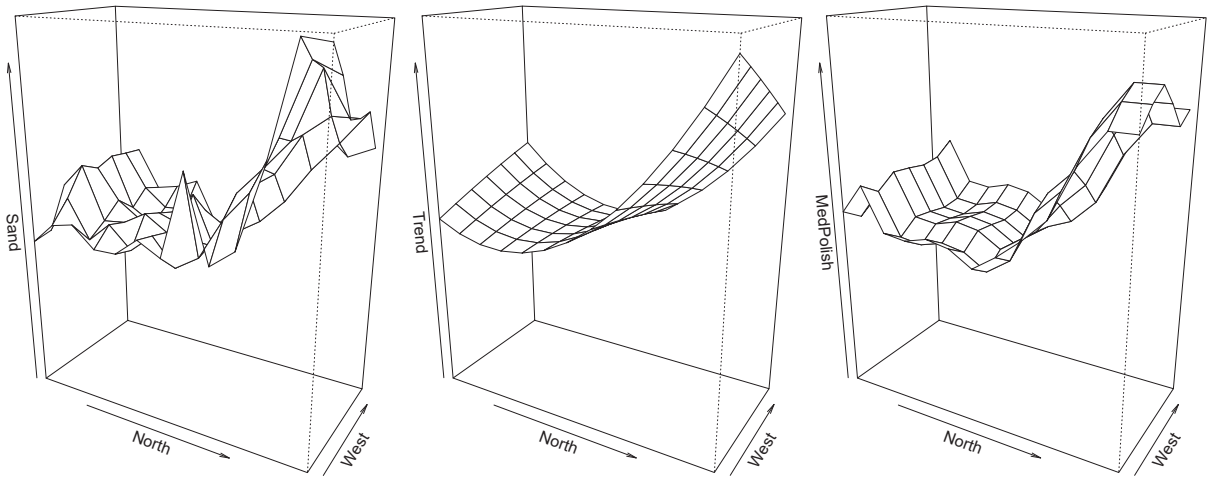
Spatial Data Analysis in Ecology and Agriculture Using R

SECOND EDITION



Richard E. Plant

CRC Press
Taylor & Francis Group

Spatial Data Analysis in Ecology and Agriculture Using R

Taylor & Francis

Taylor & Francis Group

http://taylorandfrancis.com

Spatial Data Analysis in Ecology and Agriculture Using R

Second Edition

Richard E. Plant

Departments of Plant Sciences and Biological and Agricultural Engineering
University of California, Davis

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed on acid-free paper

International Standard Book Number-978-0-8153-9275-0 (Hardback)

Library of Congress Cataloging-in-Publication Data

Names: Plant, Richard E., author.
Title: Spatial data analysis in ecology and agriculture using R / by Richard E. Plant.
Description: Second edition. | Boca Raton, Florida : CRC Press, [2019] | Includes bibliographical references and index.
Identifiers: LCCN 2018040387| ISBN 9780815392750 (hardback : alk. paper) | ISBN 9781351189910 (e-book)
Subjects: LCSH: Agriculture--Statistical methods. | Spatial analysis (Statistics) | R (Computer program language)
Classification: LCC S566.55 .P53 2019 | DDC 338.1072/7--dc23
LC record available at https://lccn.loc.gov/2018040387

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

To Kathie, Carolyn, Suzanne, and Eleanor, with all my love

Taylor & Francis

Taylor & Francis Group

http://taylorandfrancis.com

Contents

Preface to the First Edition

This book is intended for classroom use or self-study by graduate students and researchers in ecology, geography, and agricultural science who wish to learn about the analysis of spatial data. The book originated in a course entitled "Spatial Data Analysis in Applied Ecology" that I taught for several years at UC Davis. Although most of the students were enrolled in Ecology, Horticulture and Agronomy, or Geography, there was a smattering of students from other programs such as Entomology, Soil Science, and Agricultural and Resource Economics. The book assumes that the reader has a background in statistics at the level of an upper division undergraduate applied linear models course. This is also, in my experience, the level at which ecologists and agronomists teach graduate applied statistics courses to their own students. To be specific, the book assumes a statistical background at the level of Kutner et al. (2005). I do not assume that the reader has had exposure to the general linear model or modern mixed-model analysis.

The book is intended for those who want to make use of these methods in their research, not for statistical or geographical specialists. It is always wise to seek out a specialist's help when such help is available, and I strongly encourage this practice. Nevertheless, the more one knows about a specialist's area of knowledge, the better able one is to make use of that knowledge. Because this is a user's book, I have elected on some occasions to take small liberties with technical points and details of the terminology. To dot every *i* and cross every *t* would drag the presentation down without adding anything useful.

The book does not assume any prior knowledge of the R programming environment. All of the R code for all of the analyses carried out in this book is available on the book's companion website, https://psfaculty.plantsciences.ucdavis.edu/plant/sda.htm. One of the best features of R is also its most challenging: the vast number of functions and contributed packages that provide a multitude of ways to solve any given problem. This provides a special challenge for a textbook, namely, how to find the best compromise between exposition via manual coding and the use of contributed package functions, and which functions to choose. I have tried to use manual coding when it is easy or there is a point to be made, and save contributed functions for more complex operations. As a result, I sometimes have provided "homemade" code for operations that can also be carried out by a function from a contributed package.

The book focuses on data from four case studies, two from uncultivated ecosystems and two from cultivated ecosystems. The data sets are also available on the book's companion website. Each of the four data sets is drawn from my own research. My reason for this approach is a conviction that if one wants to really get the most out of a data set, one has to live with it for a while and get to know it from many different perspectives, and I want to give the reader a sense of that process as I have experienced it. I make no claim of uniqueness for this idea; Griffith and Layne (1999), for example, have done it before me. I used data from projects in which I participated not because I think they are in any way special, but because they were available and I already knew them well when I started to write the book.

For most of my career, I have had a joint appointment in the Department of Biological and Agricultural Engineering and a department that, until it was gobbled up in a fit of academic consolidation, bore the name Agronomy and Range Science. Faculty in this second department were fortunate in that, as the department name implies, we were able to

work in both cultivated and uncultivated ecosystems, and this enabled me to include two of each in the book. I was originally motivated to write the book based on my experiences working in precision agriculture. We in California entered this arena considerably later than our colleagues in the Midwest, but I was in at the beginning in California. As was typical of academic researchers, I developed methods for site-specific crop management, presented them to farmers at innumerable field days, and watched with bemusement, as they were not adopted very often. After a while I came to the realization that farmers can figure out how best to use this new technology in the ways that suit their needs, and indeed they are beginning to do so. This led to the further realization that we researchers should be using this technology for what we do best: research. This requires learning how to analyze the data that the technology provides, and that is what this book is about.

I have been very fortunate to have had some truly outstanding students work with me, and their work has contributed powerfully to my own knowledge of the subject. I particularly want to acknowledge those students who contributed to the research that resulted in this book, including (in alphabetical order) Steven Greco, Peggy Hauselt, Randy Horney, Claudia Marchesi, Ali Mermer, Jorge Perez-Quezada, Alvaro Roel, and Marc Vayssières. In particular, Steven Greco provided Data Set 1, Marc Vayssières provided Data Set 2, and Alvaro Roel provided Data Set 3. I also want to thank the students in my course who read through several versions of this book and made many valuable suggestions. In particular, thanks go to Kimberley Miller, who read every chapter of the final draft and made many valuable comments. I have benefited from the interaction with a number of colleagues, too many to name, but I particularly want to thank Hugo Firpo, who collected the data on Uruguayan rice farmers for Data Set 3, Tony Turkovich, who let us collect the data for Data Set 4 in his fields, Stuart Pettygrove, who managed the large-scale effort that led to that data set and made the notes and data freely available, and Robert Hijmans, who introduced me to his raster package and provided me with many valuable comments about the book in general. Finally, I want to thank Roger Bivand, who, without my asking him to do so, took the trouble to read one of the chapters and made several valuable suggestions. Naturally, these acknowledgments in no way imply that the persons acknowledged, or anyone else, endorses what I have written. Indeed, some of the methods presented in this book are very *ad hoc* and may turn out to be inappropriate. For that, as well as for the mistakes in the book and bugs in the code that I am sure are lurking there, I take full responsibility.

Davis
California

Preface to the Second Edition

I am grateful for the opportunity to write a second edition of this book. R is very dynamic, and there have been sufficient changes to degrade the book's usefulness for learning R. There have also been major advances on several fronts in spatial data analysis. The most dramatic changes have been in the analysis of spatiotemporal data. These have been sufficient for me to completely revise that chapter. Major advances have also been made in the application of Bayesian methods to spatial data. The use of the generalized additive model has been rapidly gaining ground in ecology. Finally, but not least importantly, the use of packages associated with the tidyverse of Hadley Wickham and his colleagues have made graphical analysis much simpler. I have incorporated all of these, although, for reasons I have elaborated in Chapter 2, I have decided to continue to use the traditional R graphics to construct the figures in this book. Of course this means that some things have had to be deleted to make room. Most prominent has been the removal of the section on principal components analysis. The substitution of a section on the generalized additive model has prompted me to switch the order of this chapter and the one on linear models. I have placed the section on principal components analysis on the books companion website, https://psfaculty.plantsciences.ucdavis.edu/plant/sda2.htm in a section called "Additional Topics." Additional material to accompany the book can be accessed at https://psfaculty.plantsciences.ucdavis.edu/plant/additionaltopics.htm. I hope to add discussions of other relevant topics there as well. I have also had the opportunity to correct numerous errors in the first edition; I hope I have not introduced too many new ones into the second. In addition to the people I acknowledged in the first edition, I want to thank Meili Baragatti, James Graham, and Andrew Latimer for their thoughtful reviews of the first edition. I would also like to thank my editors, John Sulzycki and Alice Oven, for their guidance and assistance. They have been a true pleasure to work with.

Davis
California