Fixed & Random Effects

Fall 2012

9 October 2012

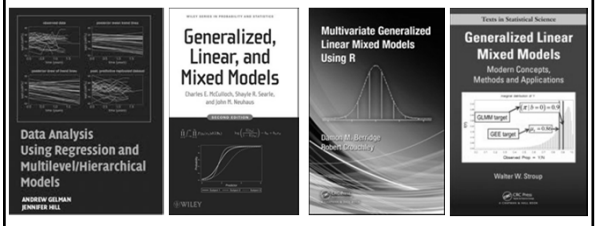
Lecture Goals

- Introduce mixed models
- Explore what we might consider as "random effects" in a statistical analysis of ecological data

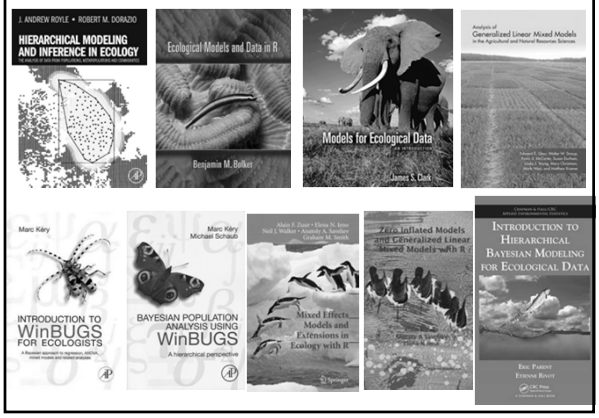
Some Terminology

- Oftentimes, the terminology is confusing
- Similar models may be described as
  - Mixed models
  - Mixed-effect models
  - Hierarchical models
  - Multilevel models
  - Random-effect models
  - Random-coefficient regression

Many Recent Publications…



(… rapidly increasing coverage in ecological modeling texts)



Bolker's take on GLMMs…

- Some challenges:
  - "GLMMs are cutting edge, and the methods for solving them are evolving rapidly." p. 326
  - "The distinction between fixed effects and random effects is murky in any case…" p. 326
  - "… fitting mixed models can be difficult." p. 326
  - "Mixed models are hard to implement." p. 333
- "They are clearly the wave of the future in ecological statistics." p. 334

Bolker 2008

Mixed Models

- Contains ≥ 1 "fixed-effect" parameters and ≥1 "random-effect" parameter
- "Fixed-effect" parameters
  - are still estimated parameters
  - may be associated with categorical or continuous predictor variables
- "Random-effect" parameters
  - associated with ≥1 "random" factors

Fixed vs. Random

- Hypothesis testing
  - H_0: there is no difference between the means among the factor levels → probably "fixed"
  - H_0: there is no variability among the factor levels → probably "random"

Fixed vs. Random

- Scope of inference
  - Are all of the levels of a factor included in the model? Yes? → probably "fixed"
  - In other words, can the levels of a factor be viewed as representing a random sample from a larger population of factors?

    Yes? → probably "random"

Fixed vs. Random Effects

- Definitions in the literature are confusing!
- "it depends" – many definitions do not always provide adequate guidance as to whether or not fixed and/or random effects should be used
- …think carefully about what it is you are trying to do

Just Ignore Random Effects?

"Finally, you can try to convince yourself (and your reviewers, readers, or supervisor) that between-group variation is unimportant by fitting the model ignoring blocks and then examining the variation of the residuals between the blocks both graphically and statistically."

"To justify ignoring between-group variation in the model, you must show that the between-group variation in the residuals is both statistically and biologically irrelevant."

"Biologically relevant variation is an important warning sign even if it is not statistically significant."

Bolker 2008, p. 326

Wagner's 4 "key" questions

1. *Can the factor(s) be viewed as a random sample from a probability distribution?*
2. *Does the intended scope of inference extend beyond the levels of a factor included in the current analysis to the entire population of a factor?*
3. *Are the coefficients of a given factor going to be modeled?*
4. *Is there a lack of statistical independence due to multiple observations from the same level within a factor over space and/or time?*

(If any answer "yes", then you might be thinking about random effects.)

Fixed vs. Random Effects

- Fixed-effects parameters describe the relationship between the predictor and response variable for the *entire population*.
- Random effects are specific to groups, clusters, or individuals within a population.
  - Thus, random effects can be used to model variation at different levels of the data.

A Simple Random-Effects Model

- Assume samples have some underlying structure that we are trying to account for
- e.g., samples within a "family", "block", or "site" are correlated

A Simple Random-Effects Model

$$Y_{ij} = \varepsilon_i + \varepsilon_{ij}$$

$$\varepsilon_i \sim N\left(0, \sigma_b^2\right)$$
$$\varepsilon_{ij} \sim N\left(0, \sigma_w^2\right)$$

ε_i random effect for the i^{th} block

ε_{ij} difference of the j^{th} individual of the i^{th} block from the block's mean

Bolker 2008

Relevant References

Bennington, C. C., and W. V. Thayne. 1994. Use and misuse of mixed model analysis of variance in ecological studies. Ecology 75(3):717-722.

Bolker, B. 2008. Ecological Models and Data in R. Princeton University Press, Princeton, NJ.

Bolker, B. M., and coauthors. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology & Evolution 24(3):127-135.

Venables, W. N., and C. M. Dichmont. 2004. GLMs, GAMs and GLMMs: an overview of theory for applications in fisheries research. Fisheries Research 70:319-337.

Wagner, T., D. B. Hayes, and M. T. Bremigan. 2006. Accounting for multilevel data structures in fisheries data using mixed models. Fisheries 31(4):180-187.

Wagner, T., and coauthors. 2007. A multilevel modeling approach to assessing regional and local landscape features for lake classification and assessment of fish growth rates. Environmental Monitoring and Assessment 130:437-454.