Alain F Zuur

Elena N Ieno

Beginner's Guide to

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

Volume II: GAM and zero-inflated models

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

ISBN: 978-0-9571741-4-6
First published in August 2018

www.highstat.com

Opas and Omas pass away, memories to them stay!

– Alain F Zuur –

To Juan Carlos, Norma and Walter who have constantly encouraged me to achieve more challenging tasks in life. Thank you!

– Elena N Ieno –

Preface

Thanks to the likes of Netflix, Amazon Video and Hulu we can now binge-watch a whole television series in one weekend. In front of you, you have Volume II of *Beginner's Guide to Spatial, Temporal, and Spatial-Temporal Ecological Data Analysis with R-INLA*. Volume II, entitled *GAM and zero-inflated models*, is a continuation of Volume I. Volumes I and II consist of a total of 24 chapters. You can binge-read the whole thing in one weekend!

There are analyses in Volume II that we could not perform 10 years ago, simply because the required software did not exist. Thanks to R (R Core Team (2018), R-INLA (Rue et al. 2009) and a large number of packages in R we can now easily apply generalised linear models (GLM), generalised additive models (GAM), generalised linear mixed-effects models (GLMM), and generalised additive mixed-effects models (GAMM) on count data, continuous data, proportional data, and their zero-inflated cousins, with spatial, temporal and spatial-temporal correlation. We can do this for geo-statistical data and for areal data. We can even deal with natural barriers like islands or coastlines.

What was lacking was an explanation and illustration of these techniques for scientists not familiar with, or not interested in, detailed mathematics. There is where our two volumes fill a gap.

The data sets that are analysed in this volume are all real data sets, and each data set comes with its own problems. Some of these data sets were a major challenge even for statisticians to analyse. Yet, they are typical of what biologists tend to sample.

Acknowledgements

We are greatly indebted to all scientists who supplied data for this book. Federico Cortés provided the skate data. Alexandre Roulin supplied the owl data. We thank Rijkswaterstaat for allowing us to use the sandeel data. We thank Jagger et al. (2015) for making their data and R code publicly available. We thank Johan Craeymeersch, Ingrid Tulp and Nicola Tien for providing the sandeel data and cooperating on the sandeel chapter. We also thank David Fifield, April Hedd and Carina Gjerdrum for providing the seabird data and assisting with the seabird chapter. And finally, we thank Adel Heenan and Ivor Williams for making their data available online and working together with us on the coral reef chapter.

We greatly appreciate the efforts of those who wrote R (R Development Core Team 2018) 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 would also like to thank Haakon Bakka for commenting on drafts of Chapters 18 and 20.

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 files with R code, use the password: **< Omitted in online version>**

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

August 2018

Contents

Contributors

Johan Craeymeersch
Wageningen Marine Research
P.O. Box 77, 4400 AB Yerseke
The Netherlands

Carina Gjerdrum
Canadian Wildlife Service
Environment and Climate Change Canada
Dartmouth, NS
Canada

David A. Fifield
Wildlife Research Division
Science and Technology Branch
Environment and Climate Change Canada
Mount Pearl, NL
Canada

April Hedd
Wildlife Research Division
Science and Technology Branch
Environment and Climate Change Canada
Mount Pearl, NL
Canada

Adel Heenan
School of Ocean Sciences
Bangor University
Anglesey, LL59 5AB
United Kingdom

Nicola SH Tien
Wageningen Marine Research
P.O Box 68, 1970 AB Ijmuiden
The Netherlands

Ingrid Tulp
Wageningen Marine Research
P.O Box 68, 1970 AB Ijmuiden
The Netherlands

Ivor D Williams
NOAA Pacific Islands Fisheries Science Center
Honolulu, HI 96818
USA

About the Authors

Alain F Zuur

I was born 2 years before man landed on the moon. When I went to high school, I was not good in much except for mathematics and sport. The nice part about mathematics is that you either 'see it' or you don't. Tears for those who don't; unlimited opportunities for those who do!

I went to university in Groningen, which is in The Netherlands. Life is sometimes directed by chance. A guy named 'Duurt' once gave me a ride to the bus station. Even in Dutch 'Duurt' is a funny name! Just 3 seconds before the bus arrived, Duurt shouted, 'Why don't you try NIOZ for a traineeship; biology is fun!' I hadn't fancied biology much at high school. Still, I decided to give it a chance. I took a 4-hour bus trip from Groningen to the island of Texel and knocked on the door of The Netherlands Institute for Sea Research: 'Anyone interested in a mathematician looking for a traineeship?'

I ended up working with some nice folks at NIOZ and they even funded my PhD. I was based in Aberdeen (Scotland) while doing my PhD. The PhD itself was a huge learning curve. Everything I learned at university didn't work for biological data. ARMA models? Waste of time. Fourier analysis? Come back in 500 years. PCA? Doesn't work.

After finishing my PhD in the millennium year, I got a job at the Marine Laboratory in Aberdeen. Not that I fancy Aberdeen, but its surroundings are beautiful. At the Marine Lab the learning curve was even steeper. People came to my office with complicated designs and expected an answer by yesterday. My line manager at the lab made the catastrophic mistake of allowing me to do a 'homer', a private consultancy job. What started with one 'homer' quickly became multiple 'homers', including statistics courses. After a (short) while this became a second full-time job, so in the year that Dr Who returned on the BBC it was 'adios' and I started to work full time for my own business, Highland Statistics. Core business activities are teaching courses, writing books (and the occasional paper) and consultancy.

Together with my colleague, Elena Ieno, Highland Statistics runs 20–25 statistics courses (covering a wide range of topics) per year. This sounds like a lot but some of these courses are in very (!) nice places. We wrote three books with Springer, and when we got stuck in Northern Europe because of a volcanic eruption we decided to write a 'Beginner's Guide to …' book series as self-publishing authors. That 'self-publishing' means paying statistical referees, editors, printing companies, artists, etc. But the best referees were our course participants who asked 10+ questions about every paragraph that we wrote. Writing these books and running these courses makes me feel that I never stopped my PhD work. Virtually every month there are new statistical methods and new packages in R that are relevant.

Since 2017, my wife Nandani and I are the proud parents of twins Aidan and Naila. Having twins hasn't

slowed me down. Most of this book was written at airports and up in the sky ... and between nap times and feedings!

Alain F Zuur
Highland Statistics Ltd.
9 St Clair Wynd
AB41 6DZ Newburgh
United Kingdom

Elena N Ieno

I have undergraduate and doctoral degrees in biology with a specialisation in marine ecology. I worked on wader feeding ecology on intertidal mudflats and carried out field research in Argentina (South America) and Texel (The Netherlands). Shortly thereafter I joined a bio-diveristy and ecosystem functioning study group as part of my postdoctoral research experience in Scotland, UK. I was also an honorary research fellow in the School of Biological Sciences, Oceanlab at the University of Aberdeen.

It was when I first began my PhD studies that I became aware of the importance of statistical methods in designing, collecting and analysing data. My career continued with a strong focus on applied statistics. After finishing my postdoc in 2003, I joined Highland Statistics. It was there that I shifted from active research work to being exclusively dedicated to the teaching of statistics to undergraduate and postgraduate biology students.

My experience as a course instructor has taken me to a vast number of countries where I have learned ways to close the gap between statistics and biology. This shared worldwide teaching experience gave me the strength and motivation for writing books as part of a contribution to my team. I am the co-author of 10 books on the analysis of ecological data.

My other interests include traveling, trekking and sharing my passion for wildlife photography and conservation with my partner, Walter.

Elena N Ieno
Highland Statistics Ltd.
Box No. 82
Avda. Escandinavia 72
Local 6, BQ 4
Gran Alacant
03130 Santa Pola
Alicante
Spain