Reviewer: Ulrike Grömping
Beuth University of Applied Sciences Berlin

Multilevel Modeling Using **R**

This book is from the "Statistics in the Social and Behavioral Sciences" series of Chapman and Hall/CRC, which has quite a broad scope. The book's preface states its goals as follows: "The goal of this book is to provide [...] a comprehensive resource for the conduct of multilevel modeling using the **R** software package." "Our goal is to provide you with a guidebook that will serve as the launching point for your own investigations into multilevel modeling". The content suggests that the book is written for **R** novices with weak statistical background – nothing is taken for granted except for means, variances, correlations, and perhaps the normal distribution. For this audience, the "own investigations into multilevel modeling" are a real stretch. It will be discussed at the end of this review, which audiences might benefit from the book for which purposes.

The book consists of nine chapters: an introduction to the basics of linear models (21 pages) is followed by an introduction to the basics of multilevel data structures and analysis techniques (19 pages). The subsequent three chapters cover fitting two-level models (23 pages), models at three and more levels (20 pages) and models for longitudinal data (15 pages) with **R**. After a chapter on graphing data (20 pages), Chapter 7 (17 pages) briefly introduces the generalized linear model without multilevel structure, followed by a chapter on multilevel generalized linear models (26 pages) and finally a chapter on Bayesian multilevel modeling (31 pages). Each chapter concludes with a brief summary. An appendix (8 pages) provides a basic introduction to **R**. The data sets used in the book examples are downloadable from the book website. I will separately discuss the statistical and **R** aspects of the book, starting with the statistical part.

The book starts with the basics of the linear model, initially presenting most concepts for a simple regression model (Sections 1.1 to 1.4 and 1.6), using the specific sum formulae for that model and refraining completely from introducing matrix calculus. Multiple regression and the variance inflation factor (VIF) are introduced in Section 1.5, and the implementation of linear models in **R** follows in Section 1.7 after a manual calculation of a simple linear model in Section 1.6 (Sections 1.5 and 1.6 should have been swapped for an improved line

of thought). Interactions, incorporation of categorical explanatory variables and model diagnostics are covered in the implementation section. As many very basic things are explained in detail, it seems strange to me that the book does not include a proper explanation for the concept of an interaction: interactions are introduced in a purely technical fashion using two quantitative variables (before categorical variables have even been mentioned), and an interpretation of an interaction is not exemplified anywhere in the book. I think that interactions with at least one categorical variable are reasonably straightforward to explain; such explanation would have also helped to further the understanding of multilevel modeling, e.g., for random coefficient models. For me, this omission in introducing the basics is one of the most unfortunate aspects of this book. The subsequent chapter "Introduction to Multilevel Data Structure" gives a basic introduction to multilevel modeling, which starts with a very brief discussion of the data structure and continues presenting the intraclass correlation coefficient (ICC) in some detail using a manually calculated example (no R code in the whole chapter). Subsequently, the chapter briefly mentions further topics in multilevel data analysis (why necessary, random intercept, random slope, centering of data, estimation with ML and REML, assumptions for multilevel models, overviews of the linear two-level, three-level and longitudinal models). In the chapter summary, the authors claim to have introduced "the basic theoretical underpinnings of multilevel modeling, but not [...] an exhaustive technical discussion". I find the chapter too vague for providing any theoretical underpinnings, basic or not. Rather, it provides a framework of reference for readers with a patchy statistical background. Take for example Section 2.5 on parameter estimation that briefly covers ML and REML: the expression "likelihood" is used without further explanation, assuming an intuitive understanding of the term. The difference between REML and ML is explained as boiling down to a difference in degrees of freedom, analogous to divisors $n - 1$ or n in the variance estimate. Chapter 7 is another mainly statistical chapter. Here, the book explains several variants of logistic regression and Poisson regression and usage of deviance and AIC for model assessment. In Chapter 9, the concept of Bayesian statistics is superficially explained, and the necessity to check for convergence and autocorrelation when using MCMC methods is mentioned. In this section, all analysis examples are accompanied by graphs of MCMC series, presumably in order to stress the importance of checking these. In summary, after some detailed explanation of very basic concepts in Chapter 1, the statistical information is kept at a superficial level. This holds also true for interpretations of R output in the more implementation-oriented chapters.

I now turn to the R aspects of the book (R Core Team 2014). The appendix on R basics is a bit too scant to be useful, in my opinion. Within the book text itself, the first R section on function `lm` should be quite useful for readers with little experience with R: not only does it explain the (very basics of the) model interface, but also the structure of the resulting `lm` object, even though this is done a bit unconventionally, using the `attributes` rather than the `str` function. The graphics chapter (Chapter 6) gives similarly basic advice regarding the `plot` function, and discusses lattice functionality for stratified graphing, which is of course very suitable for nested data. Unfortunately, the chapter stops there and does not provide ideas or implementation hints for further suitable plots in the multilevel or longitudinal data contexts. The multilevel analysis chapters mainly cover usage of the suitable functions from packages **nlme** (Pinheiro, Bates, DebRoy, Sarkar, and R Core Team 2014, function `lme`), **lme4** (Bates, Maechler, Bolker, and Walker 2014, functions `lmer` and `glmer`), **MASS** (Venables and Ripley 2002, function `glmmPQL`) or **MCMCglmm** (Hadfield 2010, function `MCMCglmm`). The general

syntax of functions `lme` and `lmer` is explained in Chapters 3 and 4; the simplicity of the exposition is helpful to readers who might get stuck with more comprehensive presentations. Later chapters – apart from the Bayesian one – build on this base; it surprises me that the Bayesian chapter uses package **MCMCglmm** only, instead of also employing **blme** (Dorie 2014), whose functions `blmer` and `bglmer` directly build on package **lme4** and are thus closer to the previous material. I like the ease with which usage of functions `lme` and `lmer` can be absorbed through many code examples. However, I do not like the fact that the book almost completely ignores the post processing functionality in R. For example, there is no mention of accessor functions like `fitted` or `residuals` – rather it is suggested to readers to access attributes of the output object by using the `$` extractor. Very useful post processing packages like **effects** (Fox 2003) are not mentioned at all. Package **effects** could help interpret interactions, which is an important aspect in data analysis that the book does not explain. Also, unlike many other R books, this book does not provide any additional functions that fulfill tasks not immediately doable with established R code. The book does not even do calculations within R – hand calculations from printed output rather than R calculations from output objects do not set good examples of reproducible research. Also, I would have hoped for little code examples that make it easy to create special graphs for multilevel data, for example, to visualize individual developments in longitudinal data. Instead, the reader is informed that everything is possible in R but is very much left alone as to the "how" ("Many useful resources are available for graphics in R. An internet search will reveal many great (and free) online resources").

Below, I discuss two specific cases for which the book makes wrong statements and misses opportunities for helpful advice. Both are at least partly connected to the above criticism regarding lack of usage of R post processing facilities and of doing calculations within R.

The book entirely avoids the interpretation and visualization of interaction effects. Visualization would have helped in explaining the concept and would (of course) also help in interpreting model results whenever interactions are involved. It would have been easily possible with package **effects** (as previously mentioned) or – for two quantitative variables – by showing contour plots of the response surface that could be obtained from function `contour` after creating predicted values using the `predict` function. The book's ignorance of interactions is an issue in itself, but is also causal for a very misleading statement in Chapter 3: Based on the data example "Achieve" that contains data on various achievements of pupils nested in schools, model 3.5 regresses the dependent variable `geread` (general reading achievement) on the two quantitative explanatory variables `gevocab` (general vocabulary achievement) and `age` and their interaction (product column), allowing a random intercept for each school. The same model is rerun later as model 3.5.C with centered predictor variables. Commenting on the comparison between the two models, the authors first observe the identical model fit and then wrongly conclude the following: "Looking now to the fixed effects of the model, we see some changes in their interpretation. These differences are likely due to multicollinearity issues in the original uncentered model." There is indeed multicollinearity there, as evidenced by the VIF (as an aside, it would have been nice to use this opportunity for demonstrating how to obtain the VIF with R or, even better, to provide a function that calculates the VIF from the output object of function `lme`). In spite of the large VIFs for `age` and the `gevocab` by `age` interaction in the uncentered model, the estimated coefficients of the centered and uncentered models are perfectly compatible; of course, the centered and uncentered explanatory variables themselves have different interpretations, and the presence of the interaction

complicates things. It is ill-advised to interpret individual numeric coefficents in the presence of interactions without considering everything in combination; at least this basic message should have been conveyed to the readers. The model comparison would also have been a great opportunity to introduce response surfaces (functionally and graphically), which would have revealed that both response surfaces perfectly coincide. It would also have been instructive and straightforward to demonstrate to readers how the coefficients of the uncentered model formula $b_0 + b_1 X_1 + b_2 X_2 + b_3 X_1 X_2$ arise from those of the centered model formula $b_{0c} + b_{1c} X_{1c} + b_{2c} X_{2c} + b_{3c} X_{1c} X_{2c}$ by spelling out the centered explanatory variables X_{jc} as $X_j - m_j$ with means m_j, $j = 1, 2$ for the two explanatory variables. As the response surface from both models is exactly the same, correct interpretations of both models must of course be compatible, contrary to the statement in the book!

The second problem case is at least partly reflected in the book's errata list: In Chapters 3 and 4, there are various places where standard deviations output by function `lme` are mistaken for variances (in spite of the unambiguous labeling in the output). This also affects hand-calculation of ICC and R^2 values in some places (p. 45-49, p. 72-74), leading to numerically wrong results. The authors missed several opportunities to avoid this consequential mistake: Better coverage of R post processing facilities would have helped, as the post processing function `VarCorr` from package **nlme** would have extracted the desired variances, both for reporting and for the hand calculations (or – better – manual calculations within R). Referring back from the Bayesian model 9.1 to its frequentist version, model 3.1, would have been an interesting additional consideration for the readers and – as a side effect – would have revealed a strong disconnect in estimated variances which could have made the authors realize their mistake in Chapters 3 and 4. Had the authors also used a function for ICC calculation as provided, e.g., in package **psychometric** (Fletcher 2010), this double-check would have revealed the error in hand calculated ICC values (the function `ICC1.lme` calculates the ICC desired by the authors).

Let us return to an overall perspective. The book is not a treasure trove of insights, neither statistically nor regarding R implementation of multilevel analyses; information density is relatively low, and there are some annoying omissions, particularly regarding data structures and graphing. So, for what purposes is this book worth reading? For one, the book is quite accessible for readers with little statistical and R knowledge. Assuming that these readers will not actually run analyses themselves but rather be members of a team in which someone else is responsible for the analyses, this book will presumably suffice to make them confident team members who can participate in discussing analysis strategies and results. A further audience for whom this book might be useful consists of readers who want to get a general idea of what is available in R for multilevel models and already know the basics of both R and multilevel modeling. For them, the book gives a useful first orientation for getting started with using R for multilevel modeling, especially because of the many different code and output examples. Most readers will need more than this one book for getting started with multilevel modeling in R: for R novices, additional material on using R will certainly be needed. Readers who want to conduct multilevel analyses themselves and are statistically literate without prior knowledge on multilevel modeling might use this book as a starting point for R implementation but should certainly consult one of the various books on multilevel modeling in addition, e.g., Kreft and De Leeuw (1998), Pinheiro and Bates (2000), Bickel (2007), De Leeuw and Meijer (2008), Hox (2010) or Snijders and Bosker (2012).

References

Bates D, Maechler M, Bolker B, Walker S (2014). *lme4: Linear Mixed-Effects Models Using Eigen and S4.* R package version 1.1-7, URL http://CRAN.R-project.org/package=lme4.

Bickel R (2007). *Multilevel Analysis for Applied Research: It's Just Regression!* Methodology in the Social Sciences. Guilford Press, New York.

De Leeuw J, Meijer E (eds.) (2008). *Handbook of Multilevel Analysis.* Springer-Verlag, New York.

Dorie V (2014). *blme: Bayesian Linear Mixed-Effects Models.* R package version 1.0-2, URL http://CRAN.R-project.org/package=blme.

Fletcher TD (2010). *psychometric: Applied Psychometric Theory.* R package version 2.2, URL http://CRAN.R-project.org/package=psychometric.

Fox J (2003). "Effect Displays in R for Generalised Linear Models." *Journal of Statistical Software*, **8**, 1–27. URL http://www.jstatsoft.org/v08/i15/.

Hadfield JD (2010). "MCMC Methods for Multi-Response Generalized Linear Mixed Models: The **MCMCglmm** R Package." *Journal of Statistical Software*, **33**(2), 1–22. URL http://www.jstatsoft.org/v33/i02/.

Hox J (2010). *Multilevel Analysis: Techniques and Applications.* Quantitative Methodology Series, 2nd edition. Taylor & Francis.

Kreft I, De Leeuw J (1998). *Introducing Multilevel Modeling.* Sage Publications, London.

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team (2014). *nlme: Linear and Nonlinear Mixed Effects Models.* R package version 3.1-117, URL http://CRAN.R-project.org/package=nlme.

Pinheiro JC, Bates DM (2000). *Mixed-Effects Models in S and S-PLUS.* Statistics and computing. Springer-Verlag, New York.

R Core Team (2014). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Snijders TAB, Bosker RJ (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling.* 2nd edition. Sage Publishers, London.

Venables WN, Ripley BD (2002). *Modern Applied Statistics with S.* 4th edition. Springer-Verlag, New York.

Reviewer:

Ulrike Grömping
Beuth University of Applied Sciences Berlin

Department II
D-13353 Berlin, Germany
E-mail: groemping@bht-berlin.de
URL: http://prof.beuth-hochschule.de/groemping/