Introduction to mixed effects models in R

Alfonso Allen-Perkins

12/01/2021

Forest Ecology and Management
Volume 212, Issues 1–3, 1 July 2005, Pages 171-183

Structural indicators of spider communities across the forest plantation cycle

Anne G. Oxbrough [a], Tom Gittings [a], John O'Halloran [a], Paul S. Giller [a], George F. Smith [b]

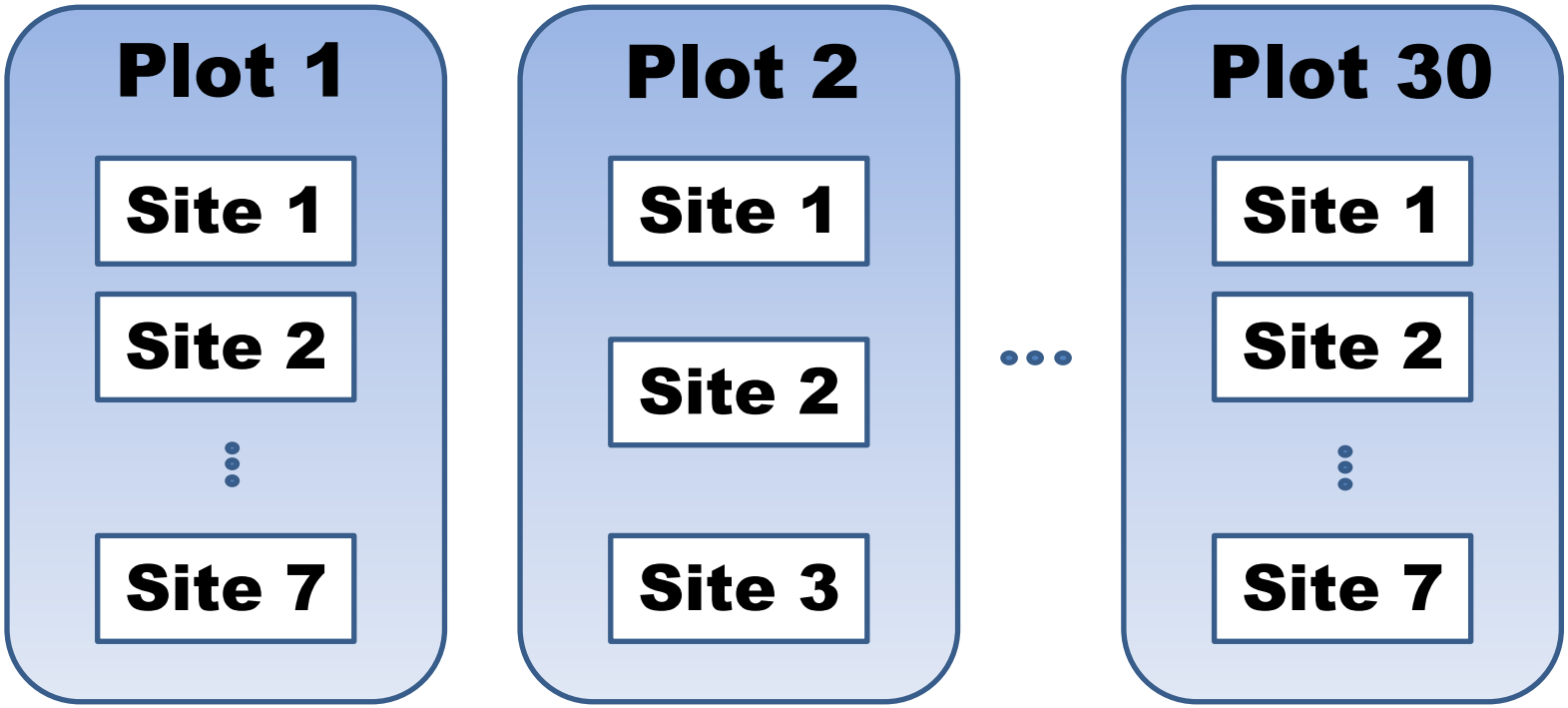
Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.foreco.2005.03.040

Get rights and content

Study sampling design



Linear regression

$$Shannon_i = \beta_1 + \beta_2 HerbLayer_i + \varepsilon_i$$

$$\varepsilon_i \sim N(0, \sigma^2)$$

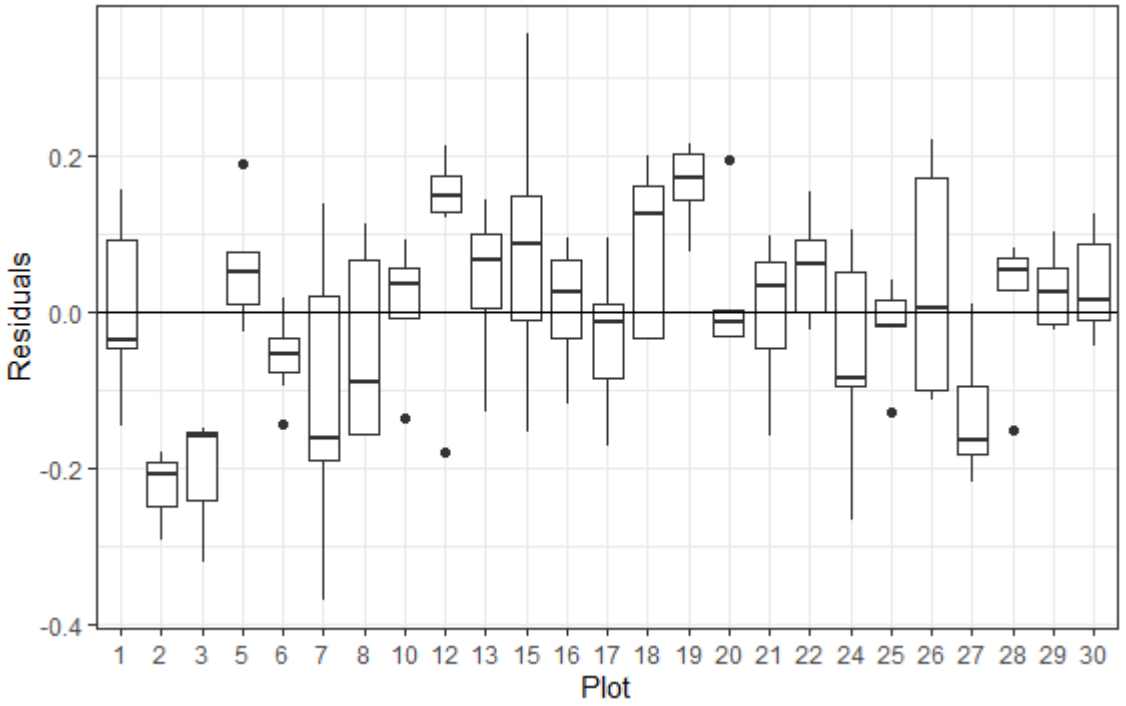
$$Shannon_i \sim N(\mu_i, \sigma^2)$$

$$\mu_i = \beta_1 + \beta_2 HerbLayer_i$$

Model validation

$$Shannon_i = \beta_1 + \beta_2 + HerbLayer_i + \varepsilon_i$$

$$\varepsilon_i \sim N(0, \sigma^2)$$



There is plot effect in the residuals!!

Alternative model

$$Shannon_i = \beta_1 + \beta_2 HerbLayer_i + Plot_i + \varepsilon_i$$

$$= (\beta_1 + Plot_i) + \beta_2 HerbLayer_i + \varepsilon_i$$

$$\varepsilon_i \sim N(0, \sigma^2)$$

$$Shannon_i \sim N(\mu_i, \sigma^2)$$

$$\mu_i = \beta_1 + \beta_2 HerbLayer_i + Plot_i$$

Complete pooling VS no-pooling

Complete pooling

There is a single intercept for all the observations.

$$Shannon_i = \beta_1 + \beta_2 HerbLayer_i + \varepsilon_i$$

No pooling

There is a single intercept for each site.

$$Shannon_i = (\beta_1 + Plot_i) + \beta_2 HerbLayer_i + \varepsilon_i$$

Complete pooling VS no-pooling

Complete pooling

There is a single intercept for all the observations.

$$Shannon_i = \beta_1 + \beta_2 HerbLayer_i + \varepsilon_i$$

Partial pooling

The intercepts are correlated

No pooling

There is a single intercept for each site.

$$Shannon_i = (\beta_1 + Plot_i) + \beta_2 HerbLayer_i + \varepsilon_i$$

Linear mixed model

$$Shannon_{ij} = \beta_1 + \beta_2 HerbLayer_{ij} + a_i + \varepsilon_{ij}$$

$$a_i \sim N(0, \sigma^2{}_{Plot})$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

Linear mixed model

$$Shannon_{ij} = \beta_1 + \beta_2 HerbLayer_{ij} + a_i + \varepsilon_{ij}$$

$$a_i \sim N(0, \sigma^2{}_{Plot})$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

Random

Fixed Part: $\beta_1 + \beta_2 HerbLayer_{ij}$

Random Part: $a_i + \varepsilon_{ij}$

A "mixed model" has both fixed and random parts

Random versus Fixed Effects

Random effects are estimated with partial pooling, while fixed effects are not.
(Andrew Gelman and Jennifer Hill)

Varies by group
Variation estimated with **probability model**

Random intercept model

$$Shannon_{ij} = \beta_1 + \beta_2 HerbLayer_{ij} + a_i + \varepsilon_{ij}$$

$$= (\beta_1 + a_i) + \beta_2 HerbLayer_{ij} + \varepsilon_{ij}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

$$a_i \sim N(0, \sigma^2{}_{Plot})$$

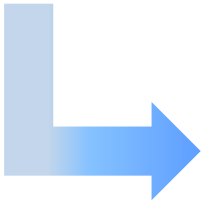
Random intercept model

$$Shannon_{ij} = \beta_1 + \beta_2 HerbLayer_{ij} + a_i + \varepsilon_{ij}$$

$$= (\beta_1 + a_i) + \beta_2 HerbLayer_{ij} + \varepsilon_{ij}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

$$a_i \sim N(0, \sigma^2{}_{Plot})$$



It imposes a correlation structure on all the Shannon values of the same plot.

Random intercept model

$$Shannon_{ij} = \beta_1 + \beta_2 HerbLayer_{ij} + a_i + \varepsilon_{ij}$$

$$= (\beta_1 + a_i) + \beta_2 HerbLayer_{ij} + \varepsilon_{ij}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

$$a_i \sim N(0, \sigma^2{}_{Plot})$$

$$cor(Shannon_{ij}, Shannon_{ik}) = \varphi = \frac{\sigma^2{}_{Plot}}{\sigma^2{}_{Plot} + \sigma^2}$$

Protocol to fit a mixed model

1. Based on prior knowledge of the dependency structure in the data, select the random structure (i.e., the random effects) a priori.

2. Fit the model and investigate which covariates in the fixed part are important.

3. When optimal model has been found, present the numerical output and provide a graphic representation of the model fit.

Zuur, Alain.F, Joseph M. Hilbe, and Elena N Ieno (2013). A Beginner's Guide to GLM and GLMM with R: a frequentist and Bayesian perspective for ecologists, Highland Statistics

Our final random intercept model

$$Shannon_{ij} = 0.994 + 0.034 HerbLayer_{ij} - 0.031 LitterCov_{ij} + a_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \sim N\left(0, 0.106^2\right)$$

$$a_i \sim N\left(0, 0.063^2\right)$$

Random intercept and slope model

$$Shannon_{ij} = \beta_1 + \beta_2 HerbLayer_{ij} + a_i + b_i HerbLayer_{ij} + \varepsilon_{ij}$$

$$= (\beta_1 + a_i) + (\beta_1 + b_i) + HerbLayer_{ij} + \varepsilon_{ij}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

$$a_i \sim N\left(0, \sigma^2{}_{1,Plot}\right)$$

$$b_i \sim N\left(0, \sigma^2{}_{2,Plot}\right)$$

Generalized linear mixed models

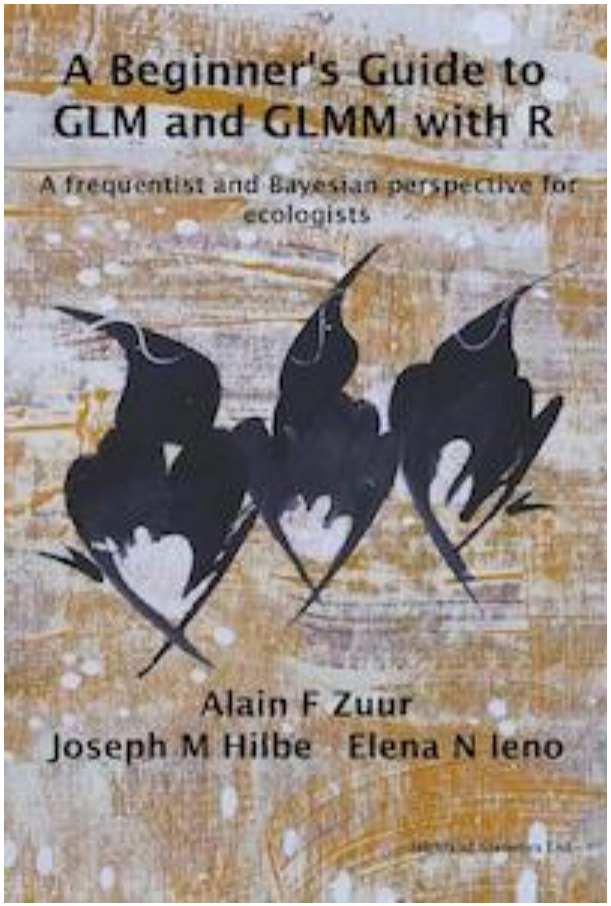
$$Richness_{ij} \sim Poisson(\mu_{ij})$$

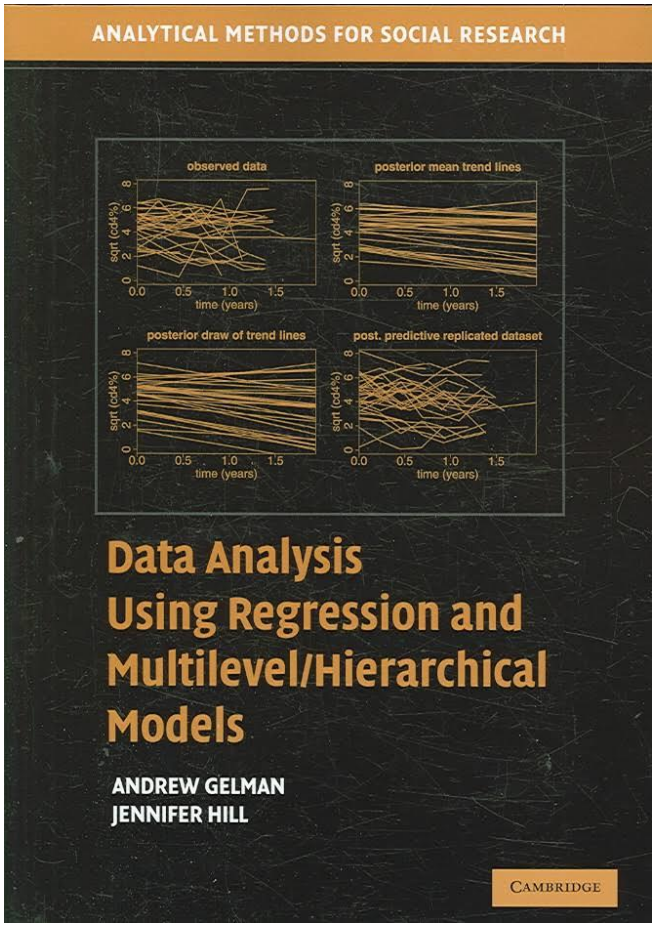
$$E(Richness_{ij}) = \mu_{ij} \qquad var(Richness_{ij}) = \mu_{ij}$$

$$\mu_{ij} = \beta_1 + \beta_2 HerbLayer_{ij} + a_i$$

$$a_i \sim N(0, \sigma^2{}_{Plot})$$

Recommended references





Recommended references

PeerJ

A brief introduction to mixed effects modelling and multi-model inference in ecology

Xavier A. Harrison[1], Lynda Donaldson[2,3], Maria Eugenia Correa-Cano[2], Julian Evans[4,5], David N. Fisher[4,6], Cecily E.D. Goodwin[2], Beth S. Robinson[2,7], David J. Hodgson[4] and Richard Inger[2,4]

[1] Institute of Zoology, Zoological Society of London, London, UK
[2] Environment and Sustainability Institute, University of Exeter, Penryn, UK
[3] Wildfowl and Wetlands Trust, Slimbridge, Gloucestershire, UK
[4] Centre for Ecology and Conservation, University of Exeter, Penryn, UK
[5] Department of Biology, University of Ottawa, Ottawa, ON, Canada
[6] Department of Integrative Biology, University of Guelph, Guelph, ON, Canada
[7] WildTeam Conservation, Padstow, UK

Thanks!!