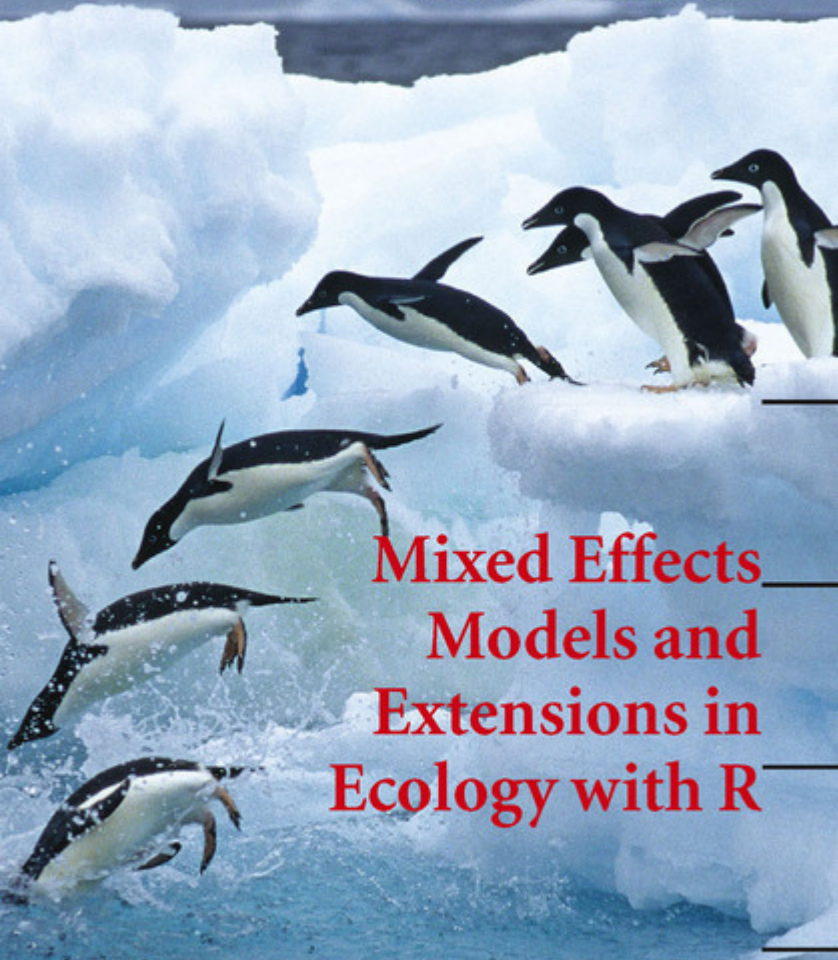
Alain F. Zuur • Elena N. Ieno
Neil J. Walker • Anatoly A. Saveliev
Graham M. Smith



# Mixed Effects Models and Extensions in Ecology with R

Springer

**Statistics for Biology and Health**

*Series Editors:*

M. Gail

K. Krickeberg

J. M. Samet

A. Tsiatis

W. Wong

# Statistics for Biology and Health

*Bacchieri/Cioppa:* Fundamentals of Clinical Research
*Borchers/Buckland/Zucchini:* Estimating Animal Abundance: Closed Populations
*Burzykowski/Molenberghs/Buyse:* The Evaluation of Surrogate Endpoints
*Duchateau/Janssen:* The Frailty Model
*Everitt/Rabe-Hesketh:* Analyzing Medical Data Using S-PLUS
*Ewens/Grant:* Statistical Methods in Bioinformatics: An Introduction, 2nd ed.
*Gentleman/Carey/Huber/Irizarry/Dudoit:* Bioinformatics and Computational Biology Solutions Using R and Bioconductor
*Hougaard:* Analysis of Multivariate Survival Data
*Keyfitz/Caswell:* Applied Mathematical Demography, 3rd ed.
*Klein/Moeschberger:* Survival Analysis: Techniques for Censored and Truncated Data, 2nd ed.
*Kleinbaum/Klein:* Survival AnalysisL A Self-Learning Text, 2nd ed.
*Kleinbaum/Klein:* Logistic Regression: A Self-Learning Text, 2nd ed.
*Lange:* Mathematical and Statistical Methods for Genetic Analysis, 2nd ed.
*Lazar:* The Statistical Analysis of Functional MRI Data
*Manton/Singer/Suzman:* Forecasting the Health of Elderly Populations
*Martinussen/Scheike:* Dynamic Regression Models for Survival Data
*Moyé:* Multiple Analyses in Clinical Trials: Fundamentals for Investigators
*Nielsen:* Statistical Methods in Molecular Evolution
*O'Quigley:* Proportional Hazards Regression
*Parmigiani/Garrett/Irizarry/Zeger:* The Analysis of Gene Expression Data: Methods and Software
*Proschan/LanWittes:* Statistical Monitoring of Clinical Trials: A Unified Approach
*Siegmund/Yakir:* The Statistics of Gene Mapping
*Simon/Korn/McShane/Radmacher/Wright/Zhao:* Design and Analysis of DNA Microarray Investigations
*Sorensen/Gianola:* Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics
*Stallard/Manton/Cohen:* Forecasting Product Liability Claims: Epidemiology and Modeling in the Manville Asbestos Case
*Sun:* The Statistical Analysis of Interval-censored Failure Time Data
*Therneau/Grambsch:* Modeling Survival Data: Extending the Cox Model
*Ting:* Dose Finding in Drug Development
*Vittinghoff/Glidden/Shiboski/McCulloch:* Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models
*Wu/Ma/Casella:* Statistical Genetics of Quantitative Traits: Linkage, Maps, and QTL
*Zhang/Singer:* Recursive Partitioning in the Health Sciences
*Zuur/Ieno/Smith:* Analysing Ecological Data
*Zuur/Ieno/Walker/Saveliev/Smith:* Mixed Effects Models and Extensions in Ecology with R

Alain F. Zuur · Elena N. Ieno · Neil J. Walker ·
Anatoly A. Saveliev · Graham M. Smith

# Mixed Effects Models and Extensions in Ecology with R

Springer

Alain F. Zuur
Highland Statistics Ltd.
Newburgh
United Kingdom
highstat@highstat.com

Elena N. Ieno
Highland Statistics Ltd.
Newburgh
United Kingdom
bio@highstat.com

Neil J. Walker
Central Science Laboratory
Gloucester
United Kingdom
n.walker@csl.gov.uk

Anatoly A. Saveliev
Kazan State University
Kazan
Russia
saa@ksu.ru

Graham M. Smith
Bath Spa University
Bath
United Kingdom
graham.smith@myotis.co.uk

*Series Editors*
M. Gail
National Cancer Institute
Rockville, MD 20892
USA

K. Krickeberg
Le Chatelet
F-63270 Manglieu
France

J. Samet
Department of Preventive Medicine
Keck School of Medicine
University of Southern California
1441 Eastlake Ave. Room 4436, MC 9175
Los Angeles, CA 90089

A. Tsiatis
Department of Statistics
North Carolina State University
Raleigh, NC 27695
USA

W. Wong
Department of Statistics
Stanford University
Stanford, CA 94305-4065
USA

ISSN 1431-8776
ISBN 978-0-387-87457-9 e-ISBN 978-0-387-87458-6
DOI 10.1007/978-0-387-87458-6

Library of Congress Control Number: 2008942429



Printed on acid-free paper

springer.com

*Thanks to my parents for sharing the burden of my university fees – Alain F. Zuur*

*To my friends, colleagues, and former students who are actively committed to the protection and care of the environment – Elena N. Ieno*

*Thanks to my wife Tatiana for her patience and moral support – Anatoly A. Saveliev*

*I would like to thank all family and friends for help and support through times good and bad during the writing of this book – Neil J. Walker*

*To my parents who, even now, continue to support me in everything I do – Graham M. Smith*

# Preface

No sooner, it seems, had our first book *Analysing Ecological Data* gone to print, than we embarked on the writing of the nearly 600 page text you are now holding. This proved to be a labour of love of sorts – we felt that there were certain issues sufficiently common in the analysis of ecological data that merited more detailed description and analysis. Thus the present book can be seen as a 'sequel' to *Analysing Ecological Data* but with much greater emphasis on these very issues so commonly encountered in the collection of, and analysis of, ecological data. In particular, we look at different ways of analysing nested data, heterogeneity of variance, spatial and temporal correlation, and zero-inflated data.

The original plan was to write a text of about 350 pages, but to do justice to the sheer range of problems and ideas we have well exceeded that original target (as you can see!). Such is the scope of applied statistics in ecology. In particular, partly on the back of reviewer's comments, we have included a chapter on Bayesian Monte-Carlo Markov-Chain applications in generalized linear modelling. We hope this serves as an informative introduction (but no more than an introduction!) to this interesting and increasingly relevant area of statistics.

We received lots of positive feedback on the approach and style we used in *Analysing Ecological Data*, especially the combination of case studies and a theory section. We have therefore followed the same approach with this book. This time, however, we have provided the R code used for the analysis. Most of this R code is included in the text, but where the code was particularly long, it is only available from the book's website at www.highstat.com. In the case studies, we also included advice on what to write in a paper.

Newburgh, United Kingdom — Alain F. Zuur
Newburgh, United Kingdom — Elena N. Ieno
Gloucester, United Kingdom — Neil J. Walker
Kazan, Russia — Anatoly A. Saveliev
Bath, United Kingdom — Graham M. Smith
December 2008

# Acknowledgements

The material in this book has been taught in various courses in 2007 and 2008, and we are greatly in debt to all participants who helped improving the material. We would also like to thank a number of people who read and commented on parts of earlier drafts, namely Chris Elphick, Alex Douglas, and Graham Pierce. The manuscript was reviewed by Loveday Conquest (University of Washington), Sarah Goslee (USDA), Thomas Kneib (LMU Munich), Bret Larget (University of Wisconsin), Ruth Salway (University of Bath), Jing Hua Zhao (University of Cambridge), and several anonymous referees. We thank them all for their positive, encouraging, and useful reviews. Their comments and criticisms greatly improved the book.

The most difficult part of writing a book is finding public domain data which can be used in theory chapters. We are particularly thankful to the following persons for donating data sets. Sonia Mendes and Graham Pierce for the whale data, Gerard Janssen for the benthic data, Pam Sikkink for the grassland data, Graham Pierce and Jennifer Smith for the squid data, Alexandre Roulin for the barn owl data, Michael Reed and Chris Elphick for the Hawaiian bird data, Tatiana Rogova for the Volzhsko-Kamsky forestry data, Robert Cruikshanks, Mary Kelly-Quinn and John O'Halloran for the Irish (sodium dominance index) river data, Chris Elphick for the sparrow and California bird data, Michael Penston for the sea lice data, Joaquín Vicente and Christian Gortázar for the wild boar and deer data, Ken Mackenzie for the cod data, and António Mira for the snake data. The proper references are given in the text. We also would like to thank all people involved in the case study chapters; they are credited where relevant.

Michelle Cronin provided the seal photo on the back cover, Joaquin Vicente the deer photo, and Malena Sabatino gave us the bee photo. The photograph of the koalas was provided by Australian Koala Foundation (www.savethekoala.com). The photo on the front cover is from © Wayne Lynch/Arcticphoto.com.

Finally, we would like to thank John Kimmel for giving us the opportunity to write this book and for patiently accepting the 6-month delay. Up to the next book.

# Contents