Alain F Zuur
Elena N Ieno
Anatoly A Saveliev

Beginner's Guide to

Spatial, Temporal, and Spatial-Temporal Ecological Data Analysis with R-INLA

Volume I: Using GLM and GLMM

Published by Highland Statistics Ltd.
Highland Statistics Ltd.
Newburgh
United Kingdom
highstat@highstat.com

ISBN: 978-0-9571741-9-1
First published in February 2017

www.highstat.com

I would like to thank my wife Nandani for giving me the most beautiful present a man can wish for. They say that storks bring babies. It is true!

– Alain F Zuur –

To Norma, Juan Carlos, and Walter for their constant support

– Elena N Ieno –

Special thanks to my wife, friends, and colleagues who make my life more interesting

– Anatoly A Saveliev –

Preface

The authors of this book have been giving statistics courses to ecologists for 15 years. We have taught more than 8,000 scientists. During our courses we cover topics such as R, data exploration, data visualizsation, multiple linear regression, generalised linear models, linear mixed-effects models, generalised linear mixed-effects modelling (GLMM), generalised additive models (GAM), generalised additive mixed-effects models (GAMM), Bayesian analysis and MCMC, and multivariate analysis, among many other topics. Over the years a large number of participants have asked us to teach a module that covers the analysis of spatial, temporal, and spatial-temporal data. Although random effects in GLMM and GAMM can be used to deal with dependency, such an approach is not optimal for spatial, temporal or spatial-temporal data. Although there were various tools available in R, they either required expertise knowledge or required extensive computing time (e.g. MCMC in WinBUGS or OpenBUGS). We therefore elected to stay away from teaching and writing about spatial, temporal, and spatial-temporal data analysis.

It was only after we became aware of material described in Lindgren et al. (2011) that we realised that GLMs and GLMMs, and all their zero-inflated cousins and smoothing cousins, can be extended to spatial, temporal, and spatial-temporal data.

Unfortunately, the literature describing the approach (Integrated Nested Laplace Approximation, abbreviated as INLA) is rather technical. A book published in 2015 by Blangiardo and Cameletti helped us understand the INLA world better. Although we find it an excellent book, it still requires a fair amount of statistical knowledge in order to fully comprehend the material.

Availability of the software package R-INLA has put the application of GLMs and GLMMs on spatial, temporal, and spatial-temporal data within the reach of every scientist. We therefore decided to extend our *Beginner's Guide* book series with a book on the use of R-INLA to analyse spatial, temporal, and spatial-temporal data.

Acknowledgements

We are greatly indebted to all scientists who supplied data for this book. Alexandre Roulin supplied the owl data. Bob Steidl provided the osprey data. Robert Cruikshanks allowed us to use the Irish pH dataset. Christophe Barbraud gave us the Adelie penguins data. Boudjéma Samraoui provided the White Storks data. Juan Timi gave us the Brazilian sand perch parasite data. Yusuke Fukuda provided the crocodile data. Matias Maggi supplied the honey bee mites data. Allesandro Ligas provided the crayfish data. Mette Mauritzen gave us the polar bears data. Chris Smeenk supplied the sperm whale data. Michael Reed gave us the Hawaiian bird data, and Helen Sofaer supplied the orange-crowned

warbler data. We also thank the following authors for making their data publically available: Petty et al. (2015) for the subnivium temperature data, Hopkins et al. (2013) for the chimpanzees data, Sturrock et al. (2015) for the otolith plaice data, Irl et al. (2015) for the plant richness data on La Palma (and who kindly emailed a modified data set), Crozier et al. (2011) for the sockeye salmon, Etheridge et al. (1998) for the historical CO_2 records, and Muller and van Woesik (2014) for the white-pox disease data.

We greatly appreciate the efforts of those who wrote R (R Development Core Team 2016) and its many packages. This book would not have been possible without the efforts of the R-INLA programmers (Rue et al. 2009; www.r-inla.org; Lindgren et al. 2011). We hope that they will keep up the excellent work.

We thank Joseph Hilbe and Thierry Onkelinx for helpful comments on an earlier draft. Special thanks to Christine Andreasen for editing this book.

Data sets and R code used in this book

All data sets used in this book may be downloaded from www.highstat.com/books.htm. All R code also may be downloaded from the website for this book. To open the ZIP file with R code, use the password **[omitted in this file]**

Cover art

The cover drawing is by Jon Thompson (www.yellowbirdgallery.org). Mr Thompson was born in 1939 to Irish parents and has lived most of his life in Scotland. In the 1980s, he was drawn to the Orkney Islands. He is continually inspired by the landscape and bird life of Orkney. He has been creating bird art for 30 years in a variety of media, including drawing, painting, sculpture, and jewellery, never attempting to reproduce nature, but to draw parallels with it. A close-up view of a bird feather is all the inspiration he needs.

Alain F Zuur,
Newburgh, Scotland

Elena N Ieno,
Alicante, Spain

Anatoly A Saveliev,
Kazan, Russia

February 2017

Contents

Spatial models with INLA. Analysis of lattice data. Point patterns. Geostatistics. R-INLA and other packages for Bayesian spatial modelling. Extending R-INLA to fit new models. Implementation.Â Although this model. 6 Spatial Data Analysis with R-INLA with Some Extensions. is not spatial, it can be combined with other spatial models. Using log(Ï„u) instead of simply Ï„u provides some advantages as log(Ï„u) is not constrained to be positive.