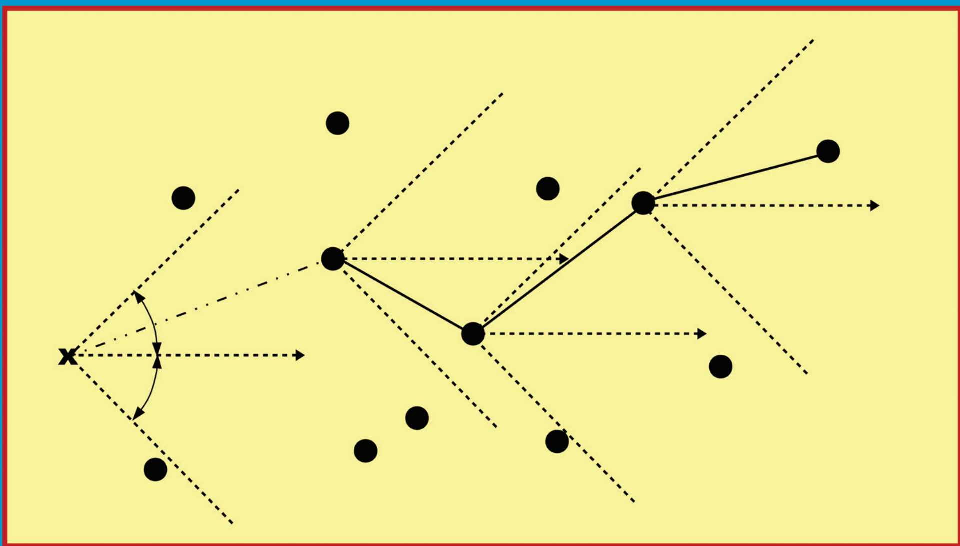
Chapman & Hall/CRC
Applied Environmental Statistics

Introduction to Ecological Sampling



Edited by
Bryan F. J. Manly
Jorge A. Navarro Alberto

CRC Press
Taylor & Francis Group
A Chapman & Hall Book

Introduction to Ecological Sampling

CHAPMAN & HALL/CRC
APPLIED ENVIRONMENTAL STATISTICS

Published Titles

Michael E. Ginevan and Douglas E. Splitstone, **Statistical Tools for Environmental Quality**

Timothy G. Gregoire and Harry T. Valentine, **Sampling Strategies for Natural Resources and the Environment**

Daniel Mandallaz, **Sampling Techniques for Forest Inventory**

Bryan F. J. Manly, **Statistics for Environmental Science and Management, Second Edition**

Bryan F. J. Manly and Jorge A. Navarro Alberto, **Introduction to Ecological Sampling**

Steven P. Millard and Nagaraj K. Neerchal, **Environmental Statistics with S Plus**

Wayne L. Myers and Ganapati P. Patil, **Statistical Geoinformatics for Human Environment Interface**

Éric Parent and Étienne Rivot, **Introduction to Hierarchical Bayesian Modeling for Ecological Data**

Song S. Qian, **Environmental and Ecological Statistics with R**

Thorsten Wiegand and Kirk A. Moloney, **Handbook of Spatial Point-Pattern Analysis in Ecology**

Chapman & Hall/CRC
Applied Environmental Statistics

Introduction to Ecological Sampling

Edited by
Bryan F. J. Manly
Western EcoSystem Technology, Inc.
Laramie, Wyoming, USA

Jorge A. Navarro Alberto
Universidad Autónoma de Yucatán
Mérida, Mexico

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group an **informa** business
A CHAPMAN & HALL BOOK

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

Version Date: 20140603

International Standard Book Number-13: 978-1-4665-5515-0 (eBook - PDF)

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

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

Contents

Preface

The origins of this book go back about 20 years when Lyman McDonald and Bryan Manly started giving courses for biologists on environmental and ecological sampling. For this purpose, they wrote notes about the different topics for use by the students in the courses and gave lectures based on the contents of the notes. These notes were subsequently used in 2006 when Bryan Manly and Jorge Navarro organized a workshop at the University of Yucatan, Mexico, directed to a biologically inclined audience. Then, in 2007 Bryan Manly was asked to set up an Internet course on environmental and ecological sampling for the Institute for Statistics Education (http://www.statistics.com). These experiences helped us realize that the range of methods covered were reachable to biologists seeking a compendium of sampling procedures for ecological and environmental studies, and their corresponding analyses, with only the necessary mathematical derivations for their comprehension. During all those years, the notes were updated, and essentially a draft version of a book was produced. This draft book was used for the Internet course until 2012, but at that stage the decision was made to publish a book not only covering standard sampling methods but also including chapters on the many recent developments in environmental and ecological sampling methods. This then required the involvement of more people in the production of what would become an edited book.

First, Jorge Navarro became a second editor for the book because of his experience in the application of statistics to many biological problems. It was then decided that the book required chapters on some specialized topics written by leading researchers on those topics. The book in its final form therefore covers standard sampling methods and analyses followed by chapters on adaptive sampling methods by Jennifer Brown, line transect sampling by Jorge Navarro and Raúl Diaz-Gamboa; removal and change-in-ratio methods by Lyman McDonald and Bryan Manly; plotless sampling by Jorge Navarro; mark-recapture sampling of closed populations by Jorge Navarro, Bryan Manly, and Roberto Barrientos; mark-recapture sampling on open populations by Bryan Manly, Jorge Navarro, and Trent McDonald; occupancy models by Darryl MacKenzie; sampling designs for environmental modeling by Trent McDonald; and trend analysis by Timothy Robinson and Jennifer Brown.

This book could not have been produced without the assistance of the seven authors who produced all or part of chapters. We are therefore grateful for their help. Finally, the contribution of Jorge Navarro as a coeditor would not have been possible without the support and resources provided by West Incorporated and the University of Wyoming during a sabbatical stay in Laramie, for which he is very grateful.

Some chapters have supplementary material that can be found at the Google site https://sites.google.com/a/west-inc.com/introduction-to-ecological-sampling-supplementary-materials/home.

Bryan Manly
Western EcoSystems Technology Inc.
Laramie, Wyoming, USA

Jorge A. Navarro Alberto
Universidad Autónoma de Yucatán
Mérida, Yucatán, México

About the Editors

Bryan Manly was a professor of statistics at the University of Otago, Dunedin, New Zealand until 2000, after which he moved to the United States to work as a consultant for Western EcoSystems Technology. His interests are in all aspects of statistics applied to biological problems but in recent years particularly in analyses related to organisms in seas and rivers.

Jorge Navarro Alberto is a professor at the Autonomous University of Yucatan, Mexico, where, since 1986, he has taught statistics and sampling design courses for undergraduate biology students and since 1994 has taught graduate marine biology and natural resource management students. His current research involves the development of statistical methods in community ecology, biodiversity conservation, and biogeography.

Contributors

Roberto Carlos Barrientos-Medina
Universidad Autónoma de Yucatán
Departamento de Ecologia Tropical
Campus de Ciencias Biológicas y Agropecuarias
Mérida, México

Jennifer Brown
University of Canterbury
School of Mathematics and Statistics
Christchurch, New Zealand

Raúl Díaz-Gamboa
Universidad Autónoma de Yucatán
Departamento de Recursos Marinos Tropicales
Campus de Ciencias Biológicas y Agropecuarias
Mérida, México

Darryl MacKenzie
Proteus Wildlife Research Consultants
Dunedin, New Zealand

Lyman McDonald
Western EcoSystems Technology Inc.
Laramie, Wyoming, USA

Trent McDonald
Western EcoSystems Technology Inc.
Laramie, Wyoming, USA

Timothy Robinson
University of Wyoming
Department of Statistics
Laramie, Wyoming, USA

1

Introduction

Bryan Manly and Jorge Navarro

1.1 Why a Book on Ecological Sampling and Analysis?

Some biological sciences, like agronomy, physiology, and the like, share concepts and research methods with the science of ecology, but there are obvious differences regarding the way data are gathered. In the former subjects, designed experiments with the sampling processes carried out under controlled conditions are more often performed. Although the use of experiments in ecology has always been suggested in order to gain a better idea of cause-and-effect relationships (Underwood, 1997), this is often difficult or impossible, and ecologists are compelled to use observational methods of sampling. Moreover, there are multidisciplinary areas combining the ecological component and socioeconomic approaches (e.g., ethnobiology, natural resource management) in which the only sampling strategy available is a nonexperimental, observational approach.

Many ideas behind sampling in ecology have their roots in methods provided by a classic discipline in statistics, namely, survey sampling of finite populations. Simple random sampling, stratified sampling, and systematic sampling are all firsthand tools applicable to ecological studies. However, there are often particular problems faced by ecologists for sampling real animal or plant populations that have made both ecologists and statisticians develop special sampling methods that take into account the peculiarities of the situation of interest. There are numerous examples of such methods, including mark-recapture sampling, adaptive sampling, removal sampling, and so on. Our purpose in presenting this book is to cover both the classic approaches and the methods needed by working ecologists.

Although the sampling procedures covered in this book are diverse, they are unified by the widespread interest in ecological studies of estimating biological (e.g., population size and density) and environmental (e.g., the concentration of chemical elements) parameters. Thus, this book emphasizes how particular ecological sampling methods in different settings are combined with estimation procedures that are justified by statistical theory.

1.2 The Scope and Contents of the Book

This book introduces ecological sampling methods and the analysis of the data obtained with the assumption that readers start with basic knowledge of standard statistical methods based on one or two introductory courses but know little about how these methods are applied with ecological data and know nothing about the more specialized methods that have been developed specifically for ecological data. The book is only an introduction, so that use of many of the methods described in the chapters may require reading a text that is more specialized or in some cases even attending a course on the use of a special statistical computer package.

There are ten chapters in the rest of the book. The remainder of this chapter briefly describes what is in these chapters and the reason why each chapter was included in the book.

Chapter 2, “Standard Sampling Methods and Analyses,” covers the traditional methods that have been employed with ecological and other data for more than 50 years, plus some newer developments in this area. These methods all assume that there is a specific population of interest, and that the population consists of many items of interest. For example, the population could be all of the plants of a certain species in a national park, and the interest is in the number of plants per square meter for the entire park. In these types of situations, it is usually the case that obtaining information for the whole population is not possible because this would be far too expensive. Therefore, it is necessary to sample the population and estimate variables of interest using the sample results. With the plant population, this could be done by sampling meter-square plots at random throughout the national park and estimating the density for the whole park using the observed density on the sampled plots, with a measure of how large the difference between the observed density and the true density is likely to be. Chapter 2 discusses estimation using random sampling methods like this and a number of variations on this that are intended to improve the estimation, such as stratified sampling, by which different parts of the population are sampled separately.

Chapter 3, “Adaptive Sampling Methods,” covers the methods that have been proposed for which there is an initial sampling of a population and the results from this are used to decide where to sample after that, with the idea that this should lead to more efficient sampling. Adaptive cluster sampling, which is covered in more detail than other adaptive sampling methods, involves dividing an area of interest into quadrats, taking a random sample of those quadrats, and then taking more samples adjacent to the quadrats where the items of interest are either present or the number exceeds some threshold. Similarly, two-phase adaptive sampling involves taking samples from different geographical parts of a population (strata) and then deciding where to do more sampling based on what is found in the initial sampling. The idea is then that the second round of sampling can be in those parts of the population where more sampling will be most beneficial. These and the other methods considered in Chapter 3 are generally concerned with estimating either the total abundance or density per unit area of animals or plants in a region.

Chapter 4, "Line Transect Sampling," is mainly concerned with the situation when one or more observers travel along a line in a region of interest and record the number of objects of interest seen (again animals or plants) and the distance of these objects from the line. Typically, this may involve following a line on the ground, flying in an aircraft and observing objects on the ground, or moving along a path in a boat over a sea or lake and recording the objects seen. It is generally assumed that the probability of detecting an object depends on its distance from the line, with a low probability of detection for objects far away. The data are then used to estimate a detection function and then estimate the total number of objects within a certain distance from the line and the density of objects per unit area around the line. This method of sampling is usually used when alternative methods such as randomly sampling quadrats in the region of interest are not practical for some reason.

Chapter 5, "Removal and Change-in-Ratio Methods," discusses two methods for estimating the size of animal populations. With the removal method, an animal population is sampled a number of times, and the animals caught are removed or possibly marked so that if they are caught again it will be known that they were seen before so that they are effectively "removed." Assuming a closed population (i.e., a population in which animals are not entering and leaving between samples), the number of animals left in the population will decrease every time a sample is taken, so that the number of animals available for removal will decrease with time and will be zero if enough samples are taken. At that time, the population size will consist of the total number of animals removed in the earlier samples. However, it is not necessary to remove all of the animals to obtain an estimate of the total population size; Chapter 5 describes how the results from several samples can be used to estimate the total number of animals in a population even though some of these animals have not been seen. The change in ratio method is similar but involves a population in which there are two or more recognizable types of animal in a population, such as males and females or juveniles and adults. Then, a sample is taken, and a fixed number of one of the animal types is removed or possibly marked so that a marked animal is considered removed. If a second sample is taken, then it becomes possible to estimate the population sizes at the time of the first sample and at the time of the second sample (which is just the first sample size minus the number of animals removed). Both the removal method and the change in ratio method are closely related to the mark-recapture methods described in Chapter 7.

Chapter 6, "Plotless Sampling," is concerned with methods for estimating the density of objects, such as the trees in a defined area, without dividing the area into plots of a certain size and then randomly sampling the plots or sampling the area by one of the other methods discussed in Chapters 2 to 5. Instead, points are selected in the area of interest either randomly or with a systematic pattern, then the distance to the nearest object is measured for each point and possibly the distance from that object to its nearest or *k*th-nearest neighbor. There are various methods that have been proposed for that type of sampling, and Chapter 6 describes two of these in some detail. The first is T-square sampling. In this case, once a point in the area of interest is selected for sampling, then the distance from this point to the nearest object is measured. A line is then set passing through the position of the object at a right angle to the line from the initial point to the object. The distance from the object to its nearest neighbor on the side of the line away from the initial point is then measured. Based on the two measured distances (from the point to the first object and the first object to its nearest neighbor) for a number of points in the area of interest, it is possible to estimate the density of objects in the whole area, assuming that the location of objects is at least approximately random in the area. The second method described in Chapter 6 is wandering quarter sampling. This starts with a randomly selected point in the region of interest. A direction is then selected, such as west, and the nearest object within the 90° angle from southwest to northwest is found. The distance from this object to the next object in the southwest-to-northwest direction is then measured. This procedure is continued until *n* distances have been measured and it is possible to estimate the density of points in the sampled area based on these distances.

Chapter 7, "Introduction to Mark-Recapture Sampling and Closed-Population Models," covers these methods for closed populations (with no losses or gains of animals during the sampling period). With closed population methods, there is a first sample time when captured animals are marked and released, and then one or more further samples are taken, with captured animals again marked and released. Then, it is possible to estimate the population size subject to certain assumptions about the capture process.

Chapter 8, "Open-Population Mark-Recapture Models," covers situations when the population size can change between sample times because of losses (deaths and emigration) and gains (births and immigration). With open populations, the sampling process is similar to that for closed populations, but there are generally more samples. It becomes possible to estimate the population size at the sample times other than the first and last, the survival rates between sample times except between the last two samples, and the number of new entries to the population between the sample times except between the last two samples. This is again subject to certain assumptions about the capture process and the survival rates and the new entries to the population.

Chapter 8 describes the traditional methods for analyzing data and developments that are more recent. In addition, methods for analyzing data based on the recovery of dead animals are considered.

Chapter 9, "Occupancy Models," covers situations where there are data on the recorded presences and absences of species in different locations but some of the absences might be because the species was present but not detected. This then leads to the idea of having a model for the probability that a species is present at a location based on its characteristics and a model for the probability of detecting a species if it is present, again based on the characteristics of the location. These models then require more than one sample to be taken from each location because if a location is sampled several times and the species is seen at least once, then this provides information on the probability of detecting a species when it is really present. Although site occupancy models were originally just used when there were presences and absences, they have now been extended to situations with more than two possibilities regarding presence, such as absence, presence with breeding, and presence without breeding. Chapter 9 therefore also considers these types of situations and situations where the status of a location may change with time.

Chapter 10, "Sampling Designs Used for Environmental Monitoring," describes the various types of sampling schemes that are being used for studies to track changes in environmental variables at local, national, and international levels. It notes the difference between these types of studies and research studies that are usually carried out over a shorter period of time to examine possible changes related to an issue of concern (e.g., a study to measure the adverse effects of an oil spill). The chapter considers the different types of spatial designs used with environmental monitoring and the possible designs that allow for repeated sampling at individual locations and changes with time of these locations.

Finally, Chapter 11, "Models for Trend Analysis," considers in detail the types of analyses that are appropriate when there are repeated observations over time at a number of sampling sites and there is interest in both changes at the individual sites and changes for the whole geographical area covered by the sites. The chapter considers the use of various simple analyses and graphical methods for exploring trends in the data and two approaches for modeling. One approach involves using linear regression methods for examining the trends at individual sampling sites and then combining these results to examine the overall trend for all sites together. The other approach does one analysis using the data for all of the sites together based on what is called a mixed model, which essentially allows for individual sites to display random differences from an overall trend. This chapter uses a data set on mercury concentrations in fish to illustrate the different methods considered, with the data collected from the yearly sampling of fish from the same 10 randomly chosen locations in a lake for 12 years.