

## Book reviews

Author(s)/Editor(s): Steven P. Millard, Nagaraj K. Neerchal  
Year of Copyright: 2001  
Title: *Environmental Statistics with S-PLUS*  
Publisher: CRC Press  
Pages: 778  
Hardcover: Yes  
Paperback:  
ISBN: 0-8493-7168-6

*Reviewed by:* Shivanand Balram, Suzana Dragičević; Department of Geography, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, V5A 1S6 Canada

Statistics is traditionally a difficult subject to master. Usually statistical methods are presented to us in a predominantly piece-meal approach with hints about the inter-relationships between the concepts, and just about when we are beginning to come to grips with these concepts, the course and/or textbook suddenly comes to an end. This leaves us unfulfilled and searching for the bigger picture on how the statistical pieces fit together.

*Environmental Statistics with S-PLUS* by Millard and Neerchal is one of the few statistics reference textbooks that offer us a seemingly complete journey into the intricate methods, definitions, and concepts of the statistics world. The comprehensiveness and logical flow of the content from the basics to advanced statistics applications provides a breath of fresh air for the reader as the bigger statistical picture unfolds through 778 pages of theory, applications and exercises. The thirteen chapters of the book begin with the issues involved in quantitative data conceptualisation, proceeds through data sampling, and conclude with statistical data analysis and decision-making. In each of the chapters, statistics concepts are treated in an exhaustive manner, and an abundant supply of exercises and examples using data (such as soil sampling, groundwater contamination, air quality monitoring) from environmental management are presented to illustrate the concepts. The data is available from the CRC Publishers website at [www.crcpress.com](http://www.crcpress.com). Moreover, the *EnvironmentalStats*, *S + SpatialStats* and *S-PLUS for Arcview* add-in modules of the S-PLUS software are used where appropriate to demonstrate how to implement the statistics methods in their user-friendly GUI-driven software environments.

The book begins with an introduction and overview in Chapter 1. Here, the inter-relationships between the statistical methods that are discussed in the book are broadly linked by the United States Environmental Protection Agency's Data quality objectives (DQO) process of environmental monitoring and analysis. The DQO flowchart brings the statistics pieces together and clearly defines the contribution of each piece towards achieving the decision outcome.

Chapter 2 discusses the seven planning steps of the DQO process and their links to the scientific method of hypothesis formation, experiment, data analysis, and decisions. Beginning the discussion from the basics of population and samples, random sampling strategies and their associated descriptive statistics measures are gradually outlined to set

the framework for effective data collection. Ways of describing and understanding the data values becomes important once the data has been collected. Chapter 3 gives a detailed overview of descriptive statistics measures and their interpretations together with exploratory data analysis techniques for interrogating the data. A useful feature in this chapter is a comprehensive summary of the characteristics of graph types used for representing single and multiple variables.

The varied issues relating to using sample data to describe the population form the focus of Chapters 4–8. Here probability distributions, measurement and sampling errors, parameter distributions, prediction intervals, hypothesis testing, and sample design are covered in a fair bit of detail. This emphasizes the important role these components play in producing reliable information and output decisions. The development of linear data relationships is discussed in Chapter 9. As such regression analysis and associated concepts are discussed in their entirety. One key challenge in monitoring studies is that the measurement data is usually censored observations—that is the data is above or below some threshold level. While this limits the information content of the observations, Chapter 10 outlines statistics techniques for dealing with the analysis and reporting of censored data.

Monitoring environmental events demand consideration of both spatial and temporal variations. In handling the temporal variations, a single spatial unit is monitored across time. Chapter 11 outlines the details of conducting time series analysis and discusses this in light of autocorrelations among temporal variables. Dealing with spatial variations in time has been the focus of statistical scientists and geographers since the 1960s. As a result of trans-disciplinary research efforts, geostatistics has been demonstrated to provide a reliable means for dealing with spatio-temporal data. Chapter 12 is about spatial statistics and how geostatistics and spatial autocorrelation analysis can be performed on geostatistical, lattice and point pattern data using the *S + SpatialStats* add-in module.

How do we make wise decisions using the information we obtain from our observations and subsequent statistical analysis? The answers to these are contained in Chapter 13 that deals with Monte Carlo simulation and risk assessment. In addition, uncertainty and sensitivity analysis is used to interrogate the reliability of the simulation models leading to model validity and risk assessments. This provides decision-makers an objective way of examining the consequences of their decisions. Unfortunately, this chapter focuses exclusively on Monte Carlo simulations to the exclusion of other approaches.

It is initially a surprise to read in the Preface that a statistics book of this nature dealing with complex concepts such as Monte Carlo simulations and quantitative risk assessment does not demand advanced statistical knowledge from the reader. It is it easy to see why this is the case. Each concept is developed from the basics to the advanced and in many ways the book is self-contained. In summary, this book is an excellent resource that will find favor with basic and advanced statistics audiences across many disciplines, but especially those from environmental/physical science and engineering. Basic statistics readers will enjoy the comprehensive problem-solving approach used to introduce the statistical techniques, while advanced readers will appreciate the more demanding topics that are presented in the later chapters. This book should be necessary reading for all those involved in some way or another with the collection, communication, and analysis of quantitative data.

Author: Mark R.T. Dale  
Year of Copyright: 2000  
Title: *Spatial Pattern Analysis in Plant Ecology*  
Publisher: Cambridge University Press  
Pages: 326  
Paperback:  
ISBN: 0-521-79437-4

*Reviewed by:* Alfred Stein, Wageningen University.

This book is the paperback version of a hardback edition, published in 1999. It addresses various approaches towards sampling and analyzing of spatial patterns, with a clear emphasis on plant ecology patterns. The book contains nine chapters. After an introductory chapter and a chapter on sampling, the next three chapters concern an analysis of one-dimensional patterns, i.e., patterns along a transect. The material is presented starting from one species, through two species towards multiple species. Chapter 6 concerns 2-dimensional patterns, Chapter 7 is on point patterns and Chapter 8 on patterns following an environmental gradient. The book ends with a conclusion chapter and some ideas on future directions. The book is well illustrated with a range of plant ecological examples and illustrations.

The book starts with the work of A.S. Watt in the 1940s. Central in data analysis procedures are (i) the two term local quadrat variance, (ii) the variogram as a measure for continuous data and (iii) Ripley's  $K$ -function for point patterns. Much emphasis is put on the local quadrat variance with some generalizations, whereas the other two analysis procedures are treated in less detail, as they can be found elsewhere. Equations are given for the 2-term local quadrat variance equations, that are generalized and applied to both real and artificial data. The authors claim that it is a useful descriptive statistic to discover at which scale the largest variation occurs. The equations are not particularly transparent, though, relying on multiple summation indices. Possibly, a presentation from an ANOVA point of view might have been preferred. Remarkably as well is the absence of the  $F$ - and the  $G$ -statistics when describing spatial point patterns, of which no motivation is given.

The chapter on sampling is rather isolated in the book, although clearly necessary, and contains only some general guidelines without becoming very specific. No attention is paid to modern optimal sampling methods, and only some general common sense directives are presented.

The question is then whether the book is useful to the plant ecological community. The positive point about the book is that it summarizes in a single volume much of the relevant literature. It is therefore relatively easy to find one's way, and also to find an abundance of challenges for methodological research. But I also have some doubts, as the book may put a researcher on the wrong track. When studying patches, for example, I would be most interested to find relations with for example topography, hydrology or soil variables. This might require some form of regression, and this is another topic that is not addressed.

Author(s)/Editor(s): W. Piegorsch and A.J. Bailer  
Year of Copyright: 1997  
Title: *Statistics for Environmental Biology and Toxicology*  
Publisher: Chapman & Hall/CRC  
Pages: 579  
Hardcover: \$74.95  
ISBN: 0-412-04731-4

*Reviewed by:* Emili Garcia-Berthou, Department of Environmental Sciences, University of Girona, Spain.

The goal of this book is “to work from problems in environmental toxicology that have motivated statistical/environmental advances, and present a selection of corresponding statistical techniques.” The authors “have targeted an audience that includes both environmental scientists and statisticians. These include: advanced graduate students in applied environmental fields, intermediate graduate students and advanced undergraduates in statistics, and also postgraduate applied researchers [ . . . ]” (pp. xi–xii). The book begins by introducing probability distributions (Chapter 1), fundamentals of statistical inference (Chapter 2), experimental design (Chapter 3), two-sample tests and confidence intervals (Chapter 4), and methods for multiple comparisons, including analysis of variance and methods under binomial and Poisson samplings (Chapter 5). However, most of the book focuses on tests for describing trends (Chapter 6), dose-response modeling (Chapter 7), an introduction to generalized linear models (Chapter 8), analysis of cross-classified categorical data (Chapter 9), the incorporation of historical control information (Chapter 10), and survival-data analysis (Chapter 11). If you would like to see more details before buying this book, 27 sample pages, including the table of contents and the full index, are available free at <http://www.amazon.com>.

This is an excellent book with many virtues: very careful writing and editing; a summary of each chapter; most data are real (not fake) and with reference to sources (Cobb, 1987); there is a combination of theory, application and computer implementation; many examples and exercises are included (Cobb, 1987); and there is an extensive up-to-date bibliography. Another pleasant feature of the book, which should be almost compulsory these days but it is still not the rule, is that a list of errata and additional SAS and S-Plus programs are available at the authors’ website (<http://www.users.muohio.edu/baileraj/book/codetable.htm>).

If I had to point out some weak points of the book, I would say that it is quite mathematical and will be not very user-friendly for applied environmental scientists. In particular, I missed more practical advice for choosing among methods and more papers from biological journals. There are plenty of exercises (383) but many are for statisticians only, particularly in the first chapters (most of them marked with an asterisk). Unfortunately, solutions to exercises are given in a separate volume (Piegorsch and Bailer, 1999) of limited circulation (available free from the publishers for scholars adopting the book). The book will not be so helpful if you are not using SAS or S-Plus. Finally, most examples in the book are toxicological, with dose-response data as the paradigm, and many methods and topics important in environmental biology (e.g., spatial and temporal data, multivariate analysis, ANOVA designs beyond the oneway) are not dealt with. This is unavoidable, as is acknowledged in the preface (pp. xv), but the title

should be less ambitious to reflect this fact. Also inevitable are some discrepancies with the suggested analyses; for instance, I would recommend an analysis of covariance to the data of Table 5.1 (instead of an analysis of variance of a ratio).

In short, this is an excellent book, an essential reference for toxicologists and readers of *Environmental and Ecological Statistics*. If you are an applied environmental scientist, unafraid of mathematical details and willing to use SAS or S-Plus, this is also a useful book.

## References

- Cobb, G.W. (1987) Introductory textbooks: a framework for evaluation. A comparison of 16 books. *Journal of the American Statistical Association*, **82**, 321–39.
- Piegorsch, W.W. and Bailer, A.J. (1999) *Solutions Manual for Statistics for Environmental Biology and Toxicology*, Chapman & Hall/CRC, Boca Raton, FL.