R Users Group
Linear Mixed - Effect Models

Content

Brief Introduction
M.Sc. Data Case Study
Additional Resources

Jordan Graham – April 6th , 2018

What Are Mixed-Models?

Lecture 10: Linear Mixed Models
(Linear Models with Random Effects)

Claudia Czado

TU München

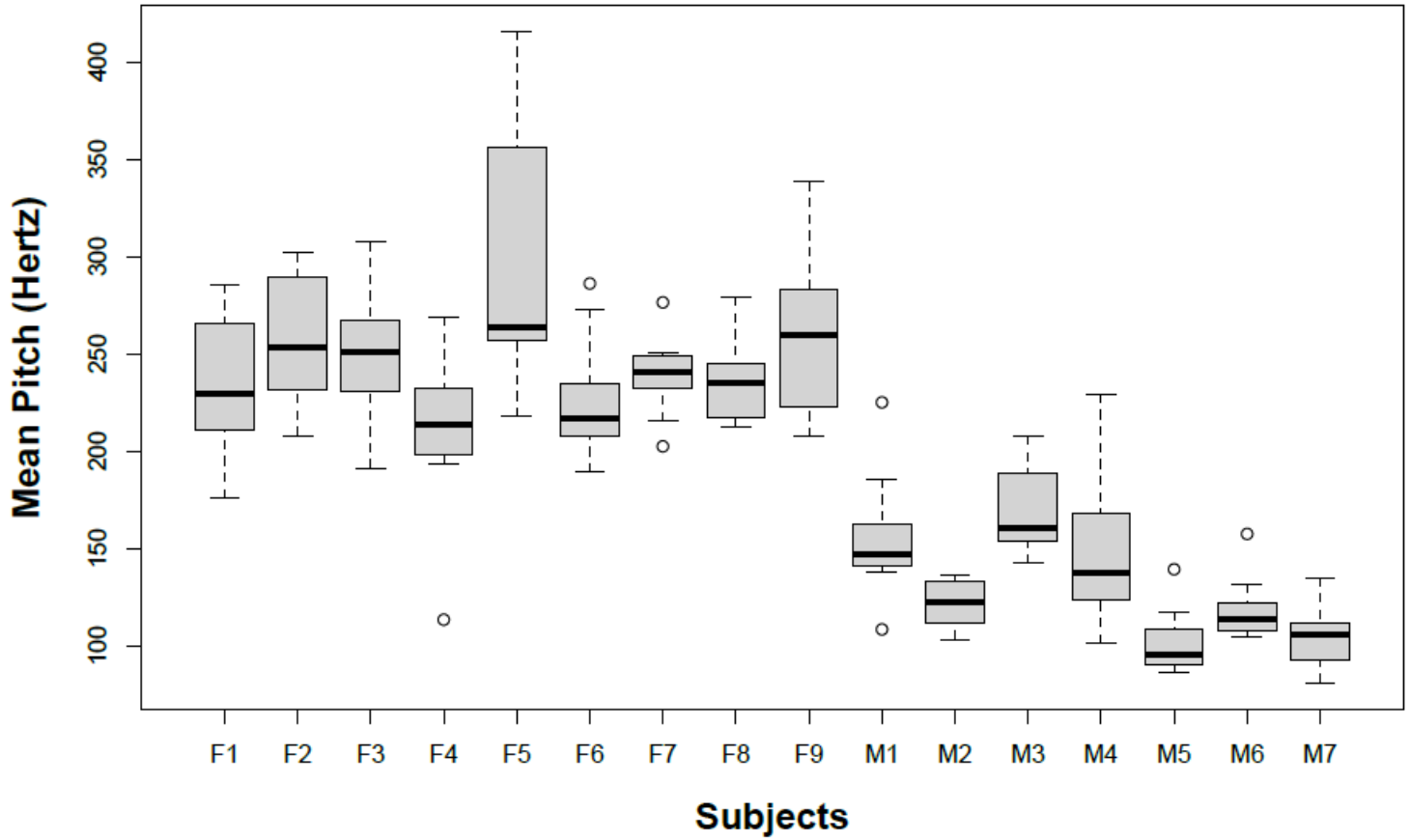
- An extension of linear models that include both **fixed** and **random** effects.
- A random effect is a quantitative variable whose levels are randomly sampled from a population of levels being studied

What Are The Benefits?

- Incorporation of random effects allows for some inter-dependence related to a model factor to be dealt with.

- Random effects allow the model to assume a different baseline response value for each factor.

LMER Tutorial 2 Example



Useful Situations

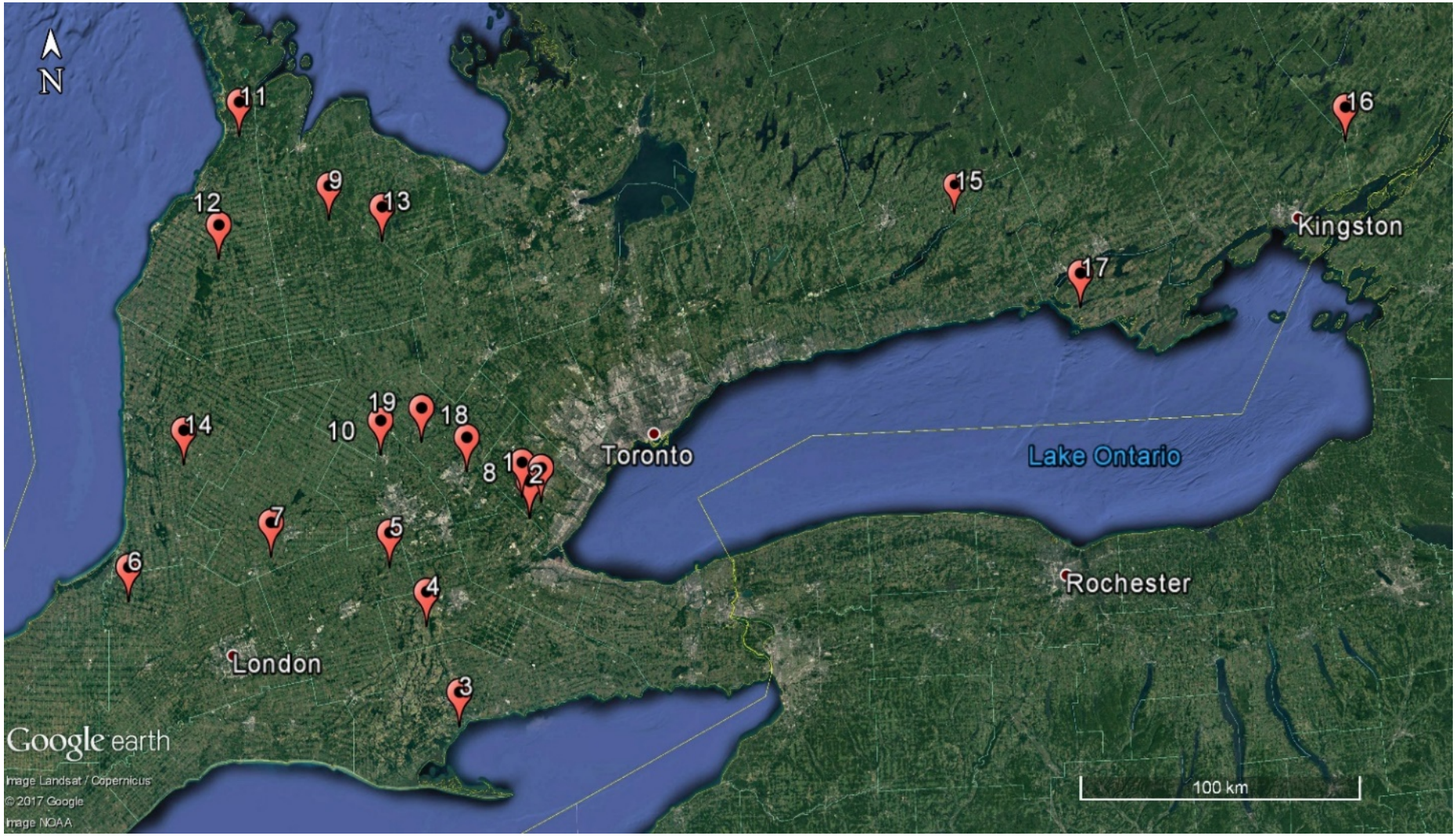
- Experiments with inter-dependence.
- When the experiment does not follow a balanced design or has missing data.
- Studies with repeated measures.
  - Offers greater flexibility

Case Study

Major Objectives

- Obtain current levels of soil organic carbon (SOC) in herbaceous biomass fields (Switchgrass and Miscanthus) to track changes in SOC overtime.
- Compare levels of SOC in soils associated with biomass, agricultural and forest land-uses and then determine the C sequestration potential of biomass crops.

Study Sites



UNIVERSITY of GUELPH

Getting Started

- Required Packages
  - .lme4
  - .LmerTEST
  - .lsmeans (or. emmeans)
  - .MultComp
  - .MultCompView
- Data must be in **long-format**

Creating the Model

- Lu2 = lmer (soc_percent ~ field_type + (1 | farm_id), data = socdata2)
- Summary (landuse_model)

- Soc_percent is soil organic carbon by concentration
- Farm_id indicates the farm location
- Land-uses or “field types” included:
  - 9 Miscanthus fields
  - 21 Switchgrass fields
  - 16 woodlots
  - 14 Agricultural fields

Model Output

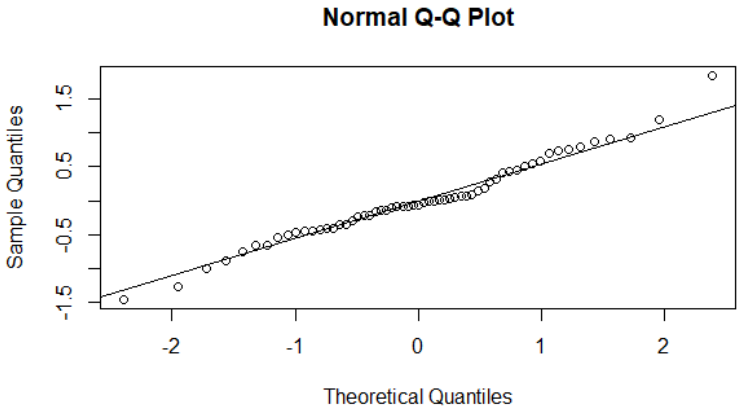
```
Linear mixed model fit by REML
t-tests use  Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: soc_percent ~ field_type + (1 | farm_id)
   Data: socdata2

REML criterion at convergence: 169.1

Scaled residuals:
     Min       1Q   Median       3Q      Max
-1.99508 -0.49472 -0.02504  0.40801  2.35384

Random effects:
 Groups   Name        Variance Std.Dev.
 farm_id  (Intercept) 0.9463   0.9728
 Residual             0.5564   0.7459
Number of obs: 60, groups:  farm_id, 19

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)
(Intercept)   2.2149     0.3055 36.4400   7.250 1.44e-08 ***
field_typem   0.2890     0.3456 41.4900   0.836    0.408
field_types   0.2830     0.2787 40.3000   1.015    0.316
field_typew   2.0488     0.2815 38.5800   7.277 9.55e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) fld_typm fld_typs
field_typem -0.398
field_types -0.522  0.368
field_typew -0.499  0.405    0.571
```

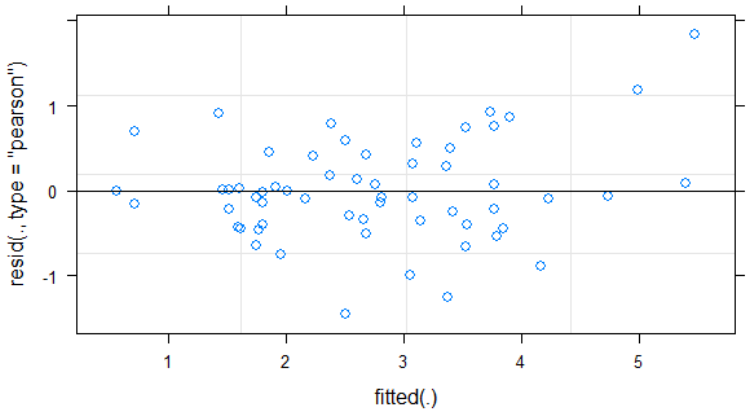
Analysis of Residuals

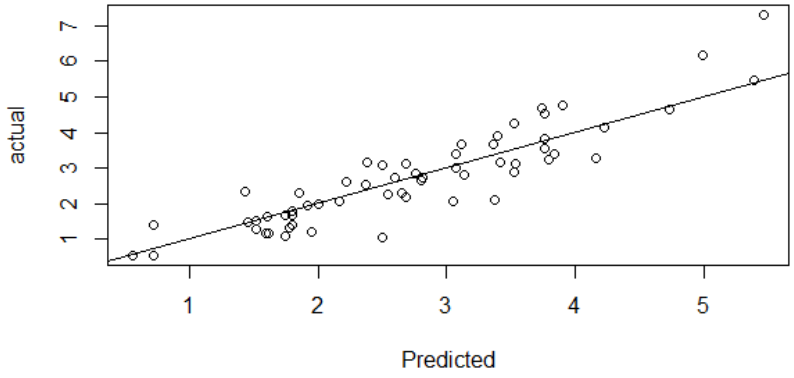
- Assumptions
  - The explanatory variables are related linearly to the response.
  - The errors have constant variance.
  - The errors are independent.
  - The errors are normally distributed.

Analysis of Residuals

- Refer to sample code to get the graphs.



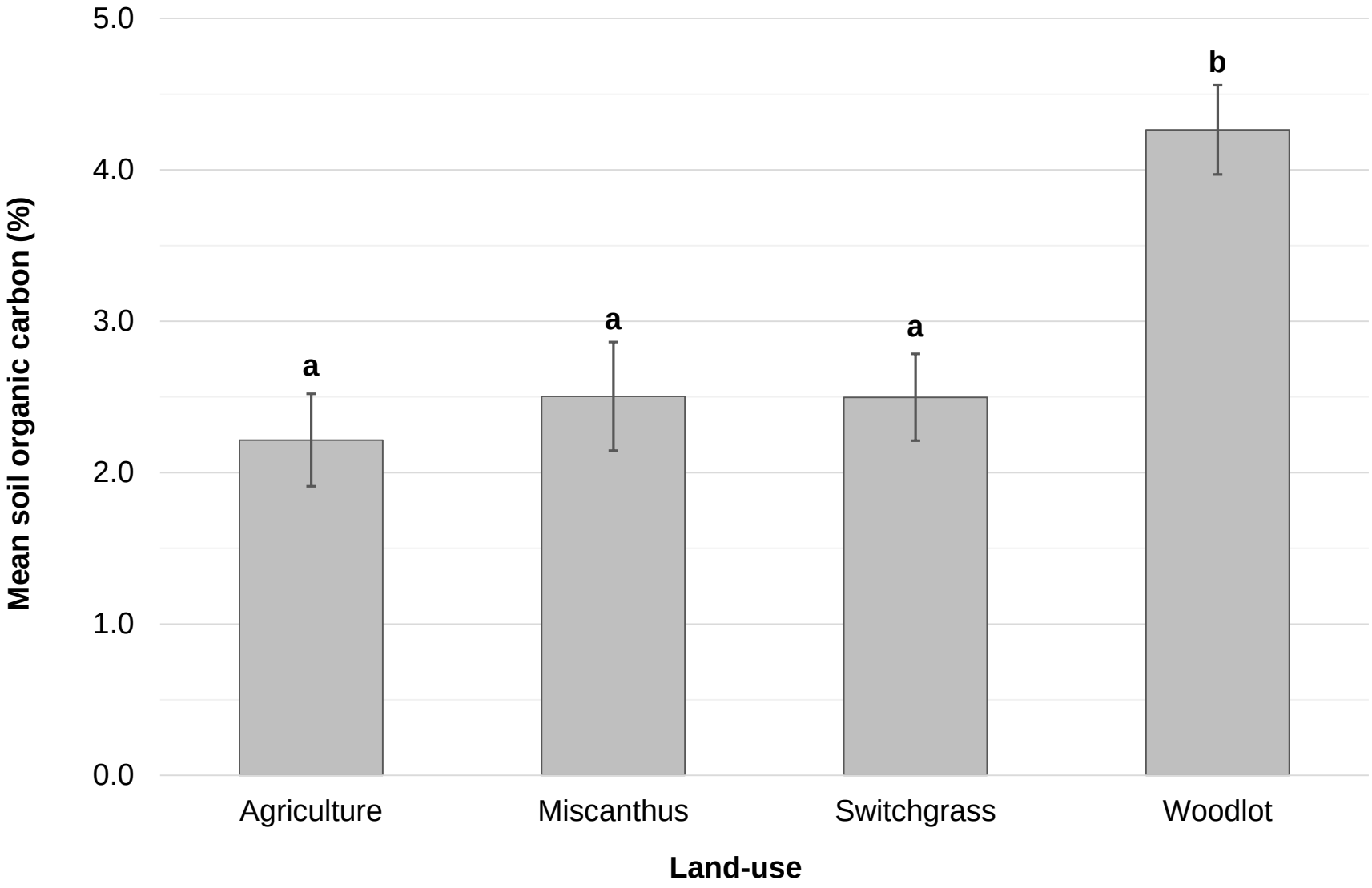




Least Square Mean Comparisons

- lspercent= lsmeans(lu2, list(pairwise ~ field_type), adjust='tukey')

```
> summary(lspercent)
$`lsmeans of field_type`
 field_type   lsmean        SE    df lower.CL upper.CL
 a          2.214885 0.3054888 36.44 1.595582 2.834188
 m          2.503880 0.3587991 48.06 1.776504 3.231256
 s          2.497872 0.2864205 29.42 1.917226 3.078519
 w          4.263644 0.2946079 33.05 3.666400 4.860889

Degrees-of-freedom method: satterthwaite
Confidence level used: 0.95

$`pairwise differences of contrast`
 contrast     estimate        SE    df t.ratio p.value
 a - m    -0.288994629 0.3455779 41.49  -0.836  0.8369
 a - s    -0.282986990 0.2786853 40.30  -1.015  0.7414
 a - w    -2.048758897 0.2815492 38.58  -7.277  <.0001
 m - s     0.006007639 0.3552005 45.02   0.017  1.0000
 m - w    -1.759764268 0.3462133 42.70  -5.083  <.0001
 s - w    -1.765771907 0.2594792 39.09  -6.805  <.0001

P value adjustment: tukey method for comparing a family of 4 estimates
```



Additional Resources

- R LMER tutorials
  - http://www.bodowinter.com/tutorial/bw_LME_tutorial1.pdf
  - http://www.bodowinter.com/tutorial/bw_LME_tutorial2.pdf
- Intro to mixed model paper
  - https://peerj.com/preprints/3113/
- Ben Bolker Interview
  - https://cesess.wordpress.com/2015/08/03/on-the-appropriate-use-of-statistics-in-ecology-an-interview-with-ben-bolker/
- Statistical mistakes (covers GLMMs)
  - https://esajournals-onlinelibrary-wiley-com.subzero.lib.uoguelph.ca/doi/epdf/10.1002/ecs2.1394
- Useful lecture slides / material found off of google
  - www2.stat.duke.edu/~sayan/Sta613/2017/lec/LMM.pdf
  - https://www.st-andrews.ac.uk/media/capod/students/mathssupport/mixedeffectsknir.pdf
  - https://biologyforfun.wordpress.com/2014/04/16/checking-glm-model-assumptions-in-r/
- Advantages of GLMM's over ANOVA's
  - https://onlinelibrary.wiley.com/doi/pdf/10.1002/dev.20245
  - https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1439-037X.2004.00120.x
  - http://www.introspective-mode.org/repeated-measures-anova-versus-linear-mixed-models/