Preface

Martin Wegmann, Benjamin Leutner and Stefan Dech

Dear Reader,

The aim of this book is to make remote sensing data and tools more accessible to ecologists. It provides a starting point for applying remote sensing and geographic information systems (GIS) in ecology using open source software (OSS), hence this book is not aimed at remote sensing specialists unless they are interested in learning how to apply OSS in remote sensing. We cannot cover the full range of GIS and remote sensing methods within this book, but will rather focus on ecologically relevant examples instead.

There are several books around that provide a more in-depth introduction to the theory, methods and history of remote sensing in ecology and conservation, such as:

- Horning, N., Robinson, J. A., Sterling, E. J., Turner, W. and Spector, S. (2010) *Remote Sensing for Ecology and Conservation: A Handbook of Techniques*. Oxford, OUP.
- Pettorelli, N. (2013) *The Normalized Difference Vegetation Index*. Oxford, OUP.
- Skidmore, A. and Prins, H. (2002) *Environmental Modelling with GIS and Remote Sensing*. London, Taylor & Francis.

We can also recommend more general remote sensing books:

- Campbell, J. B. and Wynne, R. H. (2011) *Introduction to Remote Sensing*. New York, Guilford Publications.
- Jensen, J. R. (2015) *Introductory Digital Image Processing: A Remote Sensing Perspective*. Glenview, IL, Pearson Education Inc.
- Lillesand, T., Kiefer, R. W. and Chipman, J. (2015) *Remote Sensing and Image Interpretation*. Hoboken, NJ, Wiley.

This book does not intend to compete with these books. We want to provide practical guidance by offering hands-on examples of ecological remote sensing approaches and OSS solutions.

Our target user community is situated between beginner and advanced user. We expect some prior knowledge about spatial data, remote sensing and the software used. We will guide you through the whole process from acquiring and pre-processing data, statistical analysis and modelling up to interpreting the results. By the end, we hope that you will have learned what remote sensing and GIS, as well as R and Quantum GIS (QGIS), can do for you but also what its limitations are.

In the past, remote sensing has been occasionally oversold. Ecologists were promised a novel tool and had high expectations. But all data sets and tools have limitations and remote sensing per se is not a panacea. Even the most accomplished researcher will be limited in the amount of accuracy and richness of detail they can extract because of noise in the data, physical constraints and the inherent uncertainty in data and methods – similar to ecological field work. We will introduce the concept of *accuracy* within remote sensing, what it means in practice and that it is impossible to achieve 100% accuracy.

We will use the term *remote sensing* to refer exclusively to space borne satellite imagery, not in the broader sense of any measurement without physical contact. Consequently, all data sets used in this book are acquired by sensors on space borne satellites. The term GIS will refer to spatial data handling and analysis software in general.

We will introduce the software, spatial data and especially remote sensing data structure (Chapters 1 and 2) before delving into data acquisition and the first steps in data analysis (Chapters 3 and 4). Data pre-processing is covered in Chapter 5 and is followed by a discussion of approaches in field data collection (Chapter 6). The first remote sensing-derived parameters we will discuss are vegetation indices (Chapter 7) and land cover classifications (Chapter 8). The analysis of environmental change is discussed in Chapter 9. Alternative ways to describe the landscape by using a gradient approach will be shown in Chapter 10. In-depth analysis of time series data is provided in Chapter 11. Spatial metrics are addressed in Chapter 12 and followed by spatial statistical analysis of species distribution (Chapter 13) and animal movement (Chapter 14). This set of topics will give you a good start to spatial data analysis for ecologists. However, many more details are needed to conquer remote sensing and GIS in its totality. Thus, please consider studying some of the literature listed at the end of each chapter (where available).

In this sense, the book is arranged to comply with the normal approach of ecologists:

1. in situ data sampling and adding the spatial component;
2. acquiring data (remote sensing and auxiliary data);
3. pre-processing the data sets;
4. combining field and spatial data sets;
5. deriving relevant spatial environmental information;
6. extracting landscape information relevant for ecologists;
7. carrying out statistical analysis of ecological and remote sensing data.

We will use two different programs, namely QGIS and R, which will be introduced in the following chapters. You might want to discuss approaches as well as technical issues with fellow ecological remote sensing and GIS users. For this purpose, we have provided a mailing list or **listserv** (RS-ecology) for discussions, news or updates. RS-ecology can be found at http://lists.uni-wuerzburg.de/mailman/listinfo/rs-ecology. Please feel free to post any question on this list. The authors are also subscribed and will try to help you as quickly as possible. Any bugs or out-of-date functionality will also be discussed on this list.

Textbook vs Sourcebook

In this book we clearly want to provide you with a working knowledge on remote sensing and GIS handling for ecological research. We have focused on explaining how to approach spatial data analysis in a practical fashion, and hence we have tried to keep references and detailed discussions on specific approaches to a minimum. However, the methods and tools shown here are just some of many appropriate ones. For instance, we will use *Random Forest* for classification, but there are many alternative methods which could be applied. We strongly argue for the need for a sourcebook on standards for ecological spatial data analysis that can provide a baseline to allow comparison between analyses, but this is not that book. Hence, please use this book as a textbook – which is what it is – and not as a sourcebook providing standards.

Why Open Source

The two programs used in this book, QGIS and R, are OSS programs. The term open source (OS) refers to a licensing practice for software. At its core, OS licensing means that you are free to study the underlying source code, modify it and redistribute it under the same conditions. In very few cases constraints for redistribution might exist, so please check the respective web pages of the software or, for a general overview, the http://creativecommons.org/ website. Usually OSS, like the software we are using in this book, can be used free of charge.

This liberal licensing has given rise to highly active communities around the use and development of OSS. This means that there are thousands of developers around the world writing and improving the software and countless people using and testing the software. User feedback is highly appreciated and if you report a bug in the software there is a high chance it will be fixed quickly. If you miss a certain functionality, you can submit a feature proposal or even better, provide your own solution – which is possible because you have access to the source code. Then, in a future release, the core software may well be distributed with your desired feature.

However, it is not only about software development. You will also find a community of people willing to help if you encounter problems. Open source software projects typically also provide mailing lists where you can ask questions, or dedicated websites, such as Stack Overflow (http://stackoverflow.com/), and more specialized ones like R, QGIS user mailing lists or Cross Validated (http://stats.stackexchange.com) where you will find solutions, inspiration and can post your own questions and comments.

Open source is also a necessary step towards truly reproducible research. Being able to read and understand the software allows you to verify that your computer does indeed do what you think it does, and if you share your code it also allows other researchers to repeat your analysis and hence enables scientific progress.

Another typical feature of OSS is that you will find a high degree of interoperability and standardization that facilitates switching from one program to another. Unlike vendors of proprietary software, no one in the OS community has an interest in restricting you to use their specific product or product version. It is also very likely that you will be able to open OS file formats created today in 50 years' time, which can be difficult with proprietary software if a vendor stops supporting an old format and no longer offers its old products. As a side benefit, most OSS is platform independent, meaning you can run it on many different operating systems, be it Linux, FreeBSD, Mac OS X or Windows.

The amount of OSS is increasing steadily, catering to all levels of users from beginners to power users. There are intuitive products with graphical user interfaces (GUIs), like QGIS, or specialized software like GDAL which can be run directly from the command line. A good overview of some geospatial OSS packages can be found at the OSGeo website (http://www.osgeo.org).

Planning

Any approach that endeavours to look at our environment with the aim of understanding it and apply appropriate conservation methods requires thorough planning. The subsequent planning process should proceed along these lines:

- Define your idea, what is needed, and why.
- Come up with one or several hypotheses.
- Design your approach in a way to test each of your hypotheses. Thereafter, you will know which data you have to collect or acquire.
- Make sure you know which methods are needed to achieve interpretable results of your idea.
- Check which data sets are available and if they are not available repeat the process.

In *Statistics for Ecologists Using R and Excel*, Gardener (2011) outlines the planning structure in much more detail for general statistics. Therefore, read through the planning stages described in his book if you are not familiar with this approach.

Software, Scripts and Online Resources

Links, general information about scripts and relevant documents will be updated on the website that accompanies this book: http://book.ecosens.org. Also, we provide a *cheatsheet* for R with packages and commands used throughout this book on the book's website. Additionally, various websites provide helpful overviews of extensions, for example, the CRAN Task View: Analysis of Spatial Data website (http://cran.r-project.org/web/views/Spatial.html).

Please install the following software programs for this book:

- R: http://www.r-project.org/; and
- QGIS: http://www.qgis.org/en/site/.

We use R version 3.1.1 and above and QGIS version 2.6 and above.

Study Area

We wanted to have a study area with clear patterns of land cover change, which shows the effects of a protected area on forest and deforestation patterns. Remote sensing data sets had to be available for this area also, with no or minimum cloud cover. One area that suited our needs was located in Brazil, the Parakanã Indigenous Reserve in the state of Pará, North Region of Brazil (Figure 0.1). This reserve, founded in 1975, is listed as an indigenous area but has not been assigned to an International Union for Conservation of Nature (IUCN) category (Protected Planet, http://www.protectedplanet.net). Large-scale deforestation in the region is due to agricultural expansion and cattle ranching and has modified this region considerably in the past. Mainly, the Parakanã Indigenous Reserve remains forested (Figure 0.2). A reservoir on the Rio Tocantins (Tocantins River) covers the north-east part of the area. This large body of water (*Lago Tucuruí*) is an artificial lake created by the construction of a hydroelectric dam called *Represa de Tucuruí* (Tucuruí Dam) which was finished in 1984. Construction of the dam is closely related to the observed deforestation and degradation patterns, as workers began settling down after the construction of the dam was complete. Moreover the area is bisected by the Transamazônica highway.

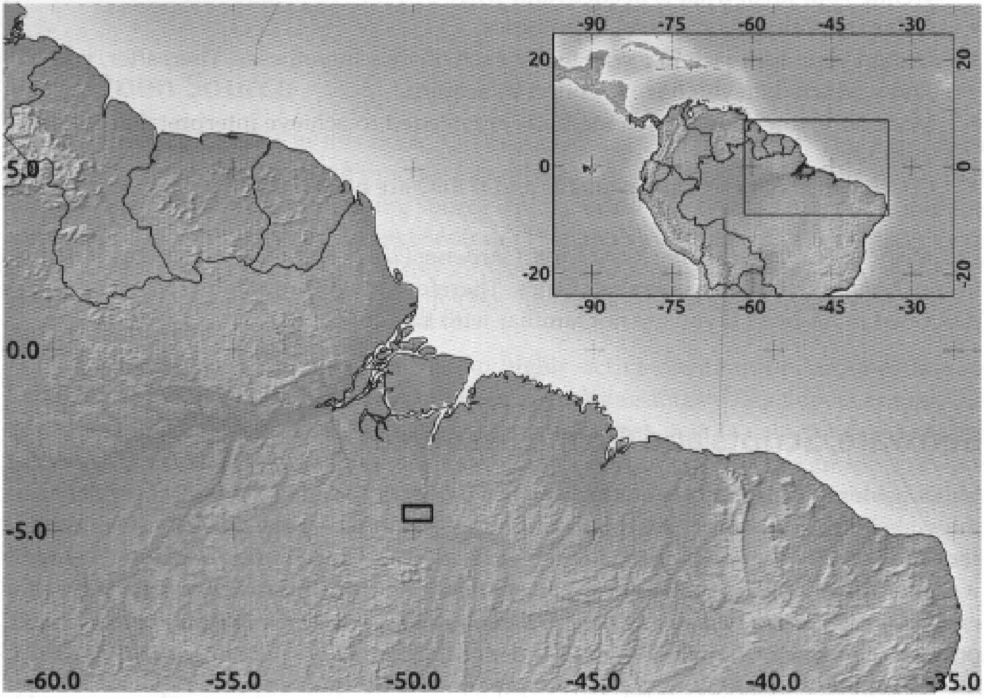

Figure 0.1 The study area in the state of Pará, North Region of Brazil.

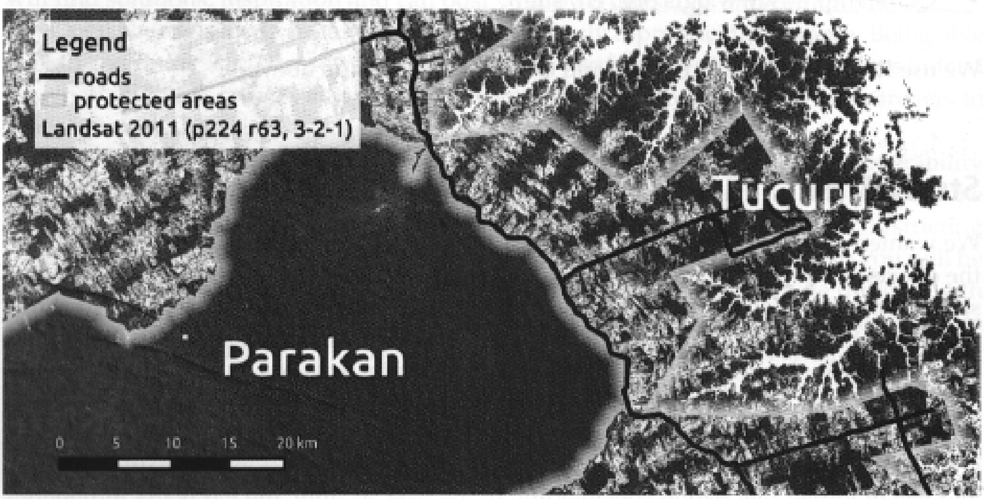

Figure 0.2 Landsat coverage (2011) of the study area showing the protected areas and roads. (The names of the protected area are from the Protected Planet website.)

Graphics in this Book

Note that in this book, we used the `ggplot2` package for creating the graphics which are displayed. However, `ggplot2` comes with its own syntax which can easily fill a book on its own (Wickham, 2009: ggplot2 Elegant Graphics for Data Analysis, Springer). To avoid adding this additional layer of complexity, we have opted to show only the code needed to create graphics with the basic plotting system. Hence your maps will look slightly different than those shown here. A short outline on how to create your maps with `ggplot2` is given in Chapter 4.

Get Connected

Many different communities use remote sensing for conservation, biodiversity research and many other purposes. Several activities are already ongoing to bring remote sensing, ecology and conservation communities together. GEO BON (Biodiversity Observation Network, http://geobon.org) is aimed at biodiversity academics, the Remote Sensing Conservation network (CRSnet, http://remote-sensing-biodiversity.org/networks/crsnet) at conservationists and the Group on Remote Sensing for Biodiversity and Conservation (CEOS Biodiversity) is targeting the remote sensing or Earth Observation community (http://remote-sensing-biodiversity.org/networks/ceos-biodiversity). These initiatives try to improve the use of remote sensing when applied to biodiversity and conservation. You are very welcome to join these initiatives and to improve the use of remote sensing and GIS for ecological applications. Moreover, several courses were born out of these activities, such as Animal Movement and Remote Sensing for Conservation (AniMove, http://animove.org) or EcoSens: Remote Sensing and GIS in Ecology & Conservation (http://ecosens.org), which provides GIS and remote sensing training for ecological applications. These provide intense professional training as well as networking opportunities.

We highly recommend using this book for what it is: as a starting point to learn about the advantages of applying remote sensing and GIS in ecology, and to develop your interest in collaborating with colleagues from other communities.

Reference

Gardner, M. (2011) *Statistics for Ecologists Using R and Excel.* Exeter, Pelagic Publishing.