Use R!

Series Editors

Robert Gentleman Kurt Hornik Giovanni Parmigiani

More information about this series at http://www.springer.com/series/6991

Daniel Borcard • François Gillet • Pierre Legendre

Numerical Ecology with R

Second Edition

Springer

Daniel Borcard
Université de Montréal
Département de sciences biologiques
Montréal, Québec, Canada H3C 3J7

François Gillet
Université Bourgogne Franche-Comté
UMR Chrono-environnement
Besançon, France

Pierre Legendre
Université de Montréal
Département de sciences biologiques
Montréal, Québec, Canada H3C 3J7

ISSN 2197-5736 ISSN 2197-5744 (electronic)
Use R!
ISBN 978-3-319-71403-5 ISBN 978-3-319-71404-2 (eBook)
https://doi.org/10.1007/978-3-319-71404-2

Library of Congress Control Number: 2017961342

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Ecology is sexy. Teaching ecology is therefore the art of presenting a fascinating topic to well-predisposed audiences. It is not easy: the complexities of modern ecological science go well beyond the introductory chapters taught in high schools or the marvellous movies about ecosystems presented on TV. But well-predisposed audiences are ready to make the effort. *Numerical* ecology is another story. For some unclear reasons, a majority of ecology-oriented people are strangely reluctant when it comes to quantifying nature and using mathematical tools to help understand it. As if nature was inherently non-mathematical, which it is certainly not: mathematics is the common language of all sciences. Teachers of biostatistics and numerical ecology thus have to overcome this reluctance: before even beginning to teach the subject itself, they must convince their audience of the interest and necessity of it.

During many decades ecologists, be they students or researchers (in the academic, private or government spheres), used to plan their research and collect data with few, if any, statistical consideration, and then entrusted the "statistical" analyses of their results to a person hired especially for that purpose. That person may well have been a competent statistician, and indeed in many cases the progressive integration of statistics into the whole process of ecological research was triggered by such people. In other cases, however, the end product was a large amount of data summarized using a handful of basic statistics and tests of significance that were far from revealing all the richness of the structures hidden in the data tables. The separation of the ecological and statistical worlds presented many problems. The most important were that the ecologists were unaware of the array of methods available at the time, and the statisticians were unaware of the ecological hypotheses to be tested and the specific requirements of ecological data (the double-zero problem is a good example). Apart from preventing the data to be exploited properly, this double unawareness prevented the development of methods specifically tailored to ecological problems.

The answer to this situation is to form mathematically inclined ecologists. Fortunately, more and more such people have appeared during the recent decades. The result of their work is a huge development of statistical ecology, the availability

of several excellent textbooks, and the increasing awareness of the responsibility of ecologists with regard to the proper design and analysis of their research. This awareness makes the task easier for teachers as well.

Until the first years of this millennium, however, a critical ingredient was still missing for the teaching to be efficient and for the practice of statistics to become generalized among ecologists: a set of standard packages available to everyone, everywhere. A biostatistics or numerical ecology course means nothing without practical exercises. A course linked to commercial software is much better, but it is bound to restrict future applications if the researcher moves and loses access to the software that he or she knows. Furthermore, commercial packages are in most cases written for larger audiences than the community of ecologists and they may not include all the functions required for analysing ecological data. The **R** language resolved that issue, thanks to the dedication of the many researchers who created and freely contributed extensive, well-designed, and well-documented packages. Now the teacher no longer has to say: "this is the way PCA works... on paper;" she or he can say instead: "this is the way PCA works, now I will show you on-screen how to run one, and in a few minutes you will be able to run your own, and do it anywhere in the world on your own data!"

Another fundamental property of the **R** language is that it is meant as a self-learning environment. A book on **R** is therefore bound to follow that philosophy, and must provide the support necessary for anyone wishing to explore the subject by himself or herself. This book has been written to provide a bridge between the theory and practice of numerical ecology, that anyone can cross. Our dearest hope is that it will make many happy teachers and happy ecologists.

Since they are living entities, both the field of numerical ecology and the **R** language evolve. As a result, much has happened in both fields since the publication of the first edition of *Numerical Ecology with R* in 2011. Therefore, it was time not only to update the code provided in the first edition, but also to present new methods, provide more insight into existing ones, offer more examples and a wider array of applications of the major methods. We also took the opportunity to present the code in a more attractive way, generated by R Markdown in RStudio®, with different colours for functions, objects, arguments and comments.

Our dearest hope is that all this will make many more happy teachers and happy ecologists.

Montréal, QC, Canada — Daniel Borcard
Besançon, France — François Gillet
Montréal, QC, Canada — Pierre Legendre

Contents

About the Authors

Daniel Borcard is lecturer of Biostatistics and Ecology and researcher in Numerical Ecology at Université de Montréal, Québec, Canada. His research interests include Numerical Ecology, Ecology of communities, and Soil Ecology/Zoology.

François Gillet is professor of Community Ecology and Ecological Modelling at Université Bourgogne Franche-Comté, Besançon, France, and visiting professor at École Polytechnique Fédérale de Lausanne, Switzerland. His research deals with the structure, diversity, ecology and dynamics of plant communities.

Pierre Legendre is professor of Quantitative Biology and Ecology at Université de Montréal, fellow of the Royal Society of Canada, and Web of Science Highly Cited Researcher in Environment/Ecology. He is the founder of the field of numerical ecology.

Supplementary Material

All the necessary data files, the scripts used in the chapters, as well as the **R** functions and packages that are not available through the CRAN web site, can be downloaded from our web page http://adn.biol.umontreal.ca/~numericalecology/numecolR/.