

Mixed models

```
library(car)
library(glmmTMB)
```

Overview

Your array of modeling tools is now quite extensive. This lab sets you to work on reproducing the results of some complicated analysis.

Setting

The paper you will analyze is Price et al. (2016). It is available from the Journal of Applied Ecology. The data are in the `glmmTMB` package as `Salamanders` and are available from the Dryad data repository. For a data dictionary and other information about the data see `?Salamanders`.

Linear models

We can begin by asking the data a very simple question: “Were salamanders more abundant in areas that were not mined?” We’ll use `glmmTMB` even though all we are doing is making a linear model with Gaussian errors (one-way ANOVA) so that when we compare models, all of them will have been built with the same tool.

```
Msal1 <- glmmTMB(count~mined, data=Salamanders)
summary(Msal1)
```

- (1) What does this model tell us?

We can add in other predictors from the data.

```
Msal2 <- glmmTMB(count~mined+cover+DOP+Wtemp+DOY+spp, data = Salamanders)
Anova(Msal2)
```

- (2) Based on which variables appear significant in this model, make a graphic or graphics showing the relationships between some of these predictors and the response.

Extending to a GLM

Since the response is a count, it might be better to model this with a Poisson or negative binomial GLM. In `glmmTMB`, there are two negative binomial families. What most others call negative binomial is `negbinom2`, and `negbinom1` is often called quasi-poisson.

```
Msal3 <- glmmTMB(count~mined+cover+DOP+Wtemp+DOY+spp, data = Salamanders,
                 family = "poisson")
Msal4 <- glmmTMB(count~mined+cover+DOP+Wtemp+DOY+spp, data = Salamanders,
                 family = "nbinom1")
Msal5 <- glmmTMB(count~mined+cover+DOP+Wtemp+DOY+spp, data = Salamanders,
                 family = "nbinom2")
```

- (3) Decide which distribution family seems most appropriate, and justify your decision.

Adding a random effect

The data contains repeated samples of the same site, so there is pseudoreplication in the models we have built so far. This can be accounted for by including `site` as a random effect.

```
Msa16 <- glmmTMB(count~mined+cover+DOP+Wtemp+DOY+spp + (1|site),  
                 data = Salamanders,  
                 family = "nbinom1") # Match this to your choice above.
```

- (4) Report the results of including the random effect, and how this changed the model.

Zero inflation

In the paper, they describe modeling occupancy distinctly from abundance. We can include step this via a zero inflation term. We'll follow the authors of the study and include only the environmental covariates `cover`, `DOP`, `Wtemp` and `DOY` in the occupancy part of the model.

```
Msa17 <- glmmTMB(count~mined+cover+DOP+Wtemp+DOY+spp + (1|site),  
                 data = Salamanders,  
                 zi = ~cover+DOP+Wtemp+DOY,  
                 family = "nbinom1") # Again, match this to your choice above.
```

Finding an optimal model and matching the results of the paper

In the paper, they consider the interaction between mining and species `mined*spp` for predicting counts. Additionally, they include the environmental covariates `cover`, `DOP`, `Wtemp` and `DOY` only in the occupancy part of the model, not the abundance part. The conclusions, as summarised on the first page of the paper are:

... mean occupancy probabilities were > 0.85 for all groups in reference reaches, whereas mean occupancy probabilities were relatively lower in MTR/VF reaches (ranging from 0.23 to 0.66).
... means of the difference in occupancy between site types were negative across all groups, although MTR/VF stream reaches were at least 95% less likely to be occupied by spring salamander *Gyrinophilus porphyriticus*, adult southern two-lined salamander *Eurycea cirrigera* and larval dusky salamanders *Desmognathus* compared to reference reaches.

... means of the difference in conditional abundance between MTR/VF and reference stream reaches were negative across all groups; 95% credible interval for difference in conditional abundance covered zero for only one species (red salamander *Pseudotriton ruber*). ... Additionally, MTR/VF reaches had higher ion concentrations, total organic carbon and specific conductance compared to reference reaches.

- (5) Find your optimal model, using all of the ingredients discussed above. Do your results agree with those in the paper? (A Bayesian procedure was used to produce the published results.)

References

Price, Steven J., Brenee' L. Muncy, Simon J. Bonner, Andrea N. Drayer, and Christopher D. Barton. 2016. "Effects of Mountaintop Removal Mining and Valley Filling on the Occupancy and Abundance of Stream Salamanders." *Journal of Applied Ecology* 53 (2): 459–68. <https://doi.org/https://doi.org/10.1111/1365-2664.12585>.