More Mixed Effect "ANOVAs" (GLMMs)

We analyze a fairly simple experiment conducted in complex conditions. A county sprayed for mosquitoes when their own monitoring indicated high counts. Flying beetles were sampled at multiple places before and after spraying occurred, for ~ 1.5 years. Matching samples were also collected where spraying did not happen – as controls in a BACI experiment (Before After Control Impact). In that design, a Night*Sprayed interaction is key because sprayed sites could see reduced beetles on Night 2, whereas unsprayed sites should not (Fig. 1).

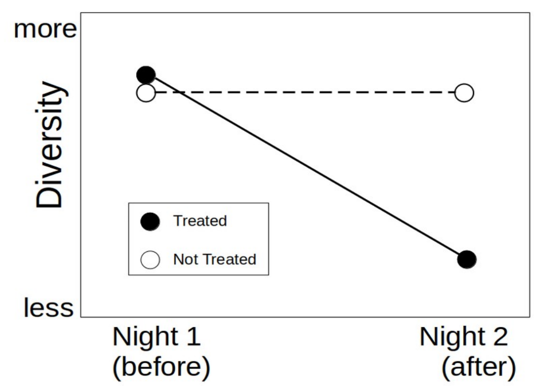

Figure 1: The basic idea of a BACI design, where treatment (here spraying for mosquitoes) reduces diversity but no such effect occurs without spraying.

The question: Does mosquito spraying detectably reduce beetle diversity (here IDed to families), or is variation in beetle diversity due to other factors? We use family richness (FR), which is the count of the number of beetle families in a sample. To answer the question, we have to account for weather, landscape conditions, and random effects of sample location and time during the experiment.

Your mission: Build a model that evaluates the effects of mosquito spraying on beetle FR and that efficiently handles spatiotemporal variation around this real—world experiment.

<u>Getting Organized</u>: These packages will help.

```
library(tidyverse) # dplyr, readr, ggplot, etc. etc.
library(glmmTMB) # to run glmm models. You could use lme4 instead
library(MuMIn) # for some R2 computations and even AIC stuff
library(bbmle) # for AICs
library(performance) # for assumptions and R2 values
```

Get the beetlesdata.csv file on the course web site. I assume below you call it "data." Columns in there include:

- **Site** (shorthand codes), and **Date** of sampling
- **FR**: Family Richness, or the number of beetle families represented in a sample
- **Night** (1 or 2) and **Sprayed** (0 or 1): codes showing if it was the night before (Night 1) or the night after (Night 2) spraying, and whether or not actual spraying happened (0 or 1)
- **AWND** = average wind speed (mph), because high winds may reduce flying beetles
- **PRCP** = precipitation (in), because rain may reduce flying beetles
- **TMAX** and **TMIN** = maximum and minimum temperature that day. These should be highly correlated – *choose one*. Warmer temperatures should have more beetles about.
- **PCs 1-4** summarize land cover around sample points and capture 80% of that variation.
  - PC1: separated residential vs. forested wetlands (34% of variation)
  - PC 2 separated non-forested wetlands and low-density residential lands from other land uses (24% more variation)

- PC3 separated low- and medium-density residential lands from other land (11% more variance)
- PC4 separated aquatic habitats and medium-density residential lands from other land uses (11% more variance). Total land cover variation represented = 80%.

Multiple sample sites and dates were used - place and time are both random effects. So we make those factors. Same for Night and Sprayed – those numbers indicate categories.

```
data$fSite <- factor(data$Site)
data$fDate <- factor(data$Date)
data$fNight <- factor(data$Night)
data$fSprayed <- factor(data$fSprayed)
```

And a base model looks like this:

```
FRbasemodel <- glmmTMB(FR ~ fSprayed*fNight + (1|fSite) + (1|fDate), data=data)
```

Note: This represents a repeated-measures BACI design in varied sites, but doesn't include coavariates

But: This base model has (at least) three problems:

a) It assumes residuals have a gaussian distribution and homogeneous variance (default family=gaussian), but FR is a count. You will have to try other families and evaluate those (e.g., with check_model in performance)
b) It ignores other variables that should also affect flying beetles: weather and habitat (here land cover measures PC1-4).
c) It needs to be compared to a null model and one that includes only random effects of site ID and date. If the model does not represent data better than those, it is clear that spraying has no demonstrable effect on beetle family richness.

YOUR JOB has several parts, listed here in an approximate sequence for your work:

1. Make null and random-effects-only models.
2. Make other models representing alternative hypotheses. For example: a) null; b) random effects only; c) the BACI design for spraying (base model above); d) weather only; d) land cover only; e) combinations of the above, including a full model.
3. Scale weather predictors because they have different units - so you can compare model coefficients later. PCs 1-4 are already scaled.
4. Compare those models with AICc to evaluate evidence for mosquito-spraying effects vs. alternative explanations for beetle diversity patterns.
5. Ensure that your "winning" model meets assumptions, or adjust family until it does. If you must adjust, then do so for all compared models and re-compare. You cannot fairly compare a Poisson-based to a Gaussian, for example.
6. Also ensure that your "winning" model does not have excessive collinearity. If you must drop predictors, make matching adjustments to competing models for fair comparisons.
7. Examine details of the most-plausible model, including relative effects of predictors on beetle family richness. Random effect Standard Deviations are comparable to Estimates below.
8. "Criticize" that model for its R^2 values – how well does your "best" model represent data?
9. Tell the mosquito spraying agency – are they inadvertently affecting nontarget beetles?