Generalized Linear Mixed Model (GLMM) Regressions

You have already used mixed-effect "ANOVAs", with categorical, fixed-effect treatments and continuous covariates. Here we continue GLMMs, but with a more exploratory regression focus - there was not a planned experiment here. We use data from Cronin et al. (2007) describing Harbor seals beached in two coves (sites) counted at multiple dates from 2003 – 2005. These data were analyzed in Ch. 23 of Zuur et al. (2009) as a glm with Poisson distribution and these terms:

Abundance ~ Month + Month^2 + TimeofDay + TimeofDay^2 + Month x TimeofDay + WindDir + Site

They did not try a mixed effects model, **but we will!**

1. Get the Seals.txt (it is tab-delimited) data file from the course web page. I assume you call it "Seals" below.

   Potential predictors of seal abundance (Abun) are:

- weather variables (numerical scores for wind direction, wind speed, and weather)
- Site (haul-out beaches where seals rest). Only two and not random.
- Time, which includes Year/Month/Week/Timeofday (i.e., lots of potential detail). A combination of all those was even made as digital years (see WeekTime in the data).

2. Load packages:

```
library(tidyverse)
library(lattice)
library(glmmTMB)
library(performance)
library(bbmle)
```

3. Make a grid of graphs to see potential predictors of Abun: see any patterns? Maybe quadratic functions to capture curves like in the model above, *such as:*

```
Seals$Month2 <- Seals$Month^2
```

4. Scale any remaining continuous predictors and make Site a factor.

5. Now use the following process to develop a most-efficient mixed effect model. First we find a most-efficient random effects structure. Then we find a most-efficient fixed effects structure while using the "best" random effects. This process comes from Zuur et al. (2009).

   A) Make a "beyond-optimal" model that uses all fixed effects in the data, including effects you made (such as `Month + Month2`), and any interactions you expect. You could use the Zuur et al. model above

   B) Use that fixed effects structure to then find an optimal random effects structure where you:

I. Include `REML = T` in your model code (e.g., `glmmTMB(Abun ~ . . . REML = T`)
II. look at random effects in the `summary()` output and in your `check_model` output: random effects with very low Std. Dev. values and flat QQ plots do not help.

C) Now keep that “optimal” random effects structure BUT re-run you “best” random effectst model after removing the `REML=T` to use the default maximum likelihood estimation.

D) Now work on a most parsimonious fixed effects structure (keep the random effects structure the same here) by comparing alternative models with AIC. Also compare models at this stage to a null model and a random-effects-only model.

Want to graph the random effects? There’s a package for that (it may require lots of other packages first):

```
library(sjPlot)
plot_model(YourModelHere,type = "re",facet.grid = FALSE)
```

What is your most parsimonious model to predict seal abundance in this data set?

INCENTIVE: The person with the most parsimonious and predictive model today gets 2% extra credit added to their final exam score.