -0.4320     0.2557  -1.689 0.108417
## ScRPLength   -0.6701     0.1582  -4.237 0.000496 ***
## TreatmentVIE  0.3790     0.3909   0.970 0.345034
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.6339407)
##
##     Null deviance: 44.276  on 23  degrees of freedom
## Residual deviance: 35.674  on 18  degrees of freedom
##   (1 observation deleted due to missingness)
## AIC: 376.28
##
## Number of Fisher Scoring iterations: 22
```

```r
# Males
summary(AUCLMM)
```

```
##
## Call:
## glm(formula = AUC ~ ScBehavVig + ScVarvelBef + ScResidPLSMI +
##     ScRPLength + Treatment, family = Gamma(link = "log"), data = IndBehavIM)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.21005    0.25048  28.785 3.28e-15 ***
## ScBehavVig  -0.69585    0.26406  -2.635   0.0180 *
## ScVarvelBef  0.51982    0.25857   2.010   0.0616 .
## ScResidPLSMI -0.27102    0.19305  -1.404   0.1795
## ScRPLength   0.03954    0.27736   0.143   0.8884
## TreatmentVIE -0.30223    0.44366  -0.681   0.5055
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.701784)
##
##     Null deviance: 23.688  on 21  degrees of freedom
## Residual deviance: 18.147  on 16  degrees of freedom
##   (1 observation deleted due to missingness)
## AIC: 368.34
##
## Number of Fisher Scoring iterations: 9
```

```r
# Testing for significance for females
Anova(AUCLMF, type = 3, test = "LR")
```

```
## Analysis of Deviance Table (Type III tests)
##
## Response: AUC
##            LR Chisq Df Pr(>Chisq)
## ScBehavVig    1.5676  1   0.210552
## ScVarvelBef   1.5190  1   0.217771
## ScResidPLSMI  3.1025  1   0.078170 .
## ScRPLength    8.3693  1   0.003816 **
## Treatment     0.7356  1   0.391066
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Testing for significance for males
Anova(AUCLMM, type = 3, test = "LR")
```

```
## Analysis of Deviance Table (Type III tests)
##
## Response: AUC
##            LR Chisq Df Pr(>Chisq)
## ScBehavVig    6.3953  1   0.01144 *
## ScVarvelBef   3.9527  1   0.04680 *
## ScResidPLSMI  1.3761  1   0.24077
## ScRPLength    0.0158  1   0.90005
## Treatment     0.3521  1   0.55291
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Examining change in activity for only infected individuals

```r
# Setting some variables as factors to confirm they are this way for the
# analysis
IndBehavChI$Sex <- as.factor(IndBehavChI$Sex)
IndBehavChI$Infection <- as.factor(IndBehavChI$Infection)
IndBehavChI$Treatment <- as.factor(IndBehavChI$Treatment)

# Subsetting our dataframe to relevant data and making some
IndBehavChI <- IndBehavChI %>%
    mutate(ScWormaf = scale(Wormaf)) %>%
    mutate(ScChBehav = scale(ChBehav))
```

```
## mutate: new variable 'ScWormaf' (double) with 43 unique values and 53% NA
```

```
## mutate: new variable 'ScChBehav' (double) with 135 unique values and 7% NA
```

```r
pairs(~ChBehav + Sex + Infection + TrialTime + PreSMI + AUC + Treatment + NRatebf +
    Wormaf, lower.panel = panel.smooth, diag.panel = panel.hist, upper.panel = panel.cor,
    data = IndBehavChI)
```

```
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter

## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero
## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero

## Warning in par(usr): argument 1 does not name a graphical parameter

## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero
## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero

## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter

## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero
## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero

## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter

## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero
## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero

## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter

## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero
## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero

## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
```
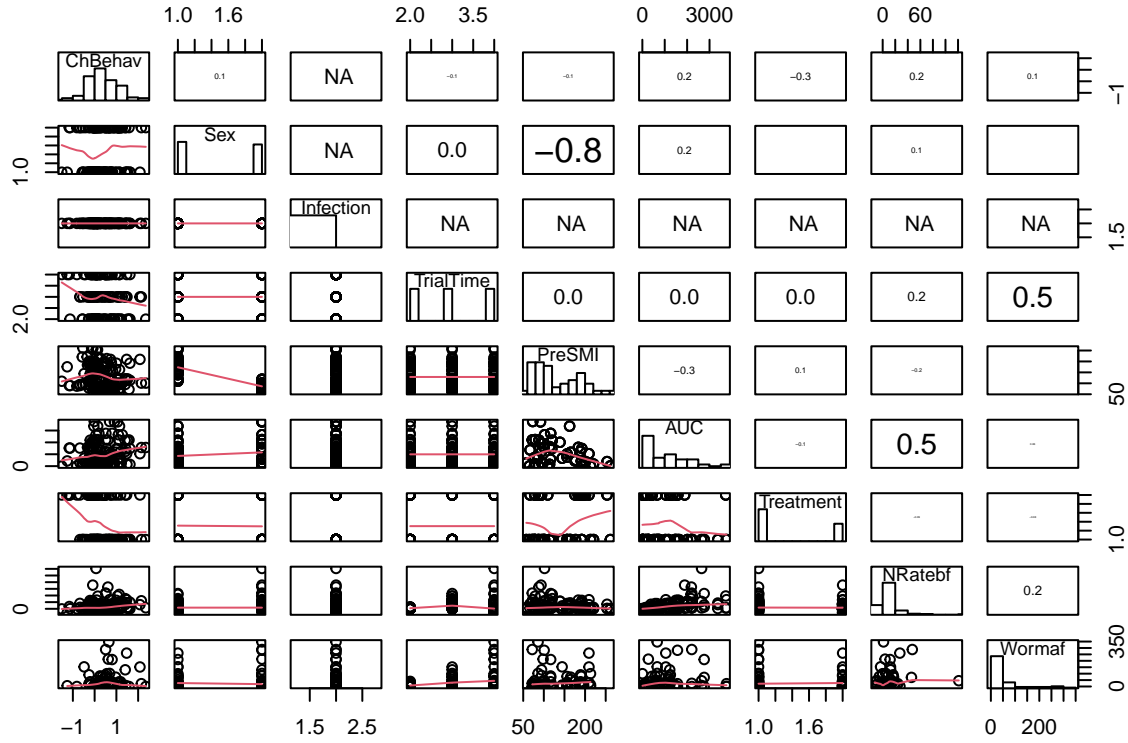
```
## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero
## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero


## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter


## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero
## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero


## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter


## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero
## Warning in cor(x, y, use = "pairwise.complete.obs"): the standard deviation is
## zero


## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
```

cpsex = c("darkseagreen4", "lightsalmon3")

We will use a linear mixed model to analyze how Change in activity differs by sexual variation and how infected an individual fish was. FishID is included as a random term to allow for non-independence of individuals due to multiple measurements per individual across time.

- Deterministic

- $Log_{10}(InfInt_{det} + 1) = a + b_1 \text{TrialTime} + b_2 \text{ChBehav} + b_3 \text{Sex} + b_4 \text{ScPSMI} + b_5 \text{Treatment} + b_6 \text{wormJump} + b_7 \text{ScChSMI} + b_8 \text{BehavVig} + a_i$

- Stochastic

  - $\text{ChVel} \sim N(Log_{10}(InfInt_{det} + 1),\ \sigma^2)$
  - $a_i \sim N(0,\ \sigma^2_{fishID})$

- Fixed

  - TrialTime
  - Change in behavior between time points
  - Sex
  - Scaled pre-infection scaled mass index
  - VIE treatment of the fish
  - number of worms that initially started the infection
  - Scaled change in scaled mass index
  - Behavior vigor of each fish

- Random

  - fishID

```
# Fit a linear model for checking what explanatory factors are important for
# Variance in activity Note this is a linear mixed model because we have
# multiple measures per fish and therefore, need to account for
# non-independence between measures.
InfIntLMInf <- lmer(log10(Wormaf + 1) ~ Sex + ScPSMI + ScChSMI + TrialTime + ScBehavVig +
    Treatment + ChBehav + wormJump + (1 | fishID), IndBehavChI)


# Summary to see the relationship of the variables.
summary(InfIntLMInf)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: log10(Wormaf + 1) ~ Sex + ScPSMI + ScChSMI + TrialTime + ScBehavVig +
##      Treatment + ChBehav + wormJump + (1 | fishID)
##    Data: IndBehavChI
##
## REML criterion at convergence: 129.3
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.03509 -0.69817 -0.04025  0.72865  1.60854
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  fishID   (Intercept) 0.005027 0.0709
##  Residual             0.406932 0.6379
## Number of obs: 61, groups:  fishID, 21
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  0.55319    0.41074   1.347
## SexM        -0.11606    0.18735  -0.619
## ScPSMI      -0.06753    0.14011  -0.482
## ScChSMI     -0.07540    0.13309  -0.567
## TrialTime2   0.77036    0.19797   3.891
## TrialTime3   1.35524    0.20857   6.498
## ScBehavVig   0.07273    0.13822   0.526
## TreatmentVIE 0.20701    0.22032   0.940
## ChBehav      0.45893    0.19621   2.339
## wormJump    -0.18333    0.17373  -1.055
##
## Correlation of Fixed Effects:
##            (Intr) SexM   ScPSMI ScCSMI TrlTm2 TrlTm3 ScBhvV TrtVIE ChBehv
## SexM        0.015
## ScPSMI      0.228 -0.281
## ScChSMI     0.193 -0.100  0.559
## TrialTime2 -0.197  0.018 -0.008 -0.020
## TrialTime3 -0.334 -0.011  0.007  0.016  0.444
## ScBehavVig  0.149 -0.018  0.523  0.334 -0.049  0.095
## TreatmntVIE -0.377  0.004 -0.440 -0.412  0.007 -0.023 -0.541
## ChBehav     -0.388 -0.168  0.072  0.190 -0.105  0.240  0.468 -0.068
## wormJump    -0.860 -0.256 -0.073 -0.051 -0.029  0.074 -0.118  0.175  0.277
```
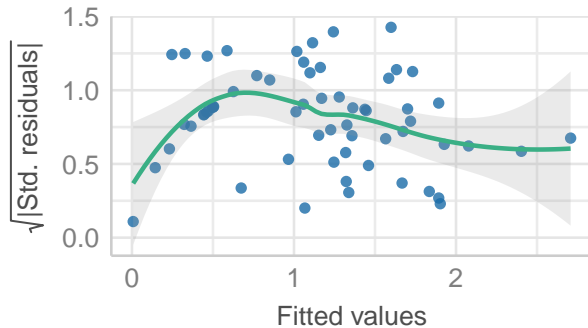
Validate that the model fits well and there are no problems

```
# Using the check_model function from the perforamnce package to check the
# model validation

check_model(InfIntLMInf, check = c("qq", "normality", "homogeneity"))
```
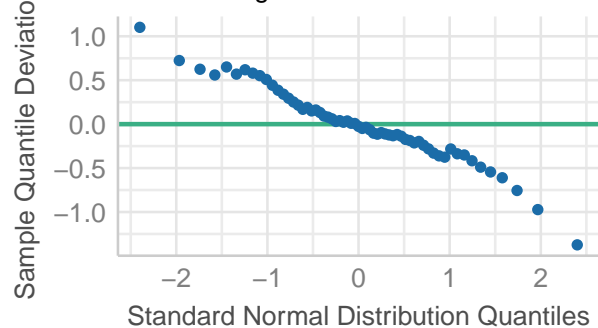
Homogeneity of Variance
Reference line should be flat and horizontal
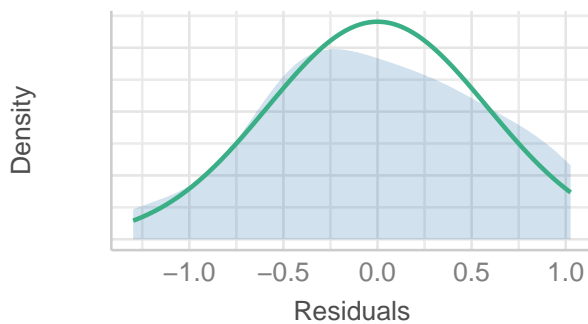

Normality of Residuals
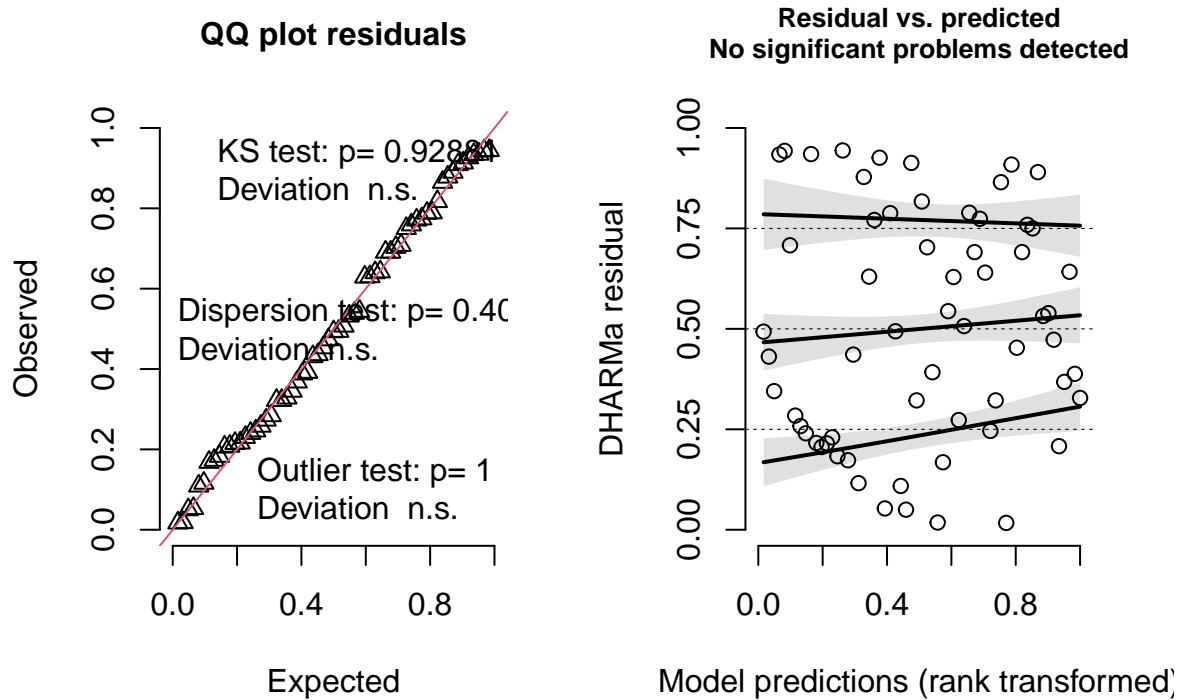Dots should fall along the line

Normality of Residuals
Distribution should be close to the normal curve


```
# Using the Dharma package to check quantile residuals First simulating the
# quantile residuals
sim_residuals_InfIntLMInf <- simulateResiduals(InfIntLMInf, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_InfIntLMInf)
```

DHARMa residual

QQ plot residuals

KS test: p= 0.92881
Deviation n.s.

Dispersion test: p= 0.40
Deviation n.s.

Outlier test: p= 1
Deviation n.s.

Observed (y-axis)
Expected (x-axis)

Residual vs. predicted
No significant problems detected

DHARMa residual (y-axis)
Model predictions (rank transformed) (x-axis)

```
# Testing for dispersion
testDispersion(sim_residuals_InfIntLMInf)
```

DHARMa nonparametric dispersion test via sd of
residuals fitted vs. simulated

Frequency (y-axis)
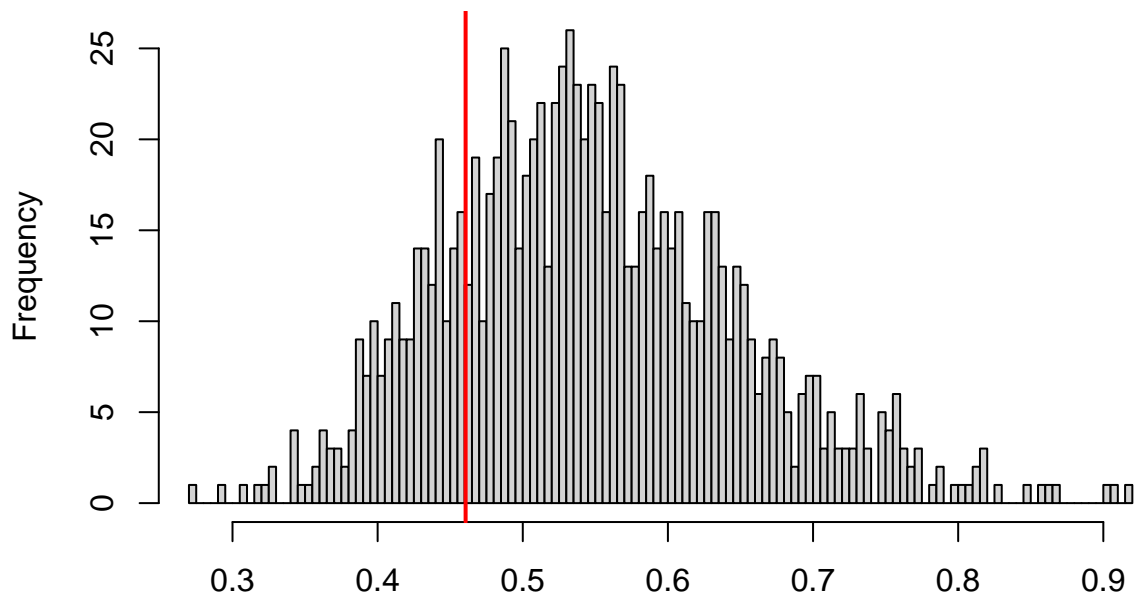
Simulated values, red line = fitted model. p−value (two.sided) = 0.408

```
##
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
##
## data:  simulationOutput
## dispersion = 0.84272, p-value = 0.408
## alternative hypothesis: two.sided
```

```
# There are some problems with this model validation. It doesnt look model
# breaking but definitely should look at other model error structures to
# resolve the issues.
```

Testing the significance of factors in our model using a Kenward-Rodgers F test for infected only infection intensity model

```
# F test to test for signficance of slope of variables
Anova(InfIntLMInf, test = "F", type = 3)
```

```
## Analysis of Deviance Table (Type III Wald F tests with Kenward-Roger df)
##
## Response: log10(Wormaf + 1)
##                   F Df Df.res    Pr(>F)
## (Intercept)  1.7844  1 17.834   0.19841
## Sex          0.3819  1 13.514   0.54686
## ScPSMI       0.2320  1 12.774   0.63816
## ScChSMI      0.3180  1 15.068   0.58110
## TrialTime   21.5646  2 40.186 4.341e-07 ***
## ScBehavVig   0.2706  1 16.492   0.60982
## Treatment    0.8818  1 12.696   0.36524
## ChBehav      4.9482  1 42.669   0.03146 *
## wormJump     1.1043  1 13.713   0.31148
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Visualize the important explanatory factors for infection intensity for infected individuals

```
# Change in behavior by Extracting fit and residual
ChVelbyRate = visreg(InfIntLMInf, scale = "response", "ChBehav", partial = T, overlay = TRUE,
    plot = FALSE)
# Extracting fit
ChVelbyRatefit <- ChVelbyRate$fit
# Extracting residual
ChVelbyRateres <- ChVelbyRate$res
# Constructing the graph
ChVelbyRategraph <- ggplot(ChVelbyRateres, aes(x = ChBehav, y = visregRes)) + geom_smooth(method = "lm")
    geom_point() + theme_classic() + theme(legend.position = "none") + ylab("Log (number of parasites +
    xlab("Change in activity") + theme(text = element_text(size = 18))

# Print the graph
print(ChVelbyRategraph)
```

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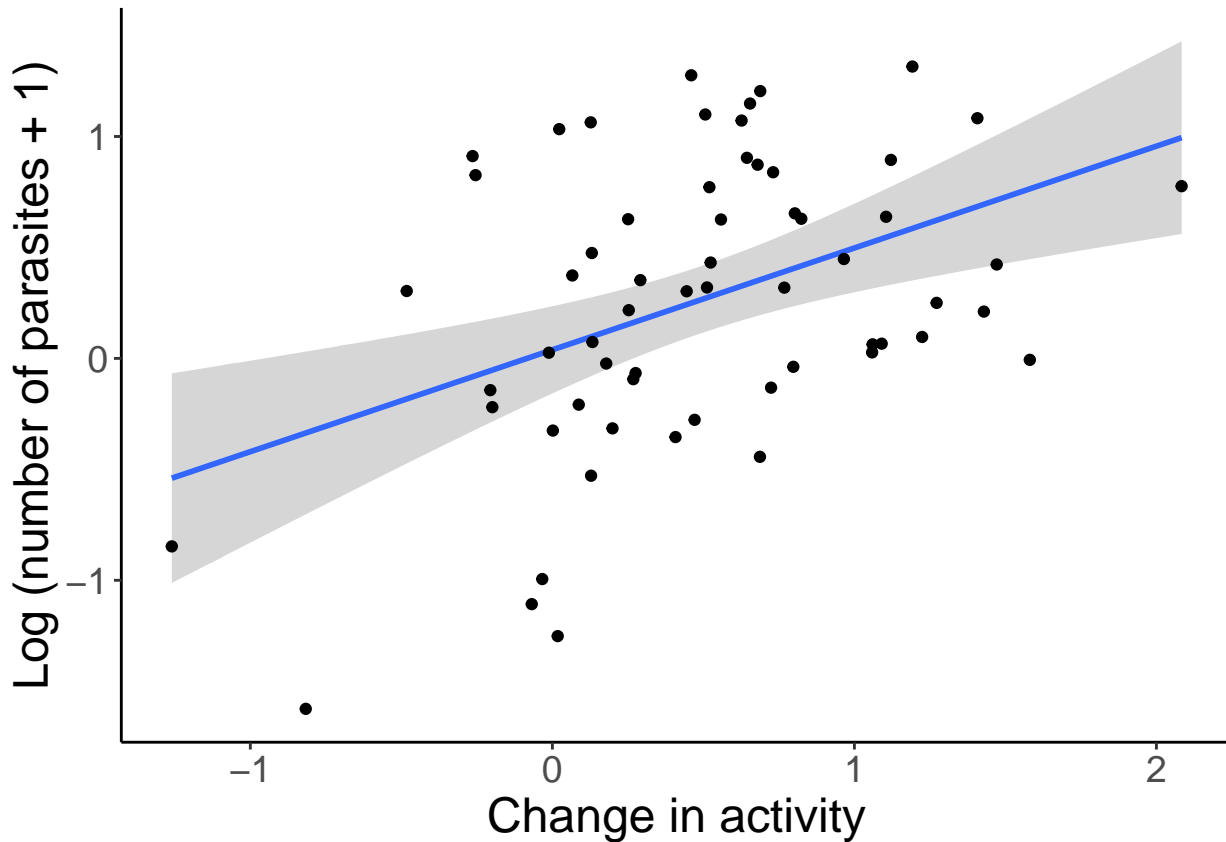


```
# Subset to only infected females
IndBehavChIF <- IndBehavChI %>%
    filter(Sex == "F" & Infection == "1") %>%
    mutate(ScWormaf = scale(Wormaf))
```

```
## filter: removed 69 rows (48%), 75 rows remaining
```

```
## mutate: changed 32 values (43%) of 'ScWormaf' (0 new NA)
```

```
# Subset to only infected males
IndBehavChIM <- IndBehavChI %>%
    filter(Sex == "M" & Infection == "1") %>%
    mutate(ScWormaf = scale(Wormaf))
```

```
## filter: removed 75 rows (52%), 69 rows remaining
```

```
## mutate: changed 36 values (52%) of 'ScWormaf' (0 new NA)
```

```
# Fit linear mixed models to post hoc check for significance for male and
# female slopes
ChVelLMInfF <- lmer(ChBehav ~ ScPSMI + ScChSMI + TrialTime + ScBehavVig + Treatment +
    ScNRatebf + (1 | fishID), IndBehavChIF)
```

```r
# Fit linear mixed models to post hoc check for significance for male and
# female slopes
ChVelLMInfM <- lmer(ChBehav ~ ScPSMI + ScChSMI + TrialTime + ScBehavVig + Treatment +
    ScNRatebf + (1 | fishID), IndBehavChIM)

# Summary to see the relationship of the variables for female model.
summary(ChVelLMInfF)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: ChBehav ~ ScPSMI + ScChSMI + TrialTime + ScBehavVig + Treatment +
##     ScNRatebf + (1 | fishID)
##    Data: IndBehavChIF
##
## REML criterion at convergence: 115.2
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.40025 -0.59721  0.01767  0.39831  2.49141
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  fishID   (Intercept) 0.1516   0.3893
##  Residual             0.1465   0.3827
## Number of obs: 74, groups:  fishID, 25
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   0.35892    0.13500   2.659
## ScPSMI        0.03171    0.19273   0.165
## ScChSMI      -0.12455    0.20790  -0.599
## TrialTime2   -0.12062    0.12073  -0.999
## TrialTime3   -0.08185    0.11226  -0.729
## ScBehavVig   -0.29924    0.09676  -3.093
## TreatmentVIE -0.14657    0.20917  -0.701
## ScNRatebf     0.20238    0.07599   2.663
##
## Correlation of Fixed Effects:
##            (Intr) ScPSMI ScCSMI TrlTm2 TrlTm3 ScBhvV TrtVIE
## ScPSMI      0.239
## ScChSMI    -0.110 -0.849
## TrialTime2 -0.419 -0.040  0.004
## TrialTime3 -0.423 -0.024  0.008  0.525
## ScBehavVig  0.035 -0.134  0.264 -0.057 -0.030
## TreatmntVIE -0.553 -0.435  0.344 -0.012  0.010  0.015
## ScNRatebf   0.134  0.090 -0.008 -0.443 -0.210  0.128  0.026
```

```r
# Summary to see the relationship of the variables for male model.
summary(ChVelLMInfM)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: ChBehav ~ ScPSMI + ScChSMI + TrialTime + ScBehavVig + Treatment +
##     ScNRatebf + (1 | fishID)
```

```
##     Data: IndBehavChIM
##
## REML criterion at convergence: 87.2
##
## Scaled residuals:
##      Min       1Q    Median       3Q       Max
## -1.77079 -0.56125 -0.08803  0.54617  2.27980
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  fishID   (Intercept) 0.1907   0.4366
##  Residual             0.1050   0.3240
## Number of obs: 60, groups:  fishID, 22
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   0.55518    0.14960   3.711
## ScPSMI       -0.13224    0.18568  -0.712
## ScChSMI      -0.10664    0.17887  -0.596
## TrialTime2    0.13984    0.10301   1.358
## TrialTime3   -0.28284    0.12219  -2.315
## ScBehavVig   -0.42170    0.10922  -3.861
## TreatmentVIE -0.19015    0.24076  -0.790
## ScNRatebf    -0.02104    0.03640  -0.578
##
## Correlation of Fixed Effects:
##            (Intr) ScPSMI ScCSMI TrlTm2 TrlTm3 ScBhV TrtVIE
## ScPSMI     -0.175
## ScChSMI    -0.120  0.769
## TrialTime2 -0.335 -0.007 -0.010
## TrialTime3 -0.302 -0.027 -0.029  0.518
## ScBehavVig  0.243  0.240 -0.007 -0.040 -0.094
## TreatmntVIE -0.603  0.134  0.221  0.003  0.013 -0.400
## ScNRatebf   0.078  0.021  0.032 -0.317 -0.439  0.127 -0.010
```

```
# Anova to test for significance for females
Anova(ChVelLMInfF, test = "F", type = 2)
```

```
## Analysis of Deviance Table (Type II Wald F tests with Kenward-Roger df)
##
## Response: ChBehav
##                 F Df Df.res   Pr(>F)
## ScPSMI     0.0271  1 19.801 0.870991
## ScChSMI    0.3589  1 19.547 0.555988
## TrialTime  0.5239  2 47.777 0.595551
## ScBehavVig 9.5583  1 20.050 0.005741 **
## Treatment  0.4909  1 19.840 0.491657
## ScNRatebf  6.8075  1 61.242 0.011393 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Anova to test for significance for males
Anova(ChVelLMInfM, test = "F", type = 2)
```

```
## Analysis of Deviance Table (Type II Wald F tests with Kenward-Roger df)
##
## Response: ChBehav
##                 F Df Df.res   Pr(>F)
## ScPSMI      0.5070  1 16.615 0.486299
## ScChSMI     0.3553  1 16.990 0.558975
## TrialTime   7.1061  2 36.341 0.002484 **
## ScBehavVig 14.8948  1 17.070 0.001250 **
## Treatment   0.6236  1 16.776 0.440748
## ScNRatebf   0.3246  1 41.558 0.571916
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```