

Mixed-effects models  
*Solutions to Exercises*

May 1, 2019

# Contents

- 1 Mixed-effects models **2****
- 1.10 Exercises 2
  - 1.10.1 PREF Canopy data 2
  - 1.10.2 Mouse metabolism 8
  - 1.10.3 Litter decomposition data 17
  - 1.10.4 Repeated measures in leaf photosynthesis 23
  - 1.10.5 EucFACE ground cover data 29
  - 1.10.6 Logistic regression 35

# Chapter 1

## Mixed-effects models

### 1.10 Exercises

In these exercises, we use the following colour codes:

- **Easy:** make sure you complete some of these before moving on. These exercises will follow examples in the text very closely.
- ◆ **Intermediate:** a bit harder. You will often have to combine functions to solve the exercise in two steps.
- ▲ **Hard:** difficult exercises! These exercises will require multiple steps, and significant departure from examples in the text.

We suggest you complete these exercises in an **R** markdown file. This will allow you to combine code chunks, graphical output, and written answers in a single, easy-to-read file.

#### 1.10.1 PREF Canopy data

1. ■ Fit and compare mixed-effects models testing the fixed effects of `species` and `dfromtop` on `narea` (nitrogen concentration in grams per square metre). Evaluate whether the model fit is improved when including a random slope for each individual. Evaluate the significance of the fixed effects and their interaction.

```
library(lme4)
library(visreg)
library(car)

# read in data
pref <- read.csv("prefdata.csv")

# Random intercept only
lmer1 <- lmer(narea ~ species + dfromtop + species:dfromtop + (1|ID), data=pref)

# Random intercept and slope
lmer2 <- lmer(narea ~ species + dfromtop + species:dfromtop + (dfromtop|ID), data=pref)

# significance of random slope? Not significant
```

```

anova(lmer1, lmer2)

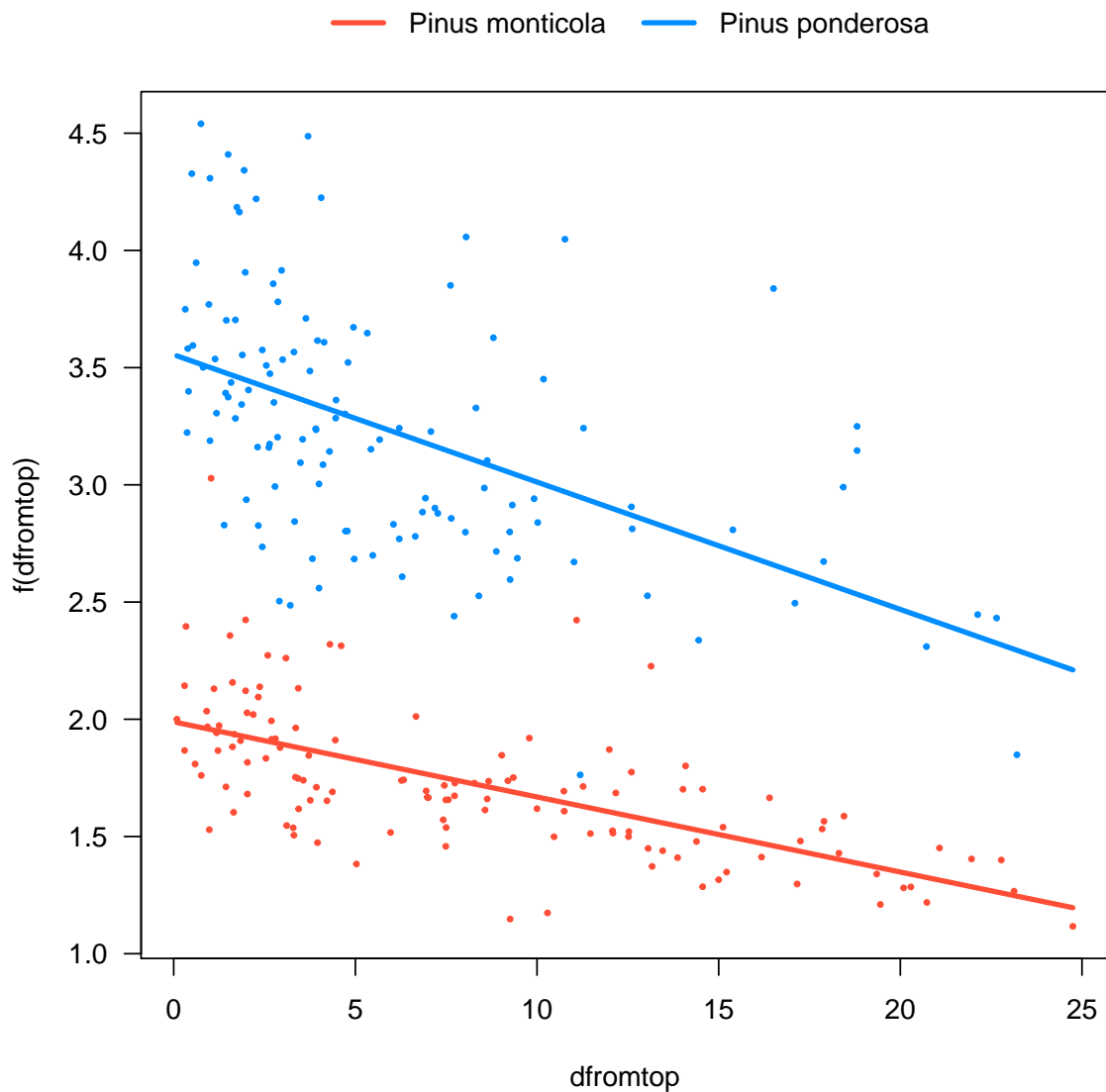
## Data: pref
## Models:
## lmer1: narea ~ species + dfromtop + species:dfromtop + (1 | ID)
## lmer2: narea ~ species + dfromtop + species:dfromtop + (dfromtop | ID)
##      Df    AIC    BIC logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer1  6 321.72 342.82 -154.86  309.72
## lmer2  8 325.36 353.49 -154.68  309.36 0.3652      2      0.8331

# significance of fixed effects?
Anova(lmer1, test='F')

## Analysis of Deviance Table (Type II Wald F tests with Kenward-Roger df)
##
## Response: narea
##              F Df  Df.res    Pr(>F)
## species      108.9832  1  33.061 5.336e-12 ***
## dfromtop       57.6841  1 238.204 6.983e-13 ***
## species:dfromtop  4.1127  1 241.119  0.04366 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

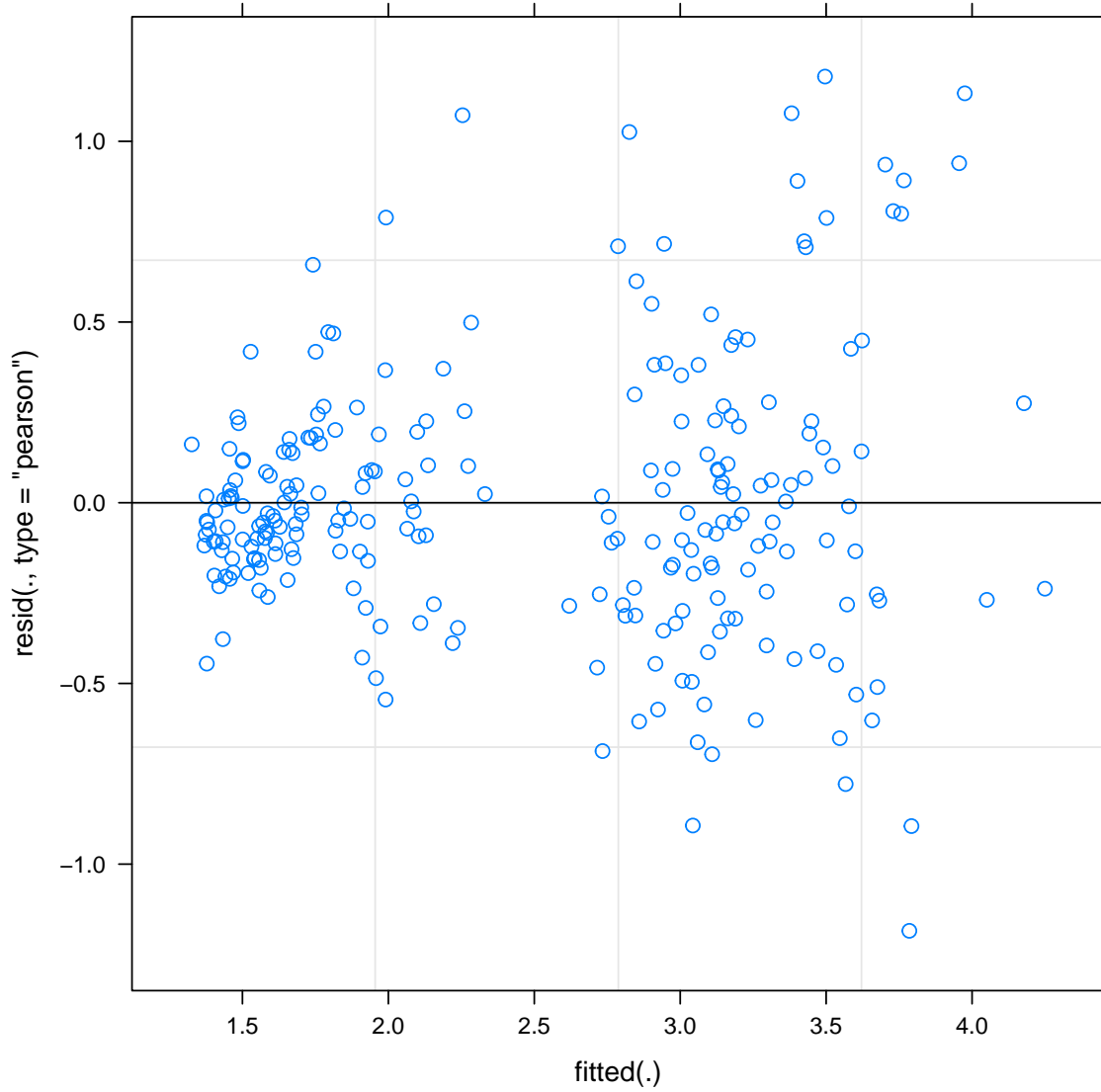
# visualise model prediction
visreg(lmer1, xvar='dfromtop', by='species', overlay=T)

```

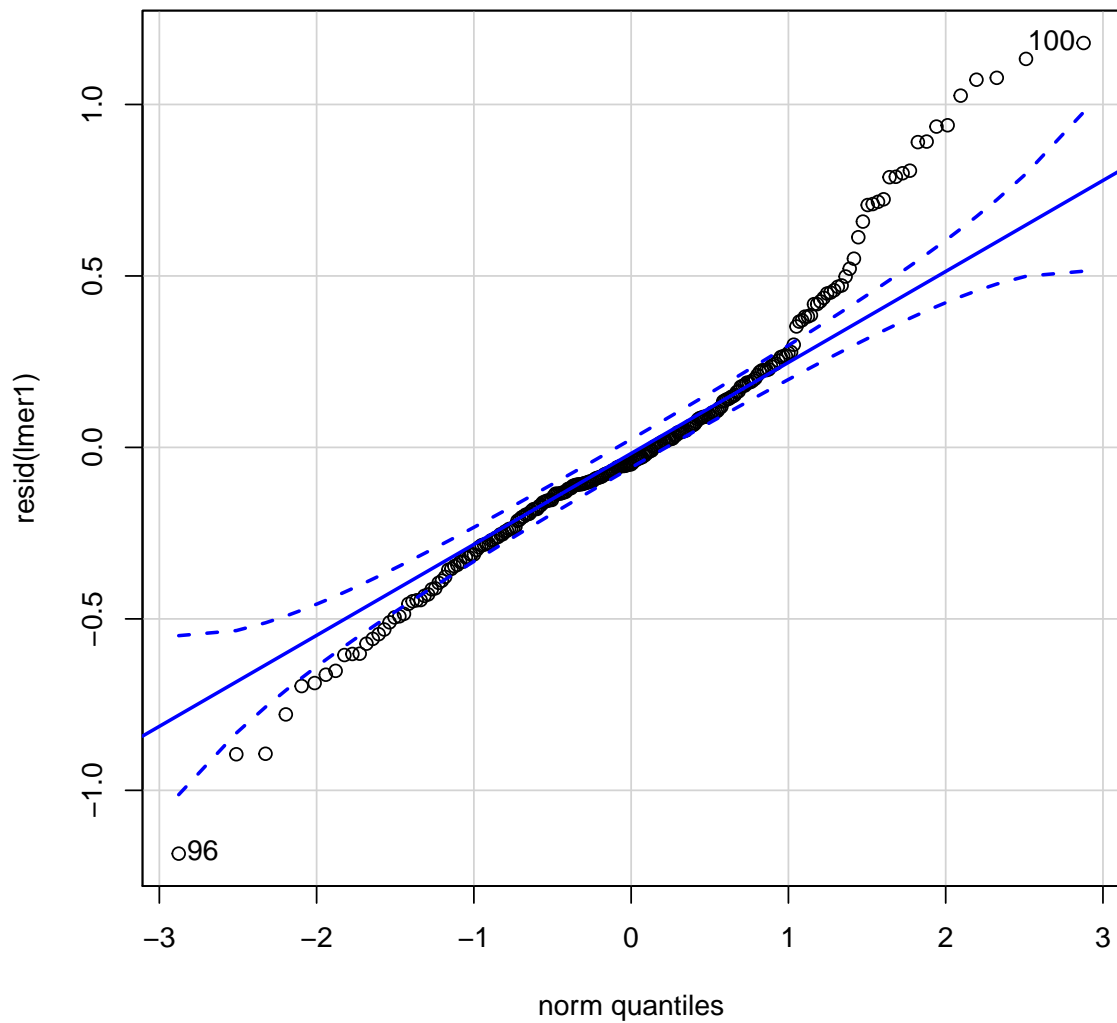


2. ■ Generate diagnostic plots for the output of the previous analysis to evaluate whether assumptions of normality and heteroscedasticity are violated. Does a sqrt- or log-transformation improve model fit? Repeat the analysis on the transformed data, using the most appropriate transformation.

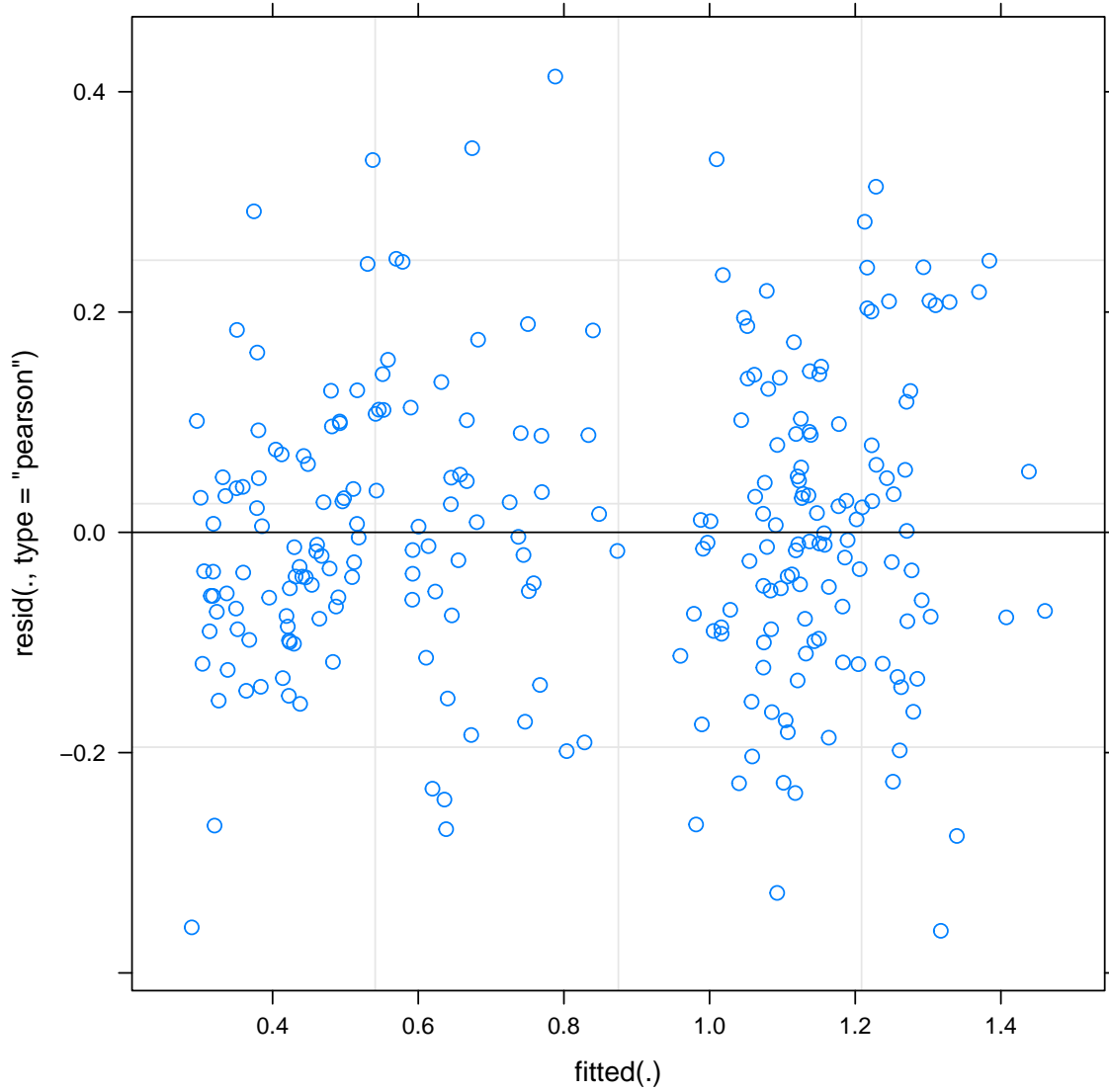
```
# residuals versus fitted plot  
plot(lmer1)
```



```
# qqplot  
qqPlot(resid(lmer1))
```

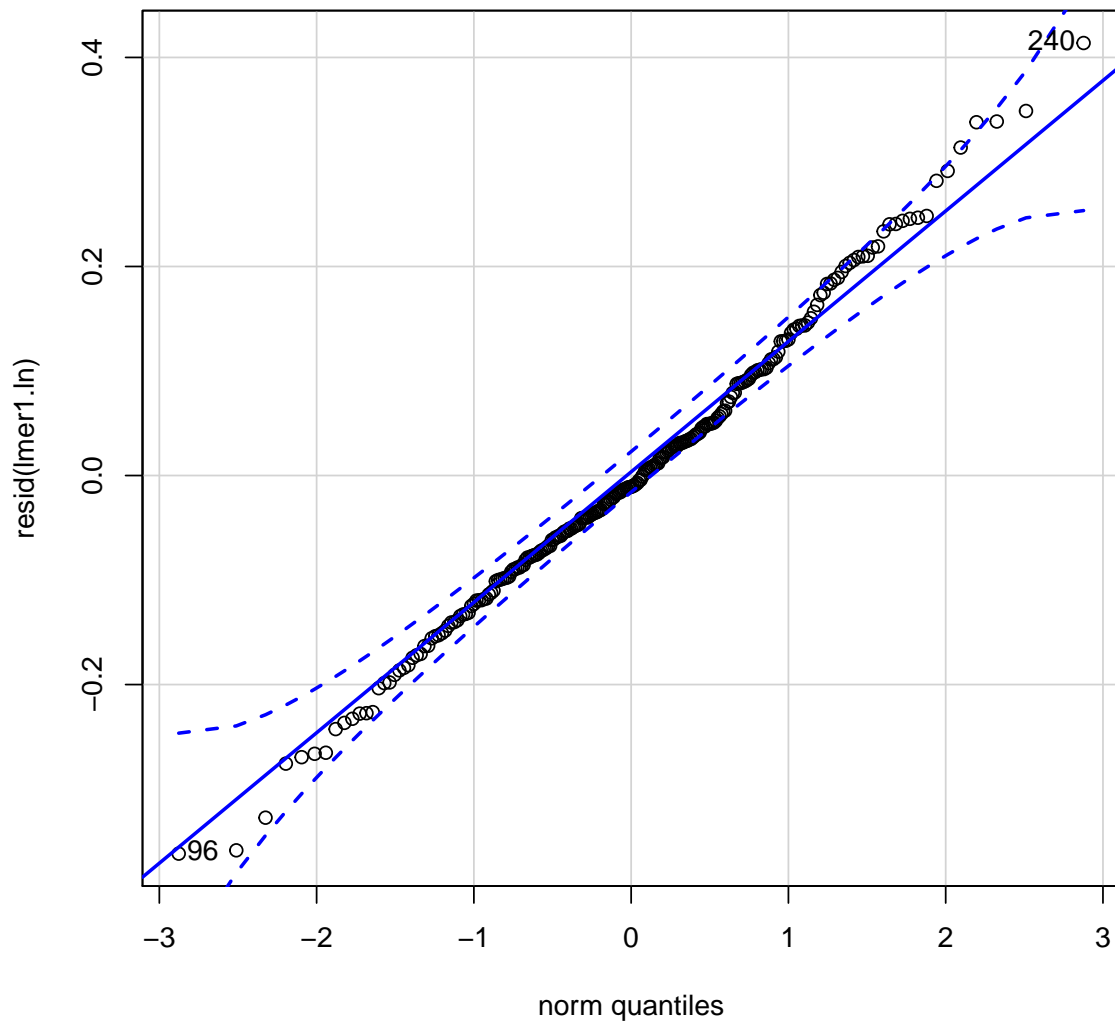


```
## [1] 96 100  
# example with log-transformation  
lmer1.ln <- lmer(log(narea) ~ species + dfromtop + species:dfromtop + (1|ID), data=pref)  
# residuals versus fitted plot  
plot(lmer1.ln)
```



```
# qqplot  
qqPlot(resid(lmer1.ln))
```





```
## [1] 240 96
```

## 1.10.2 Mouse metabolism

In this exercise you will practice with the mouse metabolism data ('wildmousemetabolism.csv').

1. ■ For the `mouse_m5` model, it was determined that `wheel` (whether or not the mice were using the wheel for exercise) was a significant predictor. What was the effect size? Use both `summary` and `visreg`.

```
mouse <- read.csv("wildmousemetabolism.csv")

# Make sure the individual label ('id') is a factor variable
mouse$id <- as.factor(mouse$id)

# recode intercept for temperature
```

```

mouse$temp31 <- mouse$temp - 31

# fit model
mouse_m5 <- lmer(rmr ~ bm*temp31 + wheel + (temp31|id/run), data=mouse)

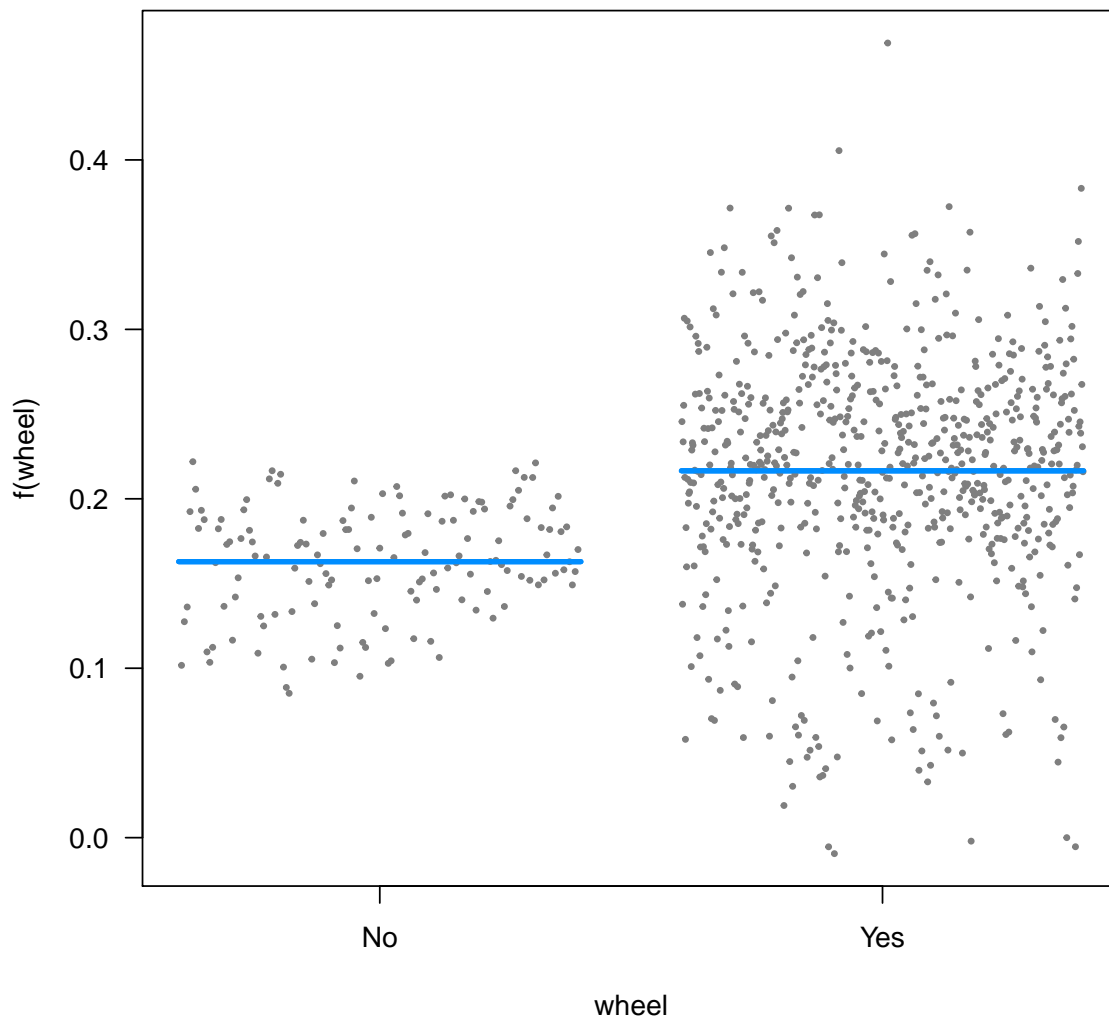
# model summary
summary(mouse_m5)

## Linear mixed model fit by REML ['lmerMod']
## Formula: rmr ~ bm * temp31 + wheel + (temp31 | id/run)
## Data: mouse
##
## REML criterion at convergence: -1992.5
##
## Scaled residuals:
##   Min       1Q   Median       3Q      Max
## -3.3436 -0.4558  0.0601  0.5774  3.7363
##
## Random effects:
##   Groups   Name                Variance Std.Dev. Corr
## run:id    (Intercept) 2.015e-04 0.014195
##           temp31      3.216e-06 0.001793 -1.00
## id       (Intercept) 0.000e+00 0.000000
##           temp31      1.454e-06 0.001206  NaN
## Residual                4.566e-03 0.067575
## Number of obs: 834, groups: run:id, 48; id, 16
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  0.0032002  0.0296136  0.108
## bm           0.0064455  0.0019549  3.297
## temp31       0.0013840  0.0033899  0.408
## wheelYes     0.0536174  0.0078597  6.822
## bm:temp31    -0.0005032  0.0002252 -2.235
##
## Correlation of Fixed Effects:
##      (Intr) bm      temp31 wheelYs
## bm      -0.979
## temp31   0.232 -0.242
## wheelYes -0.140  0.001  0.078
## bm:temp31 -0.241  0.246 -0.985  0.001
## convergence code: 0
## unable to evaluate scaled gradient
## Model failed to converge: degenerate Hessian with 1 negative eigenvalues

# visualise effect size
visreg(mouse_m5, xvar='wheel')

## Conditions used in construction of plot
## bm: 14.6
## temp31: -11
## id: 99
## run: 2

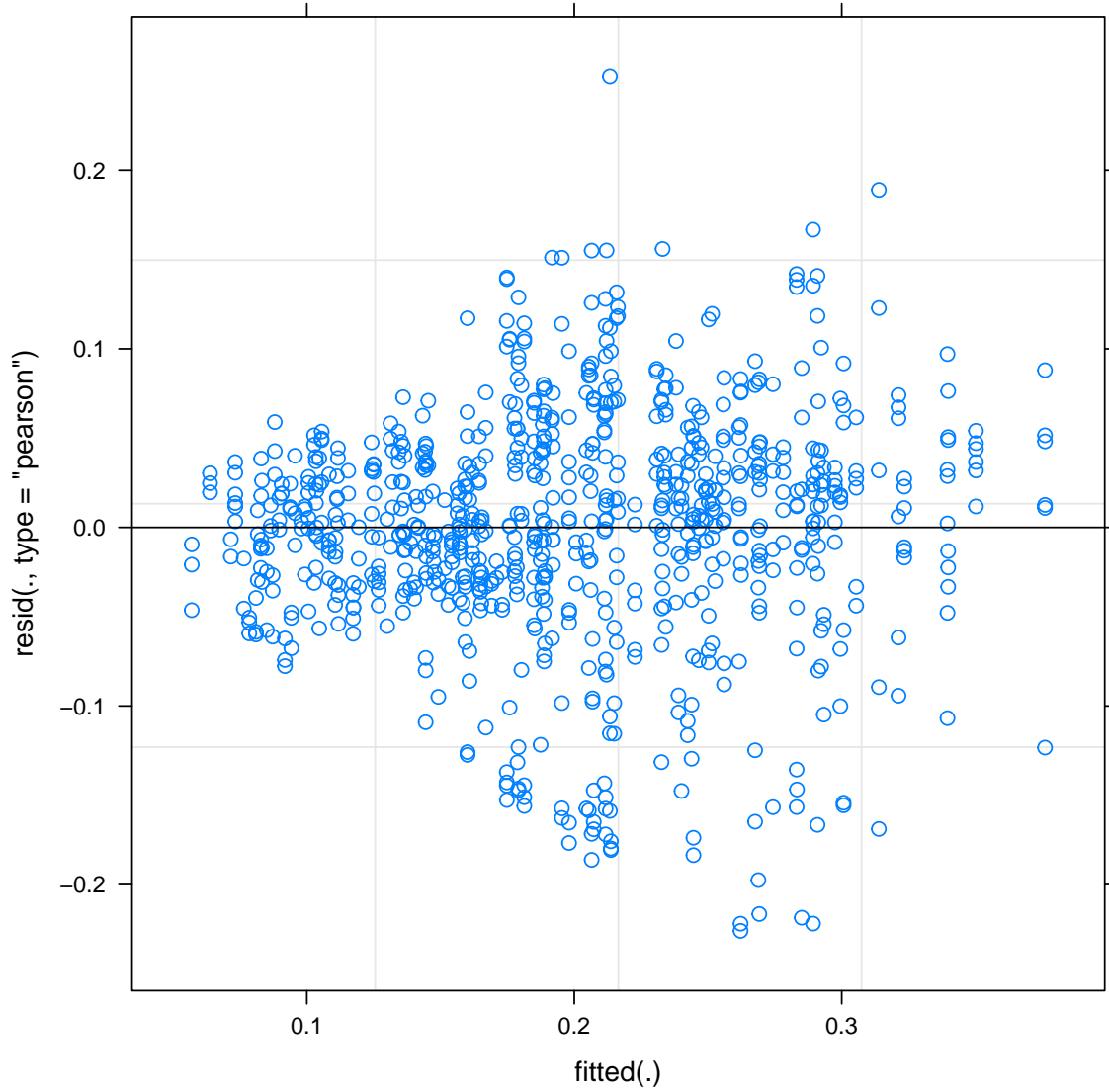
```



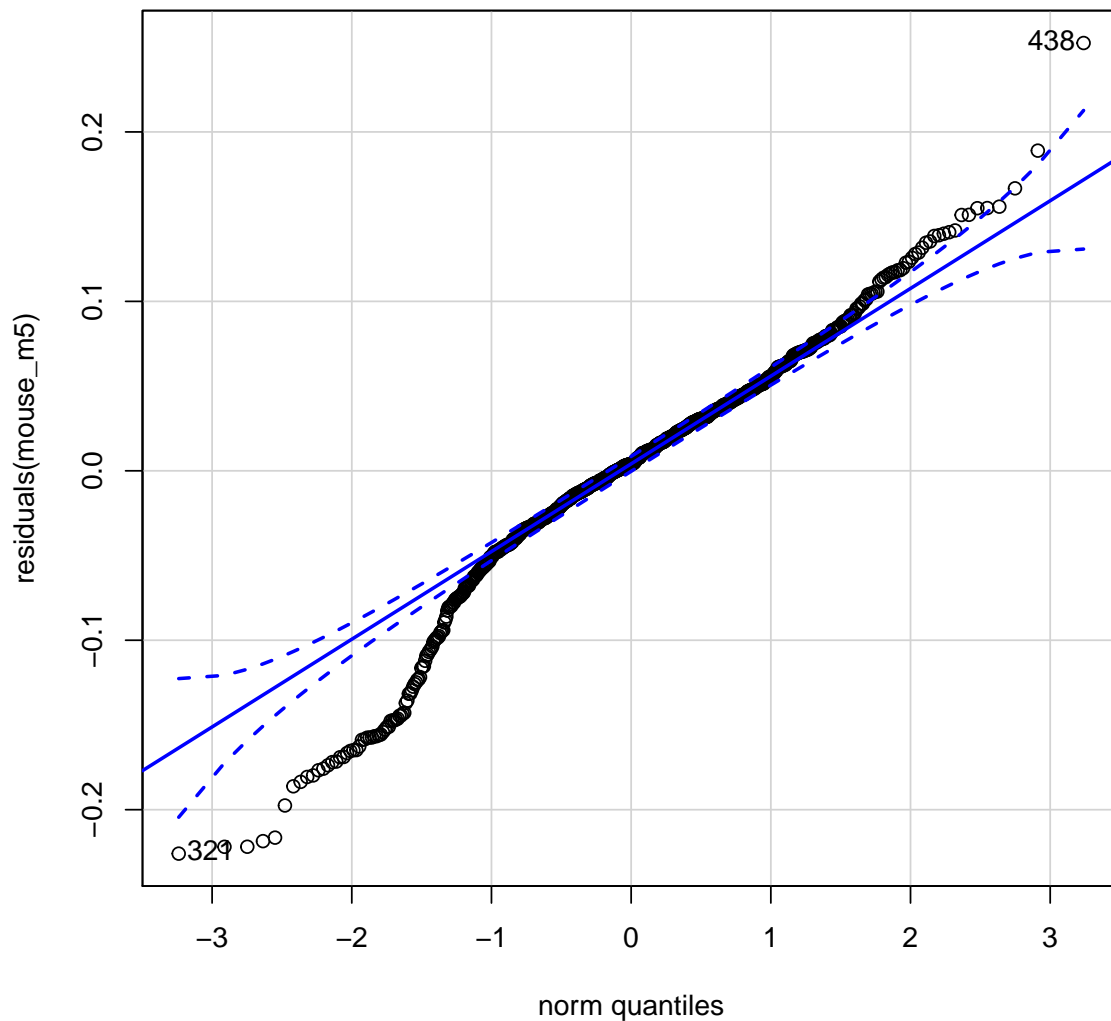
2. ♦ Inspect diagnostics for the best of the mouse models (use AIC on all models to choose the best one), including QQ plot, and residuals versus fitted. Does a logarithmic transformation (either predictor or response) improve the diagnostics?

```
mouse_m5 <- lmer(rmr ~ bm*temp31 + wheel + (temp31|id/run), data=mouse)
```

```
plot(mouse_m5)
```



```
library(car)  
qqPlot(residuals(mouse_m5))
```



```
## 438 321
## 429 318
```

3. ♦ The `mouse` dataframe also contains a variable indicating the day on which the measurements took place. Each mouse was exposed to each temperature once on each of six sequential days. Include day of measurement as another random effect, check whether the model fit is improved, and evaluate the significance of each of the fixed effects.

```
mouse_m5 <- lmer(rmr ~ bm*temp31 + wheel + (temp31|id/run), data=mouse)
mouse_m6 <- lmer(rmr ~ bm*temp31 + wheel + (temp31|id/day/run), data=mouse)
anova(mouse_m5, mouse_m6)

## Data: mouse
## Models:
## mouse_m5: rmr ~ bm * temp31 + wheel + (temp31 | id/run)
## mouse_m6: rmr ~ bm * temp31 + wheel + (temp31 | id/day/run)
##           Df      AIC      BIC logLik deviance  Chisq Chi Df Pr(>Chisq)
## mouse_m5 12 -2024.8 -1968.1 1024.4 -2048.8
```

---

```

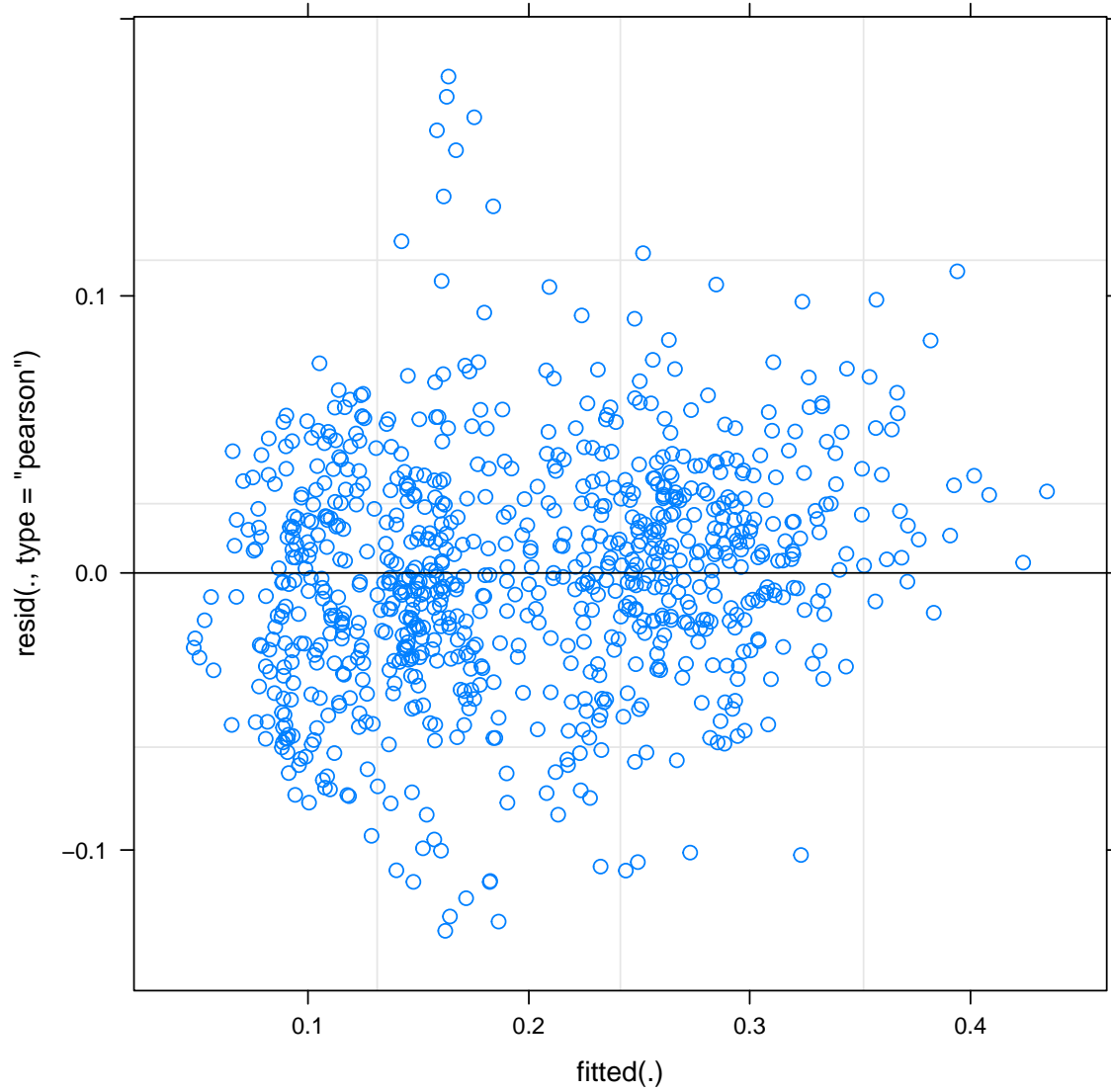
## mouse_m6 15 -2221.1 -2150.2 1125.5 -2251.1 202.27      3 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

mouse_m7 <- lmer(rmr ~ bm*temp31 + wheel + (temp31|id/day), data=mouse)
anova(mouse_m6, mouse_m7)

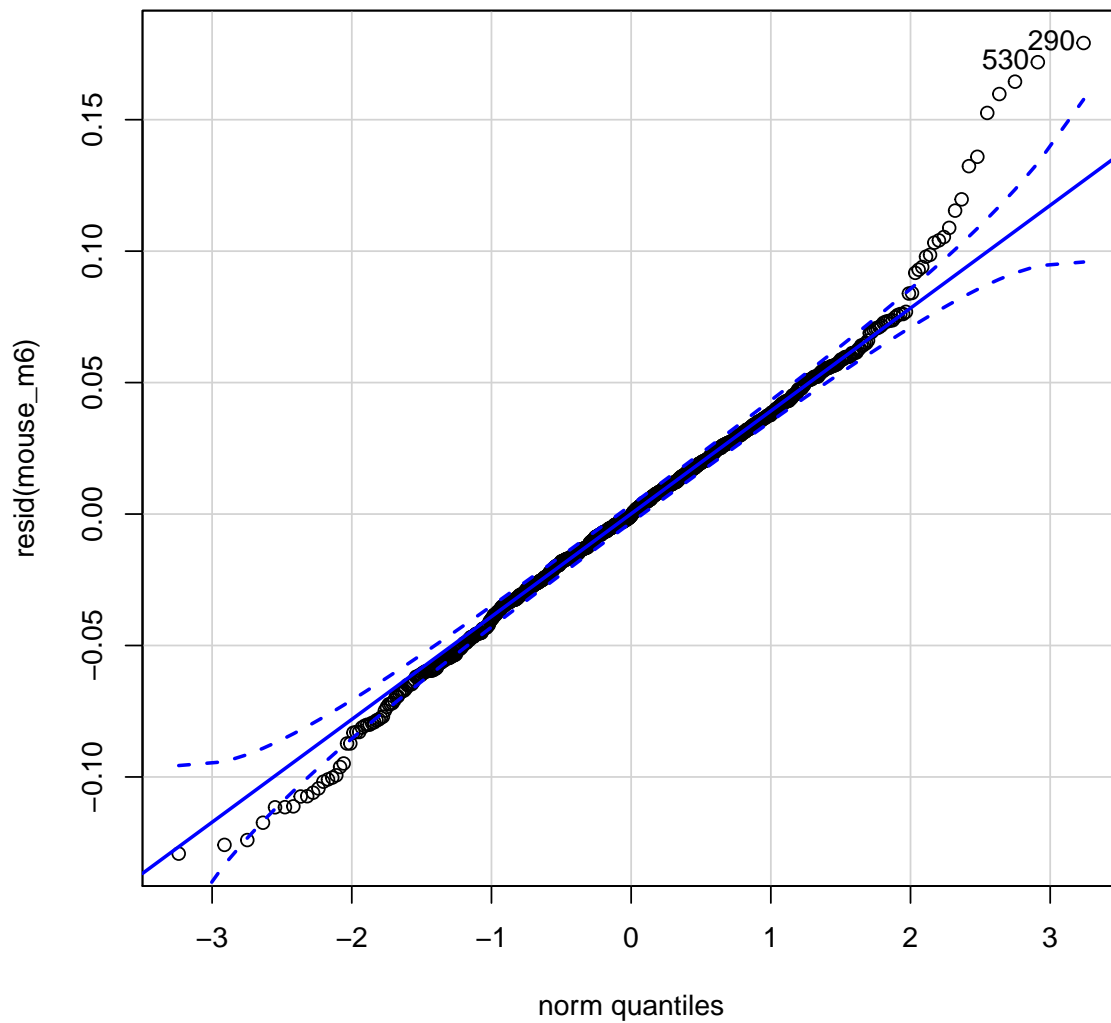
## Data: mouse
## Models:
## mouse_m7: rmr ~ bm * temp31 + wheel + (temp31 | id/day)
## mouse_m6: rmr ~ bm * temp31 + wheel + (temp31 | id/day/run)
##           Df      AIC      BIC logLik deviance Chisq Chi Df Pr(>Chisq)
## mouse_m7 12 -2146.6 -2089.9 1085.3 -2170.6
## mouse_m6 15 -2221.1 -2150.2 1125.5 -2251.1 80.51      3 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# residuals versus fitted plot
plot(mouse_m6)

```



```
# qqplot  
qqPlot(resid(mouse_m6))
```



```
## 290 530
## 287 512
# fixed effects test
Anova(mouse_m6, test='F')
## Analysis of Deviance Table (Type II Wald F tests with Kenward-Roger df)
##
## Response: rmr
##           F Df  Df.res    Pr(>F)
## bm       18.9720  1  31.873 0.0001286 ***
## temp31   77.6848  1  17.495 7.562e-08 ***
## wheel   118.9145  1 102.668 < 2.2e-16 ***
## bm:temp31  2.6467  1  51.324 0.1098908
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



- 
4. ♦ Take a subset of the mouse data where temperature (`temp`) is 31. Run a mixed-effects model of resting metabolic rate with only `sex` (Male or Female) as predictor. Choose appropriate random effects. Does `rmr` vary with sex of the mouse?

```
mouse31 <- subset(mouse, temp==31)

m31_1 <- lmer(rmr ~ sex + (1|run/id), data=mouse31)
Anova(m31_1)

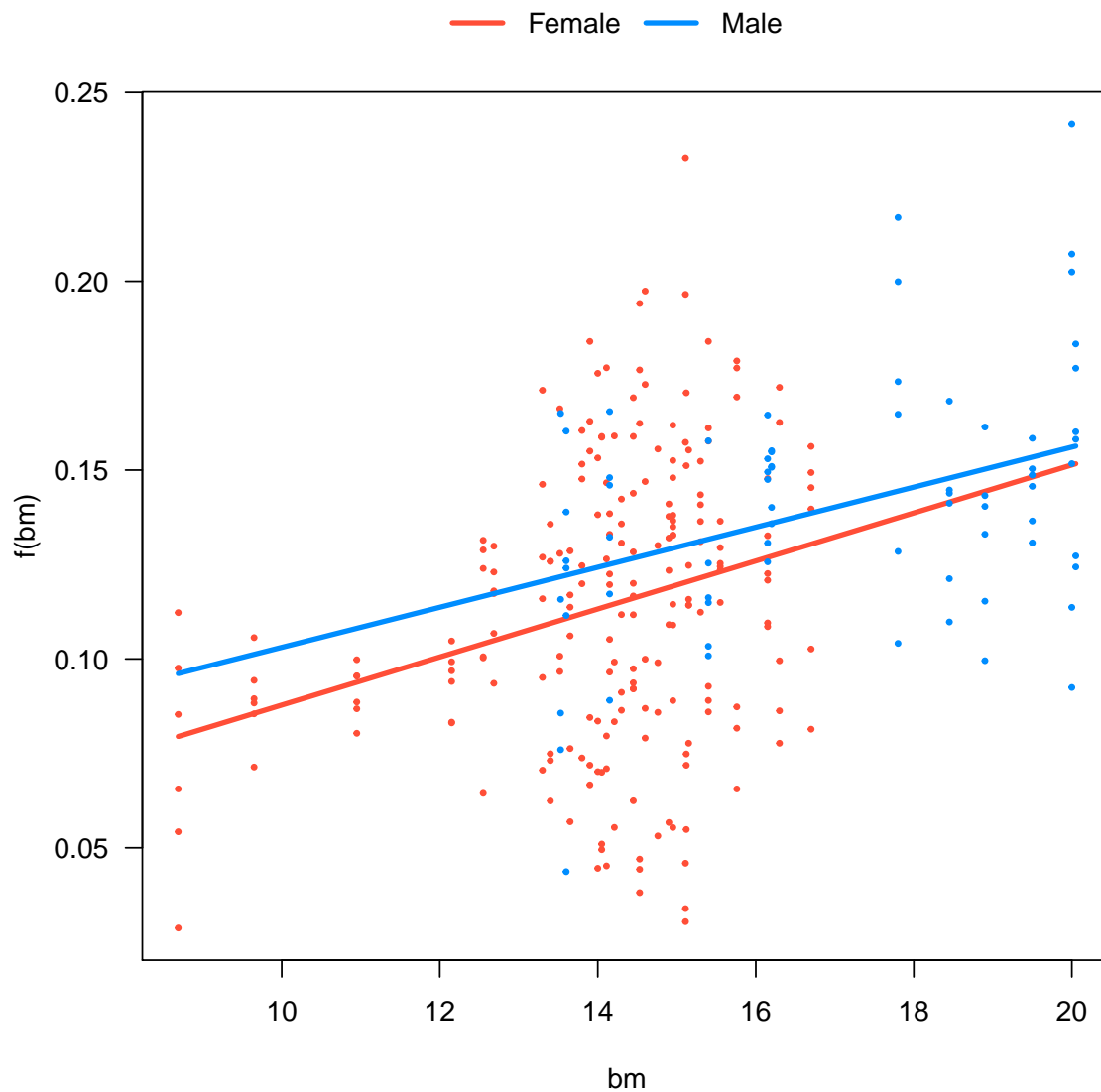
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: rmr
##      Chisq Df Pr(>Chisq)
## sex  9.7379  1  0.001805 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5. ♦ Now add body mass (`bm`) to the model with `sex`, as a main effect and interaction. Is `sex` still significant? Use `visreg` to understand how body mass, `sex`, and resting metabolic rate are related. If it helps, also test whether body mass varies with `sex`.

```
m31_2 <- lmer(rmr ~ bm*sex + (1|run/id), data=mouse31)
Anova(m31_2)

## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: rmr
##      Chisq Df Pr(>Chisq)
## bm    10.4487  1  0.001227 **
## sex     0.9337  1  0.333891
## bm:sex  0.0791  1  0.778544
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

visreg(m31_2, "bm", by="sex", overlay=TRUE)
```



### 1.10.3 Litter decomposition data

1. ■ Generate some diagnostic plots for the `litter_m1` model described in Section ??.

```
# Read data
litter <- read.csv("masslost.csv")

# Make sure the intended random effects (plot and block) are factors
litter$plot <- as.factor(litter$plot)
litter$block <- as.factor(litter$block)

# Represent date as number of days since the start of the experiment
library(lubridate)

##
```

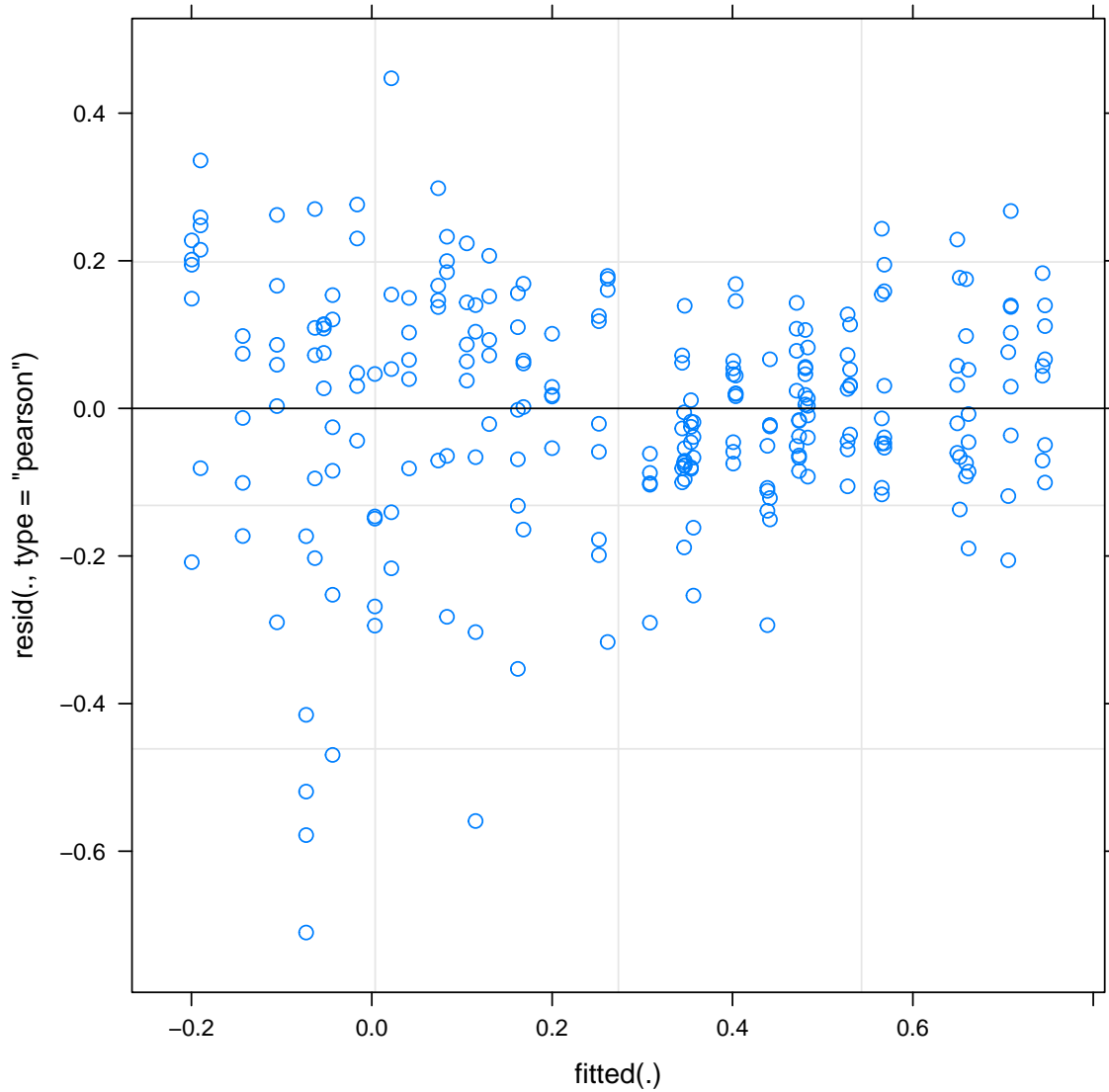
```
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':
##
##   date

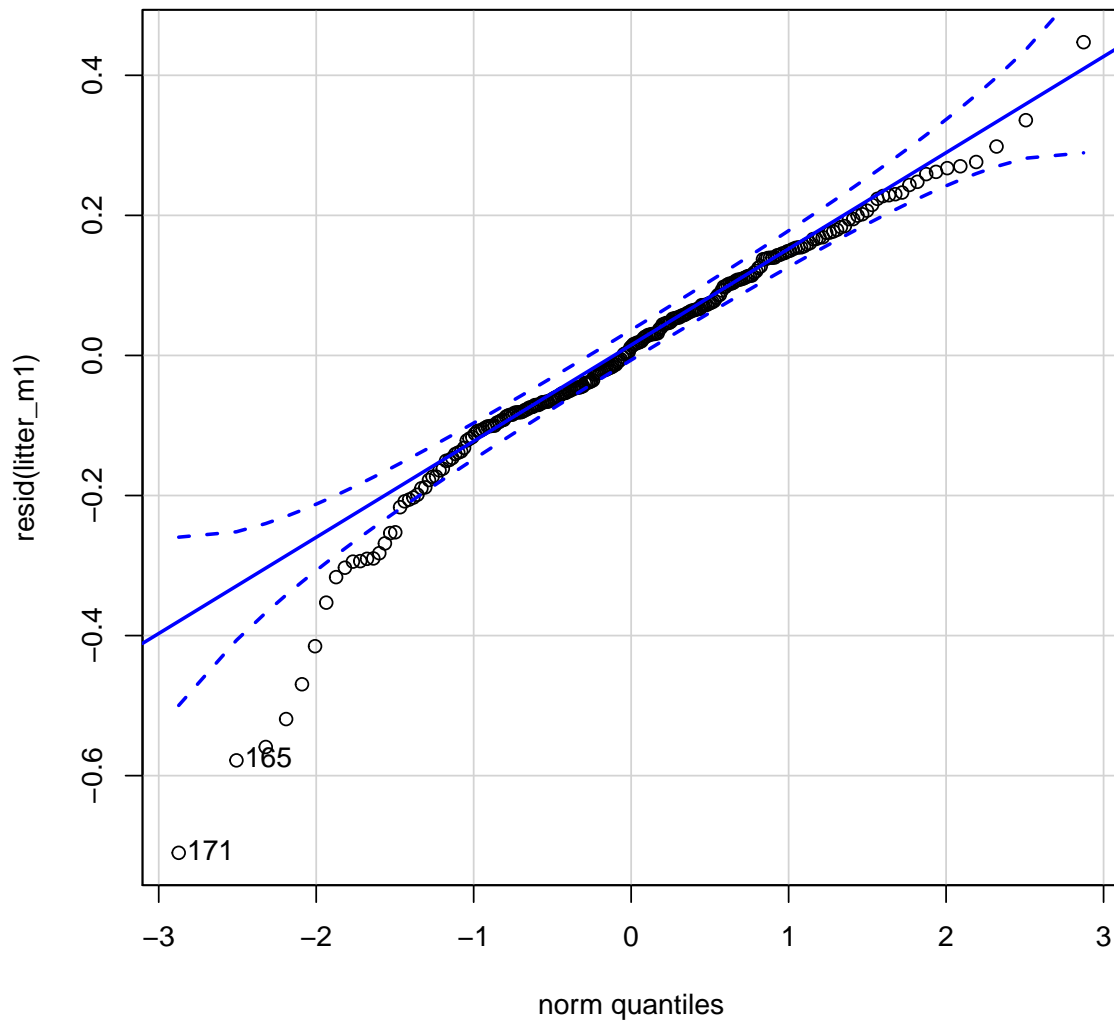
litter$date <- as.Date(mdy(litter$date))
litter$date2 <- litter$date - as.Date("2006-05-23")

# fit model
litter_m1 <- lmer(masslost ~ date2 + herbicide * profile + (1|block/plot),
  data = litter)

# residuals versus fitted plot
plot(litter_m1)
```



```
# qqplot
qqPlot(resid(litter_m1))
```



```
## [1] 171 165
```

2. ♦ The litter data contain a factor (variety) describing whether the litter is derived from a genetically modified (gm) or conventional (nongm) soy variety. Plot the data to observe the effect of variety. Use lmer to test the effect of variety, in addition to the other significant variables, on litter decomposition.

```
# Read data and get summary
litter <- read.csv('masslost.csv')

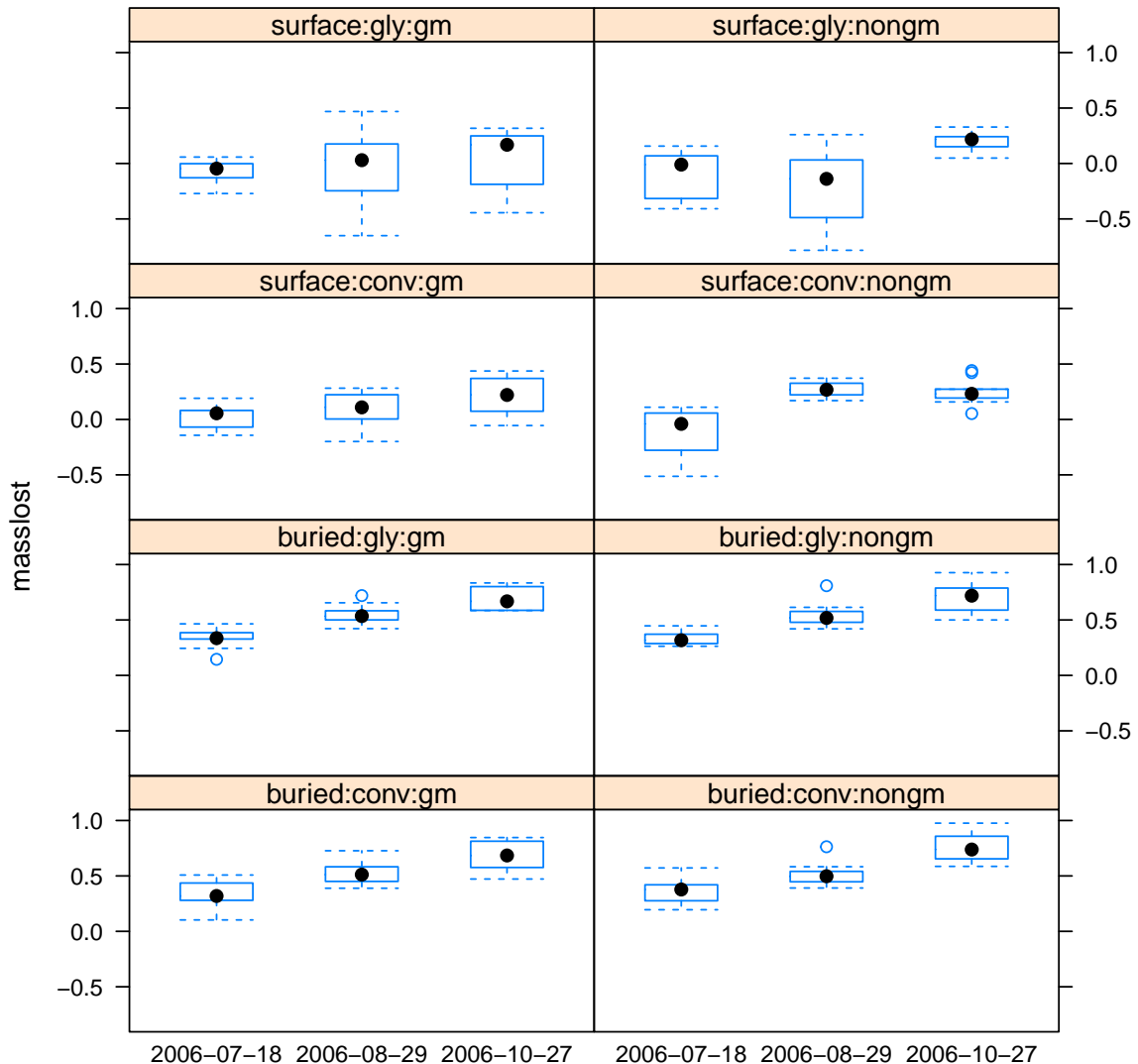
# Change random effects (plot and block) to factors
litter$plot <- as.factor(litter$plot)
litter$block <- as.factor(litter$block)
```

```

# Represent date as number of days since the start of the experiment
library(lubridate)
litter$date <- as.Date(mdy(litter$date))
litter$date2 <- litter$date - as.Date('2006-05-23')

# look for treatment effects
library(lattice)
bwplot(masslost ~ factor(date) | profile:herbicide:variety, data=litter, layout=c(2,4))

```



```

# there does not look to be much effect of 'variety' (look across two
# columns - 'layout' argument controls this)

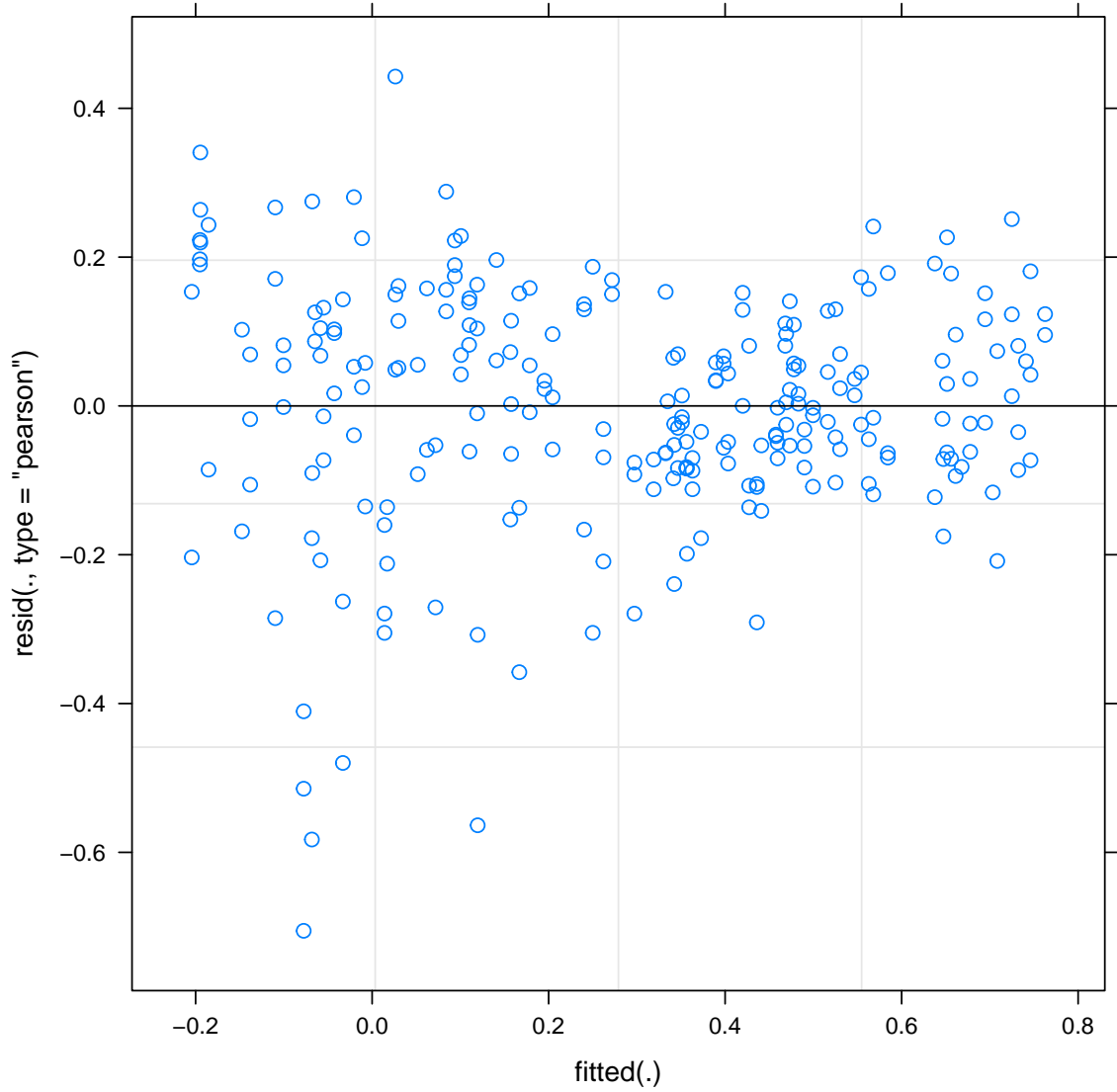
library(lme4)
library(car)
m1 <- lmer(masslost ~ date2 + herbicide * profile * variety + (1|block/plot), data = litter)

```

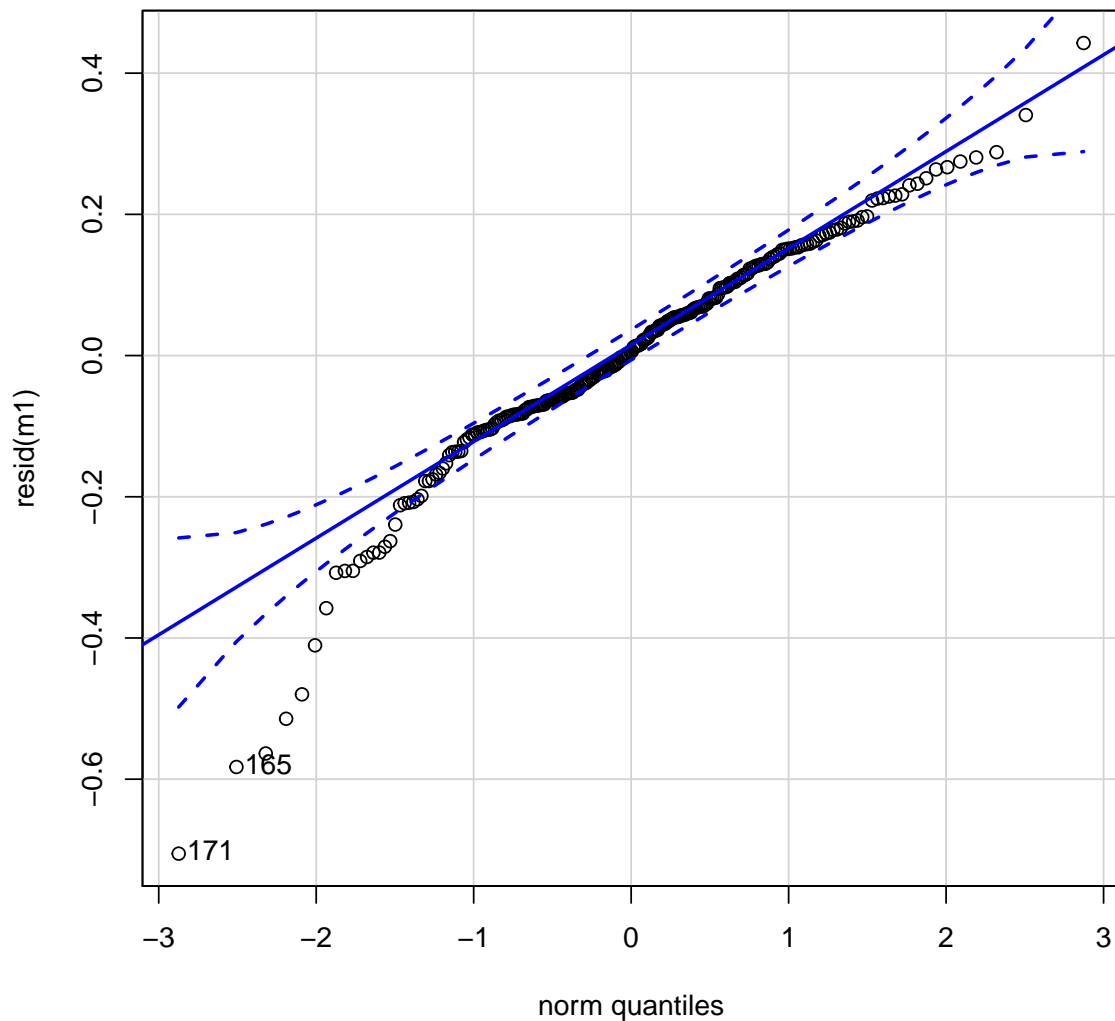
---

```
Anova(m1, test='F')
## Analysis of Deviance Table (Type II Wald F tests with Kenward-Roger df)
##
## Response: masslost
##
##           F Df  Df.res    Pr(>F)
## date2      140.0675  1 231.436 < 2.2e-16 ***
## herbicide   12.7977  1   3.005 0.0372580 *
## profile    519.2389  1 231.498 < 2.2e-16 ***
## variety     0.3392  1 231.424 0.5608858
## herbicide:profile  11.7097  1 231.277 0.0007351 ***
## herbicide:variety  0.4590  1 231.820 0.4987471
## profile:variety    0.0758  1 231.795 0.7832914
## herbicide:profile:variety 0.0051  1 231.304 0.9431549
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# residuals versus fitted plot
plot(m1)
```



```
# qqplot  
qqPlot(resid(m1))
```



```
## [1] 171 165
```

### 1.10.4 Repeated measures in leaf photosynthesis

For this exercise you will use the EucFACE leaf gas exchange dataset. First walk through Section ?? before starting this exercise.

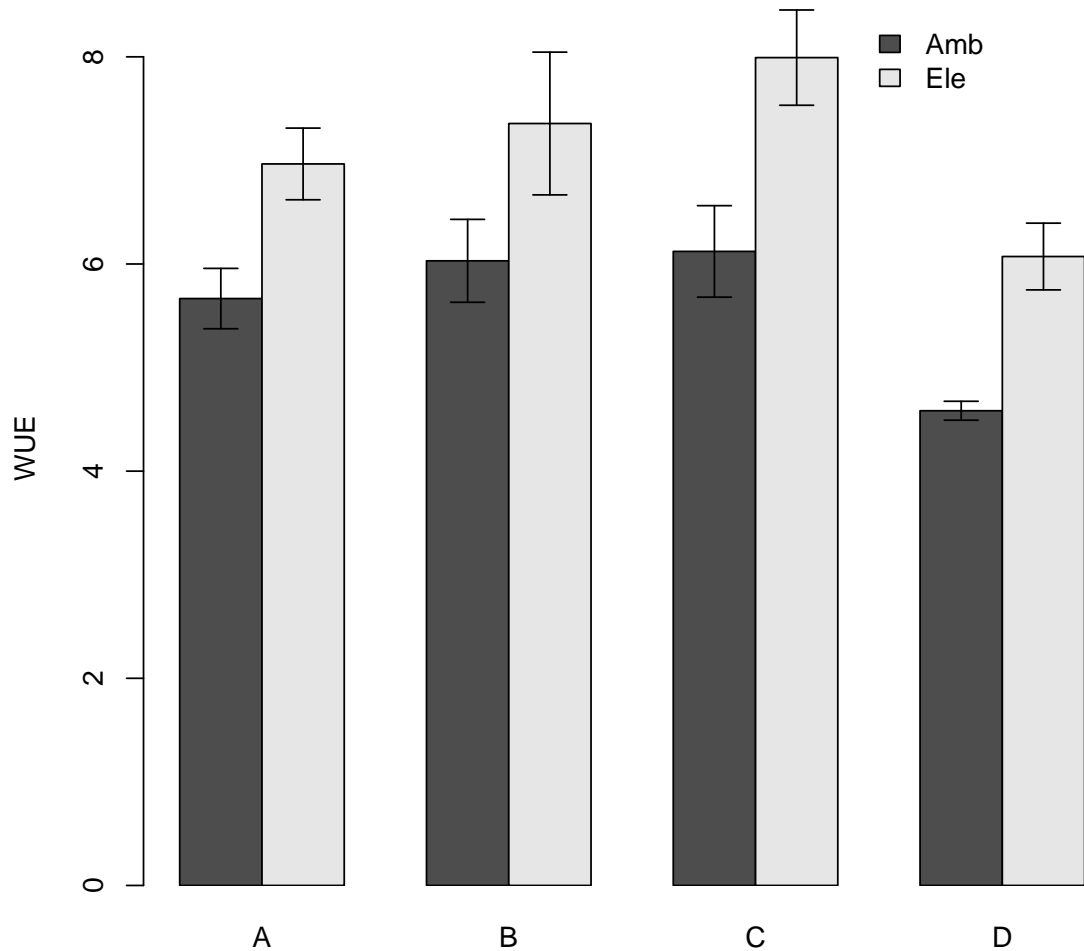
1. Read the data, and add a new variable called WUE (water-use efficiency), calculated as  $Photo$  divided by  $Trmmol$ . This new variable represents the amount of  $CO_2$  taken up per unit water lost.

```
eucgas <- read.csv("eucface_gasexchange.csv")
eucgas$WUE <- with(eucgas, Photo / Trmmol)
```

2. ■ Make a barplot of WUE by Date and  $CO_2$  treatment, as shown in Section ??.



```
library(sciplot)
bargraph.CI(Date, WUE, CO2, data=eucgas, legend=TRUE, ylab="WUE")
```



- Fit a linear mixed-effects model with `Tree` within `Ring` as the random effects structure, and `Date` and `CO2` as fixed effect. Fit a second model with additionally the interaction between `Date` and `CO2`; is the interaction significant (use both `anova` on the two models, and `Anova` on the model with the interaction).

```
wue_m0 <- lmer(WUE ~ Date + CO2 + (1|Ring/Tree), data=eucgas)
wue_m1 <- lmer(WUE ~ Date * CO2 + (1|Ring/Tree), data=eucgas)

# Likelihood ratio test
anova(wue_m0, wue_m1)

## Data: eucgas
## Models:
## wue_m0: WUE ~ Date + CO2 + (1 | Ring/Tree)
```

```
## wue_m1: WUE ~ Date * CO2 + (1 | Ring/Tree)
##           Df      AIC      BIC logLik deviance Chisq Chi Df Pr(>Chisq)
## wue_m0   8 291.00 310.44 -137.50  275.00
## wue_m1  11 296.25 322.99 -137.12  274.25 0.7458      3    0.8624

# Chi-square test
Anova(wue_m1)

## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: WUE
##           Chisq Df Pr(>Chisq)
## Date       23.9240 3 2.591e-05 ***
## CO2        17.6809 1 2.612e-05 ***
## Date:CO2   0.6917 3    0.8752
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4. ♦ Inspect the standard deviation of the random effects with `VarCorr`. Notice that the standard deviation for the term `Tree:Ring` (trees within rings) is very small compared to the `Ring` random effect. A naive user might remove this random effect since it explains little variation. Why is that a bad idea?

```
VarCorr(wue_m0)

## Groups      Name          Std.Dev.
## Tree:Ring (Intercept) 0.00000
## Ring       (Intercept) 0.28862
## Residual                    1.25981

# The trees are the experimental units. If we remove it from the random effects structure, we assume
# all measurements within rings are independent, and are thus committing pseudoreplication.
```

5. ▲ Look at the summary statement of the model for WUE that does not contain an interaction. Notice that Date D is significantly lower than the intercept (Date A). Refit the model without an intercept, and use `glht` to test whether Date D was lower than all of Date A, B and C. *Hint: this follows closely the example shown at the end of Section ?? (p. ??).*

```
wue_m0_re <- lmer(WUE ~ Date + CO2 -1 + (1|Ring/Tree), data=eucgas)

library(multcomp)

## Loading required package: mvtnorm
## Loading required package: survival
## Loading required package: TH.data
## Loading required package: MASS
##
## Attaching package: 'TH.data'
## The following object is masked from 'package:MASS':
##
##   geyser
##
## Attaching package: 'multcomp'
## The following object is masked _by_ '.GlobalEnv':
```

```
##
## litter
summary(glht(wue_m0_re, linfct=c("DateD - DateA = 0",
                                "DateD - DateB = 0",
                                "DateD - DateC = 0"
                                )))
```

6. ♦ A colleague tells you that water-use efficiency may be lower on Date D because air temperature was higher, resulting in a lower air humidity. This is indicated by the variable `VpdL` in the dataset, which represents the 'vapour pressure deficit' - higher values indicate drier air. Refit the model for WUE with the `CO2*Date` interaction, and now add `VpdL` as a fixed effect to the model (but no interactions between `VpdL` and `Date` or `CO2`). Is VPD significant (use multiple methods to test this)?

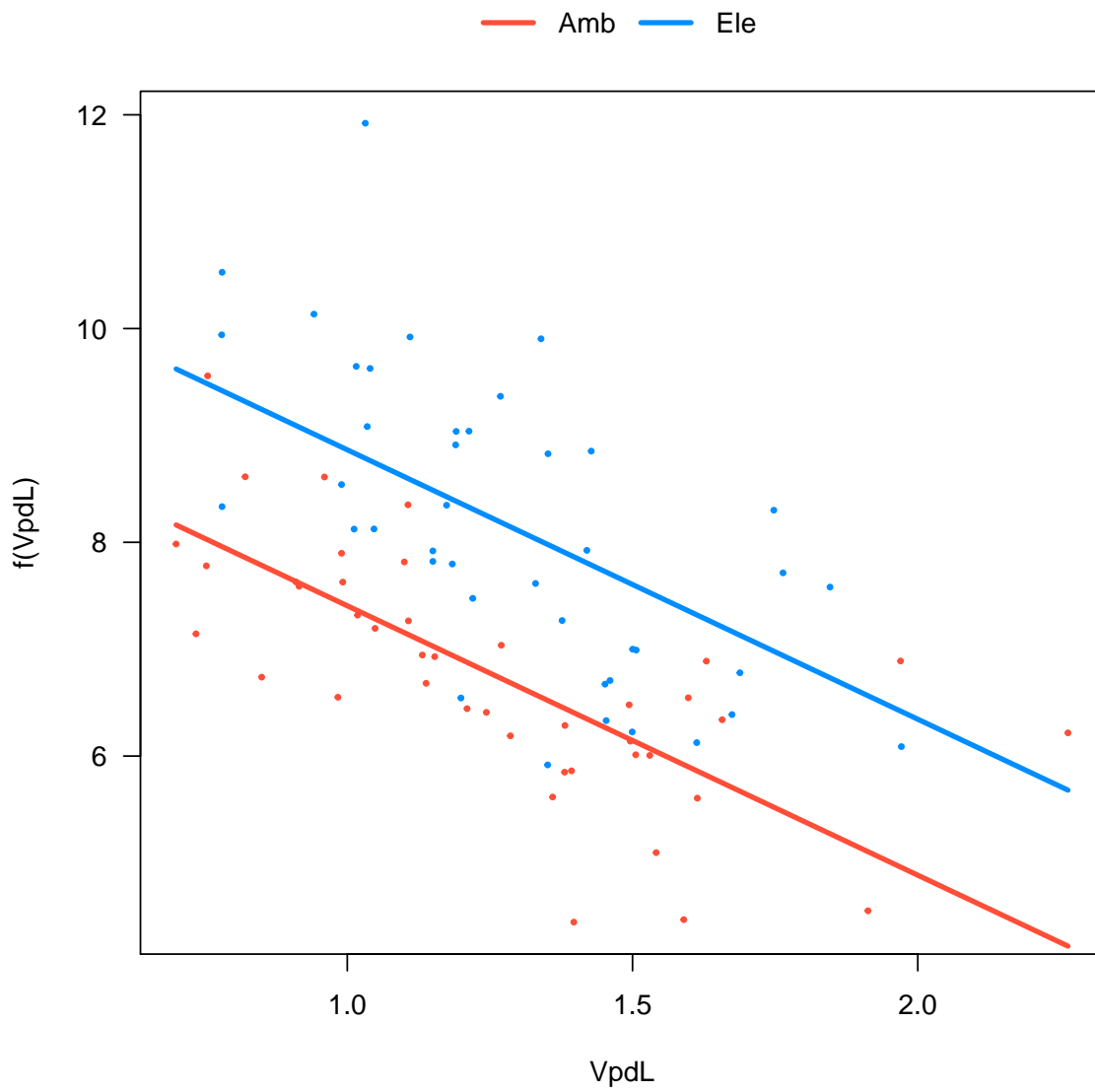
```
wue_m2 <- lmer(WUE ~ VpdL + Date * CO2 + (1|Ring/Tree), data=eucgas)

Anova(wue_m2)

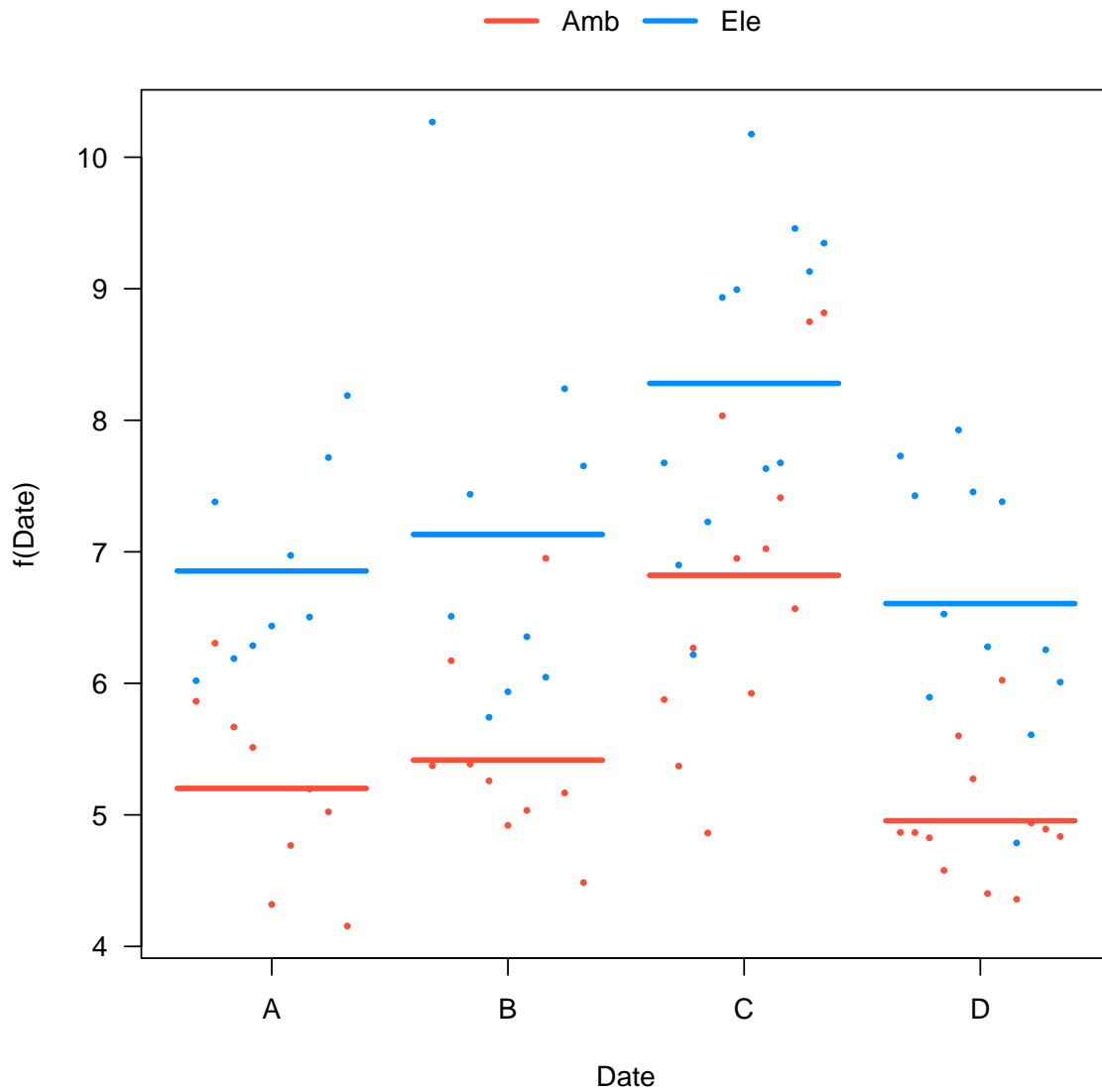
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: WUE
##           Chisq Df Pr(>Chisq)
## VpdL       28.8962  1 7.636e-08 ***
## Date       36.7936  3 5.088e-08 ***
## CO2        16.0599  1 6.137e-05 ***
## Date:CO2    0.1697  3  0.9823
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

7. ▲ Use `visreg` to visualize a) the `VpdL` effect (make a `visreg` plot with VPD as the predictor, showing predictions for both `CO2` treatment), b) The `Date` effect. For this latter plot, use `Date` as the `xvar` argument to `visreg`, and `CO2` as the `by` argument (with `overlay=TRUE`). For this last plot, `visreg` will show the `Date` effect *at a median value of VpdL*, that is, the effect of `VpdL` has been removed first. Is `Date D` still different from the other `Dates`? Can you explain this in terms of the `VpdL` covariate?

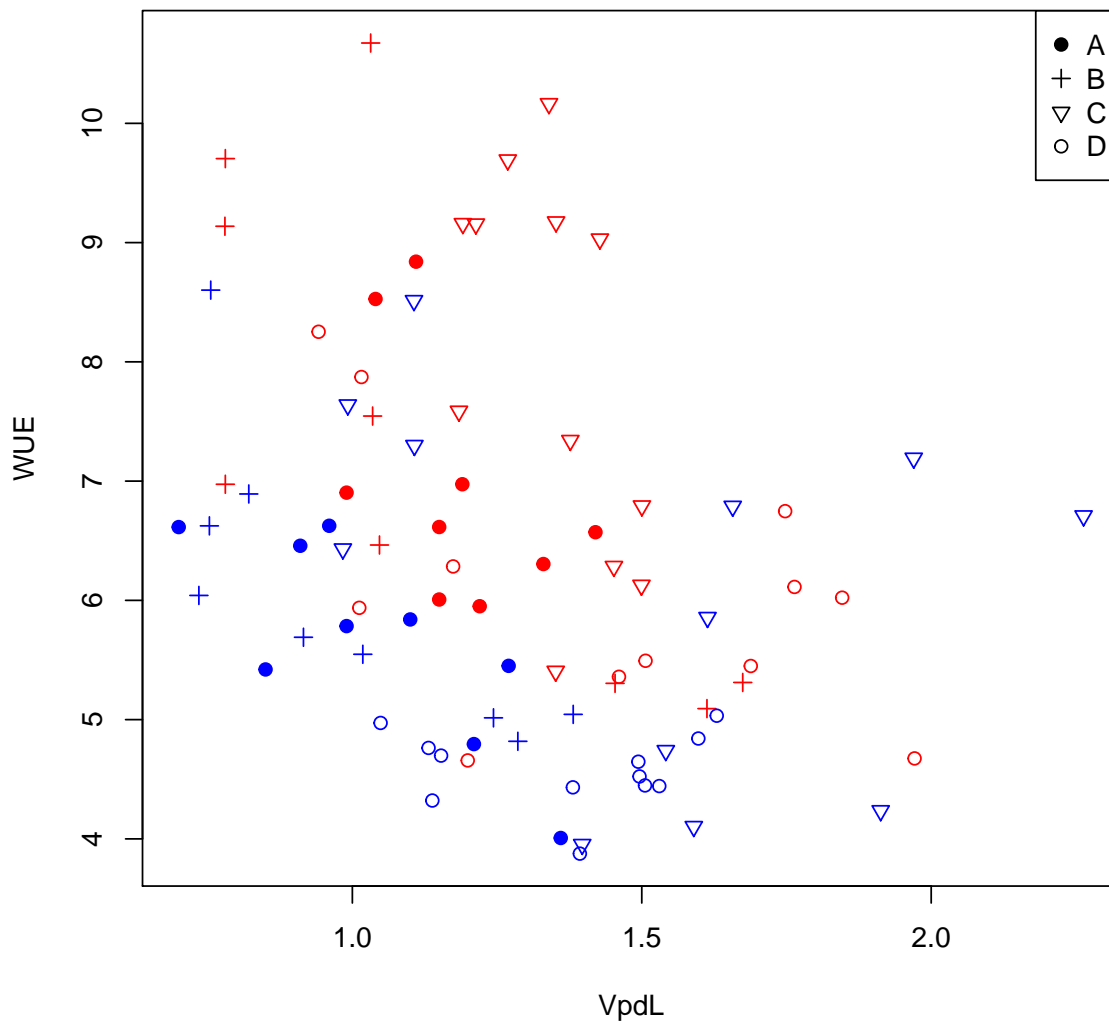
```
# VPD effect - visreg will take the first Date for the predictions
visreg(wue_m2, "VpdL", by="CO2", overlay=TRUE)
```



```
# The Date effect. visreg will show the predictions at a median value of VPD.  
visreg(wue_m2, "Date", by="CO2", overlay=TRUE)
```



```
# The Dates are now very similar, and Date D is no longer lower. This means that VPD explained the
# WUE on Date D, because in the above model its effect has been accounted for.
# We can visually check this in the figure below. Note that Date D is lower, but also has higher VP
# all data for a CO2 treatment are more or less on one line.
with(eucgas, plot(VpdL, WUE, pch=c(19,3,6,1)[Date], col=c("blue","red")[CO2]))
legend("topright", levels(eucgas$Date), pch=c(19,3,6,1))
```



### 1.10.5 EucFACE ground cover data

The file `eucfaceGC.csv` contains estimates of plant and litter cover within the rings of the EucFACE experiment, evaluating forest ecosystem responses to elevated CO<sub>2</sub>, on two dates; the data description can be found in Section ?? (p. ??).

1. ♦ Convert the variables indicating the nested sampling design to factors, then use `glmer` in `lme4` to test for an interaction between `Trt` and `Date` on `Litter` cover. Litter cover represents counts (use the `poisson` family).

```
# read data
gc <- read.csv('eucfaceGC.csv')
str(gc)

## 'data.frame': 192 obs. of 8 variables:
## $ Date : Factor w/ 2 levels "11/06/13","3/09/13": 1 1 1 1 1 1 1 1 1 1 ...
```

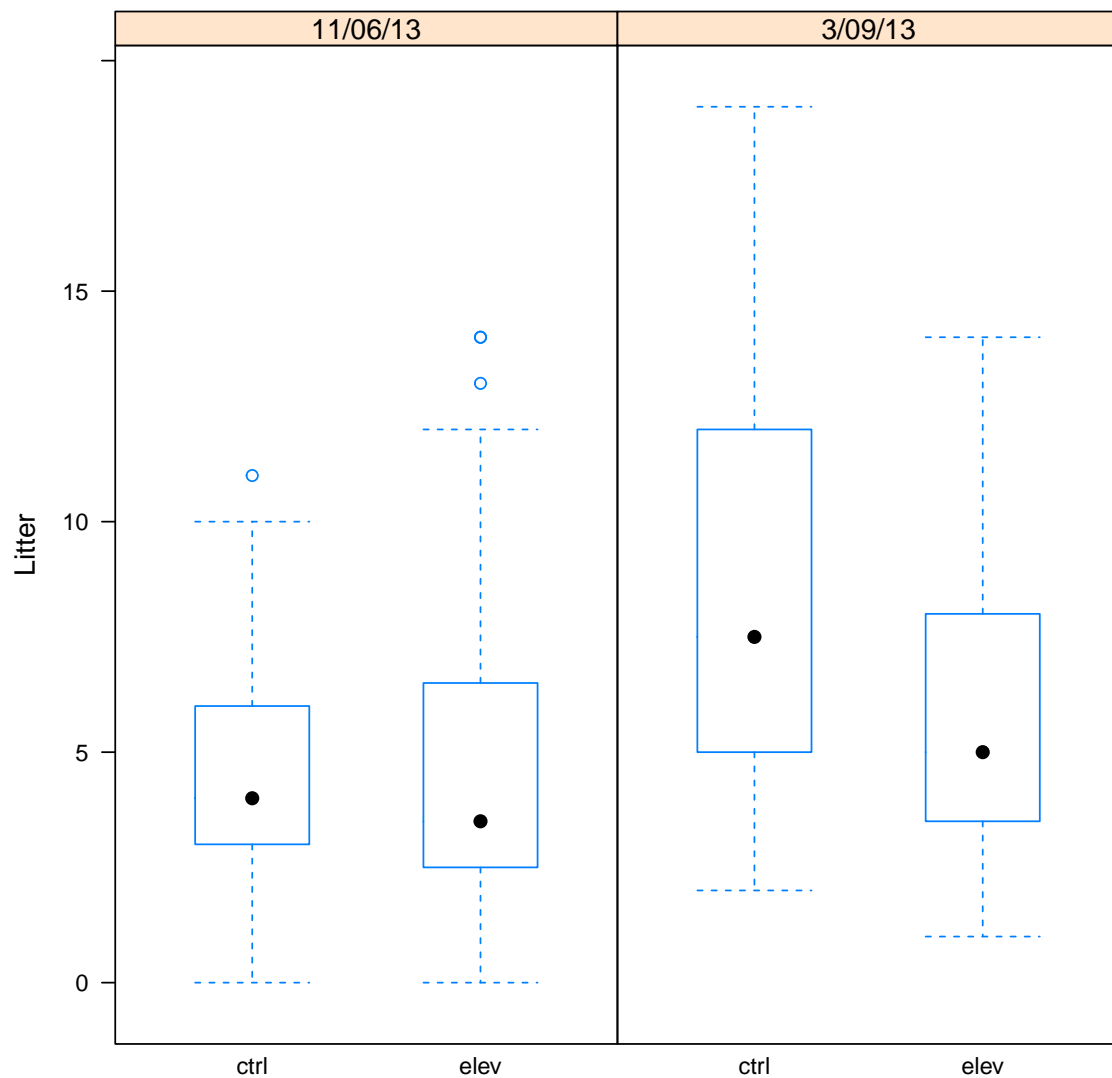
```

## $ Ring : int 1 1 1 1 1 1 1 1 1 1 ...
## $ Plot : int 1 1 1 1 2 2 2 2 3 3 ...
## $ Sub   : int 1 2 3 4 1 2 3 4 1 2 ...
## $ Forbes: int 0 2 0 0 2 5 4 0 1 3 ...
## $ Grass : int 16 14 16 16 15 16 16 16 16 13 ...
## $ Litter: int 5 6 2 2 3 3 4 3 0 4 ...
## $ Trt   : Factor w/ 2 levels "ctrl","elev": 2 2 2 2 2 2 2 2 2 2 ...

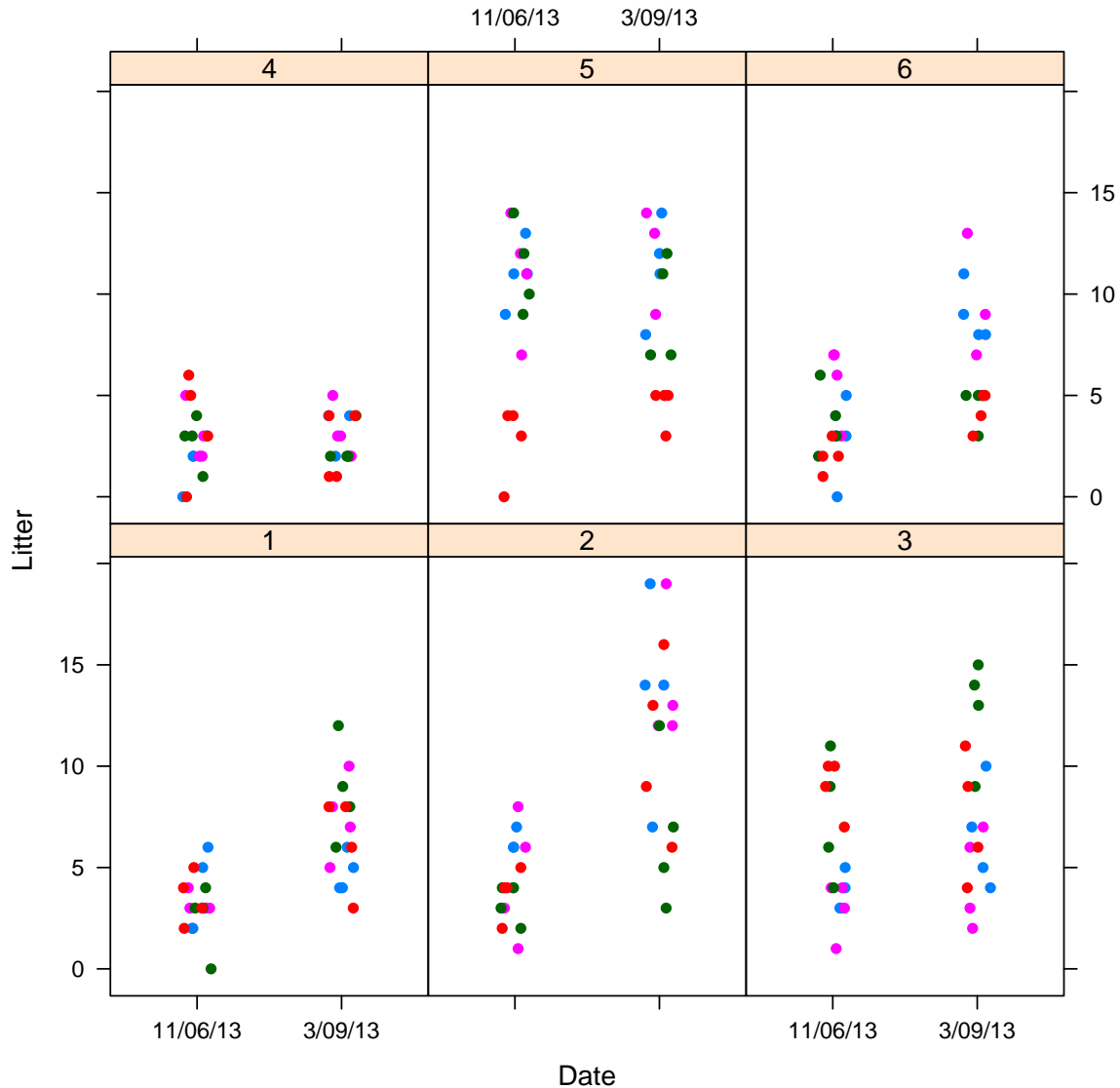
# convert random variables to factors
gc$Ring <- factor(gc$Ring)
gc$Plot <- factor(gc$Plot)
gc$Sub <- factor(gc$Sub)

# Plot variation associated with 'Trt' and date
bwplot(Litter ~ Trt | Date, data = gc)

```



```
# Plot variation associated with plots and subplots
xyplot(Litter ~ Date | Ring, groups = Plot,
       data = gc, pch = 16, jitter.x = TRUE)
```



```
# fit model and test for effects
m2 <- glmer(Litter ~ Trt * Date + (1|Ring/Plot), data=gc, family=poisson)
summary(m2)

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: poisson ( log )
## Formula: Litter ~ Trt * Date + (1 | Ring/Plot)
## Data: gc
##
##      AIC      BIC  logLik deviance df.resid
##  923.2   942.8  -455.6   911.2     186
##
```



```

## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.22194 -0.66709  0.02852  0.54770  2.40787
##
## Random effects:
##      Groups      Name              Variance Std.Dev.
## Plot:Ring (Intercept) 0.1023     0.3199
## Ring      (Intercept) 0.0867     0.2945
## Number of obs: 192, groups: Plot:Ring, 24; Ring, 6
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.45991   0.20529   7.111 1.15e-12 ***
## Trtelev           0.01786   0.29009   0.062 0.950914
## Date3/09/13      0.62952   0.08334   7.554 4.22e-14 ***
## Trtelev:Date3/09/13 -0.42318   0.12015  -3.522 0.000428 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) Trtelv D3/09/
## Trtelev      -0.707
## Date3/09/13 -0.265  0.187
## Tr:D3/09/13  0.184 -0.248 -0.694

library(car)
Anova(m2)

## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: Litter
##              Chisq Df Pr(>Chisq)
## Trt           0.7057  1  0.400889
## Date         50.3421  1  1.291e-12 ***
## Trt:Date     12.4057  1  0.000428 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

2. ♦ The interaction between `Trt` and `Date` is significant. Looking at our box and whisker plot from the previous exercise, there are two possible explanations for this significant interaction: (i) the effect of `Trt` differs for each level of `Date`, or (ii) the effect of `Date` differs for each level of `Trt`. Perform multiple comparisons using functions in the `emmeans` package to test these two hypotheses. *Hint: see Section ?? (p. ??) for guidance.*

```

# fit model
m2 <- glmer(Litter ~ Trt * Date + (1|Ring/Plot), data=gc, family=poisson)

# load emmeans library to perform multiple comparisons
library(emmeans)

##
## Attaching package: 'emmeans'

## The following object is masked from 'package:multcomp':
##
##      cld

```

```

# estimate marginal means and view 'Trt' contrasts within each level of 'Date'
m2.trt <- emmeans(m2, ~ Trt | Date)
pairs(m2.trt) # both contrasts are not significant, it's not this option

## Date = 11/06/13:
## contrast      estimate      SE  df z.ratio p.value
## ctrl - elev -0.01785786 0.2900915 Inf  -0.062  0.9509
##
## Date = 3/09/13:
## contrast      estimate      SE  df z.ratio p.value
## ctrl - elev  0.40532419 0.2850728 Inf   1.422  0.1551
##
## Results are given on the log (not the response) scale.

# estimate marginal means and view 'Date' contrasts within each level of 'Trt'
m2.date <- emmeans(m2, ~ Date | Trt)
pairs(m2.date) # both contrasts are significant, but the effect size is larger for 'Trt = ctrl'.

## Trt = ctrl:
## contrast      estimate      SE  df z.ratio p.value
## 11/06/13 - 3/09/13 -0.6295201 0.08333513 Inf  -7.554  <.0001
##
## Trt = elev:
## contrast      estimate      SE  df z.ratio p.value
## 11/06/13 - 3/09/13 -0.2063381 0.08654955 Inf  -2.384  0.0171
##
## Results are given on the log (not the response) scale.

```

3. ▲ Make model diagnostics plot for the model of Grass with Trt and Date as fixed effects and Ring and Plot as random effects. Use the DHARMA package as shown in Fig. ???. Now refit this model with the poisson family - which is a wrong model specification because Grass count is clearly bounded by a maximum (16 quadrants in each plot). Make the model diagnostic plot again, and note how it has changed.

```

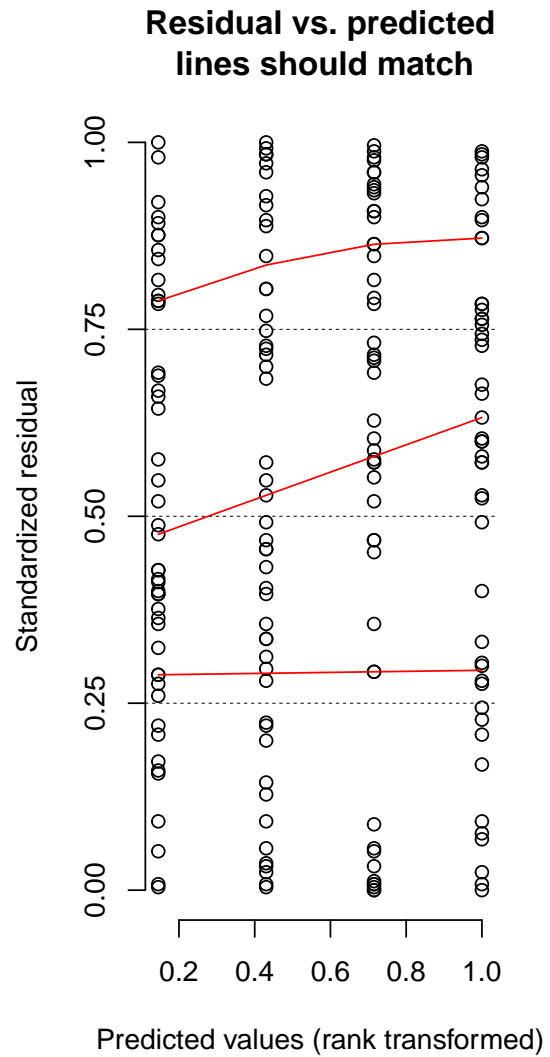
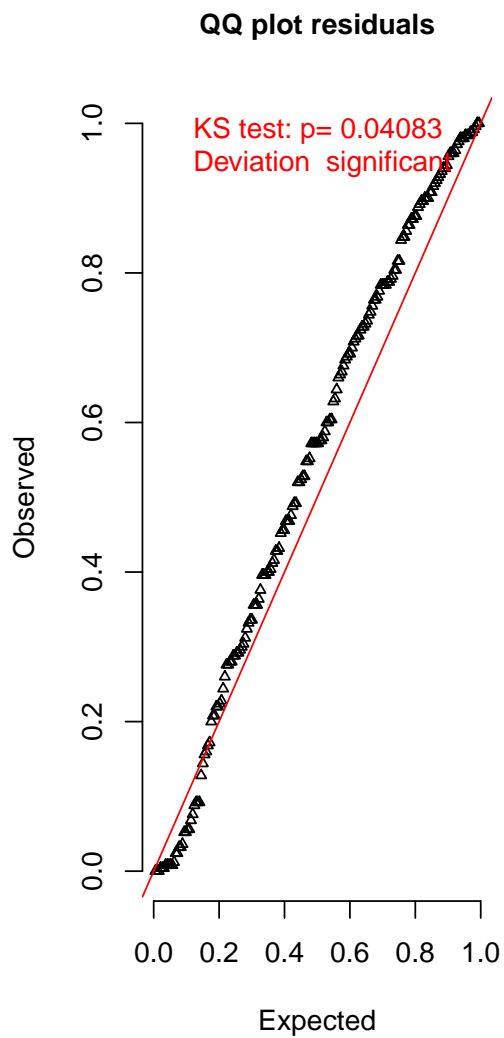
library(DHARMA)

# 'correct' model
m1 <- glmer(cbind(Grass, 16-Grass) ~ Trt * Date + (1|Ring/Plot),
            data=gc, family=binomial)

# Does not look great, though.
plot(simulateResiduals(m1))

```

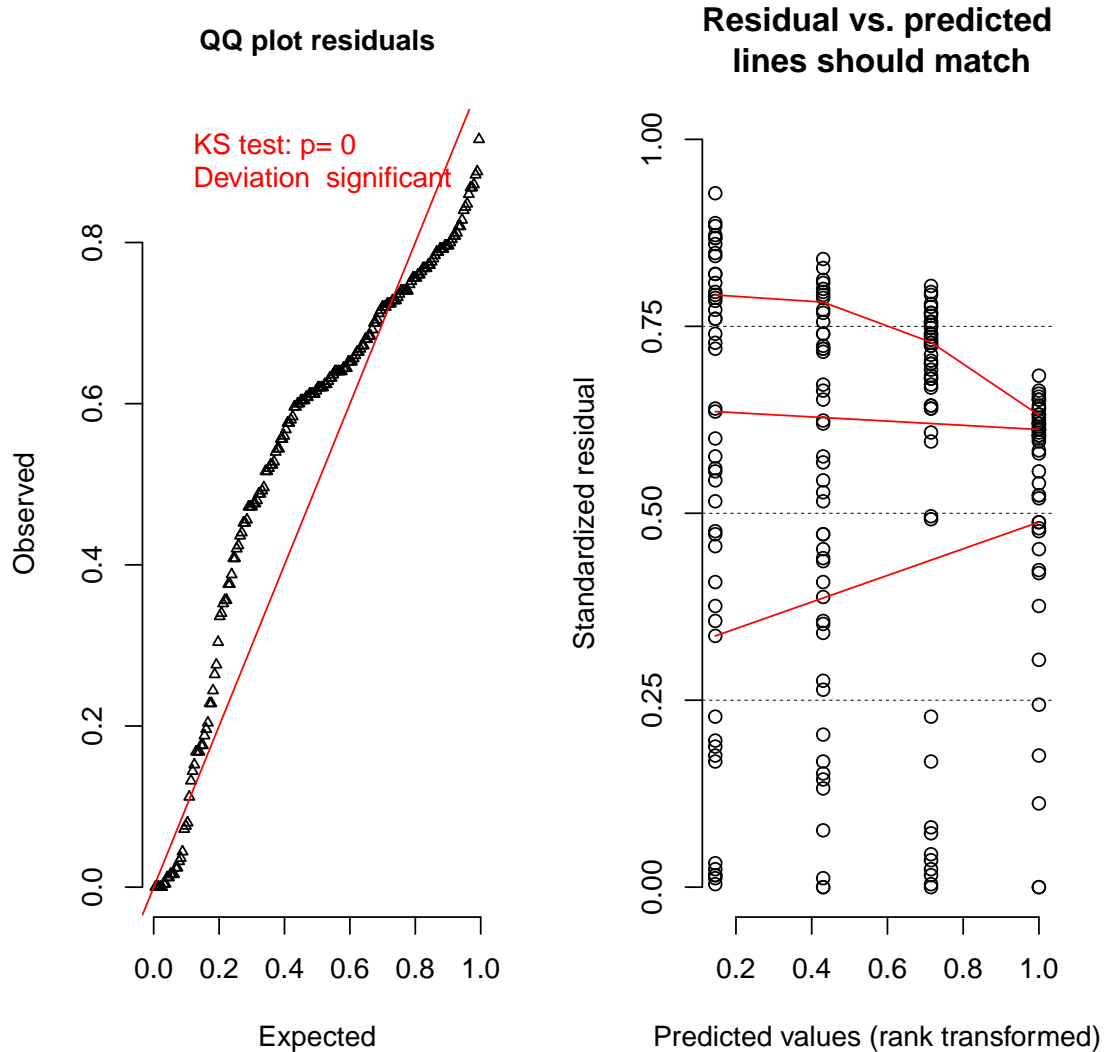
DHARMA scaled residual plots



```
# incorrect model
m1_bad <- glmer(Grass ~ Trt * Date + (1|Ring/Plot),
               data=gc, family=poisson)

# Very bad!
plot(simulateResiduals(m1_bad))
```

## DHARMA scaled residual plots



### 1.10.6 Logistic regression

For this exercise, we will use the seed germination data (fire experiment). First walk through Section ??.

1. ♦ For the seedfire dataset, the examples in the text use `temp` as a factor variable. Attempt to instead use it as an (untransformed) continuous variable in the model. Compare the model with `temp` as a factor and `temp` as a numeric variable, in terms of AIC. Now plot diagnostic plots for both models (see Section ??). Ignore the warning.

```
seedfire <- read.csv("germination_fire.csv")  
  
# Make sure temperature treatment is a factor  
seedfire$temp <- as.factor(seedfire$temp)  
  
# temp as factor
```

---

```
firefit1 <- glmer(cbind(germ, n-germ) ~ species + temp +
                 (1|site), data=seedfire, family=binomial)

# temp as numeric predictor
seedfire$temp_c <- as.numeric(as.character(seedfire$temp))

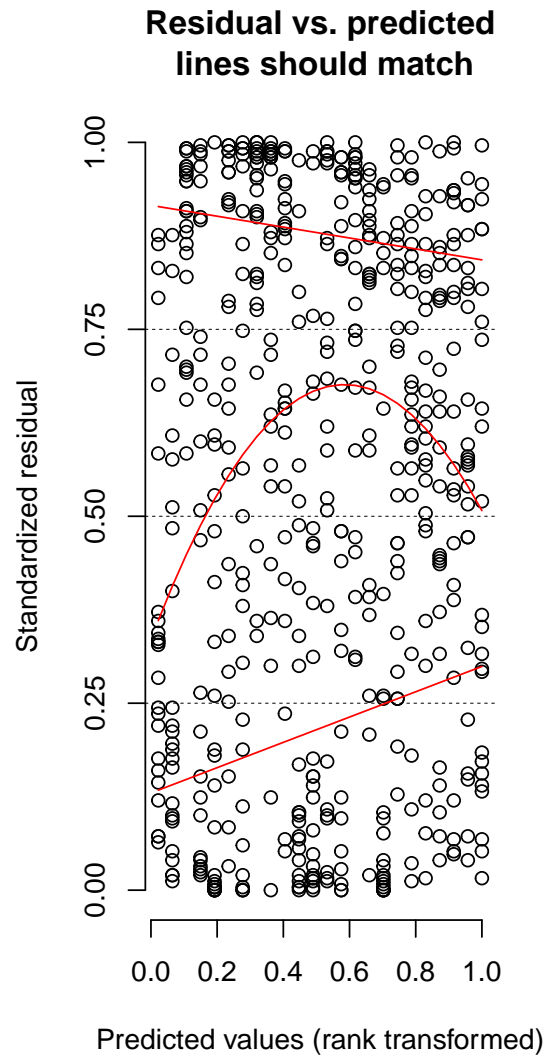
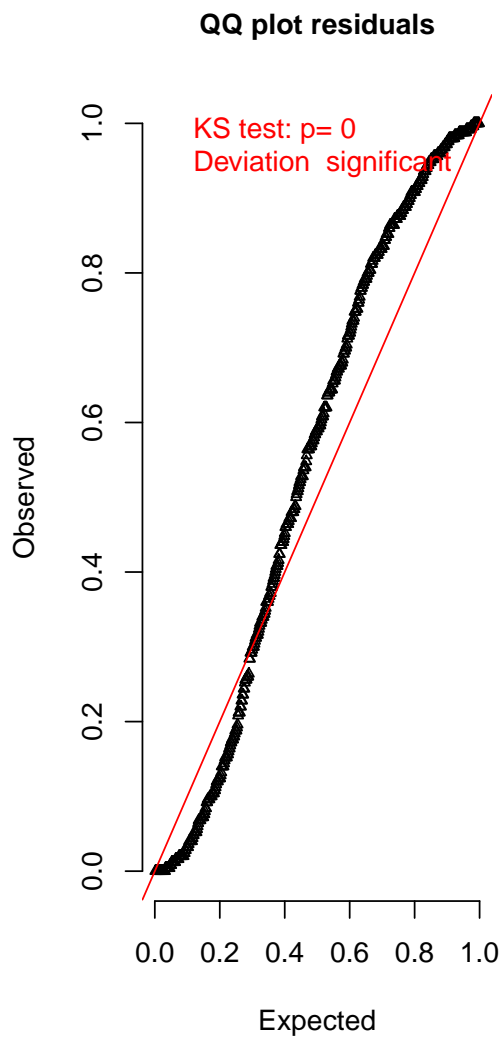
firefitc1 <- glmer(cbind(germ, n-germ) ~ species + temp_c +
                 (1|site), data=seedfire, family=binomial)

# compare model fits, the model with temp as factor is better
AIC(firefit1, firefitc1)

##           df      AIC
## firefit1  10 2712.551
## firefitc1   6 4606.511

# view diagnostic plots
library(DHARMA)
plot(simulateResiduals(firefit1))
```

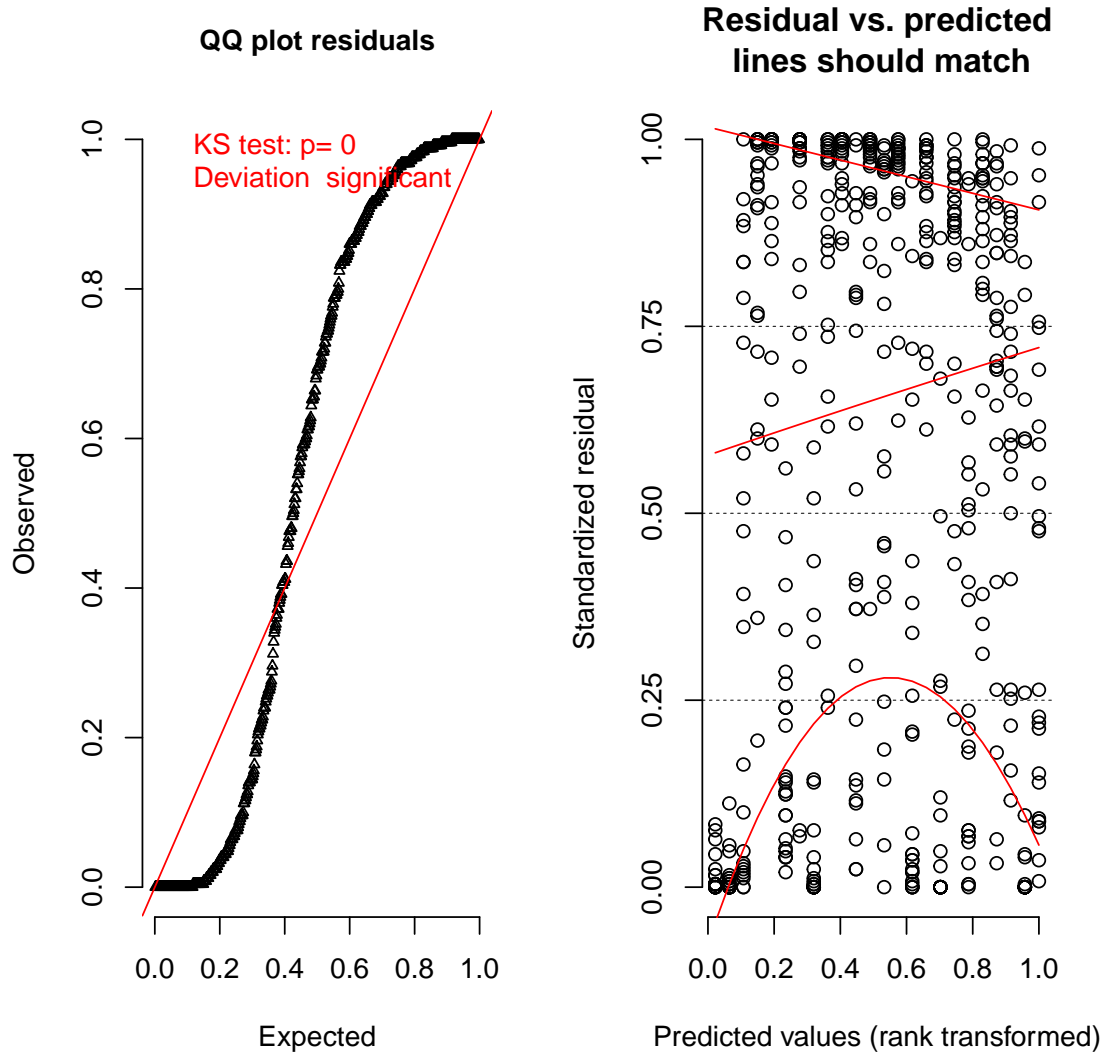
DHARMA scaled residual plots



```
plot(simulateResiduals(firefitc1))
```

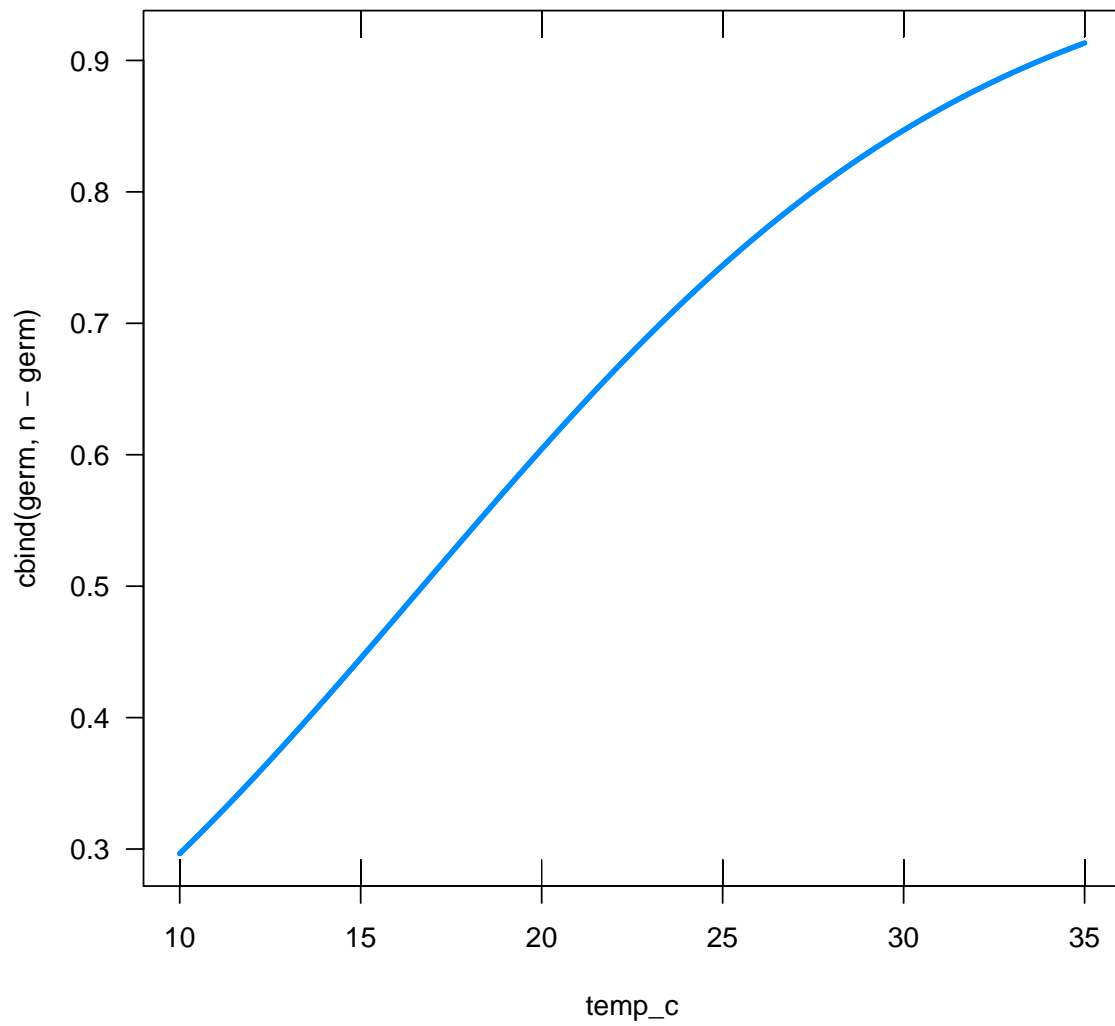
---

DHARMA scaled residual plots



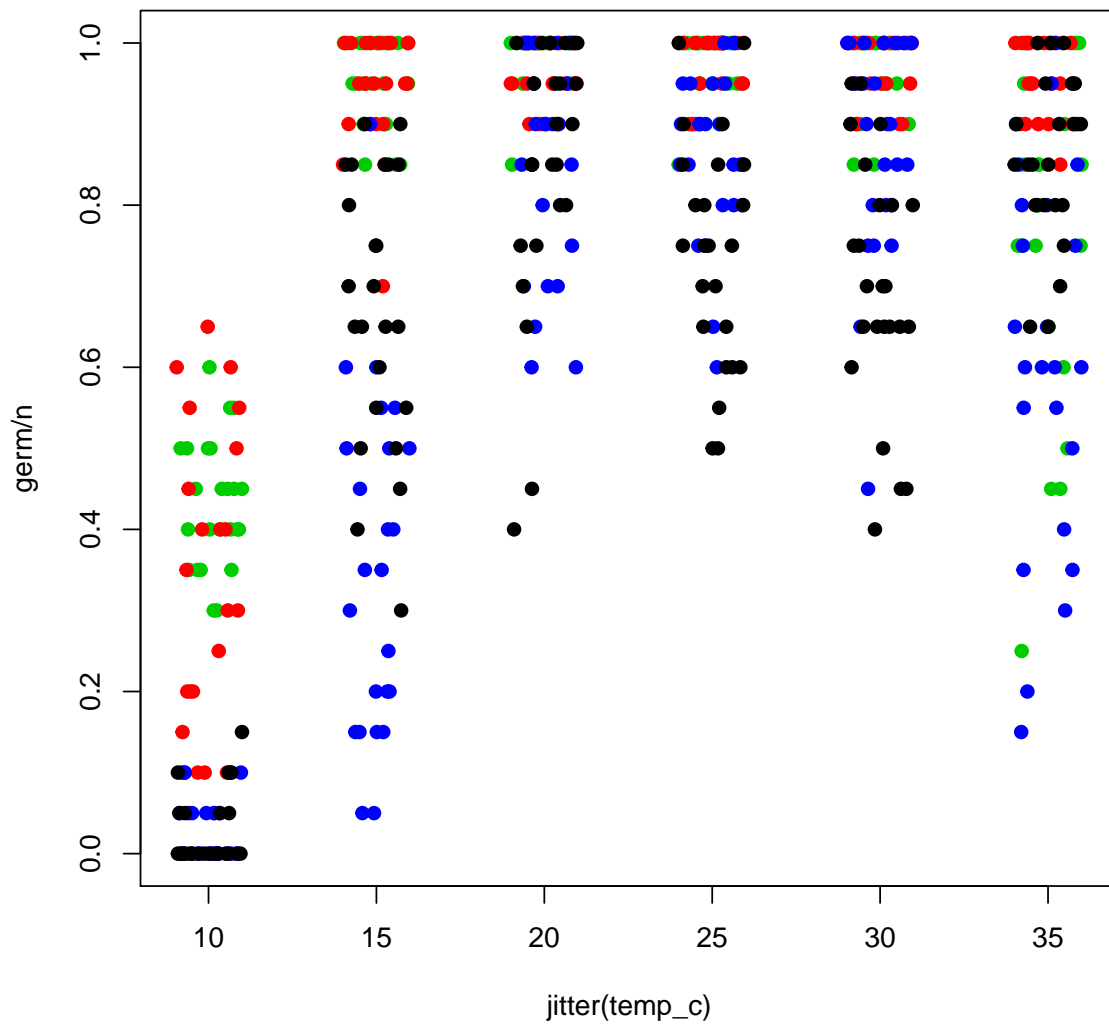
2. ▲ Try to understand why `temp` as a continuous variable does so badly, using `visreg` (set `temp` as the `xvar`, and also use `scale="response"`). Compare the fitted response to the data (i.e., also make a plot of germination success versus temperature, and compare to the `visreg` plot).

```
visreg(firefitc1, "temp_c", scale="response")
```



```
# Compare to:  
palette("default")  
with(seedfire, plot(jitter(temp_c), germ/n, col=species, pch=19))
```





```
# The data show an abrupt increase in germination success with temperature, but the predictor  
# was added as a linear term, which is much too gradual.
```