Data Analysis in Vegetation Ecology

SECOND EDITION

Otto Wildi

WILEY-BLACKWELL

free companion website

Data Analysis in Vegetation Ecology

Data Analysis in Vegetation Ecology

Second Edition

Otto Wildi

WSL Swiss Federal Institute for Forest, Snow and Landscape Research
Birmensdorf, Switzerland

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

This edition first published 2013

Wiley-Blackwell is an imprint of John Wiley & Sons, formed by the merger of Wiley's global Scientific, Technical and Medical business with Blackwell Publishing.

Registered office: John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial offices: 9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK
111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Library of Congress Cataloging-in-Publication Data

Wildi, Otto.
Data analysis in vegetation ecology / Otto Wildi.
pages cm
Includes bibliographical references and index.
ISBN 978-1-118-38404-6 (cloth) – ISBN 978-1-118-38403-9 (pbk.) 1. Plant communities–Data processing. 2. Plant communities–Mathematical models. 3. Plant ecology–Data processing. 4. Plant ecology–Mathematical models. I. Title.
QK911.W523 2013
581.70285 – dc23

2012047729

A catalogue record for this book is available from the British Library.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Cover image: Image supplied by Author
Cover design by Steve Thompson

Typeset in 10.5/13 Times by Laserwords Private Limited, Chennai, India

First Impression 2013

Plants are so unlike people that it's very difficult
for us to appreciate fully their complexity and sophistication.

Michael Pollan, *The Botany of Desire*

Contents

Preface to the second edition

Successful attempts to include instructions in $\mathcal{R}$ motivated me to prepare a second edition of the book while keeping it basically unchanged in style and content. Hence, I hoped to circumvent yet another introduction to a software environment as done earlier for MULVA-5 (Wildi and Orlóci 1996), which I previously used in many of my examples. I found the syntax of $\mathcal{R}$ to be close to ordinary mathematical notation allowing technical instructions to be minimized. Finally, this book is not an introduction to $\mathcal{R}$. There are many others providing this, such as Crawley (2005), Venables and Ripley (2010), or for advanced users Borcard *et al*. (2011), all highly recommended and referenced. The instructions I included in this second edition are aimed to serve the inexperienced in $\mathcal{R}$, getting technical help from colleagues or experts in installing and initializing $\mathcal{R}$ and loading some packages and functions, including the one I specifically provide for this book (package `dave`). Unintendedly, doing the examples explained in this second edition may even act as a beginners course in $\mathcal{R}$, hopefully with minimum effort.

Writing this second edition was a delicate task too. First, various results had to be reproduced by an entirely different or newly developed software. Only after revising the very last chapter was it clear that all this could be done in $\mathcal{R}$. It is well known that many scientists using $\mathcal{R}$ love it, those who avoid it, fear it. My objective is to encourage newcomers to do the examples and I put every effort into most parsimonious solutions. The instructions and functions I prepared for the book look and hopefully feel simple, hiding the tremendous complexity of the $\mathcal{R}$ environment. In this context I thank my colleagues who gave me technical advice, Thomas Dalang, Dirk Schmatz, Meinrad Küchler and Alan Haynes. The attendees of a course held with

an early version of the book, namely Angéline Bedolla, Elizabeth Feldmeyer, Ulrich Graf, Julia Haas, Alan Haynes, Caroline Heiri, Martina Hobi, Christine Keller, Meinrad Küchler, Helen Küchler, Mathieu Lévesque, Anna Pedretti, Kathrin Priewasser, Anita C. Risch, Marcus Schaub, Martin Schütz, Anna Schweiger, Andreas Schwyzer, Bastian Ullrich and Sonja Wipf, helped me to identify bugs and traps. Again, Anita C. Risch and Martin Schütz were willing to read the whole text critically.

All examples in the book are derived in $\mathcal{R}$ version 2.15.2 (R Development Core Team 2012). Whenever a specific method was missing I wrote a new function to avoid overloading readers with cumbersome code. On the downside every new function represents yet another black box. In the current state the reader will find solutions for all methods presented in the book, although figures may appear a little different: for the book I adapted these to layout requirements using an extended set of plot parameters explained in $\mathcal{R}$ when typing `?plot.default` and further screening for `par`. In the end I devise an $\mathcal{R}$ package for this book: `dave`, the name composed of the initials of the book title (Appendix A). An integrated part of `dave` consists of the many data sets listed in Appendix B. I would like to express my thanks to all authors cited there for giving the right to access these, as far as yet unpublished. Many are real world examples, although, with respect to ongoing research, fairly aged.

While elaborating this second edition I got trapped by the temptation to extend the panoply of methods where functions of other packages were ready to use. This concerns, for example, resemblance measures, classification techniques and ordination methods. In the modelling part I replaced my old fashioned heuristic approach by the now widely used logistic regression techniques including instructions for scenario building, considered important in the time of global change. For newcomers in $\mathcal{R}$ I highly recommend following the instructions quite carefully: $\mathcal{R}$ is very much like a programming language and for the average human brain it is extremely difficult to exactly remember all the details to get the examples running. To support proper use I extended the index considerably to facilitate quick access to all major methods covered in this book. The later will work only when all packages required are loaded, namely `dave`, `labdsv`, `tree` and `vegan` and all considered 'related' upon downloading from a CERAN repository found on the Internet.

I would again like to thank the publications team of Wiley-Blackwell for all the encouragement and support I have received throughout this revision. We agreed that the new edition shall serve users not only in theory but

now also in practice, a combination adding to the complexity of publication. Finally, I express my thanks to my host institution, the Swiss Federal Institute for Forest, Snow and Landscape Research WSL, for providing access to its computer network and literature databases needed to complete this work.

Birmensdorf, 1 October 2012

Preface to the first edition

When starting to rearrange my lecture notes I had a 'short introduction to multivariate vegetation analysis' in mind. It ended up as a 'not so short introduction'. The book now summarizes some of the well-known methods used in vegetation ecology. The matter presented is but a small selection of what is available to date. By focussing on methodological issues I try to explain what plant ecologists do, and why they measure and analyse data. Rather than just generating numbers and pretty graphs, the models and methods I discuss are a contribution to the understanding of the state and functioning of the ecosystems analysed. But because researchers are usually driven by their curiosity about the functioning of the systems I successively began to integrate examples encountered in my work. These now occupy a considerable portion of this book. I am convinced that the fascination of research lies in the perception of the real world and its amalgamation in the form of high-quality data with hidden content processed by a variety of methods reflecting our model view of the world. Neither my results nor my conclusions are final. Hoping that the reader will like some of my ideas and perspectives, I encourage them to use and to improve on them. There is a considerable potential for innovation left.

The examples presented in this book all come from Central Europe. While this was not intended originally, I became convinced the topics they cover are of general relevance, as similar investigations exist almost everywhere in the world. An example is the pollen data set: pollen profiles offer the unique chance to study vegetation change over millennia. This is the time scale of processes such as climate change and the expansion of the human population. Another, much shorter time series than that of pollen data is found in permanent plot data originating from the Swiss National Park that I had the opportunity to look at. The unique feature of this is that it dates back to the year 1917, when Josias Braun-Banquet personally installed the first wooden poles, which are still in place. Records of the full set of species

have been collected ever since in five-year steps. A totally different data set comes from the Swiss Forest Inventory, presented in the last chapter of this book. Whereas many vegetation surveys are merely preferential collections of plot data, this data set is an example of systematic sampling on a grid encompassing huge environmental gradients. It helps to assess which patterns really exist, and whether some of those described in papers or textbooks are real or merely reflect the imagination or preference of researchers scanning the landscape for nice locations. In this case the data set available for answering the question is still moderate in size, but handling of large data sets will eventually be needed in similar contexts. I used the Swiss wetland data set as an example for handling data of much larger size, in this case with $n = 17\,608$ relevés. Although this is outnumbered by others, it resides on a statistical sampling design.

Some basic knowledge of vegetation ecology might be needed to understand the examples presented in this book. Readers wishing to acquire this are advised to refer, for example, to the comprehensive volumes *Vegetation Ecology* by van der Maarel (2005) and *Aims and Methods of Vegetation Ecology* by Mueller-Dombois and Ellenberg (1974), presently available as a reprint. The structure of my book is influenced by Orlóci's (1978) *Multivariate Analysis in Vegetation Research*, which I explored the first time when proofreading it in 1977. Various applications are found in the books of Gauch (1982), Pielou (1984) and Digby and Kempton (1987) and many multivariate methods used in vegetation ecology are introduced in Jongman *et al.* (1995). To study statistical methods used in this book in more detail, I strongly recommend the probably most comprehensive textbook existing today, the second edition of *Numerical Ecology* by Legendre and Legendre (1998). Several books provide an introduction to the use of statistical packages, which are referred to in the appendix. For many reasons I decided to omit the software issue in the main text; upon the request of several reviewers I added a section to the appendix where I reveal how I calculated my examples and mention programs, program packages and databases.

I would like to express my thanks to all individuals that have contributed to the success of this book. First of all Rachel Wade from Wiley-Blackwell, who strongly supported the efforts to print the manuscript in time and organized all the technical work. I thank Tim West for careful copy-editing, and Robert Hambrook for managing the production process. My colleagues Anita C. Risch and Martin Schütz revised the entire text, providing corrections and suggestions. Meinrad Küchler helped in the computation of several examples. André F. Lotter provided the pollen data set. I cannot remember all the people who had an influence on the point of view presented here:

many ideas came from László Orlóci through our long lasting collaboration, others from Madhur Anand, Enrico Féoli, Valério de Patta Pillar, Janos Podani and Helene Wagner. I particularly thank my family for encouraging me to tackle this work and for their tolerance when I was working at night and on weekends to get it completed.

Birmensdorf, 1 December 2009

List of figures

List of tables

About the companion website

This book is accompanied by a companion website:

www.wiley.com/go/wildi/dataanalysis

The website includes:

- Powerpoints of all figures from the book for downloading
- PDFs of tables from the book
- Links to the associated statistical package