Sarah E. Gergel
Monica G. Turner
Editors

Learning Landscape Ecology

A Practical Guide to Concepts and Techniques

Second Edition

Springer

Learning Landscape Ecology

Sarah E. Gergel • Monica G. Turner
Editors

Learning Landscape Ecology

A Practical Guide to Concepts and Techniques

Second Edition

Springer

Editors
Sarah E. Gergel
University of British Columbia
Department of Forest and Conservation Sciences
Vancouver, BC, Canada

Monica G. Turner
University of Wisconsin-Madison
Department of Zoology
Madison, WI, USA

ISBN 978-1-4939-6372-0 ISBN 978-1-4939-6374-4 (eBook)
DOI 10.1007/978-1-4939-6374-4

Library of Congress Control Number: 2017933462

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Science+Business Media LLC
The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.

Preface

Landscape ecology continues to grow as an exciting discipline with much to offer for solving pressing and emerging problems in environmental science. Much of the strength of landscape ecology lies in its ability to address challenges over large areas, over spatial and temporal scales at which decision-making often occurs. As the world tackles issues related to sustainability and global change, the need for this broad perspective has only increased. Furthermore, spatial data and spatial analysis (core methods in landscape ecology) are critical for analyzing land-cover changes worldwide. While spatial dynamics have long been fundamental to terrestrial conservation strategies, land management, and reserve design, mapping and spatial themes are increasingly recognized as important for ecosystem management in aquatic, coastal, and marine systems. For these reasons, there is great demand for training in spatial analysis tools accessible to a wide audience.

The first edition of this book, *Learning Landscape Ecology: A Practical Guide to Concepts and Techniques*, was the first "hands-on" teaching guide for landscape ecology. The book introduced a diversity of tools and software in the field. The text sold over 5000 copies worldwide, was used at more than 55 universities, and had its second printing in 2006. However, landscape ecology has grown and quantitative methods have advanced substantially in the ensuing 15 years. In addition, this revised second edition of *Learning Landscape Ecology* complements the release of the second edition of *Landscape Ecology in Theory and Practice* (Turner and Gardner 2015), which pairs nicely with this updated "hands-on" teaching guide.

This second edition of *Learning Landscape Ecology* is purposefully more applied and international in its examples, approaches, perspectives, and contributors. It includes new advances in quantifying landscape structure and connectivity (such as graph theory), as well as labs that incorporate the latest scientific understanding of ecosystem services, resilience, social-ecological landscapes, and even seascapes. Of course, as before, the exercises emphasize easy-to-use, widely available software. We have also included introductory exposure to spatial analyses using R programming language in several labs.

What remains similar to the first edition is our dedication to making seemingly complex ideas easy to understand and use for scientists from diverse intellectual backgrounds and particularly for those early in their careers.

Sarah E. Gergel
Vancouver, BC, Canada
Monica G. Turner
Madison, WI, USA

Advice for Instructors

All necessary files, data, software, as well as any corrections and updates can be downloaded from the book website: http://landscape.forestry.ubc.ca

For a complete copy of the Instructor's Notes (with Answer Key), please email Sarah E. Gergel (SEG): sarah.gergel@ubc.ca and supply your course number, title, and affiliation.

The first edition is accessible via the Springer website. Teaching materials originally supplied on the CD-ROM from the first edition can be obtained directly from SEG.

Audience and Difficulty Levels

The content in the second edition spans a range of difficulty levels. These difficulty levels were assessed based on a combination of factors: the complexity of content, any assumed prior knowledge and technical expertise of students, and the amount of time likely needed to complete a lab. All of these factors also influence the amount of in-class supervision and oversight a lab likely requires, as well as instructor "prep" time prior to class. Each chapter is built around the idea of occupying a 2–3 hour lab period, with various exceptions noted below. We've also made suggestions for portions of labs suitable for "take-home" assignments, recognizing that in-class laboratory time at the computer can often be limited.

While there are far too many chapters for use in any one course, the variety of courses that could benefit from, or be built around, these labs include:

- Landscape Ecology
- Watershed Management and Monitoring
- Wildlife Conservation
- Forest Disturbance

- Conservation Planning
- Landscape Modeling and Spatial Analysis
- Landscape Sustainability and Resilience
- Spatial Statistics in Ecology
- Biogeography, Spatial Ecology or Macroecology
- Marine Spatial Planning
- Social-Ecological Systems and Ecosystem Services

As with the first edition, we have extensively beta-tested all of these labs and asked all contributors to create detailed Instructor's Notes (available by email directly from SEG). However, various glitches are always a possibility—thus we strongly suggest instructors spend the necessary time in advance of teaching to "click" their way through the material prior to use in the classroom as well as check the website for the latest updates.

To assist in understanding the suggested audience for each lab, we have grouped chapters into four categories based on their level of difficulty. These suggestions are meant to assist instructors with course planning and time allotment. We also noted any chapter prerequisites as well as suggested corresponding readings from the companion text, *Landscape Ecology in Theory and Practice* (*LETP*).

Quick and Fun

These shorter and/or low-tech, technically simpler chapters are especially suited for a shorter class period (perhaps 1–2 hours) or a course without a dedicated computer lab. These may even be suitable to use as a take-home assignment. Generally, little to no computer skills are required with the exception of entry-level familiarity with programs likely available on the laptops of all students (such as Excel, Google Earth, or a web browser). Very little prior knowledge of landscape ecology is assumed. Students may also appreciate these more straightforward labs as a break from the otherwise very challenging chapters in the rest of the book!

Undergraduate

These labs are suitable for upper-level undergraduate students from a wide variety of environmental sciences (e.g., Geography, Ecology, Forestry, Zoology, and Botany). These labs might also be very reasonable choices for a graduate course (e.g., a course-based MSc program) depending on the background of the students. These labs might be a good way to begin a graduate course and ensure students from different disciplinary backgrounds are all "on the same page."

Graduate

These lab topics and tools include those used in research and applied conservation situations. These exercises assume a higher level background in environmental and/or ecological science as well as knowledge of basic statistics. There are options for using R software (if desired) in addition to options utilizing freeware/shareware with provided data. Instructors are wise to expect some software installation and plan for basic troubleshooting prior to teaching as versions of operating systems and permissions may change.

Advanced

These labs primarily explore research-oriented tools and fit well with a PhD-level pedagogical approach in terms of the levels of independence and critical thinking required. These exercises assume substantive prior knowledge of students *as well as instructors*, including one or more of the following: statistics including multivariate statistics, basic working ability in R, and/or comfort with GIS and geomatics tools (such as Arc). Most also have one or more lab prerequisites (completion of other chapters in this book). For these labs, instructors should also be well prepared in advance and anticipate troubleshooting for the particulars of their computer lab setup. These labs can easily occupy two full weeks of a regular (2–3 hour) lab period.

Chapter	Title	Contributors	Suggested audience				Software requirements	*LLE* Chapter pre-requisite	Chapters in *LETP*
			Quick and fun	U	G	A			
1	Introduction to Remote Sensing	Coops and Tooke		X			ArcGIS, Google Earth		1
2	Historical Aerial Photography for Landscape Analysis	Morgan et al.	Part 1	X					1, 2, parts of 4 and 6
3	Citizen Science for Assessing Landscape Change	Cardille and Jackson	X	X			Google Earth, Google Docs, browser with Internet connection	2 is a nice complement	
4	Understanding Landscape Metrics	Cardille and Turner		X	X		Fragstats		4
5	Scale Detection Using Semivariograms and Autocorrelograms	Palmer and McGlinn			X		Excel (R is optional)		5
6	Characterizing Categorical Map Patterns Using Neutral Landscape Models	Gardner			X		QRULE (R is optional)		3
7	What Constitutes a Significant Difference in Landscape Pattern?	Remmel and Fortin				X	R	5 or 6 with R option is helpful	Part of 3, as well as 4, 5
8	Introduction to Markov Models	Urban and Wallin		X	X		Markov.exe executable (R is optional)		2, 3
9	Simulating Management Actions and Their Effects on Forest Landscape Pattern	Gustafson		X			Harvest Lite		3
10	Regional and Continental-Scale Perspectives on Landscape Pattern	Cardille and Turner			X		Google Docs, browser with Internet connection	6	4, 6

11	Using Spatial Statistics and Landscape Metrics to Compare Disturbance Mosaics	Turner and Simard			X	X	Excel and GS+	5	5, 6
12	Assessing Multi-Scale Landscape Connectivity Using Network Analysis	Lookingbill and Minor	X						4, 7
13	Systematic Conservation Planning with Marxan	Watts et al.		X			Marxan		7
14	Connectivity as the Amount of Reachable Habitat: Conservation Priorities and the Roles of Habitat Patches in Landscape Networks	Saura and de la Fuente				X	Conefor (optional exercise with ArcGIS)	12	Parts of 4, 7, 9
15	Linking Landscapes and Metacommunities	Bennett and Gilbert (each contributed equally)				X	R	Either prior R labs (5, 7)	5, 7
16	Modeling Spatial Dynamics of Ecosystem Processes and Services	Gergel and Reed	Part 1–2	Parts 3–4, Ex4	Part 4, Ex5		Excel		3, 8, part of 9
17	Heterogeneity in Ecosystem Services: Multi-Scale Carbon Management in Tropical Forest Landscapes	Kirby et al.	Part 1	Part 2–3			Excel	8 not required, but pairs nicely	8, 9
18	Regime Shifts and Spatial Resilience in a Coral Reef Seascape	Selgrath et al.			X		Excel		6
19	Understanding Land-Use Feedbacks and Ecosystem Service Trade-Offs in Agriculture	Schulte and Tyndall			X		Google Chrome or Mozilla Firefox with Internet connection		3, 8, part of 9
20	Social Networks: Uncovering Social-Ecological (Mis)matches in Heterogeneous Marine Landscapes	Bodin and Crona				X	NetDraw trial version of Ucinet	12	Part of 4, 10

Acknowledgments

It takes a village, and a very patient village, to produce a book such as this. At the risk of forgetting someone, we'd like to express our deepest gratitude to the many people who helped make this book possible.

For starters, SEG thanks Garry Peterson from the Stockholm Resilience Centre (SRC) for hosting her during her sabbatical when this book really began. Several primordial chapters were used as part of an SRC short course for PhD students from Canada and Sweden. Students in that short course helped test, shape, improve, and in some cases help create these materials. Big thanks to Juan-Carlos Rocha, Johanna Måård, Marat Murzabekov, Elin Jantze, Kate Kirby, Fernando Jaramillo, Angelina Sanderson, Jenny Selgrath, Matilda Thyresson, Rory Tooke, Magnus Tuvendal, and Tiina Vinter.

Feedback from very dedicated external reviewers also improved every chapter. Thank you to Raphael Chavardes, Karl Cottenie, Graeme Cumming, Jeff Cymerman, Amelie Davis, Paul Galpern, Robert Gardner, Corey Garza, Luis Gilarranz, Ellen Hines, Falk Huettman, Michelle Jackson, Richard Lathrop, Scott Mitchell, Trisalyn Nelson, Jennifer Powers, John Rogan, Bob Smith, Marcela Suarez-Rubio, Eric Treml, Ting Ting Zhao, and Benjamin Zuckerberg.

SEG is also greatly indebted to the many graduate (and a few brave undergraduate) students who have taken Landscape Ecology and Management (FRST504) at UBC since preparation for the second edition began. They deserve enormous thanks for their refreshing and untiring enthusiasm toward trial runs of material at different levels of polish. You remind me how much I love teaching and make me better at it. Thanks to the Class of 2012: Tom Bruderer, Raphael Chavardes, Arnaud DeGrave, Greg Greene, James Hehre, Julian Helfenstein, Todd James, Hannah Kent, Sarah Klain, Elly Knight, Andreas Magnusson, Simon Nadeau, Paul Pickell, Richard Schuster, and Stephanie Tomscha; and the Class of 2014: Wesley Brookes, Kirsten Dales, Ian Eddy, Crystal Ell, Tanya Gallagher, Marc-Antoine Leclerc, Bernardo Ranieri, and Xiong Zhang and the and the classes of 2016: Tamara Bonnemaison, Martina Clausen, Aria Guo, Karly Harker, David Williams, Jiali (Emily) Jin, Tanvi, as well as participants in the Biodiversity Centre Spatial Data Workshop. Kevin Yang, Ian Eddy, and Jenny Huang's contributions to copyediting were also greatly appreciated.

Much thanks goes to Brad Cardinale for hosting SEG as part of his Restoration Ecology course at the University of Michigan which helped beta-test the aerial photography materials. Teaching at the Fujian Agriculture and Forestry University (FAFU) in China also helped clarify the wording and the messages in several lab exercises for an international audience.

Monica G. Turner (MGT) especially thanks all students who were members of her Advanced Landscape Ecology course at the University of Wisconsin-Madison during the spring 2012 and 2014 semesters. Many of the exercises in the book were tested on these students, and their written feedback led to substantial improvements in the prose and in the structure of several chapters. We offer heartfelt appreciation to the following young landscape ecologists in the Class of 2012: Patricia Alexandre, Amy Alstad, Dunbar Carpenter, Sarah Carter, Jimmy Doyl, Jessica Gorzo, Brian Harvey, Trip Hook, Chris Johnson, Rachel Mallinger, Catalina Munteanu, Kristi Nixo, Jessica Prices, Jiangxiao Qiu, Carlos Ramirez, Naparat Suttidate, and Ryan Sword. Class of 2014: Luana Centoducatte, Etienne Fluet-Choiuinard, Cat Frock, Nick Galleguillos Katz, Rose Graves, Diana Guzmancolon, Winslow Hansen, Jeremy Hemberger, Chris Latimer, Jesse Miller, Ali Paulson, Ricardo Rivera, Gail Rosen, Paul Schilke, Amy Uhrin, and Kaitlin Stack-Whitney.

We'd both also like to give a huge thanks to the many beta-testers who have run many of these labs with their own classes and provided detailed feedback. Several beta-testers over the years have been particularly helpful in finding, reporting, and even fixing glitches prior to publication so that these labs function as smoothly as possible. Thank you Russell Hedberg, Jennifer Powers, Jeanine Rhemtulla, Rebecca Rowe, and Erica Smithwick.

SEG would like to give long overdue thanks to the late Professor Larry Harris for introducing her to the field of landscape ecology as an undergraduate student at the University of Florida. He was likely very unaware of how well it would "stick." I thank you for helping me find a career I love, and your humor, sarcasm, and thick-skinned scientific bravery still inspire me greatly.

We also want to thank the many of you—instructors and students of landscape ecology—for the gratitude, words of encouragement, and patience you've shown for both editions. Your words helped us maintain enthusiasm for the long process of devising this almost completely new second edition. The support and patience of Janet Slobodien from Springer has also been of fantastic help. Last but not least, SEG also thanks poet Mark Hoadley for his encouragement especially during the very final stages of this book.

Financial Support

Support for SEG as well as for geospatial data adapted for use in these labs was graciously provided by the University of British Columbia's Teaching and Learning Enhancement Fund (TLEF10/11-Global Stewards), National Science and Engineering Research Council—Discovery Grants Program (NSERC RGPIN-2014-05012), British Columbia Ministry of Forests and Range (RES07RCO), National Oceanic and Atmospheric Administration—National Marine Fisheries Service (NOAA-NFFP7100-8-41691), as well as the Stockholm Resilience Centre. Support for MGT came from a variety of funding sources over recent years, and she acknowledges funding from the National Science Foundation, including the Long-term Ecological Research Program (Coweeta LTER: DEB-0823293 and North Temperate Lakes LTER: DEB-0822700) and Water Sustainability and Climate Program (DEB-1038759); the Joint Fire Science Program (Grants 09-1-06-3, 09-3-01-47, 11-1-1-7); the Andrew W. Mellon Foundation; and the Department of Zoology, Graduate School, and Vilas Trust of the University of Wisconsin-Madison.

Contents

Part I
What Is a Landscape? Basic Concepts and Tools

This first module explores three distinct ways to identify, represent, and quantify a landscape. A landscape is an area that is heterogeneous in at least one aspect of interest. The concept of a landscape can include other ideas, an area that is very large in extent, or the inclusion of multiple different ecosystem types. In practice, however, landscape data are often derived from various geospatial data sources; thus, understanding the benefits, assumptions, and limitations of these diverse sources is fundamental to correct analysis of a landscape and interpretation of much of the published literature. The exercises in this module have no prerequisites, and by design, are meant to introduce the basic concepts of remote sensing to an audience with very little (or no) technical background in these topics. Chapter 1 explores the basic components of satellite imagery and how the sun's energy (the electromagnetic spectrum) can be converted into a representation of the Earth's surface (aka "a landscape"). Chapter 2 explores the special role of aerial photography—which has been in use since well before the advent of satellite imagery—in assessing long-term landscape change. Lastly, Chapter 3 introduces one of the newest and rapidly evolving ways to collect landscape-level data using crowd-sourced approaches that are amenable to citizen science. Depending on your background in geospatial technologies, we hope that this introduction to fundamental concepts helps you understand ways that maps are created and used to represent landscape information.

Chapter 1
Introduction to Remote Sensing

Nicholas C. Coops and Thoreau Rory Tooke

OBJECTIVES

Remote sensing is the science of gathering spatial information about the Earth's surface (as well as the oceans and atmosphere) from a distance, using either handheld, aircraft, or satellite sensors. Such data are routinely used in landscape ecology to map, monitor, and manage landscapes. It is important to understand and fully appreciate the different types of electromagnetic radiation used to create geodata derived from remote sensing systems, the spectral and spatial properties of natural and manufactured materials, as well as the characteristics of airborne and satellite sensor systems. Understanding these fundamental aspects of remote sensing will assist landscape ecologists in understanding and distinguishing the diversity and heterogeneity of land cover types in their study regions and better assess how landscapes might have changed over time. This chapter will enable students to:

1. Understand, explain, and quantify aspects of the electromagnetic spectrum (EMS) and how it can be used to describe different land cover;
2. Explore the four basic resolutions of remote sensing imagery and consider how they impact the choice of imagery for specific applications;
3. Learn how to display and conduct basic imagery analysis using GIS software; and
4. Calculate and understand the role of vegetation indices in landscape monitoring.

This chapter is targeted to upper-level undergraduate students with little to no exposure to remote sensing. Some knowledge of GIS and/or Google Earth is helpful for a few exercises. In Exercises 1–3, students will not need a computer and will explore basic remote sensing concepts via pen-and-paper exercises. For Exercise 4, students will need access to a computer with **ArcGIS** (version 9 or higher) to

N.C. Coops (✉) • T.R. Tooke
University of British Columbia, Vancouver, BC, Canada
e-mail: nicholas.coops@ubc.ca

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_1

explore digital remote sensing imagery and how it can reveal different land cover features. In Advanced Exercise 5 (which also requires **Google Earth**), students will learn how to prepare quantitative representations of vegetation using **NDVI**, a vegetation index often used in landscape monitoring. Synthesis Exercise 6 promotes additional exploration of the remote sensing literature. Prior to starting this lab, students are encouraged to read Chapters 3 and 4 of *Remote Sensing for GIS Managers* by Aronoff (2005). If in-class time is limited, consider completing Exercise 1 prior to arrival in class.

INTRODUCTION

Remote sensing has become an essential tool in many fields such as ecology, geography, geomatics, and resource monitoring as images captured from the air provide important and often unique information on the spatial patterns on the Earth's surface (Colwell 1960). As early as the 1910s, researchers were using remote sensing in forestry to better understand forest extent and condition. The advent of aerial photographic cameras in the 1920s resulted in the development of campaigns by many countries to acquire imagery to survey agricultural lands and to map land cover and land cover change. Satellite-based remote sensing began in the late 1950s and early 1960s, and since then over 100 satellite-based sensors have been launched as part of national and international remote sensing programs.

The sensors used in remote sensing can be categorized as either active or passive. **Passive, or optical, remote sensing systems** rely on energy and illumination from the sun and utilize sensors which are sensitive to radiation reflected from the 400–2500 nm region of the electromagnetic spectrum. This range includes the visible, near-infrared, shortwave, and mid- and long-infrared regions of the spectrum (Figure 1.1).

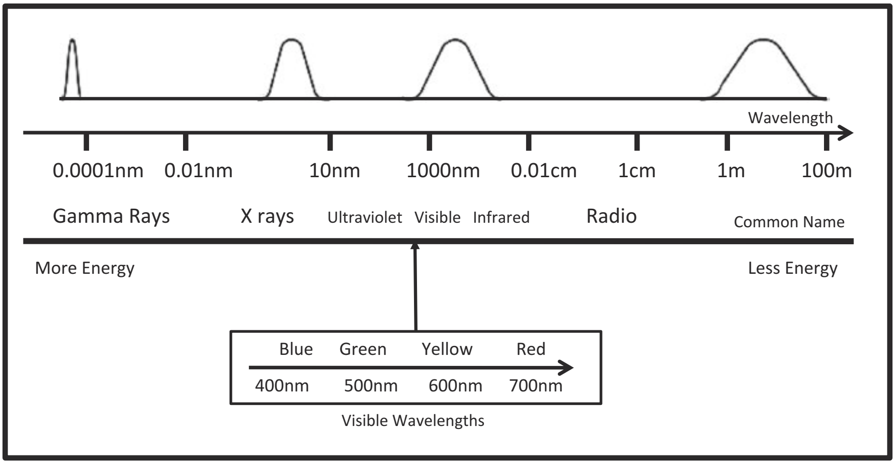


Figure 1.1 Schematic representation of the electromagnetic spectrum (EMS)

Active remote sensing systems are different from passive systems in that energy is emitted from the sensor and either the return time or amount of energy reflected back is measured by the sensor. Having their own power source, such instruments can operate day or night and often under a range of weather conditions. For example, microwave systems can obtain data through cloud cover. The most common active remote sensing sensors are RADAR, which sends and detects microwave wavelengths between 1 mm and 1 m, and **light detection and ranging (LIDAR)**, a more recent technology which most commonly sends and receives near-infrared laser pulses.

Some Remote Sensing Basics

When considering remote sensing imagery for a given application, one must carefully consider image resolutions. These resolutions refer to the four key characteristics of the imagery, including size of individual pixels, overall spatial extent, time interval between acquisitions, and lastly, the region(s) of the electromagnetic spectrum in which the sensor can acquire data and the level of detail (or discrimination) the sensor can provide. Each of these image resolution characteristics is briefly addressed below.

Spatial Resolution

The spatial resolution of a remotely sensed image is the size of the minimum area that can be resolved by the sensor (Strahler et al. 1986) and is generally equated to the pixel size or the "grain" of an image. Depending on the sensor, the spatial resolution can range from submeter to hundreds of meters. For aerial photography, spatial resolution is based on the film speed and the size of the silver halide crystals (Nelson et al. 2001). In the case of digital sensors, an image with a spatial resolution of 30 m resolves a 30×30 m area into a single reflectance response. For satellite sensors, spatial resolution is set at the design phase of the spacecraft, whereas for airborne data, the spatial resolution is governed by the height of the aircraft above the ground.

The spatial resolution provides an indication of what type of detail can be observed on an image. High (or fine) spatial resolution imagery (<5 m) can provide information on small objects such as individual trees, buildings, and cars, whereas low or coarse spatial resolution (>100 m) is more appropriate for observing broad-scale phenomena such as ocean color, broad vegetation phenological responses, and cloud patterns. Historically, medium spatial resolution sensors (10–100 m) (such as **Landsat Thematic Mapper (TM)** and Système Probatoire d'Observation de la Terre (SPOT) multispectral imagery) have provided the optimal resolution for characterizing land cover change and regional disturbance (Franklin and Wulder 2002).

Table 1.1 A sample of spatial resolutions, scene sizes, and potential applications

Spatial resolution	Multispectral sensor	Spectral resolution	Spatial resolution (m)	Spatial extent (km)	Potential applications
Broad	MODIS (MODerate Imaging Spectroradiometer)	405 nm–14.385 µm	250–1000	2330	Ocean color Cloud characteristics Vegetation productivity Phenology
	SPOT (Système Probatoire d'Observation de la Terre) VEGETATION	430–1750 nm	1000	1200	
Moderate	Landsat Thematic Mapper (TM)/ Enhanced Thematic Mapper+ (ETM+)	450–2350 nm	30	185	Land cover Vegetation characteristics Coastal and water
	SPOT	480–1750 nm	5–20	60	
	Indian Resources Satellite (IRS)	520–1700 nm	23–70	142	
Fine	IKONOS	445–853 nm	4	11	Infrastructure mapping Individual tree delineation Disaster monitoring
	QuickBird	450–890 nm	2.4	22	
	WorldView-2	400–900 nm	1.85	16	

The spatial resolution of a sensor is linked to the swath width or instantaneous field of view of the sensor that ultimately determines the spatial extent of the captured image (Lillesand and Kiefer 2000) (Table 1.1). Sensors with a coarse spatial resolution can acquire data over much larger areas, when compared to sensors with very high spatial resolution. As an example, a Landsat TM scene has an image extent of 185 × 185 km (at 30 m spatial resolution), while the MODIS sensor with spatial resolutions of 250 m and larger has an image extent of several thousand kilometers. The spatial extent of data sources must also be considered along with data costs. Coarse spatial resolution data typically cover larger spatial extents and are therefore less expensive per unit area than high spatial resolution data sources. Some medium-resolution imagery (i.e., MODIS, SPOT Vegetation, and Landsat imagery) is freely available, whereas very high spatial resolution satellite systems are often run by private companies resulting in high per unit costs. Increasing spatial resolution presents challenges as image files tend to have large data storage requirements and longer computation processing times. Furthermore, the increased spectral variability within an image with high spatial resolution imagery can confound many commonly used image classification methods such as when individual tree shadows are recorded (Wulder et al. 2004).

Temporal Resolution

The temporal resolution indicates the time required for a sensor to return to the same location on the Earth's surface (i.e., revisit). In the case of satellite systems, temporal resolution is a function of the orbit, image extent, and the capacity of the sensor to tilt and obtain images at requested sites. For Landsat Thematic Mapper imagery, the temporal resolution is 16 days, whereas MODIS with larger extent scene and pixel size has a 1-day revisit. Satellites such as IKONOS and QuickBird use sensors that can point in different directions with short revisit times (varying from 1 to 3.5 days). Such images, however, will be acquired at an angle (known as "off-nadir"). The temporal resolution of airborne sensors is often less critical as image collection via planes is often "on demand" (e.g., coincident with insect outbreaks or fires) (Stone et al. 2001).

Spectral Resolution

Spectral resolution can be considered in three components: the number, width, and location of the spectral wavelength bands detected by the sensor. Some sensors acquire images similar to black and white photography using a single band (or channel) which captures the full range of the visible spectrum (and a small component of the infrared). Known as **panchromatic**, such images are often very useful as they can provide clear and precise spatial information. Detectors with multiple bands (i.e., **multispectral**) have separate spectral bands in the visible (such as blue, green, or red), near-infrared, and mid-infrared regions of the spectrum. As the number of bands increase, **bandwidth** (the range of wavelengths a band detects) often decreases. Sensors with hundreds of narrow spectral bands are known as **hyperspectral**. Currently, most operational remote sensing systems have a small number of broad spectral channels. For example, Landsat Enhanced Thematic Mapper+ (ETM+) has 7 spectral bands in the visible to infrared portions of the spectrum, whereas MODIS has 32 spectral bands. Use of hyperspectral data is increasing because it allows greater discrimination of attributes, such as tree species.

A simplified explanation of spectral resolution for three different sensors is shown in Figure 1.2. The first four bands of Landsat ETM+, SPOT XS, and QuickBird are depicted. Each spectral band covers a slightly different spectral range (or width of band), and the number of bands and the regions of the spectrum detected can vary among sensors.

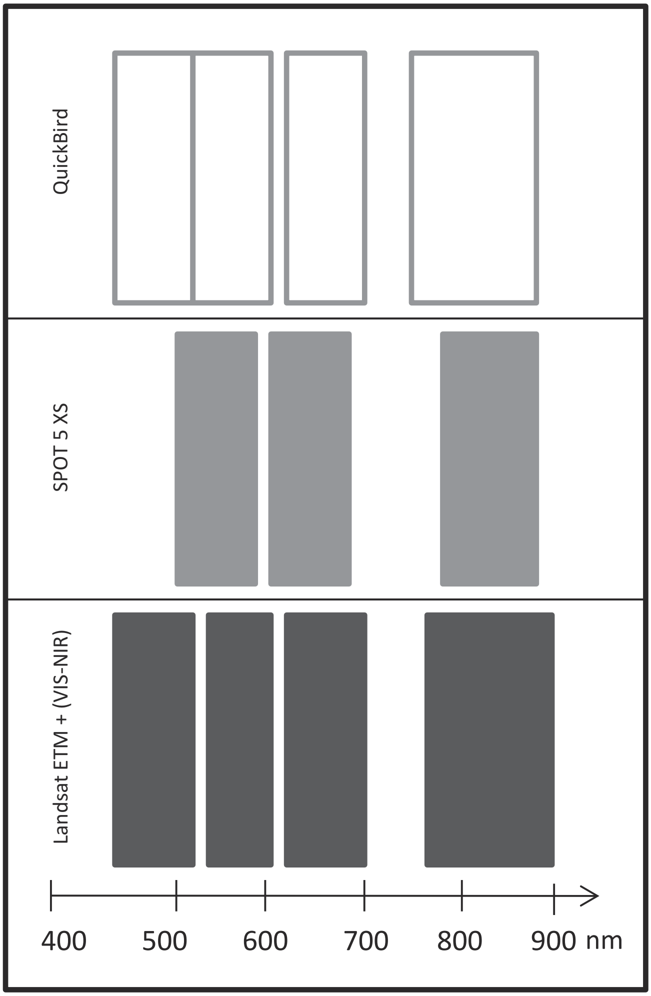


Figure 1.2 The concept of spectral resolution for a subset of bands in three commonly used sensors

Radiometric Resolution

Radiometric resolution provides an indication of the information content of an image. It is often interpreted as the number of intensity (or gray) levels that a sensor uses to quantify the detected reflectance. Generally, the finer the radiometric resolution, the greater sensitivity to detecting small differences in reflectance.

Figure 1.3 provides a simple example of three different radiometric resolutions. In the case of 1 bit, the result would be a binary image of simply pure black and white pixels. 2 bit provides 4 gray levels, whereas 8 bit provides 256 different levels of gray. Most broad- and medium-resolution sensors are 8-bit radiometric resolution. High spatial resolution data such as QuickBird can be up to 11 bit.

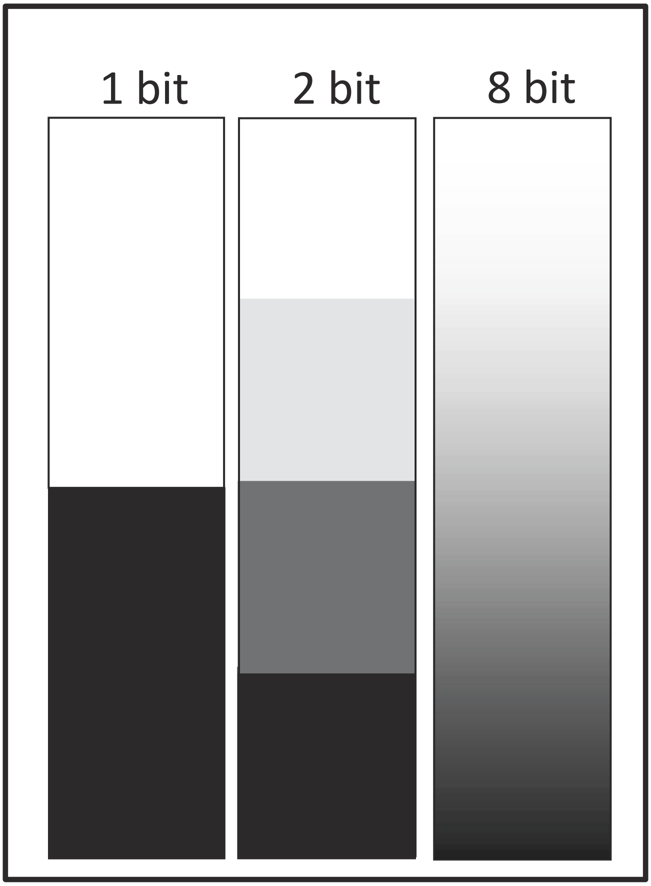


Figure 1.3 The concept of radiometric resolution

EXERCISES

EXERCISE 1: Understanding Spectral Reponses

As discussed earlier, most remote sensing systems measure the amount of reflected radiation from an object within a range of wavelengths of the electromagnetic spectrum. Objects that humans see as bright white, such as clouds or snow, have very high reflectance across all parts of the visible spectrum. For vegetation, leaves reflect more green light than blue or red. Human eyes are unable to see the near-infrared (and many other) regions of the electromagnetic spectrum; however, spectral reflectance curves can be used to show the pattern of reflectance for objects using parts of the spectrum invisible to the naked eye. Figure 1.4 shows a simple spectral curve for six different land cover classes. The *X*-axis ranges from 400 to 1400 nm wavelengths, spanning the visible to the near-infrared region. The *Y*-axis

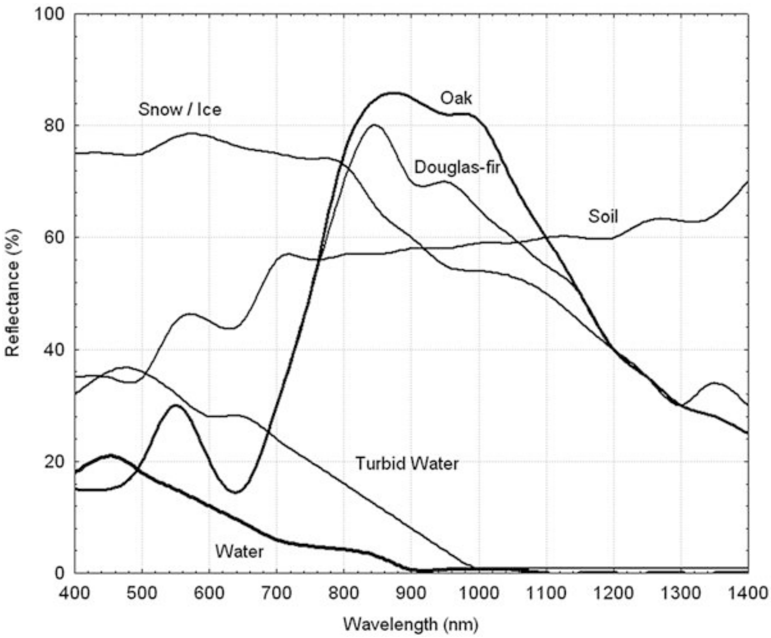


Figure 1.4 Spectral curves for several land cover classes

indicates reflectance where 100% is when all light is reflected in a given wavelength from an object and 0% is total absorption of light by the object.

Q1 What wavelength range has the highest reflectance for snow and ice?

Q2 Identify two wavelength ranges where soil is differentiable from vegetation.

Q3 What is the reflectivity of water at 1200 nm? What are the implications of this for mapping vegetation and water?

Q4 Which wavelengths show the greatest differentiation in reflectance for the two vegetation types?

Q5 The visible part of the spectrum is from 400 to 700 nm. Why is there no separation between the two vegetation types across these wavelengths?

Q6 What wavelength region exhibits the largest change in reflectance for vegetation? Why is this potentially important when looking at different vegetation characteristics?

Q7 What land cover types look almost identical at 700 nm? What about at 1100 nm? What challenges would this pose if you are trying to use these spectral bands to map these land cover types?

Table 1.2 Reflectance characteristics of three unknown cover classes

nm	Class 1 (%)	Class 2 (%)	Class 3 (%)
400	17	10	80
450	18	12	80.5
500	15	18	81
550	12	32	81.5
600	5	21	82
650	0	25	82
700	0	24	82
750	0	40	82
800	0	70	82
850	0	85	82.5
900	0	85	83
950	0	78	82.5
1000	0	75	82
1050	0	65	81.5
1100	0	55	81
1150	0	50	81
1200	0	38	81
1250	0	31	81
1300	0	29	80
1350	0	27.5	80
1400	0	25	80

EXERCISE 2: Detecting the Unknown

Now that you have an understanding of how features and land classes reflect/absorb electromagnetic radiation across the visible and near-infrared regions of the spectrum, in this exercise you will use this background to try and identify some unknown land cover types using only their spectral signatures.

Q8 Use the wavelength and percent reflectance data in Table 1.2 to plot the reflectance curves by hand onto the provided plot on your handout (Figure 1.4).

Q9 Based on the known reflective properties of various classes and what you have learned about spectral signatures, which classes refer to cedar forest, deep water, and cloud? Explain how you determined this.

EXERCISE 3: What Can Be Seen from Space?

Most satellite systems have a limited number of spectral bands with which to detect a spectral signature. In the following exercise, you will draw the location and width of the spectral bands of a number of satellite sensors and estimate what each land cover types looks like when viewed by different sensors types.

- Using Table 1.3 and several printed copies of the vegetation spectral curve handout as shown in (Figure 1.5), draw boxes that represent the spectral channels of the Landsat ETM+, Spot 5 XS, and QuickBird satellite sensors.
- Complete a new graph for each sensor and work in teams if appropriate. *NOTE*: The spectral range (*X*-axis) of this figure differs from Figure 1.2.

Q10 For each sensor, draw each spectral channel as a box (similar to Figure 1.2) using a separate copy of Figure 1.5 for each.

Q11 Once you have drawn the boxes for each sensor, draw a line across each box to approximately represent the average spectral response of vegetation within each spectral channel.

Q12 For each sensor, connect the average spectral responses for each box/channel to draw the plot of the vegetation spectral signature as seen by each sensor.

Q13 How does the averaged spectral curve differ between sensors? Which sensor reproduces the vegetation spectral curve the best? Which is more important, the number of bands or where they occur across the spectrum?

EXERCISE 4: Visual Representations of Satellite Imagery

Now that you have completed some manual interpretation of spectral signatures, we will move on to exploring actual satellite imagery using computer software. In the following exercise, you will use ArcMap, a powerful GIS analysis program for use with both vector and raster file formats. While sample Landsat satellite images for select locations have been supplied for this exercise, the entire Landsat archive covering the entire Earth is available for free and can be accessed using the United States Geological Survey's Global Visualization Viewer (http://glovis.usgs.gov). Here, you will examine a Landsat scene covering the city of Vancouver in British Columbia, Canada.

In this section, you will explore basic computer visualization techniques, allowing you to produce various color representations of a single satellite image. Computer displays typically use a combination of red, green, and blue colors to display images on the screen. Each of these colors is assigned a **color channel**, and the different combinations of values for each channel are used to produce a range of colors. When viewing satellite imagery, we assign a different satellite band to each of the color channels. Since many sensors collect more than three bands of spectral information, as we learned earlier, there is a wide range of color combinations that can be displayed. The first combination of satellite bands to display will visualize the imagery as it normally appears to the human eye. This is what is referred to as a **true color composite**.

1. Double-click **Exercise4.mxd** to start ArcMap with the Vancouver Landsat imagery.

Table 1.3 Spectral ranges of selected satellite sensors

Sensor	Wavelength range (nm)
Landsat ETM+	
Blue	450–510
Green	520–605
Red	630–690
Near infrared	750–900
Shortwave infrared	1550–1750
Thermal	10,400–12,500
Shortwave infrared	2080–2350
SPOT 5 XS	
Green	500–590
Red	610–680
Near infrared	790–890
Shortwave infrared	1580–1750
QuickBird	
Blue	450–520
Green	520–600
Red	630–690
Near infrared	760–900

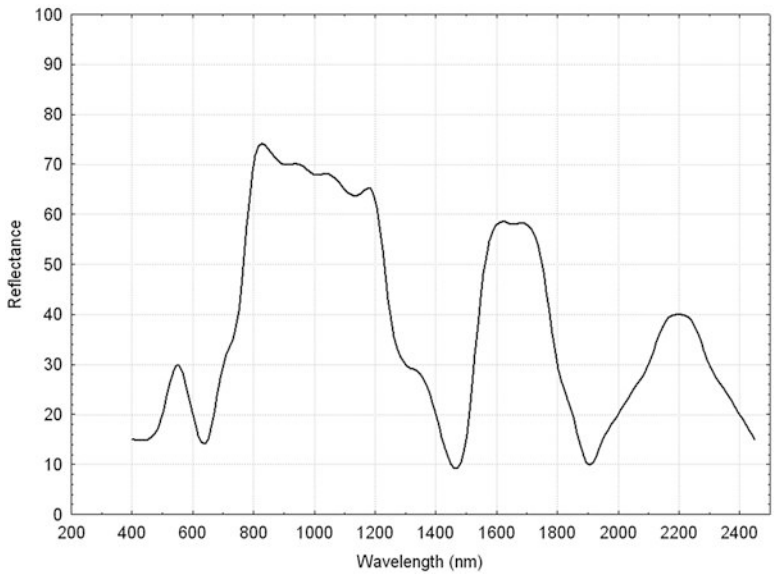


Figure 1.5 Typical vegetation spectrum

2. Right-click the **LandsatYVR** raster, choose **Properties**, and then select the **Symbology** tab from the top menu.

Notice the bands and pull-down menus associated with each color channel. Bands 1, 2, and 3 of the Landsat imagery correspond with the blue, green, and red region of the spectrum, respectively.

3. Now click on the **Histograms** button and notice the statistics for each of the color channels.

Q14 What is the range of values for each color channel?

In this imagery, there are 256 values (0–255) that can be stored for each band. Recall the earlier description of radiometric resolution, in terms of bits, which are the basic unit of information in computing and can only be one of two values (0 or 1). Notice that $2^8 = 256$; therefore this imagery has an 8-bit radiometric resolution.

Satellite imagery captured across different regions of the EMS allows us to examine information not visible to the human eye. Interestingly, many organisms perceive different regions of the EMS than humans can detect (e.g., ultraviolet light is visible to some birds and bees). Humans can, however, create alternative views of a landscape that highlight various features using different combinations of satellite bands with different color channels, creating a **false color composite**. There is a wide range of false color composites that can be created from combining different satellite bands.

4. Select one of the three unique regions (Dubai, Haiti, oil sands) and open the associated **.mxd** file. Explore different, unique combination of bands that highlight different feature attributes.

Q15 Create two separate representations of your chosen scene. Try to create contrasting representations which highlight different features. Create your new representations by associating different bands with different color channels. List some of the observable features in each of your two new images.

5. Lastly, test your feature identifications using higher spatial resolution imagery. Open the program **Google Earth** from the start menu, and direct the viewer to the same location as your Landsat image. Use the provided ***.xml** files, if needed, which are compatible with Google Earth.

Q16 How accurate were your observations? Did anything in particular surprise you? Find someone in the class exploring a different location. Discuss some of the similarities and differences between your regions, the features visible in your satellite-image color composites, and how they compare with higher spatial resolution imagery.

ADVANCED EXERCISE 5: Using Satellite Indices - the Normalized Difference Vegetation Index (NDVI)

To represent specific attributes, remote sensing scientists often compute the difference between image bands to produce a new layer. The resulting layer is referred to as an index, due to its ability to facilitate the extraction of specific information related to ground features. A common index in remote sensing is the **Normalized Difference Vegetation Index (NDVI)**. This index utilizes the difference between the near-infrared region of the spectrum (where vegetation displays high reflectance) and the red region (where vegetation has a very low reflectance). The utility of NDVI comes from its ability not only in identifying vegetation classes but also in differentiating vegetation species and assessing the condition and health of vegetation. NDVI has been used extensively for a wide range of ecological mapping applications, in harvesting operations, as well as in conservation planning and vegetation assessment.

Calculating NDVI

In this exercise, you will learn how to create an NDVI image using Landsat satellite imagery and understand how it can be used as an index of vegetation condition. You will see how different land cover types produce different NDVI values. Vegetated areas generally have higher NDVI values, making the development of a vegetation mask (or layer) a relatively routine and easy task.

1. Double-click **AdvancedExercise.mxd** to start ArcMap with the same Vancouver Landsat imagery used in the previous exercise.

Notice that each of the bands from the Vancouver Landsat scene is now loaded independently. Each band represents a region of the electromagnetic spectrum as indicated in Table 1.3 starting with band 1 as blue.

2. Compare the differences in the radiance between bands by turning each of the bands on and off by clicking the check box beside each layer.

Q17 From your knowledge of spectral signatures, which two of the Landsat satellite band numbers correspond with the highest and lowest reflectance for vegetation?

3. From the top menu, choose **Geoprocessing** and select **Search For Tools**.
4. In the search space, type **Raster Calculator** and then hit **Enter** and select the **Raster Calculator (Spatial Analyst)** tool from the top of the search list.

NOTE: The Spatial Analyst extension must be enabled before using the Raster Calculator tool. You can turn on the extension (if available with your license) by choosing **Customize** from the top menu, selecting **Extensions**, and ensuring that the **Spatial Analyst** box is checked.

5. In the map algebra expression space, enter the following:

   ```
   Float(“LandsatYVR.tif-Band_4” - “LandsatYVR.tif-Band_3”)/Float(“LandsatYVR.tif-Band_4”+“LandsatYVR.tif-Band_3”)
   ```

   - Then give the *Output raster* a location and name you will remember (and where you have permissions!) and click **OK**.
   - The new layer should display a grayscale image of NDVI values
6. Use the identify tool to examine NDVI values for individual pixels in your NDVI layer.

Q18 Clearly some features have higher NDVI values than others. Using a combination of the color composite image and the NDVI layer, what cover types look to be at the extreme ends of the range of NDVI values?

Q19 Provide some ideas as to why different vegetative cover types display different NDVI values

Q20 Return to Google Earth and check your work from *Q18*. Are there any cover types or features that presented NDVI values that seem surprising or incongruous? Why might that be?

Q21 What are some ways that you might use NDVI to map vegetation cover? What other data are needed to accomplish this?

Image Thresholding

In the next few steps, you will learn to use the NDVI layer to classify vegetation using a **thresholding technique**. A threshold is a limit used to divide a continuous set of values. Thresholding is a common approach used in remote sensing to classify one or more land cover features from an image. Thresholding can be done both manually with user-determined threshold values or automatically using statistical methods.

7. Using the identify tool to explore the NDVI layer you created in the previous step. Find the lowest value that you consider to be vegetation. Feel free to use Google Earth as a reference.
8. Open the Raster Calculator tool again, and in the map algebra expression space, enter

   ```
   raster_NDVI>x
   ```

   where *raster_NDVI* is the name of the NDVI layer you created in the previous step and *x* is your determined threshold value. Click **OK**.

The resulting layer is a binary image which contains two values. A value of 1 indicates that a condition has been met, and 0 indicates where a condition has not been met.

Q22 In the binary image that you created, what do the values 1 and 0 represent?

Q23 What benefits does a binary image like this offer for analysis? What are the disadvantages?

Exporting Your Images

The last step is to export your binary image so you can use it at a later date or with different software packages. One of the most common storage file formats for spatial raster datasets is a GeoTIFF, which you will use here to store your classified vegetation layer.

9. Right-click the most recent layer representing the vegetation extraction in the **Table of Contents pane**, choose **Data**, and select **Export Data**.
10. From the **Format** drop-down menu, select **TIFF**.
11. Choose a location folder and enter a name for your export layer (ensuring that it ends with the extension *.tif*) and then click **Save**.

Q24 Do a quick web/article search and list several of the various satellite-image-based indices that have been used to identify land cover features or phenomena. Select an index from your list and apply it to one of the three other locations (Dubai, Haiti, oil sands). Discuss how the index you selected might be useful for informing land use management at your location.

SYNTHESIS EXERCISE 6: Remote Sensing Applications

From your understanding of this brief introduction to remote sensing and vegetation analysis, find a research article that interests you and is pertinent to your studies which utilizes remote sensing. Read and prepare a brief review. The purpose of doing this is to familiarize yourself with remote sensing literature and relate the remote sensing concepts above to your specific interests.

You may choose any paper on remote sensing of vegetation; however, the paper should be no more than 5 years old. The best way to conduct this search and find a paper is online or by browsing through the remote sensing journals available in university libraries. In addition, there are several remote sensing journals, all of which contain papers that would be quite suitable for this exercise. In your search, consider searching for the term NDVI.

Your summary should be about two pages long and include the following:

- Describe why this paper is of interest to you.
- How did you find the paper?
- What type of remote sensing imagery has been used?
- Was the NDVI, or a similar vegetation index, calculated?

- How was it used in the study?
- Location/region of the study.
- What are the general outcomes/results?
- What were some of the sources of error in the study? Were they discussed?
- A paragraph indicating if you would have undertaken the study any differently.

Be prepared to also hand in a copy of your chosen journal article when you hand in your write-up.

CONCLUSIONS

Remote sensing is a dynamic and rapidly evolving field. You are encouraged to explore the recommended readings for additional ideas and applications for remote sensing. In addition, a subsequent chapter in this book (see Chapter 11 Using Spatial Statistics and Landscape Metrics to Compare Disturbance Mosaics explores the use of other vegetation indices for mapping disturbances such as fire and insect outbreaks. This advanced chapter also explores the implications of using thresholding techniques versus binary maps to represent and analyze landscape disturbance, building on the foundations of remote sensing you examined here.

REFERENCES AND RECOMMENDED READINGS[1]

*Bannari A, Morin D, Bonn F et al (1995) A review of vegetation indices. Remote Sens Rev 13:95–120. *This classic paper reviews the basis of vegetation indices and provides some context and decision points users can follow on the section of indices for their particular application.*

Colwell RN (1960) *Manual of photographic interpretation*. ASP&RS, Falls Church

Franklin SE, Wulder MA (2002) Remote sensing methods in large-area land cover classification using satellite data. Prog Phys Geog 26(2):173–205

*Gergel SE (2007) New directions in landscape pattern analysis and linkages with remote sensing. In: Wulder MA, Franklin SE (eds) *Understanding forest disturbance and spatial pattern: remote sensing and GIS approaches*. Taylor and Francis, Boca Raton, pp 173–208. *This review paper provides a nice segue from remote sensing imagery to its ecological uses for pattern analysis in landscape ecology.*

*Hudak AT, Evans JS, Smith AMS (2009) LiDAR utility for natural resource managers. Remote Sens 1:934–951. *Provides an easy to understand practical description of LIDAR technology, its processing, and exposure to a broad range of applications.*

Lillesand TM, Kiefer RW (2000) *Remote sensing and image interpretation*. Wiley, New York

Nelson T, Wulder MA, Niemann KO (2001) Spatial resolution implications of digitising aerial photography for environmental applications. J Imag Sci 49:223–232

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Stone C, Kile G, Old K, Coops NC (2001) Forest health monitoring in Australia, national and regional commitments and operational realities. Ecosyst Health 7(1):48–58

Strahler AH, Woodcock CE, Smith JA (1986) On the nature of models in remote sensing. Remote Sens Environ 20:121–139

*Unsalan C, Boyer KL (2011) Review on land use classification. Part 3. In: *Multispectral satellite image understanding: from land classification to building and road detection*. Springer, London, pp 49–64. *This paper provides background on conventional land cover and land use classifications for remote sensing imagery.*

*Wulder MA, Hall RJ, Coops NC, Franklin SE (2004) High spatial resolution remotely sensed data for ecosystems characterization. BioScience 54(6):511–521. *An overview on the use of high spatial resolution data and how it differs from broader scale imagery with examples of potential applications of this type of data.*

Wulder MA, Kurz WA, Gillis MD (2004) National level forest monitoring and modeling in Canada. Progr Plann 61:365–381

Chapter 2
Historical Aerial Photography for Landscape Analysis

Jessica L. Morgan, Sarah E. Gergel, Collin Ankerson, Stephanie A. Tomscha, and Ira J. Sutherland

OBJECTIVES

Historical patterns and spatial heterogeneity can greatly influence dynamics of contemporary landscapes. Historical conditions lay the foundation for contemporary management options and can help guide restoration goals. While historical spatial data sources are not generally common, historical aerial photography provides the longest available, spatially contiguous record of landscape change. Aerial photography has been routinely collected since the 1930s in many parts of the world and has aided land management for over 75 years. Aerial photography often forms the basis of a variety of maps routinely used by managers, including forest ecosystem inventories and digital elevation models (or DEMs). Aerial photographs generally provide higher spatial resolution information than widely available (and free) satellite imagery (e.g., Landsat). Thus, aerial photographs have unique value for mapping historical landscape baselines and assessing long-term landscape change. For these reasons, understanding how information is derived from aerial photography is enormously important for landscape ecologists. The objectives of this lab are to help students:

1. Understand how landscape heterogeneity can be characterized using aerial photographs;
2. Gain introductory exposure to the benefits and challenges associated with interpretation of aerial photography; and
3. Explore the utility of historical spatial data for characterizing baseline conditions and understanding landscape change.

J.L. Morgan • S.E. Gergel (✉) • C. Ankerson • S.A. Tomscha • I.J. Sutherland
University of British Columbia, Vancouver, BC, Canada
e-mail: sarah.gergel@ubc.ca

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_2

This lab is divided into several parts designed for teams of students to analyze aerial photography and then compare and discuss their results. Part 1 provides a fun introduction to viewing aerial photography "in stereo". Part 2 explores the comparative heterogeneity and fragmentation seen in historic and contemporary landscapes. Students manually photointerpret contemporary and historic images from the same landscape and then compare results with their classmates and to those from a professionally trained interpreter. Part 3 introduces additional considerations including potential sources of error in maps, how such uncertainties can impact their utility, and how terrain and productivity patterns can impact photo interpretation and landscape change. While our examples primarily focus on coastal temperate forests, it is important to note that aerial photographs are used to assess and monitor a diversity of landscapes all over the world, including aquatic, marine, tropical and polar environments. Several of the exercises can be adapted to use imagery from your local region and additional more advanced exercises are available on the book website.

At a minimum, students will need printed copies of the images in the **StereoPairs** and **OrthoPhotos** folders as well as the tables which are provided digitally (see book website), along with a few colored pens/pencils and a calculator. Your instructor may also wish to provide some plastic overlay transparencies on which to draw. A hair tie, for holding one's hair back, may also be helpful. A computer is not necessary if ALL the images in the rest of the **Spatial Data** folder are printed; otherwise, students will need to view these additional images on-screen. Your instructor may wish to provide a stereoscope, which is useful but not required, for demonstrating 3-dimensional (3-D) concepts in the lab.

INTRODUCTION

Aerial photography is used extensively in environmental monitoring and management (Morgan et al. 2010). Captured over a variety of spatial scales, aerial photographs are used for a wide variety of purposes in resource management, from detailed surveys of individual trees to general land cover mapping over broad extents. Common uses include creation of forest inventories, disturbance mapping, estimating productivity, and wildlife management (Avery and Berlin 1992). The fine detail (high spatial resolution) of some aerial photography is particularly well suited for mapping small features or ecosystems (Fensham and Fairfax 2002; Tuominen and Pekkarinen 2005). For example, aerial photographs have been used to identify canopy gaps and forest structures important for wildlife (Fox et al. 2000). Additional examples are shown in Figure 2.1. Examine these four images and try to identify some recognizable features.

Archival historical aerial photography can also provide valuable information on prior or baseline landscape conditions, making the imagery useful for mapping and monitoring change over time (Morgan and Gergel 2013; Cohen et al. 1996; Fensham and Fairfax 2002). The first known aerial photograph was captured in 1858 from a balloon over France. However regular collection did not begin until World War I, primarily for military reconnaissance (Lillesand and Kiefer 2004). Because historical aerial photography offers a window into the past, it has been invaluable for detecting encroachment of invasive species over time (Hudak and Wessman 1998; Laliberte et al. 2004; Mast et al. 1997).

Part 1. Viewing Stereo pairs

Aerial photographs are captured with an airborne camera and represent the reflectance (or relative brightness) of features on the ground. Aerial photographs are often acquired along a **flightline** (i.e., a path flown in a constant direction over a targeted area). A critical component of collection along flightlines is that adjacent

Figure 2.1 Examples of landscape features distinguishable from fine scale aerial photography. Shown is an area from Washington State, USA in 2006. More details can be found in Tomlinson et al. 2011 who contrasted this imagery with similar locations in 1949 to examine changes in fish habitat. (**a**) Sinuosity (curvature) of rivers and the extent of riparian zones (1:5000)

Figure 2.1 (continued) (**b**) Agricultural type (hay field vs. orchards) (1:5000). Orchards are indicated with their dark green, regularly spaced tree crowns. Hay fields are beige with a smooth texture. (**c**) Sediment loads and relative depth in aquatic environments (1:10,000)

Figure 2.1 (continued) (**d**) Urban–wildland interface and urban density (number of houses or roads in an area) (1:10,000)

photographs possess some degree of spatial overlap (often up to 60%). This overlap presents the landscape from two different viewpoints, and thus can be used to view various features in 3-dimensions. Any two adjacent photographs with overlap are referred to as **photo pairs** or **stereo pairs** and are most easily viewed in 3-D with the aid of a stereoscope.

EXERCISE 1: Seeing in Stereovision

Example stereo pairs have been provided in the folder labeled **StereoPairs**. If your instructor is able to provide a stereoscope, you can follow these steps. If no stereoscope is available, skip to the *"Low-Tech" Method 2*.

Method 1: Stereoscope

To view with a stereoscope, the simplest approach is:

1. Examine the two photographs and notice the zone of overlap (i.e., the portion of the landscape captured in both images).
2. Place the photographs within the field of view of the scope. Be sure to place the left and right images under the corresponding left and right eyepieces.
3. Within this overlapping zone, identify the same notable feature (or location) in each image with a finger.
4. Now looking through the scope, align your two fingers so they match up within your field of view. The notable features should then also be close to aligned and thus appear 3-D.

Viewing in stereo is not easy for everyone, particularly for people with unequal vision in each eye. For those who find it easy with a stereoscope, you may even be able to view photo pairs in stereo without one.

Method 2: Low-Tech

1. Using Figure 2.2, place an index card (or piece of folded paper ~20 cm high) on the line between the photographs.
2. Position your forehead directly on top of the card. The index card forces your left eye to focus on the left photograph and your right eye to focus on the right photograph.

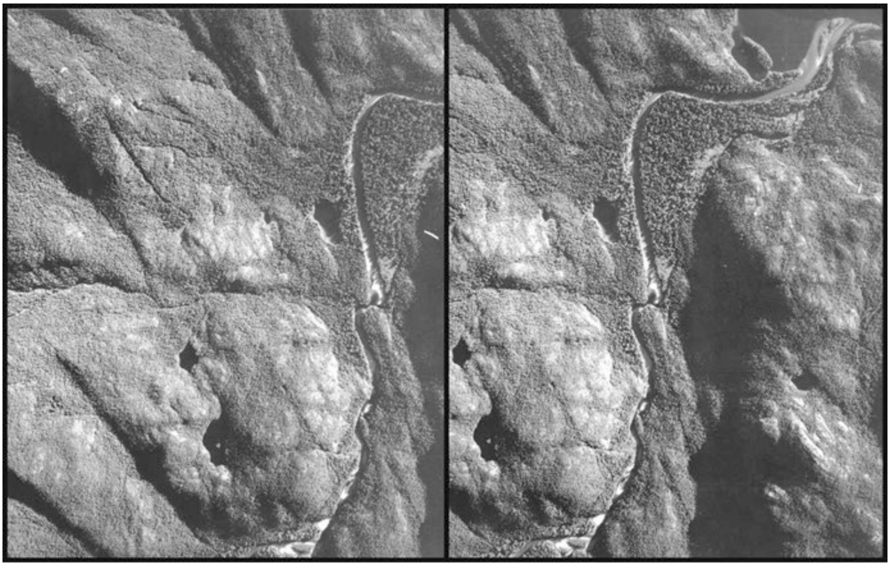


Figure 2.2 Stereo pair from coastal British Columbia captured in 1937. A printable version is available from the book website in the **StereoPairs** folder (see **Site 1**)

3. Concentrating your vision (and "relaxing" your focus), imagine bringing the two images together so they align in the middle of your view. With some patience, hopefully the image will "pop" for you at some point, giving you a deep view of the terrain of the valley.
4. It may also be helpful to try and focus your eyes "through" or "past" the images and then pick a feature (such as the river), and attempt to bring it together into focus.
5. Remember, only the area of photo overlap will be visible in 3-D. You will also see the outer parts of the two images (but blurry and not in 3-D) on either side.
6. This may not work for everyone, so move on after trying for a few minutes.

Viewing stereo pairs in 3-D without a stereoscope requires practice and patience, but once your skills become more advanced you will find it much easier to achieve stereovision. You might also wish to try again at the end of the lab after your eyes rest.

Modern aerial photography is commonly captured in color which provides more information than **panchromatic** (black and white) historic photographs, particularly for species classification and assessment of vegetation health. Conventionally, most aerial photographs were captured with a film-based camera and then converted into digital format via scanning (Wolf and Dewitt 2000). However, a recent shift towards digital cameras has aided instantaneous capture of photographs in digital format with integration of geographic positional system (GPS) data. Unmanned aerial vehicles (UAVs) or "drones" are providing novel opportunities for capturing high resolution digital photography in ways that link extremely well with spatial ecological questions (Getzin et al. 2014) and connect well with other monitoring approaches such as satellite imagery, fieldwork, and citizen science (Turner 2014).

Next, you will examine aerial photographs over two time periods and explore how different methods of analysis can be used to extract a diversity of information useful in answering important landscape ecological questions.

Part 2. Exploring Manual Photointerpretation

As much an art as a science, manual interpretation has been the primary technique used to derive ecological information from aerial photographs for eight decades (Morgan et al. 2010). While techniques have evolved greatly, from the use of plastic overlays to complex computer software, the basic approach remains similar (Avery and Berlin 1992).

First, the process of **polygon delineation** creates a series of polygons on an image (perhaps drawn "freehand") in order to delineate homogeneous areas (or patches) with similar properties. In this lab, we will be focusing on **forest patches** (or forest stands), areas which are relatively homogenous with respect to tree size and species mix. Forest polygons are routinely delineated for inventory of timber, wildlife habitat, and other features of interest to management and research.

Second, the characteristics within each polygon (e.g., dominant species or disturbance type) are interpreted and a general classification is assigned. **Classification** is based on convergence of evidence, meaning the interpreter uses a variety of characteristics on the photograph to identify features on the ground (see Table 2.1). In addition to what the interpreter can extract visually, general knowledge of the area as well as on-the-ground experience with the local habitats and ecosystems contributes greatly to the interpretive process.

Table 2.1 Eight primary characteristics used in manual interpretation of aerial photographs, adapted from Morgan et al. (2010)

Characteristic	Definition	Use in manual interpretation
Tone/Color	Relative brightness or hue of pixels	Natural and anthropogenic feature identification
Size	Area (or number of pixels) of a feature or patch	Vegetation age and structure, habitat suitability, urban features/land use
Shape	Relative complexity of a feature/patch border or edge	Identification of natural (irregular shapes) and anthropogenic (geometric shapes) features
Texture	Frequency of change in tone among pixels; smoothness or roughness	Vegetation identification, biodiversity estimates, surface properties of a feature/patch
Pattern	Spatial arrangement and repetition of features or patches across an area	Land use, disturbance, habitat suitability, landscape structure
Shadow	Dark or "shadow" pixels caused by difference in elevation of a feature relative to surroundings	Feature identification and orientation
Site	Environmental conditions of the delineated feature/patch	Microclimate, species, local habitat suitability
Context	Conditions adjacent to, or surrounding, a feature or patch	Land use

EXERCISE 2: Manual Classification of Contemporary Forests

The purpose of Exercises 2 and 3 is to gain a general understanding and appreciation for the basic approach used by interpreters to analyze aerial photographs when creating forest cover maps. This exercise requires the use of colored pens and printed copies of the aerial photographs from the **OrthoPhotos** folder.

The imagery you will analyze was assembled as part of a long-term ecological research project in Clayoquot Sound, British Columbia, Canada, near Tofino, BC (Gergel et al. 2007; Morgan and Gergel 2013; Thompson and Gergel 2008). The region has changed greatly due to decades of harvest (Figure 2.3). Extensive restoration projects are currently underway in the area with a primary goal of restoring riparian forests and fish habitats. Increasing interest in spiritual and aesthetic values of these forests also supports a tourism economy. Dominant tree species can reach hundreds of years in age. Viewing the broader region in the 1970s shows the patterns of forest harvest (Figure 2.3). Using a much smaller spatial extent, you will examine forest cover change in the area using more contemporary imagery as well as historical data from several decades prior.

Figure 2.3 Regional view of the Clayoquot Sound landscape near Tofino and Ucluelet, British Columbia, surrounding the smaller area you will be examining with more recent (1990s) and historical (1930s) photographs. Here, an orthophoto has been created demonstrating landscape condition in the 70s/80s

Working as small teams (or groups of two) you will start by classifying the contemporary (circa 1996) photographs of the Kennedy Lake, British Columbia using the categories described in Table 2.2 and Figure 2.4. The area of this modern orthomosaic is 6.85 km^2. Read the series of steps (1–6) below, *before* you begin.

1. Within the folder entitled **OrthoPhotos**, print hard copy of the image entitled **Modern**.
2. As a first step, use a colored pen to delineate the most obviously disturbed patches. These areas might include disturbances such as roads and recently logged areas. You may also find it helpful to refer to Table 2.1 to remind yourself of the generally useful characteristics for photointerpretation.
3. Using your marker, delineate all polygons (patches) which appear visually similar.
4. Next, carefully examine Table 2.2 and its accompanying visual in Figure 2.4. Together they explain and illustrate some basic forest types found in the region.
5. Next, assign a class to each polygon. Try to discriminate late seral and second-growth forest patches. Late seral patches refer to older forest stands which have never been harvested. Second-growth stands have younger smaller trees.
6. The above exercise should take no more than 25–30 min. You will need to exercise your own judgment and make a surprising number of decisions and "rules" as you complete this task—so take good notes of any decisions you make along the way.

(*NOTE:* Keep in mind variation is common even amongst trained, experienced interpreters.)

Table 2.2 Basic classification scheme for **modern** aerial photographs in coastal BC

Class	Description
Water	• dark grey/black or light grey/white color • smooth or "flat" appearance • possibly rippled texture • rivers with linear shape • lakes with round/oblique shape
Roads	• distinct linear shapes • bright (white) in tone • often adjacent to (or within) harvested areas
Recently Logged	• lighter grey/white color • irregular shapes • sharp, well-defined borders • often adjacent to or enveloping roads
Late Seral Western Redcedar	• trees are light grey in color (the brightest conifer) but patches are dark due to open distribution of trees • rough texture • open distribution of trees • patch edges often occurring as gradients
Late Seral Western Hemlock	• lighter grey color • smooth texture • small patches with indistinct edges
Second Growth	• medium grey color • smooth/fine texture • smaller, inconspicuous tree crowns • often irregular shapes with fairly well-defined borders

See Figure 2.4 for examples. This classification scheme can be modified by teams as they see fit. The simplest features to interpret are general classes such as water, forest, roads, and recently logged areas. Forested areas can further be delineated into patches (or stands) based on the dominant species, age, or other forest characteristics

Once you have completed the above steps, summarize your data as suggested below:

1. Complete Table 2.3 using the row labeled **Your Team's Result**. Remember that depending on the goals of a given project, second growth may be considered "disturbed" forest. Also, you will need to visually estimate % Landscape Disturbance.
2. Compile results on the chalkboard (in a table similar to Table 2.3), so that the results from all teams are available to the entire classroom.
3. Calculate the mean and standard deviation for the classroom and enter in Table 2.3.
4. *Only* when your interpretation is complete, examine the results of an interpretation performed by a professionally trained interpreter located in a folder entitled **Modern Interpretation**.
5. Tally results from the Professional Interpreter in Table 2.3.

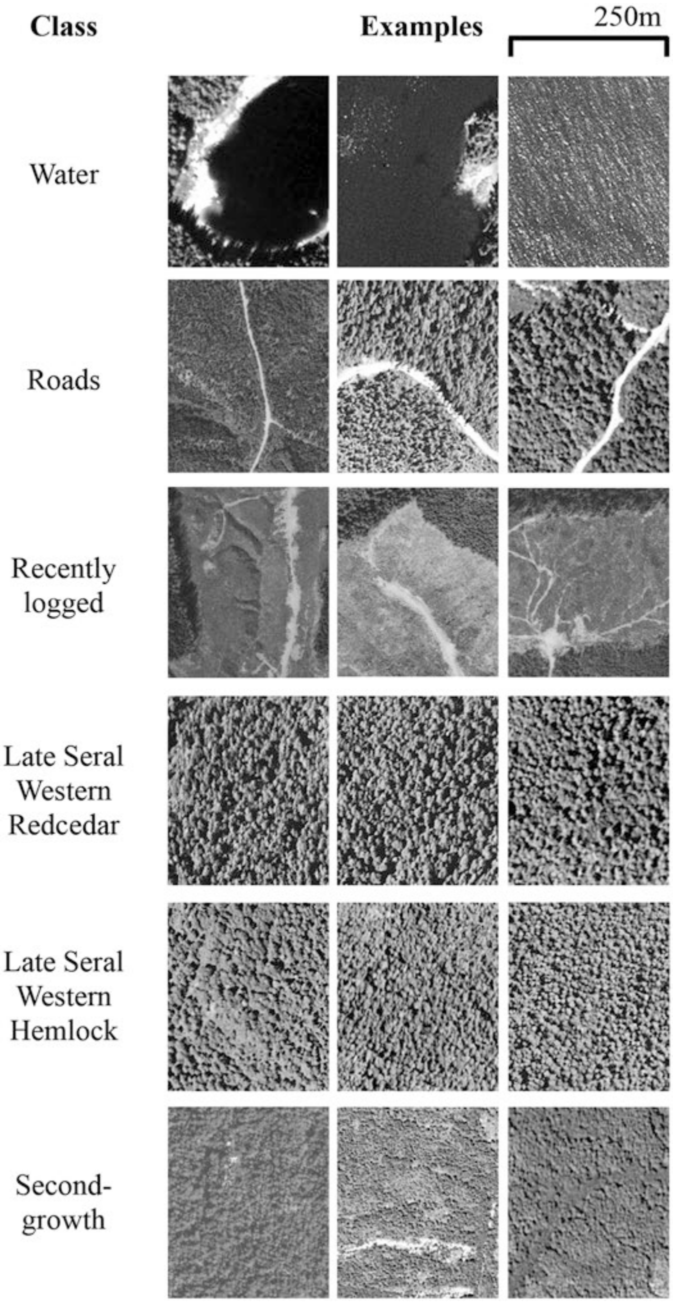


Figure 2.4 Example images for interpretation of **modern** aerial photographs. Refer to classification scheme in Table 2.2 for criteria to assist your interpretation

Table 2.3 Summary of results for **modern** aerial photograph interpretations

Modern	Number of Classes Identified	Total Number of Patches	% of Patches Disturbed	% of Landscape Disturbed
Your Team's Results				
Mean for Classroom				
Std Deviation for Classroom				
Professional Interpreter				

Q1 Explain some of the easier aspects of manual interpretation and also some of the challenges you encountered when using this technique. Were you forced to make some key decisions and assumptions? Explain.

Q2 What are the major differences between your team's interpretation and that of the professional interpreter? What are the similarities?

Q3 How do the results of the professional interpreter compare to the average classroom results? What are some potential reasons for the similarities and differences?

Q4 Considering the standard deviation of the results (Table 2.3), what do you notice about the variability of this technique? Which measures are most and least variable (# classes, # patches, % disturbed)? What might be some reasons for this?

EXERCISE 3: Reconstruction of Historical Forests

Photographs can also be defined based on their geometry as either **vertical** (captured parallel to the ground) or **oblique** (captured at an angle). Oblique photographs captured from airborne cameras or high points on the landscape (such as mountain peaks) can predate vertical aerial photographs by several decades. However, analysis techniques for oblique photos are not nearly as well developed due to the extreme difficulty in systematically extracting information from such photographs. Historical photos, in general, can be challenging to use but do provide some unparalleled advantages for landscape analyses (Morgan et al. 2010; Morgan and Gergel 2013; Jackson et al. 2016; Nyssen et al. 2016).

For the next section of the lab, we are fortunate to take advantage of historical vertical photos which have been orthorectified to help correct for distortion and ter-

Table 2.4 Basic classification scheme for **historical** aerial photographs in coastal BC

Class	Description
Water/Wetland	• dark grey/black or light grey color • smooth texture or "flat" appearance • linear or round/oblique shape
Low Productivity Western Red Cedar	• light grey color • patchy or rough texture • open distribution of trees
High Productivity Western Red Cedar	• dark grey color • coarse texture • individual tree crowns may be visible • often located in floodplains
Low Productivity Western Hemlock	• light grey color • smooth texture • often in smaller patches
High Productivity Western Hemlock	• medium grey color • smooth texture • small patches • often located in floodplains

rain. Here, you will conduct a manual classification at the identical location examined in Exercise 2 (also 6.85 km^2) using historical photographs from 1937.

1. Utilize a printed version of the image entitled **Historical** in the **OrthoPhotos** folder.
2. Classify this image using slightly different categories, as explained in Table 2.4 and shown in Figure 2.5.
3. Using the same general approach as for the modern imagery, fill in the required information for **Your Team's Results** in Table 2.5 based on your interpretation of the historical imagery.
4. Share your results (on the chalkboard) with the entire classroom.
5. Calculate the mean and standard deviation for the combined classroom results and enter in Table 2.5.

Q5 Again, *only when your interpretation is complete*, refer to the interpretations by trained interpreters within the **Historical Interpretation** folder and complete the last row of Table 2.5. Discuss the major similarities and differences between the interpretation of your team, the entire class, and the professional interpreter.

Q6 What challenges did you encounter when using this technique (manual interpretation) on the historic photographs? How did the process compare to the modern imagery?

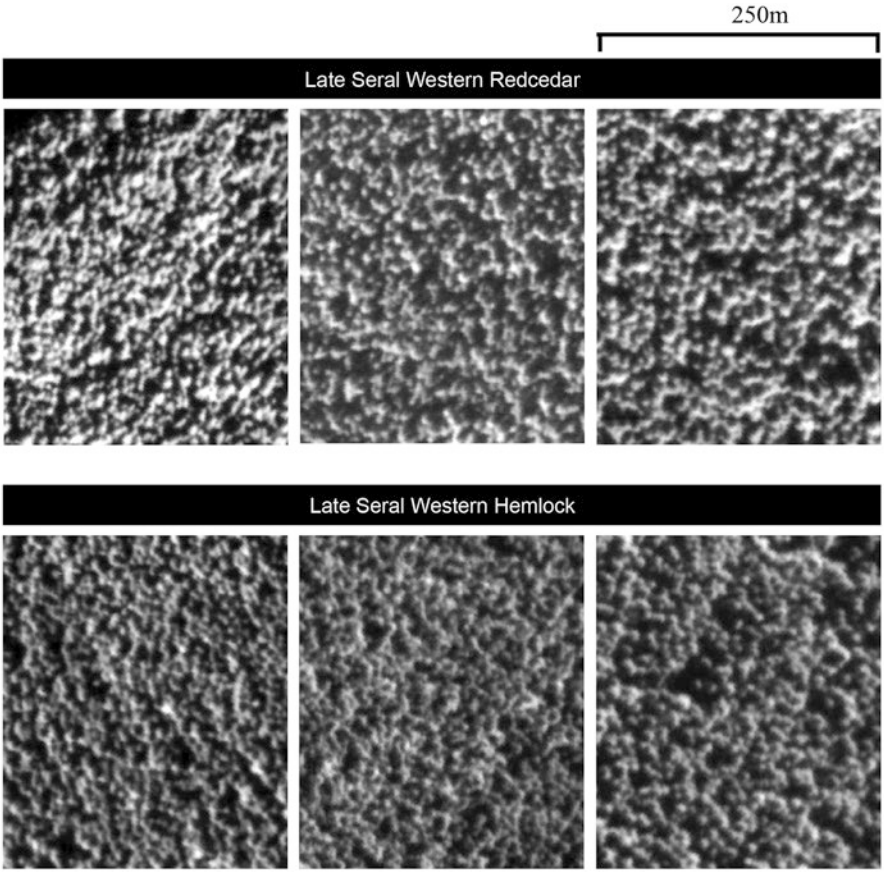


Figure 2.5 Examples of a basic classification scheme for interpretation of **historic** aerial photographs showing some subtle differences between historic forest stands of different species composition. Also see accompanying description in Table 2.4

Table 2.5 Summary of results for historical aerial photograph interpretations

Historic	# Classes	# of Patches	% Patches Disturbed	% Landscape Disturbed
Your Team's Results				
Mean for Classroom				
Std Deviation for Classroom				
Professional Interpreter				

Q7 Which of the eight characteristics of manual interpretation (Table 2.1) were most useful in guiding your interpretation? Which of the characteristics would be the most useful to track within the context of management?

Q8 Considering both Tables 2.3 and 2.5, what do you notice about the changes in the % of disturbed patches and % landscape disturbed between the two time periods? What are the strengths and limitations of such information for examining long-term variability in disturbances?

Q9 Has heterogeneity changed over time in this landscape? How would you quantify heterogeneity in order to answer this question? Does your answer change when you consider "within-patch" heterogeneity as opposed to landscape heterogeneity viewed "among" different patches?

Q10 Are the answers to the two previous questions changed greatly by the assumptions you (and other teams) made? Describe how and why.

Part 3. Additional Considerations for Improving Aerial Photo Analysis

Impact of Errors

Despite the utility of vertical aerial photographs for environmental analysis, errors can hinder interpretation and analysis (Cohen et al. 1996; Tuominen and Pekkarinen 2005). **Geometric errors** refer to positional inaccuracies which can impact both the perceived location of features as well as the size of features on a photograph (Paine and Kiser 2003; Wolf and Dewitt 2000). **Relief displacement** occurs on landscapes with high topographic variability and causes areas closer to the camera lens to appear larger than they actually are, thus misrepresenting the size of features. Before most aerial photographs can be utilized within digital applications (such as a GIS), they must be orthorectified to correct for major geometric errors and provide photographs with an appropriate spatial reference. **Orthorectification** essentially refers to the process by which vertical map coordinates (x, y, and z) are assigned to the photograph to accurately represent distances, angles, and areas (Lillesand et al. 2004). The images you used in Part 1 were stereo pairs (essentially raw imagery) whereas the imagery in Part 2 were orthorectified photographs. **Radiometric errors** refer to incorrect representation of tone/color on a photograph (Jensen 2000) and can sometimes be addressed by adjusting the contrast of the photograph.

Furthermore, errors can arise from the interpretation process. Interpretation errors can include **positional error** (errors in the location and placement of polygons), as well as **classification error** (incorrect assignment of classes). With relatively recent imagery, one can assess the accuracy of a classification through ground verification (or ground-truthing) and collect the data needed to conduct a formal accuracy assessment. When using historic imagery, however, such ground verification is often challenging, if not impossible. As an alternative to ground verification of historic imagery, we can examine uncertainty by asking a professional photointerpreter to quantify their certainty about their classification results, which we examine next.

EXERCISE 4: Uncertainty in Classification

1. Examine the images in the folder entitled **Uncertainty**. Polygons labeled 85, 90, or 100 represent those where the interpreter was confident (or highly certain) of their classification.
2. Identify the areas deemed less certain by the professional interpreter. Note any perceptible characteristics or peculiarities of these polygons.

Q11 Do these "uncertain" areas coincide with any of the areas you found trouble interpreting? Why do you think such areas were hard to interpret?

Q12 Misclassification rates for forest inventories derived from manual interpretation of aerial photography can reach as high as 60% (Thompson et al. 2007). As a team, brainstorm about some potential implications of, and solutions for, a high rate of map misclassification for resource management, conservation, and/or restoration. Prepare to share your answers with the entire class. If your instructor gives you additional time, read Thompson et al. 2007 and/or Gergel et al. 2007 for ideas.

Historic Harvest Patterns and Topography

The fundamental influence of terrain (topographic relief and landscape position) on ecological processes has long been appreciated. Despite the wealth of information obtained solely from visual (tonal, textural) characteristics of aerial photographs, additional insights regarding landscape disturbance patterns can be obtained by accounting for topography using the three-dimensional perspective obtained from stereoscopic photos. Such 3-D information can greatly help improve the process of interpretation.

EXERCISE 5: Benefits of Terrain

For this exercise, you will revisit your interpretations from previous exercises regarding forest harvest patterns. The purpose of this exercise is to understand how the inclusion of topography and terrain information can be key for understanding disturbance patterns across a landscape.

1. Familiarize yourself with the topographic data in the folder entitled **Terrain**.
2. Use the classification scheme outlined in Table 2.6 along with the topographic images, and try to identify terrain classes on your image. (This classification scheme can be applied to both the historic and the modern aerial photographs)

Q13 Can you identify any new features due to the inclusion of topography? What features now become obvious or more easily identified? Are there any changes you would make to the borders of your earlier interpretations based on these terrain classes?

Table 2.6 Topographic classification scheme adapted from the Vegetation Resources Inventory Photo Interpretation Procedures (Province of British Columbia 2002)

Class	Description
Upper Slope	• Upper portion of a hillslope including the crest or ridge of the hill/mountain • This feature is usually convex
Middle Slope	• Area of a slope with a straight profile • Located in between the upper and lower slope features
Lower Slope	• Bottom portion of a hill • Usually concave and characterized by an abrupt decrease in the gradient of the hill's slope
Flat	• Area with a relatively flat/horizontal surface profile not adjacent to a hill base
Wetland/Water	• Area with visible water features • Usually found in areas at the lowest relative elevation • Wetlands are often characterized by a depression (an area that is concave in all directions)

EXERCISE 6: Forest Productivity in Historical Forests

Tree height is an important characteristic used in management because not only is it associated with the general productivity of forest stands but it also influences forest structure, total biomass, potential wildlife habitat and, of course, timber. Well-trained interpreters can estimate tree height for a forest stand using stereo pairs. Most often, interpreters will assign an average tree height value within a homogeneous polygon. Productivity values can also be assigned to polygons by considering a combination of characteristics (in addition to tree height) such as soil moisture, aspect (exposure to sun), and slope.

1. Examine the contents of the folder entitled **Historical Tree Heights & Harvest Patterns** which includes photo-interpreted maps of historic productivity and tree height. Familiarize yourself with these images.
2. Using the historic **tree height** and historic **productivity** maps, determine the number of polygons with tree heights exceeding 30 m, as well as the number of polygons with productivity levels of "good" or "very good." Enter the total number of each in Table 2.7.
3. Compare the locations of historic polygons with tall tree heights and high productivity to the same locations in the modern photograph. Using the modern photograph (and your modern interpretation), estimate how many of these historic polygons have been logged. Enter your results in the final column of Table 2.7.
4. If you find step 3 challenging, examine the file **Logging** providing an interpretation of logging (based on the modern photo) located in the same folder.

Table 2.7 Summary of results for historic forest productivity and subsequent logging patterns according to topography

	Historic productivity		Subsequent harvest	
	# Polygons	# Polygons	# Polygons	# Polygons
Topographic Class	Tree heights >30 m	Good or very good	(from column 1) logged	(from column 2) logged
Upper Slope				
Middle Slope				
Lower Slope				
Flat				
Wetland or Water				

Topographic classes are explained more fully in Table 2.6

Q14 What trends in logging patterns do you notice from the results in Table 2.7?

Q15 Consider some potential ecological (or other) consequences of these patterns of historic harvest. Explain two potential implications for management.

Q16 Discuss how your results are influenced by the uncertainty maps from Exercise 4. Are you more or less confident of your results and interpretation after incorporating the uncertainty maps?

SYNTHESIS

Q17 Consider a landscape you know well. Perhaps it is close to your home or where you have done research. Devise an interesting question for this area utilizing historical aerial photography. Explain why your question is important and briefly explain your expected results (your proposed hypotheses). Explain how aerial photographs (and any auxiliary datasets) would be used in the project.

REFERENCES AND RECOMMENDED READINGS[1]

Avery TA, Berlin GL (1992) *Fundamentals of remote sensing and air photo interpretation*. Prentice Hall, Upper Saddle River

Blaschke T (2004) Object-based contextual image classification built on image segmentation. IEEE Trans Geosci Remote Sens, pp 113–119

Cohen WB, Kushla JD, Ripple WJ, Garman SL (1996) An introduction to digital methods in remote sensing of forested ecosystems: focus on the Pacific Northwest, USA. Environ Manag 20:421–435

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Fensham RJ, Fairfax RJ (2002) Aerial photography for assessing vegetation change: a review of applications and the relevance of findings for Australian vegetation history. Aust J Bot 50:415–429

Fox TJ, Knutson MG, Hines RK (2000) Mapping forest canopy gaps using air-photo interpretation and ground surveys. Wildl Soc Bull 28:882–889

*Gergel SE, Stange Y, Coops NC, Johansen K, Kirby KR (2007). What is the value of a good map? An example using high spatial resolution imagery to aid riparian restoration. Ecosystems 10:688–702. *Introduces user-friendly ways to assess map accuracy in a restoration context, converting map accuracy to potential restoration expenditures. A clear demonstration of the ecological and economic value of high spatial resolution imagery.*

Getzin S, Nuske RS, Wiegan K (2014) Using unmanned aerial vehicles (UAV) to quantify spatial gap patterns in forests. Remote Sens 6:6988–7004

Hay GJ, Blaschke T, Marceau DJ, Bouchard A (2003) A comparison of three image-object methods for the multiscale analysis of landscape structure. ISPRS J Photogramm Remote Sens 57:327–345

Hudak AT, Wessman CA (1998) Textural analysis of historical aerial photographs to characterize woody plant encroachment in South African Savanna. Remote Sens Environ 66:317–330

Huston MA (1999) Local processes and regional patterns: appropriate scales for understanding variation in the diversity of plants and animals. Oikos 86:393–401

*Jackson MM, Topp E, Gergel SE, Martin K, Pirotti F, Sitzia T (2016) Expansion of subalpine woody vegetation over 40 years on Vancouver Island, British Columbia, Canada. Canad J Forest Res 46:437–443. *An example of how historical photography can be used to assess changes in alpine montane communities within the context of changing climate.*

Jensen JR (2000) *Remote Sensing of the Environment: An Earth Resource Perspective*. Prentice Hall, Upper Saddle River.

Laliberte AS, Rango A, Havstad KM, Paris JF, Beck RF, McNeely R, Gonzalez AL (2004) Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. Remote Sens Environ 93:198–210

Landres PB, Morgan P, Swanson FJ (1999) Overview of the use of natural variability concepts in managing ecological systems. Ecol Appl 9:1179–1188

*Li H, Reynolds JF (1995) On definition and quantification of heterogeneity. Oikos 73:280–284. *An enduring classic on ways to think about and quantify landscape heterogeneity—a refreshing read for advanced students after students wrestle with the vagaries of photo-interpretation and how these heterogeneity ideas might translate to aerial photography and other image data sources.*

Lillesand TM, Kiefer RW, Chipman JW (2004) *Remote sensing and image interpretation*. Wiley, New York

Mast JN, Veblen TT, Hodgson ME (1997) Tree invasion within a pine/grassland ecotone: an approach with historic aerial photograph and GIS modeling. For Ecol Manage 93:181–194

*Morgan JL, Gergel SE, Coops NC (2010) Aerial photography: a rapidly evolving tool for ecological management. BioScience 60:47–59. *A user-friendly introduction to aerial photography, especially useful for novice readers less familiar with the technical specifications and terminology used in the literature.*

*Morgan JL, Gergel SE (2013) Automated analysis of aerial photographs and potential for historic forest mapping. Canad J Forest Res 43:699–710. *An overview of the utility of historical photography and ways to compare manual and automated image classification techniques with implications for assessing landscape pattern. This reading pairs well with the advanced optional exercises available on the book website.*

Morgan JL, Gergel SE (2010) Quantifying landscape heterogeneity from historic aerial photographs using object-based analysis. Landsc Ecol 25:985–998

*Nyssen J et al (2016) Recovery of the aerial photographs of Ethiopia in the 1930s. J Cult Herit 17:170–178. *Great example of the global utility of aerial photography for assessing landscape change along with explanation of the sometimes painstaking process of building an archive. In* *some cases, occupying countries are the source of information for historical imagery, so understanding the military history of a region can help in locating sources of previously unknown and under-utilized imagery.*

Paine DP, Kiser JD (2003) *Aerial photography and image interpretation*, 2nd edn. Wiley, Hoboken

Province of British Columbia (2002) *Vegetation Resources Inventory: Photo interpretation process.* Terrestrial Information Branch, Resources Inventory Committee, Victoria, B.C.

Swetnam TW, Allen CD, Betancourt JL (1999) Applied historical ecology: using the past to manage for the future. Ecol Appl 9:1189–1206

*Thompson ID, Maher SC, Rouillard DP, Fryxell JM, Baker JA (2007) Accuracy of forest inventory mapping: some implications for boreal forest management. Forest Ecol Manag 252:208–221. *A very readable exploration of practical issues in using aerial photography for forest inventories. Yields some surprising results with implications for management and research.*

Thompson SD, Gergel SE (2008) Conservation implications of mapping rare habitats using high spatial resolution imagery: Recommendations for heterogeneous and fragmented landscapes Landscape Ecology 23(9):1023–1037

Tomlinson MJ, Gergel SE, Beechie TJ, McClure MM (2011) Long-term changes in river-floodplain dynamics: implications for salmonid habitat in the Interior Columbia Basin, USA. Ecol Appl 21(5):1643–1658

*Tomscha SA, Gergel SE (2016) Ecosystem service trade-offs and synergies misunderstood without landscape history. Ecol Soc 21(1):43. *Using maps of ecosystem services derived from aerial photography, this work demonstrates the value of historical information in understanding landscapes. This reading also sets up the reader well for Chapters 16 or 19 this book.*

Tuominen S, Pekkarinen A (2005) Performance of different spectral and textural aerial photograph features in multi-source forest inventory. Remote Sens Environ 94:256–268

Turner W (2014) Sensing biodiversity. Science 346(6207):301–302

Wiens JA (1989) Spatial scaling in ecology. Funct Ecol 3:385–397

Wolf PR, Dewitt BA (2000) *Elements of photogrammetry with applications in GIS*. McGraw-Hill, New York

Wulder MA (1998) Optical remote sensing techniques for the assessment of forest inventory and biophysical parameters. Progr Phys Geogr 22:449–476

Chapter 3
Citizen Science for Assessing Landscape Change

Jeffrey A. Cardille and Michelle M. Jackson

OBJECTIVES

Citizen science is increasingly recognized as a powerful tool for addressing ecological problems across large areas. Although not a new phenomenon, certain types of citizen science rely on advanced web-based technology not previously available. "Crowdsourcing" refers to a type of citizen science in which large data-collection tasks are allocated to volunteers using the Internet. Such tasks may require minimal time or effort on the part of the volunteer but—when combined with the efforts of many others—can produce enormous datasets that are extremely useful in research and monitoring. By encouraging public participation from people who may not be experts in a given scientific subject, crowdsourcing citizen science aims to gather and collate useful scientific information from a larger number of individuals than would otherwise be feasible. In this lab, you will participate in a crowdsourcing project and explore some of the basic components of citizen science as it can be applied to landscape ecology. This lab is designed to enable students to:

1. Learn about citizen science and its application in landscape ecology;
2. Explore the use of cloud-based forms and spreadsheets for tracking and summarizing results from hundreds of citizen scientists;
3. Become familiar with using Google Earth for scientific purposes; and
4. Gain experience interpreting existing land-cover classifications and aerial photography in order to contrast historic and current land use/land cover.

J.A. Cardille (✉)
McGill University, Sainte Anne de Bellevue, QC, Canada
e-mail: jeffrey.cardille@mcgill.ca

M.M. Jackson
University of British Columbia, Vancouver, BC, Canada

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_3

Working independently, you will assess land-use/land-cover (LU/LC) change over the past 50 years in the Montérégie region of Quebec, Canada. In Part 1, you will familiarize yourself with the region by exploring aerial photos and land-cover classifications from previous decades using Google Earth. In Part 2, for a set of randomly selected points within the study area, you will compare land use and land cover at those points. As you work, your individual efforts will be pooled with the work of your classmates today and from earlier lab sessions at other institutions around the world. You'll need a web browser and the Google Earth software, along with an Internet connection, to complete the lab. Because the web interface associated with this lab will be continually updated and adapted over time, please continue with the lab online, starting here:

http://goo.gl/FQdwXk

Part II
Fundamentals of Quantifying Landscape Pattern

This module builds on the previous module in that it assumes a basic understanding of how maps are created and used to represent landscapes. First, Chapter 4 introduces you to pattern analysis using FRAGSTATS software, the long-standing workhorse of pattern analysis. This hugely popular lab from the first edition still combines hand calculations with computer analyses but has been adapted to incorporate the latest version of the software. Another fundamental challenge in landscape ecology is understanding patterns at multiple scales. Chapter 5 introduces the use of semivariograms for scale detection and for relating known patterns to measures of spatial autocorrelation. These first two labs are helpful prerequisites to several other chapters. Chapter 6 presents the concepts and tools for creating and using neutral landscape models. Exposure to QRule software helps underscore the impact of different patch-definition rules on landscape metrics and the appropriate use of landscape expectations that are spatially neutral. Chapter 7 is an important new addition to the book, providing guidance for the eternally vexing question of "What constitutes a significant difference in landscape pattern?" Here, students will learn how to assign statistical significance when comparing pattern metrics among landscapes.

Chapter 4
Understanding Landscape Metrics

Jeffrey A. Cardille and Monica G. Turner

OBJECTIVES

An extensive set of landscape metrics exists to quantify spatial patterns in heterogeneous landscapes. Developers and users of these metrics typically seek to *objectively* describe landscapes that humans assess *subjectively* as, for example, "clumpy," "dispersed," "random," "diverse," "fragmented," or "connected." Because the quantification of pattern is fundamental to many of the relationships we seek to understand in landscape ecology, a basic familiarity with the most commonly used metrics is extremely important. Several software programs evaluate maps quickly and cheaply, but there are no absolute rules governing the proper use of landscape metrics. To help foster the appropriate use of landscape metrics, in this lab students will:

1. Become familiar with several commonly used metrics of landscape pattern;
2. Distinguish metrics that describe landscape composition from those that describe spatial configuration;
3. Understand some of the factors that influence the selection and interpretation of landscape metrics;
4. Gain experience with landscape pattern analysis using Fragstats; and
5. Observe the correlation structure among some commonly used landscape metrics.

J.A. Cardille (✉)
McGill University, Sainte Anne de Bellevue, QC, Canada
e-mail: jeffrey.cardille@mcgill.ca

M.G. Turner
University of Wisconsin-Madison, Madison, WI, USA

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_4

This lab explores the calculation and interpretation of metrics commonly used in landscape ecology. Emphasis is placed on the understanding gained from actually calculating select metrics by hand rather than only using a metric-calculation package. In Parts 1 and 2, you will manually calculate several landscape metrics for a small landscape to ensure that you understand their underlying mathematics. Although the landscapes used for the hand calculations are much smaller than those typically input to metric-calculation software packages, the concepts and equations learned are the same as those used for full-sized images. Once you have a basic understanding of several metrics, a section using Fragstats (Part 3), the most widely used analysis program McGarigal and Marks (1993) and larger landscape images (Part 4) will help you investigate the behavior of landscape metrics in more realistic settings. In Part 5, you explore the capabilities and limits of using landscape metrics for real-world landscape change at different time periods. Parts 1 and 2 can be completed using only pen and paper (and perhaps a calculator). Parts 3–5 require a computer with the latest version of Fragstats. All files needed to complete the lab are accessible online via links you can find on the website for this book.

INTRODUCTION

The quantification of landscape pattern has received considerable attention since the early 1980s, in terms of both development and application (Romme and Knight 1982; O'Neill et al. 1988; Turner et al. 1989; Baker and Cai 1992; Wickham and Norton 1994; Haines-Young and Chopping 1996; Gustafson 1998; Cardille and Lambois 2010). Along with terrestrial landscapes, metrics are also applied in aquatic systems and marine "seascapes" (e.g., Teixido et al. 2007; Boström et al. 2011). Several of the most commonly used landscape metrics were originally derived from percolation theory, fractal geometry, and information theory (the same branch of mathematics that led to the development of species diversity indices). The increased availability of spatial data, particularly over the past two decades, has also presented myriad opportunities for the development, testing, and application of landscape metrics. To a large degree, metric development has stabilized, caveats about proper use and interpretation are understood (e.g., Li and Wu 2004; Corry and Nassauer 2005; Turner 2005; Cushman et al. 2008), and newly developed methods have improved statistical interpretations of metric values (e.g., Fortin et al. 2003; Remmel and Csillag 2003).

Why are methods for describing and quantifying spatial pattern such necessary tools in landscape ecology? Because landscape ecology emphasizes the interactions among spatial patterns and ecological processes, one needs to understand and quantify the landscape pattern in order to relate it to a process. Practical applications of pattern quantification include describing how a landscape has changed through time; making future predictions regarding landscape change; determining whether patterns on two or more landscapes differ from one another, and in what ways; evaluating alternative land management strategies in terms of the landscape patterns that may result; and determining whether a particular spatial pattern is conducive to movement by a particular organism, the spread of disturbance, or the redistribution of nutrients. In all of these cases, the calculation of landscape metrics is necessary to rigorously describe landscape patterns. However, relating these metrics of pattern to dynamic ecological processes still remains an area in need of further research.

In this lab, you will examine and manually calculate several commonly used landscape metrics for a small landscape to ensure that you understand their underlying mathematics (Parts 1 and 2). Then, once you have a basic understanding of several metrics, two computer-based exercises (Parts 3 and 4) are provided to allow you to calculate metrics using Fragstats and larger landscape images. Finally (Part 5), you explore the capabilities and limits of using landscape metrics for the same real-world landscape at different time periods. During the course of the lab, you will calculate a wide range of metrics of landscape composition and configuration, including Proportion, Dominance, Shannon Evenness, Number of patches, Mean Patch Size, Edge:area ratios, Probability of adjacency, Contagion, Patch Density, Edge Density, Landscape Shape Index, Largest Patch Index, and Patch Richness.

Part 1. Metrics of Landscape Composition

The simplest landscape metrics focus on the composition of a landscape (e.g., which categories are present and how much of the categories there are), ignoring the specific spatial arrangement of the categories on the landscape. In this section, you will examine three metrics designed to assess the composition of a landscape: (1) the proportion of the landscape occupied by each cover type, (2) Dominance, and (3) Shannon Evenness.

Proportion (p_i) of the landscape occupied by the *i*th cover type is the most fundamental metric and is calculated as follows:

$$p_i = \frac{\text{Total number of cells of category } i}{\text{Total number of cells in the landscape}}$$

Proportions of different landscape types have a strong influence on other aspects of pattern, such as patch size or length of edge in the landscape (Gardner et al. 1987; Gustafson and Parker 1992), and p_i values are used in the calculation of many other metrics. Several metrics derived from information theory use the p_i values of all cover types to compute one value that describes an entire landscape. First developed by Shannon (1948), information theoretic metrics were first applied to landscape analyses by Romme (1982) to describe changes in the area occupied by forests of varying successional stage through time in a watershed in Yellowstone National Park, Wyoming. Romme reasoned that indices used to quantify species diversity in different communities could be modified and applied to describe the diversity of landscapes. Dominance and Shannon Evenness are two such metrics that characterize how evenly the proportions of cover types occur within a landscape.

Dominance (*D*) (O'Neill et al. 1988) can be calculated as:

$$D = \frac{\ln(S) + \sum_i \left[p_i * \ln(p_i) \right]}{\ln(S)}$$

where ***S*** is the number of cover types, p_i is the proportion of the *i*th cover type, and *ln* is the natural log function. The maximum value of this index, given *S* cover types, is ln(*S*); dividing by the maximum value scales the index to range between 0 and 1. Values of *D* near 1 indicate a landscape dominated by one or few cover types, while values near 0 indicate that the proportions of each cover type are nearly equal.

Shannon Evenness Index (*SHEI*) (Pielou 1975) can be calculated as:

$$SHEI = \frac{-\sum_i \left[p_i * \ln(p_i) \right]}{\ln(S)}$$

where *S* is the number of cover types, p_i is the proportion of the *i*th cover type, and *ln* is the natural log function. Values for *SHEI* range between 0 and 1; values near 1 indicate that the proportions of each cover type are nearly equal; values near 0 indicate a landscape dominated by one or few cover types.

A very important detail to note in the formulations of information theoretic metrics is whether or not a particular metric has been normalized to a standard scale. Some early applications of Dominance and Shannon Evenness were not normalized (e.g., O'Neill et al. 1988). The non-normalized forms of these metrics are very sensitive to the number of cover types *S* in the landscapes, and thus comparisons among landscapes that differed in *S* were problematic. Normalizing a metric ensures that its values fall within a standardized range, such as from 0 to 1 (and not from 0 to 157, for example!). With *D* and *SHEI*, the normalization involves dividing the numerator by the maximum possible value of the index (ln *S*), as shown above.

CALCULATIONS

To understand these metrics and calculate them by hand within a reasonable time frame, you will calculate the metrics for two small hypothetical landscapes represented as 10 × 10 grids (Figure 4.1). It may be useful to print paper copies of these small landscapes for your hand calculations.

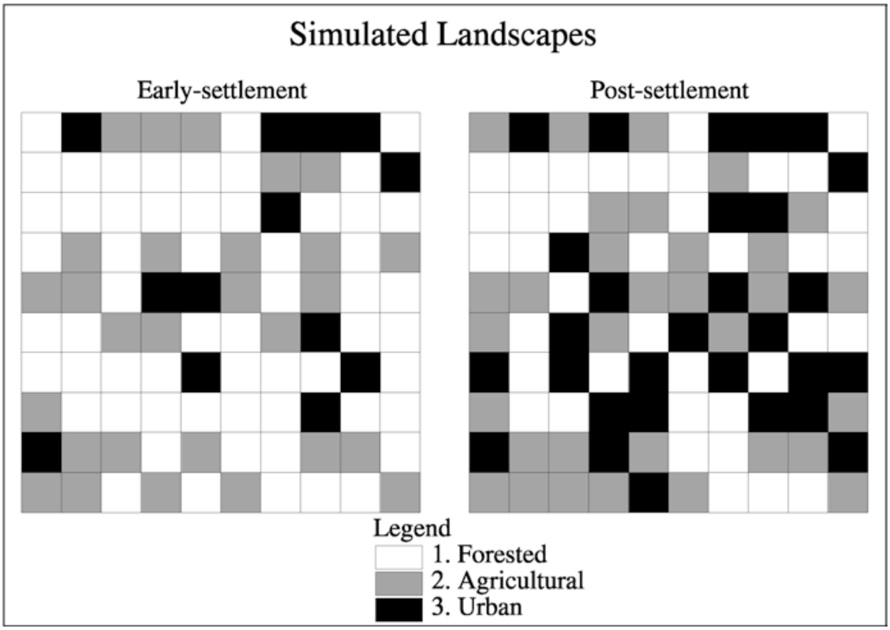


Figure 4.1 Hypothetical early-settlement and post-settlement landscape classifications

Metrics of Landscape Composition in an Early-Settlement Landscape

An invented "early-settlement" landscape is shown on the left in Figure 4.1. This image is intended to represent an area that was previously fully forested, but has lost some forest to agricultural and urban uses. The landscape is composed of a 10×10 grid with each grid cell representing an area of 1 km^2 (1000 m $\times$ 1000 m; 10^6 m^2).

Calculation 1: Calculate the proportions occupied by each of the three land covers in the early-settlement landscape. Record the values in Table 4.1.

Table 4.1 Metrics of landscape composition in an **early-settlement** landscape

Proportion occupied by:	Result
Forested	
Agricultural	
Urban	
Dominance	
Shannon Evenness Index	

Calculation 2: Calculate Dominance for the early-settlement landscape and record in Table 4.1.

Calculation 3: Calculate Shannon Evenness for the early-settlement landscape and record in Table 4.1.

Metrics of Landscape Composition in a Post-settlement Landscape

A "post-settlement" landscape is shown on the right in Figure 4.1. This image represents the exact same area as the early-settlement landscape, but much later in time. Note that more of the forest has been converted to agricultural use. Additionally, some of the agricultural and forest land in the early-settlement image has been converted to urban use, while some of the early-settlement agricultural land has been reverted to forest in the post-settlement image.

Calculation 4: Calculate the proportions occupied by each of the three land cover types in the post-settlement landscape. Record the values in Table 4.2.

Table 4.2 Metrics of landscape composition in a **post-settlement** landscape

Proportion occupied by:	Result
Forested	
Agricultural	
Urban	
Dominance	
Shannon Evenness Index	

Calculation 5: Calculate Dominance for the post-settlement landscape and record it in Table 4.2.

Calculation 6: Calculate Shannon Evenness for the post-settlement landscape and record it in Table 4.2.

Given the answers you obtained for both the early- and post-settlement landscapes, consider the following questions:

Q1 How would you interpret/describe the changes in this landscape between the two time periods?

Q2 Explain the relationship between Dominance and Shannon Evenness. If you were conducting an analysis of a real landscape, would you report both D and $SHEI$? Why or why not?

Q3 Use your calculator to perform some additional calculations of D assuming the proportions listed in Table 4.3.

Table 4.3 Proportion of the landscape occupied by three different cover types in four different landscapes

Landscape	p_{Forested}	$p_{\text{Agricultural}}$	p_{Urban}	Dominance
W	0.10	0.80	0.10	
X	0.80	0.10	0.10	
Y	0.65	0.20	0.15	
Z	0.15	0.20	0.65	

Q4 Which of these hypothetical landscapes might be considered "similar" when only comparing *D*?

Q5 Under what conditions could interpretation of Dominance (or other similar metrics) be problematic?

Q6 Considering your interpretation of the data in Table 4.3, what other types of information and/or metrics would be necessary to distinguish these landscapes?

SYNTHESIS QUESTIONS

Q7 Is there an upper and lower limit of *S* beyond which *D* and *SHEI* will not work?

Q8 To compare *D* or *SHEI* across two or more landscapes, does *S* need to be the same for each landscape in the comparison? Why or why not?

Q9 The developers of the normalized versions of these metrics chose to normalize them using the maximum possible number of cover types that could ever appear in a landscape. What are some other ways that a metrics could be normalized, and how might this change the results?

Part 2. Metrics of Spatial Configuration

A variety of landscape metrics are sensitive to the specific spatial arrangement of different cover types on a landscape. In this section, we will consider four components of landscape configuration: (1) patches, (2) edges, (3) probability of adjacency, and (4) contagion.

The **total number of patches** in a landscape results from first defining connected areas (i.e., patches or clusters) of each cover type *i*. Patches are commonly identified by using either of two rules for evaluating which cells belong to the same patch. A patch may be identified using the **4-neighbor rule**, where two grid cells are considered to be part of the same patch *only* if they are of the same cover type and share a flat adjacency (i.e., horizontal or vertical) between them. Alternatively, the

8-neighbor rule specifies that two grid cells of the same cover type are to be considered as part of the same patch if they are adjacent *or diagonal* neighbors. In reporting the number of patches (or any other patch-based characteristic) it is important to distinguish whether the calculation is for all patches of all cover types or whether it is only for patches of a certain cover type *i*. In addition to the total number, patches can be described in terms of their size (i.e., area) and edge:area ratio, which will be discussed later.

Mean Patch Size (MPS) is the arithmetic average size of each patch on the landscape or each patch of a given cover type. It is often calculated separately for each cover type as follows:

$$\mathrm{MPS} = \frac{\sum_{k=1}^{m} A_k}{m}$$

where m = the number of patches for which the mean is being computed and A_k = the area of the kth patch. The units of area are defined by the user and should always be specified.

Edge calculations provide a useful measure of how dissected a spatial pattern is and can be calculated in a variety of ways. An edge is shared by two grid cells of different cover types when a side of one cell is adjacent to a side of the other cell. The 4-neighbor rule is used for edge counting: diagonals are not used for this aspect of landscape configuration. The total number of edges in a landscape can be calculated by counting the edges between different cover types for the entire landscape. When considering the edges surrounding a given cover type, every edge in the landscape is counted once per cover type. As a result, an edge between a forest and cornfield will be counted once as part of forest edge and once as part of cornfield edge. Edges are sometimes considered with respect to the type of adjacency; in this case, a given forest-cornfield edge would be counted once.

Edge calculations are sometimes used to compute an **edge:area ratio**. Edges may be computed in a variety of ways for a given landscape. For example, the total linear edge in a landscape can be divided by the area of the landscape to provide a single edge:area estimate, or edge density. More useful, however, are computations of edge:area ratios by cover type or for individual patches.

Edge calculations are sensitive to several factors. Whether the actual borders of the landscape image are considered as edges influences both the edge counts and edge:area ratios. (*NOTE:* In this exercise, the landscape border will not be considered edge for your calculations) Computer programs may use slightly different algorithms for totaling edges. It is extremely important to be consistent in both algorithm and units within a set of analyses. Additionally, although edge counts are relatively simple to compute from a landscape map, they can be very sensitive to the grain of the map.

CALCULATIONS

Metrics of Spatial Configuration in an Early-Settlement Landscape

Refer back to Figure 4.1. Recall that the early-settlement landscape is meant to represent an area which was formerly fully forested, but where some of the land has been converted for agricultural and urban use.

Calculation 7: Using the 4-neighbor rule, calculate the total number of patches for each cover type in the early-settlement landscape. Enter your results in Table 4.4.

Table 4.4 Number of patches and mean patch size (in grid cells) using the 4-neighbor rule for categories in the **early-settlement** landscape

Cover type	Number of patches	Mean patch size
Forested		
Agricultural		
Urban		

Calculation 8: Using the 4-neighbor rule, calculate the mean patch size for each cover type in the early-settlement landscape. Enter your results in Table 4.4.

Calculation 9: Calculate the number of edges for each category in the early-settlement landscape of Figure 4.1. Be sure to count both horizontal and vertical edges between cover types. This count is done for cells (not patches), and you may find it useful to mark edges in pencil in your lab manual as you count. Do not count the borders of the map for this exercise. Enter your results in Table 4.5.

Table 4.5 Number of edges and edge:area ratio for the **early-settlement** landscape

Cover type	Number of edges	Edge:area ratio
Forested		
Agricultural		
Urban		

Calculation 10: Using the results from Calculation 9, compute the edge:area ratio for each cover type and enter into Table 4.5.

Q10 What characteristics of a landscape will influence the result you obtain for the number of patches and the average patch size?

Probability of adjacency ($q_{i,j}$) is the probability that a grid cell of cover type *i* is adjacent to a cell of cover type *j*. This metric is sensitive to the fine-scale spatial distribution of cover types and can be computed as:

$$q_{i,j} = \frac{n_{i,j}}{n_i}$$

where $\boldsymbol{n_{i,j}}$ = the number of adjacencies between grid cells of cover type *i* and cover type *j*, and $\boldsymbol{n_i}$ = the total number of adjacencies for cover type *i*.

Probabilities of adjacency are often reported in an *S* x *S* matrix referred to as the **Q matrix**. Because they are probabilities, values for $q_{i,j}$ range from 0 to 1. High $q_{i,j}$ values indicate that the cells of cover type *i* have a high probability of being adjacent to cells of cover type *j*, while low $q_{i,j}$ values indicate a low probability. Values along the diagonals of the Q matrix (the $q_{i,i}$ values) are useful measures of the degree of clumping found *within* each cover type. High $q_{i,i}$ values indicate a highly aggregated, clumpy cover type, and low $q_{i,i}$ values indicate that the cover type tends to occur in isolated, dispersed grid cells or small patches.

The calculation of probabilities of adjacency may be performed in only the horizontal or only the vertical direction to detect directionality (referred to as anisotropy) in a pattern. For example, imagine a landscape composed of alternating ridges and valleys oriented in a north south direction and in which forest cover occupies the ridges and agriculture occupies the valleys. The probabilities of adjacency would be different depending on whether you moved from north to south or from east to west across this landscape. In this lab, the horizontal and vertical values are averaged into a single measure of adjacency.

Contagion (C) (O'Neill et al. 1988; Li and Reynolds 1993, 1994) uses the **Q matrix** values to compute an index of the overall degree of clumping in the landscape. Just as *D* and *SHEI* used all p_i values for all cover types to compute one metric, contagion incorporates all $q_{i,j}$ values into one metric for the entire landscape. The Contagion metric is intended to capture relatively fine-scale differences in pattern that relate to the "texture" or "graininess" of the map. The equation is given by:

$$1 + \frac{\sum_i \sum_j \left[\left(p_i * q_{i,j} \right) * \ln \left(p_i * q_{i,j} \right) \right]}{C_{\max}}$$

where $\boldsymbol{q_{i,j}}$ = the adjacency probabilities defined above, and $C_{\max} = 2 * \ln(S)$, which gives the maximum value of the index for a landscape with *S* cover types.

Values for Contagion range from 0 to 1. A high Contagion value indicates generally clumped patterns of landscape categories within the image, while values near 0 indicate a landscape with a dispersed pattern of landscape categories. Note that Contagion can be computed differently if the $q_{i,j}$ probabilities are computed by another algorithm (Li and Reynolds 1993; Riitters et al. 1996). Because the Contagion metric is computationally intensive, for this exercise it would be tedious to determine this value by hand for even a relatively tiny landscape like the early-settlement landscape. Thus, for illustration purposes, you will compute the Contagion value for only a subset of that landscape.

CALCULATIONS

Metrics of Spatial Configuration in an Early-Settlement Landscape (Continued)

Calculation 11: To begin calculating Contagion, use Figure 4.2 to calculate the proportions occupied by each of the three land cover types in the *subset* of the early-settlement landscape. Record the values in Table 4.6.

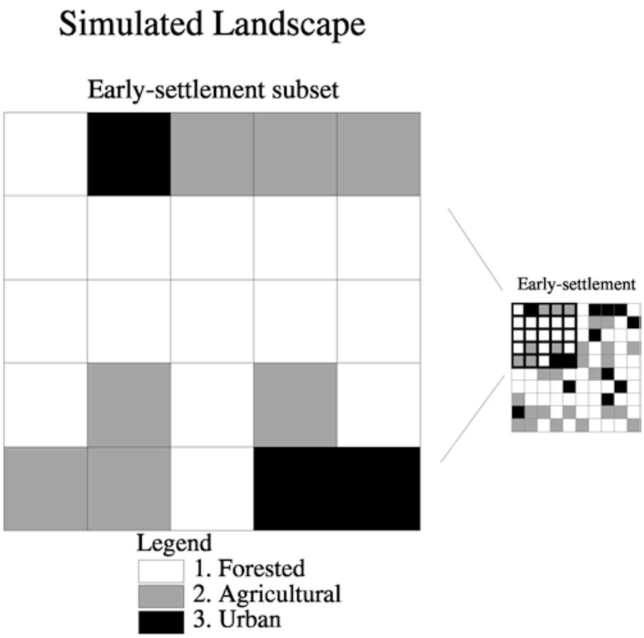


Figure 4.2 **Subset** of the early-settlement landscape used for calculating the Contagion index

Table 4.6 Proportion of the landscape occupied by three different cover types in the **subset** of the early-settlement landscape

Cover type	Proportion (p_i)
Forested	
Agricultural	
Urban	

Calculation 12: Count the adjacencies for all cover types for the *subset* of the early-settlement landscape, as seen in Figure 4.2. Enter the results in Table 4.7. Do not count the borders of the map for this exercise. (*HINT:* If you mark each adjacency once as it is counted, you will mark 40 adjacencies)

Table 4.7 Adjacency counts for the **subset** of the early-settlement landscape

	Category *j*:		
Category *i*:	Forested	Agricultural	Urban
Forested			
Agricultural			
Urban			

Calculation 13: Note the values along the diagonal in Table 4.7. In effect, we have counted most, though not all, of the adjacencies twice. In particular, diagonal elements, which represent adjacencies between cells of the same type, have been counted only once. So that each adjacency is counted the same number of times, double the values from the diagonal elements of Table 4.7 and enter them in Table 4.8, the **N** matrix. For the non-diagonal elements of Table 4.8, use the same value seen in Table 4.7.

Table 4.8 **N** matrix for the **subset** of the early-settlement landscape

	Category *j*:			Row total (n_i)
Category *i*:	Forested	Agricultural	Urban	
Forested	$n_{1,1}$	$n_{1,2}$	$n_{1,3}$	
Agricultural	$n_{2,1}$	$n_{2,2}$	$n_{2,3}$	
Urban	$n_{3,1}$	$n_{3,2}$	$n_{3,3}$	

Calculation 14: Use the values of the **N** matrix (Table 4.8) to compute the elements of the **Q** matrix (Table 4.9).

Table 4.9 **Q** matrix for the **subset** of the early-settlement landscape

	Category *j*:		
Category *i*:	Forested	Agricultural	Urban
Forested	$q_{1,1}$	$q_{1,2}$	$q_{1,3}$
Agricultural	$q_{2,1}$	$q_{2,2}$	$q_{2,3}$
Urban	$q_{3,1}$	$q_{3,2}$	$q_{3,3}$

Calculation 15: Calculate the Contagion value for the subset of the early-settlement landscape using the elements of the **Q** matrix.

The Contagion value for the subset of the early-settlement landscape is: ___________

Q11 If you were considering a real landscape, do you think it would be reasonable, in general, to save computer time by calculating the Contagion value for only a subset? What characteristics of a real landscape might inhibit or encourage you to make your decision?

Q12 Imagine a landscape of large extent for which you couldn't easily calculate this metric. If you could partition the landscape into tiles small enough to compute Contagion in each, could you combine the results in each tile to represent Contagion in the entire extent? What would be the conceptual and practical limits to this approach?

Q13 Suppose that you are given the task of describing how a landscape changed between two time periods, t_1 and t_2. The map of the first time period contains five cover types; the map from the second time period contains seven cover types because "forest" in t_2 was mapped in more detail—as deciduous, coniferous, and mixed forest. How should you proceed with your comparison, and why?

SYNTHESIS

Q14 Two landscapes are the same size and both contain the same amount of a given cover type. Landscape A has four patches of that cover type, and Landscape B has 17 patches of the same cover type. Which of the landscapes will have the greater length of edge of that cover type?

Q15 What characteristics of the landscape appear to have influenced the Contagion value calculated in this section? How would you change the values of the grid cells to raise the Contagion value?

Q16 From your set of calculations, do you think after calculating a large number of metrics for a single landscape, additional metrics would provide little new information? How might you attempt to objectively determine an upper limit to the number of useful metrics?

Part 3. Using Fragstats for Automated Landscape Metric Calculation for the Early- and Post-settlement Landscapes

In this section, you will use Fragstats (McGarigal et al. 2012) to analyze the landscapes you examined in Parts 1 and 2. Fragstats is available for free, computes a wide variety of metrics, is available in versions to analyze both raster and vector maps, and is probably the most widely used program for landscape pattern analysis. Fragstats can be run in a variety of ways, including from a graphical user interface as a stand-alone program, as a plug-in to ArcGIS, and from the command line. Information about Fragstats is available in the student material for the book, or can be provided by your instructor.

INPUT AND SETTINGS

Before calculating a given set of metrics, Fragstats requires settings for the suite of metrics it calculates for your image. Some of the major settings to consider and understand are given below. Each has an impact on how Fragstats interprets the landscape in its calculation of metric values.

- **Grid cell size**: The size of cells for each image is given in each of the calculations for this section.
- **Diagonals in patch finding**: You must specify in Fragstats whether to use the 4-neighbor or 8-neighbor rule for finding patches.
- **Scale of Analysis**: Fragstats can output calculations at the landscape level (i.e., considering all the cover types together), class level (reported by each of the cover types in the map), and patch level (calculated for each patch).

  To complete these sections, we ask you to select the landscape-level and class-level metrics. In this section, we are not interested in knowing details about each patch, but instead are primarily interested in metrics that summarize the entire image. Although we will not directly use the information contained in the summaries of each landscape category, it is useful to note that some metrics can be calculated for each class.

OUTPUT

Fragstats outputs information in several files. In this lab, we are concerned with the **.land** file, a text file that can be viewed with any text editor. Information about each landscape category is at the beginning of the file, and metrics for the entire landscape are at the end of the file. In these landscapes, Category 1 = Forested, Category 2 = Agricultural, and Category 3 = Urban.

CALCULATIONS

You will input text files containing the land-cover categories for the early- and post-settlement landscapes. You will then use Fragstats to specify your output file name and landscape metrics to calculate.

Calculation 16: Early-Settlement Landscape with the 4-Neighbor Rule

- Run Fragstats using the **esett** landscape file and the **4-neighbor** rule. This is a 10×10 landscape where one side of a cell represents 1000 m on the ground. Use **early4** as the base for output file names. You might make a new folder to contain the results. You may choose which metrics to compute, but you should include several of the metrics you calculated by hand (e.g., number of patches, mean patch size, contagion, and Shannon evenness).

- To verify that you are using Fragstats correctly and that your answers calculated by hand were correct, compare the calculations for the early-settlement landscape from the previous section. You should get the same answers (*NOTE:* Fragstats does not calculate Dominance).

Calculation 17: Early-Settlement Landscape with the 8-Neighbor Rule
Run Fragstats using the **8-neighbor** rule for the early-settlement landscape. Again, use the **esett** landscape file. As the base for naming output files, enter **early8**.

Calculation 18: Post-settlement Landscape with the 4-Neighbor Rule
Run Fragstats using the **psett** landscape file. This is a 10×10 landscape where one side of a cell represents 1000 m on the ground. As the base for output files, enter **post4**.

Calculation 19: Post-settlement Landscape with the 8-Neighbor Rule
Run Fragstats using the 8-neighbor rule for the post-settlement landscape. Again, use the **psett** landscape file. As the base for output files, enter **post8**.

Q17 Organize the results obtained for the four runs (early- and post-settlement landscapes, 4- and 8-neighbor rules). Describe how the metrics are affected by the choice of 4- and 8-neighbor rules. Taken as a whole, how do the metrics indicate that this landscape has changed from the early-settlement to post-settlement period?

Part 4. Automated Landscape Metric Calculation for Real Landscapes and Interpretation of Multiple Metrics

In this section, we use Fragstats to compute landscape metrics for real landscapes. Calculate at least the following metrics with Fragstats for each of the maps described below. Use the 8-neighbor rule for each of the analyses.

- Contagion
- Patch Density (the average number of patches per 100 ha)
- Edge Density (an expression of edge:area relationships)
- Landscape Shape Index (a measure of shape complexity)
- Largest Patch Index (an indicator of connectivity)
- Patch Richness (the number of patch types)

CALCULATIONS

Calculation 20: Madison, Wisconsin, USA
We present two classifications of the same satellite image produced by two different users of the same landscape processing software. Subjectivity inherent in the classification process inevitably produces differences among resultant maps. The two

landscapes are referred to **mad1** and **mad2**. Each landscape has 575 rows and 800 columns, and one side of a grid cell represents 30 m on the ground. Comparing the results of these analyses illustrates that differences or errors in classification will influence landscape metrics.

Calculation 21: New England Landscape #1 [Latitude = 40.71754, Longitude = –76.81646]
This landscape is referred to as **x632y165s2** according to its index in the Metaland software (see Chapter 10). This landscape has 216 rows and 216 columns, and one side of a grid cell represents 30 m on the ground.

Calculation 22: New England Landscape #2 [Latitude = 40.77141, Longitude = –75.29400]
This landscape is referred to as **x651y160s2** according to its index in the Metaland software. This landscape has 216 rows and 216 columns, and one side of a grid cell represents 30 m on the ground.

Calculation 23: New England Landscape #3 [Latitude = 41.32851, Longitude = –72.06994]
This landscape is referred to as **x689y141s2** according to its index in the Metaland software. This landscape has 216 rows and 216 columns, and one side of a grid cell represents 30 m on the ground.

SYNTHESIS

Q18 Using your Fragstats results, plot the values of the metrics specified above to assess their relationships. For each pair of metrics, graph a scatter plot (metric a on the *Y*-axis, metric b on the *X*-axis); your plots will have five points, one for each landscape.

Q19 To compare metrics across the five landscapes, you can make a bar graph with metric values on the *Y*-axis and each landscape map on the *X*-axis. When looking at the maps and the metrics, which of the landscapes above appears to be the most fragmented, and which appears least fragmented? How did you determine this? Use the results of your quantitative analyses to support your interpretations.

Q20 How would the correlation among landscape metrics influence your choice of what to report in an analysis that describes landscape pattern or quantifies differences between two landscapes or changes in a single landscape through time?

Q21 What criteria would you use to select the “best” set of metrics to describe a landscape?

Part 5. Understanding Landscape Change Through Metrics

In this section, you will explore some of the challenges of using landscape metrics to assess landscape change through time. While it is easy to generate large amounts of data quantifying the landscape patterns of a given area, it is quite challenging to make credible comparisons across time periods. Data sets of the same area for two time periods are often produced with different classification techniques and philosophies, which may make comparisons challenging, at least for some metrics.

You will draw on what you have learned in the previous sections: for example, interpreting and reflecting on the equations that are used to calculate landscape metrics; exploring how some landscape metrics respond principally to landscape composition, while others are more clearly responsive to a landscape's configuration. You will study four landscapes from the **National Land Cover Data Set (NLCD)**, a continental-scale land-cover assessment program using satellite data and ancillary information to track and update land-cover change and stability through time (Vogelmann et al. 2001; Homer et al. 2004; Jin et al. 2013). The landscapes are taken from 6.5 km × 6.5 km regions in New England, USA. For each landscape at multiple times, you will be given the land-cover images, and the values of a large number of landscape-level and class-level metrics from Fragstats runs.

After you have downloaded the data, investigate by viewing the images of the same landscapes at different times and by exploring the landscape metric data using the associated **sandbox** spreadsheet. The spreadsheet allows you to quickly collate the output from multiple runs of Fragstats.

SYNTHESIS

Q22 What are some of the practical obstacles to comparing the landscapes from these two time periods? In your estimation, to what extent are differences in landscape metrics likely driven by differences in landscape data and data-processing approaches, rather than in true changes in the real world?

Q23 Choose two cover types and compare the class-level metrics in these landscapes across time periods. Are some land-cover classes more readily comparable than others? If so, which ones?

Q24 According to the computed landscape metric values, which landscapes have changed the most between the two time periods? Which have changed the least? For your analyses, you should include both landscape-level and class-level metrics, which may be more informative in answering particular questions than those that are computed for all cover types simultaneously.

Q25 Given your experience in this chapter, how well (or poorly) do landscape metric values support your subjective assessment of land-cover change and stability in these real-world landscapes?

REFERENCES AND RECOMMENDED READINGS[1]

Baker WL, Cai Y (1992) The r.le programs for multi-scale analysis of landscape structure using the GRASS geographical information system. Landsc Ecol 7:291–302

Boström C, Pittman SJ, Simenstad C, Kneib RT (2011) Seascape ecology of coastal biogenic habitats: advances, gaps, and challenges. Mar Ecol Prog Ser 427:191–217

*Burnicki AC (2012) Impact of error on landscape pattern analyses performed on land-cover change maps. Landsc Ecol 27:713–729. *Accuracy of the data used in any landscape analysis will influence the results, and this is especially important when you want to use metrics to quantify how landscapes change over time.*

Cardille JA, Lambois M (2010) From the redwood forest to the Gulf Stream waters: human signature nearly ubiquitous in representative US landscapes. Front Ecol Environ 8:130–134

Corry RC, Nassauer JI (2005) Limitations of using landscape pattern indices to evaluate the ecological consequences of alternative plans and designs. Landsc Urban Plann 72:265–280

*Cushman SA, McGarigal K, Neel MC (2008) Parsimony in landscape metrics: strength, universality and consistency. Ecol Indicators 8:691–703. *Many landscape metrics are correlated with one another, and this paper emphasizes the unique contributions of metrics that are independent of one another and associated with qualitatively different aspects of pattern.*

Fortin M-J, Boots B, Csillag F, Remmel TK (2003) On the role of spatial stochastic models in understanding landscape indices. Oikos 102:203–212

Gardner RH, Milne BT, Turner MG et al (1987) Neutral models for the analysis of broad-scale landscape pattern. Landsc Ecol 1:19–28

*Gustafson EJ (1998) Quantifying landscape spatial pattern: what is the state of the art? Ecosystems 1:143–156. *A synthetic overview of the ways in which landscape pattern is quantified, this classic paper emphasizes conceptual issues and distinguishes between metrics, including patch-based metrics, calculated from categorical data and approaches from spatial statistics.*

Gustafson EJ, Parker GR (1992) Relationships between landcover proportion and indices of landscape spatial pattern. Landsc Ecol 7:101–110

*Haines-Young R, Chopping M (1996) Quantifying landscape structure: a review of landscape indices and their application to forested landscapes. Progr Phys Geogr 20:418–445. *A good review that includes examples of how different landscape metrics are used in questions associated with forested landscapes.*

Homer C, Huang C, Yang L et al (2004) Development of a 2001 National Land-Cover Database for the United States. Photogramm Eng Remote Sens 70(7):829–840

Jin S, Yang L, Danielson P et al (2013) A comprehensive change detection method for updating the national land cover database to circa 2011. Remote Sens Environ 132:159–175

*Li H, Reynolds JF (1995) On definition and quantification of heterogeneity. Oikos 73:280–284. *An excellent discussion of what is meant by heterogeneity. This seminal paper emphasizes understanding what is being quantified and should be read by all those beginning to consider the causes or consequences of spatial pattern.*

Li H, Reynolds JF (1993) A new contagion index to quantify spatial patterns of landscapes. Landsc Ecol 8:155–162

Li H, Reynolds JF (1994) A simulation experiment to quantify spatial heterogeneity in categorical maps. Ecology 75:2446–2455

*Li H, Wu J (2004) Use and misuse of landscape indices. Landsc Ecol 19:389–399. *Useful synthesis of issues associated with quantifying landscape patterns.*

McGarigal K, Cushman SA, Ene E (2012) FRAGSTATS v4: spatial pattern analysis program for categorical and continuous maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. http://www.umass.edu/landeco/research/fragstats/fragstats.html

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

McGarigal K, Marks BJ (1993) FRAGSTATS. Spatial analysis program for quantifying landscape structure. USDA Forest Service General Technical Report PNW-GTR-351

O'Neill RV, Krummel JR, Gardner RH et al (1988) Indices of landscape pattern. Landsc Ecol 1:153–162

Pielou EC (1975) *Ecological diversity*. Wiley-Interscience, New York

Remmel TK, Csillag F (2003) When are two landscape pattern indices significantly different? J Geograph Syst 5:331–351

Riitters KH, O'Neill RV, Wickham JD et al (1996) A note on contagion indices for landscape analysis. Landsc Ecol 11:197–202

Romme WH (1982) Fire and landscape diversity in subalpine forests of Yellowstone National Park. Ecol Monogr 52:199–221

Romme WH, Knight DH (1982) Landscape diversity: the concept applied to Yellowstone Park. BioScience 32:664–670

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423. (July), with corrections pp. 623–656 (October)

*Símovsá P, Gdulová K (2012) Landscape indices behavior: A review of scale effects. Appl Geogr 34:385–394. *There are a fair number of empirical papers documenting the consequences of changing grain and extent on landscape metrics, and this paper provides an overview.*

Teixido N, Garrabou J, Gutt J et al (2007) Iceberg disturbance and successional spatial patterns: the case of the shelf Antarctic benthic communities. Ecosystems 10:142–157

Turner MG (2005) Landscape ecology: what is the state of the science? Ann Rev Ecol Evol Syst 36:319–344

*Turner MG, Gardner RH (2015) Chapter 4, Landscape metrics. In: *Landscape ecology in theory and practice*. Springer, New York, pp 97–142. *We highly recommend reading this chapter from the landscape ecology text, as it provides an introduction and overview of why and how to use landscape metrics in spatial pattern analysis.*

Turner MG, Costanza R, Sklar FH (1989) Methods to compare spatial patterns for landscape modeling and analysis. Ecol Model 48:1–18

Vogelmann JE, Howard SM, Yang L et al (2001) Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogramm Eng Remote Sens 67(6):650–662

Wickham JD, Norton DJ (1994) Mapping and analyzing landscape patterns. Landsc Ecol 9:7–23

Chapter 5
Scale Detection Using Semivariograms and Autocorrelograms

Michael W. Palmer and Daniel J. McGlinn

OBJECTIVES

The evolution and ecology of all organisms are contingent on the complex variation seen in nature. Landscape ecology differs from most other branches of ecology in that it explicitly involves spatial variation. Therefore, one of the goals of landscape ecology is to describe spatial variation. The purpose of this exercise is to:

1. Introduce two tools for describing this variation: semivariance and autocorrelation; and
2. Give students experience creating and interpreting semivariograms and autocorrelograms.

In this lab, you will collect field data from quadrats arranged along a transect (or alternatively, you will use supplied data). You will then calculate and graph semivariograms and autocorrelograms using a spreadsheet, and you will use these graphs to determine how spatial patterns vary as a function of scale in your system. For this lab, you will need access to a spreadsheet program (such as Excel) and the file **vario.xlsx** provided on the book's website. If you choose the fieldwork option, you will also need two 100-m measuring tapes and one 1 × 1-m sampling quadrat. We've also provided code (see book's website) if you'd like to try the lab using R software.

M.W. Palmer (✉)
Department of Plant Biology, Ecology and Evolution,
Oklahoma State University, Stillwater, OK, USA
e-mail: mike.palmer@okstate.edu

D.J. McGlinn
College of Charleston, Charleston, SC, USA

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_5

INTRODUCTION

Nature is intrinsically variable, and the evolution and ecology of all organisms are contingent on such variation. Landscape ecology is concerned not only with the magnitude of this variation, but also with its geometry. Most patterns in nature are far more complex than the simple polygons and curves of Euclidean geometry. For example, forest edges are rarely straight lines, animal home ranges are not rectangles, and trees are not cones. Therefore, we need special methods to describe the shape of nature.

The discipline of spatial statistics has diversified and matured (see Cressie 1991; Bailey and Gatrell 1995), and it is not possible here to give a full summary of the wealth of methods available. Instead, the purpose of this exercise is to describe two different methods for characterizing variation in a variable as a function of position in the landscape. This variable could be a soil nutrient, a measure of vegetation height, an index of species composition, or anything else of interest. In spatial statistics, we term variables with known locations **regionalized variables**, and we label them z (so as not to confuse them with x and y, typically reserved for the spatial coordinates, or for independent and dependent variables, respectively).

The two methods covered in this exercise are variography, which is part of the discipline of geostatistics (see Isaaks and Srivastava 1989), and autocorrelation, which is derived from the familiar correlation coefficient (Sokal and Rohlf 1981). Recall that the **correlation coefficient**, ***r***, is a number that varies between −1 and +1 and reveals the nature of the relationship between two variables. It is close to −1 for two variables that are strongly negatively related, close to 0 for unrelated variables, and close to +1 for positively related variables. In contrast, variography is derived from the *variance*, which must be a positive number but can otherwise take any value.

One of the most important properties of almost all regionalized variables is **spatial dependence**. Spatial dependence (as assessed by **spatial autocorrelation**, or the tendency of a random variable to be correlated with itself at finite distances) means that a variable measured at one location *depends*, in one way or another, on the same variable measured at a different location. Spatial dependence arises for a number of different reasons, but let us consider two examples.

If you examine mean annual temperature as a function of position on the globe, you will note that (with many important and interesting exceptions) there is a gradient from warm temperatures at the equator to cool temperatures at the poles. If you have two sites that are almost at the same latitude, they will have similar temperatures. On the other hand, two sites that are on different latitudes will have different temperatures, and the amount of the difference in temperature will be positively (and gradually) related to the difference in latitude. **Spatial dependence** occurs when information available at one location allows you to infer information about the other location.

Another example is in a savanna landscape where widely spaced trees provide islands of shade in an otherwise sunny landscape. Two sites that are centimeters apart are likely to have a similar amount of sunshine. However, two sites that are several meters apart may, or may not, have similar amounts of sunshine—a lot depends on the size and spacing of trees. If the sites are hundreds of meters apart, you may not be able to predict the sunlight regime very well. So in this case, we have spatial dependence at fine scales, but not necessarily at coarse scales. Also, unlike the example of global temperature, our regionalized variable consists of fairly discrete *patches* of sun and shade.

The first column of graphs in Figure 5.1 displays a variety of made-up regionalized variables with identical means and variances. These hypothetical variables have been constructed to illustrate the diversity of patterns that could potentially be found in nature. Note that regionalized variables can consist of a variety of features such as patches (i.e., homogeneous regions), noise (random, independent variation), random walks (a random walk is when a value at a given location equals the value at an adjacent location, plus or minus a small random number), or some combination of these. Also, note that the different variables behave differently as a function of scale. For example, patches can be large, small, or intermediate. Stretches of linear behavior can also be large, small, or intermediate. Also, noise can operate at any scale. If the graphs in Figure 5.1 were based on real data, we would seek biological explanations for the different scales. Such explanations might involve the size and shape of underlying geomorphology, the average size of plant clones, the average size of a natural disturbance, the home range size of the dominant mammal species, or the average farm size.

Except for variable A in Figure 5.1, there is some spatial dependence. That is, nearby locations are, *on average*, more similar than distant locations. Since similarity typically decreases as a function of distance of separation, we also call this phenomenon **distance decay**. Distance decay has important consequences for living things. For example, if soil conditions are very similar at nearby locations (as for variables C and F in Figure 5.1), then natural selection *might* favor plants with short dispersal distances. If, on the other hand, soil conditions were spatially unpredictable (as for variable A), a long dispersal distance *might* be advantageous. Similarly, the foraging behavior of animals, the growth of plant roots, the spread of fire, the flow of water, and the behavior of many other ecological phenomena all depend on the nature of distance decay in environmental factors.

In statistics, spatial dependence has both desirable and undesirable attributes (Legendre 1993). It means that one can predict variables (to some degree) based on geographic location, which can aid in mapping the environment. However, spatial dependence also violates the standard statistical assumption of independent observations (even if samples are randomly located). Thus, unless specifically corrected for, many statistical methods are invalid if your data exhibit distance decay. Fortunately, there are tools to evaluate the *degree* and the *scales* of spatial dependence. The two tools we introduce in this laboratory exercise are the semivariogram and the autocorrelogram.

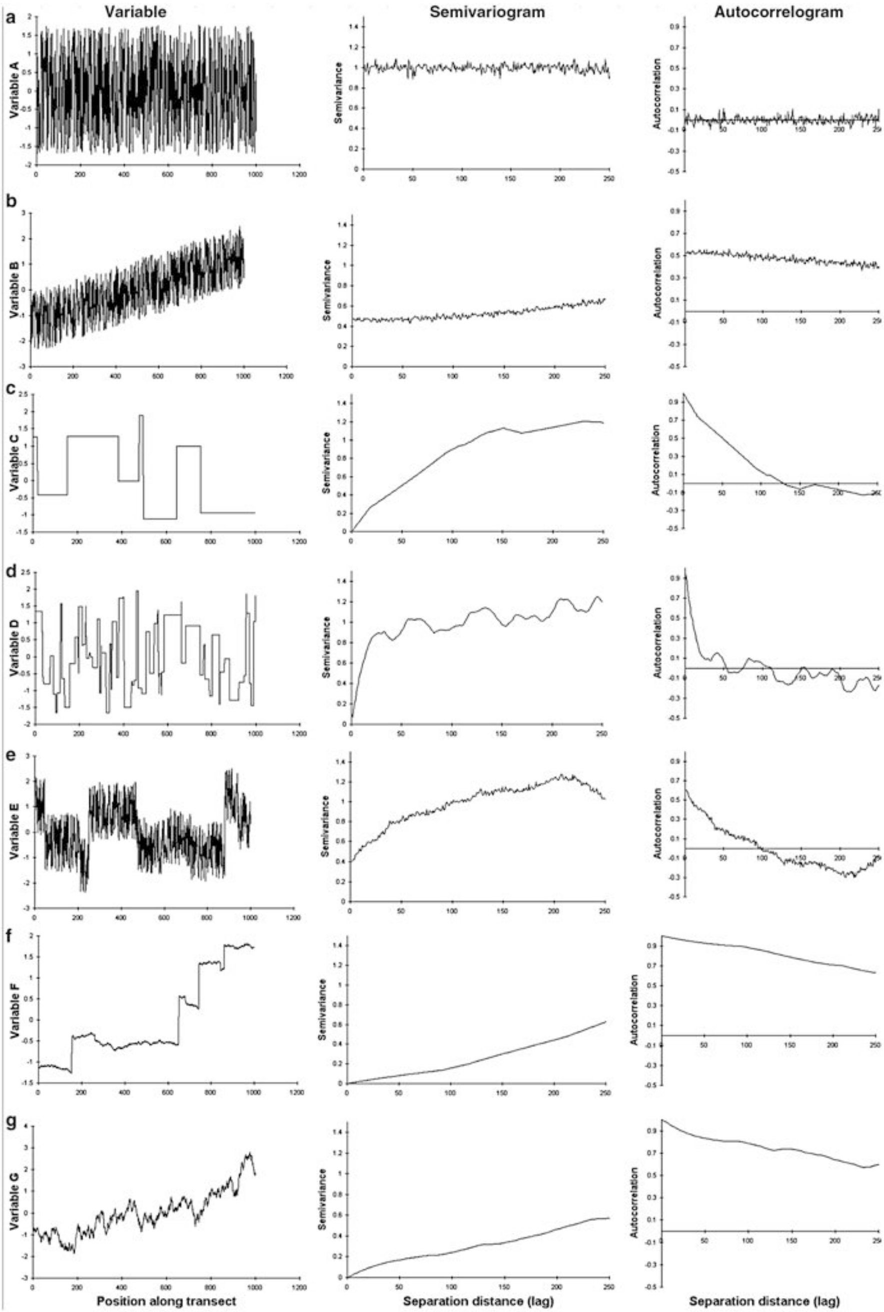


Figure 5.1 Seven artificial regionalized variables (column 1) as a function of position along a transect, along with their corresponding semivariograms (column 2) and autocorrelograms (column 3). All variables have identical means and variances. The variables can be described as follows: (**a**) pure noise; (**b**) fine-scale noise superimposed on a linear trend; (**c**) large patches; (**d**) small patches; (**e**) noise superimposed on large patches; (**f**) patches with "drift" in their mean values, plus fine-scale noise; (**g**) random walk

Variography

Variography is the discipline of using semivariograms (and related graphs such as covariograms) to uncover the degree to which the variance in a regionalized variable depends on distance (Rossi et al. 1992). The geographic distance between two samples is termed the **spatial lag**. Recall that the variance is the square of the standard deviation and is a measure of the spread or variation of data. The word **semivariance** is derived from "half of the variance," and indeed it is a measure of the variance of the regionalized variable, *z*. But what is special about the semivariance is that it changes as a function of distance. The semivariance is computed as follows:

$$\gamma(h) = \left\{ \Sigma \left[z(i) - z(i+h) \right]^2 \right\} / 2N(h)$$

where $\gamma(h)$ is the semivariance of a **lag** of distance h, $z(i)$ is the value of a regionalized variable z at location i, $z(i+h)$ is the value of z at a location separated from i by lag h, and $N(h)$ is the number of pairs of points separated by lag h. The summation is over all pairs of points separated by distance h. In plain English, the semivariance is half of the average squared difference of all pairs of points separated by a given distance. A semivariogram is a plot of semivariance versus the lag distance. As with the variance, the semivariance cannot be less than zero, but it is not bounded on the top.

An idealized, hypothetical semivariogram is given in Figure 5.2. Since the semivariance is directly related to variance, a high value indicates high variation, and a low value indicates low variation. Almost always, variance increases as a function of lag

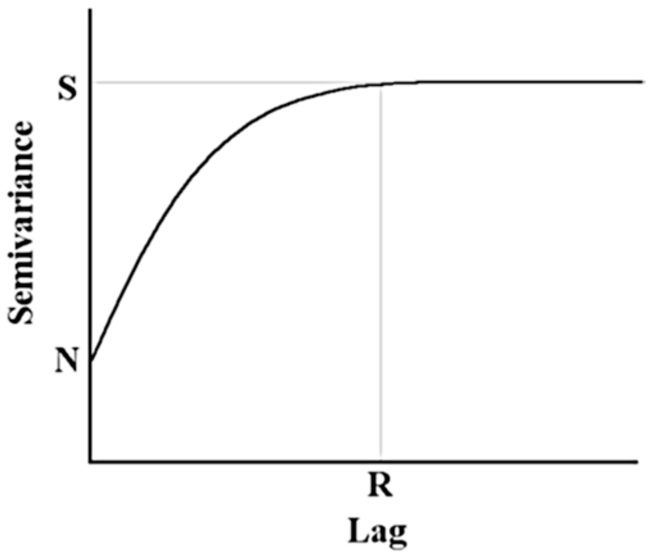


Figure 5.2 An idealized semivariogram. N = nugget, R = range, S = sill

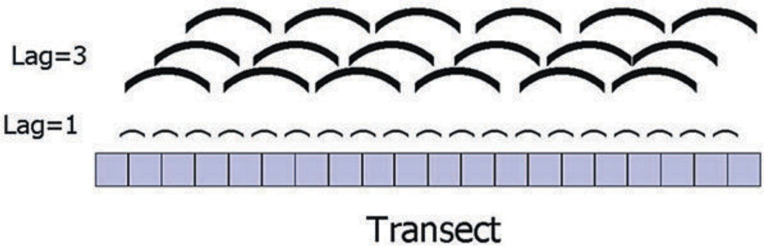


Figure 5.3 A 20-quadrat-long transect illustrating all of the pairs of points separated by two selected lags: 1 and 3

distance. In other words, the larger the area you study, the more variable your conditions are. It is important to reiterate that the lag distance is not the same as the distance from the origin or starting point. Rather, it is calculated for all pairs of points (Figure 5.3).

At distances less than *R* (the **range**), we have spatial dependence (Figure 5.2). That is, closer samples are more similar than distant samples. At distances of at least *h*, we have spatial independence; therefore, samples separated by longer distances would be valid for conventional statistics. Any area with linear dimensions of at least *R* would have as much variance as the landscape as a whole. A horizontal asymptote at distances greater than *R* is known as the **sill**. The sill indicates the amount of "background" variation.

For a very smooth, regionalized variable, two samples that are infinitesimally close to each other may have almost identical values. Elevation of the ground surface almost always behaves this way. However, most variables are not so smooth, if for no other reason than measurement error. Such unresolved variation at very fine scales is termed the **nugget effect**, and is indicated by *N* (Figure 5.2). The term derives from the original use of variography in gold mining: at fine spatial scales, you either find the gold nugget or miss it. Soil variables such as pH or nutrient concentrations typically have very high nugget effects.

The semivariogram specifically address how variance increases as a function of scale. Although it can only describe *patterns*, we often hope to infer the *processes* that generate such patterns. If we find a distinct range, or even a pronounced inflection in the semivariogram, we suspect that there are different processes operating at different scales. For example, if we discover that the range equals approximately 10 m, we need to seek an underlying process that operates on a scale of 10 m. In a forest, this scale could represent the average canopy gap size or the average size of a canopy tree crown. In the arctic tundra, the range could represent the average size of a permafrost polygon. Of course, you are never guaranteed to actually find a range. It is possible (and indeed, likely) that variation in nature increases continuously as a function of scale.

The second column of Figure 5.1 shows the semivariograms for the hypothetical regionalized variables. Since very few pairs of points represent very far distances, it is usually not advisable to plot $\gamma(h)$ for large *h*. A general rule of thumb (adopted here) is to plot only up to half of the maximum distance between samples. The bumps and wiggles at such far distances in the semivariograms are due to chance variation in the data, and not to the underlying process generating the patterns. Note that only three of the variables (C, D, and E) have semivariograms remotely resembling those in Figure 5.2. The range of variable C (≈150 units) is much larger than the range of variable D (≈25 units), which reflects the differences in patch size. Variable E seems to have an inflection at 50 m, which marks the difference between noise within patches, and the differences between patches.

The semivariograms of variables B, F, and G are continuously increasing functions. Therefore, there is spatial dependence at broad spatial scales. Three variables (A, B, and E) have much fine-scale noise, and hence a substantial nugget effect. Note that variable F has very little fine-scale noise, and hence has a negligible nugget. Variable A represents pure noise and hence pure spatial independence. For such variables, the nugget equals the sill.

Autocorrelation

Autocorrelograms are plots of the correlation coefficient, r, as a function of lag:

$$r(h) = \mathrm{corr}\left[z(i), z(i+h)\right]$$

It is called "auto"-correlation because the variable is correlated with itself. Autocorrelation can take values from −1 to +1 although for most applications positive values are most common. In situations with distance decay, autocorrelograms are declining functions and often look like upside-down semivariograms (third column of Figure 5.1). If there is little fine-scale noise, the y-intercept will be close to 1.0. In situations in which the semivariogram displays a nugget effect, the y-intercept of the autocorrelogram will be less than 1. An autocorrelation of 0 means there is no spatial predictability; this is related to the concept of the sill.

This describes only the simplest kind of autocorrelogram. More complex (and usually more appropriate) ways to calculate autocorrelograms, as well as testing their statistical significance, are described by Legendre and Fortin (1989), Bailey and Gatrell (1995), and Legendre and Legendre (1998). Autocorrelograms are also used in the analysis of change through time (also known as "time series").

Comparing Autocorrelation to Semivariance

The interpretation of autocorrelograms is very similar to that of semivariograms, so the choice between them is largely a matter of taste. Since the correlation coefficient is a dimensionless number (i.e., it is standardized), autocorrelograms are useful in comparing variables with different units (e.g., plant density and soil calcium). Semivariance has a dimension of units squared (so if the regionalized variable is in parts per million or ppm, semivariance is in ppm^2). Thus, it is useful in comparing different commensurate variables or (more commonly) the same variable in different locations. However, semivariance can be standardized for comparing variables measured in different units (see Rossi et al. 1992).

Since it is derived from the correlation coefficient, autocorrelation is closely related to classical statistical theory. Variography, on the other hand, is a branch of geostatistics. This discipline was largely developed for the mining industry to help predict the locations of mineral deposits. Variography is a precursor to geostatistical interpolation (for mapping) or "kriging" (see Isaaks and Srivastava 1989) and to fractal geometry (Burrough 1983; Palmer 1988).

EXERCISES

EXERCISE 1: Data Collection

Option 1: Field Exercise Using Vegetation Height

1. Choose a field site in which the maximum height of the vegetation is about 2.5 m or less, and in which it is possible to fit a 200-m transect. If the site is large enough, randomly choose a starting location and compass direction. If it is too small for this, choose an appropriate direction but randomize the starting point, so you are not biased by particular plants.
2. Extend two (or more) 100-m tapes end to end along the chosen compass direction. Ideally, you would do this with a surveyor's compass or level. It is crucial that the transect be as straight as possible and not influenced by the vegetation!
3. Beginning at 0 m, establish a 1×1-m quadrat (this can consist of three meter sticks plus the meter tape as the fourth boundary).
4. Within this plot, measure and record the height of the tallest plant.
5. Now repeat the process with an adjacent plot at 1 m, then at 2 m, and continue to the end of the transect.

If you do not have the luxury of a large enough field site, it is possible to perform this exercise with smaller contiguous quadrats. Also, the regionalized variable need not be height: you can perform this same exercise using stem density, biomass, species richness, ordination scores, percent cover of bare soil, elevation, percent sunlight, soil parameters, and more. It may be possible to derive a regionalized variable from a map or a remotely sensed image, but be aware that data on such images *may* have already been "smoothed" or interpreted for ease of display, and hence your analyses would be inappropriate. Regardless of the overall length and quadrat size, try to have at least 200 quadrats in your transect (spatial analyses typically require large sample sizes). Another option is to split the class up into two or more groups, with each group studying either a different regionalized variable along the same transect, or the same variable in a different vegetation type.

Option 2: Using Provided Data Sets

Some example data sets are provided on the book's website, in case there are no opportunities for collecting new data. The file is entitled **vario.xlsx** and contains a worksheet with three different **example data sets**.

EXERCISE 2: Data Analysis

You will analyze the data using a spreadsheet. The example given here is for Microsoft Excel, but similar commands exist in other spreadsheets. Before beginning this exercise, review absolute and relative cell references, how to graph data, as well as the following Excel functions: **OFFSET**, **SUMXMY2**, **CORREL**. Make sure that automatic calculation of formulas is in effect (this is the most usual default; it means that the results will be continuously updated. Check under **File - Options - Calculation - Automatic**).

1. Enter your data in the blank worksheet labeled **Vegetation Heights**. The following description assumes that the transect is 200 quadrats long; if not, substitute "200" with the correct number.
2. In row 1, label columns **A–F** as follows:

A	B	C	D	E	F
POSITION	VALUE	LAG	SEMIVARIANCE	LAG	AUTOCORRELATION

3. In column **A**, fill rows **2–201** with the numbers 1–200.
   (*HINT:* One quick way to do this is to put "`1`" in cell **A2**, and then put the formula: "`=A2+1`" in cell **A3**. Then copy the contents of **A3** and paste them into cells **A4–A201**. Since there were no dollar signs ($) in the original formula, the cell reference of **A2** is copied as a relative location. Therefore, each one of the cells will equal the cell above it plus 1).
4. In column **B**, fill rows **2–201** with the data you collected (or copied from the provided data sets) in the correct spatial sequence.
5. In column **C**, fill rows **2–101** with the numbers 1–100. (Recall the rule of thumb that it is best not to plot semivariograms for more than half of the maximum lag distance). Repeat this for column **E**.

Two different ways to calculate the semivariance follow. Method 1 is conceptually easier, but method 2 is less labor intensive. Therefore, read and understand method 1, but use method 2. It is important to keep in mind that in a transect Q units long, the number of pairs of quadrats separated by a given lag distance h will equal $Q-h$. In our example, there are 199 pairs of quadrats separated by 1 m, 198 separated by 2 m, and 1 pair separated by 199 m (i.e., the first quadrat and the last quadrat). Before continuing, review the equation for semivariance.

Semivariance Method 1: Read and understand this method

1. In cell **D2**, put the formula:
   "`=SUMXMY2(B2:B200,B3:B201)/(2*(200-1))`"
   The formula SUMXMY2 means "sum of (*x* minus *y*) squared". The two selected blocks (**B2:B200**) and (**B3:B201**) are actually the same data, but shifted by a lag of 1 unit. The denominator is two times the number of pairs of points separated by distance *h*. It is, of course, possible to put the number 198 in the denominator, but writing the formula out often helps with troubleshooting.
2. In cell **D3**, write the formula:
   "`=SUMXMY2(B2:B199,B4:B201)/(2*(200-2))`"
   This is the semivariance for a lag of 2. The formula for **D4** should be:
   "`=SUMXMY2(B2:B198,B5:B201)/(2*(200-3))`"
3. Continue filling in column **D** until you reach a lag of 100.

Semivariance Method 2: Use this method

1. Instead of typing in a unique formula for each cell of column **D**, it is more time-efficient to type in a generic formula in cell **D2:**
   "`=SUMXMY2(B$2:OFFSET(B$2,200-C2-1,0),OFFSET(B$2,C2,0)`
   `:B$201)/(2*(200-C2))`"

NOTE: This is precisely the same formula as in method 1, except for how we specify addresses. The dollar signs ($) before the row means a reference to that exact row, no matter where you copy and paste the formula. **OFFSET** returns a new cell address and has three arguments: a cell address, the number of rows of separation, and the number of columns of separation. Therefore, the block B$2:OFFSET(B$2,200–C2–1,0) refers to a column of data beginning at cell B2 and ending (199–C2) cells below B2. Since cell C2 indicates a lag of 1, the column of data will be the same as B2:B200, as desired. The second block, OFFSET(B$2,C2,0):B$201, means a block beginning at C2 below B2, and ending at B201. This will be the same as B3:B201. The denominator of the equation will equal 2*(200–1).

2. Copy cell **D2** and paste it into cells **D3–D101.** Note that when you do so, the formula remains identical in all cells *except* that the reference to the lag, column **C,** changes. Thus, the formula will return the semivariance for whatever lag is indicated in the same row of column **C**.

You can calculate autocorrelation by similar methods to those described earlier for semivariance.

Autocorrelation Method 1: Read and understand this method

1. In cell **F2,** type: "`=CORREL(B2:B200,B3:B201)`". This will return the correlation between the variable and itself, with a lag of 1.
2. In cell **F3**, type "`=CORREL(B2:B199,B4:B201)`" for a lag of 2, and continue filling column **F** until lag 100.

Autocorrelation Method 2: Use this method

1. Following the same reasoning as method 2 for the semivariance, type the following in cell **F2:**
"`=CORREL(B$2:OFFSET(B$2,200-E2-1,0),OFFSET(B$2,E2,0):B$201)`"
2. Copy this formula and paste it into cells **F3–F101.**

Before proceeding further, make sure to save your results.

EXERCISE 3: Results

Using your spreadsheet program, create the following plots:

1. Vegetation height as a function of transect position
2. Create a semivariogram as follows:
   a. Plot the semivariance as a function of lag. The data will be in the block **C2:D101**.
   b. Label the *X*-axis "Lag (meters)", and the *Y*-axis "Semivariance".
   c. Drag the graph immediately under the graph of the raw data.

HINT: First, make an X,Y (scatter) plot under **Insert - Charts - Scatter**. Then, double-click on any point in the graph and select **Format Data Series**. In Excel 2013 onwards, select the icon resembling a paint can (aka Fill & Line) and choose **Solid Line**. In older versions of Excel, select **Patterns - Line - Automatic**.

3. Create an autocorrelogram as follows:
   a. Graph the data in **E2:F101** and drag the graph under the semivariogram.
   b. Label the axes appropriately.

Interpretations and Rules of Thumb

As with any bivariate (two-variable) graph, the scaling of the *Y*-axis relative to the *X*-axis should not affect our interpretation, but it often does. A short, long graph often appears less "noisy" than a tall, narrow one. It is generally best to choose a scaling relatively close to 1:1 (that is, square), or at most 1.5:1 or 1:1.5. Of course, there may be exceptions (e.g., if one wants to display the results of numerous transects, one graph on top of the other, it might be useful to have them short and long).

For semivariograms, it is conventional and advisable for both the *x*-minimum and the *y*-minimum to be zero. The *x*-minimum should be zero for the autocorrelogram, but a case can be made that the *y*-minimum and the *y*-maximum should be −1.0 and +1.0, respectively. If part of the goal of the research is to compare the results from different transects, *x*-axes and *y*-axes of the same kind of plot should be scaled identically.

The plot of vegetation height as a function of transect position will typically have some sort of broad-scale pattern, immediately detectable upon inspection, in addition to fine-scale variation. The details will vary markedly depending on the nature of your plant community. The fine-scale variation may be partially measurement error, but in most cases it is predominantly caused by natural variation. Increasing the number of samples (i.e., transect length) will not reduce the magnitude of this fine-scale variation.

The semivariograms and autocorrelograms will also have an overall shape, summarizing the spatial patterns of the community, as well as fine-scale variation. However, in contrast to the graph of height, increasing the sample size (transect length) will tend to decrease the finer-scale patterns. This means that we are increasingly confident that we have described how spatial pattern (variance or correlation) is related to scale (spatial lag).

Q1 How does height behave as a function of distance along the transect? Is this generally consistent with your impression of the field site?

Q2 Examine the semivariogram. Is there an identifiable nugget? Range? Sill?

Q3 Does the regionalized variable (height) exhibit spatial dependence?

Q4 Examine the autocorrelogram. Is there spatial autocorrelation?

Q5 How would you describe the nature of your spatial variation? Does your pattern consist of patches? Noise? A dominant trend? Nested patterns of variation? Random walk? A random walk (also known as "drift") is when there is spatial dependence, but the difference between each number and the previous number is random. The term *random walk* derives from a plot of distance from the starting point as a function of time for an animal whose direction of movement is purely random.

Q6 Is there periodicity in your data (i.e., did the response change regularly at several spatial intervals)? How would you know this from the shape of the semivariogram or autocorrelogram? Note that both the semivariogram and the autocorrelogram can describe variance as a function of scale, but neither can completely summarize the *nature* of spatial variation (e.g., patches, gradients, or a combination). This is akin to the observation that variance does not fully describe the statistical distribution of data (e.g., whether it is normally distributed), and that the correlation coefficient does not fully describe the nature of the relationship between two variables (e.g., whether they might have a nonlinear relationship).

Q7 Suppose a rodent species requires tall vegetation for cover. Does the nature of the spatial pattern you observe have implications for this species?

Q8 Suppose a predator only hunts in relatively short vegetation. Does the nature of spatial variation have implications for foraging behavior?

Q9 Can you think of any other biological ramifications of your results?

Q10 If you collected data from more than one site or variable, how do their spatial patterns compare?

SYNTHESIS

Q11 How does your variable behave in comparison to the supplied data sets? The supplied data sets (on the page **example data sets** in the spreadsheet accompanying this laboratory) can be pasted into the data column (column **B**) in the worksheet **ready-to-go blank**, and the semivariograms and autocorrelograms will be recalculated automatically (however, note that you may need to change the *Y*-axis scaling on the graphs).

Q12 Refer to Figure 5.1. Choose two or three of the variables and describe what natural phenomena might lead to those patterns.

Q13 If you find spatial dependence, what does this imply for the use of conventional statistics?

Q14 Are some spatial scales better than others for studying your system? Why or why not?

Q15 In theory, regionalized variables are measured at points. However, you have measured them in a quadrat. What do you expect would happen to the semivariogram if you reduced the size of the quadrat?

Q16 What is noise? Is it a useful concept?

OPTIONAL EXERCISES

EXERCISE 4: Correlation and Variation

As their names imply, the auto**correl**ograms and semi**vari**ograms stress correlation and variance, respectively. Therefore, they are likely to behave differently in data sets with different variance. In the supplied spreadsheet accompanying this exercise, locate a worksheet entitled **2 hypothetical variables**. Examine the two variables carefully.

Q17 How do these variables differ?

Q18 How do you expect their semivariograms and autocorrelograms to differ?

Q19 Now copy one of the variables and paste it in the data column (column **B**) in the sheet labeled **ready-to-go blank**. Examine both the semivariogram and autocorrelogram. Now repeat with the second variable. Were you right in your answer to the previous question?

EXERCISE 5: Variography and Fractals

Plot your semivariogram on a double logarithmic scale. Do this by left-clicking on the *X*-axis of your semivariogram. Then right-click and choose **Format axis**. Select **Scale** and click **Logarithmic**. Repeat the same procedure for the *Y*-axis.

Q20 Is the semivariogram a straight line? If so, we can say that the variable is *statistically self-similar*. This means that fine-scale patterns are indistinguishable from scaled-down versions of broader-scale patterns. The concept of "self-similarity" is intrinsic to the study of *fractal geometry*. The fractal dimension D can be determined from the slope m of the log-log semivariogram with the formula $D=(4-m)/2$. The interpretation of the fractal dimension is beyond the purpose of this chapter; see Burrough (1983) and Palmer (1988) for more details.

Q21 Are there multiple plateaus? If so, we have a hierarchy of spatial patterns. This would imply that we have distinctly different processes operating at distinctly different *scale domains*.

Q22 Would you predict that most spatial patterns in nature are self-similar, hierarchical, or neither?

EXERCISE 6: Variography using R

R code for semivariograms and autocorrelograms is provided on the website for this book. Repeat the same analyses as presented in the main lab, but using the supplied code instead.

FURTHER STUDY

This exercise only considered one-dimensional patterns. However, ecologists typically study spatial patterns in two dimensions. The same formulas for semivariance and autocorrelation hold, but the calculations are a bit more complicated because distances no longer fall in discrete lag intervals. Therefore, we typically average semivariance over a certain range of lags. A further complication arises if the patterns are not **isotropic** (statistically the same in all compass directions). In such cases, we usually calculate different semivariograms and autocorrelograms for different directions.

Furthermore, sampling need not be in a perfectly sampled transect as in this lab. It is perfectly legitimate for samples to have locations that are random, on interrupted transects and grids, or any other objective method. When samples are located at irregular intervals (and/or in two dimensions), many of the spatial lags are not a simple multiple of the minimum spacing. We deal with this by creating “lag classes” (e.g., 0–1 m, 1–2 m, 2–3 m) much in the same way as we would generate a histogram. Although there are no firm rules about how many pairs of points should fall within a lag class for an accurate semivariogram or autocorrelogram, a general rule of thumb is that it should be at least 80.

In this lab, we interpreted semivariograms and attempted to find the range, sill, and nugget by eye-balling. However, it is common to use a curve-fitting procedure such as nonlinear regression to actually obtain estimates of these parameters (see Legendre and Legendre 1998), as in Chapter 11. Such curve-fitting is an essential step for procedures such as kriging, discussed next.

Variography is often a precursor to a geostatistical interpolation procedure known as **kriging** (Hohn 1988; Isaaks and Srivastava 1989; Cressie 1991). By interpolation, we mean that we estimate the value of a regionalized variable at an unsampled location, based on knowledge from sampled locations. The most common product of kriging is a map (usually a contour map) of the variable of interest. Kriging performs best when the nugget is small relative to the sill and when the average distance between nearby samples is less than the range. See Legendre and Fortin (1989), Halvorson et al. (1994), Marinussen and Van Der Zee (1996), and Carroll and Pearson (1998) for examples of kriging in ecology.

One may have noticed a resemblance between the semivariogram and the well-known species–area relationship (Scheiner 2004). This resemblance is more than casual. As Wagner (2003) elegantly illustrates, distance decay in species distributions scales up to distance decay in species richness, one of the root causes of the species–area relationship (Palmer and White 1994). Indeed, Wagner (2003) develops numerical techniques by which one can separate how much of the semivariogram for species richness is due to intraspecific autocorrelation, and how much is due strictly to interspecific co-occurrence. Such techniques elevate variography beyond mere description of pattern, and into the realm of uncovering fundamental properties of biodiversity such as those explored in Chapter 15. While we have only discussed univariate patterns in this lab, bivariate or multivariat patterns are often of interest. If so, we can use covariograms or cross-correlograms to determine whether the relationships between variables change as a function of spatial scale. Additional multivariate approaches are also explored further in Chapter 15.

Lastly, for a simple analysis of spatial pattern along a transect (as in this lab), it is possible to perform basic calculations on a spreadsheet. However, a spreadsheet becomes cumbersome for more complex sampling designs such as interrupted or two-dimensional sampling and for complex analyses such as detection of anisotropy, significance testing, nonlinear curve-fitting, multivariate patterns, or kriging. Fortunately, a wide range of software exists for such analyses, including packages in R and GIS software (Chapters 11 and 15).

Acknowledgements We thank Steven Thompson, Sam Fuhlendorf, Anne Cross, Sophonia Roe, Marie-Josée Fortin, and four anonymous reviewers for comments on earlier versions. We especially thank Sam Fuhlendorf for allowing the use of his data, and José Ramón Arévalo for collecting new data for the lab.

REFERENCES AND RECOMMENDED READINGS[1]

Bailey TC, Gatrell AC (1995) *Interactive spatial data analysis*. Longman Scientific & Technical, Essex *An excellent text on spatial statistics; includes some demonstration software*

Burrough PA (1983) Multiscale sources of spatial variation in soil. Application of fractal concepts to nested levels of soil variations. J Soil Sci 34:577–597

*Carroll SS, Pearson DL (1998) Spatial modeling of butterfly species richness using tiger beetles (*Cicindelidae*) as a bioindicator taxon. Ecol Appl 8:531–543. *A good example of the application of spatial statistics in biodiversity studies.*

*Cohen WB, Spies TA, Bradshaw GA (1990) Semivariograms of digital imagery for analysis of conifer canopy structure. Remote Sens Environ 34:167–178. *A good example of the use of variography in digital imagery.*

*Cressie NAC (1991) *Statistics for spatial data*. Wiley Interscience, New York. *Comprehensive and authoritative work on spatial statistics.*

*Fortin M-J, Dale MRT (2014) *Spatial analysis: a guide for ecologists.* Cambridge University Press, Cambridge. *Excellent survey of different spatial analysis tools for ecologists including variograms and autocorrelation with many helpful visual aids.*

*Halvorson JJ, Bolton Jr H, Smith JL et al (1994) Geostatistical analysis of resource islands under Artemisia tridentata in the shrub-steppe. Great Basin Naturalist 54:313–328. *An example of ecological applications of geostatistics on a fine scale.*

*Hohn ME (1988) Geostatistics and petroleum geology. Van Nostrand Reinhold, New York. *A good, clear text for the use of geostatistics for interpolation and mapping.*

*Houle G (1998) Seed dispersal and seedling recruitment of Betula alleghaniensis: spatial inconsistency in time. Ecology 79:807–818. *Another good example of geostatistics in fine-scale plant ecology.*

*Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press, New York. *A popular text on the use of geostatistics for interpolation.*

Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? Ecology 74:1659–1673

*Legendre P, Fortin M-J (1989) Spatial pattern and ecological analysis. Vegetatio 80:107–138. *A good, brief introduction to spatial methods available to ecologists.*

Legendre P, Legendre L (1998) Numerical ecology, 2nd English edn. Elsevier, Amsterdam

*Marinussen MPJC, Van Der Zee SEATM (1996) Conceptual approach to estimating the effect of home-range size on the exposure of organisms to spatially variable soil contamination. Ecol Model 87:83–89. *A good applied example of the use of spatial statistics.*

*Palmer MW (1988) Fractal geometry: a tool for describing spatial patterns of plant communities. Vegetatio 75:91–102. *Follows Burrough (1983) to describe how variography can be used as an introduction to the world of fractals.*

Palmer MW, White PS (1994) Scale dependence and the species-area relationship. Am Nat 144: 717–740

*Rossi RE, Mulla DJ, Journel AJ et al (1992) Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecol Monogr 62:277–314. *A good general introduction to variography, geared toward the ecologist.*

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Scheiner SM (2004) A mélange of curves—further dialogue about species-area relationships. Global Ecol Biogeogr 13:479–484

Schlesinger WH, Raikes JA, Hartley AE et al (1996) On the spatial pattern of soil nutrients in desert ecosystems. Ecology 77:364–374 *A good geostatistical study of soil nutrients*

Sokal RR, Rohlf FJ (1981) *Biometry*. Freeman, New York

*Wagner HH (2003) Spatial covariance in plant communities: integrating ordination, geostatistics, and variance testing. Ecology 84:1045–1057. *Seminal paper uniting variography with biodiversity statistics.*

Chapter 6
Characterizing Categorical Map Patterns Using Neutral Landscape Models

Robert H. Gardner

OBJECTIVES

Spatial patterns of landscapes are the result of numerous biotic, abiotic, and anthropogenic processes, and every landscape is in some way unique. Neutral landscape models—models that lack the explicit consideration of the particular processes generating landscape pattern (Gardner et al. 1987; Gardner and Engelhardt 2008) have proven to be a helpful first step in characterizing pattern in the absence of specific ecological processes and thus serve as a null hypothesis, or baseline, for comparison with actual landscapes. Neutral landscape models have led to new understanding about habitat connectivity thresholds and the influence of landscape composition on spatial configuration (see Gardner and Urban 2007 for a review), and they offer a practical means of generating multiple landscape maps with similar statistical properties. This lab is designed to:

1. Illustrate the methods used for generating neutral landscape models;
2. Explore methods for analyzing patch structure with particular emphasis on the use of different neighborhood rules for identifying patches;
3. Explore the factors influencing connectivity in landscapes as well as threshold effects in connectivity; and
4. Examine the use of neutral models for formulating hypotheses regarding the relationship between pattern and process in actual landscapes.

Before data are collected or experiments are performed, the analysis of landscape pattern requires (at least) two things: (1) a clearly stated, testable question or

R.H. Gardner (✉)
Appalachian Lab, University of Maryland Center for Environmental Science, Frostburg, MD, USA
e-mail: rhgardner99@gmail.com

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_6

hypothesis, and (2) robust quantitative methods to address that question (Gardner and Urban 2007). Throughout this lab, a variety of neutral models will be generated to create a broad range of landscape patterns. Students will also become familiar with a number of common metrics used to quantify patterns in these landscapes. The concept of connectivity will be addressed, particularly with respect to the neighborhood rules used to define "patches," or "clusters," of habitat in a landscape.

The four exercises of this lab help students develop testable questions and interpret quantitative results. The first exercise provides familiarity with computational methods and can be completed entirely by hand or in Excel. The second investigates the surprising degree of structure present in simple random models and illustrates threshold effects. The third exercise uses multifractal maps to examine contagion effects. The fourth exercise compares metrics of real landscapes with those of a neutral model. You will be using QRULE software for Exercises 2–4, with a series of R files to analyze and display results. All software for this lab is free and can be downloaded from the book website! Some familiarity with R as well as Chapters 4, 5, and 7 is a nice complement to these exercises.

INTRODUCTION

Neutral, or null, models in ecology provide a useful baseline for comparison when examining potential cause-and-effect relationships. In terms of landscape pattern, a **neutral model** is one that exhibits characteristic spatial patterns in the absence of processes that may affect patterns in actual landscapes (e.g., topography, resource gradients, and disturbance regimes; Gardner et al. 1987; With and King 1997; Gardner and Urban 2007). In the neutral models examined here, landscape pattern is an emergent property of either simple random processes or via algorithms derived from fractal geometry that create random but auto-correlated patterns (e.g., multifractal maps). Comparing patterns and landscape indices for real landscapes with those from neutral landscape models can provide insight into the effects of ecological processes on landscape patterns; if a real landscape differs significantly from an appropriate neutral model, it is quite likely that some important ecological process is driving observed patterns. This insight allows the investigators to focus efforts on specific landscape processes and attributes (rather than a broad "shotgun" approach) to possibly reveal pattern–process relationships operating in heterogeneous landscapes.

Landscape pattern analysis usually begins by converting continuous land-cover data (e.g., derived from satellite images) into a gridded map of land-cover categories for analysis by computer programs. Most analysis methods involve identification of habitat patches (or clusters) and description of their sizes, shapes, and spatial arrangements. Although the clustering of habitat into patches may be visually obvious,

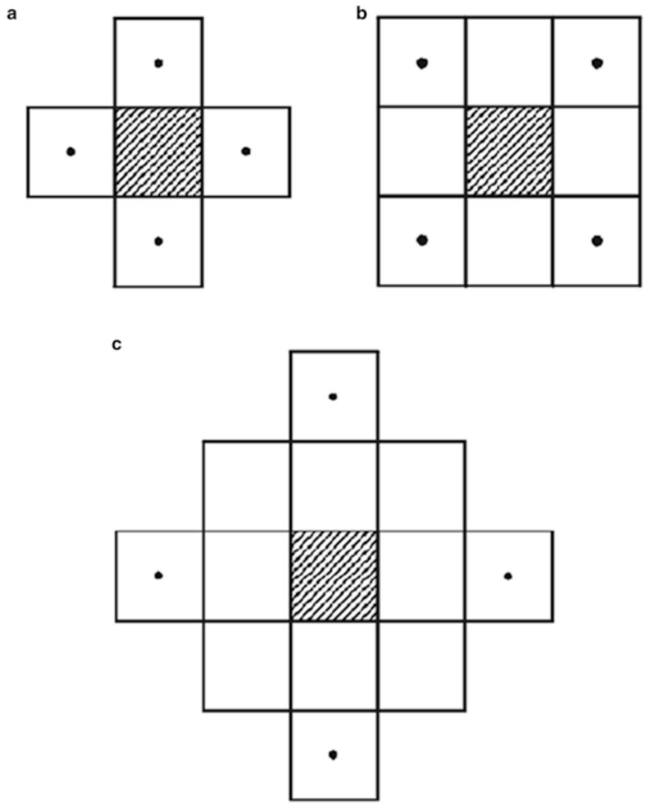


Figure 6.1 Three primary neighborhood rules: (**a**) the nearest (4) neighbor rule; (**b**) the next-nearest (8) neighbor rule; and c. the third-nearest (12) neighbor rule. The additional neighbors added to this sequence of increasing neighborhoods for each shaded pixel are indicated by the dot in the pixel center

clear patch-definition rules are needed for computers to identify habitat patches uniquely and unambiguously.

The most basic "rule" for patch definition is referred to as the "nearest-neighbor rule" (Figure 6.1a). The **nearest-neighbor rule** states that if two similar sites have one edge along one of the four cardinal directions in common (i.e., adjacent pixels), then they are "joined" and are members of the same patch. Iterative application of this rule to each "joined" site results in the identification of members of a single patch. This rule requires sites to touch along one edge to be members of the same patch, and thus a single row arranged diagonally (along a non-cardinal direction) will *not* be identified as a single continuous patch!

More commonly used in ecology is the **eight-neighbor rule** (Figure 6.1b), also called the next-nearest neighbor rule. It states that similar habitat cells are members of the same patch if they touch along one of their four edges (cardinal directions) *or* four corners (the diagonal directions). While corner sites are not considered members of the same patch with the nearest-neighbor rule, they *are* members of the same patch with the eight-neighbor rule.

Changing the patch-definition rule, such as by increasing the neighborhood that is searched for patch members, alters the metrics used to characterize landscape structure—something we will explore here. The use of different rules for defining patches for a landscape analysis is, in part, how QRULE gets its name. The user may define any rule he/she wishes with three rules conveniently "hard-wired" into QRULE code. The third "hard-wired rule" is the **third-nearest-neighbor rule** (Figure 6.1c) which extends consideration to sites that may not directly touch! Although we do not emphasize the third-nearest-neighbor rule in this lab, it can be useful for identifying habitat patches for an organism that effectively ignores a single cell gap of non-habitat within an otherwise continuous patch.

Once patches are identified, computer analysis quantifies patch attributes including size, shape, and spatial arrangement. Drawing inference from these results is problematic because, since so many metrics may be calculated, some will be statistically significant by chance alone (a Type II statistical error); a theme also explored further in Chapter 7. Neutral models were developed, in part, to avoid this problem by providing a standard against which the patterns of actual landscapes could be compared (Gardner et al. 1987; Gardner and Urban 2007). When hypotheses are clearly stated before the analysis begins, using a limited set of specific metrics helps avoid obtaining spurious, but apparently significant, results.

The simplest **neutral landscape model (NLM)** is a random map generated by assigning to each grid cell a probability of the cell being occupied by "habitat." Such a **simple random map** contains only two land-cover categories (habitat and non-habitat) and the proportion of the landscape occupied by habitat is similar to the probability of the cell being occupied by habitat. Before embarking on the computer-based generation and analysis of neutral landscape models, we begin with an exercise that demonstrates the basic procedure used by the computer algorithm.

EXERCISE 1: Simple Random Map(s) Analyzed with Three Different Neighborhood Rules

The purpose of this exercise is to become familiar with the method for generating and analyzing patch structure in random maps and the effect of changing neighborhood rules for defining patches. Your first step is to generate "by-hand" a simple random map with rows and columns equal to ten and the proportion of a cell being occupied, $\boldsymbol{p} = 0.5$. This exercise should be done with paper and pencil according to the following steps (Alternatively, see the instructions for using spreadsheet software, listed *after* the "by-hand" instructions).

Instructions for Generating and Analyzing a Simple Random Map "By-Hand"

1. Use graph paper to create a grid with ten rows and columns.
2. Repeatedly flip a coin to determine the habitat type of each cell. If heads, then the habitat type equals 0. If tails, then the habitat type equals 1.
3. Analyze the map by coloring in all sites with habitat type = 1.
4. Count the total number of colored cells, the total amount of edge, the number of clusters as defined by the nearest-neighbor rule, and the size of the largest cluster.
5. Using the next-nearest-neighbor rule, recalculate the number of clusters and the size of the largest cluster. (*NOTE:* the total number of colored cells and the total amount of edge will not be affected by this change in neighborhood rule.)
6. Record your results in tabular form.

Instructions for Using a Spreadsheet to Generate a Matrix of Random Numbers

Open Excel and then:

1. Open a new worksheet
2. Type this equation in the first cell: "`=rand()`" (this produces a single random number between the interval 0.0–1.0).
3. Copy this cell to a 10 × 10 grid of cells
4. Analyze the map by coloring all sites with random numbers ≤0.5
5. Print the resulting matrix and go to step 4 of the "by-hand" directions

Q1 The generation of maps by hand is a tedious exercise that results in a small, inadequate sample size. Does the number of habitat sites of type 1 equal exactly 50% of the map? How many sites with habitat of type 1 touched the edge of the map? How many of these sites that touched the edge of the map would have adjoined another site of habitat type 1 if the map size was increased? (*HINT*: See Gardner et al. 1987, for a discussion of cluster truncation effects.) How big would clusters be if the map size were increased?

Q2 Combine your results with those of other students and statistically summarize (e.g., mean, standard deviation, minimum, maximum) the number of cells of habitat, total amount of edge, number of clusters, and the size of the largest cluster. What are the most reliable statistics (i.e., which ones have the lowest coefficient of variation)?

Q3 Do you expect the results from actual landscapes to be more or less variable than random maps? Explain your rationale.

Using QRULE to Generate and Analyze Neutral Landscape Models

The remainder of this lab will be performed using QRULE, a program written in Fortran with separate versions that run in either DOS or Linux. The latest version of QRULE makes several improvements over older versions, including the output of statistics in metric units rather than pixel units. See documentation, *Qdocumentation.pdf* for details. QRULE was developed to be a research tool—and it still is used as such! Consequently only minimal attention has been devoted to making QRULE "user friendly." QRULE does not have a GUI (graphical user interface); nor does it produce instant graphical output or data displays. It may "crash" if you input conflicting or incorrect information (i.e., a file name that does not exist in the directory specified). The good news is that a few simple "tricks" detailed below will allow you to run QRULE in a remarkably efficient and flexible manner.

Several types of maps can be generated by QRULE. The ones of interest for this exercise are simple random maps and multifractal maps (Figure 6.2). Maps created by other programs—especially those developed from remotely sensed images—may also be read into QRULE for analysis of spatial patterns. The algorithms are explained briefly below.

Simple random maps (Figure 6.2a) may be created by specifying the number of rows and columns in the map, the number of habitat types to be generated, and the probabilities, ***pi***, associated with each habitat type ***i***—including the probability, ***p0***, for areas lacking any habitat at all. Table 6.1 provides a sample dialog for QRULE execution producing a random map with 128 rows and columns and two habitat types. A uniform random number (URN, a computer-generated random number ranging from 0.0 to 1.0) is iteratively used to randomly and independently assign a habitat type to each grid site.

In Table 6.1, the example specifies the value of ***p0*** = 0.1, ***p1*** = 0.3, and ***p2*** = 0.6. If URN ≤ ***p0***, then the site is set to "non-habitat"; if URN is between 0.1 and 0.4, the site is set to habitat type 1; and if the UNR > 0.4, it is set to habitat type 2 (also notice the cumulative probability distribution, ***CumP*** in Table 6.1 and the realized probabilities for habitat types 1 and 2 were 0.301 and 0.5991, respectively. The definitions for each landscape statistic calculated by QRULE are given in Table 6.2.

Multifractal maps produce patterns that are quite realistic (Figure 6.2b–d) because a fractal algorithm is used to produce spatially correlated patterns of land cover. Fractal maps have been frequently used by investigators wishing to use random but more realistic maps to simulate biological and physical processes (e.g., With 1994; Plotnick and Prestegaard 1995; Wiens et al. 1995). Multifractal maps (Figure 6.2b–d) are generated in QRULE by the midpoint displacement algorithm (MidPointFM2d, Saupe 1988). This algorithm creates a map of real numbers by iterative interpolation to locate the midpoint of a line, followed by perturbation of the line's midpoint by a Gaussian random value (GRV). Successive reductions of the variance of the GRV as the distance between points becomes finer and finer produces correlated patterns. Two parameters are used by QRULE to control this process:

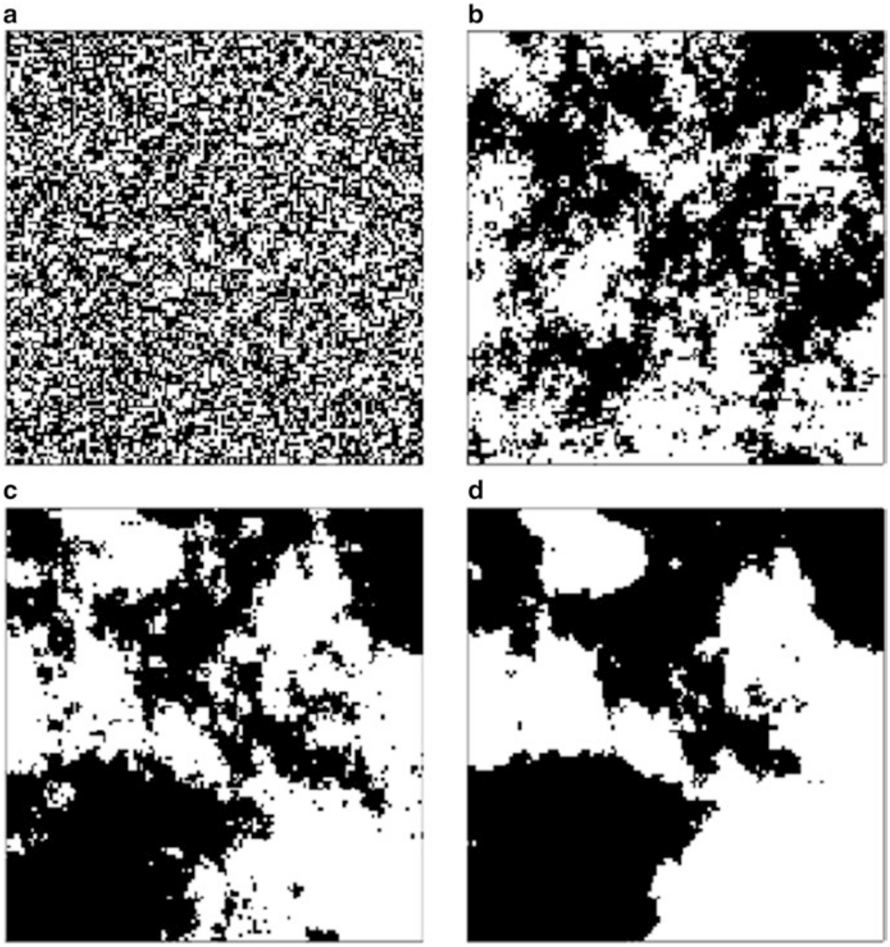


Figure 6.2 Sample maps produced in RULE: (**a**) a simple random map, and multifractal maps with (**b**) $H=0.1$, (**c**) $H=0.5$, and (**d**) $H=0.9$. In each instance, the value of p, the proportion of black cells within the map, is equal to 0.5

- ***L***, the "number of levels" or iterations of the midpoint displacement algorithm. The size of the map will always equal to 2^L. For instance when $L=4$, then the dimensions of the map (number of rows and columns) $=16$; when $L=6$, map dimensions $=64$, etc.
- ***H***, the parameter that controls the rate of reduction of the GRV in successive iterations of the midpoint displacement method (H may range in value from 0.0 to 1.0).

The generation of successive finer Gaussian increments results in the variance between points separated by distance $\boldsymbol{x}$ that is approximately equal to $\boldsymbol{x^{2H}}$ (assuming

Table 6.1 Sample Qrule dialog producing a Simple Random Map with two habitat types
User response to questions by Qrule are given in *Bold Italic*. All Output by Qrule is also written to a disk file: *rulerun.log*

./Qrule.exe

```
    Qrule (v 4.1) Landscape Pattern Analysis   20131113

    Enter map type to be analyzed:

       <I> Input existing map file

       <R> Generate a random map (with replacement)

       <S> Generate a simple random map

       <M> Generate a multifractal random map

       <G> Generate a multifractal random map with a gradient

       <X> Use input map as mask, generate "seeded" map

       <Y> Use input map as mask, generate siple random map
```

s

```
    Map choice: S

    Enter number of map rows and columns (max = 20000 ea.)
```

128 128

```
    Rows x Columns =  128 x  128

    Enter a negative random number seed
```

-191827

```
    Random number seed:    -191827

    Enter the neighborhood rule

       1 - nearest neighbor (N_nb = 4)

       2 - next nearest neighbor (N_nb = 8)

       3 - 3rd nearest neighbor (N_nb = 12)

       4 - user defined
```

(continued)

Table 6.1 (continued)

1

```
    Rule choice is: 1

    Enter the number of map classes
```

2

```
    Map classes =     2

    Enter the   3 probabilities, starting with p(0)
```

0.1

0.3

0.6

```
    The normalized probabilities are:

       P          CumP

     0    0.1000     0.1000

     1    0.3000     0.4000

     2    0.6000     1.0000

    Enter the number of replications
```

10

```
    N_Reps =     10

    Create an output maps?

     N = None

     G = generated map

     S = cluster Size map

     C = cluster ID map
```

(continued)

Table 6.1 (continued)

n

```
Map output choice = N

Perform map analysis?

   <N>o analysis

   <L>acunarity analysis

   <R>ule analysis

   <A>ll (both Rule and Lacunarity)
```

r

```
Analysis method: RULE

What is the resolution of each grid element?

  (length of the side of a grid element, in meters)
```

30

```
 Resolution:  30.0000     meters

          Mean Association Matrix

   Avg ChiX = 8.35187 w/ 4 df (FXceed (9.4480= 0.3000)

      0        1        2

  0   0.010298 0.030043 0.059652

  1   0.030043 0.090493 0.180387

  2   0.059652 0.180387 0.359046

 p's  0.099992 0.300923 0.599085
```

(continued)

Table 6.1 (continued)

```
          STATISTICAL SUMMARY (N=    10; Resolution= 30.0000 meters)
--Cover Type 0 (non habitat)--    [p =  0.1000   Cum. p =  0.1000]
--Land Cover Type  1--            [p =  0.3010   Cum. p =  0.4009]
Variable  Units  Mean        St.Dev.       C. V.      Minimum      Maximum
 L.C.size  ha  2.44800     0.510529      20.8549      1.89000      3.51000
 L.C.edge  m   1566.00      286.287      18.2814      1260.00      2160.00
 L.C.fract -   1.51835     0.710262E-01  4.67785      1.41497      1.66736
 L.C._rms  m   119.670      12.6254      10.5502      96.9325      134.728
 TTL clstr N   2118.20      46.5422      2.19725      2048.00      2210.00
 TTL edgs  m   415206.      5080.47      1.22360      404580.      423300.
 Sav size  ha 0.451091     0.201442E-01  4.46567     0.424208     0.485755
 S_Freq    N   4931.10      50.0132      1.01424      4833.00      4989.00
 Cor_len   m   73.7511      3.46446      4.69750      68.6342      79.3984
 Perc      %   0.00000      0.00000      0.00000      0.00000      0.00000

--Land Cover Type  2--            [p =  0.5991   Cum. p =  1.0000]
Variable  Units  Mean        St.Dev.       C. V.      Minimum      Maximum
 L.C.size  ha  367.317      140.730      38.3128      212.310      655.110
 L.C.edge  m   178968.      68017.9      38.0056      102360.      317880.
 L.C.fract -   1.75286     0.437641E-01  2.49672      1.68642      1.83279
 L.C._rms  m   1270.99      169.286      13.3192      976.709      1483.23
 TTL clstr N   463.000      36.6667      7.91937      386.000      512.000
 TTL edgs  m   477390.      3784.05     0.792653      469200.      482160.
 Sav size  ha  220.843      111.444      50.4632      129.076      487.170
 S_Freq    N   9814.90      50.0354     0.509791      9757.00      9913.00
 Cor_len   m   1193.55      196.669      16.4776      869.864      1480.94
 Perc      %  0.500000     0.527046      105.409      0.00000      1.00000
```

Table 6.2 Indices of spatial patterns produced by QRULE for each habitat type

Index	Definition
L.C.size	The size of largest cluster (total number of grid units making up the largest cluster)
L.C.edge	Number of edges of largest cluster sites adjacent to a different habitat type
L.C.fractal	Fractal index of largest cluster estimated as *ln*(L.C.edge) / *ln*(average diameter of the cluster)
L.C._rms	Mean squared radius of largest cluster (also known as the radius of gyration, Stauffer and Aharony 1992). If r_i is the ith of s sites in the cluster, then L.C._rms = $\sum(r_i - r_j)^2 / s^2$. Diffuse sites of size s will have a larger L.C._rms than more compact sites
TTL clusters	Total number of clusters on the map
TTL edges	Total edge of all clusters
Sav size	Area weighted average cluster size. If S_i is the size of the i^{th} cluster, then $S_{av} = \frac{\sum s_i^2}{\sum s_j}$
S_Freq	Total number of sites of current habitat type. ***P***, the fraction of sites of the current habitat type are estimated as: ***P*** = S_Freq/(***nr*** * ***nc***), where ***nr*** and ***nc*** are the number of rows and number of columns of the map, respectively
Cor_len	Average mean squared radii of all clusters
Perc/freq	Frequency (percent of all maps) with a cluster large enough to span the dimensions of the map

NOTE: Units for each index are indicated in program output and Table 6.1. See Gardner (1999) for additional details concerning the calculation of each index

that σ^2, the variance of the Gaussian process, is equal to 1.0). Thus, extremely fragmented patterns are produced when caused by a fractal algorithm that causes negative correlations among sites (i.e., *H less than 0.5*; Figure 6.2b) while positive correlations of differences produce highly aggregated patterns (i.e., *H greater than 0.5*; Figure 6.2d). For further details regarding the use of multifractal maps, see Plotnick and Gardner (1993), Pearson and Gardner (1997), and With and King (1997).

Analyzing real landscapes with QRULE simply requires that the map type be defined as **I** indicating that a map will be **input** rather than generated. Then, the full name of the map file (e.g., "C:/foldername/mapname") is entered and then the number of rows and columns and the number of habitat types is specified. The landscape map file must be a space delimited sequence of ASCII integers representing each habitat type. An example input map file, ***anti_128.map***, is provided with this exercise.

Instructions for Using QRULE

Acquiring software. All exercises require QRULE. A text editor will also be needed for handling scripts (see explanation and example below) and examining output. There are many good choices for a text editor, but a particularly useful one is **Notepad++**. Statistical and graphical analysis may be performed either with **R** or using Excel. The current version of Qrule (V4) may be acquired at either the website for this book or the Qrule website (http://www.umces.edu/al/program/gardner/qrule). The current version of Notepad++ can be downloaded at www.notepad.todownload.com and R (R Development Core Team 2010) from www.r-project.org. Excel is, of course, part of the Office software distributed by Microsoft.

(*NOTE*: An important word on operating systems. There are hundreds of operating systems (and multiple versions of each OS), but Qrule runs on only two: DOS (**D**isk **O**perating **S**ystem, a part of Microsoft Windows distributions more correctly referred to as MS-DOS), and Linux. We provide here descriptions for running Qrule under DOS. Those using Linux will have little difficulty adapting the following instructions for this OS.)

Unpacking Qrule. Download Qrule into a directory you have created—ideally, one at the top of your directory tree (e.g., c:/Qlab). Extract all files from the zip file (see Table 6.3 for a list of files contained in this zip file). This action will create a series of subdirectories. Next, go through the following steps.

1. Assuming you are running Windows software, open the **start** menu in the lower left corner of the screen (a separate application in Windows 8 titled "Command prompt" provides this functionality). Next click on the **run** icon and then type **cmd** in the window that opens and check **OK**. This action has opened a DOS window (see Table 6.4 for some useful DOS commands).

Table 6.3 Overview of files for this lab exercise

File Name	Function	Location
QruleV4.exe	Executable file for Qrule	Qlab
sample.scr	Script file for generating random map	Qlab
sample.scr	Copy of above script file	Qlab/Ex1
multifract.scr	Script file for generating multifractal map	Qlab/Ex2
Qcfd.R	R program for generating cumulative frequency distributions	Qlab/Ex2
Qcfdfun.R	Function called by Qcfd.R	Qlab/Ex2
anti_128.map	Actual map for analysis	Qlab/Ex3
anti_128.scr	Script file for analysis of the anti_128.map	Qlab/Ex3
Zview.R	R program for map display	Qlab/Ex3
mapview.R	Function called by Zview.R	Qlab/Ex3

Table 6.4 Useful DOS Commands

Command	Definition	Example
Cd	change directory	cd C:/Qlab/Ex1
Dir	list directory	dir
Copy	copy file	copy rulerun.log rulerun.abc.log
Del	delete file	del rulerun.log
Edit	simple editor	edit sample.scr

2. Navigate in the DOS window to the directory where you have placed the QRULE files. If you created a directory called C:/Qlab in your user area, then simply type `cd C:/Qlab` to reach that directory. Then, use `cd` to navigate to the appropriate subdirectory (e.g., cd Ex1 will take you to the Ex1 subdirectory under *Qlab*). The executable QRULE (file name is QruleV4.exe) exists under C:/Qlab. It will be easiest if you copy QruleV4.exe into each subdirectory.
3. Then, to execute, type `QruleV4.exe` and answer the questions that the program asks (if you can! Explanations for each input are provided subsequently.)

A test run using a script file. The program executable file for DOS is QruleV4.exe. The series of interactive questions required to run QRULE is tedious and error-prone. Many errors in input may cause QRULE to crash. No harm is done when it crashes (I told you it isn't user friendly), but you do have to start over. A more efficient, error-free method of running QRULE is to assemble all required inputs into a **script file** and run all analyses in batch mode. A sample script file, ***sample.scr***, is provided to illustrate running QRULE in batch mode (this is not part of the lab—just a practice run to see how script files are used). To run in batch mode:

1. Open the sample script file (sample.scr) with a text editor such as Notepad++. See **Qdocumentation.pdf** for description of the contents of *sample.scr* which has all the interactive answers to a QRULE run. Examine this file. You also may want to use this as a base from which to make modifications for future runs by changing and resaving the file.
2. Run QRULE using *sample.scr* by typing in the DOS window:

```
QruleV4.exe<sample.scr
```

Program results and output. Once QRULE executes, the screen output has a lot of valuable information, which is difficult to print and save. Therefore, all output is automatically written to a text file in the directory from which you have been running QRULE. The name of the output "log" file is ***rulerun.log***. You can look at this file in Notepad++ and print it if you like. [*IMPORTANT*: You should save it to a unique name before running Qrule a second time because each execution of Qrule will overwrite this file. Type `copy rulerun.log yourpreferredname.log`].

In the test run example, the screen and logfile contain the statistics for 10 landscape metrics for each of the land-cover types simulated (in this case, there are two—the habitat and non-habitat). The meaning of these metrics is given in Table 6.2. Because ten map iterations were performed, QRULE provides a statistical summary of each metric—its mean, standard deviation, C.V. (coefficient of variation), minimum and maximum values. The statistical results are also saved in a separate disk file, *stats.csv* in the directory in which the Qrule is located. This file may be viewed in Notepad++ or Excel. The file contents are described in Qdocumentation.pdf.

In addition to *rulerun.log*, five other files are created each time QRULE is executed: *assmat.dat*, *stats.csv*, *patch_cfd.dat*, *sample.map*, and *arcgrid.map* (descriptions of each file can be found in Qdocumentation.pdf.). We ignore *assmat.dat* for now but will use the other files with R programs to illustrate and examine the QRULE results.

EXERCISE 2: Random Maps and Critical Thresholds

A central concept to emerge from neutral landscape models (which were themselves derived initially from a branch of physics called percolation theory; Stauffer and Aharony 1992) is that of **critical thresholds**. In short, small changes in p can result in sudden changes in spatial patterns, and in particular, whether habitat is connected from one edge of the map or not. The value of p at which spatial patterns on a random map change qualitatively is called the **percolation threshold**, typically abbreviated as p_c or p_{crit}. In this exercise, we use QRULE to generate and analyze a series of simple random maps as a function of p, which will range from 0.1 to 0.9. The results of four landscape metrics (total number of patches, total edge, area-weighted mean patch size, and frequency of percolation among the map replicates) will be examined to determine if and at what values of p a percolation threshold may exist.

1. You and/or your team will generate a series of random maps using a specific neighborhood rule.
   a. If your last name begins with a letter in the range A–L then use **neighborhood rule 1** (nearest-neighbor); M–O then use **neighborhood rule 2** (next nearest-neighbor); P–W then use **neighborhood rule 3** (third nearest-neighbor).
   b. If you do not have a last name you are excused from this exercise.

2. Open a command window and navigate to the Qlab directory. Under the subdirectory **Ex1** you will find a script file, *sample.scr*, which is set to generate a random map with two habitat types. Each habitat type will have a value of p of 0.5. Run this file by typing:

```
QruleV4.exe < sample.scr
```

3. Save the *rulerun.log* file in Windows Explorer or by typing in the command window:

```
copy rulerun.log rulerun.ran55.log
```

4. Save the *patch_cfd.dat* and *stats.csv* files in a similar manner producing files named *patch_cfd.ran55.dat* and *stats.ran55.csv* by typing:

```
copy patch_cfd.dat patch_cfd.ran55.dat
copy stats.csv stats.ran55.csv
```

5. Now open *sample.scr* in Notepad++ (or another suitable editor) for a series of changes. First, change the two probabilities (i.e., habitat, non-habitat) for land-cover generation from 0.5 to 0.1 and 0.9 (but leave ***p0*** unchanged). Save the edited file to *sample.91.scr*. Edit again, changing probabilities to 0.2 and 0.8, saving the script file as *sample.82.scr.* And a third time with probabilities of 0.3 and 0.7, saving as *sample.73.scr.* One last time changing to 0.4 and 0.6, saving as *sample.64.scr.* You will now have four new script files.

6. Repeat the above process outlined in steps 2–4, running QRULE separately for each script file and then renaming the log files and the *patch_cfd.dat* and *stats.csv* files after each run (and before the next sequence of runs begins!).

```
QruleV4.exe<sample.91.scr
copyrulerun.logrulerun.ran91.log
copypatch_cfd.datpatch_cfd.ran91.dat
copystats.csvstats.ran91.csv
```

7. When you have finished you will have five log files and five corresponding data sets, which contain the results from QRULE for ten different values of ***p*** analyzed using one of the three neighborhood rules. You may use a text editor (e.g., Notepad++) to print and examine the results of each log file. The details of each simulation have been preserved in the csv files, which may be opened with Excel directly (or R which isn't quite as simple).
8. Using separate graph windows, plot the values from "TTL clstr", "TTL edgs", "Sav size", and "Perc" on the *Y*-axis as a function of ***p*** on the X-axis.

Q4 Were the resulting fractions of each land-cover type in the results (see the log files) equal to what was specified in the script files? If not, why not?

Q5 What is the meaning of ***po***? And why is it always set to zero in the above sequence?

Q6 Inspect the plots of "Sav size" and "TTL edge" versus ***p*** and describe the resulting relationships.

Q7 Plot "Sav size" against "L.C. size". What does this relationship show?

Q8 Inspect the histogram of "Perc" as a function of ***p***. What do these results indicate?

Q9 Compare your results with the other class teams that used a different neighborhood rule. How does the neighborhood rule affect the location (i.e., ***p***) of a critical threshold changes in metric values? Explain why you observe such changes?

EXERCISE 3: Spatial Contagion

In real landscapes, habitats are seldom (if ever) distributed in a completely random manner. Instead, land-cover categories have some degree of **spatial autocorrelation**, or contagion, in which nearby locations are likely to be similar to one another. Hopefully you explored this concept already in Chapter 5. Autocorrelated patterns also can be represented using neutral landscape models, but a more complex algorithm is required to produce the spatial contagion. The patterns with autocorrelation are still neutral landscape models because no particular generating process besides spatial autocorrelation is specified. However, they produce patterns that appear more realistic and allow the user to control the amount of spatial autocorrelation in the habitat. Neutral landscape models with spatial contagion have been used for a variety of studies, including landscape–genetics relationships (Graves et al. 2012), nitrogen cycling (Gergel et al. 2005), and animal movement (With et al. 1999).

This exercise will compare the distribution of patch sizes produced from simple random maps with those generated with a significant degree of spatial autocorrelation. The maps are produced using the multifractal algorithm described above in which the parameter **H** controls the level of autocorrelation. Higher values producing greater spatial autocorrelation, i.e., a more clumped distribution of the habitat. The instructions below will allow you to run QRULE once to a single map for your selected values of *p* and H. Your simulations will comprise only one combination of *p* and H, but comparisons with the results of other class members will allow an evaluation of the full factorial experiment for multiple values of *p* and H.

1. The files needed for this exercise are in the subdirectory **Exercise 3**, including a script file for generating a multifractal map (see *multifract.scr* for map parameters used in this simulation). Open this script file and change the neighborhood rule to the one assigned by your last name in the above exercise. Save this modified file. Only a single execution of QRULE is needed using the *multifract.scr* script file.
2. Use your version of *multifract.scr* to generate a multifractal map with QRULE (*HINT:* it will be very convenient to copy QruleV4.exe to this subdirectory).
3. Rename all resulting files from QRULE (e.g., *rulerun.log to rulerun.mf.log,* etc.) before running QRULE again.
4. We'll next use R to compare results of this map with a random map with the same probabilities. Copy the patch file from the previous exercise, *patch_cfd.ran55.dat*, to this directory.
5. Open R, then open the file *Qcfd.R* in the **New script** dropdown option under **File** menu. Reset to the proper directory in the **setwd** command of Qcfd.R, and run the first part of this program in R (through to the "STOP" comment). The subsequent statistical tests listed in this script (which you may wish to run, but are not required to do) will only work for data sets with one iteration.
6. Copy the plot of the cumulative frequency distributions here.

Q10 What is the value of "Sav" for the random and multifractal case? What differences do you see in the cumulative frequency distributions (CFDs) for the random and mulitfractal maps? (*HINT:* Compare values on the X-axis for the 1, 5, 50, 95, and 99 percentiles shown on the Y-axis)

Q11 What effect does the H parameter have on the patterns produced in the multifractal maps? (*HINT:* multiple runs of Qrule will be necessary to answer this question).

Q12 What will happen if you rerun the mulitfractal case with H = 0.9? (Those that are ambitious might try it).

Q13 What is the advantage of using the cumulative frequency distribution over simple landscape metrics?

Q14 How does the neighborhood rule affect the differences between random and multifractal maps (*HINT:* this requires comparison of your results with those from the other end of the alphabet)?

EXERCISE 4: Neutral Models and Actual Landscapes

One use of neutral landscape models is to produce multiple maps that capture spatial attributes of a real landscape so that the effect of patterns on processes of interest can be evaluated. For example, one may wish to generate spatially neutral maps that have a particular habitat amount and level of spatial autocorrelation that is based on an actual landscape. Gergel (2002) used this approach to generate replicate floodplain landscapes that had set proportions and spatial autocorrelation in different elevation classes. These neutral landscapes then provided the foundation for modeling effects of flooding on nitrogen processing (Gergel et al. 2005). This exercise will compare an actual landscape with random and multifractal neutral landscape models.

1. The "actual landscape" is an image derived from Landsat (30 m resolution) of Antietam, Maryland and rescaled to 120 m resolution to reduce map size and computational expense. This map (***anti_128.map***) has 128 rows and columns and 4 land-cover types. The script file for map analysis (***anti_128.scr***) provides the necessary input to QRULE. Run this script file by typing:

```
QruleV4.exe < anti_128.scr
```

2. Rename files as you have done before.
3. We would like to compare the patterns in this map with a simple random map and with a multifractal map. If your last name begins with letters in the range A–L, then you will make a random vs. actual landscape comparison. Otherwise, you will compare a multifractal map with the actual landscape.
4. The maps you will generate should have the same number of habitat types and values of $\boldsymbol{p}$ as the actual landscape. For the random map, this means you must

specify ***p*** for each habitat type based on the QRULE analysis of the Antietam map (step 1). For the multifractal maps, you must also select values of H such that your neutral landscape maps have similar levels of habitat clumping. Create the appropriate script file for your case, using the same values of ***p*** that are in the *runrule.anti.log* file. Save the files produced by appropriately renaming of each.

5. You decide how best to compare your neutral maps with the actual landscape.

Q15 List the script file that you used to generate a neutral model for this exercise.

Q16 What metrics are the same, what metrics differ? What differences are statistically significant (i.e., the metric value for the actual landscape lies outside the range in the neutral landscape model) or would be ecologically important?

Q17 Given what you have learned about simple random maps and multifractal maps, are there particular situations in which one or the other might be the most appropriate neutral model? Would you ever expect a real landscape to have random spatial patterns? Explain your reasoning.

Q18 Develop a question from your own research that could be answered by using QRULE and an NLM approach. Provide the rationale for your question (why is it interesting and important?), explain how you would use QRULE to answer your question (i.e., design your simulation experiment), and describe how you would evaluate the results. If you have more time for the assignment, you can even carry out your study!

CONCLUSIONS

Neutral landscape models will continue to play an important role in ecological studies that seek to understand the effects of landscape composition and configuration on ecological processes. NLMs are useful for determining the extent to which landscape metrics deviate from some theoretical expectation and for studies of how ecological processes respond to variation in landscape structure (With and King 1997). NLMs are also important for evaluating the statistical behavior and interpretability of landscape metrics (e.g., Wang and Malanson 2007); any newly introduced metric should be fully evaluated by applying it to a suite of neutral landscape models in which ***p*** and ***H*** are varied. The tools introduced here can be used in a variety of different contexts. However, it is important to remember that NLMs do not represent actual landscapes (and were never expected to do so), rather, they provide a standard against which actual landscapes may be compared, and provide a baseline against which the effects on processes can be evaluated.

REFERENCES AND RECOMMENDED READINGS[1]

Gardner RH (1999) RULE: a program for the generation of random maps and the analysis of spatial patterns. In: Klopatek JM, Gardner RH (eds) *Landscape ecological analysis: issues and applications*. Springer, New York, pp 280–303

Gardner RH, Engelhardt KAM (2008) Spatial processes that maintain biodiversity in plant communities. Perspect Plant Ecol Evol Syst 9:211–228

*Gardner RH, Milne BT, Turner MG, O'Neill RV (1987) Neutral models for the analysis of broad-scale landscape pattern. Landsc Ecol 1:19–28. *This paper introduces the concept of using neutral landscape models to explore the effects of ecological processes on landscape pattern. Comparisons between randomly generated maps and USGS land use data maps reveal more aggregated clusters in the latter, suggesting the importance of processes such as disturbance in determining landscape structure. The existence of critical thresholds in landscape connectivity was also explored in the analysis of neutral landscapes.*

*Gardner RH, Urban DL (2007) Neutral models for testing landscape hypotheses. Landsc Ecol 22:15–29. *Led by the author of the original NLM paper published in 1987 in Landscape Ecology, this is a good summary of QRULE and the current state of these models.*

*Gardner RH (2011) Neutral models and the analysis of landscape structure. In: Jopp, Reuter H, Breckling B (eds) *Modelling complex ecological dynamics*. Springer, New York, pp 215–229. *In addition to providing an overview of QRULE software, this chapter explains hypothesis testing and options for statistical tests with NLMs. This chapter is one of many in a book explaining a wide variety of ecological models.*

Gergel SE (2002) Assessing cumulative impacts of levees and dams on floodplain ponds: a neutral terrain approach. Ecol Appl 12:1740–1754

Gergel SE, Carpenter SR, Stanley EH (2005) Do dams and levees impact nitrogen cycling? Simulating the effects of flood alterations on floodplain denitrification. Glob Chang Biol 11(8):1352–1367

Graves TA, Wasserman TN, Ribeira MC et al (2012) The influence of landscape characteristics and home-range size on the quantification of landscape-genetics relationships. Landsc Ecol 27:253–266

Pearson SM, Gardner RH (1997) Neutral models: useful tools for understanding landscape patterns. In: Bisonnette JA (ed) *Wildlife and landscape ecology: effects of pattern and scale*. Springer, New York, pp 215–230

*Phillips JD (2007) The perfect landscape. Geomorphology 84:159–169. *Argues that the particular combination of factors affecting landscape development (e.g., topography, geology, climate, landscape and disturbance history, etc.) are diverse and unique producing outcomes that can not be deterministically predicted. Consideration of the suite of possible is required if landscape form and function is to be understood*

Plotnick RE, Gardner RH (1993) Lattices and landscapes. In: Gardner RH (ed) *Lectures on mathematics in the life sciences: predicting spatial effects in ecological systems*. American Mathematical Society, Providence, RI, pp 129–157

Plotnick RE, Prestegaard KL (1995) Fractal and multifractal models and methods in stratigraphy. In: Barton CC, LaPointe PR (eds) *Fractals and their use in the petroleum industry*. Plenum, New York, pp 73–96

R Development Core Team (2010) *R: a language and environment for statistical computing*. R Development Core Team. R Foundation for Statistical Computing, Vienna

Saupe D (1988) Algorithms for random fractals. In: Petigen HO, Saupe D (eds) *The science of fractal images*. Springer, New York, pp 71–113

Stauffer D, Aharony A (1992) *Introduction to percolation theory*. Taylor & Francis, London

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Wang Q, Malanson GP (2007) Patterns of correlation among landscape metrics. Phys Geogr 28:170–182

Wiens JA, Crist TO, With KA et al (1995) Fractal patterns of insect movement in microlandscape mosaics. Ecology 76:663–666

With KA (1994) Using fractal analysis to assess how species perceive landscape structure. Landsc Ecol 9:25–36

*With KA, King AW (1997) The use and misuse of neutral landscape models in ecology. Oikos 79:219–229. *The applications of neutral models to problems in landscape ecology are reviewed here. The authors discuss the appropriateness of using neutral landscape models to infer relationships between pattern and process. The paper also discusses misuses of neutral models, describing analyses that mistakenly conclude a direct causal connection between pattern and process.*

With KA, Cadaret SJ, Davis C (1999) Movement responses to patch structure in experimental fractal landscapes. Ecology 80:1340–1353

Chapter 7
What Constitutes a Significant Difference in Landscape Pattern?

Tarmo K. Remmel and Marie-Josée Fortin

OBJECTIVES

Characterizing landscape spatial structure can provide insights about the underlying mechanisms that generate pattern. Quantifying spatial structure enables analysis of landscape change over time as well as comparisons among different locations. Although numerous landscape metrics (LMs) exist to quantify spatial structure and characterize a landscape, how do we know when two landscapes significantly differ? As a single landscape represents only one replicate, its metrics are not statistics; thus, testing for differences between two landscapes becomes difficult. To address this problem, randomization procedures can help assess statistical significance using simulation approaches that assess whether the observed spatial structure could have occurred by chance alone. In this chapter, exercises will allow students to accomplish the following objectives:

1. Perform significance tests of landscape metrics based on a randomization procedure using a simulation model;
2. For one landscape, assess whether LM values are significantly different than those from landscapes of similar class proportions and spatial autocorrelation;
3. Determine where an LM value falls within its potential distribution, after controlling for class proportions and spatial autocorrelation;
4. Evaluate the statistical differences of LM values from two landscapes; and
5. Graphically produce, present, and discuss results in an R software environment.

Exercise 1 uses a simple, simulated binary landscape to explore null hypothesis testing, whereas Exercise 2 addresses statistical significance for more than one

T.K. Remmel (✉)
York University, Toronto, ON, Canada
e-mail: remmelt@yorku.ca

M.-J. Fortin
University of Toronto, Toronto, ON, Canada

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_7

landscape. In Exercise 3, an applied example is explored within the context of making landscape restoration decisions. All of our implementation, analysis, graphics, and exercises are produced within the R statistical software environment (R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/, 2010) and students are provided with all coding and files needed (from the book website). This lab requires installation of R (or R Studio) as well as loading the following libraries (*SDMTools, fields*). Some basic familiarity with R is helpful for completing this lab, and knowledge of concepts from Chapter 5 is assumed. Where relevant, advice is provided for adapting the code to the user's own research.

INTRODUCTION

Landscape metrics describe and quantify the spatial pattern of categorical landscape data (McGarigal 2002, as well as Chapters 4 and 10). LMs are often used to compare spatial patterns of landscapes from distinct regions or within the same landscape through time. However, such comparisons based on LM are valid however, only when the proportions, p_i, of each category, i, are the same between the landscapes (Turner et al. 2001). The reason for comparing only landscapes with similar values of p is because of the strong relationship between landscape metrics and the proportion of the landscape occupied by that cover type.

Another key factor affecting spatial pattern is spatial autocorrelation (i.e., values at nearby locations are more similar than by chance; Fortin and Dale 2005) which you explored in Chapter 5. Clumped, aggregate patterns occur when correlations are positive whereas dispersed, disaggregate patterns occur when correlations are negative. Spatial autocorrelation is also confounded with class proportions. As a result, in comparisons among landscapes of different proportions, one cannot often determine whether the differences are due to class proportions, spatial autocorrelation, or some other process (e.g., fragmentation).

Prior to comparing LM among landscapes, one should first determine whether the observed LM values for a single landscape could have occurred by chance. This is not a trivial task because probability distributions of LM are largely unknown, and cannot be analytically derived, especially when both proportion and spatial autocorrelation need to be accounted for (Fortin et al. 2003; Remmel and Csillag 2003). As a result, the distributions of LM must be constructed empirically using randomization and simulation procedures. The lack of analytically derived distributions in ecological studies often requires computer-intensive randomization procedures, such as resampling, Monte Carlo methods, or bootstrapping (Manly 2006; Fortin et al. 2012), in order to perform significance testing.

While randomization procedures are very flexible methods, they are usually based on the assumption of data independence. In a spatial context, this independence assumption corresponds to a complete spatial randomization (CSR), where the values of a variable are equally likely to be distributed over the entire area. In the presence of spatial autocorrelation, this assumption is unlikely to be valid, and hence a null distribution should be generated using stochastic simulation approaches. Stochastic simulations can produce highly replicated landscapes with known levels of spatial autocorrelation and class proportions in order to generate the empirical distributions needed for significance testing (Fortin et al. 2003; Remmel and Csillag 2003). Once you complete the lab, we recommend you re-read this introductory material to help solidify these concepts and new terminology.

For simplicity, the examples we use in this lab are binary landscapes, and their proportions and spatial structure affect the values of the LM (Proulx and Fahrig 2010). Figure 7.1 (after Remmel and Csillag 2003) illustrates the dependence

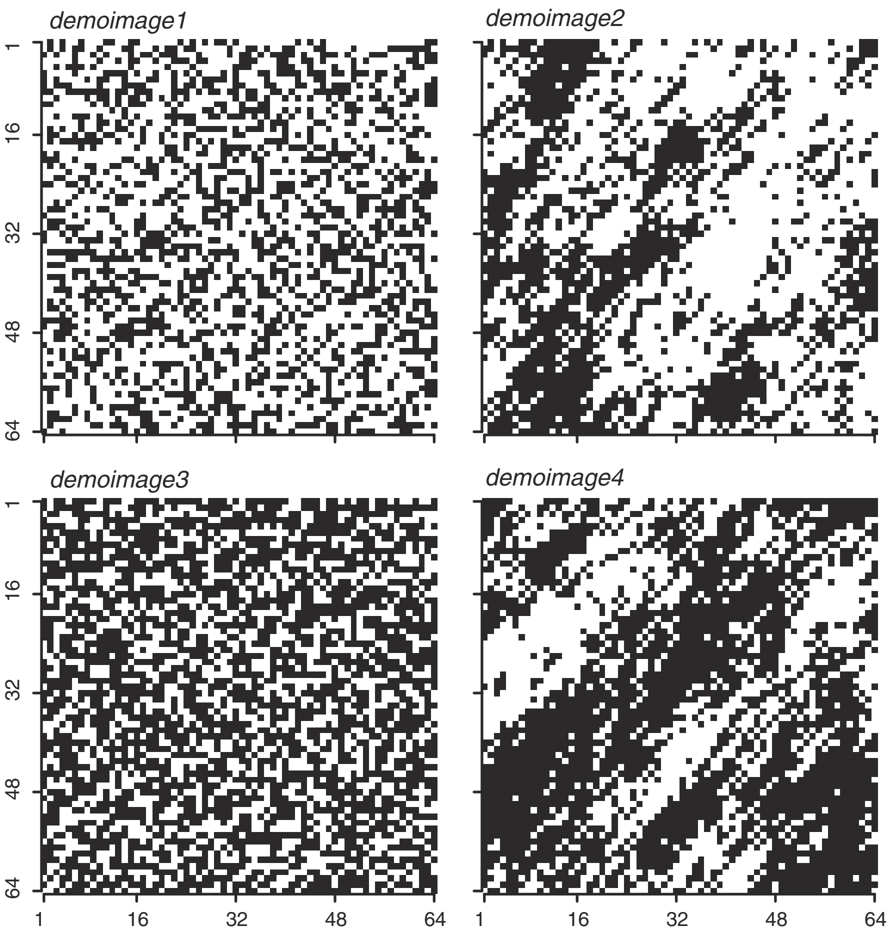


Figure 7.1 Four simulated landscapes (64×64 pixels) with specified class proportions and degrees of spatial autocorrelation. Landscapes have the following class-level proportions: in the *top row* Black = 0.40 and White = 0.60; while in the *bottom row* Black = 0.60 and White = 0.40. The levels of spatial autocorrelation are fixed: random ($\rho_W = 0.00$) in the left column and distinct patches ($\rho_W = 0.49$) in the right column

between p and spatial autocorrelation: when the proportion increases, the degree of spatial autocorrelation increases as well; therefore, LM should be compared between landscapes with similar proportions occupied by the category of interest.

Simulated Landscapes: A Stationary Stochastic Random Field Simulator

Several algorithms exist to generate binary landscapes with a known degree of spatial autocorrelation (Hargrove et al. 2002; Fortin et al. 2003; Remmel and Csillag 2003; Gardner and Urban 2007; as well as Chapter 6 in this book) but none exist (with the same degree of control) for landscapes with multiple categories, due to the inherent complexity of spatial patterns possible. Hence, the exercises in this chapter focus only on binary landscapes and significance testing of class-level metrics.

We implement **conditional autoregressive (CAR) simulations** (Cressie 1993) for generating null distributions of **binary** (two land cover classes: black and white), **isotropic** (the spatial autocorrelation parameters do not vary with direction), **stationary** (the spatial autocorrelation parameters are constant across the entire map) landscapes that rely on the stochastic random field model (Remmel and Csillag 2003). The CAR model utilizes a covariance matrix, **C**:

$$\mathbf{C} = \left(\mathbf{I} - \rho\mathbf{W}\right)^{-1}$$

where **I** is an **identity matrix** (a matrix of all zeros except for the diagonal that is filled with ones), and **W** is a **connectivity matrix** that defines which cells are considered neighbors (nearest neighbors = 1) or not (=0). Also, $\boldsymbol{\rho}$ is the **spatial autocorrelation** parameter, similar to Moran's *I* (Moran 1950) that ranges from −1 (repulsion) to 1 (attraction). Our implementation ranges from complete spatial randomness ($\rho_W = 0.00$) to highly spatially autocorrelated, where ρ approaches 1. For continuous data, ρ can be estimated using Whittle's algorithm (Whittle 1954) which is explained in more detail in Appendix A. As this correction factor requires intensive computation, it has already been performed and stored as a lookup table in the provided Remmel–Fortin code as object *DIFF50* and is used internally when estimating ρ_W.

Landscape Metrics

The LM computed in the provided Remmel–Fortin code are class-level metrics (Table 7.1) from FRAGSTATS (McGarigal and Marks 1995) and implemented within the *SDMTools* (VanDerWal et al. 2011) library in R. Here, numbers 1–38 to refer to specific class metrics corresponding to the black category, with 39–76 for white as computed by *SDMTools* (Table 7.1). All LMs, where required, are computed using nearest neighbors (4-neighbor rule).

Table 7.1 Explanation of 38 class-level landscape metrics

Low	High	Metric acronym	Definition of class-level metric
1	39	`class`	Particular patch type from the original input data
2	40	`n.patches`	Number of patches of a particular patch type or in a class
3	41	`total.area`	Sum of the areas (m^2) of all patches of the corresponding patch type
4	42	`prop.landscape`	Proportion of the total landscape represented by this class
5	43	`patch.density`	Numbers of patches of corresponding patch type divided by total landscape area (m^2)
6	44	`total.edge`	Total edge length of a particular patch type
7	45	`edge.density`	Edge length on a per unit area basis that facilitates comparison among landscapes of varying size
8	46	`landscape.shape.index`	A standardized measure of total edge or edge density that adjusts for the size of the landscape
9	47	`largest.patch.index`	Largest patch index quantifies the percentage of total landscape area comprised by the largest patch
10	48	`mean.patch.area`	Average area of patches
11	49	`sd.patch.area`	Standard deviation of patch areas
12	50	`min.patch.area`	Minimum patch area of the total patch areas
13	51	`max.patch.area`	Maximum patch area of the total patch areas
14	52	`perimeter.area.frac.dim`	Perimeter-area fractal dimension equals 2 divided by the slope of regression line obtained by regressing the logarithm of patch area (m^2) against the logarithm of patch perimeter (m)
15	53	`mean.perim.area.ratio`	Mean of the ratio patch perimeter. The perimeter-area ratio is equal to the ratio of the patch perimeter (m) to area (m^2)
16	54	`sd.perim.area.ratio`	Standard deviation of the ratio patch perimeter
17	55	`min.perim.area.ratio`	Minimum perimeter area ratio
18	56	`max.perim.area.ratio`	Maximum perimeter area ratio
19	57	`mean.shape.index`	Mean of shape index
20	58	`sd.shape.index`	Standard deviation of shape index
21	59	`min.shape.index`	Minimum shape index
22	60	`max.shape.index`	Maximum shape index
23	61	`mean.frac.dim.index`	Mean of fractal dimension index
24	62	`sd.frac.dim.index`	Standard deviation of fractal dimension index
25	63	`min.frac.dim.index`	Minimum fractal dimension index
26	64	`max.frac.dim.index`	Maximum fractal dimension index
27	65	`total.core.area`	Sum of the core areas of the patches (m^2)
28	66	`prop.landscape.core`	Proportional landscape core
29	67	`mean.patch.core.area`	Mean patch core area
30	68	`sd.patch.core.area`	Standard deviation of patch core area

(continued)

Table 7.1 (continued)

Low	High	Metric acronym	Definition of class-level metric
31	69	`min.patch.core.area`	Minimum patch core area
32	70	`max.patch.core.area`	Maximum patch core area
33	71	`prop.like.adjacencies`	Calculated from the adjacency matrix, which shows the frequency with which different pairs of patch types (including like adjacencies between the same patch type) appear side-by-side on the map (measures the degree of aggregation of patch types)
34	72	`aggregation.index`	Computed simply as an area-weighted mean class aggregation index, where each class is weighted by its proportional area in the landscape
35	73	`lanscape.division.index`	Based on the cumulative patch area distribution and is interpreted as the probability that two randomly chosen pixels in the landscape are not situated in the same patch
36	74	`splitting.index`	Based on the cumulative patch area distribution and is interpreted as the effective mesh number, or number of patches with a constant patch size when the landscape is subdivided into S patches, where S is the value of the splitting index
37	75	`effective.mesh.size`	Equals 1 divided by the total landscape area (m^2) multiplied by the sum of patch area (m^2) squared, summed across all patches in the landscape
38	76	`patch.cohesion.index`	Measures the physical connectedness of the corresponding patch type

Data Input Format

Currently, the Remmel–Fortin code is constrained to square binary landscapes having 64×64 pixels. While computation for larger landscapes is possible, processing times become prohibitive for demonstration purposes. Remember that some possible bias may result in computations for some LMs due to constraints on patch size using a 64×64 landscape. If importing your own landscapes for future analyses, ensure that all landscape representations are in numerical matrix format.

Step 0: Initialize Workspace by Loading Libraries, Demo Data, and Lookup Tables

If the package **`SDMTools`** has not been downloaded and installed on your computer, you will need to install it from the Comprehensive R Archive Network (CRAN) along with all dependencies. Then, load the library and source the code provided for this chapter.

```
> library(SDMTools)
> load("Remmel-Fortin.save")
```

EXERCISES

EXERCISE 1: Analysis of a Single Landscape

Using a single landscape, one can assess the significance of observed LMs relative to expectations based on null empirical distributions. The null hypothesis would be that the LM value could occur due to random chance, given the composition and configuration of that landscape. The expected distribution (and its variability) is fabricated by simulating landscapes with identical extent, spatial resolution, composition, and configuration, from which LM values would be quantified. If the observed LM falls within either tail of this distribution, it is then considered significantly different than what could be expected by chance.

In this exercise, you will compute a series of landscape metrics (see Table 7.1) on a landscape (represented as the `demoimage3` object). You will then determine whether observed LM values are significantly different from those based on null empirical distributions generated using the CAR simulator. We use $n = 100$ simulations to aid the feasibility of teaching as well as precedence (Remmel and Csillag 2003); though, this value could be adjusted in the code.

Use the following steps:

Step 1: Plot the Original Landscape Dataset

```
> plot.new()
> par(pty="s", mfrow=c(1,1))
> imaks(demoimage3)
> title("demoimage3")
```

Step 2: Compute Parameters from the Observed Binary Landscape

Two parameters need to be estimated from the input binary landscape data:

- The proportion of each category (black/LOW, white/HIGH)
- The estimated degree of spatial autocorrelation (recall the Whittle's algorithm)

Then, these two estimated parameters are used to generate the empirical distributions for 38 class-level metrics for both classes using $n = 100$ landscapes.

```
> result1 <- singlemap(IMG = demoimage3, VERBOSE = TRUE, reps = 100)
```

Notice the total number of landscape pixels is 4096 (64×64 pixels = 4096) or alternatively 2458 Black + 1638 White = 4096 total pixels. The proportion of black pixels is (2458/4096) = 0.60 and the corresponding proportion of white pixels is (1638/4096) = 0.40. The Whittle estimate of rho ($\rho_W = 0.00$) is zero, or very close to it, indicating the absence of spatial autocorrelation.

Step 3: Assess the Significance of Each Class-Level Metric

The significance of each metric can be assessed by computing its associated probability. In other words, the number of times the observed metric is greater or smaller than the empirical null distribution. The smallest probability that can be obtained is 0.01. Here, one can assess the significance of the LM by comparing the computed probability, P, against $\alpha = 0.05$. These results are returned with the `singlemap()` function call above.

Next, plot the observed metric values (as indicated by a red dot) for a select subset of the distributions of metrics (as specified by

`"metrics = c(2,7,18,20,21,22)"`; here 2 = `LOW.n.patches`, 7 = `LOW.edge.density`, 18 = `LOW.max.perim.area.ratio`, 20 = `LOW.sd.shape.index`, 21 = `LOW.min.shape.index`, 22 = `LOW.max.shape.index`).

NOTE: see Table 7.1 for how LMs link to these short-hand numbers. Recall that black, B, or LOW refers to the lower value of the two categories, and white, W, or HIGH for the other.

```
> singleplotter(data=result1, img = demoimage3, metrics=c
(2,7,18,20,21,22), rows=2, cols=3, addactual=TRUE, colour=TRUE)
```

If `addactual = FALSE`, then the red dot (indicating the original LM value being analyzed, shown within the simulated distribution) will not be added to each boxplot.

NOTE: Due to the stochastic nature of this approach, your results may differ from this, as well as differ from your classmates. Furthermore, `"NA"` in the results below indicates all simulated values were identical.

```
Actual Metric Value ( LOW.n.patches ):  3
88 higher values, 4 lower values, and 8 identical values as the map.
Probability of map having a value <= to expectation: P=0.9600
Probability of map having a value >= to expectation: P=0.1200

Actual Metric Value ( LOW.edge.density ):  0.9897461
25 higher values, 73 lower values, and 2 identical values as the map.
Probability of map having a value <= to expectation: P=0.2700
Probability of map having a value >= to expectation: P=0.7500

Actual Metric Value ( LOW.max.perim.area.ratio ):  4
0 higher values, 5 lower values, and 95 identical values as the map.
Probability of map having a value <= to expectation: P=0.9500
Probability of map having a value >= to expectation: P=1.0000

Actual Metric Value ( LOW.sd.shape.index ):  11.10245
5 higher values, 95 lower values, and 0 identical values as the map.
Probability of map having a value <= to expectation: P=0.0500
Probability of map having a value >= to expectation: P=0.9500
```

```
Actual Metric Value ( LOW.min.shape.index ):  1
1 higher values, 0 lower values, and 99 identical values as the map.
Probability of landscape having a value <= to expectation: P=1.0000
Probability of landscape having a value >= to expectation: P=0.9900

Actual Metric Value ( LOW.max.shape.index ):  20.23
22 higher values, 78 lower values, and 0 identical values as the map.
Probability of map having a value <= to expectation: P=0.2200
Probability of map having a value >= to expectation: P=0.7800
```

The probabilities above indicate how likely the observed map LM is relative to the simulated values (the null, empirical distribution). Two probabilities are given. First shown is the probability that the map value is less than or equal to the expected value obtained from the empirical null distribution. Second is the probability that the map value is greater than or equal to the expected value obtained from the empirical distribution. These values are computed as (e.g., `LOW.n.patches`): $P=(88+8)/100=0.9600$, where the 8 is the number of simulated values equal to the observed one, and 88 is the number of simulated values greater than the observed one. Thus, $P=0.9600$ that the observed LM value for the map is less than what would be expected from the empirical distribution (because there are many more simulated landscapes that had higher values).

These probabilities can be considered for one- or two-tailed tests, depending on the context of the question posed. If the question asks simply whether an observed LM computed for a map differs from an expected value, the test would be two-tailed, and it is possible to specify a probability level that identifies how far in the tail of the distribution the observed LM resides. However, if the question posed implies directionality (e.g., is this landscape more fragmented than expected?), the test would be one-tailed. Thus, the probability that of the LM is in the upper-tail would be assessed because the specified LM measuring fragmentation would have this as an implied directionality.

Step 4: Explore Additional Metrics

Try this exercise using different class-level metrics (e.g., select five new class-level LMs; see Table 7.1) and different landscape datasets (e.g., `demoimage1`, `demoimage2`, `demoimage3`, and `demoimage4`).

Q1 Are the class-level LMs significantly different than expected under the null hypothesis of random chance, given null empirical distributions constructed based on simulations with identical composition and spatial autocorrelation? You will need to run `singlemap()` and `singleploter()` functions for each landscape you wish to assess (Table 7.2).

Q2 Are the probabilities different than one would expect given how the original data were generated (see Figures 7.1 and 7.2)?

Q3 How do the probabilities differ from what one would expect based on the proportion of the classes in the landscapes (e.g., `demoimage1` and `demoimage2` or `demoimage3` and `demoimage4`) versus the degree of spatial autocorrelation (e.g., `demoimage1` and `demoimage3` or `demoimage2` and `demoimage4`) (see Figure 7.1)?

Table 7.2 Sample LM results for **`demoimage3`** with actual values and then probabilities in parentheses. The probabilities shown for **`demoimage3`** are probabilities of having an LM value less than or equal to expectations. Entries for **`demoimage2`** are to be filled by students

LM	`demoimage3`	`demoimage2`
Number of patches	3 (0.9600)	
Edge density	0.989746 (0.2700)	
Maximum perimeter/area ratio	4 (0.9500)	
Standard deviation shape index	11.10245 (1.0000)	
Minimum shape index	1 (1.0000)	
Maximum shape index	20.23 (0.2200)	

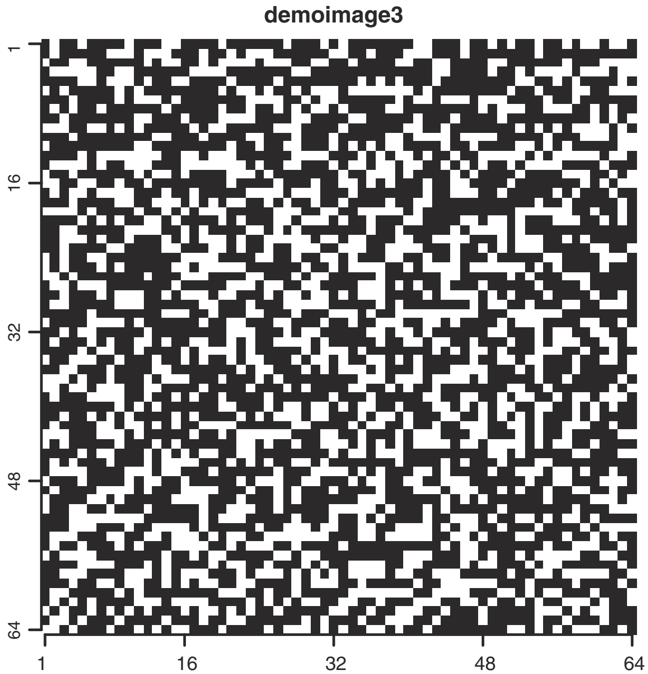


Figure 7.2 Plot of landscape **`demoimage3`** with two categories (LOW in black and HIGH in white) and 64×64 pixels

EXERCISE 2: Comparing Two Different Landscapes

For two different landscapes (either the same landscape at two different time periods, or two landscapes at different locations), it is also important to assess whether observed LM values are similar based on null empirical distributions. The null hypothesis is that the mean expected LM value is equal to the same LM from a simulated landscape with identical extent, spatial resolution, composition, and spatial autocorrelation. We will test the null hypothesis by assessing whether distributions for the specified LM overlap by a specified amount; if the overlap is large, the two values are not considered significantly different.

In this exercise, you will compute a series of landscape metrics (Table 7.1) on two landscapes (`demoimage2` and `demoimage3`) to determine whether the spatial structures of the landscapes are similar or not. You will learn how to determine whether or not class-level metrics from two landscapes are significantly different based on expectations from null empirical distributions ($n = 100$ simulated landscapes) generated using the CAR simulator. To do so, perform the following steps.

Step 1: Plot the Two Original Landscape Datasets (Figure 7.3)

```
> plot.new()
> par(pty="s", mfrow=c(1,2))
> imaks(demoimage2)
> title("demoimage2")
> imaks(demoimage3)
> title("demoimage3")
```

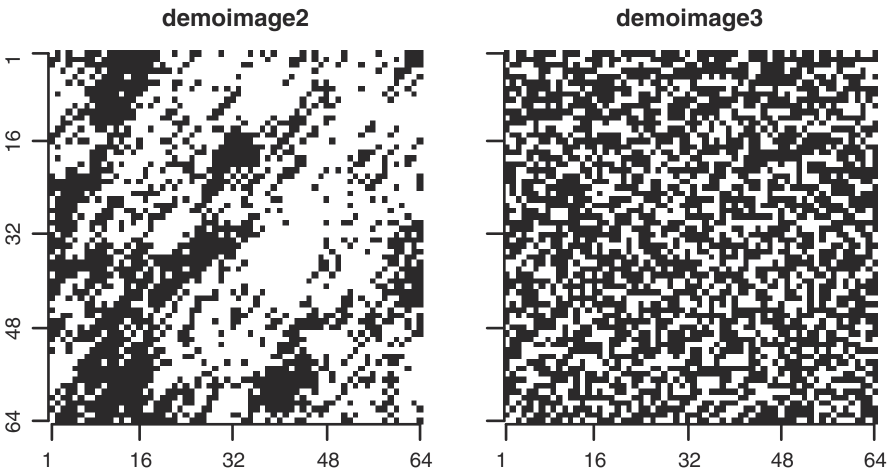


Figure 7.3 Plot of landscapes **`demoimage2`** and **`demoimage3`**

Step 2: Compute Parameters for Each Observed Binary Landscape

As in the first exercise, it is necessary to estimate pattern parameters for both input maps: (1) the proportion of each category (black, white), and (2) the estimated degree of spatial autocorrelation based on Whittle's algorithm (as in Appendix A). Then, the two estimated parameters are used to generate the empirical distributions for all the 38 class-level metrics for both categories using 100 simulated landscapes based on the CAR simulator (as described earlier).

```
> result1 <- singlemap(IMG = demoimage3, VERBOSE = TRUE,
reps = 100)
```

You should produce the same results here as in Step 2 of Exercise 1 (see explanation in Exercise 1).

```
> result2 <- singlemap(IMG = demoimage2, VERBOSE = TRUE,
reps = 100)
```

Step 3: Assess the Significance of Each Class-Level Metric

The significance of each class-level metric can be assessed by computing its probability (i.e., the number of times the observed class-level metric is greater or smaller than the expected metric under the empirical null distribution based on 100 replicates). Plot the observed class-level metrics (as indicated as a red dot) and a selected subset of the empirical distributions of the class-level metrics indicated as: "`metrics =c(2,7,18,20,21,22)`" where 2=`LOW.n.patches`, 7=`LOW.edge.density` and 18=`LOW.max.perim.area.ratio`, 20=`LOW.sd.shape.index`, 21=`LOW.min.shape.index`, 22=`LOW.max.shape.index`).

```
> singleplotter(data=result1, img=demoimage3, metrics=c(2,7,18,
20,21,22), rows=2, cols=3, addactual=TRUE, colour=TRUE)
> singleplotter(data=result2, img=demoimage2, metrics=c(2,7,18,
20,21,22), rows=2, cols=3, addactual=TRUE, colour=TRUE)
```

For explanation of these results, see Step 3 in Exercise 1.

Step 4: Plot Side-By-Side Boxplots

For comparison purposes, it is useful to produce side-by-side boxplots for each metric, contrasting the range of variability between map pairs. The range of LM variability comes from the *n* simulated landscapes, forming the empirical null distribution and metric expectation.

```
> doubleplotter(data1 = result1, data2 = result2, img1 = demoimage3,
img2 = demoimage2, metric = 8)
> doubleplotter(data1 = result1, data2 = result2, img1 = demoimage3,
img2 = demoimage2, metric = 15)
```

Step 5: Assess the Significance of Each Class-Level Metric

Assess whether there is a significant difference for each class-level metric for each observed landscape and interpret the results to determine whether the confidence intervals of each metric are overlapping. This can be most easily determined visually by looking at the boxplots for overlap of the notched region, but could be done numerically by extracting the simulated values stored in the result objects (where exact value ranges of the notches could be computed and compared) (Figures. 7.4 and 7.5).

Step 6: Additional Lab Exercises

Try this exercise using different class-level metrics (e.g., select five new class-level LMs; see Table 7.1) and different landscape datasets (e.g., `demoimage1`, `demoimage2`, `demoimage3`, and `demoimage4`).

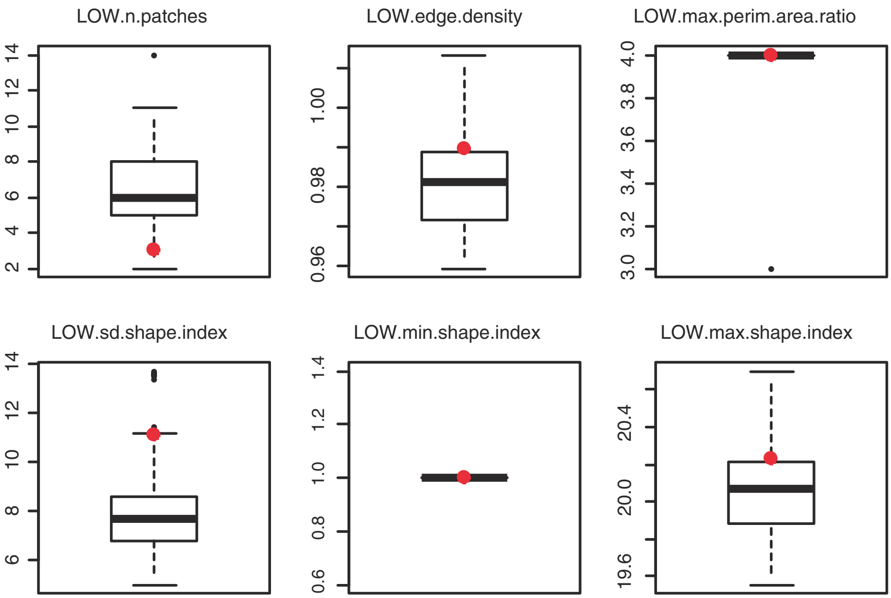


Figure 7.4 Boxplots of null empirical class-level metrics based on 100 simulated landscapes. The observed LM value is indicated by *red dot*

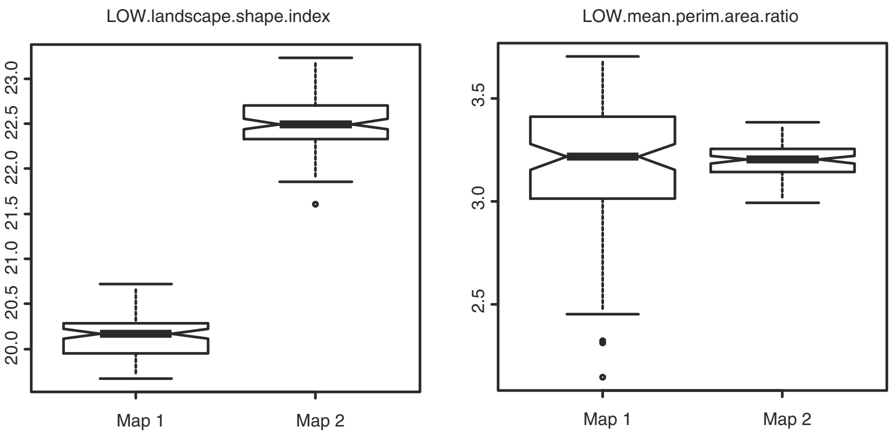


Figure 7.5 Boxplots of the class-level metrics for the two landscapes (Map 1 = **`demoimage3`** and Map 2 = **`demoimage2`**). *Left panel* shows confidence intervals for the landscape shape index and *right panel* shows mean perimeter area ratio. When the boxplot notches of two landscapes do not overlap, one can conclude that the two medians differ (Chambers et al. 1983)

Q4 Are the class-level LMs significantly different than expected under the null hypothesis of random chance, given the null empirical distributions constructed based on simulations with identical composition and spatial autocorrelation? You will need to run `singlemap()` and `singleplotter()` functions for each landscape you wish to assess. This is done individually for each landscape selected; when comparing two landscapes, it must be done individually for each of these two landscapes.

Q5 Are the probabilities different than one would have expected given how the original data were generated (see Figure 7.1)?

Q6 How do the probabilities differ from how one would expect according to the proportion of the classes in the landscapes (e.g., `demoimage1` vs. `demoimage2`, or `demoimage3` vs. `demoimage4`) compared with the degree of spatial autocorrelation (e.g., `demoimage1` vs. `demoimage3`, or `demoimage2` vs. `demoimage4`) (see Figure 7.1)?

Q7 For each pair of landscapes compared, and each LM you are interested about, are the landscapes significantly different based on the overlap of the boxplots (as was illustrated in Figure 7.5)?

EXERCISE 3: Determine a Landscape's Position Within the Distribution of Possible Class-Level Metric Values

Landscape pattern assessment can help inform an intervention or manipulation for management purposes. Restoring (or enhancing or adjusting) landscape spatial structure may be needed to influence provisioning of ecosystem services. In the context of conservation, one might wish to reduce fragmentation to help maintain biodiversity. To accomplish either, one must first determine the extent to which changing composition vs. spatial autocorrelation might change LM values. After deciding on one or more useful, robust, and informative LM(s), the next step would be to determine where the LMs reside within their null empirical distributions.

The goal of this third exercise is to determine where the observed landscape exists within the class-level metric space, given its proportion and estimated degree of spatial autocorrelation. Observing LM values among the joint influences of composition and configuration permits us to identify which aspect of spatial pattern (when manipulated) would most efficiently lead to the desired LM value change. To do so, perform the following steps.

Step 1: Perspective Plots of Class LM Median and Variance

Begin by producing a perspective plot for the median of a selected LM as it varies with class proportion and degree of spatial autocorrelation. The example provided has prop = 0.72 and $\rho_W = 0.49$ relative to a given class-level metric; the median value is based on 100 simulated landscape replicates.

```
> tempmed <- apply(surfaces[9,,,], MARGIN=c(1,2), median)
> persp(tempmed, ticktype="detailed", cex.axis=0.7, zlab="Metric",
ylab="Proportion", xlab="Rho", theta=-45)
```

The variance of this surface can also be computed to indicate the amount of variability at each point on the surface for a selected class LM along the identical axes of proportion of the category and the degree of spatial autocorrelation. The example below is for the ninth class metric: largest patch index for the black (or LOW) category

```
> tempvar <- apply(surfaces[9,,,], MARGIN=c(1,2), var)
> persp(tempvar, ticktype="detailed", cex.axis=0.7, zlab="Metric",
ylab="Proportion", xlab="Rho", theta=-45)
```

Step 2: Drop a Point Onto the Perspective Plot Indicating Observed Landscape's Position

Draw the perspective plot with a point indicating the location of a specific proportion and ρ relative to a single metric (Figure 7.6).

```
> surfplot(surfaceobj=tempmed, prop=0.72, rho=0.49, colour=TRUE,
drop=TRUE, cross=FALSE)
```

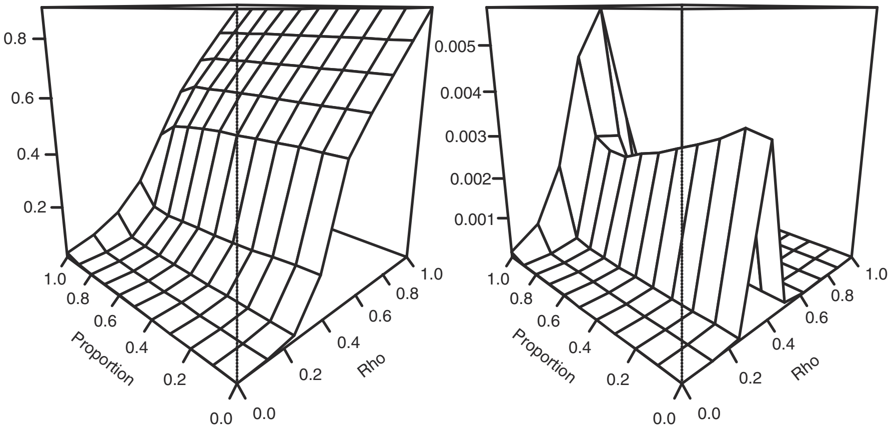


Figure 7.6 Surface response for class-level metric (z-axes; *left panel*: median, stored in `tempmed`; *right panel*: variance, stored in `tempvar`) according to proportion and spatial autocorrelation. These indicate how the expected value (median) and the variability (variance) of the largest patch index for the black (or LOW) category behave as composition and configuration parameters are incrementally changed

Step 3: Produce Boxplots of Surface Variability Along Axes Crossing Through the Observed Landscape Position Within the Perspective Plot

Include boxplots for Step 1 at the level of the observed proportion and spatial autocorrelation (ρ_W) axes on the surface to depict the variability (to do so set `cross=TRUE`).

```
> surfplot(surfaceobj=tempmed, prop=0.72, rho=0.49, colour=TRUE,
drop=TRUE, cross=TRUE)
```

Step 4: Perspective Plot with Both Median and Variance of Selected LM

Start by plotting the perspective plot of the median value of the class-level metric. This surface shows the variance (in color) indicating the variability at each point on the surface along axes of proportion and the degree of spatial autocorrelation (ρ_W).

The example below (Figure 7.7) demonstrates the ninth class metric: largest patch index for the LOW category. To color the median surface based on the variance values, the library *fields* must be installed and loaded along with dependencies.

```
> library(fields)
```

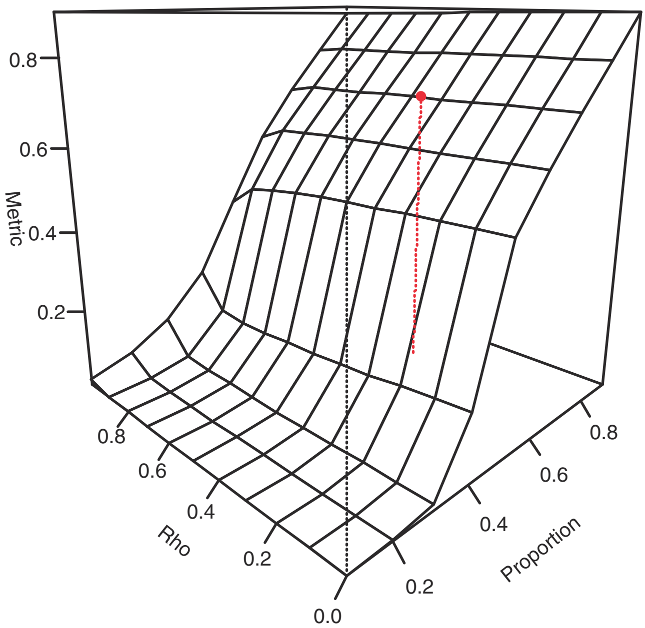


Figure 7.7 Surface response of the class-level metric median according to incremental variation of proportion and spatial autocorrelation for the category largest patch index as computed for the LOW value. The dropped red point indicates the position of the observed landscape within the range of variation that varying proportion and spatial autocorrelation can produce

Then, the plot can be produced:

```
> drape.plot(seq(0.1,0.9,by=0.1), seq(0,0.2499999, by=0.2499999/10)*4,
tempmed, tempvar, ticktype="detailed", col=topo.colors(50), theta=-25,
phi=15, cex.axis=0.5, xlab="Proportion", ylab="Rho", zlab="Metric")
```

Step 5: Additional Lab Exercises

Try this exercise using different class-level metrics (e.g., select five new class-level LMs; see Table 7.1) and different landscape datasets (e.g., `demoimage1`, `demoimage2`, `demoimage3`, and `demoimage4`).

Q8 Given the observed class-level LM value (i.e., the red dot in Figure 7.8) for the largest patch index within the context of all possible values of composition and spatial autocorrelation (Figures 7.8 and 7.9), what strategy (or strategies) would you recommend to reduce fragmentation (i.e., to increase the largest patch

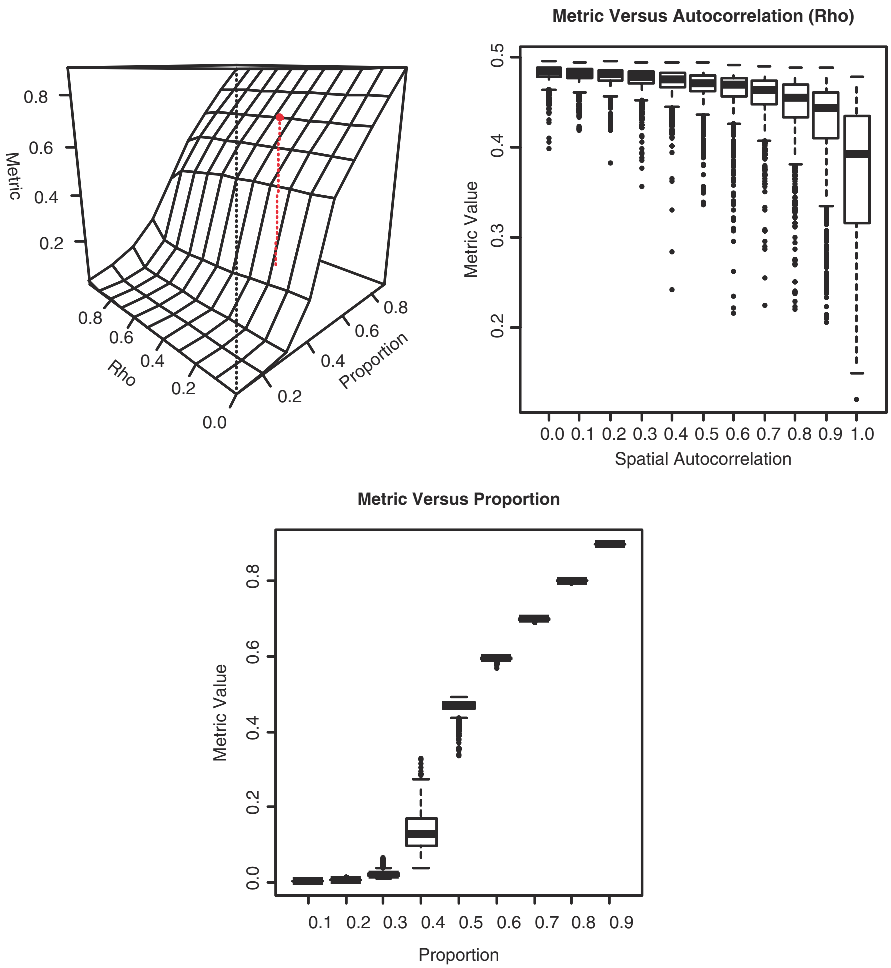


Figure 7.8 Class-level metric values according to category proportion and spatial autocorrelation: *left panel* as in Figure 7.6; middle panel: boxplot across spatial autocorrelation for the observed proportion level; *right panel*: boxplot across proportion level for the observed degree of spatial autocorrelation

index value)? Would it be easier to adjust the proportion or the level of spatial autocorrelation to achieve the goal? Imagine adding or removing land cover (either the white or black category) or simply rearranging them. Along which axis (composition or spatial autocorrelation) would LM values change more rapidly? How much change would be required to effect a 0.1 change to the LM value?

Q9 Repeat the previous question by examining another LM. You will need to begin at step 1 and adjust the value for the LM from 9 to the metric you select (Table 7.1).

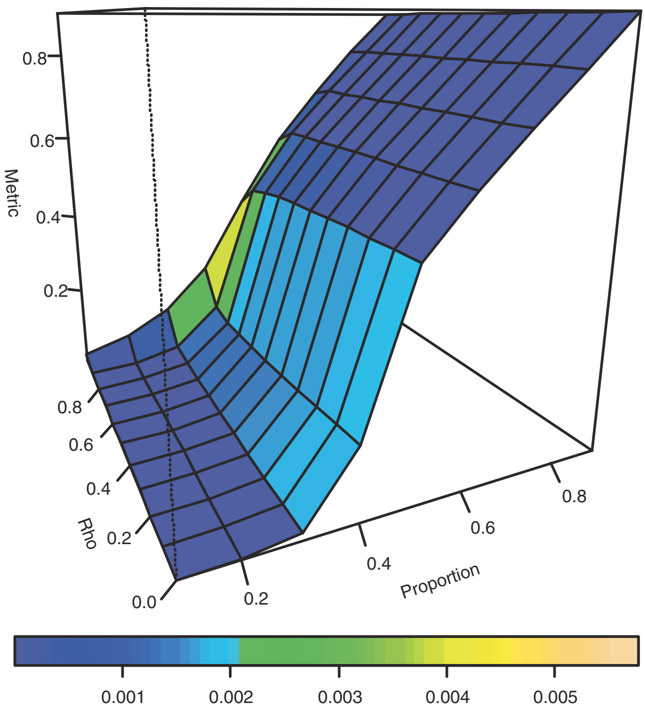


Figure 7.9 Class-level metric surface (as in Figure 7.8) combining the median LM value with the variability shown in colour across the ranges of proportion and configuration

Q10 Note and describe how the generated surfaces differ (along with the cross-sectional boxplots). Indicate the easiest way to decrease that LM value; would it be easier to alter composition or spatial autocorrelation?

CONCLUSIONS

While landscape metrics have been computed for decades, a major issue revolves around the difficulty in testing their significance. By using a simulation approach as presented here, it is now possible to assess whether or not the observed spatial structure of the landscape could have occurred by chance alone. Assessing the significance of LM is a major step in relating spatial structure to underlying process(es) that generated it. Furthermore, the knowledge of where in the range of proportion and spatial autocorrelation observed LM values lie can be used to propose restoration strategies for conservation purposes.

Acknowledgements We appreciate the contributions from Drs. Sándor Kabos and Ferko Csillag for their roles in providing some of the earlier programming work that we expanded upon. The chapter benefited from comments by R.H. Gardner, S.E. Gergel, M.G. Turner, and an anonymous reviewer.

APPENDIX A. EXPLANATION OF WHITTLE'S ALGORITHM (WHITTLE 1954)

For continuous data, ρ can be estimated using Whittle's algorithm (Whittle 1954) that extends the convention of time-series analysis to spatial processes reflected as collections of linear transects in geographic space. There is however the chance of bias in the estimated value of spatial autocorrelation when applying this algorithm to categorical data. This bias varies according to the composition, pi, such that around an even proportion of two classes, the bias is relatively small; however, it can be quite strong when the proportions differ greatly. Therefore, a correction factor needs to be applied to adjust the spatial autocorrelation estimate, resulting in the "true" ρ_W for categorical data. This true ρ_W needs to be multiplied by 4 to compensate for the isotropy of the algorithm and to scale the estimation to a range between 0 and 1. As this correction factor requires intensive computation, it has already been performed and stored as a lookup table in the provided Remmel–Fortin code as object DIFF50 and is used internally when estimating ρ_W.

REFERENCES AND RECOMMENDED READINGS[1]

Chambers JM, Cleveland WS, Kleiner B et al (1983) *Graphical methods for data analysis*. Wadsworth & Brooks/Cole, Pacific Grove

Cressie NAC (1993) *Statistics for spatial data*, revised edn. Wiley, New York

*Fortin MJ, Boots B, Csillag F et al (2003) On the role of spatial stochastic models in understanding landscape indices in ecology. Oikos 102:203–212. *This paper puts into perspective the use of stochastic models and simulations for studying complex ecological systems, specifically spatial patterns. The authors present a theoretical framework for the implementation of stochastic models to examine the confidence intervals around various landscape metrics.*

Fortin MJ, Dale MRT (2005) *Spatial analysis: a guide for ecologists*. Cambridge University Press, Cambridge

Fortin MJ, Jacquez GM, Shipley B (2012) Computer-intense methods. In: El-Shaarawi AH, Piegorsch WW (eds) *Encyclopedia of environmetrics*, 2nd edn. Wiley, Chichester, pp 489–493

Gardner RH, Urban DL (2007) Neutral models for testing landscape hypotheses. Landsc Ecol 22:15–29

*Haines-Young R, Chopping M (1996) Quantifying landscape structure: a review of landscape indices and their application to forested landscapes. Progr Phys Geogr 20(4):418–445. *One of few papers that provide a balanced summary of landscape pattern indices available for assessing spatial patterns. The authors elucidate uses and constraints on various metrics.*

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Hargrove WW, Hoffman FM, Schwartz PM (2002) A fractal landscape realize for generating synthetic maps. Conserv Ecol 6(1):2

Manly BFJ (2006) *Randomization, bootstrap and Monte Carlo methods in biology*. 3rd edn. CRC Press: Boca Raton, Florida

McGarigal K (2002) Landscape pattern metrics. In: El-Shaarawi AH, Piegorsch WW (eds) *Encyclopedia of environmetrics*, vol 2. Wiley, Chichester, pp 1135–1142

McGarigal K, Marks BJ (1995) FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. Gen. Tech. Rep. PNW-GTR-351, p 122

Moran PAP (1950) Notes on continuous stochastic phenomena. Biometrika 37(1/2):17–23

Proulx R, Fahrig L (2010) Detecting human-driven deviations from trajectories in landscape composition and configuration. Landsc Ecol 25:1479–1487

R Development Core Team (2010) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R--project.org/

*Remmel TK, Csillag F (2003) When are two landscape pattern indices significantly different? J Geogr Syst 5(4):331–351. *The authors present an extensive simulation experiment that demonstrates how empirical distributions of landscape metrics can be used for statistically comparing spatial patterns expressed by binary gridded maps. The procedure improves on traditional hypothesis testes that assume normally distributed values for landscape metrics.*

*Remmel TK, Fortin M-J (2013) Categorical class map patterns: characterization and comparison. Landsc Ecol 28(8):1587–1599. *Provides much of the theoretical and statistical background and detail related to the exercises in this chapter. Here, the authors explore the use of landscape simulations to compare binary landscape patterns using class-level metrics.*

*Riitters KH, O'Neill RV, Hunsaker CT et al (1995) A factor analysis of landscape pattern and structure metrics. Landsc Ecol 10(1):23–39. *This keystone paper summarizes the redundancies present among many landscape pattern metrics and proposes new compound variables for characterizing landscape patterns. The new variables are produced by ordination of many commonly computed metrics.*

Turner MG, Gardner RH, O'Neill RV (2001) *Landscape ecology in theory and practice*. Springer, New York

Van Der Wal J, Falconi L, Januchowski S et al (2011) SDMTools: species distribution modelling tools: tools for processing data associated with species distribution modelling exercises. 1.1–5. The Comprehensive R Archive Network (CRAN). GPL 2.0, http://cran.r-project.org/

Whittle P (1954) On stationary processes in the plane. Biometrika 41:434–449

*Wu JG, Shen WJ, Sun WZ et al (2002) Empirical patterns of the effects of changing scale on landscape metrics. Landsc Ecol 17(8):761–782. *This pivotal paper examines the effect of spatial scale on landscape metrics. While scale is an important concept in ecology, the effects of grain size and extent are not always clearly understood in the analysis of spatial patterns.*

Part III
Landscape Change and Disturbance

This module provides a compelling sequence of activities designed to solidify your basic knowledge of landscape change by providing exposure to landscape models and advanced methods for analyzing disturbance mosaics. Chapter 8 presents a simple approach for Markov landscape change modeling with a well-loved "favorite" from the first edition. Use of the Markov.exe program (a stand-alone DOS window executable) still works well, and many students find it a quaint throwback; however, we now include additional (optional) capability using R software. Chapter 9 presents Harvest Lite, a fun simulation model allowing the user to explore the fundamentals of spatial modeling and sensitivity analysis with a forest harvest model that incorporates edge effects and forest recovery. Two new advanced labs explore new modes of examining landscape disturbance. Chapter 10 guides students through analyses of replicated landscapes across continental scales, which can improve comparative studies of landscape dynamics, with a crafty web interface for the METALAND tool. Lastly, Chapter 11 explores the rich mosaic of landscape patterns created by three different types of disturbance and contrasts categorical and continuous measures of disturbance (i.e., landscape metrics and spatial statistics). Patterns produced by fire, insect outbreaks, and forest harvest are explored with the user-friendly software package GS+, building on knowledge from prior chapters on spatial statistics and landscape metrics, but with the complexity and subtleties of realistic landscape data.

Chapter 8
Introduction to Markov Models

Dean L. Urban and David O. Wallin

OBJECTIVES

Models of landscape change are important tools for understanding the forces that shape landscapes. One motivation for modeling is to examine the implications of extrapolating short-term landscape dynamics over the longer term. This extrapolation of the status quo can serve as a frame of reference against which to assess alternative management scenarios or test hypotheses. There is a spectrum of ways to consider landscape change, ranging from simple and readily interpretable, to more realistic and less tractable. The goals of this lab are to:

1. Provide an introduction to the mathematics of simple Markov models;
2. Enable students to build a simple model of land use change based on transition probabilities;
3. Explore the process of model creation, verification, and validation; and
4. Encourage creative speculation as to how Markov models might be extended to incorporate more complex and realistic mechanisms of and constraints on landscape change.

In this exercise, you will build a simple model of landscape change, evaluate it, and use it as a point of departure to consider more realistic (but more complicated) models. Raster maps of Pacific Northwest forests are compared over three time periods to summarize the rates of transition between cover types. A simple model of

D.L. Urban (✉)
Duke University, Durham, NC, USA
e-mail: deanu@duke.edu

D.O. Wallin
Western Washington University, Bellingham, WA, USA

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_8

landscape change is built from these transition probabilities. Lastly, this model is projected forward in time to verify and validate the model. In order to complete this lab, you will need a PC, the program **markov.exe**, the landscape images (**pnw72.gif**, **pnw85.gif**, **pnw91.gif,** and **samp200.gif**) and the accompanying data file (**samp200**) which can be obtained from the book's website.

INTRODUCTION

Perhaps the most fundamental observation on landscape change arises from measurements of the state of a landscape at two time periods. For example, we might have land cover maps classified from satellite images obtained for two dates 10 years apart, and note that some of the cells (pixels) changed type over that time interval.

One way to summarize landscape change is to simply tally all the instances, on a cell-by-cell basis, in which a cell changed cover types over that time interval. A concise way of summarizing these tallies is a raw **tally matrix**, which for m cover types is an m x m matrix. The elements, n_{ij}, of the tally matrix tally the number of cells that changed from type i to type j over a time interval. A raw tally matrix is often converted into proportions by dividing each of the elements by the row total to generate a **transition matrix P**. The elements, p_{ij}, of the transition matrix P summarize the proportion of cells of each cover type that changed into each other cover type during that time interval. The diagonal elements of this matrix, p_{ii}, are the proportions of cells that did not change.

While there are a variety of approaches to modeling landscape change (see Weinstein and Shugart 1983; Baker 1989; Sklar and Costanza 1991; Mladenoff and Baker 1999, for reviews), many of these begin with a tally matrix or the transition matrix **P**. Here, you will examine the simplest of such models based on a transition matrix. This simple model will serve as a point of departure for contemplation of more realistic but more complicated models.

Markov Models

A **first-order Markov model** (Usher 1992) assumes that to predict the state of the system at time $t+1$, one need only know the state of the system at time t. The heart of a Markov model is the **transition matrix P**, which summarizes the probability that a cell in cover type i will change to cover type j during a single timestep. The timestep is the interval over which the data were observed to change (i.e., the time interval of the two maps).

Markov models, while simple, have a number of appealing properties. In particular, they can be solved by iteration to project the state of the system. Writing the state of the system as a vector

$$x_t = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots \end{bmatrix} \quad (1)$$

where x_i is the proportion of cells in type *i* at time *t*, a Markov model is projected:

$$x_{t+1} = x_t P \quad (2)$$

that is, the state vector post-multiplied by the transition matrix. The next projection, for time $t+2$ is continued:

$$x_{t+2} = x_{t+1} P = x_t PP = x_t P^2 \quad (3)$$

and in general, the state of the system at time $t=t+k$ is given by:

$$x_{t+k} = x_t P^k \quad (4)$$

where x_t is the initial condition of the map. Thus, the model can be projected into the future simply by iterating through the matrix operation (see Exercise 2 for details on how to do this manually).

The **steady-state** or equilibrium state of the system is given by the eigenvector of the transition matrix; thus, there is a closed-form solution to the model. Recall, the **eigenvector** of the matrix is defined such that the matrix multiplied by the eigenvector yields the vector again:

$$\tilde{x} = \tilde{x} P \quad (5)$$

That is, the system does not change once it reaches this state. There are some computational tricks for estimating steady-state solutions (Usher 1992), or you could use a math package (e.g., Mathematica™, MatLab™) to do this. But for simple models, the solution often converges rapidly and you can estimate the solution simply by projecting the model a few times.

Graphical Representation. The model implied by the transition matrix **P** can also be represented as a graph (a "box-and-arrow" diagram). An example with three cover types could be illustrated as in Figure 8.1. Casual inspection of the graph reveals the direction of flow in the system, and suggests a succession from type 1, through type 2 to type 3, with some recycling (possibly a disturbance) to the initial cover type.

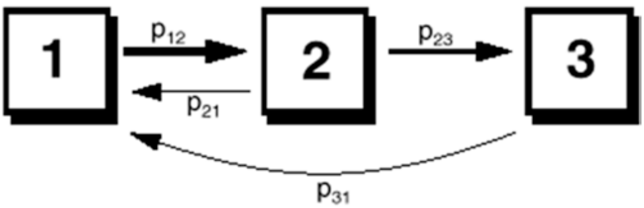


Figure 8.1 A schematic box-and-arrow diagram of a transition matrix **P** with three cover types. The thickness of the *arrows* indicates the magnitude of the transition rates between the different cover types (the *arrows* for self-replacement are not shown). The diagram shows flow from cover type 1 to 2, to 3, with some recycling to previous cover types via disturbance

Model Projection. To explore a Markov model, it is initialized with a state vector and then projected for one or more timesteps. The vector of cover types produced at each iteration is the prediction of overall landscape composition for that timestep. In the following exercises, we will show you how this is accomplished.

EXERCISES

Modeling Landscape Change in the Pacific Northwest, USA

Much debate over the management of Pacific Northwestern (PNW) forests occurred in the early 1990s. The debate centered on the effects of intensive logging on old-growth forests and on old-growth-dependent wildlife species such as the northern spotted owl (*Strix occidentalis*) and the marbled murrelet (*Brachyramphus marmoratus*) (Hansen et al. 1991; Ruggiero et al. 1991), primarily on U.S. Forest Service land. Since most of the old-growth on private lands was already gone, most of the old-growth harvested in the 1980s came from these federal lands (Harris 1984; Robbins 1988). By the mid-1990s, harvests were reduced by over 90% on federal lands relative to the peak harvests of the late 1980s (FEMAT 1993; USDA Forest Service and USDI Bureau of Land Management 1994; Marcot and Thomas 1997). A central question during this debate was, *How long can current rates of harvest be sustained before the old growth is virtually gone*?

The reduction in harvest levels in the early 1990s has reduced the urgency of this question for the PNW (Moeur et al. 2011), but the relevance of being able to assess trajectories of change is still timely. In PNW forests, the assessment of landscape change has expanded from the initial emphasis on stand-replacing disturbances such as timber harvesting and wildfire (Cohen et al. 2002), to partial disturbances (Healey et al. 2006) and efforts to quantify variation in the rate of regrowth following disturbance (Schroeder et al. 2006, 2007). Initially, these analyses relied on the detection of a difference between pairs of Landsat imagery at intervals of approximately 5 years. More recently, Kennedy et al. (2007) have developed an automated approach that is based on the use of a long time series (>20 years) with an annual timestep to describe trajectories of change in forested environments. This approach is based on the notion that many natural and anthropogenic phenomena in forests can be identified on the basis of unique trajectories both before and after the event. This approach is being used to identify a much wider range of disturbance processes, including insect outbreaks, disease, and windthrow and also provides a more nuanced description of gradients in disturbance severity (tree retention levels following partial harvest; variation in tree mortality following wildfires).

Study Area. The Oregon Cascades were at the center of this debate over the management of PNW forests. The study area is on federally managed lands where timber harvesting has been conducted using a dispersed (staggered setting) system of

Table 8.1 Definition of cover types in the Pacific Northwestern forest landscape

Class	Age (yr)	Cover type
0	Background	(Non forest)
1	0–20	Recent clear-cut
2	21–40	Early seral
3	41–80	Mid seral
4	81–170	Mature
5	>170	Old growth

10–20 ha clear-cut patches. The rate and pattern of these disturbances is somewhat different than those on private lands (Spies et al. 1994) and is quite different than disturbances generated by wildfire during the presettlement era (Wallin et al. 1996b).

Spatial data for this area were derived from Landsat Thematic Mapper data using methods outlined in Cohen et al. (1995, 1998, 2002) and Wallin et al. (1996a). Forest cover was classified into six approximate age classes (Table 8.1). Images for three time periods are included here: 1972, 1984, and 1991. The images for the different time periods (**pnw72.gif**, **pnw84.gif,** and **pnw91.gif**) can be examined by using a web browser (use **File**, then **Open** on your web browser's pull-down menu). Yellow areas denote young stands, successively darker greens are older forest, and brown areas are recent clear-cuts. The gray spot in the images is rock. Each image is 500 × 500 cells (15,625 ha) with a cell size of 25 m.

The file **samp200.gif** shows 200 random locations on the 1972 image. The cover type at each of these locations was tallied at time 1 (1972), 2 (1984), and 3 (1991) on the three maps. This information is tallied in columns of a **primary data matrix** and can be viewed by opening the text file **samp200.dat**.

Next, you will use the cover type data from the 200 sample points to build your own Markov model. In the following exercises, you will perform three main steps: model development, model verification, and model validation.

EXERCISE 1: Model Development

Here, you will use the data from the 200 sample points on the PNW images to calculate transition probabilities for a Markov model. The transition probabilities will be based on landscape change from 1971 to 1984.

1. From the primary data matrix (**samp200.dat**), construct a **raw tally matrix** that summarizes the number of the 200 cells that underwent a transition from type i to type j during the time period t_1 (1972) to t_2 (1984). Recall that each element, n_{ij}, in the tally matrix is the number of times where a cell changed from type i to type j during the time interval. Enter your results in Table 8.2 (*NOTE:* You may also wish to work on this in a spreadsheet using the file **samp200.xls**).

Table 8.2 Raw tally matrix (1972–1984)

	To (*j*): 1	2	3	4	5
From (*i*): 1					
2					
3					
4					
5					

Table 8.3 Transition probability matrix (P) over a **12-year** timestep

	To (*j*): 1	2	3	4	5
From (*i*): 1					
2					
3					
4					
5					

To avoid rounding errors later, perform your calculations to five decimal places

2. Divide each element in the raw tally matrix by its row total to yield a matrix of transition probabilities p_{ij}, the probability (or rate) of change from type *i* to *j*. These probabilities are on a **12-year timestep** (1972–1984). Enter your results in Table 8.3.
3. Next, convert the transition matrix to an **annual timestep**. This is partly cosmetic (transient dynamics will look smoother), but will also make it possible to reconcile the 12-year timestep of the first period with the 7-year timestep of 1984–1991. To convert the transition probability matrix P to an annual timestep, do the following:
   a. Divide each of the **off-diagonal** elements p_{ij}, $i \neq j$, by 12.
   b. Adjust the **diagonal** elements p_{ii}, to be $1.0 - \sum_j p_{ij}$

   In other words, all rows must sum to 1.0. Enter the results in Table 8.4.

In this matrix (Table 8.4), the off-diagonal rates are now annual transitions (probabilities). The diagonal elements are now larger than in Table 8.3 because on an annual timestep fewer of the cells actually change, and again, the rows of the matrix still must sum to 1.0.

4. To finish building the model, summarize the state of the map at each timestep. The state of the map is defined by a row vector, the elements of which are the proportion of cells in each cover type in each of the maps. Construct three summary vectors from the primary data matrix (**samp200.dat**). To construct these vectors, simply sum up all of the cells in each state in each of the 3 years of sampling. Convert these numbers to proportions by dividing each element of the table by its row sum. (These vectors can also be derived from the transition matrices. How?). Enter your results in Table 8.5.

With these state vector tallies and the transition probability matrix **P**, you have constructed a simple model of landscape change. All that remains is to evaluate the

Table 8.4 Transition probability matrix (P) over an annual timestep

	To (j): 1	2	3	4	5
From (i): 1					
2					
3					
4					
5					

The calculations along the diagonal cells (the probabilities of remaining the same) must be calculated last

Table 8.5 Summary vectors at three timesteps

	Class 1	2	3	4	5
1972					
1984					
1991					

Remember, enter your data as proportions

model. The 1972 data will be used as initial conditions for the model. The second time period (1984) will be used to verify the model. The third time period (1991) will be reserved to validate the model.

EXERCISE 2: Model Verification Via Matrix Projection By Hand

Model verification consists of testing the model against the data used to construct it (Haefner 1996). In this case, the test is of the model projection from 1972 to 1984, compared to the actual data from 1984. Because the model was built from these data, this is *not* an independent test of the model; the model *should*, in fact, match these data. You will verify the model by initializing it with 1972 data and projecting it to 1984, both by hand (here) and by using a computer program MARKOV in Exercise 3. Do the first matrix projection (from 1972 to 1984) by hand following the example below. To do this in one iteration, use the transition probabilities in Table 8.3, which are for a single 12-year timestep.

This example shows a generic matrix projection using only three cover types. The first step entails multiplying the transition probability matrix **P** by the state vector.

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \cdot \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} =$$
$$\begin{bmatrix} x_1 p_{11} + x_2 p_{21} + x_3 p_{31}, & x_1 p_{12} + x_2 p_{22} + x_3 p_{32}, & x_1 p_{13} + x_2 p_{23} + x_3 p_{33} \end{bmatrix}$$

Note the subscripts to ensure that the proper elements are being used: the inner subscripts must always match (i.e., the subscript on x must match the first subscript on p).

An example using actual data with **only three cover types** would look like this.

a. Suppose the following transition probability matrix (**P**) is used:

$$P = \begin{bmatrix} 0.90 & 0.10 & 0.00 \\ 0.10 & 0.80 & 0.10 \\ 0.10 & 0.00 & 0.90 \end{bmatrix}$$

b. Next, suppose all of the landscape is assigned to cover type 1 to begin:

$$x_0 = \begin{bmatrix} 1.00 & 0.00 & 0.00 \end{bmatrix}$$

c. The first projection to t_1 is:

$$\begin{bmatrix} 1.0 \cdot 0.9 + 0.0 \cdot 0.1 + 0.0 \cdot 0.1, & 1.0 \cdot 0.1 + 0.0 \cdot 0.8 + 0.0 \cdot 0.0, & 1.0 \cdot 0.0 + 0.0 \cdot 0.1 + 0.0 \cdot 0.9 \end{bmatrix} = \begin{bmatrix} 0.9 & 0.1 & 0.0 \end{bmatrix}$$

The second projection to t_2 uses the resulting vector:

$$\begin{bmatrix} 0.9 & 0.1 & 0.0 \end{bmatrix}$$

and the original transition probability matrix, and produces:

$$\begin{bmatrix} 0.82 & 0.17 & 0.01 \end{bmatrix}$$

You should verify this by hand. Also, note that these same data are used in the demo data file supplied with the lab (**demo.txt**).

EXERCISE 3: Matrix Projection Using the Program Markov

Not surprisingly, matrix projections are often accomplished with the use of a computer program. Here, you will use a simple Fortran program called Markov. The program (**markov.exe**) can be run from a DOS prompt or by double-clicking directly on its icon. Demo input data are available (**demo.txt**).

MODEL INPUT

Markov expects to read a user-provided ASCII data file containing the transition matrix and a vector of initial conditions. For these exercises (with 5 cover types), these must be formatted as follows:

- Rows 1–5: the elements of the transition matrix (from Table 8.4)
- Row 6: the initial conditions (the 1972 row from Table 8.5)

The data values themselves can be delimited by spaces, a comma, or tabs, and you should use enough significant digits to minimize round-off error (say, 5 decimal places).

MODEL OUTPUT

The output written by MARKOV consists of one line per timestep, reporting timestep (in column 1) followed the proportion of the landscape in each cover type at that time (in this case, columns 2–6). The output file from MARKOV is formatted so that it can be imported directly into a spreadsheet or graphics package. The program will report either the timestep at which the solution converged to steady-state, or that the model did not converge during the simulation. In the latter case, you need to rerun the model for a longer time so that it has time to converge.

A session with MARKOV is as follows:

```
Project a markov Model?
 Name of file with input data?
 And name of file for output data?
 Number of patch types in model?
 And number of timesteps to project?
Model failed to converge in 100 timesteps
```

When asked for the number of patch types, this refers to the number of cover types in the model.

Using the program MARKOV, repeat your model projection from 1972 to 1984 using the following steps:

1. Make an input file (a text file) that includes your matrix and vector data. Format your input data as explained in the *MODEL INPUT* section. (*NOTE*: you can refer to the file **demo.txt** for guidance, but remember that these data are for the example with only three cover types.)
2. Run the program MARKOV.
3. Compare your results from Exercises 1 and 2. Does the model projection reproduce the data used to build the model? If it doesn't reproduce the 1984 data, what might explain the discrepancy?

HINTS: When running Markov, enter the full file name (including the filename extension). Be sure your input file is not open in some other program. Be sure your input file is in same directory as Markov.exe. Only use a simple text editor (like WordPad or NotePad) to edit your input files (don't use Word or Excel).

EXERCISE 4: Model Validation

Model validation consists of testing a model against data that were *not* used to construct the model (Haefner 1996). This is important as it is an *independent* test of the model.

1. Still using the 1972 data as your initial condition vector, use the program MARKOV to project the model to 1991.
2. Compare the predicted landscape composition to the actual composition tallied from the primary data table. This is a test to **validate** your model. Does the model projection match the 1991 data? If not, what might explain the discrepancy?
3. Continue to project the model into the future, until it converges to a steady-state or until there is less than 10% of the landscape in old-growth forest, whichever comes first. How long will the old-growth last, or when will it equilibrate?

NOTE: You can complete all of these tasks with a single projection of the model. Simply run the model for a very long time (say, 1000 years). If it converges in less time, simply delete the extraneous years from the model output file using a text editor (the output is only interesting while the landscape is changing).

CONCLUSIONS

This concludes the development, verification, and validation of a simple Markov model of landscape change. In some applications, such a simple model is sufficient (e.g., Johnson and Sharpe 1976; Johnson 1977; Hall et al. 1991). But in many cases, this simple model serves as a point of departure for more complicated models (e.g., Turner 1987; Baker 1989; Acevedo et al. 1995, 1996; Wear and Bolstad 1998; Wu and Webster 1998; Hong and Mladenoff 1999a, b; Mladenoff and Baker 1999; Urban et al. 1999). In particular, some consideration of the assumptions and limitations of Markov models can be a useful aid in interpreting the behavior and predictions of other models. These considerations are presented next.

Further Considerations in Modeling Landscape Change

The simple Markov model serves as a useful point of departure for more complicated issues in landscape modeling. Several of these are especially relevant to landscape dynamics.

Stochasticity. Models such as the one you have created are often used to project a map into the future by changing each cell in the map according to the transition probabilities. Because each cell can only change into one other state (a cell can't change fractionally), the state changes must be done probabilistically. This can create some new problems, as the new map is only one of many possible stochastic realizations of a new map (since the map was created probabilistically). Thus, any comparison to a real map would have to be based on a number of replicate maps. Note that, in the limit, the average of a large number of stochastic simulations would be a map of the transition probabilities themselves.

Importance of history. Another issue concerns the assumption that to predict the future state of a system one need only know its current state. In cases where this is true, the process is truly first-order, also known as a Markov chain. In reality, there may be cases where information about additional prior states is needed. These cases would lead to higher order Markov models (e.g., in a second-order model, one would need to know the state of the system at time t and $t-1$ to predict its state at time $t+1$). Systems with even longer "memory" would require still higher order models. For cases where the memory is very long, it might become more convenient to envision (and model) the dynamics in terms of a "time since" variable, such as "time since abandonment" or "time since disturbance," instead of keeping track of a large number of previous states.

Stationarity. Because we have three maps, two first-order Markov models could be derived from our study landscape. We could derive transition matrices from 1972 to 1984, and another from 1984 to 1991. There is a formal test for stationarity of these matrices (Usher 1992); nonstationary transition probabilities would vary between time periods. Nonstationary transition matrices would suggest the forces (or rules) governing landscape change were changing over time. Certainly, the drivers of landscape change might vary through time in regions with historical variation in socioeconomic drivers, true of most of the United States over the past several decades.

In the case of nonstationary transition rules, two alternatives are possible. In the discrete case, separate transition matrices can be computed for each time period of interest. For example, given a sequence of airphotos taken every 10 years for 50 years, one could derive four separate transition matrices. Each matrix would be used to project from one time period to the next. Alternatively, the transitions could be specified explicitly as functions of time, so that the rules governing landscape change would vary with time. This approach would generate "smoother" dynamics, but would require some sort of curve-fitting for the time functions (as well as the data to support that curve-fitting!).

Spatial dependencies. A fourth complication arises if some of the transitions appear to have spatial dependencies. For example, certain kinds of transitions might tend to occur in certain topographic settings or in certain spatial configurations as defined by a cell's immediate neighbors. These complications drive the modeling approach away from a simple Markov framework, toward models where the transition probabilities depend not only on the current state of the system, but also on some other stated conditions. That is, the transition matrix contains **conditional probabilities** such as "if the cell is type i and meets condition k, then its probability of becoming type j is $p_{ij|k}$." The condition might relate to site conditions (e.g., soil) or neighborhood effects (e.g., contagious disturbance). It is a relatively straightforward procedure to tally transition matrices as conditional probabilities. One simply constructs a multi-layered tally matrix analogous to Table 8.4, but incorporating all of the special conditions of interest. Clearly, this can become extremely data-hungry.

In any of these more complicated models of landscape change, the ability to solve the model analytically is rapidly lost; thus, complex models must be "solved" by iteration to steady-state (if such a state exists). This trade-off between simplicity and realism is common to all modeling efforts.

WRITE-UP

Your lab write-up should include the following sections:

1. The **Introduction** should state the motives of the exercise and provides some context. For example, why would we want to model landscape change?
2. The **Background** section should focus on the conceptual basis and assumptions of a first-order Markov chain as a model of landscape change. Address the following issues:
   a. Try to present a Markov model, in a narrative sense, as clearly and concisely as possible. How does it work?
   b. What assumptions do we make, implicitly or explicitly, in using such a simple model?
   c. Given the assumptions and the simplicity of such models, why use them at all? That is, what is the value of simple models in assessing landscape change?
3. The **Methods** should reiterate, concisely, the steps you followed to generate the model. This section could culminate in a presentation of the model as a transition matrix and as a graph, along with a table with the state vectors.
4. **Results** should consist of:
   a. The projection of your landscape from 1972 to 1991, and your comparison of the model projection to the actual data for 1984 and 1991. You should be able to include all of this in one figure. You need NOT include the actual model output in tabular form.
   b. Include your hand calculations for the 1-year projection.
5. The **Discussion** should address the following questions:
   a. How well does the model projection match the actual 1991 data? If it doesn't match, what possible reasons might you suggest for the discrepancy?
   b. How would you address these discrepancies (what changes to the model or what additional data would you need)?
   c. Would you expect the landscape to ever reach a steady-state? Of what interpretative value is the model solution (i.e., the steady-state composition of the landscape)?
   d. What would it take to maintain 20% of the landscape in old growth? That is, which transition rates would have to change, and how much?

Acknowledgements The classified imagery used in this lab was developed under funding provided by the USDA Forest Service New Perspectives Program, the National Sciences Foundation through the H.J. Andrews Long-Term Ecological Research (LTER) site (Grant 90-11663) and the Terrestrial Ecology Program of the National Aeronautics and Space Administration (NAGW-3745).

REFERENCES AND RECOMMENDED READINGS[1]

Acevedo M, Urban DL, Ablan M (1995) Transition and gap models of forest dynamics. Ecol Appl 5:1040–1055

Acevedo MF, Urban DL, Shugart HH (1996) Models of forest dynamics based on roles of tree species. Ecol Model 87:267–284

*Baker WL (1989) A review of models of landscape change. Landsc Ecol 2:111–133. *A general review and good introduction to a wider variety of models.*

Cohen WB, Spies TA, Fiorella M (1995) Estimating the age and structure of forests in a multiownership landscape of western Oregon, USA. Int J Remote Sens 16:721–746

Cohen WB, Fiorella M, Gray J et al (1998) An efficient and accurate method for mapping forest clearcuts in the Pacific Northwest using Landsat imagery. Photogramm Eng Remote Sens 64:293–300

Cohen WB, Spies TA, Alig RJ (2002) Characterizing 23 years (1972–95) of stand replacement disturbance in western Oregon forests with Landsat imagery. Ecosystems 5:122–137

FEMAT (1993) Forest ecosystem management: an ecological economic, and social assessment, Report of Forest Ecosystem Management Assessment Team, July 1993, Portland

*Haefner JW (1996) *Modeling biological systems: principles and applications*. Chapman and Hall, New York. *The introductory chapter and chapter 8 on model validation are great background reading.*

Hall FG, Botkin DB, Strebel DE (1991) Large-scale patterns of forest succession as determined by remote sensing. Ecology 72:628–640

Hansen AJ, Spies TA, Swanson FJ (1991) Conserving biodiversity in managed forests. Bioscience 41:382–392

Harris LD (1984) *The fragmented forest*. University of Chicago Press, Chicago

Healey SP, Yang Z, Cohen WB et al (2006) Application of two regression-based methods to estimate the effects of partial harvest on forest structure using Landsat data. Remote Sens Environ 101:115–126

Hong SH, Mladenoff DJ (1999a) Spatially explicit and stochastic simulation of forest- landscape fire disturbance and succession. Ecology 80:81–99

Hong SH, Mladenoff DJ (1999b) The effects of seed dispersal on the simulation of long-term forest landscape change. Ecosystems 2:308–319

Johnson WC (1977) A mathematical model of forest succession and land use for the North Carolina Piedmont. Bull Torrey Botan Club 104:334–346

Johnson WC, Sharpe DM (1976) An analysis of forest dynamics in the northern Georgia Piedmont. Forest Sci 22:307–322

Kennedy et al (2007) https://pubag.nal.usda.gov/pubag/downloadPDF.xhtml?id=38818&content=PDF

*Mladenoff DJ, Baker WL (eds) (1999) *Spatial modeling of forest landscape change: approaches and applications.* Cambridge University Press, Cambridge. *This book includes a good sampling of more recent approaches to modeling landscape dynamics.*

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Marcot BG, Thomas JW (1997) Of spotted owls, old growth, and new policies: a history since the Interagency Scientific Committee report. USDA Forest Service General Technical Report PNW-GTR-408. Pacific Northwest Forest and Range Experiment Station, Portland

Moeur M, Ohmann JL, Kennedy RE et al (2011) Status and trends of late-successional and old-growth forests. Northwest Forest Plan: The First 15 Years (1994–2008). General Technical Report PNW-GTR-853 http://www.fs.fed.us/pnw/pubs/pnw_gtr853.pdf

Robbins WG (1988) *Hard times in paradise, Coos Bay, Oregon, 1850–1986*. University of Washington Press, Seattle, p 194

Ruggiero LF, Aubry KB, Carey AB et al (1991) Wildlife and vegetation of unmanaged Douglas-fir Forests. USDA Forest Service General Technical Report PNW-GTR-285. Pacific Northwest Forest and Range Experiment Station, Portland

Schroeder TA, Cohen WB, Song C et al (2006) Radiometric calibration of Landsat data for characterization of early successional forest patterns in western Oregon. Remote Sens Environ 103:16–26

Schroeder TA, Cohen WB, Yang Z (2007) Patterns of forest regrowth following clear-cutting in western Oregon as determined from multitemporal Landsat data. For Ecol Manag 243(2–3): 259–273

Sklar FH, Costanza R (1991) The development of dynamic spatial models for landscape ecology: A review. In M. G. Turner and R. H. Gardner, eds. *Quantitative methods in landscape ecology*. Springer-Verlag, New York, pp. 239–288

Spies TA, Ripple WJ, Bradshaw GA (1994) Dynamics and pattern of a managed coniferous forest landscape in Oregon. Ecol Appl 4:555–568

Turner MG (1987) Spatial simulation of landscape changes in Georgia: a comparison of 3 transition models. Landsc Ecol 1:29–36

USDA Forest Service and USDI Bureau of Land Management (1994) *Record of decision for amendments for Forest Service and Bureau of Land Management planning documents within the range of the northern spotted owl*. USDA Forest Service and USDI Bureau of Land Management, Portland, p 74

Urban DL, Acevedo MF, Garman SL (1999) Scaling fine-scale processes to large-scale patterns using models derived from models: meta-models. In: Mladenoff D, Baker W (eds) *Spatial modeling of forest landscape change: approaches and applications*. Cambridge University Press, Cambridge, pp 70–98

*Usher MB (1992) Statistical models of succession. In: Glenn-Lewin DC, Peet RK, Veblen TT, (eds) *Plant succession: theory and prediction*. Chapman and Hall, London, pp 215–248. *This tutorial on Markov processes and statistical models is an excellent introduction to this type of model, and includes clearly worked examples as well as formal treatment of more formal analysis of Markov models.*

Wallin DO, Harmon ME, Cohen WB et al (1996a) Use of remote sensing to model land use effects on carbon flux in forests of the Pacific Northwest, USA. In: Gholz HL, Nakane K, Shimoda H (eds) *The use of remote sensing in the modeling of forest productivity at scales from the stand to the globe*. Kluwer Academic, Dordrecht, pp 219–237

Wallin DO, Swanson FJ, Marks B et al (1996b) Comparison of managed and pre-settlement landscape dynamics in forests of the Pacific Northwest, USA. For Ecol Manage 85:291–310

Wear DN, Bolstad P (1998) Land-use changes in southern Appalachian landscapes: spatial analysis and forecast evaluation. Ecosystems 1:575–594

Weinstein DA, Shugart HH (1983) Ecological modeling of landscape dynamics. In H. A. Mooney and M. Godron, eds. *Disturbance and ecosystems*. Springer-Verlag, New York, pp. 29–45

Wu F, Webster CJ (1998) Simulation of land development through the integration of cellular automata and multicriteria evaluation. Environ Plann B 25:103–126

Chapter 9
Simulating Management Actions and Their Effects on Forest Landscape Pattern

Eric J. Gustafson

OBJECTIVES

Landscapes are characterized by their structure (the spatial arrangement of landscape elements), their ecological function (how ecological processes operate within that structure), and the dynamics of change (disturbance and recovery). Thus, understanding the dynamic nature of landscapes and predicting their future dynamics are of particular emphasis. Landscape change is difficult to study because controlled experiments at landscape scales often are not feasible for political, economic, social and logistical reasons. Opportunistic studies of change (e.g., after a large fire) are often confounded by uncontrolled factors. For these reasons, changes in landscape pattern are often studied using simulation models. This lab will:

1. Introduce simulation modeling as an important tool of landscape ecology;
2. Show the utility of simulation models for examining landscape change at spatial and temporal scales that are not easily addressed using field methods;
3. Illustrate an applied use of simulation modeling in landscape ecology—examining changes in landscape pattern caused by timber management;
4. Discuss the assumptions and limitations of simulation models; and
5. Show how models can be used to answer questions about landscape pattern and landscape change.

This lab exercise focuses on landscape change produced by forest management, using a timber harvest simulation model. The model you will use is a simplified version of HARVEST (http://www.ncrs.fs.fed.us/4153/Harvest/v61/documentation/), which generates patterns similar to those produced by timber management

E.J. Gustafson (✉)
U.S. Forest Service, Northern Research Station, WI, USA
e-mail: egustafson@fs.fed.us

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_9

Gustafson and Crow (1999). The model allows you to change the size of timber harvest openings, the total area harvested, and the spatial distribution of harvested areas (whether harvests will be clumped or dispersed). You will determine how different harvest regimes influence the amount of forest interior, amount of forest edge, and the mean patch size of forests. The software and data files needed for this lab can be obtained from the book's website.

INTRODUCTION

Simulation Modeling

Science is a process of ruling out ideas that are *not* true, always leaving some uncertainty about the ideas we think *are* true. The things we accept as scientifically true are actually a collection of conceptual models of how we *believe* the world works and that have withstood multiple attempts to disprove them. When we formalize a conceptual model using mathematical relationships, we have constructed a simulation model capable of generating a prediction based on the initial conditions and the relationships formalized in the model. The utility of simulation models lies in their ability to show the consequences of assumptions as a result of variation in the input parameters. Model assumptions typically are based on an understanding of a process derived from empirical study of the process. Simulation models are built for varying purposes (Karplus 1983). Some are used because they have predictive capabilities (e.g., tree-growth models), some are used to improve our understanding of a newly developed theoretical model (e.g., metapopulation theory), and others illuminate how we might manage an ecological system (e.g., by timber harvest) to produce desired conditions.

Spatial simulation models specifically include the spatial arrangement of key elements of the system being studied. Simulation modeling is especially suited to answer general questions about the spatial implications of interacting processes, especially when manipulative experiments of many combinations of treatments are not feasible. Simulation models also allow control of effects that are difficult to control in empirical experiments. Although stochastic (i.e., based on a random process) spatial models may not be useful to predict the specific location of individual events, they can be used to generate replicate patterns with properties that vary in response to variation in the model inputs. These simulated patterns are assumed to be statistically indistinguishable from those that would be produced in the real world *if the real process behaves as the model has assumed*. Therefore, if comparison of the model results and empirical data reveal a significant difference, we can conclude that our model does not adequately simulate reality. Such a discrepancy also provides an opportunity to reexamine and revise underlying assumptions about the proposed mechanisms for the process that the model represents. Spatial modeling also allows identification of the parameters to which spatial pattern is most sensitive, focusing hypothesis testing and empirical model development. This use provides insight into the implications of the view of reality that is formalized in the model. As such it has heuristic value—that is, it helps clarify our thinking.

When using a simulation model it is critical to understand the sensitivity of model results to changes in the input parameters (Haefner 1996). Large changes in some parameters may have little effect on the model results, whereas small variation in other parameters may induce large effects. Alternatively, model results may be nonlinearly related to the magnitude of the parameter value. An understanding of these model properties is gained by systematically varying input parameters, a process known as **sensitivity analysis**. Here, you will conduct a limited sensitivity analysis of the HARVEST Lite model, and begin to understand the relationship between disturbance (timber management) and landscape pattern.

Because replicated landscape studies involving extensive removal of trees are generally not feasible, the study of how forest spatial pattern is affected by harvesting strategies is facilitated by spatial simulation models. The conceptual basis for landscape-scale simulation of harvest patterns can be traced back at least to the coarse-grid model developed by Franklin and Forman (1987). Other harvest pattern-generation models include LSPA (Li et al 1993), CASCADE (Wallin et al. 1994), HARVEST (Gustafson and Crow 1996), SIMPPLLE (Chew et al 2004), and LANDIS (Gustafson et al. 2000). These models differ in the input data required, and the sophistication of the scenarios they can simulate.

The model you will use is a simplified version of HARVEST (Gustafson and Rasmussen 2005). HARVEST was designed to simulate even-aged timber harvest techniques that regenerate a stand of trees of the same age (e.g., clear-cutting, shelterwood, seed-tree techniques), generating patterns similar to timber management (Gustafson and Crow 1999). HARVEST has been used to predict the effect of alternatives management strategies on forest fragmentation (Gustafson and Crow 1996) and animal habitat (Gustafson and Crow 1994; Gustafson et al. 2001), and evaluate effects of diverse owner management strategies on landscape sustainability (Gustafson et al. 2007). For this exercise, the model was simplified to minimize the input data required and allow the user to experiment with the most interesting and important parameters while minimizing confusion caused by too much complexity. The model enables one to change the size of timber harvest openings, the total area harvested, and the spatial distribution of harvested areas (whether harvests will be clumped or dispersed).

Change in Spatial Pattern

An important spatial consequence of intense disturbance (including even-age timber management techniques) or disease in forested ecosystems is the increased production of edge and reduction of **forest interior habitat**. Although a number of species appear to be sensitive to **edge habitat** (forest that is in proximity to a forest edge), it is not entirely understood why this sensitivity exists. One possible explanation is that the amount of forest habitat found within a given radius of a nest located adjacent to an open area is less than a nest surrounded by forest (King et al. 1997). Another might be the increased predation or brood parasitism rates observed in edge habitats (Brittingham and Temple 1983; Andren and Angelstam 1988). It is also unclear how far negative edge effects permeate into the forest. Effects of light and microclimate on vegetation may extend only a few tens of meters into the forest (Chen et al. 1992). For some forest interior birds, the effect may extend 100–500 m (Andren and Angelstam 1988; DellaSalla and Rabe 1987; Van Horn et al. 1995) although the strongest evidence suggests an effect of only about 50 m (Paton 1994). Conversely, some species prefer edge habitat, and their numbers respond positively to the creation of edge habitat (Litvaitis 1993; Hewison et al. 2001). Likewise, it is not known how far from an edge that the habitat will still be suitable for edge species. Because of these uncertainties, it is useful to quantify the amount of edge and interior habitat using a range of edge-buffer widths. The amount of interior present is quite sensitive to the width of the edge-buffer under certain patterns of forest openings, as you will discover.

The patch structure of landscapes is thought to have a significant effect on ecological communities (Turner 1989). Disturbance usually produces **patches** (i.e., an area with habitat conditions that are different from those surrounding it). The patchiness of a landscape mosaic is the result of the interaction of past disturbance and the heterogeneity of the abiotic environment. Disturbance has the potential to significantly alter the scale of patchiness of the landscape mosaic (Levin 1992). Consequently, monitoring change in patch-based measures of spatial pattern is an important way to assess landscape change, and spatial models provide a tool to investigate how disturbance may affect patchiness (Gustafson 1998).

Change in spatial pattern is also related to the rate of recovery after disturbance. When recovery is quick, disturbance effects are more transient. Timber harvest openings are generally ephemeral—succession occurs and forests regrow. However, the rate of recovery may vary widely depending on a number of factors, most notably climate (precipitation and temperature) and soil conditions, where colder, dryer or unproductive sites may take decades longer to recover than warmer, wetter, or productive sites. For this reason, the persistence of disturbance effects may vary markedly between different parts of the world.

The HARVEST Lite Model

HARVEST Lite is a simple, yet powerful harvest simulator that allows control of the most important determinants of spatial pattern in managed forests. HARVEST Lite allows the user to specify the definition of forest interior and the rate at which harvested areas recover to a closed canopy condition. This requires you to specify the width of the edge-buffer used to calculate forest interior, and how long harvest openings will function as openings, perforating the forest interior. You will find this useful to investigate how changes in these definitions affect spatial pattern. The patch structure of the forest age map is also analyzed by HARVEST LITE. Patches are identified using an 8-neighbor rule, meaning that cells of the same age that share a common edge or are adjacent in a diagonal direction are considered part of the same patch. The model parameters are as follows:

MODEL INPUT

- **Forest age map of initial conditions**. Managed forests are typically divided into stands. A stand is an area with a common history and is relatively homogeneous with respect to forest composition and age. The age of a stand usually reflects the time (yrs) since harvest or other disturbance such as fire or windthrow. Two forest age (stand) maps representing different disturbance histories are supplied for this exercise.
- **Mean harvest size**. This is the average size of harvests (ha) that HARVEST LITE will apply to the landscape. The model will generate harvests from a distribution with this mean size and a standard deviation 10% of this value. In real-world management, values may range from <1.0 ha to more than 300 ha, depending on the ecosystem and the management goals.
- **Percent of forested area to cut**. This is the percent of the forested area in the input map cut by the model each decade. For example, if 10% of the forest is cut each decade, 80% of the forest will have been harvested by the end of the eight-decade simulation.
- **Dispersion method**. Two spatial dispersion methods for harvests are available—dispersed (harvests openings are placed independently), or clumped (harvests are placed in clusters of nine openings.) In both cases, harvests are only permitted in forest stands older than 80 years of age.

Two additional parameters are specified for the analysis of forest interior and edge:

- **Edge-buffer width**. This is the maximum distance (m) from a forest opening that edge conditions permeate into the forest. Interior conditions are assumed to exist at distances greater than this value. HARVEST LITE must use a value that is a multiple of the map grid cell width (in this case 30 m). Other values will be converted to the nearest multiple of the cell width. A proposed definition of edge habitat for forest interior birds ranges between 50 and 500 m, which would be represented as 60–490 m in the model.
- **Opening persistence time**. This is the average time (yrs) that it takes for harvest openings to regrow to closed canopy conditions. Harvested cells younger than this value are assumed to be an opening, whereas cells exceeding this value are assumed to have a closed canopy. HARVEST LITE will round values to the nearest decade.

MODEL OUTPUT

Each simulation represents eight decades of harvest activity. Model outputs take the form of maps and map analysis reports including:

- **Forest age map.** Displayed upon completion of the simulation, this map reflects the cells harvested during the simulation, and the aging of unharvested cells. This map may be saved and used as input for other simulations.
- **Area of forest interior.** This is the area (ha) of forest interior conditions based on the input forest age map using the defined edge-buffer width described above. Forest interior habitat is shown in red.
- **Area of forest edge.** This is the area (ha) of forest edge conditions calculated based on the forest age map with interior conditions defined by the edge-buffer width. Forest edge habitat is shown as a gradient of colors other than red. A measure of linear edge (boundary) between patches of different ages is calculated as part of patch analysis, and this is different than the area of forest edge habitat calculated as part of interior analysis.
- **Mean size of patches.** This is the average size (ha) of forest patches, where patches are defined as contiguous cells of the same forest age. Some patches will be the result of simulated harvests and other remnants of uncut forest. Consequently, the mean size of patches *will not likely equal* the mean harvest size you used to simulate harvest activity.

ASSUMPTIONS

A number of simplifying assumptions were made in the development of HARVEST Lite to reduce input data requirements, and enable quick simulations over relatively large areas. The first assumption is that unless forest managers are intentionally trying to manage spatial pattern, harvest openings within areas managed for timber typically take a spatially random distribution when accumulated over the course of a decade. This assumption is based on an analysis of harvest activity on the Hoosier National Forest (Gustafson and Crow 1996). However, HARVEST Lite does include the constraint that harvests cannot be placed where the forest is younger than a specified age. This minimum age of forest that may be harvested has been fixed at 80 years in HARVEST Lite, and all simulations run for eight decades. Several other simplifications have been made to reduce model complexity for this exercise. The standard deviation around the user-specified mean harvest size has been fixed at 10%. HARVEST Lite includes an option to manage spatial pattern by producing clumped distributions of harvest openings. The nucleus of each clump is randomly placed, and then eight other harvest units are placed randomly around the initial harvest. HARVEST Lite always leaves a 1-cell buffer between harvests allocated in the same decade and adjacent to any non-forested land uses. HARVEST Lite ignores specific forest types, assuming that forest types are harvested in proportion to their availability. HARVEST Lite uses stand age as a surrogate for merchantability and ignores the density of trees and tree size class. The proximity of roads and the feasibility of conducting logging operations are assumed to be uniform across the land base.

For this exercise, two forest age maps have been provided. These maps were derived from stand maps of the Hoosier National Forest, and represent an area of almost 4000 ha, with a cell size of 30 m (0.09 ha). Non-forested areas appear black on the forest age maps; a lake occurs in the right-center of the input maps provided, and a small agricultural area in the lower right corner. One represents a managed landscape, where stands range in age from <10 to 140 years old (**managed.gis**). The other (**undistbd.gis**) contains a map with the same spatial characteristics, yet with no young stands, suggesting a lack of disturbance. Because none of the stands in the undisturbed map are too young to be harvested, there are initially no constraints on harvest.

Instructions for Using the HARVEST Lite Model

Start HARVEST Lite by double clicking on its icon (or HarvLite.exe). A help document is available from the **Help** menu.

1. Specify a base forest age map to use for simulations by selecting **Choose base map** from the **Model** menu. This will load the map into memory and allow you to analyze the initial pattern, or alternatively, immediately begin a simulation.
2. The spatial pattern of patches may be analyzed at any time using the **Analyze** menu. Analysis of forest interior requires specification of an **edge-buffer width** and an **opening persistence time**. HARVEST Lite will display a map of forest interior and edge and calculate the amount of interior and edge habitat based on these values, with results printed to the screen and written to a running log file that can be saved as a record. You may conduct multiple analyses of interior on the same forest age map using various values for interior definition.
3. Similarly, an analysis of patches calculates the **mean size of forest patches** (defined by their age) with results also written to the running log file.
4. The **running log** can be saved to a text file at any time using the **Save log file** option under the **Save** menu. The running log is cleared when you load a new base map (and when you save the log file), so if you wish to save any analyses, do so prior to loading a base map.
5. The map of interior may also be saved using the **Save** menu. Map files are saved in ERDAS 7.4 format, and may be loaded into many common GIS systems, or used as input maps for other HARVEST Lite simulations.
6. To conduct a harvest simulation, select **Execute** from the **Model** menu. A dialog will allow you to set the parameters controlling the allocation of harvests on the landscape. When the simulation is finished, an updated forest age map is displayed.
7. You may now analyze the pattern of this changed landscape using the same analysis functions described above in steps 2 and 3. Analyses will be appended

to those conducted previously. You may also wish to save the new maps for later analysis or to use as input for further simulations (as in step 5).

8. To conduct a new simulation using different parameters, reload a base map by selecting **Choose base map** under the **Model** menu. This will clear all prior maps, analyses, and parameter settings from memory.
9. To quit HARVEST LITE, choose **Exit** from the **Model** menu.

EXERCISES

Forest Harvest Simulation Scenarios

Complete the simulations for all assigned exercises before answering discussion questions to ensure you complete the simulations in the allotted time. An excel spreadsheet is provided to record and graph your results (**HarvLite.xlsx**). This is a stochastic simulation model (i.e., simulations are based on random number sequences). Thus, the model will not produce the same results on successive runs, and results will differ slightly among users. When asked to describe the relationship between a model parameter and a measure of landscape pattern, consider the possibility that they are not related.

EXERCISE 1: Effects of Mean Harvest Size on Forest Pattern

Forest managers are being compelled (either by regulation or public opinion) to reduce the size of clear-cuts and other timber harvest activities. For example, in the USA there is a 16 ha limit on the size of clear-cuts on most National Forests. In preparation for this exercise, propose hypotheses (circle options below) about how clear-cut size is related to the amount of forest interior and forest edge remaining on a landscape, given a constant area of timber harvest.

- *Hypothesis 1*: The amount of forest interior will (increase/decrease) as clearcut size increases.
- *Hypothesis 2*: The amount of forest edge habitat will (increase/decrease) as clearcut size increases.

The following exercise will allow you to test your hypotheses by simulating four different management scenarios in which mean harvest size differs. To do so, follow the sequence of steps below.

1. From the Model menu, select **Choose base map**. Use the input file **managed.gis**, found in the same directory as the HARVEST Lite program itself.
2. Use HARVEST Lite to simulate four forest management scenarios in which mean harvest sizes vary (use sizes of 1, 10, 20, and 30 ha). To do this, select **Execute** from the Model menu. Enter a **harvest size** of 1.0. The values for the

other two simulation parameters (% forest area to cut and dispersion method) will be held constant across simulation runs. For **Percent of forested area to cut**, enter a value of 3.0, and select the **Dispersed** dispersion method. Click on the **OK** button to start the simulation.

3. When each simulation run has completed, calculate the amount of forest interior by selecting **Interior (after harvest)** from the **Analyze** menu. Enter an **Edge-buffer distance** of 180 m and an **Opening persistence time** of 20 years. Be sure to use these values for each simulation in this exercise. Also conduct a patch analysis by selecting **Patches (after harvest)** from the **Analyze** menu.
4. Use spreadsheet provided (see the tab for Exercise 1) to record the mean harvest size, area of forest interior, area of forest edge, and mean patch size (all age classes) for each run. These values can be found in the **Progress and Results** window after each analysis is completed. If you wish to save the log file after each simulation, be sure to do so (select **Save log file** under **Save** menu) prior to reloading the base map.
5. Repeat steps 1–4 for the other three harvest sizes (10.0, 20.0, and 30.0 ha).
6. Recalling that Harvest Lite is a stochastic model, replicate your data two more times, and record those results in the appropriate tables in the spreadsheet. The purpose of replication is to provide a sense of the variability among model runs and to provide mean values that are more accurate than those from a single model run.
7. Examine the graphs of the area of forest interior, forest edge, and mean patch size plotted against mean harvest size. Answer the following questions:

Q1 Is there a threshold effect of mean harvest size (i.e., a small range of values where the effect changes markedly)? If so, at approximately what mean harvest size does the threshold occur?

Q2 If you were advising a forest manager who was under pressure to both minimize harvest size and to maximize forest interior habitat, what would you recommend as a policy for mean harvest size?

Q3 Would you say the variability among model runs is high or low?

Q4 Based on the graphs, would you say your hypotheses were supported or discredited?

EXERCISE 2: Effects of Percent of Forest Cut Each Decade on Forest Pattern

Timber production levels are declining on many publicly owned forests in the USA, primarily to enhance biodiversity and other non-commodity values of forests. This is implemented primarily by reducing the percentage of the land over which timber harvest is allowed. In this exercise, we will examine the effect of changing the percent of forest cut each decade. In preparation, propose a hypothesis about how the percent of forest cut each decade is related to the amount of forest interior and forest edge remaining on a landscape, given a constant timber harvest size. Upon reflection, you might likely hypothesize that increasing the area cut will decrease the amount of interior. Thus instead, consider a more subtle hypothesis about whether the relationship is linear (straight line) or nonlinear (curve) by circling an option below:

- *Hypothesis:* The amount of forest interior will decrease (linearly/nonlinearly) as clearcut size increases.

To explore your hypothesis, follow these steps:

1. From the Model menu, select **Choose base map**. Use the input file **managed.gis**.
2. Use HARVEST Lite to simulate four other forest management scenarios in which the **Percent of forested area to cut** varies, from 1 to 7% of the landscape each decade in 2% increments. Hold the other parameters constant for each of these runs. Use a **Mean harvest size** of 5.0 ha, and the **Dispersed** dispersion method.
3. Calculate the amount of forest interior using an **Edge-buffer distance** of 180 m and an **Opening persistence time** of 20 years. Be sure to use these values for each simulation in this exercise. Also conduct a **patch analysis**.
4. Record data in the Exercise 2 tab of the spreadsheet.
5. Repeat steps 1–4 for each **Percent of forested area to cut** (1, 3, 5, 7%) and replicate two more times.
6. Examine the graphs and answer the following questions:

Q5 What is different about the shape of these plots compared to those generated for the effects of mean harvest size? Was your hypothesis supported?

Q6 Does there appear to be a threshold effect of percent of forest area cut each decade? If so, at approximately what percent does the threshold occur?

OPTIONAL: You may wish to produce a 3-D surface plot combining the results of this and the previous exercise. Additional simulations will be necessary to complete the plot. Compare your results to those in Gustafson and Crow (1994).

EXERCISE 3: Effects of Spatial Dispersion on Forest Pattern

Landscape ecologists have argued that intentionally managing the spatial pattern of landscapes can improve habitat conditions. One option available to forest managers is clustering harvest activity. This exercise will examine the effects of clustering of harvests on area of forest interior, area of forest edge, and mean patch size.

1. Select the input file **undistbd.gis**.
2. Choose a **mean harvest size** between 5 and 30 ha, and a **percent of the forest area to be cut** between 1 and 7%. Run three replicate (parameters unchanged) simulations each for dispersed and clustered harvests.

3. Calculate the amount of forest interior using an **Edge-buffer distance** of 180 m and an **Opening persistence time** of 20 years. Also conduct a **patch analysis**. Record data in the tab for Exercise 3.
4. Examine the graphs and answer the following questions:

Q7 In statistics, a significant difference between groups indicates that values observed in one group are highly unlikely to be also observed in a different group. Would you say that clustering harvests significantly changes the area of forest interior habitat when compared to dispersed harvests? What about forest edge habitat? Mean patch size?

Q8 How did you judge significance from these plots? (*HINT:* look at the error bars.)

EXERCISE 4: Effects of Edge-Buffer Width and Opening Persistence Time on Forest Pattern

There is some debate about how far into the forest the effects of edge are evident. The effects related to reduced nesting bird densities and increased nest predation may extend much further into the forest than do microclimate effects. Forests also recover from harvesting at different rates in different ecosystems. Forests on good soils in moist climates may recover more quickly than forests on poor sites or in relatively dry climates. This exercise will examine how the spatial pattern of forest interior depends on how interior habitat is defined.

1. Select the input file **managed.gis**.
2. Simulate harvests with a **mean harvest size** of 1.0 ha and 4% of the **forest area cut each decade**. Use the **Dispersed** dispersion method. This will be the only simulation run for this exercise.
3. Calculate area of forest interior using no edge-width buffer (**0 m**), a **150 m** edge-buffer width, and a **300 m** edge-buffer width. Assume that openings persist for two decades. Do not conduct a new simulation between interior calculations for each edge-buffer width.
4. Record data in the tab for Exercise 4 in the spreadsheet.
5. Without running another simulation, repeat the three calculations in step c, this time using an **Opening persistence time** of three decades. Record your data.
6. Examine the graphs and answer the following questions:

Q9 Does an increase in edge-buffer width have a disproportionate effect on the area of forest interior habitat? Why or why not?

Q10 Some species are highly sensitive to the presence of edge in their habitat, and therefore a large edge buffer width would be used to calculate suitable habitat for them. What is the consequence of timber cutting for such species relative to less sensitive species?

Q11 What is the effect of a longer opening persistence time?

Q12 To maintain interior habitat in a forest with slow-growing species (i.e., openings persist longer), make recommendations about timber cutting strategies to achieve this goal based on what you have learned in this lab exercise.

SYNTHESIS

Q13 Is forest interior more sensitive to variation in "Harvest size" or "Percent of forest area cut?" To which is forest edge more sensitive? To which is patch size more sensitive? Can you think of a situation where a forest manager would find this information useful?

Q14 How did this exercise change your thinking about the spatial aspects of timber harvesting? How might the results of your simulations be used to develop a research project or advise a forest management debate about trade-offs?

Q15 Review the assumptions made by the HARVEST Lite model. Under what scenarios might they be reasonable or unreasonable? How might you test these assumptions? How does knowledge of the assumptions influence interpretation of the results?

Q16 Was there a parameter missing from the model with which you wanted to be able to experiment? What was it, and how might that parameter be related to the habitat requirements of a forest species?

Q17 Consider the process and impacts of landscape change represented by each of the parameters that can be manipulated by HARVEST Lite (mean size of harvests, % of forest cut, dispersion). How are these processes and impacts similar to disturbances in non-forested landscapes? How are they different? Do you think the principles you learned today can be applied to other ecological systems?

REFERENCES AND RECOMMENDED READINGS[1]

Andren H, Angelstam P (1988) Elevated predation rates as an edge effect in habitat islands: experimental evidence. Ecology 69:544–547

Brittingham MC, Temple SA (1983) Have cowbirds caused forest songbirds to decline? BioScience 33:31–35

Chen J, Franklin JF, Spies TA (1992) Vegetation responses to edge environments in old-growth Douglas-fir forests. Ecol Appl 2:387–396

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Chew JD, Stalling C, Moeller K (2004) Integrating knowledge for simulating vegetation change at landscape scales. Western J Appl Forestry 19(2):102–108

DellaSalla DA, Rabe DL (1987) Response of least flycatchers *Empidonax minimus* to forest disturbances. Biol Conserv 41:291–299

*Franklin JF, Forman RTT (1987) Creating landscape patterns by forest cutting: ecological consequences and principles. Landsc Ecol 1:5–18. *This paper was among the first to use a simulation model to investigate the landscape pattern effects of timber harvesting.*

*Gustafson EJ (1996) Expanding the scale of forest management: allocating timber harvests in time and space. Forest Ecol Manag 87:27–39. *In this paper HARVEST was used to simulate several clustered cutting strategies on a real landscape encompassing the entire Hoosier National Forest in Indiana. Results showed that the area harvested could be increased while also increasing the amount of forest interior when a clustered strategy was used.*

Gustafson EJ (1998) Quantifying landscape spatial pattern: what is the state of the art? Ecosystems 1:143–156

Gustafson EJ, Crow TR (1994) Modeling the effects of forest harvesting on landscape structure and the spatial distribution of cowbird brood parasitism. Landsc Ecol 9:237–248

Gustafson EJ, Crow TR (1996) Simulating the effects of alternative forest management strategies on landscape structure. J Environ Manage 46:77–94

Gustafson EJ, Crow TR (1999) HARVEST: linking timber harvesting strategies to landscape patterns. In: Mladenoff DJ, Baker WL (eds) *Spatial modeling of forest landscapes: approaches and applications*. Cambridge University Press, Cambridge, pp 309–332

Gustafson EJ, Shifley SR, Mladenoff DJ et al (2000) Spatial simulation of forest succession and timber harvesting using LANDIS. Can J For Res 30:32–43

Gustafson EJ, Murphy NL, Crow TR (2001) Using a GIS model to assess terrestrial salamander response to alternative forest management plans. J Environ Manage 63:281–292

Gustafson EJ, Rasmussen LV (2005) HARVEST for Windows v6.1: User's guide. http://www.ncrs.fs.fed.us/4153/Harvest/v61/documentation/. Accessed 8 Nov 2007

*Gustafson EJ, Lytle DE, Swaty R et al (2007) Simulating the cumulative effects of multiple forest management strategies on landscape measures of forest sustainability. Landsc Ecol 22:141–156. *HARVEST was used to predict the cumulative effects of four owner groups (two paper companies, a state forest and non-industrial private owners) with different management objectives on landscape pattern in an upper Michigan landscape managed primarily for timber production. The authors studied trends in landscape pattern metrics that were linked to Montreal Process indicators of forest sustainability. The diversity of management approaches produced a diversity of habitat conditions such that most trends pointed to a positive outcome on sustainability.*

*Gustafson EJ, Shvidenko AZ, Scheller RM (2011) Effectiveness of forest management strategies to mitigate effects of global change in south central Siberia. Canad J Forest Res 41:1405–1421. *The authors used the LANDIS-II disturbance and succession model to study how alternative timber cutting strategies affect forests under climate change in Siberia. Strategies were discovered that resulted in greater forest biomass, increased abundance of species predicted to grow well under future climates, and reduced fragmentation, but they did not concurrently reduce losses by disturbance. No single strategy appears able to achieve all possible forest management objectives.*

Haefner JW (1996) *Modeling biological systems: principles and applications*. Chapman and Hall, New York (An excellent introduction to simulation modeling)

Hewison AJM, Vincent JP, Joachim J et al (2001) The effects of woodland fragmentation and human activity on Roe Deer distribution in agricultural landscapes. Can J Zool 79:679–689

Karplus WJ (1983) The spectrum of mathematical models. Perspect Comput 3:4–14

King DI, Griffin CR, DeGraaf RM (1997) Effect of clearcut borders on distribution and abundance of forest birds in northern New Hampshire. Wilson Bull 109:239–245

Levin SA (1992) The problem of pattern and scale in ecology. Ecology 73:1943–1967

Li H, Franklin JF, Swanson FJ et al (1993) Developing alternative forest cutting patterns: a simulation approach. Landsc Ecol 8:63–75

Litvaitis JA (1993) Response of early successional vertebrates to historic changes in land use. Conserv Biol 7:866–873

Paton PW (1994) The effect of edge on avian nest success: how strong is the evidence? Conserv Biol 8:17–26

Turner MG (1989) Landscape ecology: the effect of pattern on process. Ann Rev Ecol Syst 20:171–197

Van Horn MA, Gentry RM, Faaborg J (1995) Patterns of ovenbird (*Seiurus aurocapillus*) pairing success in Missouri forest tracts. Auk 112:98–106

*Wallin DO, Swanson FJ, Marks B (1994) Landscape pattern response to changes in pattern generation rules: land-use legacies in forestry. Ecol Appl 4:569–580. *The authors used a harvest simulation model to show that landscape patterns produced by dispersed disturbances are difficult to erase, persisting in some form for a long time.*

*Zollner PA, Roberts LJ, Gustafson EJ et al (2008) Influence of forest planning alternatives on landscape pattern and ecosystem processes in northern Wisconsin, USA. Forest Ecol Manage 254:429–444. *The authors simulated harvest and succession for 250 years in response to eight alternative timber management plans drafted for the Chequamegon-Nicolet National Forest in Wisconsin, USA. These plans varied in the amount of clearcutting relative to timber harvest methods that leave some trees behind. Most habitat variables sorted across the plans along a gradient of the amount of clearcutting specified in each plan.*

Chapter 10
Regional and Continental-Scale Perspectives on Landscape Pattern

Jeffrey A. Cardille and Monica G. Turner

OBJECTIVES

Landscape patterns vary widely across Earth's surface as a result of both anthropogenic and natural causes. This variation among landscapes can be quantified by using a large number of metrics developed to capture distinctive qualities of spatial pattern. An informed understanding of pattern–process relationships involves landscape comparisons among and within regions. Despite many advances in landscape pattern analysis, informed selection of landscapes for studying pattern–process relationships in real-world situations remains challenging. This lab explores these challenges with objectives designed to enable students to:

1. Think critically about the benefits and limitations of subjective, nonquantitative landscape assessments;
2. Examine the statistical distributions of landscape metrics within or among regions by exploring histograms for commonly used metrics;
3. Learn and implement ways to improve landscape comparisons through selection of appropriate study landscapes based on specified land-cover proportions, arrangements, or gradients; and
4. Gain experience using two practical tools (Metric Finder and Metaland) within the context of realistic landscape monitoring scenarios.

In Parts 1 and 2, you will conduct your own rankings of landscapes visually and then use Metric Finder to evaluate your work. In Parts 3 and 4, you will use Metaland

J.A. Cardille (✉)
McGill University, Sainte Anne de Bellevue, Montréal, QC, Canada
e-mail: jeffrey.cardille@mcgill.ca

M.G. Turner
University of Wisconsin-Madison, Madison, WI, USA

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_10

to identify landscapes according to criteria useful in providing a continental perspective on landscape pattern. For each of these exercises, you will need to query and evaluate information from evolving online databases. Links to these tools and databases can be found in this chapter's student guide at http://goo.gl/FTc4gY or via the web site for the book.

INTRODUCTION

Landscape ecologists now have at their disposal a large number of well understood and widely used metrics that quantify landscape **composition** (the relative abundance of different land-cover categories in a landscape) and **configuration** (the spatial arrangement of those land-cover categories). Readily available data and user-friendly software, such as Fragstats (McGarigal et al. 2012) make such analyses routine. However, researchers new to landscape pattern analysis may find it difficult to understand metrics intuitively, and visual assessments of how a set of landscapes vary with respect to particular metrics can be very challenging, and even misleading. Exercises that allow students to "see" landscapes through the lens of different metrics provide a useful foundation for using and interpreting the results of numerical landscape analyses.

Because landscape ecology focuses on the causes and consequences of landscape heterogeneity, landscape pattern metrics are often used as independent variables in research. There are many, many examples of such studies in the published literature, but the type of question is general: How does landscape composition and/or configuration influence a process of interest? The response of interest may be the presence, abundance, or demography of focal taxa, species richness, nutrient loading or water quality in lake or rivers in a given watershed, rates of land conversion that has occurred over a particular time period, spread rate of invasive species, rates of encounter between predators and prey, or many other phenomena that may be affected by landscape pattern. To answer such questions, researchers often need to systematically identify replicated study landscapes that vary in composition and configuration in predetermined ways, and this task can be daunting. In many cases, researchers are left to analyze multiple maps in the hope of finding sets of study areas that vary in the desired manner. Land managers may face similar challenges, such as identifying forest-dominated landscapes in which forest patch sizes are above a threshold size required to sustain species of conservation concern.

Consider the following scenario. In many areas, urban and semi-urban environments are replacing agriculture and forests. This low-density development often greatly increases the amount of edge between forested and non-forested land. Such fragmented forest landscapes are common in New England, USA, which is a region with high prevalence of Lyme disease. Lyme disease is a bacterial, tick-borne illness found throughout eastern North America that causes skin rashes, cardiac abnormalities, and neurological problems. Since its discovery in the 1970s, rates of Lyme disease occurrence have increased steadily, and over 15,000 people are infected each year. Small mammals, especially the white-footed mouse (*Peromyscus leucopus*), are the most abundant competent hosts for the disease-causing bacterium and a key host for the larval ticks. Landscapes with numerous small patches of forest (and thus high forest edge density) tend to have high populations of white-footed mice, and these are also landscapes where humans are likely to encounter the ticks. Research has shown (e.g., LoGiudice et al. 2003) that ecosystems with a high density of competent hosts are associated with increased rates of Lyme disease, and this is where the connection to landscape pattern occurs. Imagine a study of New England that seeks to initiate a wide-ranging field campaign to survey host abundance and disease prevalence, along a tightly controlled gradient of expected risk of Lyme disease. Given a land-cover map of all of New England, how would you begin to choose 5, 10, or 100 sampling areas to sample? It is often difficult to identify and select landscapes that allow for processes or conditions to be controlled across gradients, or for random sampling among replicated landscapes that share a given set of characteristics.

To help with such spatially extensive and complex challenges, the Metaland (Cardille et al. 2005) and Metric Finder tools have been developed. Built on databases of landscape and class-level metrics generated for large data sets of same-sized landscapes, **Metaland** is designed for understanding variation in patterns across large areas, learning about the statistical distribution of real-world landscape metric values, and selecting landscapes with desired characteristics. It includes values for more than 190,000 contiguous 6.48-km × 6.48-km landscapes (at 30-m resolution) across the conterminous US, across several time periods. Despite caveats associated with their use and interpretation (e.g., Gustafson 1998; Li and Wu 2004; Langford et al. 2006; Cushman et al. 2008; Eigenbrod et al. 2011; Turner and Gardner 2015), landscape metrics have been seen to be valuable for finding differences and similarities among landscapes in this comprehensive data set. For example, Cardille and Lambois (2010) used the 1992 Metaland data to discern a widespread imprint of human activities on US landscapes, distinguishing among different types of landscapes based on the similarity of their landscape metric "signature."

This chapter focuses on subjective and objective assessments of landscape pattern, and the interactions between what we perceive and what computations of landscape metrics can tell about the world around us. This set of exercises introduces students to Metaland and explores the associated Metric Finder tool, two resources for understanding differences and similarities among different landscapes and for identifying sets of landscapes that meet predetermined criteria for landscape composition and configuration. You will investigate patterns in land-cover data derived from Landsat imagery for the National Land Cover Data Set (NLCD; Vogelmann et al. 2001; Homer et al. 2007; Fry et al. 2009; Jin et al. 2013).

Part 1. Estimating Landscape Metrics by Eye

In landscape ecology, how hard is it to quantify pattern? How straightforward is it to say that two landscapes are similar, and that another one of them is unlike the others? We begin by determining how well our visual assessments match up with quantitative measures of landscape pattern. You will describe the patterns you see in sample landscapes, and then compare your visual estimates with calculated landscape metric values.

EXERCISE 1: Visual inspection

Consider six landscapes extracted from the New England region, located in the northeastern USA (Figure 10.1). More than 15 land-cover categories are shown in these images, but you can think of them in four main categories: (1) red and pink for residential and commercial development; (2) green for forest; (3) yellow, brown, and beige for agriculture; and (4) light and dark blue for water and wetlands. Inspect these six landscapes in the figure or at the link provided in the student guide, then compare and contrast their composition and configuration.

Q1 What qualities do you see that are similar or different? You might consider the proportions of different land-cover types and how they are arranged; you might also think about the land-use history that may have driven the patterns you see, or the connections between areas of a given land-use type. Write down your observations, noting at least three similarities and three differences.

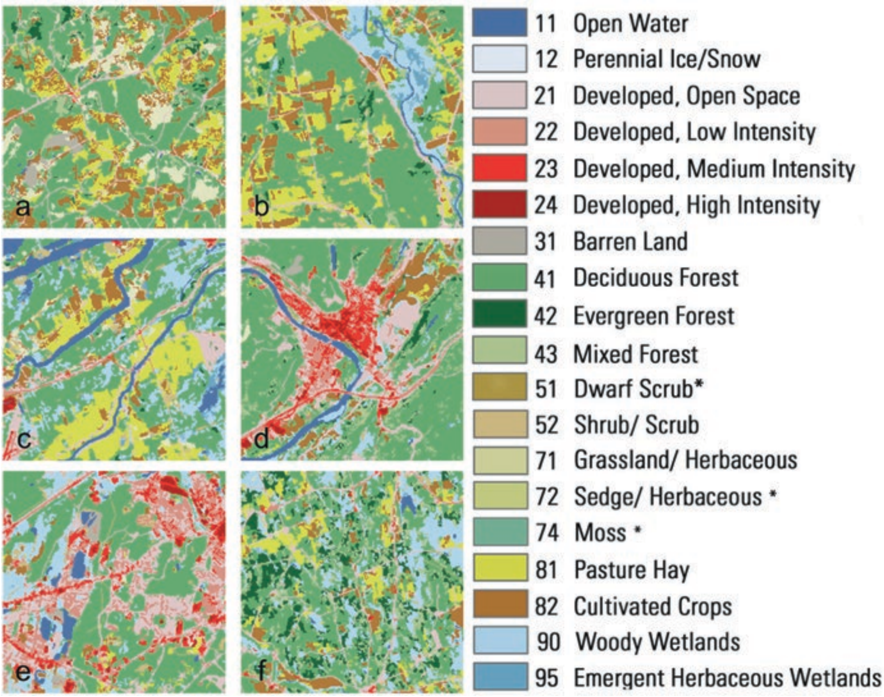


Figure 10.1 Six landscapes from New England, USA. Classification taken from the National Land Cover Data Set for 2001

EXERCISE 2: Ranking using metrics

Any of a large number of landscape metrics can be estimated in a landscape ecological study, but how well can we "see" these values on different landscapes? Some metrics are more intuitive than others, and some landscapes may have distinctive characteristics that are not easily seen. For some commonly used landscape metrics, you will see how well you can assess key characteristics of different landscapes based solely on visual inspection.

Rank each landscape in Figure 10.1 from lowest to highest value according to each of the following metrics:

- Proportion of agriculture
- Proportion of forest
- Proportion of residential and commercial development
- Proportion of water and wetlands
- Total linear amount of edge between all land-cover categories in the landscape
- Contagion
- Shannon Evenness

Q2 Describe the relative ease or difficulty in ordering these landscapes according to each metric. Which metrics were easiest to rank, which were most challenging, and what made them easy or difficult? Next, compare your ranking to the "true" order based on the numerical values, which can be provided by your instructor. How well did you do? When there is disagreement, why do you think this occurred?

Part 2. Metric Finder: Relating Visual Assessments to Landscape Metrics

In a landscape ecology analysis, we may know the type, or "look," of the landscapes we are interested in, but it can be difficult to match metrics with those mental criteria. This section explores a way to identify a suite of landscape metrics that corresponds to visual criteria that you define. The idea is that with a landscape characteristic in mind, you visually determine which two of a trio of landscapes appear most similar, and which one of the three is most different. Using the same logic as an email spam filter that learns more and more about the characteristics of unwanted messages as you identify them, **Metric Finder** learns your preferences as you choose pairs of landscapes according to the criteria you use to judge similarity. In computer science, this is known as a **labeling** approach: you provide the labels of similarity, and Metric Finder gradually reveals which metrics of composition and configuration fit the pattern of pairings that you make. Given more and more iterations of labeling, Metric Finder builds an increasingly confident view of your perception of landscape characteristics. Using this tool can help you to decide which metrics might be useful for a particular study, while also revealing whether some routinely used metrics are easy or hard to distinguish visually in real-world landscapes.

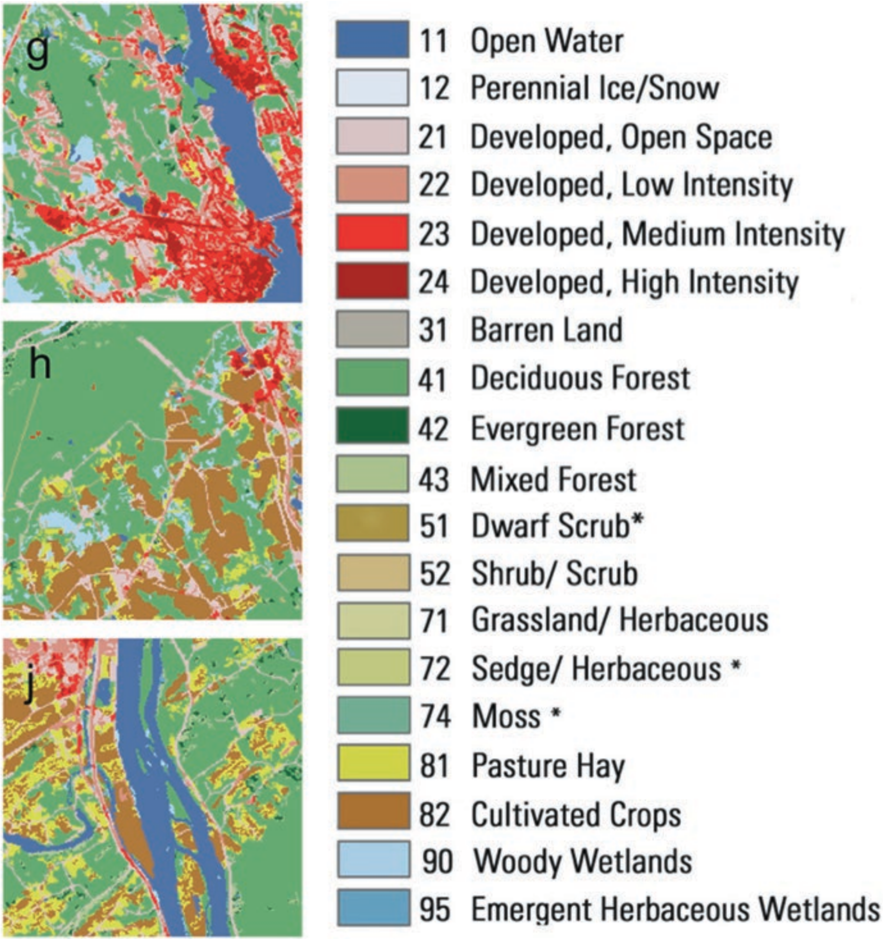


Figure 10.2 Example set of three landscapes from an iteration of Metric Finder

Imagine different ways in which the three landscapes in Figure 10.2 are similar to and different from each other. For example, Landscapes G and J each contain a river, so a user emphasizing the shape of waterways, or the impact of water on nearby development, might rate them the most similar with respect to those criteria. Alternatively, all three landscapes contain low- and medium-intensity development, and landscapes H and J appear most similar with respect to the amount (and, perhaps shape) of these developed classes. Landscapes H and J are also similar to each other with respect to the amount of core forest, so a forest criterion might consider them more similar to each other than to landscape G.

Metric Finder is designed to interpret such labels of landscape similarity to identify the pattern measures that best distinguish the landscapes, using the unique perspective of each user. To do this, the tool repeatedly presents three landscapes for

Figure 10.3 Conceptual diagram of two metrics that record different characteristics of three landscapes, G, H, and J. For users who select landscapes G and J as similar, a metric quantifying the amount or shape of Open Water corresponds more closely to their perception than does the metrics of the amount of Core Forest. Users who select landscapes H and J as similar suggests that Core Forest may be important to their perception of landscapes, and that Open Water is not

visual interpretation. While looking at the three landscapes, the user identifies which two of the three look the most similar from his or her perspective—or equivalently, which one of these landscapes is not like the others. Then, for each metric from a predetermined list, Metric Finder evaluates the relative similarity of the three values, estimating the probability of having picked those two landscapes as being similar if the metric had been a criterion considered by the user.

In the iteration of Metric Finder shown in Figure 10.3, the values of the amount of Core Forest are more similar to each other than are the values of the shape metric of Open Water, for which two of the three are much more similar. By Metric Finder's logic, a user who selects landscapes G and J as being more similar is more likely to be responding to the shape of Open Water than to the amount of Core Forest. After the user chooses which two landscapes in a trio are the most similar, Metric Finder adjusts each metric's current estimated score either up or down, depending on the weight of the evidence given the paired landscapes. By interpreting pairings of landscapes in this way, Metric Finder tries to estimate which metrics correspond best to a user's perception of landscape pattern.

In this exercise, you will use a set of more than 650 New England landscapes, each of which is 6.48 km × 6.48 km in extent and represents land cover at 30-m resolution for the year 2001 (Homer et al. 2007). We also will explore this set in a later part of this chapter, extracting the set using Metaland from a much larger set of landscapes representing the continental USA.

EXERCISE 3: Labeling Landscapes Using Metrics of Proportion

1. To begin, you will explore the basic functionality of Metric Finder. The web address of the tool can be found in this chapter's student guide. The tool will load into your browser.
2. Examine the Metric Finder interface. At the top are three landscape images chosen at random from the set of New England landscapes, with selection boxes for each. Below the images are two sets of metrics: one denoting class proportions, such as "Shrub," and one denoting landscape-level metrics, such as *te* for *Total Edge*. Near that is a two-class clustering of the landscapes shown on a map, as well as summary characteristics of the two clusters.

3. Focus first on the *Proportion of Pasture*, a prominent land-use class, whose yellow color in the classification is easy to identify.
4. For the first three landscapes presented by Metric Finder, select the two that appear to have the same amount of Pasture. You can choose a pair either because the amounts of Pasture are both high or because they are both low—the important point is to choose the two that are more similar to each other than to the third nonselected landscape. Recall that for a given Pasture proportion, Pasture pixels can appear in very different configurations.
5. If the decision is too difficult or subtle for a given trio of landscapes, click "Skip" to load a new set.

Q3 When you focus on pairing landscape using the *Proportion of Pasture* as the criterion, how many iterations does it take for Metric finder to detect this by moving it near the top of the list of proportion metrics?

EXERCISE 4: Metrics of Proportion for Rare Classes

Pasture was one of the most frequently occurring land-cover classes, and it is relatively easy to work with. Next, you will examine less-dominant classes.

- Reset the metric probabilities using the Metric Finder interface.
- Now, try to label landscapes using a rare land-cover class, *Medium-Intensity Developed*. In this set of landscapes, less than 1% of the landscape (on average) is *Medium-Intensity Developed*, compared to 12% for Pasture

Q4 Is it easier or harder for Metric Finder to detect that you are distinguishing landscapes using the *Proportion of Medium-Intensity Developed* metric? Do you find that other land-use classes move up and down with the *Proportion of Medium-Intensity Developed* metric? If so, why might this be?

EXERCISE 5: Labeling Landscapes Using Metrics of Configuration

Based on your understanding of landscape metrics from other chapters, you will use Metric Finder to try to pair landscapes with respect to their landscape metrics of configuration. First, reset the metric probabilities and try to distinguish landscapes using the *Contagion* metric, which you learn about in Chapter 4. High values of *Contagion* are associated with more aggregated land-cover patterns, and low values have land-cover patterns that are very dissected. After working to promote *Contagion* to the top of the metrics list, choose another configuration metric (e.g., *Total Edge*) that you think might be straightforward to distinguish among landscapes.

Q5 Explain the conceptual similarities and differences between *Contagion* and the other configuration metric(s) you chose. Why do you think each might be more or less difficult to recognize visually using Metric Finder?

Q6 In comparison to the class-level metrics of proportion of the previous exercise, is it easier or harder to distinguish landscapes using metrics of configuration? That is, is it easier or harder to intentionally move a certain configuration metric to near the top? Why or why not?

EXERCISE 6: Grouping Landscapes with Metric Finder

As it promotes and demotes landscape metrics based on their fit to how you have distinguished landscapes, Metric Finder dynamically groups and regroups landscapes into two clusters based on the results of your labeling. At a given iteration, the metrics chosen as clustering criteria are those having the highest estimated score: thus, as you refine and confirm your criteria by selecting pairs of landscapes, the clustering should gradually come to match the clustering you would expect for your set of criteria. As criteria for the grouping, the algorithm bases its clustering on the three highest-rated composition metrics and the three highest-rated configuration metrics. At any moment, the three highest-rated of each are shown with a green background on the Metric Finder interface.

Using Metric Finder's clustering ability, you can evaluate the usefulness of the set of metrics you are implicitly choosing as you label landscapes. The clustering is shown during this process in several ways. First, the full set of landscapes in consideration by Metric Finder is grouped into two sets with the well-known k-means algorithm; their locations are drawn on a map. Second, a selected group of 15 representative landscapes for the set (Cardille et al. 2012; Cardille and Lambois 2010; Frey and Dueck 2007) are displayed, with their current grouping shown at each iteration. You can form your opinion of the usefulness of a set of metrics using the regional-scale clustering map, the clustering of representative landscapes, and the cluster statistics shown for the high-score metrics.

- Reset the metric probabilities using the Metric Finder interface.
- Develop your own criteria for distinguishing landscapes, and explore with Metric Finder. For example, you might try to separate landscapes with a high proportion of any type of developed land.
- When distinguishing landscapes for this exercise, pay attention to the area of the Metric Finder interface that shows the clustering assignments, projected over all of the landscapes in the set. Because the clustering algorithm considers only landscape metric values and not the geographic position of each landscape, there may or may not be a coherent pattern in the clustering of the landscapes.

Q7 How successful is the mapping of the full set of landscapes according to your identified selection criteria? As you select pairs of landscapes and your desired criteria are gradually confirmed, does the resulting classification of landscapes become gradually more or less stable?

Q8 In your opinion, how successful was the clustering of the representative landscapes into two groups? That is, how well did the clustering of the landscapes reflect your identified selection criteria? Are regional-scale landscape patterns revealed? If so, what elements of the landscapes appear to drive them?

SYNTHESIS EXERCISE 7: Identifying Metrics for Management

You have been named the director of a newly created Agency for Monitoring Environmental Change (AMEC) for the USA. Your charge is to develop the means for objectively monitoring differences in landscape patterns and their connection to ecosystem processes. In your position, you want to know whether the metrics selected by your colleagues are good candidates to distinguish landscapes in proposed study areas. While working to identify metrics for national reporting of Lyme Disease risk, two colleagues have developed a serious disagreement about which metrics can best represent the New England landscape.

One colleague prefers a simple composition metric, believing that the *Proportion of Deciduous Forest* is easy to calculate and recognize, straightforward for land managers to use, and thus an excellent proxy for landscapes with high amounts of forest edge. The second colleague argues that the total amount of edge between different land-use types, a configuration metric, is of more direct ecological relevance, and that the agency does not need to use an indirect proxy to estimate Lyme Disease risk. The first colleague counters that while the *Total Edge* metric is also easy to calculate with Fragstats, it is a landscape-level metric that includes edges between land-use classes that are not relevant to the research question. She feels that even if there were a strong statistical correlation between *Total Edge* and tick density, differences in values of the *Total Edge* metric are in practice much harder to distinguish and control in real-world landscapes, suggesting, "you can't manage what you can't see." As director, you suggest that each analyst use Metric Finder to look at landscapes with respect to their metric of choice.

With a partner, assume the roles of these two analysts and work side-by-side with Metric Finder in these same New England landscapes, with one trying to distinguish landscapes using the metric of composition, *Proportion of Deciduous Forest*, and the other distinguishing them using the metric of configuration, *Total Edge*. Continue labeling landscapes until the top three metrics of each type attain stable, high scores. Use your judgment about what constitutes a stable high score.

When assessing the landscapes, each analyst should fill out the online form that will record the information about the metrics that were promoted to the top; this should allow you to put your own work in the context of others who have done this exercise. The form can be found via this chapter's student guide.

Q9 Are the two metrics equally easy to recognize in this real-world set of landscapes? When labeling landscapes according to one metric, what are the other metrics that appear to move in sync with them? Are those metrics correlated, or is there another explanation?

Q10 Are the clusters of representative landscapes similar in the work of the two analysts? Do the clusters appear to reflect the criteria you have used to classify them? Are the maps of landscape groupings similar for the two analysts?

Q11 For the two analysts, what are the statistical characteristics of the two clusters "resulting from" the labeling exercise? How might the statistical properties of the clusters be used to evaluate whether one analyst was more successful than the other in this effort?

Q12 The entire set of responses from everyone who has done this lab exercise are recorded in a growing online spreadsheet. A link to the spreadsheet can be found in this chapter's student guide. By inspecting the charts there, you can see the results of others who have played the roles of these analysts, which can help you better understand any differences between your work and that of your partner. Your report can address the following questions: Across all analysts, were the distances between the clusters significantly different between the two types of analysts? That is, is it more effective to focus on one type of metric or the other for this set of landscapes? Do some metrics show up much more frequently than others—if so, why?

Part 3. Exploring Variation in Landscape Metric Values Across a Region

You have now distinguished landscapes based on visual, qualitative assessments of landscape composition and configuration. You have also used a variety of criteria to assess similarity and differences among landscapes. What if you needed to identify a set of landscapes that met specific composition and configuration criteria for a field or modeling study? For example, what if you wanted to explore the effects of forest spatial pattern on seed dispersal, and needed to locate in your study region a set of landscapes with similar amounts of forest cover but different numbers and sizes of patches of forest? Or what if you wanted to study natural enemies of agricultural pests and needed to identify 30 landscapes having similar amounts of cropland but varied amounts of natural vegetation? Could you identify, say, 50 replicates of landscapes with 40% cropland but with high, medium, and low amounts of natural vegetation? This section explores a way to answer questions like these with the Metaland tool, which allows you to identify landscapes according to specified criteria for a regional- or continental-scale perspective on landscape pattern.

EXERCISE 8: Using Metaland to Select Landscapes with Geographic Criteria

- Access the Metaland web site. A link to the tool can be found in this chapter's student guide.
- Select **Understanding and Retrieving Statistical Distributions**. Choose the data set for the 2001 NLCD having tiles that are 6.48 km × 6.48 km and covering the continental USA.

- Select **Geographic Criteria > Locations**, then check the boxes for both Latitude and Longitude to prepare to specify landscapes covering the region of New England. New England lies in the box roughly between **Latitude >40.5 and Latitude <45.1**, and **Longitude >−80 and Longitude <−69.5**. Select **Search** using these criteria; the result will return more than 7500 landscapes. On the page showing search results, the locations shown in red are landscapes that satisfy the search criteria. They should be located over the northeastern corner of the image, where New England is located within the continental USA.

Refer to the student guide to access a folder containing images of all of the landscapes returned from the search. In preparation for the next exercise, look through the images of these New England landscapes, noting land-cover patterns and especially the abundance and distribution of Deciduous Forest, represented in green.

EXERCISE 9: Estimating Histograms and Testing Expectations

- On a piece of paper, sketch the histogram that you would imagine to represent the percentage of Deciduous Forest among the landscapes of this heavily forested region. On the *X*-axis, place the numbers 0, 10, 20, … 100. These categories will be percentage values corresponding to different levels of Deciduous Forest abundance for the 6.48 × 6.48 km landscapes. The *Y*-axis will indicate the percentage of these landscapes that have values in each interval. For example, if you think that about 10% of the landscapes of New England will have a percentage of Deciduous Forest between 50 and 60, draw a bar at $Y=10$ from $X=50$ to $X=60$. Continue for the other ranges of percentages until you have sketched your anticipated histogram.
- Return to the Metaland interface and click **Percentage of Land** to learn about the frequency of different land covers in these landscapes. The page shows the basic statistics for the percentages of each of the land-cover categories within the chosen subset of landscapes.
- Find the entry for Deciduous Forest, and click **View**. On the resulting page you can see a histogram for that metric (proportion—the metric p_i), for all landscapes in your subset.
- In the field below, note the box labeled **Find the percentile of this value**. In this box, you can enter any metric value, and the program will return the proportion of landscapes in your set that have lower values: that is, the percentile of that value. The median value in a set represents the 50th percentile; this is the value for which half the values in the set are lower, and half are higher. Look at the histogram and, through a process of trial and error, determine the median value of this set of landscapes.

(*NOTE:* Using your browser's controls, you can save an image of the histogram itself for use in a report).

Q13 Did your estimate of the histogram of Deciduous Forest differ from that of the actual distribution, and if so, how? What might account for discrepancies between one's estimation of the histogram and the evidence from the landscape metrics? In what ways does a histogram help assess and understand the makeup of a landscape?

Q14 What are the values at the 33rd, 50th, and 67th percentile in this set?

EXERCISE 10: Histograms of Other Land-Cover Proportions

- Return to the results of the landscape search, and now view the histogram for the Cultivated Crops land cover.

Q15 Consider the following:
(a) Are there substantial differences in the shapes of the histograms of Deciduous Forest and Cultivated Crops?
(b) Were you equally able (or unable) to anticipate the look of the histograms of these land use categories?
(c) Now, consider other land-cover types and their histograms. Is there a pattern to the distribution of landscape percentages among the land covers in this data set?

Q16 If you selected another set with different criteria, do you expect that the shapes of these curves would be similar, nearly identical, or entirely different? That is, do you think that the histograms reveal a basic property of land cover in real-world landscapes?

Part 4. Selecting Study Landscapes Along Gradients

In this section, you will use Metaland to identify landscapes with a wide range of amounts of edge habitat, but only within a subset of landscapes having a certain mix of proportions of deciduous forest and agriculture. With a little additional analysis in a spreadsheet program, Metaland's output can be used to rapidly identify sampling sites along a gradient of landscape metric values. You will identify landscapes according to given characteristics, and use the output to identify landscapes for potential field sampling.

EXERCISE 11: Controlling for Proportion and Total Edge

1. Return to the home page of Metaland and select the NLCD 2001 data set.
2. Use criteria of Latitude and Longitude to select landscapes that have certain landscape proportions within the same rectangular region of New England that was viewed in the previous exercise. As you did earlier, use the selection criteria to specify **Latitude >40.5** and **Latitude <45.1**, and **Longitude >−80** and **Longitude <−69.5**. For the proportion criterion (located in the **Classes** part of

the interface), request landscapes having greater than 30% and less than 50% Deciduous Forest (class 41), and greater than 5% Developed, Open Space (class 21). When done correctly, this request should return more than 650 landscapes. This is the same set used in the Metric Finder section.

3. Refer to the student guide to access a folder containing images of all of the landscapes returned from the search. Make a note of the index field for five or ten of the landscapes that look interesting to you. The index field is the string that looks like: x689y141s2.
4. Return to the Metaland interface and click **View All Landscapes** to see a list of the landscapes for individual inspection. You can view any single landscape and its landscape metric values by clicking **View Landscape** on the right side of the page.
5. For each landscape, Fragstats produces two distinct files for metric values: one for landscape-level metrics and another for class-level metrics. For this lab, we have downloaded the full set of these metrics for these landscapes for you, and done some of the formatting and graphing addressed in the following steps.
6. Refer to the student guide to access the spreadsheet for this section, which contains a number of pages with formulas and graphs that you can use to evaluate the landscapes.
7. The first sheet contains the values as they were pasted into the spreadsheet immediately after downloading them from Metaland. The formulas in Sheet 2 take the values in Sheet 1 and rank each landscape with respect to each metric value. To help you see the sizes of the values, cell values are color coded to indicate whether a given landscape metric value is small, medium, or large in comparison to the other values for that metric in the given subset. Values that are marked "low" are shown in blue and are in the lowest 33% of metric values. Mid-sized values are those between the 33rd and 67th percentile of the metric values in the set and are colored yellow. The largest one-third of values for a given metric is shown in orange.
8. We will use this set of landscapes and metrics to identify a gradient of *Total Edge*. Select the *Total Edge* metric (labeled *te*) between different land use categories of all types. Beginning with this subset, use the page with metric rank values to identify landscapes that are at percentile 20, 40, 60, and 80 with respect to the amount of *Total Edge*. (There may be several landscapes marked as being at a given percentile, due to rounding.) Note the unique identifiers of these landscapes for the next step.
9. Return to the home page of Metaland and select **View Landscapes by Identifier**. Enter the unique identifiers for your four landscapes (from percentiles 20, 40, 60, and 80) in the text box to view their images on one page.

Q17 Judging by eye, do the four landscapes selected along the gradient have substantially different amounts of *Total Edge*? If you forgot the ordering of the landscapes by *Total Edge*, could you correctly sort them by eye for that metric? What aspects of the landscapes make them easier or harder to sort, by eye, according to *Total Edge*?

Q18 Return to the page to **View Landscapes by Identifier**. To the four landscapes you entered just above, add the identifiers of the five that you earlier identified as being interesting. View the summary statistics for these nine landscapes. With respect to *Total Edge*, how do these landscapes compare to those chosen along the gradient? Judging by visual inspection alone, can you correctly insert your four chosen landscapes along the gradient of *Total Edge*?

Part 5. Synthesis

EXERCISE 12: Identifying Landscapes Along Two Gradients

Imagine a situation in which you wanted to identify landscapes along two simultaneous gradients: for example, *Total Edge* (metric *te*) and *Effective Mesh Size* (metric *mesh*). Using the values, colors, and landscape identifiers found in the spreadsheet of Exercise 4, try to find a landscape with low values for each metric—say between the 10th and 15th percentile for each. Is there a landscape listed in the spreadsheet that satisfies these two criteria?

To understand the potential relationships between landscape metrics, the spreadsheet's graphing abilities can be employed. On one of the sheets of the spreadsheet, we have made a scatterplot of the values of the two metrics in the set of New England landscapes.

Suppose you wanted landscapes for a 3×3 field study, built using landscapes selected with low, moderate, and high values for each of the two metrics. The scatterplot indicates where metric values are low (say, between the 1st and 33rd percentile), moderate (between the 34th and 66th percentile), and high (between the 67th and 100th percentile). Suppose you wanted to identify ten landscapes for each of the nine combinations: for example, landscapes with low *mesh* and moderate *te* is one of the combinations.

Q19 Are there enough landscapes for each of the categories of your study? What potential risks are there in identifying these landscapes from the set?

Q20 How could you use the Metaland interface to accomplish this task of finding landscapes?

EXERCISE 13: Complementary Work with Metric Finder and Metaland

Metric Finder is linked to Metaland, and this linkage allows you to select a subset of landscapes in Metaland and explore and compare them in Metric Finder. The option to bring a subset of landscapes to Metric Finder is found on the same page in Metaland, where metric results can be downloaded for use in a spreadsheet.

Assume the role of an analyst tasked with identifying a set of landscapes, for a purpose developed by you or your instructor. For example, you might identify a set of landscapes in Colorado with no discernable human activity within any of the landscapes, with the ultimate goal of identifying conditions in those landscapes that could make a new park. Develop a protocol for developing and using your subset, in which you:

- Use landscape metric criteria to select a subset of landscapes for consideration
- Export metrics and analyze their values, then use this analysis to select a proposed set of metrics of interest for use within the subset
- Export the results to Metric Finder and use the interface to explore which metrics help you identify landscapes of interest for your park
- Return to Metaland with these new metrics and refine your selection

Write a report in which you describe the protocol and illustrate some of the landscapes that fit your criteria. Include an assessment of the similarities and differences of the landscape sets, the metrics that best describe them, and how the similarities or differences change when subsets of these landscapes are selected by geography.

EXERCISE 14: Designing Regional Comparisons

In your position at AMEC, you are interested in understanding landscape differences among regions. Beginning with the set of New England landscapes as one of two regions to consider, design, and implement a regional comparison of landscape pattern between New England and another region in the USA. Explain your sampling design and the rationale for your selected metrics, what relationships you expected to observe, and present the results of your study of regional differences. Some statistical analysis will be required, as you will be sampling from the 6-km × 6-km landscapes that cover the USA when evaluating regional differences. What are the advantages and limitations of interpreting landscape differences using Metaland and the NLCD data?

REFERENCES AND RECOMMENDED READINGS[1]

Cardille JA, White JC, Wulder MA et al (2012) Representative landscapes in the forested area of Canada. Environ Manag 49(1):163–173

*Cardille JA, Lambois M (2010) From the redwood forest to the Gulf Stream waters: human signature nearly ubiquitous in representative US landscapes. Front Ecol Environ 8:130–134. *Mines the Metaland database for 1992 using the "affinity propagation" algorithm, which clusters landscapes while selecting a representative of each group of landscapes. Supplementary material shows other landscapes that were evaluated as similar, using landscape metrics, to each of the chosen representatives.*

*Cardille JA, Turner MG, Clayton M et al (2005) METALAND: characterizing spatial patterns and statistical context of landscape metrics. BioScience 55:983–988. *Presents the goals of the*

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Metaland database, an overview of the Metaland interface, and maps and histograms of metrics across the state of Wisconsin, USA.

Cushman SA, McGarigal K, Neel MC (2008) Parsimony in landscape metrics: strength, universality, and consistency. Ecol Indic 8(5):691–703

*Eigenbrod F, Hecnar SJ, Fahrig L (2011) Sub-optimal study design has major impacts on landscape-scale inference. Biol Conserv 144:298–305. *Excellent demonstration of how poor sampling design can compromise the inference from a study that includes landscape pattern as a predictor—even causing the direction of relationships to switch. Tools like Metaland can be used to improve study design by quantifying the ranges and correlation structure of metrics used in an analysis and assuring that sample landscapes do not overlap in space.*

Frey BJ, Dueck D (2007) Clustering by passing messages between data points. Science 315(5814):972–976

*Fry JA, Coan MJ, Homer CJ et al (2009) Completion of the National Land Cover Database (NLCD) 1992–2001 land cover change retrofit product. US Geological Survey Open-File Report 2008-1379, p 18. *This document provides details on differences between the 1992 and 2001 NLCD data, and is a necessary reading for anyone comparing the two time periods. Differences in the data can interfere with the ability to detect change.*

Fry JA, Xian G, Jin S et al (2011) Completion of the 2006 national land cover database for the conterminous United States. Photogramm Eng Remote Sens 77(9):858–864

Gustafson EJ (1998) Quantifying landscape spatial pattern: what is the state of the art? Ecosystems 1(2):143–156

Homer C, Dewitz J, Fry J et al (2007) Completion of the 2001 National Land Cover Database for the Coterminous United States. Photogramm Eng Remote Sens 73(4):337

Jin S, Yang L, Danielson P et al (2013) A comprehensive change detection method for updating the national land cover database to circa 2011. Remote Sens Environ 132:159–175

*Kupfer JA (2012) Landscape ecology and biogeography: rethinking landscape metrics in a post-FRAGSTATS landscape. Progr Phys Geogr 36:400–420. *This paper summarizes the current pitfalls and limitations of landscape pattern analysis, arguing that the research community has focused on landscape structure at the expense of landscape function. It summarizes the state of affairs of landscape structure measures from several different perspectives, and makes clear recommendations for future research.*

Langford WT, Gergel SE, Dietterich TG et al (2006) Map misclassification can cause large errors in landscape pattern indices: examples from habitat fragmentation. Ecosystems 9(3):474–488

Li H, Wu J (2004) Use and misuse of landscape indices. Landsc Ecol 19(4):389–399

LoGiudice K, Ostfeld RS, Schmidt KA et al (2003) The ecology of infectious disease: effects of host diversity and community composition on Lyme disease risk. Proc Natl Acad Sci 100(2):567–571

McGarigal K, Cushman SA, Ene E (2012) FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Computer software program produced by the authors at the University of Massachusetts, Amhers. Available at the following web site: http://www.umass.edu/landeco/research/fragstats/fragstats.html

*Proulx R, Fahrig L (2010) Detecting human-driven deviations from trajectories in landscape composition and configuration. Landsc Ecol 25:1479–1487. *This paper offers a regional analysis of the relationship between two landscape metrics based on 16 Canadian landscapes that vary in land use and human disturbance, demonstrating the kind of analysis that can be facilitated by approaches similar to Metaland.*

Turner MG, Gardner RH (2015) *Landscape ecology in theory and practice*, 2nd edn. Springer, New York

Vogelmann JE, Howard SM, Yang L et al (2001) Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogramm Eng Remote Sens 67(6):650–662

Xian G, Homer C, Demitz J et al (2011) Change of impervious surface area between 2001 and 2006 in the conterminous United States. Photogramm Eng Remote Sens 77(8):758–762

Chapter 11
Using Spatial Statistics and Landscape Metrics to Compare Disturbance Mosaics

Monica G. Turner and Martin Simard

OBJECTIVES

The causes and consequences of disturbances are major research topics in landscape ecology (Foster et al. 1998; White and Jentsch 2001; Turner 2005, 2010). Disturbances are of particular interest because of their reciprocal interactions with landscape pattern—they both respond to and create spatial heterogeneity (Turner 1987). Understanding the disturbance-created landscape mosaic is important for conserving resources and biodiversity, anticipating potential consequences of global change on disturbance regimes (Turner 2010), and managing landscapes in ways that mimic attributes of natural disturbances or keep a landscape within its historic range of variability (Perera et al. 2004; Long 2009). The spatial patterns created by disturbances can also provide novel insights into the state and dynamics of a landscape (Fraterrigo and Rusak 2008). Thus, disturbance has been a primary focus of landscape ecology for a long time (e.g., Turner 1987).

Disturbance-created heterogeneity also provides an opportunity to compare and contrast different approaches for quantifying spatial variability. Disturbances differ in severity (effects on the biota) across the landscape, and this variation can be quantified using continuous measures (e.g., the amount or proportion of tree basal area killed by disturbance) or represented categorically (e.g., high vs. low severity). These different types of data require distinct methods of analysis. Spatial statistics characterize the spatial dependence in continuous data, whereas landscape metrics quantify spatial pattern in categorical data (Gustafson 1998). How ecological

M.G. Turner (✉)
University of Wisconsin—Madison, Madison, WI, USA
e-mail: turnermg@wisc.edu

M. Simard
Université Laval, Quebec City, QC, Canada

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_11

interpretations differ based on these methods, and whether they provide complementary or redundant insights, is often not considered in landscape studies. Most studies use one or the other kind of data but not both. In this lab, students will compare spatial patterns created by three different kinds of disturbance using spatial statistics and landscape metrics and then contrast the interpretations that emerge from these complementary approaches. Upon completion of this lab, students will have accomplished the following objectives.

Part 1. Using spatial statistics to compare mosaics generated by different disturbances:

1. Quantify and compare landscape patterns generated by fire, insect outbreaks, and clear-cut harvesting in forested landscapes of Greater Yellowstone (Wyoming, USA) using spatial statistics applied to continuous measures of disturbance severity;
2. Learn to use and interpret output from GS+, a commercially available program for spatial statistics; and
3. Gain experience fitting different theoretical models to empirical semivariograms of disturbance severity and understand how model choice can influence interpretation.

Part 2. Analyzing fire patterns using landscape metrics and spatial statistics:

1. Describe and compare landscape patterns generated by fire in Greater Yellowstone using landscape metrics applied to categorical measures of fire severity using FRAGSTATS;
2. Interpret the landscape metrics and evaluate the effects of different classification schemes on the numerical results and ecological interpretations; and
3. Compare results obtained using spatial statistics (on continuous data) and landscape metrics (on categorical data) for quantifying fire patterns and determine whether these approaches provide complementary or redundant ecological insights.

This is an advanced exercise recommended for two successive class periods. This lab assumes basic understanding of semivariance and correlograms (completion of Chapter 5), as well as landscape metrics and the use of FRAGSTATS (Chapter 4). The lab also requires a geostatistics software program GS+ (© Gamma Design 2015), which can be found online at https://www.gammadesign.com/. The software is available as a free 10-day demo version. Lastly, while not essential, familiarity with vegetation indices derived from remote sensing data (as in Chapter 1) is helpful.

INTRODUCTION

Disturbance has long been recognized as an important driver of landscape heterogeneity and integral to understanding landscape dynamics (Watt 1947). A **disturbance** is defined as a relatively discrete event that disrupts the structure of an ecosystem, community, or population and changes resource availability or the physical environment (White and Pickett 1985). Recent treatments of disturbance advocate for separating out environmental drivers, initial system properties, and physical and biological mechanisms by which the system is affected so that different kinds of disturbance can be more easily compared (Peters et al. 2011). Disturbances are interesting in that they both create and respond to spatial heterogeneity in the landscape, and this is one reason disturbance has received so much attention in landscape ecology (Turner 1987, 2010). Landscape ecological studies of disturbance focus on several aspects of these relationships (Turner and Gardner 2015).

A variety of attributes are used to characterize a disturbance regime. Included among these are the spatial location of the disturbance, the size and shape of disturbed patches, and the variation in disturbance severity within the affected area. In this exercise, you will compare the spatial patterns created by three different disturbance types—fires, bark beetle (Dendroctonae) outbreaks, and clear-cuts—in the Greater Yellowstone Ecosystem (GYE).

Study Area and Disturbances

The 80,000 km^2 GYE is centered on Yellowstone National Park (YNP) and straddles portions of Wyoming, Montana, and Idaho (Figure 11.1). The GYE is unique in interesting respects—most notably the extensive geothermal features and large populations of native wildlife for which the region is famous—but in many ways it is also representative of coniferous forest ecosystems in the northern US Rocky Mountains.

YNP encompasses ca. 9000 km^2, most of which lies on a high (elevation ca. 2100–2700 m) volcanic plateau with relatively gentle relief. Surrounding the plateau are higher, rugged mountains of various crystalline, sedimentary, and volcanic substrates, as well as broad river valleys and basins characterized by a semiarid climate. Approximately 80% of YNP is dominated by lodgepole pine (*Pinus contorta* var. *latifolia*) forest, although subalpine fir (*Abies lasiocarpa*), Engelmann spruce (*Picea engelmannii*), and whitebark pine (*Pinus albicaulis*) are locally abundant at high elevations. At lower elevations, Douglas-fir (*Pseudotsuga menziesii*) and aspen (*Populus tremuloides*) forests grade into sagebrush (*Artemisia* spp.) steppe and grasslands. The climate is characterized by cold, snowy winters and dry, mild summers.

Portions of the GYE have a history of intensive resource exploitation, including logging, grazing, market hunting, and mining (as in other parts of the Rocky Mountains), as well as an expanding wildland-urban interface on private lands. The GYE differs from much of the rest of the Rocky Mountain region, however, in that much of the pre-Columbian flora and fauna remain intact, in part because the GYE contains one of the largest tracts of wild, undeveloped land in the continental USA (Gude et al. 2006). This largely pristine condition makes Yellowstone uniquely suitable for research into natural landscape patterns and processes at multiple scales and comparisons with human-created patterns.

We will consider two components of the natural disturbance regime: fire and bark beetles. **Stand-replacing fire** is an important component of the natural disturbance

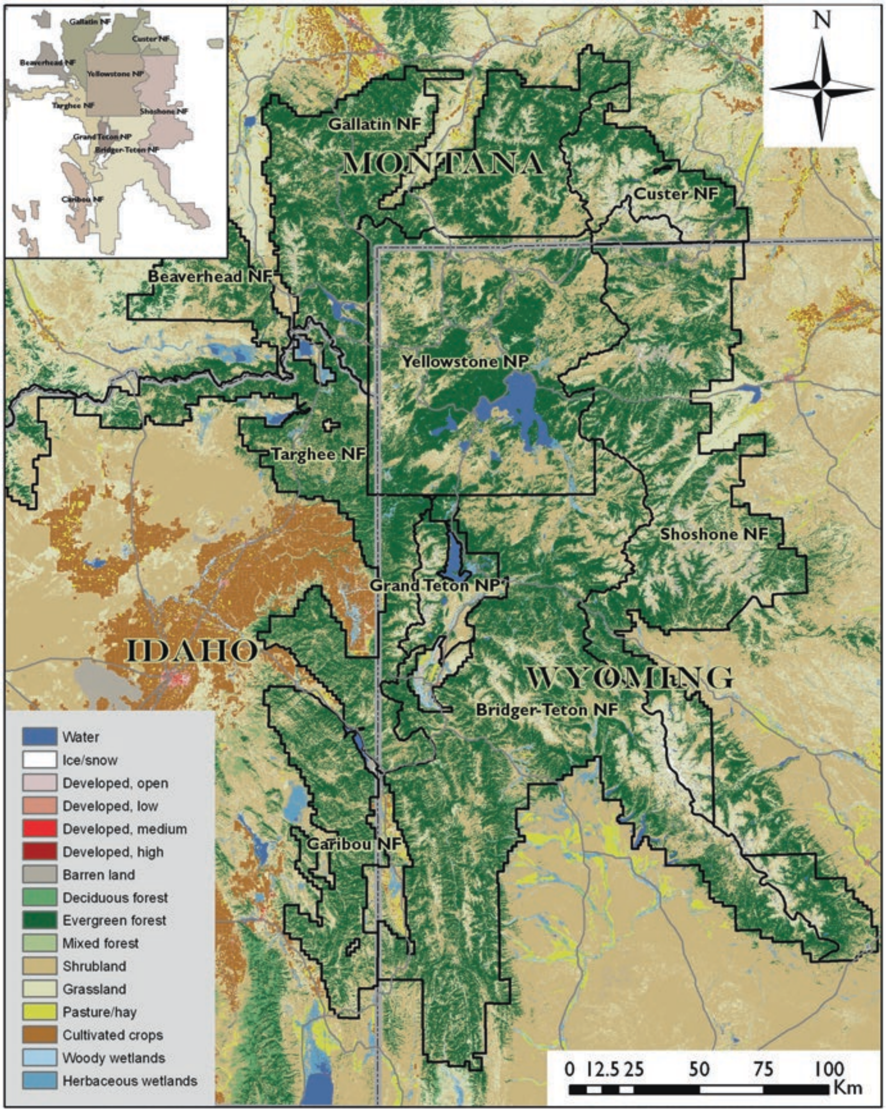


Figure 11.1 Location of the Greater Yellowstone Ecosystem, USA, from which the landscape maps of spatial patterns of disturbance used in these exercises were obtained

regime in the GYE and in many conifer forests throughout western North America (Turner and Romme 1994; Schoennagel et al. 2004). Infrequent, high-severity fires kill most of the trees either via intense surface fire or fire spread through the crowns of the trees. Climate, particularly severe regional drought, sets the stage for occasional years of extensive conflagrations. Stand-replacing fires occur in the GYE at 100–300-year intervals (Romme 1982; Romme and Despain 1989; Millspaugh et al. 2000;

Schoennagel et al. 2003). In 1988, large fires affected more than 30% of the GYE and created a complex spatial mosaic of burned patches that varied in size, shape, and severity (Turner et al. 1994). Fire weather was extreme: it was the driest summer on record in YNP (Renkin and Despain 1992), as dry cold fronts brought high wind and lightning but no rain. The 1988 fire season is considered to have ushered in a new era of fire activity in the West (Running 2006), and stand-replacing fires have continued to occur during many subsequent summers.

Bark beetles are also a key element of the natural disturbance regime in the GYE, and insect outbreaks have been recorded since 1922 (Furniss and Renkin 2003). Bark beetles are phloem-feeding specialists native to temperate and boreal coniferous forests (Raffa et al. 2008; Bentz et al. 2010). These insects are important because they kill healthy trees over extensive areas during episodic outbreaks. In contrast to fire, however, bark beetles do not kill all trees within a stand because they preferentially select larger trees. Furthermore, the forest floor and duff layer remain intact following bark beetle attack. Between 2003 and 2012, both mountain pine beetle (MPB) and Douglas-fir beetle (DFB) were active in lodgepole pine and Douglas-fir, respectively, in the GYE (Simard et al. 2012). Understanding the dynamics of both natural disturbances remains important throughout forests of western North America. The extent and severity of bark beetle epidemics have reached unprecedented levels (Raffa et al. 2008), and the frequency of large, severe fires continues to increase (Westerling et al. 2006, 2011). These trends are expected to continue because climate change—especially warmer temperatures, earlier snowmelt, and more severe summer droughts—is implicated for both disturbances.

We will also consider an anthropogenic disturbance: **clear-cut harvesting**. During the mid-twentieth century, many national forests in the GYE were subject to harvesting where merchantable trees were clear-cut in strips or patches distributed across the forest (Tinker et al. 2003). Forestry activities were particularly conspicuous along the western boundary of YNP where the differences in landscape patterns inside and outside of the national park provide an illustration of how human actions can change landscape patterns (see Figure 11.1).

Part 1. Using Spatial Statistics to Compare Mosaics Generated by Different Disturbances

In this first part of the lab, you will compare the spatial structure of the three different disturbance types (fires, bark beetle outbreaks, and clear-cut harvest) in the GYE. Five landscapes (5 × 5 km) of each disturbance type were sampled from disturbance maps generated from remote sensing data (Landsat, 30-m pixels, $n = 15$ different maps). These disturbance maps were created by analyzing differences between images taken before and after the disturbance event (Figure 11.2).

The fires and clear-cuts both date from the late 1980s, and the severity of these disturbances is represented here by a disturbance index based on the **Normalized**

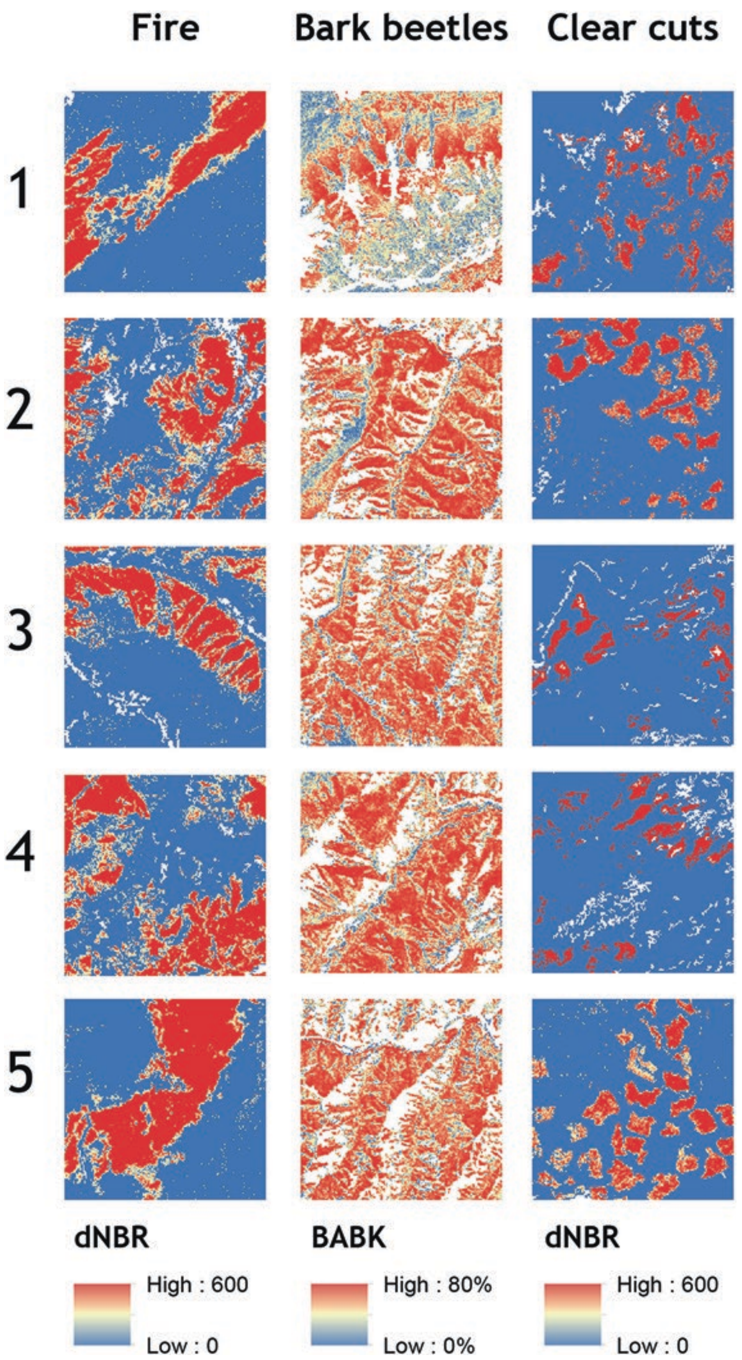


Figure 11.2 Landscape maps of spatial patterns created by three different types of disturbance in Greater Yellowstone corresponding to the data files used in Part 1. See text for description of map categories and interpretation. Areas of more severe disturbance are in redder colors

Burn Ratio (NBR). The NBR uses a band ratio (similar to the NDVI, described in Chapter 1) to measure disturbance severity within each grid cell based on the reduction in live biomass after the disturbance (Key and Benson 2006). The NBR uses two bands of Landsat that are most affected by fire and logging but in opposite ways. The near-infrared band is sensitive to decreases in forest biomass, and a shortwave infrared band captures increases in extent of bare soil. To create the index, these two bands are combined mathematically to calculate the NBR for each pre- and post-disturbance image. The post-disturbance NBR map is subtracted from the pre-disturbance NBR map to generate a difference map that contains values of **dNBR (delta Normalized Burn Ratio)**, which represents the change between the pre- and post-disturbance image. Here, we use the dNBR value as an index of disturbance severity. Unburned and unlogged areas show little change in NBR between the two image dates, and thus dNBR values are very low. Areas burned or clear-cut show large changes in NBR between the two image dates, and as a result, dNBR values are high. Greater values of dNBR correspond to higher disturbance severities (see Figure 11.2).

The bark beetle outbreak map was created using a similar approach with pre- and post-disturbance images but with a different disturbance index, the **Moisture Stress Index (MSI)**. The MSI is well suited to detect subtle changes in forest biomass, like those created by insect outbreaks that selectively kill some trees and do not directly affect the forest understory (Jin and Sader 2005). The **dMSI (delta Moisture Stress Index)** values were linked to field-based measurement of beetle-caused tree mortality to map the outbreak severity as measured by **cumulative basal area beetle-killed (% BABK)** between 1999 and 2007 (see Figure 11.2). This timeframe includes an outbreak first recorded in 2003 that lasted for several years. As with dNBR, greater values of dMSI indicate higher severity of the bark beetle outbreak (i.e., more beetle-induced tree mortality).

Although the remote sensing indices differ, dNBR and dMSI both quantify disturbance severity and return continuous values for each pixel. In all maps, non-forest pixels were removed from analysis, and undisturbed forest pixels were assigned an index value of zero.

EXERCISE 1: Describing Disturbance Patterns Qualitatively

Begin this lab by inspecting Figure 11.2 and considering those patterns while answering the next question. Your instructor may suggest that you complete this exercise prior to arriving in class.

Q1 Qualitatively describe the spatial patterns of the three disturbance types (Figure 11.2) in your own words. How would you characterize the pattern produced by each disturbance? How are they visually similar and different?

EXERCISE 2: Quantifying Disturbance Patterns Using Spatial Statistics

Next, describe the spatial patterns of each disturbance type using variograms and correlograms computed using GS+ software.

1. Thoroughly read the **handout Instructions for GS+** and familiarize yourself with the general steps and options in the software before proceeding further (Your instructor may suggest you read this prior to coming to class, to save valuable class time).
2. Generate an isotropic variogram and a correlogram (see GS+ handout) for each sample landscape using GS+. You will need to open each file separately in GS+. For each variogram, use an **exponential model** so that comparisons can be made using the same theoretical model. Remember to set the **Lag Class Distance** to 30 so that it matches the resolution of the data, which is 30 m. Be sure to record summary statistics (e.g., mean, minimum, maximum) for the response variable (Z) when you inspect the results, as well as the information described below.
3. In Table 11.1, record the parameter estimates listed below from each variogram. Save these results in an Excel worksheet so that you can produce the required graphs and use these results later in Part 2.
   - **Nugget** (C_0), the intercept of the model variogram representing variance at scales less than the minimum lag distance or variance arising from measurement error that cannot be resolved
   - **Sill** (C_0+C), or maximum semivariance
   - **Range** (A), which represents the limit of spatial dependence or the distance over which measurements are autocorrelated
   - **Proportion of structural variance** [$C/(C_o+C)$], which estimates the magnitude of spatial dependence in the data
   - **Fit** (R^2) of the model to the data

NOTE: When reporting numerical information, be consistent and appropriate in your use of significant digits. Do not report unrealistic precision, and be consistent in your summaries so the values can be easily compared.

4. Using the same landscape data file, compute the correlogram. In Table 11.1, record the parameter estimates below from each correlogram. Save the results in an Excel worksheet.
   - **Magnitude of the largest correlation** (usually r for the first lag distance)
   - **Lag distance** at which the correlogram declines to approximately $r=0.2$
   - **Lag distance** at which the correlogram declines to approximately $r=0$
5. Repeat the above steps for all landscapes ($n=15$).
6. Using the parameter estimates in Table 11.1, calculate the mean of each parameter for each disturbance type (so $n=5$ for each mean). Create a bar graph (with error bars of plus and minus one standard error for the 5 replicates) of the mean estimated **range** value (on the Y-axis) by disturbance type (classes on the X-axis). Create another bar graph for the mean values of the **proportion of structural**

Table 11.1 Results obtained for **exponential models** fit to semivariograms for the landscapes on which patterns of disturbance severity were analyzed in Part 1

		Semivariogram results					Correlogram results		
Disturbance	Replicate map	Fit (R^2)	Nugget (C_o)	Sill (C_o+C)	Prop. structural variance	Range (A), in m	Max. r	Lag (m) when $r\sim0.2$	Lag (m) when $r\sim0$
Fire	1								
	2								
	3								
	4								
	5								
Bark beetle outbreak	1								
	2								
	3								
	4								
	5								
Clear-cut	1								
	2								
	3								
	4								
	5								

variance, and the **lag distance** at which the correlogram declines below $r=0.2$ (thus, you will produce three graphs).

7. To compare results obtained from the variograms and correlograms, use the observations from Table 11.1 to create a scatterplot of the lag distance at which the correlogram declines to about $r=0.2$ (on the Y-axis) vs. the range estimate from the variogram (on the X-axis). For this plot, use all observations, irrespective of disturbance type ($n=15$). Add a linear trend line and show the coefficient of regression (R^2) and the equation.

Q2 Describe the spatial patterns of each disturbance type using the variograms and correlograms computed in GS+ and explain the relationship between the statistics and the qualitative description from ***Q1***. Use the results (Table 11.1) and graphs you produced, and reference these appropriately to support your interpretations and address these aspects:

- Make sure to discuss the overall magnitude of semivariance in the data, as well as the proportion of that variance that reflects spatial dependence (i.e., autocorrelation), the fit of the model, and the estimated distances (range) over which the response variables are autocorrelated.
- Use estimates of r from the correlograms when thinking about the strength of the autocorrelation.
- Inspect the scatterplot. Do the semivariograms and correlograms for the same map(s) lead to consistent interpretations of the spatial patterns?

Q3 Interpret the relationship between the spatial statistics and the ecological processes (fire, insect outbreak, and clear-cutting) associated with each disturbance type.

- Is there any consistent "signature" of these disturbance patterns that relates to the different mechanisms that generate them?
- Within a particular disturbance type, do the results suggest different conditions that may have influenced the disturbances? Use the parameter estimates in Table 11.1 and your graphs to answer this question.

EXERCISE 3: Understanding Differences Among Theoretical Models of Semivariance

Select **one** of the three disturbance types (your choice) to explore how quantitative results and interpretations can vary among different theoretical models (exponential, spherical, linear, and Gaussian) fit to empirical semivariance values.

1. For the five landscapes within the disturbance type that you selected, fit three additional theoretical models (spherical, linear, Gaussian) to the empirical semivariogram.
2. In Table 11.2, record the parameter estimates for each additional model (i.e., nugget, sill, range, proportion of structural variance, and fit) and in an Excel worksheet. Examine the range estimates and proportion of structural variance and compare their mean values among theoretical models.

Table 11.2 Results obtained for different theoretical models fit to empirical semivariograms of the same disturbance type.
Disturbance type = ________________________________

Replicate map	Model	Fit (R^2)	Nugget (C_o)	Sill ($C_o + C$)	Prop. structural variance	Range (A), in *m*
1	Exponential					
	Spherical					
	Gaussian					
	Linear					
2	Exponential					
	Spherical					
	Gaussian					
	Linear					
3	Exponential					
	Spherical					
	Gaussian					
	Linear					
4	Exponential					
	Spherical					
	Gaussian					
	Linear					
5	Exponential					
	Spherical					
	Gaussian					
	Linear					

Note that the results for the exponential model can be copied here from Table 11.1

Q4 How do your descriptions of spatial pattern and interpretations change with different theoretical models fit to the semivariance? Which of the theoretical models provides the best fit for the data, and is this consistent across each of the five landscapes? How will the choice of model influence the scientific conclusions that can be drawn from this analysis, and why is that important?

OPTIONAL EXERCISE 4: Kriging Disturbance-Severity Patterns

A valuable use of semivariance analysis is to interpolate expected values and generate a continuous map of the variable of interest. Select one of the disturbance types and use the results from the semivariogram analysis to krige a map of disturbance pattern. Depending on the disturbance you select, the kriged map will be a continuous landscape of dNBR or dMSI values that share the same statistical properties as the data used in computing the semivariance. How does the kriged map compare with the original map from which spatial dependence was quantified?

Part 2. Analyzing Fire Patterns Using Landscape Metrics and Spatial Statistics

Disturbance is often mapped in discrete classes (such as categories that represent both high and low severity and undisturbed areas). These classes are often derived from continuous measures, such as those used in Part 1, by setting a threshold value to convert the continuous variable to a category. The choice of threshold and how many categories to represent in a map of disturbance severity can influence the quantitative measures of spatial pattern, including patch size, patch shape, connectivity of disturbed area, distance to edge, etc.

In the exercises in Part 2, you will compute landscape metrics for fire patterns (from the same five landscapes of fire-created patterns in Part 1) using categorical data analyzed in FRAGSTATS. You will compare two categorization schemes, each with a different minimum dNBR threshold for defining burned areas (Figure 11.3), to explore the effects of classification scheme on numerical output. Binary maps of fire pattern (which use only two categories, burned vs. unburned) that were derived from the continuous maps of dNBR are provided for you. The dNBR maps were reclassified to categories using two alternate classification schemes. Maps of all burned areas used a low threshold (dNBR > 170) and include a range of burn severities, i.e., all areas affected by fire. Maps of high-severity burned areas used a high threshold (dNBR > 600) and include only the upper end of burn severities where fires would all be stand replacing.

EXERCISE 5: Selecting Landscape Metrics

For each of the ten landscapes (2 thresholds × 5 reps = 10 landscapes), you will use FRAGSTATS to characterize the spatial patterns of the burned landscapes. Choosing which metrics to quantify in any analysis is an important step, as each metric provides information about some aspect of pattern, and many metrics are redundant. You will include the five metrics listed below in your selection **plus** five others of your own choosing. You must have a rationale for selection of each and should choose indices that you think will have different (complementary) information about the spatial patterns.

- **Landscape-level metrics:** edge density (ED) and contagion (CONTAG)
- **Class-level metrics:** proportion of landscape occupied (PLAND), number of patches (NP), and mean patch size (AREA_MN)

Q5 Provide the rationale for each metric you selected for analysis. Which among these do you expect to be related to the variogram parameters, and how might they be related? Which metrics do you expect to provide new/different information about the spatial pattern? Explain your reasoning.

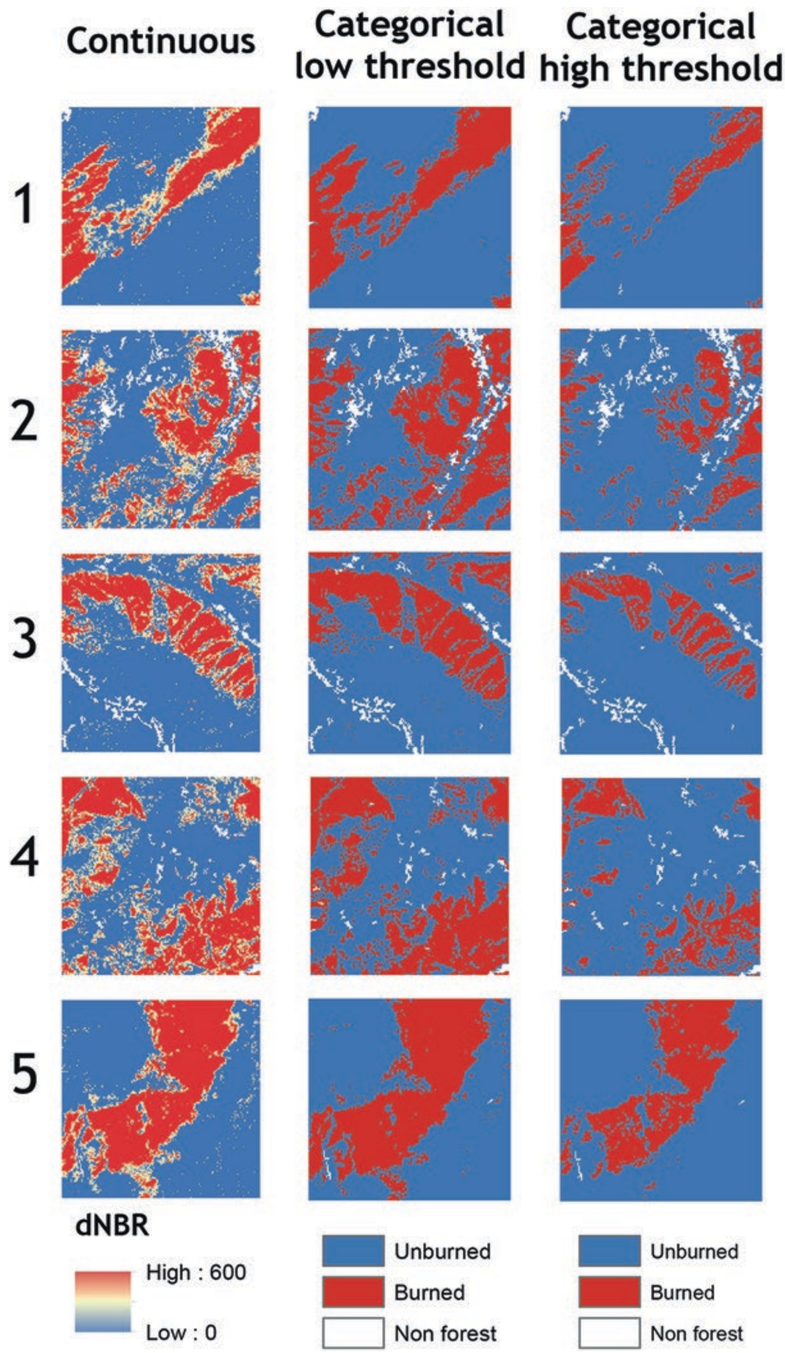


Figure 11.3 Landscape maps of spatial patterns created by fire in Greater Yellowstone. Continuous maps are the same as in Figure 11.2. The "low-threshold" maps identify all cells that were affected by fire (i.e., range of fire severities). The "high-threshold" maps identify only patches of high-severity fire and exclude areas of low-severity fire

EXERCISE 6: Quantifying Fire Patterns Using FRAGSTATS

Students are urged to consult the documentation for FRAGSTATS, which is excellent and can be downloaded from http://www.umass.edu/landeco/research/fragstats/fragstats.html.

It is assumed that students know how to run FRAGSTATS on their computer system (as you learned in Chapter 4). Use an **8-neighbor rule** for patch identification, and set the cell size to **30 m**. More details on the appropriate settings (rows, columns, cell size, etc.) are given in the **Instructions for FRAGSTATS**, which accompanies this chapter. Your output data (landscape metrics as well as class-level metrics for burned and unburned forest) should be saved in Excel for subsequent use. In these datasets, class 0 = unburned and class 1 = burned.

Once you have generated the output files from FRAGSTATS for all ten landscapes, compare the metric results for each of the classification schemes (mean with error bars, $n = 5$) using either graphs or a table. For class-level metrics, only include results for burned habitat. However, you will need metrics for burned and unburned habitat for some of the questions below.

Q6 Briefly describe the spatial pattern of the landscapes that used the low dNBR threshold vs. the maps of high-severity fire only (i.e., the high dNBR threshold) (Figure 11.3). How did the assignment of the low and high threshold affect the quantitative estimates of burn patterns? What is your ecological interpretation of these patterns, and does your interpretation change with the classification scheme?

EXERCISE 7: Using Spatial Statistics and Landscape Metrics to Compare Disturbance Mosaics

Next, you will compare and contrast the results from Part 1 (spatial statistics with continuous data) with the results generated above in Part 2 (analyzing categorical data). To do so, follow these steps:

1. Plot the landscape metrics (*Y*-axis) vs. the **range** estimates (*X*-axis) obtained from the semivariogram analyses in Part 1. When plotting class-level metrics, include both the burned and unburned categories on your plots by assigning a different symbol.
2. Plot the landscape metrics vs. the **proportion of structural variance** obtained from the semivariogram analyses in Part 1. When plotting class-level metrics, include both the burned and unburned categories on your plots by assigning a different symbol.
3. Although the sample size is small, compute correlation coefficients for each relationship and report a correlation table.

Q7 Given the pairwise comparisons generated between output from spatial statistics vs. landscape pattern indices, as well as how you interpreted disturbance patterns using these quantitative approaches separately, which of your expectations (from ***Q5***) were supported, and which were not supported? How are these analysis methods similar and different, and do they provide complementary or redundant information? Does your ecological interpretation change qualitatively with the methods used?

Q8 What spatial analysis approaches would you recommend for future studies of disturbance patterns and why? Are certain disturbance agents potentially best suited to continuous vs. categorical representation, and if so, why?

REFERENCES AND RECOMMENDED READINGS[1]

Bentz BJ, Regniere J, Fettig CJ et al (2010) Climate change and bark beetles of the Western United States and Canada: direct and indirect effects. BioScience 60(8):602–613

Despain DG (1990) *Yellowstone vegetation. Consequences of environment and history in a natural setting*. Roberts Rinehart, Boulder, p 239

*Fortin MJ, Dale MRT (2005) *Spatial analysis: a guide for ecologists*. Cambridge University Press, New York, NY, USA, p 369. *An extremely helpful reference book for applications of spatial statistics.*

*Foster DR, Knight DH, Franklin JF (1998) Landscape patterns and legacies resulting from large, infrequent forest disturbances. Ecosystems 1(6):497–510. *Classic paper describing how different natural disturbances create landscape patterns.*

*Fraterrigo JM, Rusak JA (2008) Disturbance-driven changes in the variability of ecological patterns and processes. Ecol Lett 11:756–770. *A terrific conceptual and practical overview of how to consider and quantify variance and why that is useful.*

Furniss MM, Renkin R (2003) Forest entomology in Yellowstone National Park, 1923–1957: a time of discovery and learning to let live. Am Entomol 49:198–209

Gude PH, Hansen AJ, Rasker R et al (2006) Rates and drivers of rural residential development in the Greater Yellowstone. Landsc Urban Plan 77:131–151

*Gustafson EJ (1998) Quantifying landscape spatial pattern: what is the state of the art? Ecosystems 1:143–156. *Another classic, this paper nicely contrasts continuous and categorical landscape data and analyses.*

Jin SM, Sader SA (2005) Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. Remote Sens Environ 94:364–372

Key CH, Benson NC (2006) Landscape assessment: sampling and analysis methods. In: Lutes DC, Keane RE, Caratti JF et al (eds) *FIREMON: fire effects monitoring and inventory system*. General technical report RMRS-GTR-164-CD, USDA Forest Service. Rocky Mountain Research Station, Ogden, pp LA1–LA51

Long JN (2009) Emulating natural disturbance regimes as a basis for forest management: a North American view. For Ecolo Manage 257(9):1868–1873

Millspaugh SH, Whitlock C, Bartlein PJ (2000) Variations in fire frequency and climate over the past 17 000 yr in central Yellowstone National Park. Geology 28(3):211–214

Perera AH, Buse LJ, Weber MG (eds) (2004) *Emulating natural forest landscape disturbances: concepts and applications*. Columbia University Press, New York

Peters DPC, Lugo AE, Chapin FS et al (2011) Cross-system comparisons elucidate disturbance complexities and generalities. Ecosphere 2:81

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Raffa KF, Aukema BH, Bentz BJ et al (2008) Cross-scale drivers of natural disturbances prone to anthropogenic amplification: the dynamics of bark beetle eruptions. BioScience 58(6): 501–517

Renkin RA, Despain DG (1992) Fuel moisture, forest type, and lightning-caused fire in Yellowstone National Park. Can J For Res 22(1):37–45

Romme WH (1982) Fire and landscape diversity in subalpine forests of Yellowstone National Park. Ecol Monogr 52(2):199–221

Romme WH, Despain DG (1989) Historical perspective on the Yellowstone fires of 1988. BioScience 39:695–699

Running SW (2006) Is global warming causing more, larger wildfires? Science 313:927–928

Schoennagel T, Turner MG, Romme WH (2003) The influence of fire interval and serotiny on postfire lodgepole pine density in Yellowstone National Park. Ecology 84(11):2967–2978

Schoennagel T, Veblen TT, Romme WH (2004) The interaction of fire, fuels, and climate across rocky mountain forests. BioScience 54:661–676

*Simard M, Powell EN, Raffa KF (2012) What explains landscape patterns of tree mortality caused by bark beetle outbreaks in Greater Yellowstone? Glob Ecol Biogeogr 21:555–567. *An empirical study of bark beetle outbreak severity, this paper gives background on the outbreak patterns in the landscape used here.*

*Tinker DB, Romme WH, Despain DG (2003) Historic range of variability in landscape structure in subalpine forests of the Greater Yellowstone Area, USA. Landsc Ecol 18:427–439. *Another Yellowstone paper, this study contrasts logging patterns and fire-generated patterns.*

Turner MG (ed) (1987) *Landscape heterogeneity and disturbance*. Springer, New York, p 239

Turner MG (2005) Landscape ecology: what is the state of the science? Annu Rev Ecol Evol Syst 36:319–344

Turner MG (2010) Disturbance and landscape dynamics in a changing world. Ecology 91(10): 2833–2849

Turner MG, Romme WH (1994) Landscape dynamics in crown fire ecosystems. Landsc Ecol 9(1):59–77

Turner MG, Gardner RH (2015) *Landscape ecology in theory and practice*, 2nd ed. Springer, New York

*Turner MG, Hargrove WH, Gardner RH et al (1994) Effects of fire on landscape heterogeneity in Yellowstone National Park, Wyoming. J Veg Sci 5:731–742. *One of the very first papers to quantify postfire landscape heterogeneity, this paper presents landscape metrics for areas affected by the 1988 Yellowstone fires.*

Watt AS (1947) Pattern and process in the plant community. J Ecol 35:1–22

Westerling AL, Hidalgo HG, Cayan DR et al (2006) Warming and earlier spring increase western U.S. forest wildfire activity. Science 313:940–943

Westerling AL, Turner MG, Smithwick EAH et al (2011) Continued warming could transform Greater Yellowstone fire regimes by mid-21st century. Proc Natl Acad Sci 108:13165–13170

White PS, Jentsch A (2001) The research for generalities in studies of disturbance and ecosystem dynamics. Prog Bot 62:399–450

White PS, Pickett STA (1985) Natural disturbance and patch dynamics: an introduction. In: Pickett STA, White PS (eds) *The ecology of natural disturbance an patch dynamics*. Academic Press, Orlando, pp 3–13

Part IV
Applications for Conservation and Assessing Connectivity

A variety of recent developments now better support conservation planning and connectivity assessment. This module explores these tools and presents a user-friendly start for those less familiar with these approaches. Chapter 12 introduces the basic uses of network analysis (aka graph theory) for assessing connectivity from the perspective of different species and assumes no prior knowledge of network concepts and terminology. Chapter 13 introduces the basic components of Marxan, one of the most widely used conservation planning tools for reserve design decisions. The last two chapters in this module provide advanced tools most relevant for research in this arena. Chapter 14 explores advanced graph theoretic approaches using Conefor software, which is used widely throughout Europe for conservation planning. This software can assess network connectivity while incorporating habitat quality and identify critical source areas and stepping stones for different species. Chapter 15 requires the use of R software and demonstrates the statistical ways meta-communities can be examined across a landscape, incorporating important multispecies perspectives.

Chapter 12
Assessing Multi-Scale Landscape Connectivity Using Network Analysis

Todd R. Lookingbill and Emily S. Minor

OBJECTIVES

Landscape connectivity has implications for many ecological processes, including spread of invasive species and conservation of native ones. Because species have different minimum area requirements and different movement abilities, landscape designs suitable for one species (or group of species) may be inappropriate for other species. Methods from network analysis can be used to combine information on landscape pattern and species life history characteristics for species-specific assessments of potential connectivity. The lab is intended to provide students with the following:

1. An introduction to the concepts of landscape connectivity and network analysis;
2. Practice defining the basic elements of the landscape network including nodes, links, and components;
3. Exploration of simple measures of connectivity related to dispersal, home range, and species persistence;
4. Ways to construct and compare landscape networks for different species with differing perceptions of the landscape;
5. Consider how protected areas form networks of potential connectivity; and
6. Discuss simplifying assumptions of the approach and how methods for quantifying connectivity may differ in contrasting landscapes.

T.R. Lookingbill (✉)
University of Richmond, Richmond, VA, USA
e-mail: tlooking@richmond.edu

E.S. Minor
University of Illinois at Chicago, Chicago, IL, USA

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_12

Exercises in Part 1 assess potential connectivity on a highly fragmented simulated landscape for two species with different life history characteristics, using pencils/markers and the provided map. You will then calculate and compare two simple metrics of landscape connectivity for these two networks. In Part 2, you will examine the consequences of landscape connectivity for protected lands in the Willamette Valley ecoregion of the United States in a more realistic example. You will examine three species with differing minimum area requirements and movement abilities (ranging from large carnivores to rodents). While a variety of software programs can automate most of these analysis procedures, it is worthwhile to construct your first networks "by hand" as you learn the concepts and calculations; thus, the exercises in this chapter have been simplified and do not require a computer. Two subsequent chapters in this text (see Chapters 14 and 20) provide more detailed applications of network analysis that build on and assume familiarity with the concepts in this lab.

INTRODUCTION

One of the biggest problems in conservation biology and biogeography today is the alteration of landscapes and loss of native habitat (Richardson and Whittaker 2010). Landscape fragmentation has potential implications for many ecological processes. As habitats become more fragmented and separated, and the intervening matrix becomes more dangerous and inhospitable, native populations experience a loss of genetic variation and/or permanent extinction. Barriers to immigration and recolonization may be especially detrimental as species attempt to adapt to other stressors such as those associated with climate change. As a result, it has become essential to accurately measure landscape connectivity and understand its effects on major ecological and evolutionary processes. Landscape ecology offers specific tools for the quantitative study of landscape connectivity among fragmented habitat patches.

Landscape Connectivity

Landscape connectivity is a measure of how well the landscape facilitates or impedes movement among resource patches (Taylor et al. 1993). In fragmented environments, connectivity of habitat patches is important for movement of genes, individuals, populations, and species over multiple time scales (Fahrig and Merriam 1985). Over the short term, it affects the success of juvenile dispersal and thus recolonization of empty habitat patches (Clergeau and Burel 1997). At intermediate temporal scales, connectivity affects migration, persistence of metapopulations (Hanski and Gilpin 1991; Ferreras 2001), and genetic diversity (Dixo et al. 2009; Angelone and Holderegger 2009). Over longer time frames, connectivity influences the ability of species to adapt, expand, or alter their ranges in response to climate change (Lyford et al. 2003; Opdam and Wascher 2004). Habitat connectivity is especially important when habitat is degraded, rare, fragmented, or otherwise sparsely distributed (Flather and Bevers 2002; King and With 2002; Fischer and Lindenmayer 2007).

Landscape connectivity can be defined in many ways (Calabrese and Fagan 2004). **Structural connectivity** refers simply to landscape pattern and is not necessarily associated with the movement behavior of any particular organism. **Functional connectivity**, on the other hand, includes information on the movement of organisms in response to landscape pattern; this is a species-specific measure of connectivity. Functional connectivity may take two forms: **actual connectivity**, which requires detailed observations of the movement of individuals, and **potential connectivity**, where life history data on mobility are used to estimate movement pathways. Potential measures of connectivity, such as those derived from network analysis, are thought to be the most cost effective for addressing questions of basic ecology and applied natural resource management in both terrestrial and marine ecosystems (Calabrese and Fagan 2004; Grober-Dunsmore et al. 2009).

It is also necessary to distinguish between **landscape connectivity**, where connectivity is seen as a property of the entire landscape, and **patch connectivity**, where connectivity is seen as a patch-level attribute (Kindlmann and Burel 2008). Within a landscape, each patch may have a different level of connectivity—some may be highly connected to other patches while others may be completely isolated. Methods of network analysis, based on the mathematics of graph theory, are useful here as well. The patch-based data structure lends itself naturally to assessment at both of these levels.

Network Analysis

Network analysis has been proposed as a simple solution to unify and evaluate multiple aspects of habitat connectivity (Kadoya 2009; Urban et al. 2009). Although relatively recently introduced to landscape ecology (Urban and Keitt 2001; Jordan et al. 2003), network analysis is a well-developed body of research often used in the computer and social sciences that quantifies connectivity and flow in networks (Harary 1969). A **network** is a set of **nodes** (points) connected by **links** (lines); a link between points indicates a connection between them. In the case of landscape networks, nodes represent habitat patches or local populations, and links indicate interaction or dispersal among populations (Figure 12.1). The approach can quantify either structural or functional connectivity, but, because it typically uses information on dispersal processes to define patch connections, it is especially useful for quantifying potential connectivity.

Nodes and links are the two fundamental elements that define the landscape network. Discrete patches of habitat are represented as nodes, invoking an island view of discrete habitat islands in a “sea” of nonhabitat (i.e., the matrix). The network approach connects patches with links if they are within some user-specified (and

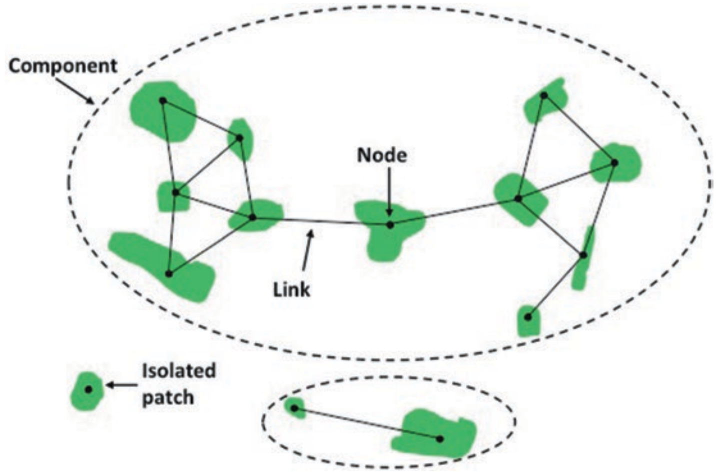


Figure 12.1 An example of a landscape network with network features identified. Links are drawn from node to node (i.e., *patch centroids*), as is conventional, even though the distances are measured from patch edge to edge. All patches (*shaded polygons*) in this example are considered large enough to be suitable habitat for the species of study. Largest component = 0.77; link density = 0.18; degree centrality of labeled node = 2 links; domain of labeled node = 10 nodes

preferably ecologically relevant) distance of each other or connected by corridors. Networks can be represented graphically, as in Figure 12.1, or in a connectivity matrix. A connectivity matrix is a table with information about the connectivity (or lack thereof) of every pair of nodes. The links might be binary (connected or not) or they might be weighted, specifying the strength of a connection between two nodes. For example, the weights might represent geographic distance or likelihood or rate of dispersal. Links also might indicate movement in multiple directions among patches or they might be directional, designating preferred pathways of flow.

Scores of metrics are available for characterizing connectivity based on properties of the network (Pascual-Hortal and Saura 2006; Bodin and Norberg 2007; Kindlmann and Burel 2008). Some of these are relevant to properties of the entire landscape—i.e., landscape connectivity. Others assess connectivity of an individual habitat patch—i.e., patch connectivity (Table 12.1). At the landscape scale, groups of connected patches are called **components** (Figure 12.1). By definition, dispersal can occur among patches within a component but not among patches of different components. One rather intuitive measure of landscape connectivity is an index of the size of the **largest component**, simply calculated as the proportion of suitable habitat on the landscape that is connected within the largest component (Ferrari et al. 2007). This metric is informative about the potential for large-scale population processes on the landscape: many small components suggest isolated subpopulations, while large components suggest a well-mixed population.

Another basic measure of landscape-level connectivity is the link density of the network. **Link density** is defined as $L/[n(n-1)/2]$ where L and n are the number of

Table 12.1 Definitions of network connectivity metrics used in these exercises

	Term	Units	Definition
Landscape-level metrics	*Largest component*	Unitless	Area of habitat contained in the largest component (H_{LC}) divided by the total amount of suitable habitat area (H_T) where only patches ≥ minimum size are considered suitable. H_{LC}/H_T
	Link density	Unitless	Number of links (L) in the network divided by the maximum number of links possible. $L/[n(n-1)/2]$; where n = number of suitable nodes
Patch-level metrics	*Degree centrality*	Links	Total number of links for a node. This is a very local measure of patch connectivity (i.e., only accounts for nearest neighbors)
	Domain	Nodes	Total number of nodes reachable from the node. This is a larger-scale measure of patch connectivity (i.e., extending to entire component)

links and nodes in the network, respectively (Royer et al. 2008). The denominator represents the maximum POSSIBLE number of links of links in the network. Landscapes with a large number of links relative to the number of patches should be well connected. The more links, the greater the redundancy in the network and the less vulnerable the landscape to the loss of any individual connection (e.g., through loss of habitat or addition of a dispersal barrier such as a road). Systematically removing either the nodes (Urban and Keitt 2001) or links (Lookingbill et al. 2010) from a network and evaluating the effects on connectivity can be an informative exercise for evaluating the vulnerability of landscapes to habitat loss and fragmentation.

Connectivity metrics also can be used to examine more localized issues of patch occupancy, population stability, and genetic diversity. Patch-level metrics are useful for this purpose because they can quantify the structural importance of habitat patches within the landscape network (Galpern et al. 2011). Two commonly used measures are degree centrality and domain (Table 12.1). **Degree centrality** is the number of direct connections for a given habitat patch (nearest neighbors). This is ecologically similar to the number of patches within a given distance or patch density (e.g., van Dorp and Opdam 1987). A **hub** is a node with very high degree (i.e., a patch with many neighbors), while an **isolated node** has no neighbors (Figure 12.1). From an applied perspective, hubs might be identified and targeted for protection to facilitate rapid species migration. **Domain** is a measure of the number of other nodes that are reachable from a node, which is equivalent to the size of the component containing the node (De Nooy et al. 2005). While these patch-level metrics describe connectivity of individual patches, they measure connectivity at different scales. Degree measures connectivity at the most local scale (the number of immediate neighbors), while domain measures how connected a patch is to the broader landscape.

Conservation Networks

Landscape ecology has become highly invested in habitat connectivity and its implications for populations (Tischendorf and Fahrig 2000; Fahrig 2003; Calabrese and Fagan 2004; Crooks and Sanjayan 2006; Fischer and Lindenmayer 2007), and the application of network analysis has erupted over the last few years. Network analysis offers a valuable set of tools for conservation that are visually intuitive, computationally efficient, and easily interpretable for conservation management. In particular, a network representation is often used to invoke a metapopulation model with subpopulations interacting across a fragmented landscape (Estrada-Pena 2005). Network analysis has been applied to the design and assessment of reserve networks (Saura and Pascual-Hortal 2007; Minor and Lookingbill 2010), the identification of important movement corridors or habitat linkages (Jordan et al. 2003; Morzillo et al. 2011), and the detection of population sources and sinks (Minor and Urban 2007; Treml et al. 2008). Because the construction of the landscape network is dependent upon species life history characteristics, most of these applications are species specific.

Species have different perceptions of the landscape in which they live. These perceptions inform what constitutes a patch of suitable habitat, as well as the willingness to traverse through the matrix to a neighboring patch (Pe'er and Kramer-Schadt 2008). In network terms, the nodes and the links in any network will likely differ for different species on the same landscape. Large-scale conservation plans should consider landscape connectivity from the perspective of all relevant species groups. Generalizations for different species groups are valuable for multi-species applications such as designing green infrastructure or marine protected-area networks. For mammals, there is a strong linear relationship between what a species perceives as a patch (a function of home range size) and its maximum possible dispersal distance (Bowman 2003). Both of these factors scale with body size such that small mammals have small home ranges and short dispersal distances; intermediate-sized mammals have intermediate home ranges and intermediate dispersal distances; and large mammals have large home ranges and longer dispersal distances. This allometric scaling relationship can be used to construct landscape networks to evaluate connectivity of protected-area networks for general classes of mammals based on their body size (see West et al. 1997 for further background on allometric scaling laws).

Spatial data about the protected areas in the United States are available from the Conservation Biology Institute (CBI 2012). Such data could represent potential patches for use in a network analysis. For example, reserves could be identified that meet the minimum home range requirements of a particular group of species (e.g., 1000 ha for large mammals). Two reserves could be considered linked if they were as close as (or closer) than the maximum dispersal distance for the species group (e.g., 100 km for large mammals). Network representations could be constructed for different species groups, and the resulting networks would differ in both the spatial configuration of patches (e.g., the reserves included in a large-mammal protected-area network would be a subset of the reserves in a small-mammal protected-area network) and the rules used to create links (e.g., large mammals have the potential to travel farther than small mammals).

To learn the basic concepts of network analysis, Part 1 begins with a simulated landscape and hypothetical species to illustrate two metrics of landscape connectivity. In Part 2, we present protected-area networks based on CBI data for three mammals in the Willamette Valley ecoregion of the western United States. The three networks facilitate comparisons among species inhabiting the same landscape in order to illustrate metrics of patch connectivity.

Part 1. Introduction to Landscape-Level Connectivity

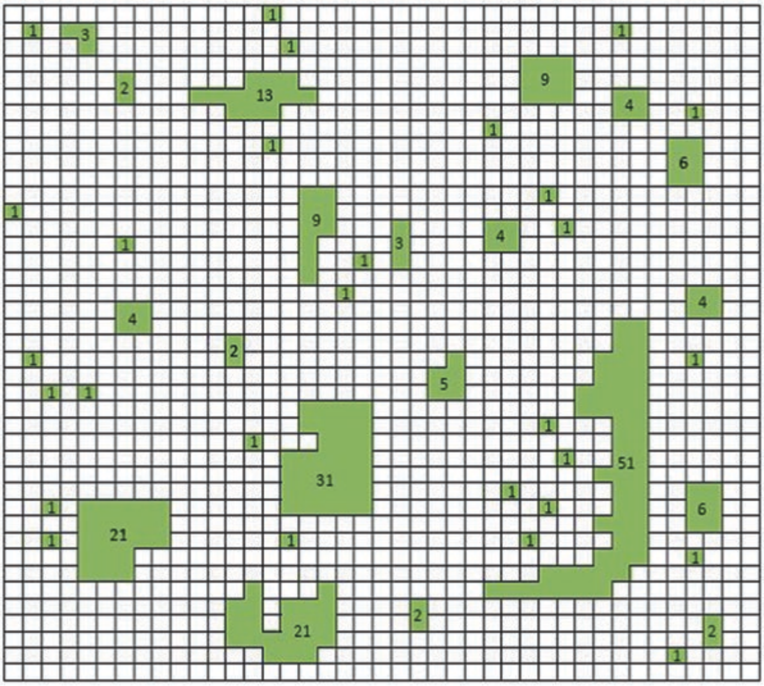


Figure 12.2 Hypothetical landscape for Exercise 1. Habitat patches are labeled according to their size. The total number of habitat cells (*shaded*) on the landscape is 230 divided among 48 patches. Note that not all of this area is suitable for both species and depends on the size of the patch

EXERCISE 1: Constructing Networks by Hand

In this exercise, you will assess potential connectivity for two species on a simulated landscape with a high level of fragmentation (13% of the habitat remains on the landscape). Your first step is to construct the landscape networks using a pencil, the map found in Figure 12.2, and the life history characteristics described below. You will then calculate two simple metrics of landscape connectivity and compare the values for the two networks. Although computer programs can be used to automate most of these processes, it is a worthwhile exercise to practice constructing a network by hand.

1. Print (or copy) Figure 12.2.
2. From this map of potential habitat, identify the suitable habitat patches (or nodes) that meet the minimum area requirement for a hypothetical small mammal (see Table 12.2). Shade or color these areas.
3. Next, on either separate pieces of paper or in different colors on the same map, identify and shade the patches that meet the minimum area requirements for a large mammal as defined in Table 12.2 (i.e., identify its suitable habitat).
4. Using the maximum dispersal distances (defined in Table 12.2), draw the links among nodes of suitable habitat for each of the two species. Links should be drawn and measured as the shortest distance between patches. To do this, start counting from the edge of a habitat patch; diagonal distances count as only one cell length so it is possible to move in eight directions from a given cell (i.e., an 8-neighbor rule). If the edge of a new habitat patch is reached within the allowable number of steps, then the two patches are connected. Dispersal can occur across any kind of cell (habitat or nonhabitat).
5. Using the definitions provided in Table 12.1, calculate the **largest component index** and **link density** for each of the two species and add these values to Table 12.2.

Q1 Which species had the greatest connectivity for this landscape (by measure 1 (the largest component index), by measure 2 (link density)? Show your work.
  (a) How do you think the differences in landscape-level connectivity would affect the long-term persistence of the two species on the sample landscape?
  (b) Landscape connectivity measures by themselves may not provide sufficient information for a species conservation plan. How might the total amount of habitat and number of patches in each of the networks also be important for the long-term persistence of a species? What other information about the landscape would be useful?

Table 12.2 Landscape-level connectivity metrics for Exercise 1

	Minimum habitat requirement (cells)	Maximum dispersal distance (cells)	Largest component index	Link density
Small mammal	1	2		
Large mammal	16	32		

The largest component index and link density should be calculated for the two species and inserted in the table

Q2 Define the concept of an umbrella species as it is applied to conservation (you may need to look this term up in an ecology text if you not familiar with it).

(a) What can you say about the ability to generalize about connectivity from one species to another based on the results from this exercise?

(b) Given what you have observed about the habitat connectivity and total amount of habitat area for these two species on the sample landscape, what specific conservation actions would you recommend for the small mammal? for the large mammal?

Q3 A developer would like to remove some of the habitat to create a subdivision on a parcel of property, four cells in area, somewhere on the landscape.

(a) Where would be the most detrimental place(s) to locate this subdivision (circle on Figure 12.2)? Would this location be the same for the two species? Provide a rationale for your choice(s).

(b) To offset development elsewhere, the developer is required by law to create four cells of new habitat somewhere on the landscape. What would be the best strategy for adding this habitat to the existing map (ignore the loss of habitat described in the previous question)? Options may include adding property to existing patches, creating new patches, random placement of new habitat, etc. You should consider multiple strategies and the potential impacts to each of the two species, but ultimately a total of only four new habitat cells will be created. Draw the four cells on Figure 12.2 and provide a rationale for your choice.

Part 2. Consequences of Connectivity

EXERCISE 2: Analyzing Landscape- and Patch-Level Connectivity

In this exercise, you will examine the consequences of landscape connectivity among protected lands in the Willamette Valley ecoregion of the United States from the perspectives of a regional planner, wildlife biologist, and pathologist. A map of the ecoregion derived from the Conservation Biology Institute (CBI) GIS-based Protected Areas Database is provided (Figure 12.3).

Only sites with a land stewardship status 1 or 2 in the USGS GAP analysis program will be considered strictly "protected" in your analysis (Scott et al. 2001), thus ensuring the highest level of biodiversity protection. Any adjacent protected areas with differing ownership, but sharing a boundary, were merged into a single unit (i.e., patch or node) for subsequent network analysis.

The potential connectivity of three different species has been analyzed for this map of protected areas (Figure 12.4). Each species has different minimum area requirements and movement abilities which are related as per Bowman (2003).

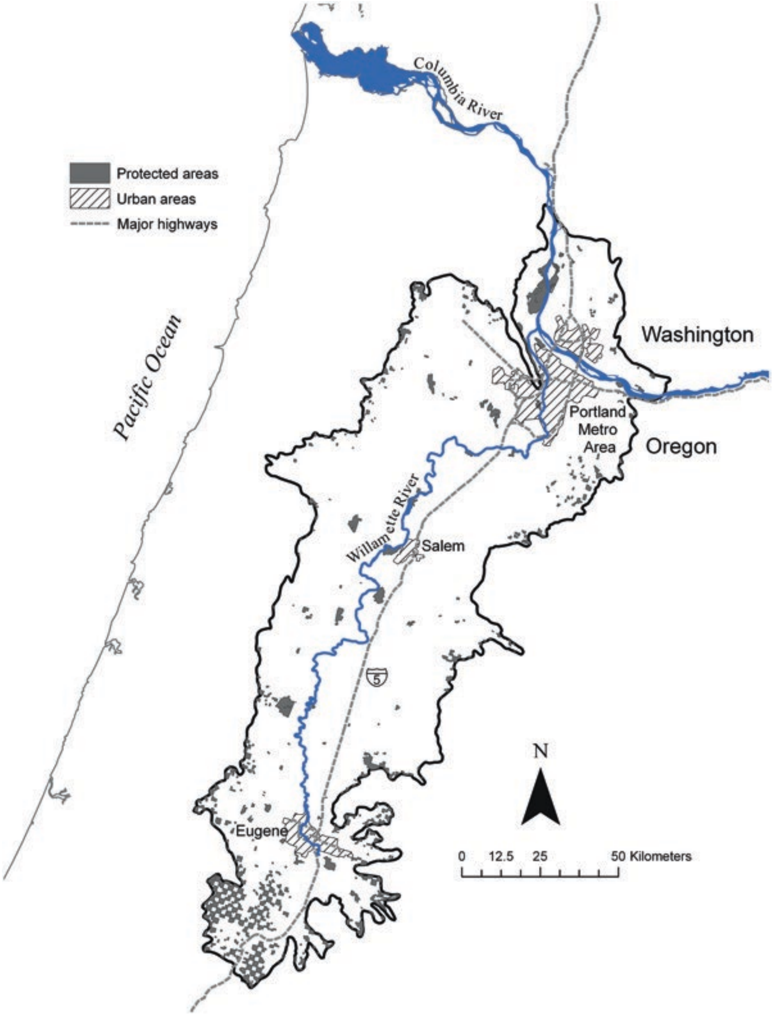


Figure 12.3 Overview map of the Willamette Valley ecoregion showing prominent landscape features and all protected lands

Tables of landscape- (Table 12.3) and patch-level (Table 12.4) connectivity metrics have been provided for each of the species. You will use this information to answer questions about the size and spatial configuration of the protected areas for the three species, and how the life history differences among species should influence their management.

1. Examine the three separate representations of potential connectivity based on the distribution of protected areas (Figure 12.4). One is for wolves with a home range of 1000 ha and a dispersal distance of 100 km; a second is for foxes with a

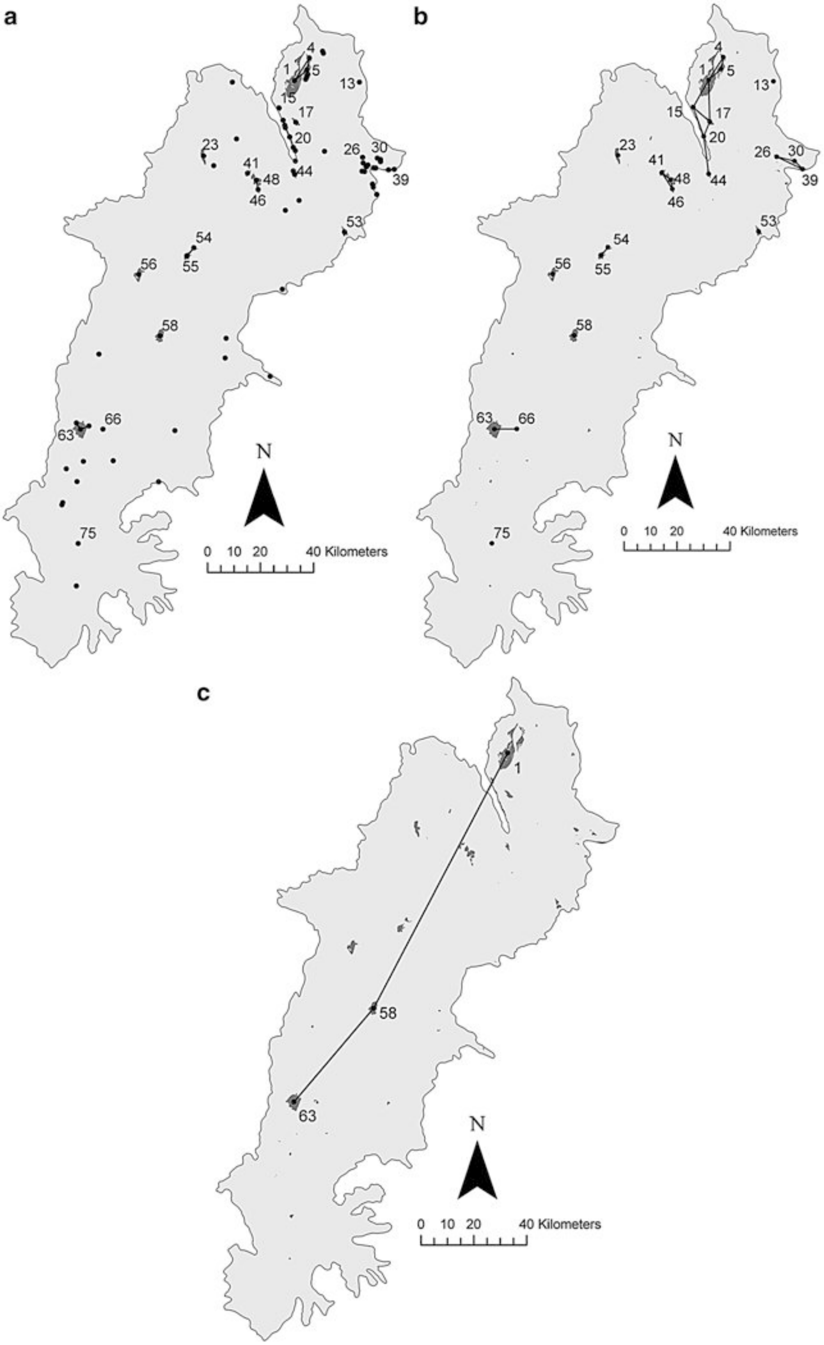


Figure 12.4 Landscape networks of potential connectivity for protected areas in the Willamette Valley ecoregion calculated for mice (*Panel a*), foxes (*Panel b*), and wolves (*Panel c*) based on the parameters provided in Table 12.3. Nodes represent protected lands of sufficient size to provide suitable habitat for the species and do not represent actual presence or absence of species

Table 12.3 Landscape-level connectivity metrics for protected areas in the Willamette Valley ecoregion calculated for three species

	Home range (ha)	Maximum dispersal distance (km)	Largest component index	Link density
Mouse	1	1	0.35	0.02
Fox	100	10	0.45	0.07
Wolf	1000	100	1.0	0.67

Table 12.4 Patch-level connectivity metrics for protected areas in the Willamette Valley ecoregion calculated for three species

		Mouse		Fox		Wolf	
Node	Size (ha)	Degree (links)	Domain (nodes)	Degree (links)	Domain (nodes)	Degree (links)	Domain (nodes)
1	4499	2	9	4	6	1	2
4	222	2	9	2	6		
5	783	9	9	2	6		
13	101	0	0	0	0		
15	166	0	0	3	6		
17	404	0	0	3	6		
20	249	2	4	3	6		
23	752	0	0	0	0		
26	206	0	0	2	2		
30	127	2	2	2	2		
39	146	1	1	2	2		
41	231	0	0	2	2		
44	238	1	1	1	6		
46	867	1	1	2	2		
48	133	1	1	2	2		
53	358	0	0	0	0		
54	184	1	1	1	1		
55	539	1	1	1	1		
56	959	0	0	0	0		
58	1108	0	0	0	0	2	2
63	2172						
66	133						
75	132						
Mean		1.2	2.0	1.5	2.5	1.3	2.0

For simplicity of presentation, values are provided for only a subset of the 76 nodes contained within the mouse-based network; mean values have been calculated over all nodes. Blanks have been left in the table for the values for the last three nodes and should be filled in as part of Exercise 2. *NOTE:* Mean values are provided as the average number of links per node and average number of nodes reachable per node home range of 100 ha and a dispersal distance of 10 km; and the third is for mice with a home range of 1 ha and a dispersal distance of 1 km.

2. Examine the tables of connectivity metrics for each of the three networks (Tables 12.3 and 12.4). In particular, note that Table 12.4 provides information for individual patches. These values have been calculated using Pajek (De Nooy et al. 2005) a freely available Windows-based program for analyzing large networks (http://pajek.imfm.si/).
3. Fill out the missing values in the three bottom rows of Table 12.4. Use the information from the maps and tables to answer the questions below.

Q4 Compare the overall landscape-level connectivity for the three different species (i.e., largest component index and link density).
- (a) Which species would have the highest connectivity assuming it relied solely on protected lands? Which species is potentially least connected?
- (b) How do these results compare to your expectations from Exercise 1? Do you see similar or dissimilar patterns?

Q5 Compare the mean patch-level connectivity metrics for the three species (i.e., degree centrality and domain).
- (a) Which species seems to be best connected by these measures? Which species is potentially least connected?
- (b) How does this compare with your assessment from Question 4? Which species would you expect to experience the most problems due to isolation in the ecoregion according to these results?

Q6 Imagine you were asked to prioritize regional spending for habitat improvement on protected lands. Consider the relative importance of a specific patch (patch 1) to connectivity.
- (a) Is the patch equally important to the overall, broad-scale connectivity of the landscape for all three species relative to other protected areas in the ecoregion? Are other patches more important?
- (b) Is the patch equally important to local dispersal movement for all three species relative to other protected areas in the ecoregion?

Q7 Imagine you were a wildlife biologist tasked with establishing a reintroduction program for an endangered species of fox. Which patch would be the most logical location to transplant new individuals to maximize rapid dispersal of the species to other nearby patches? Include in your justification a statement about which metric is most important for this type of decision.

Q8 Consider a mouse-borne pathogen that threatens humans and has begun to invade the ecoregion. Which patches would be logical locations to focus eradication efforts to try to control the spread of this disease? Include in your justification a statement about which metric is most important for this type of decision.

SYNTHESIS

Q9 For learning purposes, several simplifying assumptions were made in this exercise. For example, what implicit assumptions were made about the matrix when constructing these networks? What types of land covers and landscape features might violate these assumptions?

Q10 Patch 1 in Exercise 2 represents the Sauvie Island Wildlife Area. What are two specific challenges that might reduce the actual connectivity of this patch to the rest of the protected areas in the Willamette Valley?

Q11 Consider a landscape with a highly connected network of wetland patches.

(a) Imagine a four-lane highway planned to cross the landscape and intersect the network. What species groups are likely to be most negatively affected by the road? Why? Do not restrict your answer only to mammals.

(b) Now imagine that instead of a road, portions of the upland landscape were proposed to be logged. What effects might this disturbance have on the landscape network? What types of species would likely be most affected? Why?

(c) What strategies might be implemented to try to reduce the negative impacts of the road and/or the logging?

Q12 Methods from network analysis can also be used to evaluate connectivity of riverscapes or seascapes (e.g., Grant et al. 2007; Treml et al. 2008; Almany et al. 2009; Grober-Dunsmore et al. 2009).

(a) What types of additional considerations might be required in conducting a network analysis of freshwater mussel populations within a stream network?

(b) What additional factors should be considered in quantifying connectivity for marine environments?

(c) Consider a scenario in which multiple sites are being assessed for potential inclusion into an existing marine protected-area (MPA) network. How might an emphasis on connectivity in deciding among the different sites be at odds with other network objectives?

CONCLUSIONS

The ability to move among habitat patches is vital to ecosystem processes ranging from biological invasions to fire spread and climate adaption. Network analysis provides useful quantitative measures of potential connectivity. However, these measures are scale (and species) specific; thus, more research is needed to develop coherent strategies for the design of multi-species management plans. The development of new methods for quantifying the effect of the matrix on connectivity is another active and important research direction. In many cases, management actions taken within the matrix may be the most effective approach to promoting connectivity. With these advances in the science and theory of connectivity, network analysis will continue to provide a robust set of tools to be applied to conservation challenges across terrestrial, aquatic, and marine environments.

REFERENCES AND RECOMMENDED READINGS[1]

Almany GR, Connolly SR, Heath DD et al (2009) Connectivity, biodiversity conservation and the design of marine reserve networks for coral reefs. Coral Reefs 28:339–351

Angelone S, Holderegger R (2009) Population genetics suggests effectiveness of habitat connectivity measures for the European tree frog in Switzerland. J Appl Ecol 46:879–887

Bodin O, Norberg J (2007) A network approach for analyzing spatially structured populations in fragmented landscape. Landsc Ecol 22:31–44

Bowman J (2003) Is dispersal distance of birds proportional to territory size? Can J Zool 81:195

*Calabrese JM, Fagan WF (2004) A comparison-shopper's guide to connectivity metrics. Front Ecol Environ 2:529–536. *Very readable discussion of the different approaches available to ecologists for quantifying connectivity*

CBI (2012) The Conservation Biology Institute. October 2012. PAD-US (CBI Edition) Version 2. Corvallis, Oregon

Clergeau P, Burel F (1997) The role of spatio-temporal patch connectivity at the landscape level: an example in a bird distribution. Landsc Urban Plan 38:37

*Crooks KR, Sanjayan M (2006) *Connectivity conservation*. Cambridge University Press, Cambridge. *The interested reader would enjoy many of the articles in this book*

De Nooy W, Mrvar A, Batagelj V (2005) *Exploratory social network analysis with Pajek*. Cambridge University Press, Cambridge

Dixo M, Metzger JP, Morgante JS et al (2009) Habitat fragmentation reduces genetic diversity and connectivity among toad populations in the Brazilian Atlantic Coastal Forest. Biol Conserv 142:1560–1569

Estrada-Pena A (2005) Effects of habitat suitability and landscape patterns on tick (Acarina) metapopulation processes. Landsc Ecol 20:529–541

Fahrig L (2003) Effects of habitat fragmentation on biodiversity. Annu Rev Ecol Evol Syst 34:487–515

Fahrig L, Merriam G (1985) Habitat patch connectivity and population survival. Ecology 66:1762–1768

Ferrari JR, Lookingbill TR, Neel M (2007) Two measures of landscape-graph connectivity: assessment across gradients in area and configuration. Landsc Ecol 22:1315–1323

Ferreras P (2001) Landscape structure and asymmetrical inter-patch connectivity in a metapopulation of the endangered Iberian lynx. Biol Conserv 100:125–136

Fischer J, Lindenmayer DB (2007) Landscape modification and habitat fragmentation: a synthesis. Glob Ecol Biogeogr 16:265–280

Flather CH, Bevers M (2002) Patchy reaction-diffusion and population abundance: The relative importance of habitat amount and arrangement. Am Nat 159:40–56

*Galpern P, Manseau M, Fall A (2011) Patch-based graphs of landscape connectivity: a guide to construction, analysis and application for conservation. Biol Conserv 144:44–55. *A catalog of the different ways patch-based networks can be built and analyzed for different conservation purposes*

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

*Grant EHC, Lowe WH, Fagan WF (2007) Living in the branches: population dynamics and ecological processes in dendritic networks. Ecol Lett 10:165–175. *Excellent summary of theory and empirical studies relating to ecological dynamics in streams and other dendritic networks*

*Grober-Dunsmore R, Pittman SJ, Caldow C et al (2009) A landscape ecology approach for the study of ecological connectivity across tropical marine seascapes. In: Nagelkergen I (ed) *Ecological connectivity among tropical coastal ecosystems*. Springer, New York, pp 493–530. *The reader with marine interests would enjoy the collection of articles compiled in this book*

Hanski I, Gilpin M (1991) Metapopulation dynamics—brief-history and conceptual domain. Biol J Linn Soc 42:3–16

Harary F (1969) *Graph theory*. Addison-Wesley, Massachusetts

Jordan F, Baldi A, Orci KM et al (2003) Characterizing the importance of habitat patches and corridors in maintaining the landscape connectivity of a Pholidoptera transsylvanica (Orthoptera) metapopulation. Landsc Ecol 18:83–92

Kadoya T (2009) Assessing functional connectivity using empirical data. Popul Ecol 51:5–15

Kindlmann P, Burel F (2008) Connectivity measures: a review. Landsc Ecol 23:879–890

King AW, With KA (2002) Dispersal success on spatially structured landscapes: when do spatial pattern and dispersal behavior really matter? Ecol Model 147:23–39

Lookingbill TR, Gardner RH, Ferrari JR et al (2010) Combining a dispersal model with network theory to assess habitat connectivity. Ecol Appl 22:427–441

Lyford ME, Jackson ST, Betancourt JL et al (2003) Influence of landscape structure and climate variability on a late Holocene plant migration. Ecol Monogr 73:567–583

*Minor ES, Lookingbill TR (2010) A multiscale network analysis of protected-area connectivity for mammals in the United States. Conserv Biol 24:1549–1558. *Assessment of potential connectivity among protected area networks for three large biomes of the United States. Strategies for prioritizing lands for conservation using network analysis are discussed*

Minor ES, Urban DL (2007) Graph theory as a proxy for spatially explicit population models in conservation planning. Ecol Appl 17:1771–1782

Morzillo AT, Ferrari JR, Liu JG (2011) An integration of habitat evaluation, individual based modeling, and graph theory for a potential black bear population recovery in southeastern Texas, USA. Landsc Ecol 26:69–81

Opdam P, Wascher D (2004) Climate change meets habitat fragmentation: linking landscape and biogeographical scale levels in research and conservation. Biol Conserv 117:285–297

Pascual-Hortal L, Saura S (2006) Comparison and development of new graph-based landscape connectivity indices: towards the priorization of habitat patches and corridors for conservation. Landsc Ecol 21:959–967

Pe'er G, Kramer-Schadt S (2008) Incorporating the perceptual range of animals into connectivity models. Ecol Model 213:73–85

Richardson DM, Whittaker RJ (2010) Conservation biogeography—foundations, concepts and challenges. Divers Distrib 16:313–320

Royer L, Reimann M, Andreopoulos B et al (2008) Unraveling protein networks with power graph analysis. PLoS Comput Biol 4(7), e1000108

Saura S, Pascual-Hortal L (2007) A new habitat availability index to integrate connectivity in landscape conservation planning: comparison with existing indices and application to a case study. Landsc Urban Plan 83:91–103

Scott JM, Davis FW, McGhie RG et al (2001) Nature reserves: do they capture the full range of America's biological diversity? Ecol Appl 11:999–1007

Taylor PD, Fahrig L, Henein K et al (1993) Connectivity is a vital element of landscape structure. Oikos 68:571–573

Tischendorf L, Fahrig L (2000) On the usage and measurement of landscape connectivity. Oikos 90:7–19

Treml EA, Halpin PN, Urban DL et al (2008) Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. Landsc Ecol 23:19–36

Urban DL, Keitt T (2001) Landscape connectivity: a graph-theoretic perspective. Ecology 82:1205–1218

*Urban DL, Minor ES, Treml EA (2009) Graph models of habitat mosaics. Ecol Lett 12:260–273. *A review of recent applications of graph theory to landscape ecology*

van Dorp D, Opdam PFM (1987) Effects of patch size, isolation and regional abundance on forest bird communities. Landsc Ecol 1:59–73

West GB, Brown JH, Enquist BJ (1997) A general model for the origin of allometric scaling laws in biology. Science 276:122–126

Chapter 13
Systematic Conservation Planning with Marxan

Matthew E. Watts, Romola R. Stewart, Tara G. Martin, Carissa J. Klein, Josie Carwardine, and Hugh P. Possingham

OBJECTIVES

Conservation planning is the science of choosing which actions to take where for the purpose of conserving biodiversity. Creating a system of protected areas is the most common form of systematic conservation planning. Hence, we will focus on the process of protected area selection in this chapter. Marxan is the most widely used software in the world for creating marine and terrestrial protected area systems. Because conservation planning is an important job skill for conservation and resource managers, you should understand the principles involved even if you don't use this software in your job and even if you use software other than Marxan for systematic conservation planning. From this chapter, we would like you to:

1. Gain an understanding of the principles of conservation planning: representation, complementarity, adequacy, efficiency, and spatial compactness;
2. See and understand how these principles can be applied to a practical example; and
3. Gain familiarity with Marxan software (via the Zonae Cogito interface).

In Exercise 1, you will explore a simple reserve design problem using a spreadsheet exercise (**Exercise1.xls**) to implement the basic principles of reserve design in a simple hypothetical landscape. In Exercise 2, you will use Marxan to design systems of protected areas in Tasmania. You will run Marxan through Zonae Cogito, a decision support system through which Marxan can be run in an interactive and

M.E. Watts (✉) • R.R. Stewart • C.J. Klein • H.P. Possingham
University of Queensland, Brisbane, QLD, Australia
e-mail: m.watts@uq.edu.au

T.G. Martin • J. Carwardine
CSIRO, Brisbane, QLD, Australia

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_13

user-friendly way. Software installation for Exercise 2 requires following detailed instructions provided here: http://www.uq.edu.au/marxan/docs/Installing%20Zonae%20Cogito%20on%20your%20computer.pdf and will likely require administrator privileges on your machine to install and operate properly. All the data files needed to complete the exercises can be found on the book website, along with some options for additional advanced exercises.

INTRODUCTION

Part 1. Systematic Conservation Planning

World-class conservation planning processes for land and sea use an approach known as systematic conservation planning (Moilanen et al. 2009; Possingham et al. 2006). **Systematic conservation planning** focuses on locating, designing, and managing conservation areas that collectively represent the biodiversity of a region for the least cost. In many cases, new protected areas are selected to add to an existing set of protected areas. The systematic conservation planning approach is transparent, and the system of protected areas function together to meet clearly defined conservation goals.

Systematic conservation planning is a departure from ad hoc and opportunistic approaches used in the past. An **ad hoc** approach is one in which site selection is driven by conservation urgency, affinity, scenery, and ease of designation, often avoiding areas that are politically or economically costly. Most national parks or other places considered to be areas for "conservation" were not chosen to meet specific biodiversity objectives (Possingham et al. 2000). Many existing protected areas were selected because of their amenity value, for example, as a vacation spot. Most are located in places unsuitable for other purposes such as agriculture or urban development (Pressey et al. 1993). Other areas have been selected to protect a few charismatic flagship or umbrella species (Simberloff 1998) without any guarantee that they will adequately conserve the biodiversity of a region. This ad hoc approach has resulted in a legacy of fragmented collections of sites in which some habitats or ecosystems, like the "rock and ice" of high mountain areas, are overrepresented, while low-lying fertile plains are underrepresented (Pressey et al. 1993; Soulè and Terborgh 1999).

Systematic conservation planning is a more rigorous and accountable approach for selecting priority areas for protection compared to the opportunistic approach (Groves et al. 2002; Margules and Pressey 2000). Over the past 25 years, a systematic approach to conserving biodiversity has evolved (Moilanen et al. 2009; Pressey and Bottrill 2009) and now includes 11 well-defined stages (Table 13.1). Marxan was designed primarily to help inform stage 9, the selection of new conservation areas to complement existing ones in order to achieve the conservation objectives. Specifically, Marxan identifies potential priority areas for inclusion in a protected area network and provides other information to assist decision-makers in choosing the final selection of areas.

Table 13.1 Phases of a framework for systematic conservation planning (Margules and Pressey 2000; Pressey and Bottrill 2009). Phase 9 is the main focus of this chapter.

1	Scoping and costing the planning process
2	Identifying and involving stakeholders
3	Describing the context for conservation areas
4	Identifying conservation goals
5	Collecting data on socioeconomic variables and threats
6	Collecting data on biodiversity and other natural features
7	Setting conservation objectives
8	Reviewing current achievements of objectives
9	**Selecting additional conservation areas**
10	Applying conservation actions to selected areas
11	Maintaining and monitoring conservation areas

Fundamental Principles for Designing Conservation Areas

Here, we discuss five fundamental principles used when designing conservation areas: representation, complementarity, adequacy, efficiency, and spatial compactness (Margules and Pressey 2000; Possingham et al. 2006). Marxan can accommodate all of these principles.

Representation. Protected area systems should contain the full range of biodiversity, taking into consideration biodiversity composition, structure and function, and evolutionary processes. Incorporating as many kinds of **biodiversity (or conservation) features** as possible (such as species, ecosystems, vegetation types) will result in a more comprehensive protected area system. Protected area systems that represent all facets of biodiversity have high representativeness. For example, if you wish to protect populations of a particular species or samples of a habitat, it is best if the areas chosen cover the range of variation in that species and/or habitat. Wherever possible, the selection of areas should take into consideration any species/habitats that are rare, endangered, or considered unique.

Complementarity. Protected areas for conservation should be selected as a complementary set, where each one complements features of others. Sites with the highest species richness are not necessarily the most important for inclusion in a protected area system, because the most species rich sites may contain similar assemblages. Sites complement each other well if they contain different features of biodiversity. Consequently, their selection provides a combination of sites that achieve the goal of comprehensiveness in the most efficient way. The principle of complementarity means that planning is best informed by knowing what is already contained within existing conservation areas—an exercise referred to as **gap analysis**. The selection proceeds by iteratively reviewing how well the targets (e.g., 20% of total habitat for each species) are achieved when individual sites are added to (or removed from) the protected area system.

Adequacy. The goal of protected area system design is not to merely capture biodiversity, but to promote its persistence and long-term viability. Larger and more connected systems of conservation areas are considered to be superior to smaller and more isolated conservation areas. Larger connected systems can provide for the maintenance of ecosystems through connectivity and offset the effects of local catastrophes. A system-based approach to protected area system design—where the whole is more than the sum of its parts—recognizes the relationship between individual conservation areas, and therefore the role of each area as part of a system.

Ideally, a protected area system is designed to conserve enough of each feature of biodiversity to enable persistence. However, the minimum habitat area or population size required for the persistence of a species or ecosystem is rarely known, and often limited budgets mean that we cannot simply conserve more to be on the safe side. One general strategy proposed to address the issue of persistence in the absence of this knowledge is **redundancy**, making sure that you don't have all of one feature in one place. Replication improves the likelihood of regional persistence, spreading the risk of failure by providing greater opportunity for recolonization of empty protected areas from other viable and connected areas.

Efficiency. Efficiency describes the ability of a protected area system design process to deliver biodiversity objectives for least cost or fewest resources. Because resources available to achieve conservation goals are finite, inefficient systems are less likely to achieve their goals. By planning protected area systems efficiently, we minimize the risk of exhausting available resources before biodiversity objectives are met (Ban and Klein 2009; Carwardine et al. 2008; Klein et al. 2008; Stewart and Possingham 2005). We describe the limiting resources or limiting factors as "costs." The typical **costs** of a conservation area include:

- Area available to reserve
- Costs of ongoing management
- Costs to industry, tourism, and recreation from displaced activities
- Acquisition or land purchase costs

Marxan provides efficient solutions by incorporating these costs into the design process. A protected area system design process that ignores costs is not as practically useful as one that considers cost. Lastly, decisions about individual protected areas affect the performance of the protected area system as a whole. Efficiency is therefore also concerned with the way sites are prioritized for conservation. The most efficient solutions are obtained by selecting sites as a complementary set, rather than selecting sites one by one.

Spatial compactness. A compact protected area system, with a low edge:area ratio has three advantages over a fragmented system. First, biodiversity within a compact system is more connected, giving a greater chance of persistence compared with a fragmented system. Second, many of the most sensitive species are absent or have low population growth rates within edges. Finally, edges between a park and other areas cost money: a longer edge means more neighbors and more management costs.

Before we explore a real-world example and learn about the kind of software used by professionals, we will explore a small spreadsheet example that explores these themes.

EXERCISES

EXERCISE 1: Small-Scale Protected Area System Design

In this exercise, you will use the spreadsheet and provided handout to design protected area systems that reach conservation objectives in a cost-effective manner. Here, our objective will be to represent 20% of the total habitat area for each of three species in the study region. An additional objective will be to design protected areas with different degrees of spatial compactness. In this exercise, we consider a hypothetical landscape made up of a grid of 100 sites—referred to as planning units—arranged in a 10×10 grid:

1. Download and unzip the folder called **Exercise1**.
2. Within this folder, open the file **Exercise1.xls**. You will use the first sheet within the spreadsheet whose tab is labeled **3 features**.
3. Notice that this spreadsheet contains information on each planning unit. There are 100 total planning units, each in a separate row with a unique Planning Unit Identification Number [**PUID**].
4. Notice the additional columns in the spreadsheet that include the cost of each planning unit, as well as the area of each species contained in a given planning unit.
5. Notice the second column highlighted yellow, labeled [**SELECTIONS**]. In the spreadsheet, you can easily select a planning unit for inclusion in the protected area network by changing the value in the [**SELECTIONS**] field from 0 (unselected) to 1 (selected).
6. Also of use is the file **Exercise1_handout.pdf** within the same folder. You can use this handout to visualize the spatial configuration of your protected area system. It contains information about the cost of each planning unit and the area of each species in each planning unit.
7. Notice when you select a planning unit, summary information [**green cells**] is automatically updated for your protected area system, including the cost of the protected area system selected [**SUM COST**] as well as the amount needed to meet the targets for the protected area system [**TARGET GAP**].
8. You can also track the individual species targets [**red cells**] as you select various planning units and then determine if your target is met. Remember, our target is 20% of the total habitat area for each of these three species.
9. To answer the questions below, you will use this spreadsheet to find a protected area system that meets all of your conservation targets in a cost-effective way. When you have found a protected area system that meets your conservation goals, record the value of [**SUM COST**], the cost of your protected area system.

10. If you wish, you could also devise a simple heuristic to prioritize sites. For example, at each site, you might compute the sum of feature areas and divide by the site cost as a measure of the cost-effectiveness of a single site.

Q1 Without considering spatial compactness, what is the lowest cost of a protected area system you can design that meets the desired habitat protection objectives? Record the cost. Save the "map" of your protected area system (either using Excel or by coloring your handout).

Q2 What is the additional cost of a protected area system that meets the habitat protection objectives but with a low, medium, or high degree of spatial compactness? As you answer this question, consider the following:

- Remember that how you determine the level of compactness can be a subjective choice.
- Create a graph where the cost of the protected area system is the *X*-axis, and the boundary length (edge or compactness) is the *Y*-axis. Good protected area systems will be in the left-hand bottom corner of your plot.
- If you are working as a group, each person can create a single system but then include all of your systems together on one plot.

Part 2. Using Marxan for Conservation Planning

What Is Marxan?

Marxan is software that delivers decision support for systematic conservation protected area design (Ball et al. 2009). It was initially designed to solve a conservation problem known as the minimum-set problem, where the goal is to achieve a certain amount of every biodiversity feature for the smallest possible cost (McDonnell et al. 2002). Or put another way, the objective is to minimize costs subject to the constraint of meeting biodiversity targets (Possingham et al. 2000; Ball and Possingham 2000). An example **biodiversity target** might be to ensure at least 30% of every habitat is represented in a protected area network. A planner is likely to want to minimize the total monetary cost required to purchase and manage a conservation area that meets this constraint.

The number of possible solutions to this problem is vast and beyond the ability of the human mind or a computer to consider. For example, the number of possible solutions to Exercise 1 is 2^{100} or 1.3×10^{30}! For this reason, algorithms have been developed to support decisions around the design of conservation areas. Furthermore, not only would it would take an extremely long time to find the single optimal solution to any given real-world protected area design, but a single solution is unlikely to be the most useful. Thus, currently heuristics are preferred over exact algorithms because they provide timely solutions to complex problems and offer a range of near-optimal solutions for planners and stakeholders (Possingham et al. 2000; McDonnell et al. 2002).

Marxan can be used for a variety of purposes at different stages in the systematic conservation planning process (Table 13.1). The tool was designed primarily to help inform Stage 9: "Selecting additional conservation areas" to complement existing ones in order to achieve the conservation goals. The software identifies sets of areas that meet conservation targets for minimal "cost," and it can be used to explore trade-offs between conservation and socioeconomic objectives. In addition, it can highlight sites that occur in a large number of solutions, which can help identify priority areas for conservation action. It can also be used to measure the achievement of targets within existing conservation areas (Stage 8) (Stewart et al. 2003) and to help prioritize conservation actions and develop management plans for selected sites (Stage 11).

Problem Formulation Using Marxan

Any conservation planning problem can be formulated as an optimization problem with the following essential elements (Moilanen et al. 2009; Possingham et al. 2001; Wilson et al. 2009):

1. A clearly defined objective stating the desired outcome (e.g., maximize the number of species conserved or represent 30% of each habitat type);
2. A list of features to be targeted for conservation (e.g., species, habitats, soil types);
3. A list of actions (e.g., protect an area) and how these actions contribute to achieving the objective (e.g., how many species are conserved if the action is applied); and
4. Financial information specifying the cost of implementing each action in a site, as well as the budget available.

Clearly defining each element helps to identify conservation priorities using the software.

Marxan uses two well-accepted approaches to identify spatial conservation priorities, minimum-set and maximal coverage, and each solves a different objective. The objective of the **minimum-set strategy** is to achieve the conservation objectives while minimizing the resources expended. Less commonly, Marxan is used to solve the **maximal coverage strategy**, which is to maximize the biodiversity benefit given a fixed budget (Possingham et al. 2006). Regardless of approach, it is essential to clearly define an objective that states the desired outcome before using the software to identify priorities (Moilanen et al. 2009; Possingham et al. 2006).

The **objective function** is the mathematical formulation of the minimum-set problem. In protected area design, the problem we are trying to solve is to identify the protected area system that achieves our targets and spatial requirements for the least cost. Thus, a protected area configuration is given an objective function score to measure how well it performs. In comparing alternative solutions, those with lower scores are better. Thus, the objective function is a score that we want to minimize and is calculated as follows:

Score = Cost + Boundary Length + Penalty

where costs, boundary length, and penalties are determined as below.

Cost of the protected area system. Each planning unit (parcel of land or sea) is assigned a cost that the user defines prior to planning. The cost is summed for all planning units included in a protected area system to calculate their combined cost.

Boundary length of the protected area system. One of the practical considerations for protected area design is the spatial configuration of the protected area system (i.e., a single large system or several small systems). The protected area system boundary length is measured as the sum of the planning units that share a boundary with planning units outside the protected area system. Hence, fragmented protected area systems will have a large boundary length. The objective function addresses the issue of connectivity by using the **boundary length modifier (BLM)** which places a value on the importance of having a more compact protected area system. The BLM is important because a system that is fragmented will likely be difficult (and costly) to manage. In addition, there are increased edge effects and reduced connectivity in a fragmented solution, potentially leading to reduced biodiversity persistence. Thus, some level of "clumping" or spatial compactness is desirable for management. The BLM is a user-defined parameter and allows you to control the amount of clumping that occurs in the solutions. With a large value for BLM, the system will be more clumped.

Penalty incurred for every feature that fails to meet its target. For each alternative solution, Marxan calculates whether the target for each conservation feature is met or not. If a target is unmet, then a user-defined penalty cost called the **species penalty factor (SPF)** is applied. Making the SPF user-defined allows different weightings be given to different feature targets. For example, it may be more important to achieve targets for feature A than for feature B. Alternatively, the same SPF can be applied to all conservation features (in which case, the SPF for feature A = SPF for feature B). The higher the SPF, the higher the penalty when a conservation feature target is unmet. An appropriately high SPF will result in more costly protected areas with more targets met.

More formally, the objective function is:

$$\text{Score} = \sum_{\text{PUs}} \text{Cost} + \text{BLM} \times \text{Boundary Length} + \sum_{\text{Features}} \text{SPF for missing features}$$

where PUs are the planning units, BLM is the boundary length modifier, and SPF is the species penalty factor.

Finding Optimal Solutions Using Simulated Annealing

Marxan finds near-optimal solutions to a minimum-set problem by minimizing the objective function—a lower score means a better solution to the problem. The number of possible solutions to this problem is vast, so it is usually impossible to find the optimal solution. Instead, a metaheuristic algorithm, simulated annealing, is employed to find many near-optimal solutions (Kirkpatrick et al. 1983).

The **simulated annealing** algorithm uses a technique borrowed from statistical mechanics to find good solutions from among this vast number of possible solutions. A large number of random changes to the protected area system are attempted, typically one million or more. At the start of the process of annealing, any change in score is accepted. As the process proceeds, the acceptance probability of bad changes is progressively reduced, until finally only good changes are accepted. A bad change is one that increases the objective function score, while a good change is one that reduces the score (Moilanen and Ball 2009). This process allows the algorithm to find solutions that are close to an exact solution.

In reality, protected area design problems have many *near*-optimal solutions, none of which are significantly better or worse than the *optimal* solution. As such, it is more useful for decision-making to identify a range of near-optimal solutions that provide diverse options for a decision-maker, rather than a single optimal solution (Kirkpatrick et al. 1983). Happily, this is the way Marxan works, generating a range of options, making it useful in the real world. Some heuristic algorithms do not explore the solution space well because they get "stuck" at a local minimum nowhere near optimal. The simulated annealing algorithm avoids this problem by taking random backward steps (or bad moves), making it a useful algorithm for the purposes of conservation planning. Simulated annealing is fast, simple, and robust to changes in the size and type of problem. These advantages allow it to explore a variety of scenarios with differing constraints and parameters while producing many good solutions. Users can also access a variety of simpler, but often faster, heuristic algorithms within Marxan. More information on simple heuristic algorithms and simulated annealing can be found in the Marxan User Manual Appendix B (Game and Grantham 2008) and in the Marxan Good Practices Handbook (Ardron et al. 2010).

Lastly, while Marxan can help find efficient solutions to spatial prioritization problems, it cannot *make* decisions. The software is designed to be a decision *support* tool. As such, Marxan solutions should be used within a larger decision-making process involving stakeholders, managers, local people, etc.

MARXAN INPUTS

The information Marxan needs to run must be formally organized in input files that conform to its information management system. At a minimum, the following files are needed to run Marxan:

- Planning unit file
- Conservation feature file (species and habitat list)
- Planning unit versus conservation feature file
- Boundary length file
- Input parameter file

Examine Table 13.2 for more details on the output files from Marxan.

In Exercise 2, you will use species and habitats as **conservation features** for Marxan and use land acquisition cost for the **planning unit costs**. It is possible to use more abstract concepts for Marxan features and costs, and we illustrate some of these in the online appendix for this chapter.

Table 13.2 Description of Marxan Input Files

File name	Description	How Marxan uses the file
Planning unit file ***(pu.dat)***	This file lists all the planning units in the planning region. It usually includes additional data on each planning unit's individual cost and reserve status. This list of planning units corresponds to the spatial layer of planning units defined in your GIS. The planning unit layer may be preexisting cadastral boundaries or watersheds, or you may determine that a grid, hexagon, or other shape of planning units is more appropriate for your planning exercise. There are tools that can help you create the planning unit layer.	• Identify each planning unit • Identify if planning unit is already conserved • Calculate how much the protected area system costs when planning units are included
Conservation feature file ***(spec.dat)***	This file contains information about each of the conservation features being considered, such as their name, conservation targets, and representation requirements. The penalty that is applied, if the representation requirements for each feature are not met (or SPF), is also in this file.	• Identify each conservation feature • Determine how much of each conservation feature must be included in a given solution to meet targets • Calculate the penalty for conservation features not meeting targets
Planning unit versus conservation feature file ***(puvssp.dat)***	This file contains information on the distribution of conservation features across the planning units.	• Find planning units that contain conservation features • Calculate the amount of a given conservation feature in a planning unit • Allow Marxan to calculate the contribution a planning unit makes toward reaching the conservation feature targets
Boundary length file ***(bound.dat)***	This file contains information about the spatial relationship between planning units (e.g., the length of shared boundaries between planning units) and some other measures of the desirability or cost including adjacent planning units in a solution. This file is necessary if you wish to use the boundary length modifier to adjust the compactness of the solutions.	• Calculate the boundary length of each solution by adding up all of the boundary values on the edges of the solution
Input parameter file ***(input.dat)***	This file defines many of the parameters that control the way that Marxan works, such as the number of solutions to generate and the BLM. It is also used to tell Marxan where to find the input files containing your data and where to place the output files.	• Set input parameters • Locate the input and output files

Detailed information about inputs can be found in the Marxan User Manual (Game and Grantham 2008) and Marxan Good Practices Handbook (Ardron et al. 2010). Tools to create the files are available on the Marxan website (www.uq.edu.au/marxan) along with detailed tutorials. It is also possible to create input files using a GIS and spreadsheet application.

MARXAN OUTPUTS

The most commonly used output includes:

- Solution for each run
- Summed solution
- Missing value information
- Summary information

Review Table 13.3 for more details on Marxan output files.

Table 13.3 Description of Marxan output files

File name	Description	How Marxan uses the file
Solution for each run ***(scenario_r001.dat)***	This is a text file that lists the planning units and identifies if they were selected for inclusion in the protected area system. A planning unit may be selected because it contains conservation features that are irreplaceable, and also because its cost or location efficiently improves spatial compactness.	• Display a protected area system in a GIS • Compare protected area systems spatially • Maps can be used as part of a stakeholder involvement plan
Summed solution ***(scenario_ssoln.dat)***	This file shows the number of times each planning unit is selected across all the protected area systems. Planning units which are never selected have a selection frequency of 0, while those always selected have a selection frequency equal to the total number of runs. This file gives an indication of the relative importance of a planning unit for efficient protected area system design. It is often used to indicate the relative priority of planning units.	• GIS display of how frequently planning units are selected • It should **not** be used on its own to create protected area systems, but it can be informative to identify key areas • These maps are also used as part of a stakeholder involvement process
Missing value information ***(scenario_mv001.dat)***	This file provides detailed information about how well each solution meets the conservation feature targets, providing information such as the target amount of the feature required in the protected area system, how much of the feature was conserved, and whether the target was met.	• Find out which feature targets are not met in each solution and by how much • Helps set the species penalty factor (SPF) parameter
Summary information ***(scenario_sum.dat)***	This file shows information about each run including the objective function score, cost, number of planning units selected, boundary length, species penalty, shortfall (cumulative target gap for all features), and number of features not meeting their targets.	• Compare the performance of solutions in terms of targets met, score, cost, etc.

Detailed information about each of the output file types is available in Section 5.3 of the Marxan User Manual (Game and Grantham 2008), and the Marxan Good Practices Handbook (Ardron et al. 2010) contains information about how these output files are used.

Instructions for Zonae Cogito: Marxan Graphical User Interface

Zonae Cogito (ZC) is a decision support system through which Marxan can be run in an interactive and user-friendly way. It allows users to edit and calibrate the key input files including the planning unit file, species file (SPF), boundary length modifier (BLM), as well as change parameters such as the number of runs (NUMREPS) and number of iterations (NUMITNS). It uses an open source GIS to display Marxan solutions interactively, allowing seamless interaction with Marxan inputs and outputs. ZC has two windows: a Marxan window where parameters and input files can be edited and a GIS window where spatial outputs can be viewed.

In the **GIS window** of ZC, you can spatially view Marxan outputs. The list of items in the **Output to Map** control shows all the spatial outputs you can view:

- *Selection frequency reserved zone* corresponds to the *summed solution* output file.
- *Best solution*, *solution 1*, etc. correspond to the *solution for each run* output files for each reserve system and the best reserve system (the one with the lowest objective function score).

In the **Marxan drop-down menu** of ZC, you can use the **View Output** control to view the nonspatial output tables:

- *Summary* corresponds to the *summary information* output file.
- *Missing values bar graph* corresponds to the *missing value information* output file for each protected area system.
- *Best solution* corresponds to the *missing value information* output file for just the best protected area system (the one with the lowest objective function score).

ZC allows easy calibration of Marxan parameters. **Calibration** is the process of choosing parameters, so the software properly represents the real-world situation being analyzed. Calibration helps ensure that the protected area systems produced are close to optimal while still achieving the conservation feature targets and desired degree of clumping. If you do not calibrate the key Marxan parameters, you risk ending up with:

- Inefficient sets of solutions
- Inappropriate degree of clumping
- Inefficient running time for your analysis
- Unmet feature targets

Further reading on calibration is available in Fischer and Church (2005) and the Marxan Good Practices Handbook, Chapter 8 (Ardron et al. 2010).

A **sensitivity analysis** allows you to determine which input data and parameters most influence the solution. This can be important if, for example, there is a data layer with a great deal of uncertainty driving the results. In this case, you may want to remove the data layer from the analysis or use another data layer to represent the conservation feature. More information about sensitivity analysis can be found in Section 8.4 of the Marxan Good Practices Handbook (Ardron et al. 2010).

Additional Information

Additional documentation with detailed information is available on the Marxan website: (http://www.uq.edu.au/marxan/documentation). Also see the online Appendix for this chapter. Segan et al. 2011 provides more background on Zonae Cogito. Also see the user manual "Using the Zonae Cogito Decision Support System" for more technical information (Watts et al. 2010).

EXERCISE 2: Real-World Protected Area Design

Using the Zonae Cogito and Marxan software packages, you will generate and explore alternative protected area systems for Tasmania, an island state south of Australia. The provided Marxan data (**Exercise2.zip**) include existing protected areas, cost data (land acquisition costs), and biodiversity features (vegetation types and a single bird species) (Figure 13.1). For our purposes, the objective will be to represent 20% of the total area for each vegetation type and species in the region.

NOTE: Be sure your instructor has fully installed the required ZC software before proceeding, following the detailed instructions provided at http://www.uq.edu.au/marxan/docs/Installing%20Zonae%20Cogito%20on%20your%20computer.pdf.

You must have full write permissions (administrator privileges) in order to run the software:

1. Unzip the file **Exercise2.zip** to your computer into a folder where you have full write permissions.
2. Launch the Zonae Cogito software.
3. From the folder where you have unzipped your files, load the project **Exercise2.zcp** with Zonae Cogito.

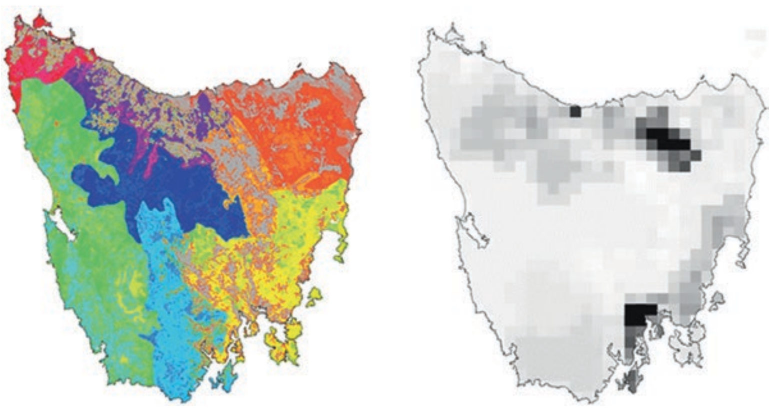


Figure 13.1 Example maps of Marxan input for Tasmania. *Panel A* represents the 63 vegetation types used. *Panel B* shows the cost surface used where darker areas are more expensive

4. In the **Marxan** window within the Zonae Cogito graphical user interface, navigate to the **Marxan Parameter To Edit** list, and locate key parameters:
   - The **NUMREPS** and **BLM** parameters are accessible directly from this drop-down menu.
   - The **SPF** value for each feature can be found with the SPEC parameters.
5. Leave the **NUMITNS** parameter set to one million throughout the exercise. It is only necessary to increase this parameter for working with broader scale datasets than the one being used for this exercise.
6. Set the **NUMREPS** parameter to 10 for sensitivity analysis, and set it to 100 for generating your final results. This means you will generate only 10 reserve systems for the parameter setting phase, and you will generate 100 reserve systems once your parameters for final results.
7. Press the **Run** button on the **Marxan window** to compute a set of reserve systems based on your input files and parameters.

Q3 The targets are set at 20% of the current habitat. Try increasing the targets to 40% and then decreasing them to 10%. What effect does this have on the size and the cost of the protected area system?

Q4 Revisit the earlier definition and utility of the SPF value (species penalty factor). What is an appropriate SPF value to use for each biodiversity feature that ensures a reserve system will capture the targets for each? For this question, generate reserve systems ignoring spatial compactness (i.e., use a BLM of zero). What is the cost of one of your representative efficient reserve systems?

Q5 Consider designing different reserve systems that meet your objectives but have low, medium, and high degrees of spatial compactness. What are appropriate boundary length modifiers (BLM) values to use? Adjust the BLM, and monitor how the spatial compactness changes. As you did in **Q2**, plot boundary length as a function of reserve system cost for low, medium, and high degrees of spatial compactness.

SYNTHESIS

EXERCISE 3: Stakeholder Report Based on Marxan Output

Using your results from Exercise 2, prepare a report to stakeholders in a hypothetical decision-making process that illustrates several distinct options for reserve system design in Tasmania. The target audience should include:

- Government agencies concerned with conservation and resource use
- Commercial organizations concerned with resource use
- Commercial ecotourism operators concerned with exploiting the natural features of the study region for tourism

- Nongovernment organizations concerned with protecting biodiversity in the study region

Include and discuss the following information in your report:
(a) Map showing one of your final solutions (or a map showing selection frequency of your final solutions).
(b) Trade-offs between planning unit cost and biodiversity protection. Find a range of SPF values or target values that illustrate this trade-off and include a trade-off curve.
(c) Explain the rationale behind the degree of spatial compactness used to generate your results. Create a trade-off curve with various BLM values to help illustrate your point.
(d) Read another scientific paper (or report) that uses these types of outputs, and then incorporate this study into your own report as context.

Your instructor will determine word/page limits depending on the amount of time you have to complete your assignment. Consider giving oral presentations of your results. See the online Appendix for this chapter for even more additional readings.

Acknowledgments We thank Patricia Sutcliffe for reviewing the chapter for us. We also thank Ian Ball, Eddie Game, Dan Segan, and many others for their work that we have drawn on in the production of this chapter. None of this work would have been possible without financial support from our funding organizations: the Department of Sustainability, Environment, Water, Population and Communities; the Australian Research Council funded Centre of Excellence for Environmental Decisions; the National Environmental Research Program funded Environmental Decisions Hub; and the Commonwealth Scientific and Industrial Research Organisation.

REFERENCES AND RECOMMENDED READINGS[1]

Ardron J, Possingham HP, Klein CJ (2010) *Marxan good practices handbook*, Version 2. University of Queensland, St. Lucia

Ball IR, Possingham HP (2000) MARXAN (V1.8.2): marine reserve design using spatially explicit annealing, a manual. marxan.net

*Ball IR, Possingham HP, Watts M (2009) Marxan and relatives: software for spatial conservation prioritisation. In: Moilanen A, Wilson KA, Possingham HP (eds) *Spatial conservation prioritisation: quantitative methods and computational tools*. Oxford University Press, Oxford, pp 185–195; Chapter 14. *This core Marxan reference presents a general overview, including mathematical equations for Marxan, as well as more recent developments, including Marxan with Zones, Marxan with Probability, and Marxan with Connectivity.*

Ban N, Klein CJ (2009) Spatial socio-economic data as a cost in systematic marine conservation planning. Conserv Lett 2(5):206–215

*Beger M, Linke S, Watts M, Game E, Treml E, Ball IR, Possingham HP (2010) Incorporating asymmetric connectivity into spatial decision making for conservation. Conserv Lett 3(5):359–368.

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Describes various types of connectivity important to conservation and how they could be applied in Marxan.

*Carwardine J, Wilson KA, Watts ME, Etter A, Klein CJ, Possingham HP (2008) Avoiding costly conservation mistakes: the importance of defining actions and costs in spatial priority setting. PLoS One 3(7): e2586. *Demonsrates how the explicit consideration of conservation actions and socioeconomic costs is critical to designing effective conservation plans.*

Fischer DT, Church RL (2005) The SITES reserve selection system: a critical review. Environ Model Assess 10:215–228

Game ET, Grantham HS (2008) *Marxan user manual: for Marxan version 1.8.10*. University of Queensland, St. Lucia

*Game ET, Watts ME, Wooldridge S, Possingham HP (2008) Planning for persistence in marine reserves: a question of catastrophic importance. Ecol Appl 18(3):670–680. *The first paper to apply the version of Marxan with Probability. It shows how marine reserves can be designed that achieve ecological targets whilst avoiding areas with a high chance of being destroyed by a catastrophic event, coral bleaching.*

*Game ET, Lipsett-Moore G, Hamilton R, Peterson N, Kereseka J, Atu W, Watts ME, Possingham HP (2010) Informed opportunism for conservation planning in the Solomon Islands. Conserv Lett 4(1):38–46. *Describes how the Lauru Land Conference of Tribal Communities and The Nature Conservancy have worked with the communities of Choiseul Province, Solomon Islands, to develop a conservation planning process that reconciles community-driven conservation opportunities, with a systematic and representation-based approach to prioritization.*

Groves CR, Jensen DB, Valutis LL, Redford KH, Shaffer ML, Scott JM, Baumgartner JV, Higgins JV, Beck MW, Anderson MG (2002) Planning for biodiversity conservation: putting conservation science into practice. Bioscience 52(6):499–512

Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. Science 220(4598):671–680

Klein CJ, Watts ME, Kircher L, Segan D, Game ET (2007–2011) Introduction to Marxan, course manuals. marxan.net/courses

Klein CJ, Chan A, Kircher L, Cundiff A, Gardner N, Hrovat Y, Scholz A, Kendall BE, Airamé S (2008) Striking a balance between biodiversity conservation and socioeconomic viability in the design of marine protected areas. Conserv Biol 22(3):691–700

*Klein CJ, Wilson, KA, Watts, ME, Stein, J, Berry, S, Carwardine, J, Stafford-Smith, DM, Mackey B, Possingham HP (2009) Incorporating ecological and evolutionary processes into continental-scale conservation planning. Ecol Appl 19:206–217. *Shows four ways of capturing ecological and evolutionary processes in Marxan, inculding identifying connected subcatchments along waterways as priorities for conservation.*

*Klein CJ, Steinback, C, Watts, ME, Scholz AJ, Possingham HP (2010) Spatial marine zoning for fisheries and conservation. Front Ecol Environ 8:349–353. *First paper to apply Marxan with Zones. Shows how to zone a region for four different uses relevant to fishing and conservation.*

Linke S, Watts ME, Stewart R, Possingham HP (2011) Using multivariate analysis to deliver conservation planning products that align with practitioner needs. Ecography 34:203–207

*Makino A, Klein CJ, Beger M, Jupiter SD, Possingham HP (2013) Incorporating conservation zone effectiveness for protecting biodiversity in marine planning. PLoS One 8(11), e78986. *Describes how to use zone effectiveness when zoning for multiple uses using Marxan with Zones.*

Margules CR, Pressey RL (2000) Systematic conservation planning. Nature 405:243–253

McDonnell MD, Possingham HP, Ball IR, Cousins EA (2002) Mathematical methods for spatially cohesive reserve design. Environ Model Assess 7:107–114

Moilanen A, Ball IR (2009) Heuristic and approximate optimization methods for spatial conservation prioritization. In: Moilanen A, Wilson KA, Possingham HP (eds) *Spatial conservation prioritisation: quantitative methods and computational tools*. Oxford University Press, Oxford

Moilanen A, Possingham H, Wilson KA (2009) Spatial conservation prioritization: past, present, future. In: Moilanen A, Wilson KA, Possingham HP (eds) *Spatial conservation prioritisation: quantitative methods and computational tools*. Oxford University Press, Oxford

Possingham HP, Ball IR, Andelman S (2000) Mathematical methods for identifying representative reserve networks. In: Ferson SB (ed) *Quantitative methods for conservation biology*. Springer, New York

Possingham HP, Andelman SJ, Noon BR, Trombulak S, Pulliam HR (2001) Making smart conservation decisions. In: Orians G, Soule M (eds) *Research priorities for conservation biology*. Island Press, Washington, DC, pp 225–244

Possingham HP, Wilson KA, Andelman SJ, Vynne CH (2006) Protected areas: goals, limitations, and design. In: Groom MJ, Meffe GK, Carroll CR (eds) *Principles of conservation biology*. Sinauer, Sunderland

Pressey RL (2002) Classics in physical geography revisited. Prog Phys Geogr 26(3):434–441

Pressey RL, Bottrill MC (2009) Approaches to landscape- and seascape-scale conservation planning: convergence, contrasts and challenges. Oryx 43:464–475

Pressey RL, Humphries CJ, Margules CR, Vane-Wright RI, Williams PH (1993) Beyond opportunism: key principles or systematic reserve selection. Trends Ecol Evol 8:124–128

Segan DB, Game ET, Watts ME, Stewart RR, Possingham HP (2011) An interoperable decision support tool for conservation planning. Environ Model Software 26(12):1434–1441

Simberloff D (1998) Flagships, umbrellas, andkeystones: is single-species management passé in the landscape era? Biol Conserv 83(3):247–257

Soulè ME, Terborgh J (eds) (1999) *Continental conservation: scientific foundations of regional reserve networks*. Island Press, Washington, DC

Stewart RR, Possingham HP (2005) Efficiency, costs and trade-offs in marine reserve system design. Environ Model Assess 10:203–213

Stewart RR, Noyce T, Possingham HP (2003) Opportunity cost of ad hoc marine reserve design decisions: an example from South Australia. Mar Ecol Prog Ser 253:25–38

Watts ME, Stewart RR, Segan D, Kircher L, Possingham HP (2010) *Using the zonae cogito decision support system, a manual*. University of Queensland, Brisbane

Wilson KA, Carwardine J, Possingham HP (2009) Setting conservation priorities. Ann New York Acad Sci 1162:237–264

Chapter 14
Connectivity as the Amount of Reachable Habitat: Conservation Priorities and the Roles of Habitat Patches in Landscape Networks

Santiago Saura and Begoña de la Fuente

OBJECTIVES

Landscape connectivity plays an important role in sustaining ecological processes at different spatial and temporal scales (e.g., Crooks and Sanjayan, 2006). Landscape connectivity can help to counteract some of the adverse effects of habitat fragmentation and to facilitate species range shifts in response to climate change. Therefore, maintaining or enhancing landscape connectivity is a key part of current biodiversity conservation efforts. A variety of metrics for analyzing connectivity have been developed, ranging from some derived from or used within metapopulation models (Hanski and Ovaskainen 2000; Moilanen and Nieminen 2002) to others based on network analysis (graph theory) (Ricotta et al. 2000; Urban and Keitt 2001; Estrada and Bodin 2008; Saura and Rubio 2010; Galpern et al. 2011; Rayfield et al. 2011). In particular, graph-based approaches have gained increasing popularity in ecological research and applied conservation planning in recent years (Calabrese and Fagan 2004; Saura and Pascual-Hortal 2007; Urban et al. 2009; Pereira et al. 2011; Awade et al. 2012; Rodriguez-Perez et al. 2014). Graphs are just a data structure, and, similarly to vector or raster data structures in geographical information systems, different outcomes of variable quality can be obtained through their use. A crucial issue

S. Saura (✉)
ETSI Montes, Universidad Politécnica de Madrid, Madrid, Spain
e-mail: santiago.saura@upm.es

B. de la Fuente
ETSI Montes, Universidad Politécnica de Madrid, Madrid, Spain

Servicio Territorial de Medio Ambiente de Segovia, Junta de Castilla y León, Segovia, Spain

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_14

is how to measure and analyze connectivity in order to capture important aspects and provide meaningful guidance for conservation decisions. Through this lab, students will:

1. Analyze potential weaknesses and limitations of several widespread connectivity metrics for establishing conservation priorities;
2. Understand why connectivity is not only related to connections between habitat patches but should also consider the contribution to connectivity coming from the amount of habitat within individual patches;
3. Measure connectivity as the amount of available (reachable) habitat in the landscape using the Integral Index of Connectivity (*IIC*) and the Probability of Connectivity (*PC*) metrics (and understand their formulas, ingredients, and behavior);
4. Learn to use the Conefor software to quantify landscape connectivity;
5. Evaluate the importance and different roles of individual habitat patches as connectivity providers; and
6. Apply these concepts and tools to a real-world conservation case study;

This lab assumes you have already gained familiarity with the basics of spatial networks in Chapter 12. Here, we go a step further with a more in-depth analysis on how (and which) connectivity metrics can be used to prioritize landscape elements for conservation planning. You will explore situations where many connectivity metrics fail to provide appropriate answers, particularly when used for identification of specific key patches or links essential to connectivity. First, you will perform exercises "by hand" to learn some new metrics (Exercises 1, 2 and 3). In Exercises 4 and 5 you will use Conefor, and in Exercise 6 you will combine Conefor with GIS for the analysis of a real-world example. Several printed handouts (available from the book website) will be helpful as you work through the exercises. While we frame this chapter within the context of a graph-theoretical approach for the analysis of landscape connectivity, the concepts here presented and illustrated are of a wider reach; they apply in general to the way landscape connectivity is conceived and measured and to the quantification of the different roles of habitat patches in landscape networks.

INTRODUCTION

Recall that a landscape graph consists of a set of nodes and links between them (Ricotta et al. 2000; Urban and Keitt 2001; Jordán et al. 2003; Pascual-Hortal and Saura 2006; Galpern et al. 2011; Rayfield et al. 2011). **Nodes** represent differentiated habitat units, which generally correspond to habitat patches (as we will assume hereafter in this lab) but may also correspond to other options such as habitat cells, river segments, management units, or protected areas. Nodes can be weighted to incorporate some characteristic (attribute) of the habitat units such as habitat area, quality, or population size. For simplicity, Parts 1 and 2 of this lab we will assume that the attribute corresponds to habitat patch area. Later, in Part 3, we will consider a case where habitat quality is incorporated in the attribute of the patches. **Links** represent ecological flow (usually the movement of an organism or species) directly between two habitat patches, without use of any intermediate stepping stone patch. Two patches not directly linked in the graph may still be connected through a **path** (a sequence of links) involving several **stepping stone** patches. A link may correspond to a physical corridor or it may represent potential of an organism to directly disperse between two patches; however, links contain no habitat area. Any landscape element containing habitat is represented as a node even when its main role is to serve as a stepping stone or connector between other habitat areas.

For simplicity in this lab, we assume **undirected graphs** where the possibility of moving from patch i to patch j is the same as moving from j to i. However, the concepts and metrics presented also apply to directed networks with asymmetric connections such as wind-driven dispersal or water flows. In Exercises 1–5 we will consider graphs with **unweighted links** (i.e., binary connection model where two patches are simply considered either directly connected or unconnected, with no intermediate modulation of the quality, strength, feasibility, or frequency of use of that connection). Later, in Exercise 6, we will consider a richer graph representation of the landscape (probabilistic connection model) in which links are **weighted** according to their ability or effectiveness in conducting a movement or ecological flows.

Connectivity: Is it Just Between Habitat Patches?

A classic definition of **landscape connectivity** is "the degree to which the landscape facilitates or impedes movement among resource patches" (Taylor et al. 1993). This definition implies that landscape connectivity is related to and can be successfully addressed by only considering the number or quality of connections among habitat patches in the landscape. Consider the two landscapes in Figure 14.1. Which landscape is more connected? Consider your answer using some of the metrics you are familiar with from the previous chapter: number of components, number of links or link density.

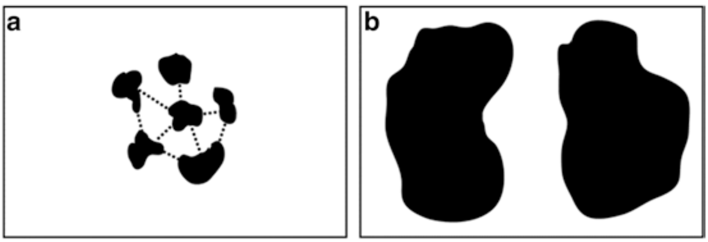


Figure 14.1 Two simple hypothetical landscapes (**a**, **b**) illustrating the concept of habitat availability (reachability) at the landscape scale. Habitat patches (nodes) are shown in *black* and the links (direct connections) between the nodes are represented by *dotted lines*. Adapted from Saura (2008)

Using many existing metrics, one would deduce that connectivity is higher in A than in B, since A contains eight links between patches while landscape B has no links. Similarly, landscape A would be regarded as more connected because all habitat occurs within one component, whereas B consists of two completely isolated components (i.e., the two big patches). However, it makes little sense to consider landscape A more connected than B. No matter how well connected patches are in A, together they comprise less amount of available (reachable) habitat than either single patch in B (Pascual-Hortal and Saura 2006). Either one of the big isolated patches in B (on its own) contains more area of connected habitat than the entire area of habitat potentially reached via the links in landscape A. In fact, habitat loss and fragmentation of either patch in landscape B could lead to a pattern similar to landscape A. From the standpoint of conservation, fragmentation of one large patch should not be deemed more beneficial (or more connected) than the original continuous habitat patch.

Clearly, some connectivity metrics indicate higher connectivity in more fragmented landscapes (Tischendorf and Fahrig 2000). In fact, some metrics indicate zero connectivity in landscapes with one contiguous patch, even when the habitat patch occupies the entire landscape (Tischendorf and Fahrig 2000). Approaches and metrics that focus only on the connections between habitat patches (**interpatch connectivity**), while ignoring **intrapatch connectivity** (e.g., area of the patches), are problematic for informing conservation planning priorities or evaluating the impacts of landscape changes (Pascual-Hortal and Saura 2006; Saura and Pascual-Hortal 2007). When habitat patterns change through time, or when networks with differing numbers and sizes of patches are compared, metrics that only consider interpatch connectivity can provide misleading results.

The concept of **habitat availability (or reachability)** at the landscape scale helps to address these deficiencies (Pascual-Hortal and Saura 2006; Saura 2008; Saura and Rubio 2010) in two ways. First, it recognizes that a habitat patch provides (by itself) some amount of connected habitat—more connectivity within bigger patches—even when the patch is completely isolated from all other patches. Second, the connected area *within* habitat patches (i.e., **intrapatch connectivity**) should be measured along with the area made available by (or reachable through) connections with other patches (i.e., **interpatch connectivity**) (Pascual-Hortal and Saura 2006; Saura and Pascual-Hortal 2007; Saura and Rubio 2010). The habitat availability concept acknowledges that species may be able to reach a larger amount of habitat resources in the landscape *either* through bigger patches (intrapatch connectivity) *or* through more or stronger connections among different patches (interpatch connectivity). More frequently, a combination of both will determine the total amount of available/reachable habitat in the landscape for a given species.

Through this chapter you will learn to understand and appropriately apply habitat reachability metrics which account for intra- and interpatch connectivity.

EXERCISES

Part 1. Understanding the Behavior and Limitations of Commonly Used Connectivity Metrics

Assume one wants to evaluate how important an individual habitat patch (or link) is to the maintenance of overall landscape connectivity, using a given landscape-level connectivity metric X. One approach would be to calculate the relative change in the metric value (dX) after the removal of that patch (or link) from the habitat network:

$$dX(\%) = 100 \bullet \frac{X_{\text{initial}} - X_{\text{removal}}}{X_{\text{initial}}}$$

where X_{initial} is the value of the connectivity metric in the initial or intact landscape and X_{removal} is the metric value after the removal of a particular patch or link from the landscape.

This calculation can be repeated for any (or every) individual element in the landscape, and the resulting dX values would quantify the importance of each patch (or link) in maintaining landscape connectivity. Therefore, ranking patches by dX values can prioritize their value to sustaining connectivity and thus provide guidance on where to concentrate conservation efforts. Similarly, dX values could also be calculated for landscape changes involving the loss of multiple patches and/or links.

There are a myriad of metrics to quantify the connectivity of landscape networks (e.g., Jordán et al. 2003; Galpern et al. 2011; Rayfield et al. 2011). Each would most likely yield different results for dX, and suggest different conservation priorities; therefore it is of utmost importance to scrutinize and understand the actual behavior, performance, and adequacy of various metrics. A key question is: *What properties should a connectivity metric fulfill to be reliable for such conservation planning purposes?*

To address this, you will examine the behavior of several commonly used connectivity metrics in response to changes in habitat networks. Your goal is to evaluate whether these metrics can be reliably used for prioritizing habitat patches and links for connectivity conservation. In particular, you will examine three metrics that are easy to understand, widely used in connectivity analyses, and similar to metrics you learned about in Chapter 12 in this book:

- **Number of Links (*NL*)**—A link is a direct connection between different habitat patches.
- **Number of Components (*NC*)**—A component is a set of connected patches in which every patch can be reached from the others through at least one path (sequence of links). There are no links or paths between patches in different components.

- **Mean Component Size (*MCS*)**—The size of a component is the sum of habitat areas for all patches within the component. *MCS* is the average size of all the components in the landscape.

Higher values of *NL* and *MCS* indicate more connectivity while the inverse interpretation applies for *NC*. Upon patch removal, *NL* will never increase (and hence $dNL \geq 0$). In contrast, *dNC* and *dMCS* can potentially yield positive or negative values. Therefore, the most important landscape elements (patches or links) for connectivity according to *NL* and *MCS* would be indicated by the highest *dNL* and *dMCS* values when these elements are removed (which for *dMCS* may correspond either to the highest positive or the least negative values). For *NC*, the most important landscape elements would be those that, when removed, produce the lowest *dNC* values (which may include negative *dNC* values).

EXERCISE 1: Response of Connectivity Metrics to Changes in Habitat Networks with Some Hypothetical Examples

Figure 14.2 shows six different hypothetical landscapes with their corresponding habitat networks (graphs with unweighted links). In each landscape, two different losses can occur—loss of either A or B—which generally correspond to different habitat patches. The exceptions are in landscape L5 where B corresponds to the link between two patches, and in landscape L6 where A corresponds to a component made up of three patches and two links (identified by a dashed line).

Next, we'll assume, due to budget constraints, that only A or B can be protected. Thus, a decision has to be made as to which element will be protected and which will be lost. When a patch is lost, all connected links are considered lost.

Q1 Visually examine the six landscapes in Figure 14.2. Based on your own qualitative visual assessment (no detailed calculations), would the loss of A or B be more detrimental to habitat connectivity and availability (reachability) in each landscape? Which element (A or B) should be prioritized for conservation? Why? Consider that the reasons might be different for each of the six cases.

Q2 Next, you will systematically and quantitatively examine the behavior of *NL*, *NC*, and *MCS* to determine which of the two losses (A or B) is more detrimental to connectivity according to these three metrics. Evaluate using *dNL*, *dNC*, and *dMCS* values or simply from the absolute difference in the metric value before and after a change (X_{initial}–X_{removal}). Recall that higher *NC* values are assumed indicative of lower connectivity with the reverse is true for *NL* and *MCS*. Use the tables provided in **Handout #1** (available on the website for the book) to organize your calculations. Example calculations are provided for landscapes L1 and L6 in Handout #1.

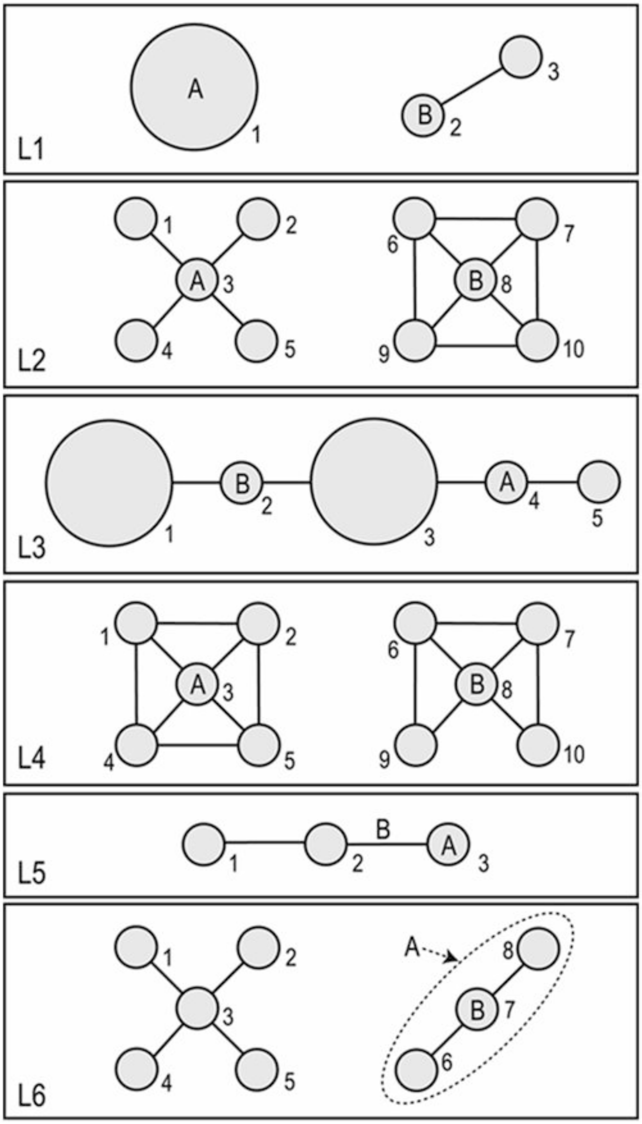


Figure 14.2 Six simple landscapes comprising different sets of habitat patches or nodes (*grey numbered circles*) and links between them (*solid lines*). In each landscape two hypothetical losses can occur indicated by *A* and *B*. The big patches in *L1* (*patch 1*) and *L3* (*patches 1 and 3*) have a habitat area of 9 ha whereas the rest of the patches have a habitat area of 1 ha. See Exercise 1 for additional explanations. Later in Exercise 2, you will need to know that the total area of each landscape is 25 ha.

Q3 What do you conclude about the behavior and reaction to spatial changes of these connectivity metrics? Would you advocate the use of *NL*, *NC*, or *MCS* as a basis for prioritizing landscape elements for the conservation of habitat connectivity and availability or for evaluating the potential impacts of landscape changes in the ecological flows related to connectivity? Why?

Part 2. Integrating Intrapatch and Interpatch Connectivity in a Single Measure: The Integral Index of Connectivity

Next you will explore new connectivity metrics which measure the amount of available (reachable) habitat in the landscape (Pascual-Hortal and Saura 2006; Saura and Pascual-Hortal 2007; Saura and Rubio 2010). These metrics integrate both intrapatch (within patch) and interpatch (between-patch) connectivity in a single measure. In doing so, these new metrics address the main primary limitations of previous metrics you have examined.

Now, in Part 2, you will only consider the habitat availability metrics that are based on a binary connection model (graphs with unweighted links). These metrics are the **Integral Index of Connectivity** (*IIC*) and the related **Equivalent Connectivity** metric (*EC(IIC)*). In the particular case where the node attribute corresponds to habitat area, as in exercises in Part 2, the latter metric is referred to as the Equivalent Connected Area (*ECA(IIC)*). These metrics are described in Table 14.1. They derive from the same concept and way of measuring connectivity but are expressed in different units and over a different range of variation. Later, in Part 3, we will introduce a probabilistic version of these metrics corresponding to graphs with weighted links.

EXERCISE 2: Understanding *IIC* and *ECA(IIC)* Calculations

1. Examine the definition of the *IIC* and *EC(IIC)* metrics given in Table 14.1. The latter metric will be referred to as *ECA(IIC)* hereafter, since we are using habitat area as the attribute of the nodes. Pay attention to the variable in these metrics that relates to the number of links in the shortest path between patches (nl_{ij}), which can take different values depending on the pair of patches considered, as described next.
   - **Interpatch connectivity** is addressed in several ways. For a **direct link** between two patches, $nl_{ij}=1$ (e.g., patches 2 and 3 in L1 in Figure 14.2). **Indirect connections** occur between patches if *i* and *j* belong to the same component (i.e., there is a path from *i* to *j*), but have no direct link. In such cases, $1<nl_{ij}<\infty$. This occurs for example for patches 9 and 10 in L4 in Figure 14.2, where $nl_{ij}=2$. For **unconnected patches** (not connected through any path), $nl_{ij}=\infty$ (no matter how many links are traversed, movement from *i*

Table 14.1 Habitat availability (reachability) metrics: formulas and description

Metrics for the **binary connection** model (graphs with unweighted links) used in **Part 2**

Metric name	Acronym	Formula	Description	Reference
Integral index of connectivity	*IIC*	$IIC = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{a_i \cdot a_j}{1+nl_{ij}}}{A_L^2} = \frac{IICnum}{A_L^2}$	*IIC* ranges from 0 to 1 and increases with improved connectivity where n is the total number of nodes (usually patches) in the landscape, a_i and a_j are attributes (e.g., habitat area, quality) characterizing habitat resources in patches i and j, and nl_{ij} is the number of links in the shortest path between patches i and j (further details on nl_{ij} are given in Exercise 2). A_L is the maximum landscape attribute, a constant with the same units as a_i and a_j and sets the reference for the maximum *IIC* attainable ($IIC = 1$). If patch attributes are habitat areas (as assumed in Exercise 2), then A_L is equal to the total landscape area. Therefore $IIC = 1$ when the landscape is fully covered by habitat.	Pascual-Hortal and Saura (2006)
Equivalent connectivity for *IIC*	*EC(IIC)*	$EC(IIC) = \sqrt{IICnum} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{a_i \cdot a_j}{1+nl_{ij}}}$	*EC(IIC)* is the amount of habitat resources (in the same units as the attribute a_i) that a single habitat patch (maximally connected) should have to provide the same *IIC* value as the actual network currently in the landscape. *EC(IIC)* may be thought of as a measure of the amount of reachable habitat resources in a landscape. *EC(IIC)* will not be smaller than the largest a_i and will not exceed the sum of a_i values for all patches in the landscape (total amount of habitat resources). The maximum possible *EC(IIC)* for a given amount of habitat resources occurs when all the habitat occurs as a single habitat patch (no fragmentation). In the particular case in which the attribute a_i corresponds to patch area, *EC(IIC)* can be named as Equivalent Connected Area for *IIC* (*ECA(IIC)*) and be defined as the size of a single habitat patch that would provide the same *IIC* value as the actual habitat network in the landscape.	Saura et al. (2011)

(continued)

Table 14.1 (continued)

Metrics for the **probabilistic connection** model (graphs with weighted links) used in **Part 3**				
Metric name	Acronym	Formula	Description	Reference
Probability of connectivity	*PC*	$PC = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} a_i \cdot a_j \cdot p_{ij}^*}{A_L^2} = \frac{PCnum}{A_L^2}$	*PC* ranges from 0 to 1 and increases with improved connectivity. *PC* is defined as the probability that two individuals of a given species, randomly placed within the landscape, are able to reach each other (i.e., fall into habitat connected to each other). This will happen either if both individuals are placed within the same patch (intrapatch connectivity) or if they are placed in different patches connected through a path (interpatch connectivity). p^*_{ij} is defined as the maximum product probability of all the paths between patches *i* and *j* (see Part 3 for further details and examples). All the rest is the same as for *IIC*.	Saura and Pascual-Hortal (2007)
Equivalent connectivity for *PC*	*EC(PC)*	$EC(PC) = \sqrt{PCnum} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_i \cdot a_j \cdot p_{ij}^*}$	*EC(PC)* is completely analogous in interpretation to *EC(IIC)* but calculated from *PCnum* rather than from *IICnum*. For example, when patch attributes correspond to areas, *EC(PC)* can be called *ECA(PC)* and is defined as the size of a single patch providing the same *PC* value as the actual habitat network. The maximum possible *EC(PC)* value for a given amount of habitat would be obtained when all habitat occurs as a single habitat patch (no fragmentation) or when the habitat is dissected into different patches but with a maximum interpatch connectivity ($p^*_{ij}=1$ between all patches).	Saura et al. (2011)

Table 14.2 Calculation of the Integral Index of Connectivity (*IIC*) and the Equivalent Connected Area (*ECA(IIC)*) for **landscape L3** in Figure 14.2. Some missing values need to be filled to complete the calculations

	Patch i	Patch j	a_i	a_j	nl_{ij}	$\frac{a_i \cdot a_j}{1+nl_{ij}}$	$\sum \frac{a_i \cdot a_j}{1+nl_{ij}}$
Intrapatch connectivity ($i=j$)	1	1	9	9	0	81	165
	2	2	1	1	0	1	
	3	3		9	0		
	4	4	1		0	1	
	5	5	1	1	0	1	
Interpatch connectivity for $i<j$	1	2	9	1	1	4.5	48.63
	1	3	9	9	2	27	
	1	4				2.25	
	1	5	9	1		1.8	
	2	3	1	9	1	4.5	
	2	4	1	1	2	0.33	
	2	5	1	1	3	0.25	
	3	4	9	1		4.5	
	3	5		1	2		
	4	5	1	1	1	0.5	
Interpatch connectivity for $i>j$	Same as above for $i<j$ (undirected graphs)						48.63
Total connectivity (*IICnum*) = intrapatch + interpatch connectivity							262.26
$IIC = IICnum / A_L^2 = 262{,}26/25^2 = 0.4196$							
Equivalent Connected Area ($ECA(IIC) = IICnum^{0.5}$) = 16.19 ha							

to j will never be achieved) and the numerator equals zero. This is the case for example for patches 1 and 2 in L1.

- **Intrapatch connectivity** is also accounted for. When $i=j$ then $nl_{ij}=0$ because no links are needed to reach a patch from itself. This relates to the habitat availability concept, in which a patch itself is considered to provide some connected habitat.

2. Examine the example (partial calculation) of *IIC* and *ECA(IIC)* given in Table 14.2. Determine generally how these numbers correspond to **landscape L3** in Figure 14.2 from which these calculations are derived. Remember that the total landscape area (A_L as in Table 14.1) is 25 ha (as in all landscapes in Figure 14.2).
3. Fill in the missing values in the table (use **Handout #2**) to ensure you understand the formulation of these calculations. Check your answers with your partner (or instructor) at each step, before proceeding.
   - Note patches 1 and 3, which are very large compared to the rest of the network. The large connected areas within these patches will make a large contribution to the value of *ECA(IIC)* due to their high **intrapatch connectivity**.

- Recall that connections are symmetric (as for undirected graphs). Thus, we can simplify the calculations for the **interpatch connectivity** by just considering the cases where $i<j$ and use identical values for $i>j$.
- Due to the contribution of both intrapatch and interpatch connectivity $ECA(IIC)=16.19$ ha, about 6 ha lower than the total habitat area (22 ha) in the network.

EXERCISE 3: Evaluating Changes in Connectivity (by-hand) Using Habitat Availability Metrics *IIC* and *ECA(IIC)*

Next you will explore some "by-hand" calculations in two of the simplest landscapes from Figure 14.2 to analyze the behavior of *IIC* and *ECA(IIC)*. Later you will learn to use Conefor software to perform such computations more rapidly and automatically which will be necessary when we move to analysis of real-world networks with a much larger number of patches.

Examine the example calculations for L1 (shown in a series of three tables in **Handout #3** and described below). Remember that the total landscape area (A_L) is 25 ha for these landscapes. After you follow and understand these calculations for L1, you will follow the same procedure for L5. Recall your work from Table 14.2, which can be a useful guide together with the example calculations given next. Your goal is to determine: Does loss of a or b represent a bigger problem for connectivity?

1. Calculations for intact landscape L1 are shown in **Handout #3** Table (a).
2. If we lose A (patch 1) from the landscape, the table simplifies to that shown in Table (b).

   By losing A the *IIC* value decreases from $IIC_{\text{initial}}=0.1344$ to $IIC_{\text{remove}}=0.0048$ (the *IICnum* value decreases from $IICnum_{\text{initial}}=84$ to $IICnum_{\text{remove}}=3$), which yields $dIIC=96.43\%$. Note that since the *IIC* and *IICnum* values are proportional (the former equals the latter divided by A_L^2), *dIIC* and *dIICnum* values will be identical, and therefore you can use *IICnum* to calculate *dIIC*.
3. If we lose B (patch 2) from the landscape, then the table simplifies to that shown in Table (c).

   By losing B the *IIC* value decreases from $IIC_{\text{initial}}=0.1344$ to $IIC_{\text{remove}}=0.1312$ ($IICnum$ value decreases from $IICnum_{\text{initial}}=84$ to $IICnum_{\text{remove}}=82$), which yields $dIIC=2.38\%$. This *dIIC* value is lower than the one obtained by losing A, which was 96.43%. Therefore, the loss of A would cause a much larger decrease in habitat connectivity and availability in this landscape, and thus A is more important than B according to *IIC*.
4. Following the above example (detailed in **Handout #3**), perform the necessary calculations to analyze the behavior of *IIC* and *ECA(IIC)* for **landscape L5**. You may find it helpful to create a similar series of data tables for L5.

Q4 Which of the two possible losses (A or B) would be more detrimental for connectivity (according to *dIIC* values) in landscape L5?

Introduction to Conefor Software

Although it is advisable to use hand calculations initially to fully understand these metrics, this is cumbersome and infeasible in realistic networks. Fortunately, these metrics have been implemented in the Conefor software available for download at http://www.conefor.org. Conefor calculates both the overall metrics for the whole landscape (e.g., *IIC* and *EC(IIC)*) and the patch-level importance (*dIIC*) for every patch in the network, among other outputs.

INPUT

Conefor requires two input files for each network you wish to analyze: the **node file** and the **connection file**. Both are simple text files (with columns separated by tabs or spaces) that can be produced with any basic text editor (later in Part 3 you learn how to automatically generate these files using a custom-made GIS extension).

The **node file** simply has one row for each node and two columns. The first column identifies each node by a distinct node ID (integer value), and the second column contains its corresponding attribute value for that node (a_i and a_j in equations in Table 14.1, usually corresponding to habitat area).

In the case of landscape L3 in Figure 14.2, the node file is simply:

```
1    9
2    1
3    9
4    1
5    1
```

The **connection file** identifies connections between each pair of nodes, presented in three columns. The first two columns contain a pair of node IDs, whereas the third column characterizes the direct connection between the node pair. In our example, where connections correspond to binary links, they are represented as 1 (linked) or 0 (not linked). Alternatively, distances or probabilities for each pair of nodes could be used. Because our connections are symmetric (undirected graphs), each pair of nodes needs to be listed only once, although both directions will be considered in the calculations. The ordering of pairs in the file has no effect on Conefor calculations.

In the case of landscape L3, the connection file would be as follows:

```
1    2    1
1    3    0
1    4    0
1    5    0
2    3    1
```

```
2    4    0
2    5    0
3    4    1
3    5    0
4    5    1
```

The above format is used in a **full connection file** where **all pairs** of nodes are listed once.

It is also possible to list only node pairs with direct connections (links, 1 values only), to create a **partial connection file**. In such a file, any missing node pairs are considered not directly connected. The partial connection file is much shorter, with one line for each link, as follows:

```
1    2    1
2    3    1
3    4    1
4    5    1
```

EXERCISE 4: Using Conefor to Calculate the Importance of Nodes and Links

1. Copy the entire folder **Exercise 4** to C:\temp\ or C:\workspace\
2. Open and examine the digital node, full, and partial connection files for L3 with any text editor. Do they match the above description?
3. See the **Conefor Instructions Handout** for steps 3–7.
   Using Conefor, be sure you can successfully run the example for L3 (using the provided files) before proceeding to subsequent steps. Calculate the *IICnum*, *IIC*, and *ECA(IIC)* values for this landscape and check that you obtain the same values as in Table 14.2 above. **Save your output files**.
4. Build node and <u>partial</u> connection files for the other five landscapes in Figure 14.2. *NOTE:* Each of these files should end with a blank line.
5. Using Conefor, calculate the *dIIC* values for each of the possible changes (A or B) in each of the six landscapes. **Save your output files.** In almost all the cases, loss of A or B corresponds to specific patches and the *dIIC* values will be calculated by Conefor in the node importance file. Exceptions include **B in L5 and A in L6**, explained next.
6. B in L5 corresponds to the loss of a **link**. Thus, the *dIIC* value for this link is in the link importance file resulting from the link removal analysis in Conefor.
7. Loss of A in L6 involves **multiple patches and links**. Conefor will not automatically calculate the corresponding *dIIC* value. In this case you need to:
   - Get the *IIC* or *IICnum* value corresponding to L6 (in the initial landscape, using the node and connection files created in step 4).

- Build node and partial connection files corresponding to landscape L6 after the entire component A has been lost.
- Use these files in Conefor to calculate the *IIC* or *IICnum* values for this modified network.
- Calculate *dIIC* for the metric values before and after the loss of component A.

NOTE: Remember that the values of *dIIC* and *dIICnum* are, by definition, identical since A_L is a constant that remains unvaried before and after any change in the landscape. You can therefore use either *IIC* or *IICnum* values to obtain the requested *dIIC* values.

8. Complete the table in **Handout #4**. The results for L1 (from Exercise 3) are already included. Compare your Conefor results for L5 with your manually calculated results from Exercise 3.

Q5 When considering the *dIIC* values, which of the two losses (A or B) is more detrimental to connectivity in each landscape? Does *IIC* identify conservation priorities in a way that is more relevant to your responses to Q1? Explain why.

Understanding Three Distinct Fractions of Landscape Connectivity

Now that you are familiar with the habitat availability (reachability) metrics, we will examine the ingredients of the *IIC* metric in more detail and explore how these can be used to gain a more thorough understanding of the role of specific habitat patches in a network.

The *dIIC* values for a given patch can be partitioned into three distinct fractions which quantify the different ways a patch can contribute to habitat connectivity and availability in the landscape (i.e., contribute to the amount of reachable habitat) (Saura and Rubio 2010):

$$dIIC = dIICintra + dIICflux + dIICconnector$$

The **intra fraction (*dIICintra*)** is the contribution of the patch in terms of **intra-patch connectivity**, corresponding to $a_i \cdot a_j/(1+nl_{ij})$ when $i=j$ and, therefore, $nl_{ij}=0$. It corresponds to the amount of habitat resources (habitat area, quality, or other attribute) provided by the patch (i.e., the amount available or reachable from within that patch). *dIICintra* is **completely independent of the patch's connections** to other patches. This metric returns the same value even if the patch is completely isolated.

The **flux fraction (*dIICflux*)** corresponds to the **dispersal flux** (weighted by the focal patch attribute) through the connections of the focal patch with all other patches in the network. It assumes the focal patch is the starting (or ending) point of the dispersal flux. This fraction depends on both the attribute of the focal patch and its position within the network. It corresponds to the sum $a_i \cdot a_j/(1+nl_{ij})$ for each node pair where $i \neq j$. This fraction measures how well connected the focal patch is to the rest of the habitat in the landscape. It does not quantify the patch's importance for maintaining connectivity among the other patches, however, which is quantified by the next fraction.

The **connector fraction (*dIICconnector*)** quantifies the contribution of the analyzed (focal) patch as a connecting element or **stepping stone** between other patches. This fraction depends only on the **topological position** of the patch in the network. The connector fraction for a focal patch is independent of the attributes of the focal patch, but accounts for attributes of other patches connected via it. Thus, the connector fraction for a patch will be higher when it connects patches with more habitat resources (higher a_i).

dIICconnector for focal patch k corresponds to a part of the sum of $a_i \cdot a_j/(1+nl_{ij})$ for each pair of patches i and j in which $i \neq k$, $j \neq k$, and k is part of the shortest path between i and j. A given patch k will contribute to *dIIC* through the connector fraction only when it is part of the shortest path between at least two other patches. The value of *dIICconnector* for patch k also depends on any alternative paths which, upon loss of patch k, still allow movement among other patches. If remaining alternative paths are nearly as good, the connector fraction for k will be low; if patch k is irreplaceable (its role cannot be compensated for by other patches or paths after its loss), then it will present a higher *dIICconnector* value. See Saura and Rubio (2010) and Bodin and Saura (2010) for further details and equations.

These three fractions allow for multifaceted, integrated connectivity analyses in which the different roles of habitat patches are measured using identical units and can be directly compared and summed (Saura and Rubio 2010). *dIICintra* measures intrapatch connectivity, while *dIICflux* and *dIICconnector* measure interpatch connectivity. A patch will be more or less important (*dIIC*) due to one or more of these three fractions, depending on its local characteristics (i.e., attributes) and its topological position within the network. Since, by definition, links do not contain any habitat area, they do not provide intrapatch connectivity (thus, *dIICintra* = 0); nor can they be the starting or ending flux of any dispersal flux (thus, *dIICflux* = 0). As such, links can only contribute to *dIIC* through the *dIICconnector* fraction. Because the connecting role of nodes and links is measured in the same way by the *dIICconnector* fraction, their contributions can be directly compared.

EXERCISE 5: Examining Results from Intra, Flux, and Connector Fractions

Q6 Without making any calculations, which of the patches in each of the six networks in Figure 14.2 contribute via the intra fraction? Which contribute through the flux fraction? And which patches contribute through the connector fraction? Why?

Q7 Go back to your results from Conefor where you calculated *dIIC* (for Q5). Examine the values for the three fractions (*dIICintra*, *dIICflux*, *dIICconnector*) for all the patches in Figure 14.2. How do these results compare to your answer to the previous Q6? Which patches present the highest values of each fraction in each landscape?

Q8 Without making any calculations, if the same procedure and fractions were used to evaluate links rather than patches, which links in the six landscapes in Figure 14.2 would have no importance according to *dIIC*? Why?

Q9 Use Conefor to calculate the *dIIC* values for each link in landscapes L2, L3, and L4, selecting the link removal mode under the link importance options (See Conefor Instructions Handout). Do these results match to your answer to previous Q8? Among the links with $dIIC>0$, which are the most important and the least important ones in each landscape? Why?

Part 3. Connectivity Conservation Planning for an Endangered Bird in Spain

The previous exercises demonstrated how *IIC* reacts to certain changes in habitat networks and how it can prioritize landscape elements for conservation, and also helped you understand what aspects of connectivity are being measured by this metric. Next, you will learn how to use habitat availability metrics and the Conefor software in a real-world case study for an endangered bird species in the region of Catalonia in Spain.

In Exercise 6, you will analyze connectivity using the probabilistic *PC* metric (Table 14.1). This metric considers more complex information about links using graphs with weighted links, but is otherwise conceptually similar to *IIC*. The *dPC* values can be partitioned in three distinct fractions as for *dIIC* (i.e., *dPCintra*, *dPCflux*, *dPCconnector*). An Equivalent Connectivity metric (*EC(PC)*) can also be calculated from the numerator of *PC* and it is denoted *ECA(PC)* when the node attribute is area (Table 14.1). You will also use Conefor with ArcGIS (or QGIS) to link GIS data to graph-based connectivity analyses. Upon completion, you will be able to adapt and apply the *IIC* and *PC* metrics and Conefor to other study areas and species for your own project.

Understanding the Probability of Connectivity (*PC*) Metric

PC is based on a probabilistic connection model. Links are weighted by $\boldsymbol{p_{ij}}$ which is the **probability of direct dispersal** between patches *i* and *j*. The **product probability of a path** is the product of all p_{ij} along the path. If intermediate patches are traversed, the product probability of a path incorporates all intermediate links (Saura and Pascual-Hortal 2007).

A key ingredient in the formulation of *PC* is the **maximum product probability** (**p^*_{ij}**) (Table 14.1). This is the only ingredient that differs from *IIC*. p^*_{ij} identifies the maximum product probability among all possible paths between patches i and j to determine the "best" path (and thus, $p^*_{ij} \geq p_{ij}$).

When patches i and j have a **strong direct connection** (e.g., close to each other), the maximum product probability path is the direct link connecting i and j, and then $p^*_{ij} = p_{ij}$. In contrast, if patches i and j are weakly or not connected through the direct link, the "best" (maximum product probability) path may follow several intermediate steps via **stepping stone patches**, yielding $p^*_{ij} > p_{ij}$. When two patches are completely isolated from each other (either due to great distance or barriers such as a road) and there is no possibility for movement between both patches, then $p^*_{ij} = 0$.

When $i = j$ then $p^*_{ij} = 1$ because it is 100% certain that a patch can be reached from itself. This relates to the habitat availability (reachable habitat) concept, which accounts for the amount of habitat resources available within a patch (the intrapatch connectivity).

The **maximum product probability path (quantified by p^*_{ij}) is not necessarily the same as the shortest path for *IIC*** where it is quantified as the number of links (nl_{ij}). An example using Figure 14.2 is instructive to consider. For example, in L2, assume $p_{ij} = 0.5$ for all links except $p_{6,7} = 0.1$. Given this, the direct link yields $p_{6,7} = 0.1$ whereas an indirect link (via patch 8) yields $p^*_{6,7} = 0.25$ (8 is a stepping stone, therefore $0.5 \times 0.5 = 0.25$). Thus, a two-link path is a better option than the direct link between 6 and 7. The direct link between patches 6 and 7 is the shortest path according to *IIC,* but for *PC,* movement is better conducted via patch 8, when p_{ij} values are incorporated.

Status of the Endangered Capercaillie in Spain

The capercaillie is one of the most endangered forest-dwelling bird species in Spain (Canut et al. 2011). We focus on the subspecies of capercaillie (*Tetrao urogallus aquitanicus*) endemic to the Pyrenees and its distribution in Catalonia (NE Spain), a region of 32,000 km^2 (Figure 14.3). Its habitat in Catalonia is concentrated in upper montane and subalpine forests of the Pyrenees and Pre-Pyrenees (Figure 14.3). Decreasing populations with poor breeding success raise concerns about long-term persistence and habitat fragmentation is a major concern (Canut et al. 2011). Common conservation measures focus on protection of vital areas (leks, hibernating and breeding sites, etc.); however, this approach may not be sufficient to meet the birds' broad spatial requirements. Sustaining functional connectivity among subpopulations to facilitate dispersal and minimizing mortality risk (Canut et al. 2011) requires identification and conservation of areas most critical for habitat connectivity and habitat availability. You will address this using the *PC* metric and actual geodata on habitat quality for this species, with some simplifications and modifications for teaching purposes.

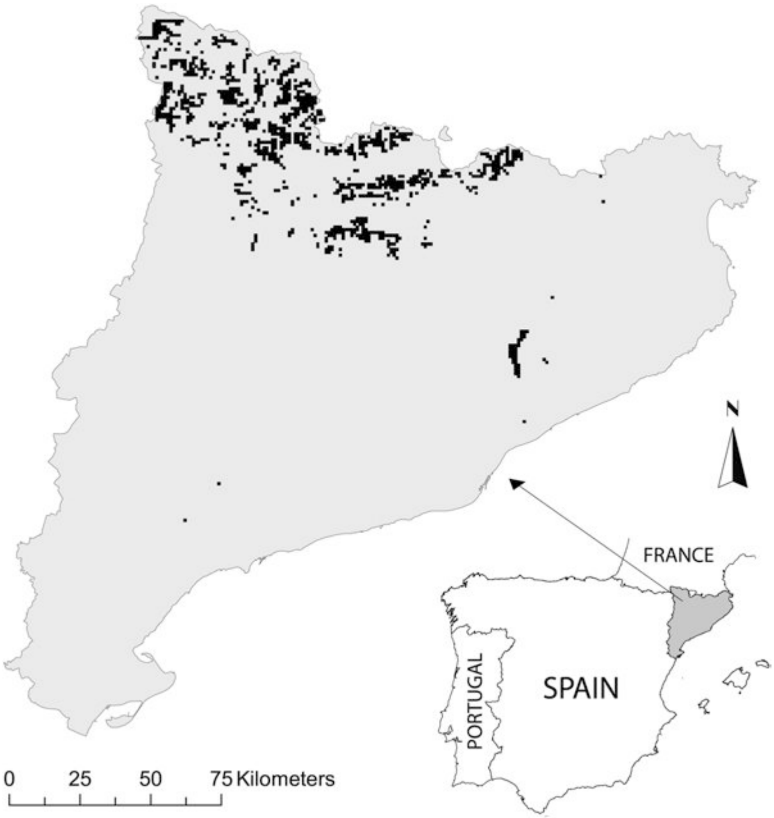


Figure 14.3 Location of the study area (Catalonia, NE Spain) and capercaillie habitat shown in *black* (see text for further details)

Habitat Mapping, Habitat Quality, and Node Creation

Habitat distribution data for the capercaillie in Catalonia (Figure 14.3) were obtained from the Catalan Breeding Bird Atlas 1999–2002 (Estrada et al. 2004). Presence–absence data were gathered in ~3000 UTM 1 × 1 km cells. Field survey data were used to build a niche-based model estimating probability of occurrence (ranging from 0 to 1) in 1 x 1 km cells (described further in Estrada et al. 2004). Probability of occurrence can be used as a measure of habitat quality, so cells with a higher probability were considered more suitable and able to support more individuals (Estrada et al. 2004). The **habitat distribution layer** (**capercaillie.shp**) is a slightly modified and simplified version of these data. Only cells with a probability of capercaillie occurrence ≥ 0.1 are considered habitat for Exercise 6. Cells were grouped into **117 habitat patches (or nodes)** comprising 1173 km^2 total habitat (Figure 14.3). **Habitat quality** was calculated as the average probability of occurrence for all the 1 × 1 km cells within a patch.

Both the amount and the quality of habitat are relevant patch characteristics for this habitat analysis. Thus, the **area-weighted habitat quality** is used as the node attribute (a_i) which corresponds to the product of habitat area × habitat quality. The values of habitat area and quality are provided for each patch in the GIS layer. Remember, since the patch attribute is not simply area, we will use *EC* (Equivalent Connectivity) rather than *ECA* (Equivalent Connected Area).

Defining Links and Direct Dispersal Probabilities

Determining the strength of the links between habitat patches (p_{ij} values) involves two basic components (explained step-by-step, later in Exercise 6):

- **Conefor Inputs extension**. This runs in ArcGIS (or QGIS) and can be used to compute distances between the edges of habitat patches (d_{ij}) which are measured as Euclidean (straight-line) distances. This extension creates the connection file in the input format required by Conefor.
- **Median dispersal distance (*d*)** relates to the median distance individuals can reach when dispersing. When running Conefor, you will need to specify this distance. You will use a **5000 m** median dispersal distance for capercaillie (derived from Menoni (1991) for the Pyrenees). By definition, $p_{ij}=0.5$ for links between patches separated by distance *d*. For each pair of patches, p_{ij} values are calculated by Conefor using a (fairly typical) decreasing exponential function based on distance (e.g., Bunn et al. 2000; Hanski and Ovaskainen 2000; Urban and Keitt 2001; Saura and Pascual-Hortal 2007) where $p_{ij}=1$ when the distance between patches is zero. The exponential decay rate is determined by the probability and dispersal distance specified by the user (here 0.5, and 5000 m, respectively). Thus, in our case, patches separated by less than 5000 m will get assigned $p_{ij}>0.5$, while patches separated by distances larger than *d* will be assigned $p_{ij}<0.5$.

EXERCISE 6: Network Analysis for Capercaillie Habitat

Now you will analyze connectivity of the capercaillie habitat network using the *PC* metric from Conefor. This will enable you to evaluate the contribution of individual patches by examining *dPC* and its three fractions *dPCintra*, *dPCflux*, and *dPCconnector* for each node. In addition to the capercaillie (with median dispersal distance $d=5000$ m), you will also consider two hypothetical species dwelling in the same habitat with different dispersal abilities ($d=500$ m and $d=50{,}000$ m). You will compare results and conservation guidelines for these three species.

To do so, follow these steps, referring to the **Conefor Instructions Handout** when needed:

1. Make sure that the point (and not the comma) is set as the decimal symbol in the regional configuration settings in your computer. Conefor will expect all the numerical values having the point as the decimal separator and no thousand separator, and will write the results in the same format. In North America, you will likely not need to worry about this step.
2. Copy the entire folder **Exercise 6** to C:\temp\ or wherever your instructor is certain you have full read/write permissions.
3. Using ArcGIS or QGIS, open the capercaillie habitat layer (**capercaillie.shp**) and a layer of the study area boundary (**catalonia.shp**), located within the Exercise 6 folder. Examine the distribution of the habitat patches.
4. In the GIS, open and examine the attribute table in the **capercaillie.shp** file and make the following changes:
   - Create two new fields in the attribute table that correspond to the two columns needed in the Conefor node file. One field will contain a unique identifier (an integer) for each node whereas the other field (floating type) will contain the patch attribute. Name these two new fields **NodeID** and **AreaQual**, respectively.
   - Fill the **NodeID** field with unique integer values for each feature (patch). In ArcGIS this can be done using the internal feature identifier of the layer (FID) as NodeID = FID + 1, so that NodeID ranges from 1 to the total number of patches. If using QGIS, use the variable $id to calculate the new field NodeID as $id + 1.
   - Calculate **AreaQual** so that it contains the value of the attribute for each patch. Recall that here, this attribute equals the product of habitat patch area × habitat quality.
5. Use the provided **Conefor Inputs extension for ArcGIS** or QGIS to generate the node and connection files in the format required for Conefor. The Conefor Inputs extension will calculate Euclidean (straight-line) edge-to-edge distances (here in meters) between all pairs of patches in the layer, and present this information in the **connection file**. (*NOTE:* In ArcMap, you may need to select **Customize—Toolbars**—and select **Conefor**):
   - In the ArcGIS extension, select to compute distances between all features.
   - Select to calculate distances from feature edges.
   - In the ArcGIS extension, use the **ASCII text** file as the output option.
   - The names of the node and connection files will typically be: **nodes_capercaillie.txt** and **distances_capercaillie.txt**.

   Check the online Conefor Inputs extension user manual (available from www.conefor.org) for further details on the usage of the extension if needed.
6. Open the two text files generated by the Conefor Inputs extension with any text editor and check for consistency before proceeding. (*NOTE:* The point "." should be the decimal separator symbol and no thousand separator should be in the numbers. For example, 1234.5 is correctly formatted whereas 1,234.5 or 1234,5 is not. Refer back to Step 1 if needed.)

7. Use Conefor to perform a network connectivity analysis with the *PC* metric using the node and connection files generated in the previous step. Specify the median dispersal distance for the capercaillie ($d=5000$ m, corresponding to a $p_{ij}=0.5$). **See the Conefor Instructions Handout** for more details on using this software.
8. Save the following two files generated by Conefor:
   - **overall metric values** (containing the *PCnum* and *EC(PC)* values for the entire habitat network)
   - **node importances** (containing the values of *dPC* and its three fractions for every patch). The node importance file will also contain the *dA* values for every node, which is the percentage (%) of the total landscape attribute within a particular node (i.e., a_i for a patch *i* divided by the sum of a_i for all nodes). *dA* is not a connectivity metric but rather a useful "network-independent" reference to compare with the connectivity metrics, as you will do later in this exercise.
9. Open the node importance file as a table in ArcGIS or QGIS and join the numerical results to your original capercaillie habitat layer, based on the common field (node ID).
10. Repeat the analyses in steps 7 and 8 for two other hypothetical species with different median dispersal distances of $d=500$ m and $d=50{,}000$ m. See the **Conefor Instructions Handout** for more details if needed.

Q10 What is the Equivalent Connectivity *EC(PC)* value for capercaillie habitat in Catalonia? How can this value be interpreted? What are the *EC(PC)* values for the other two hypothetical species with $d=500$ m and $d=50{,}000$ m?

Q11 Without using Conefor, what are the minimum and maximum possible values of *EC(PC)* for this set of habitat patches? Consider two hypothetical species with zero and infinite dispersal abilities to answer this question.

Q12 Visually examine the capercaillie connectivity results using GIS. You might consider seven classes and natural breaks (jenks) to classify patches by importance values. According to *dPC*, which habitat areas are most important in maintaining overall habitat connectivity and availability?

Q13 How do the three fractions of the *dPC* metric (*dPCintra*, *dPCflux*, *dPCconnector*) distinguish the roles of different patches in the habitat network? Which patches likely exchange a larger number of individuals with other habitat areas? Which patches are best connected to the rest of the habitat in the landscape? Which patches function as important, somewhat irreplaceable stepping stones?

Q14 How important is the contribution of the three *dPC* fractions to total habitat connectivity and availability for the capercaillie? To answer this, determine the relative contribution made by each fraction. Compute this as the ratio between the sum of the delta values for a particular fraction across all the patches, divided by the sum of the total *dPC* values across all patches. See the **Conefor Instructions Handout** for details.

Q15 Given the relative importance of these three fractions, which of the three patch roles is more appropriate for this network and species? How should management approach the spatial priorities for capercaillie conservation?

Q16 Respond to the same questions as in Q14 and Q15, but now consider the other two species using the same habitat ($d = 500$ m and $d = 50{,}000$ m). Do the primary roles of habitat patches change as different species are considered? What implications does this have for the conservation of each species?

Q17 Determine the sum of the *dPC* values for all the habitat patches for the capercaillie ($d = 5000$ m). Does this value exceed 100%? Why?

Q18 Plot the *dA* values (*X*-axis) against the *dPC* values (*Y*-axis) for capercaillie patches ($d = 5000$ m). How similarly do *dA* and *dPC* prioritize patches? That is, are the most important patches ranked similarly?

Q19 Produce three new plots as above but instead with each of the *dPC* fractions (*dPCintra*, *dPCflux*, *dPCconnector*) in the *Y*-axis. Which of the fractions are more and less related to the local patch attributes (*dA*)? Why is this so?

Q20 Produce similar plots for the two species where $d = 500$ m and $d = 50{,}000$ m. What has changed, and why? Do local attributes (of individual patches) have more or less weight on the prioritization of patches given by the metric of habitat connectivity and availability?

CONCLUSIONS

Many connectivity metrics have been developed and the choice of metric depends on the question at hand. However, to inform conservation planning or to evaluate the impacts of landscape change, metrics that only consider interpatch connectivity may provide misleading results. Connectivity should not be solely conceived of, or defined as, connectivity among patches in many cases. Rather, connectivity should be viewed as a landscape-level property describing the amount of habitat resources a species can reach: a combination of the resources within patches as well as those which can be reached via connections to other patches. Ecologically, it seems reasonable that the amount of reachable habitat may be more related to species persistence than (a) the total amount of habitat in the landscape (which ignores the likelihood of movement among the different patches) and (b) the connections among patches (which may not compensate for having less total habitat distributed in many smaller patches).

The *IIC* and *PC* metrics have been developed to address this new way of conceiving connectivity as the amount of reachable habitat. They provide enriched indicators which incorporate the role of habitat amount and local patch characteristics in influencing connectivity. At the same time, they move beyond spatially blind assessments by incorporating connectivity among patches. The *intra*, *flux*, and *connector* fractions allow for comparison of the different ways habitat patches contribute to habitat connectivity and availability. These fractions are measured in a common currency within an integrated conceptual and analytical framework enabling objective decision-making. When using *IIC* or *PC,* there is no risk of either overemphasizing or underestimating the importance of connecting elements between habitat patches when setting conservation priorities, since both alternatives are integrated and jointly considered in the same analysis.

FURTHER APPLICATIONS

Through this chapter, you have learned to understand the *IIC* and *PC* metrics and the Conefor freeware package, and applied them to a real-world case study. There are however many other ways in which these concepts can be applied. The same analytical approach and metrics can be applied to guide restoration to increase landscape connectivity, or to help to identify focal areas to halt the spread of undesired diseases, forest fires, or invasive species. Additional applications throughout the world are provided at **http://www.conefor.org/applications.html**, where more details and references are available. Additional applications include using *IIC* and *PC* in the analysis of pond or river networks, endangered species conservation plans, design of urban ecological networks, seed deposition patterns by frugivorous birds, and applications in combination with least-cost paths and circuit theory models. Other applications include assessments of directional (non-symmetrical) connectivity, evaluations of transnational protected area networks, mitigating barrier effects of transportation infrastructure, quantifying network vulnerability, evaluating long-distance spread over multiple generations, as well as assessments of the role of connectivity in influencing species diversity, distributions, colonization, or genetic diversity. The inspiration found in such applications, together with the understanding and practical skills that you have acquired throughout this chapter, should enable you to adapt and apply these approaches to other landscape management plans and research projects.

Acknowledgments Funding was provided by the Spanish Ministry of Science and Innovation through DECOFOR (AGL2009–07140) and GEFOUR (AGL2012-31099) projects. The capercaillie habitat data were provided by Instituto Catalán de Ornitología (ICO) with support from Generalitat de Catalunya (Departament d'Agricultura, Ramaderia, Pesca, Alimentació i Medi Natural) and CatalunyaCaixa (Obra Social). We thank all volunteers that made possible the collection of the information for the Catalan Breeding Bird Atlas.

REFERENCES AND RECOMMENDED READINGS[1]

Awade M, Boscolo D, Metzger JP (2012) Using binary and probabilistic habitat availability indices derived from graph theory to model bird occurrence in fragmented forests. Landsc Ecol 27:185–198

Bodin Ö, Saura S (2010) Ranking individual habitat patches as connectivity providers: integrating network analysis and patch removal experiments. Ecol Model 221:2393–2405

Bunn AG, Urban DL, Keitt TH (2000) Landscape connectivity: a conservation application of graph theory. J Environ Manage 59:265–278

*Calabrese JM, Fagan WF (2004) A comparison-shopper's guide to connectivity metrics. Front Ecol Environ 2:529–536. *Reviews and compares approaches for analyzing connectivity as well as their advantages and limitations, emphasizing practitioner needs. After evaluating the tradeoff between information content and data requirements, concludes that graph-based approaches may possess the greatest benefit to effort ratio for conservation problems requiring connectivity analysis at broad scales.*

Canut J, García-Ferré D, Afonso I (eds) (2011) *Manual de conservación y manejo del hábitat del urogallo pirenaico*. Serie de Manuales de Gestión de Especies Amenazadas. Ministerio Medio Ambiente y Medio Rural y Marino, Madrid

Crooks KR, Sanjayan M (eds) (2006) *Connectivity conservation*. Cambridge University Press, Cambridge

Estrada E, Bodin Ö (2008) Using network centrality measures to manage landscape connectivity. Ecol Appl 18:1810–1825

Estrada J, Pedrocchi V, Brotons L et al (eds) (2004) *Atles dels ocells nidificants de Catalunya 1999–2002*. Lynx Edicions, Barcelona

*Galpern P, Manseau M, Fall A (2011) Patch-based graphs of landscape connectivity: a guide to construction, analysis, and application for conservation. Biol Conserv 144:44–55. *An elegant, well-structured synthesis of different ways a landscape graph can be constructed to represent a habitat mosaic and various research questions that can be addressed using a graph-based approach.*

Hanski I, Ovaskainen O (2000) The metapopulation capacity of a fragmented landscape. Nature 404:755–758

Jordán F, Baldi A, Orci KM et al (2003) Characterizing the importance of habitat patches and corridors in maintaining the landscape connectivity of a *Pholidoptera transsylvanica* (Orthoptera) metapopulation. Landsc Ecol 18:83–92

Menoni E (1991) *Ecologie et dynamique des populations du grand tétras dans les Pyrénées*. Doctoral thesis. Université Paul Sabatier de Toulouse

Moilanen A, Nieminen M (2002) Simple connectivity measures in spatial ecology. Ecology 84:1131–1145

*Pascual-Hortal L, Saura S (2006) Comparison and development of new graph-based landscape connectivity indices: towards the priorization of habitat patches and corridors for conservation. Landsc Ecol 21:959–967. *First scrutiny of the behavior and performance of a wide set of graph metrics in the prioritization of habitat patches and links in reaction to spatial changes. It proposed that to adequately inform conservation decisions, intrapatch and interpatch connectivity should be integrated in a single measure. The analytical expression for such an index (IIC) is provided. Read this paper only after having completed the exercises in Parts 1 and 2.*

Pereira M, Segurado P, Neves N (2011) Using spatial network structure in landscape management and planning: a case study with pond turtles. Landsc Urban Plan 100:67–76

Rayfield B, Fortin MJ, Fall A (2011) Connectivity for conservation: a framework to classify network measures. Ecology 92:847–858

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Ricotta C, Stanisci A, Avena GC et al (2000) Quantifying the network connectivity of landscape mosaics: a graph-theoretical approach. Community Ecol 1:89–94

Rodríguez-Pérez J, García D, Martínez D (2014) Spatial networks of fleshy-fruited trees drive the flow of avian seed dispersal through a landscape. Funct Ecol 28:990–998

Saura S (2008) Evaluating forest landscape connectivity through Conefor Sensinode 2.2. In: Lafortezza R, Chen J, Sanesi G et al (eds) *Patterns and processes in forest landscapes: multiple use and sustainable management*. Springer, New York, pp 403–422

*Saura S, Pascual-Hortal L (2007) A new habitat availability index to integrate connectivity in landscape conservation planning: comparison with existing indices and application to a case study. Landsc Urban Plan 83:91–103. *This paper extended the developments in Pascual-Hortal and Saura (2006) to networks with weighted links and proposed a new metric (PC) based on a probabilistic definition and connection model. Advantages of PC are discussed with an application for a forest bird species in NE Spain.*

*Saura S, Rubio L (2010) A common currency for the different ways in which patches and links can contribute to habitat availability and connectivity in the landscape. Ecography 33:523–537. *Presents an integrated analytical framework that quantifies the different roles played by habitat patches and links as connectivity providers. This approach helps avoid arbitrary weighting of connectivity considerations when taking conservation decisions.*

Saura S, Estreguil C, Mouton C et al (2011) Network analysis to assess landscape connectivity trends: application to European forests (1990–2000). Ecol Indic 11:407–416

Taylor PD, Fahrig L, Henein K et al (1993) Connectivity is a vital element of landscape structure. Oikos 68:571–573

*Tischendorf L, Fahrig L (2000) How should we measure landscape connectivity? Landsc Ecol 15:633–641. *The first study to note problems with certain connectivity metrics that indicate higher connectivity in more fragmented landscapes and zero connectivity in any landscape containing just one habitat patch even if that patch fills all the landscape.*

*Urban DL, Keitt T (2001) Landscape connectivity: a graph-theoretic perspective. Ecology 82:1205–1218. *Seminal classic paper on the use of a graph-theoretical approach for the study of landscape connectivity. It inspired and stimulated much of the research and subsequent developments that came later in the field.*

Urban DL, Minor ES, Treml EA et al (2009) Graph models of habitat mosaics. Ecol Lett 12:260–273

Chapter 15
Linking Landscapes and Metacommunities

Joseph R. Bennett* and Benjamin Gilbert*

OBJECTIVES

Landscape ecology is the study of interactions between spatial landscape patterns and ecological processes, typically examining real landscapes at broad spatial scales. Metacommunity ecology focuses more specifically on how spatial processes alter species interactions and typically involves a localized spatial extent and more abstracted spatial landscapes (Bolker, 2004). These disciplines have evolved somewhat independently, despite a shared interest in how organisms respond to and interact with spatial phenomena. In this chapter, we combine perspectives from both disciplines using a suite of multivariate spatial statistical techniques designed to help understand the relative importance of abiotic factors (such as climatic gradients, geologic features, and resource availability) and biotic factors (such as predator territoriality and seed dispersal) in determining the abundances of species in communities. To illustrate these techniques, we use a well-known dataset of tropical trees. This lab will enable students to:

1. Utilize semivariograms to examine and understand spatially autocorrelation in species and environmental data;
2. Model distributions of species and communities along environmental gradients using redundancy analysis;

*Both authors contributed equally to this work.

J.R. Bennett (✉)
Department of Biology, Carleton University, Ottawa, ON, Canada
e-mail: joseph.bennett@carleton.ca

B. Gilbert
Department of Ecology and Evolutionary Biology, University of Toronto,
Toronto, ON, Canada

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_15

3. Gain familiarity with multivariate spatial regression as well as the associated tools of trend-surface analysis and eigenvector analysis; and
4. Use joint modeling of abiotic and biotic factors to determine their influences on species distributions in a metacommunity.

The lab is divided into four parts, which build sequentially. In Part 1, we explore the concepts of spatial processes in metacommunities, linking the perspectives of both community and landscape ecology using tree species information from a long-term vegetation study in Barro Colorado Island, Panama. In Part 2, we introduce how communities are measured in order to represent many species at once. We then show how multiple regression can be applied to model communities along abiotic environmental gradients. In Part 3, we introduce two types of spatial methods, called trend-surface analysis and eigenvector analysis, and apply these to the tropical tree species. Finally, in Part 4 we compare the abiotic and spatial determinants of metacommunity structure. We also use our results (from Part 3) to illustrate the advantages and disadvantages of each method, as well as important caveats in the spatial analysis of communities.

This lab aims to make several spatial statistical techniques accessible and understandable to those without extensive training in statistics. However, familiarity with basic regression, especially multiple regression, is extremely helpful. Familiarity with the concepts covered in Chapter 5 (basic semivariograms) is also assumed. This lab assumes prior familiarity with R (including installation procedures), and thus is not intended as the first exposure to the R environment for instructors or students. That being noted, the provided R code (available on the book web site) is very well documented and numbered according the various Figures, Exercises, and Steps in the lab. The lab requires a computer running R version 2.12.0 or higher and access to the datasets provided with this chapter available from the book web site.

INTRODUCTION

Landscape ecology and metacommunity ecology offer different but complementary world views and approaches. **Metacommunity ecology** generally considers entire communities of species and their interactions in a quantitative and spatially explicit (or spatially implicit) way. Metacommunity theory is important for understanding how factors such as habitat suitability and species-specific dispersal abilities impact community-level responses such as alpha and beta diversity. Landscape ecology provides tools and approaches for understanding how the structure of the landscape alters diversity, thus offering a positive feedback between the disciplines. In this chapter, we combine the benefits of both approaches by using maps of species and the environment in conjunction with multivariate analyses of community composition to model relationships among species, underlying environmental conditions, and spatial locations.

Part 1. Biotic and Abiotic Influences in a Metacommunity

Ecologists know that all species are limited in part by resources, predators, competitors, and diseases, and these items can be collectively considered to influence a species' **niche**. For our purposes, a working definition of a species' niche is *the environmental conditions that allow a species to persist*, with "environmental conditions" referring to the collective factors that influence reproduction and mortality. To understand how biotic and abiotic components of a species' niche influence its distribution, we begin by considering one tree species from a 50 ha plot on Barro Colorado Island, Panama (BCI). *Trichilia tuberculata* is a new world tree species in the mahogany family and is relatively common in this forest. A map of its distribution (Figure 15.1, Panel A) shows that *Trichilia* appears to have a patchy distribution, with some areas having high densities (light shading) and other areas having a low density (dark areas).

The distribution of *Trichilia* may reflect areas of favorable environmental conditions. Plant ecologists have hypothesized that these conditions are mainly based on available resources, such as nitrogen, an important nutrient for plants. Therefore,

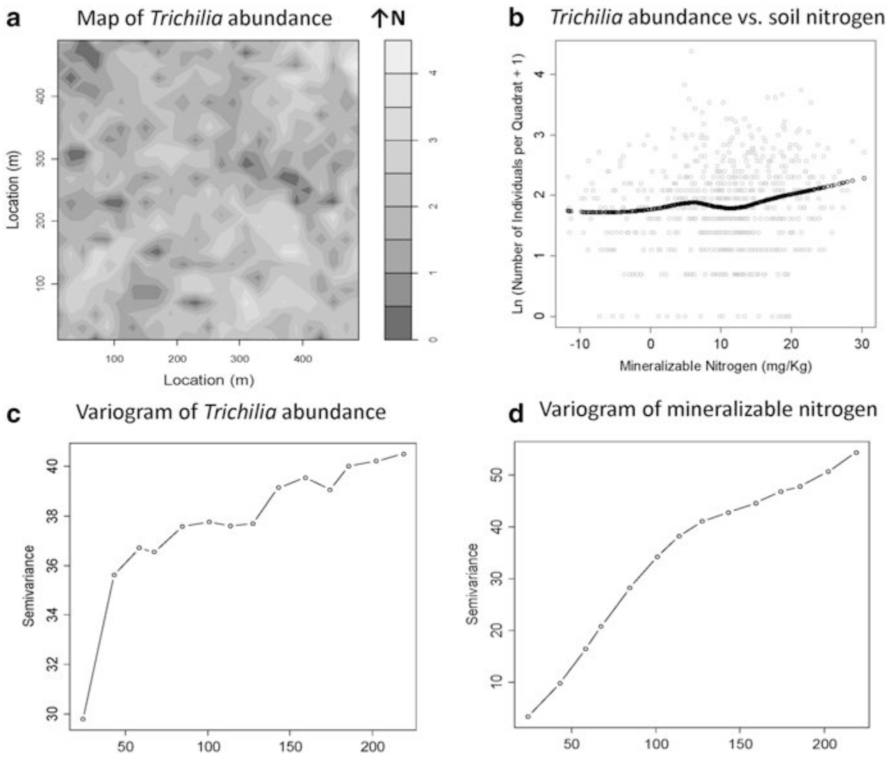


Figure 15.1 Important spatial relationships for *Trichilia tuberculata Panel a*. Abundance in the 500×500 m sampling plot. *Panel b*. Relationship to mineralizable nitrogen. *Panel c*. Spatial clustering of *Trichilia. Panel d*. Mineralizable nitrogen

one hypothesis is that the distribution of *Trichilia* is caused by available nitrogen. By plotting the abundance of *Trichilia* against mineralizable nitrogen (Figure 15.1, Panel B), we make a preliminary exploration of this hypothesis. Modeling the relationship between *Trichilia* and mineralizable nitrogen using regression is a more quantitative way of testing this hypothesis.

Q1 How would you describe the distribution of *Trichilia* in the study area?

Q2 How would you describe the relationship between *Trichilia* and mineralizable N?

Q3 What are some other potential causes for *Trichilia* abundance patterns on the landscape in addition to nitrogen availability?

There is an important caveat to simply modeling species and environmental relationships without considering space. Species distributions can also be strongly influenced by spatial processes such as inter- or intraspecific competition from neighbors, dispersal, and foraging patterns of predators or herbivores. For example, windblown seeds, or even seeds such as those of *Trichilia* that are dispersed by large birds and mammals, tend to fall close to their parent. Relatively few seeds travel long distances (Muller-Landau et al. 2008) often creating clumpy patterns of species distributions. Environmental characteristics also often show similar spatial patterns (areas in close proximity share similar characteristics), making inferential tests based on these patterns difficult.

Variograms (explained more fully in Chapter 5) help illustrate this problem. In Figure 15.1, Panels C and D show how predictable *Trichilia* abundances and mineralizable nitrogen are at different lag distances. Recall that low variation (small values on the *Y*-axis) indicate that sample locations closer together tend to have similar values. Because this variation increases with distance between sample points, we deduce that sample locations that are closer together tend to have more similar values of both mineralizable nitrogen and *Trichilia* abundance. Since the pattern of *Trichilia* based on dispersal alone would also be clumped, we cannot be sure of the extent to which mineralizable nitrogen influences *Trichilia* distributions, independent of the effect of dispersal. In the case of *Trichilia*, or any single species, mixed models are well suited to account for spatial autocorrelation that is separate from the effect of abiotic variables (Zuur et al. 2009). However, when considering the responses of an entire community, a more common approach is to use multivariate partitioning methods, which we explore next.

Part 2. RDA for Modeling Abiotic Gradients

To illustrate how multivariate partitioning methods work, we will first model the response of a community to abiotic gradients only. When modeling the entire community, we cannot simply do statistical tests for each species because the large number of tests would inflate the Type 1 error. Testing a large number of environmental variables against each species would make the problem even worse. Therefore, an analytical method that models how the abundances of individual species change along the abiotic gradients and can test this for the entire community at once using a global statistic is extremely useful.

Ordination is a way of projecting relationships among multiple samples, predictor and response variables into mathematical space in more than one dimension, where the first axis represents the greatest amount of variability and each subsequent axis is uncorrelated with the previous ones. In a nutshell, ordination algorithms find a configuration of samples (or species) in mathematical space that best represents differences among them. Mathematically, there are a number of ways in which ordination algorithms can work (see Jongman et al. (1995) and Lepš and Šmilauer (2003) for details). We use two methods we use in this chapter: **principal component analysis (PCA) and principal coordinate analysis (PCoA)** which use an analytic solution to compute the mathematical distances. The resulting **site scores** represent positions of sites relative to each other in mathematical space and will be similar for samples with similar species abundances. Likewise, species commonly found together will have similar **species scores**.

Constrained ordination is simply ordination performed after the response variables (site or species scores) are obtained from a regression on the predictor variables. **Redundancy analysis (RDA)** is a constrained ordination technique that explicitly models linear relationships among multiple predictor variables and uses a randomization approach that statistically tests the strength of these relationships. In our case, we will create an RDA where species scores are predicted by regression on the environmental (or spatial) variables, and the new site scores are those generated by these predictions. The ordination results now correspond to the greatest variability in the dataset *that can be explained by the measured variables*.

EXERCISE 1: Data Preparation

The first step involves reading in the data and creating the two matrices that we need. The data provided (see R code available from the web site) have been organized for multivariate analyses:

- Rows represent sample sites, and columns hold information on species abundances or predictor variables.
- By looking at the subsets of the data matrices provided, you can get an idea of the data characteristics. For example, sample locations are spaced along a 20 m grid, with observations at *x* or *y* locations of 10, 30, 50, and so on.
- The 25 species included in the dataset are represented by six letter codes, with the full species names for these codes given at https://repository.si.edu/handle/10088/20925. For example, *Trichilia* is represented by the code TRI2TU.
- Finally, you can see that 13 soil variables were collected. Details on these soil variables are given at: http://biogeodb.stri.si.edu/bioinformatics/bci_soil_map/.

Since RDA is essentially an extension of linear regression, it is important to note that linear relationships with independent variables are tested; thus, variables may need to be transformed to achieve linearity. In addition, in order for the ordination portion of the RDA to work correctly, species data often need to be transformed to avoid creating a "horseshoe effect" (see Lepš and Šmilauer 2003 for details). For this chapter, we use a transformation called the Hellinger transformation that has been shown to work for a wide range of communities (Legendre and Gallagher 2001). Scientists who are interested in other transformations should consult Legendre and Legendre (1998).

Each environmental variable needs to be transformed if it is related to a nonlinear change in species scores. The choice of transformation is extremely important for inferring the relationship between species and the predictor variables (Gilbert and Bennett 2010), and requires considerable thought. Species or community responses to environmental variables are often nonlinear, so for the soils data we need to build a dataset that includes transformations for each variable that can model nonlinear response. From our initial analysis of the data, using visual inspection of trends, spline graphs, and polynomial regressions, we opted to use a third-order polynomial transformation of each variable. We chose polynomials in this case because they are relatively simple and flexible transformations and opted for the smallest polynomial that appeared to fit the data sufficiently well.

EXERCISE 2: Species-Environment RDA

Now that we have the two data matrices, we can perform the species-environment RDA. The significance of the relationship between the species matrix and the environmental matrix is determined with a Monte Carlo permutation procedure that tests whether the association between the matrices is stronger than expected by chance. This is done by comparing the test statistic from the true data to the test statistic that would be generated if the data were randomly assigned to sample plots. If the test statistic from the true data is higher than 95% of the random ones, the *P*-value is 0.05; if it's higher than 99% of the random ones, then the *P*-value is 0.01, etc.

Selection of significant variables in ordination has similar problems with model selection in regular multiple regression. Here, we use a forward selection method that attempts to control for model selection errors, using an adjusted R^2 that accounts for the number of variables used (Blanchet et al. 2008). The method makes sure the adjusted R^2 from all of the selected variables never exceeds the adjusted R^2 for the full model that includes all of the variables. In particular, we first do an RDA of the entire data soil matrix versus the species matrix to get the output in Table 15.1. The *p*-value on the right (0.005) is the value from the permutation.

Q4 What does significance in the full model mean? What would lack of significance mean?

Because this first test of the full model is significant, we go on to use a constrained forward selection, which is analogous to forward stepwise multiple regression.

Table 15.1 Result from RDA that includes all soil variables as predictors

Permutation test for rda under reduced model					
Model: rda(X=species.matrix, Y=soil.matrix)					
	Df	Var	F	N.Perm	Pr(>F)
Model	39	0.058567	6.5182	199	0.005**
Residual	585	0.134776			

Table 15.2 Result from RDA that includes only significant soil variables from forward selection

Model: rda(X=species.matrix, Y=soil.selected)					
	Df	Var	F	N.Perm	Pr(>F)
Model	32	0.05659	7.6554	199	0.005**
Residual	592	0.13675			

Q5 What single soil variable explained the most variation? (*HINT:* type “fwd.sel” to see the forward selection results).

It is important to note that the number of significant variables can change from one analysis to another because the RDA uses a randomization procedure; a variable that is marginally significant in one run (say, p value of 0.047–0.053) could change to insignificant in another run, or vice versa. This can be resolved by increasing the number of randomizations, but also reflects the problem of choosing an arbitrary significance level.

We can now do the analysis with the selected variables only, obtaining the output found in Table 15.2. We can see from this output that there is a significant influence of the selected environmental variables on community composition.

EXERCISE 3: Exploring the Results—Variation Explained

An important step in constrained ordination is to determine the amount of variation explained (R^2) in the community by the included variables. The simple R^2 is calculated by dividing the variation accounted for by the environmental variables (termed the “constrained inertia”) by the total variation in the species dataset (the “total inertia”). Our uncorrected R^2 is 0.29, meaning that 29% of the variation in species distributions is explained by the selected soil variables. However, a correction is necessary to account of the number of independent variables tested, just as an adjusted R^2 is used in multiple regression. In the corrected model, the explained variation by environmental variables is 0.25 or 25%. We can also find the order of variable selection, from the first (the variable that on its own explains the most variation) to the last significant variable.

Table 15.3 The first 5 variables selected in the forward selection

	Variables	Order	R^2	R^2Cum	AdjR^2Cum	*F*	Pval
1	P1	25	0.084913	0.084913	0.083444	57.80954	0.001
2	Al1	1	0.030145	0.115058	0.112212	21.18773	0.001
3	K1	16	0.019986	0.135043	0.130865	14.34904	0.001
4	Cu1	10	0.013505	0.148548	0.143055	9.833824	0.001
5	N_min1	34	0.013006	0.161554	0.154782	9.601939	0.001

Using the R code in the Appendix, we'll look at the first five variables (Table 15.3). You'll notice in the table that the cumulative R^2 (that includes the variable and all previously selected ones) as well as adjusted cumulative R^2 are presented. Once the last variable is selected, the cumulative adjusted R^2 is equal to the adjusted R^2.

EXERCISE 4: Exploring the Results Graphically

We can also explore our results graphically by using an ordination plot (Figure 15.2), which plots species and site scores on the axes described above. In Figure 15.2, we graph only the five most important soil variables to simplify the presentation, but the R code in the online appendix can be used to plot all variables. The ordination plot includes sample sites, species, and independent variables on the first two ordination axes (Figure 15.2). There are different ways to scale these plots (see R help for CCA. plot and also Legendre and Legendre 1998). Here, we have used "species" scaling, which can be interpreted as follows. The angles between the arrows that point to each environmental variable represent the correlations between those variables. For example, potassium (K1) and aluminum (A1) are strongly negatively correlated, as they point in almost opposite directions. Similarly, if a line were drawn from the plot center to each species, the angles between each pair of lines would represent the correlation between species abundances. For example, abundances of the species labeled DRYPST and HIRTTR are also strongly correlated and are higher at high potassium and low aluminum levels. The length of the arrows along the ordination axes indicate the strength of their relationship to each axis, with phosphorus (P1) being more closely related to the first axis than mineralizable N (N_min1), for example.

Q6 Examine the distribution of *Trichilia* (TRI2TU) along the mineralizable nitrogen (N_min1) gradient in Figure 15.1. Where are TRI2TU and N_min1 in the ordination figure? What does the location of the arrow head and the species tell you?

Q7 Using the same approach as in the question above, name a species that should occur at higher abundances when there are large amounts of copper (Cu1) present. Use R to graph the abundance of this species relative to copper to see if your prediction is correct.

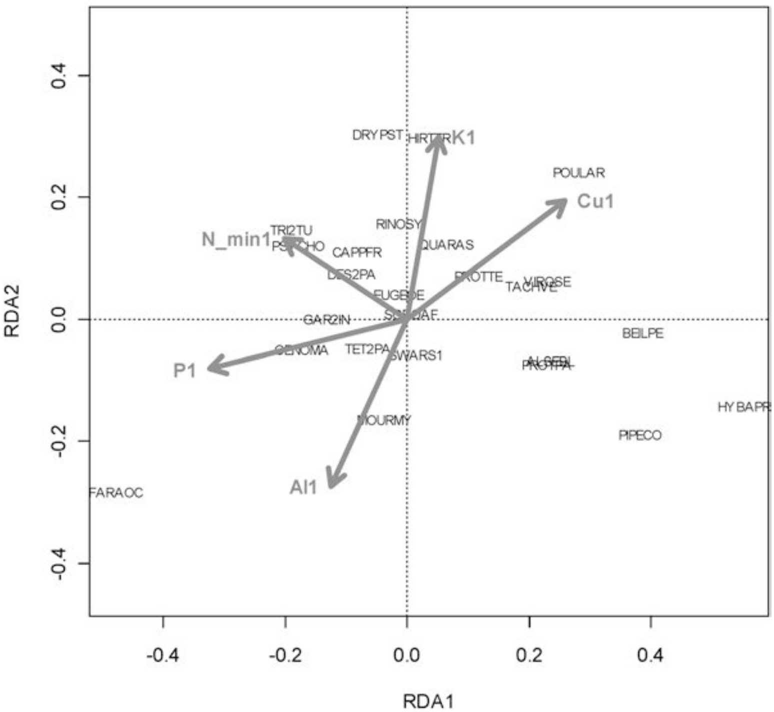


Figure 15.2 Biplot showing species (six letter codes) and soil resources (vectors with letters at end)

Q8 Species can also be negatively correlated to specific resources. Find the species VIROSE. How do you think this species' abundance will change with soil phosphorus (P1)? Use R to graph the abundance of this species relative to copper.

Part 3. Accounting for Space in Community Analyses

Quantifying spatial changes in distributions of species and communities, and then attributing these changes to specific causes, can be quite difficult. As mentioned previously, spatially correlated environmental factors are not the only possible causes of the spatial distributions of organisms. Biotic factors such as dispersal, inter- and intraspecific competition, predation, and disease can also create spatial patterns which may be distinct. For example, intraspecific competition promotes negative spatial autocorrelation among species and communities while dispersal limitation promotes positive spatial autocorrelation. However, some patterns may not be distinct; for example, disease may promote negative spatial autocorrelation (similar to intraspecific competition) since close proximity of individuals may increase risk of infection. The refinement of techniques to attribute spatial change to specific causes is an active area of ecological research (Diniz-Filho et al. 2003; Gilbert and Bennett 2010).

For the multivariate statistics that we are presenting, the common method for modeling spatial patterns is to use regression-based approaches to model the abundances of species or changes in community composition as a function of location. While the many methods used for multivariate spatial analysis are well beyond the scope of this chapter, here we present two commonly used methods: **trend-surface analysis** and **Moran's eigenvector mapping (MEM)**. Both approaches are used within the RDA framework and can help us ask what type of spatial patterns might result from a given process and then test which spatial variables can model these patterns. We outline each approach below, and explain the reasoning behind using one or the other in different situations.

EXERCISE 5: Polynomial Trend-Surface Analysis

Using polynomials to model nonlinear spatial patterns in communities is very similar to using them to model nonlinear changes due to environmental factors. Instead of polynomials of environmental variables, we generate polynomials of the centered X and Y coordinates of our spatial data. In particular, if all plots have an x and y location (in meters or UTMs), the trend-surface is a polynomial function of those x and y locations so that predictor variables are: X, Y, X^2, Y^2, XY, X^2Y, XY^2, X^3, Y^3. We can then use these variables in RDA, in a similar forward-selection process to that used for the environmental polynomials.

Q9 Use the code provided in the Appendix to conduct a forward selection of the spatial polynomials. What variables were selected in the forward-selection process? How does the number of spatial variables compare with the number of environmental variables that were selected in the last section?

Q10 How much variation was explained by the spatial variables in the analyses above? How does that compare with the variation explained by the initial analyses of environmental variables?

Q11 Do you think it is valid to add the variation explained by environmental and spatial variables, as analyzed so far, to get the full variation explained by both? Why or why not? (*HINT:* reexamine Figure 15.1).

EXERCISE 6: Moran's Eigenvector Maps

The second approach that we demonstrate, called **Moran's Eigenvector Maps (MEM)**, is an adaptation of an earlier method, called principal coordinates of neighbor matrices (PCNM) that is still frequently used. The simplest explanation of the MEM approach is that it uses a series of waves to model a spatial pattern to fit complex spatial relationships among samples. This approach is similar to spectral analysis, with a more detailed explanation found in Dray et al. (2006) and Peres-Neto and Legendre (2010). We will use a step-by-step process that generates MEM. This process can take a long time because some of the randomizations are computer intensive. If time is short, your instructor may choose to skip steps 1–5 (see "Shortcut" in appendix R code), or assign them as homework.

The MEM technique has several components:

Step 1: Create a distance matrix. This matrix should represent the spatial distances among all sample sites.

Step 2: Simplify the distance matrix. Ensure that all distances that are greater than a critical value are all assigned the same (large) value.

Step 3: Conduct a principal coordinates analysis (PCOORD) on the simplified distance matrix. PCOORD is another ordination technique that represents the distances among samples in different dimensions. These dimensions are the same as the axes of an ordination plot, as explained above, and the axes are called **eigenvectors**. The technique can be used to represent spatial relationships among samples that are not readily apparent to us, for example, wave patterns of repeating spatial clusters, and other complex, nonlinear relationships.

Step 4: Test the eigenvectors for significant spatial autocorrelation. Some of the eigenvectors are not spatially autocorrelated thus including them is equivalent to including useless predictor variables in a regression. For the purposes of this chapter, we consider only significant positive autocorrelation. We retain only the axes that indicate strong positive spatial structure by only considering those values of Moran's I that are positive and statistically significant. (*NOTE:* Depending on the size of the dataset and available computing power, the randomizations may require a lot of time (up to an hour for this particular dataset).

Q12 The "dim" function gives the dimensions of the dataset (i.e., the number of rows and columns). How many eigenvectors show positive and significant autocorrelation? How does this compare to the number of predictors in the trend-surface analysis and in the environmental analysis?

Step 5: Remove linear trends. The vectors created using MEM are good at detecting nonlinear patterns at a finer scale than the trend-surface analysis. However, they are not good at detecting linear trends, which can cause analysis problems. We therefore do a separate analysis of the linear spatial trends over the sample area and include this in the total spatial signal. In our case, we have already included linear x and y trends in the trend-surface analysis, and we know that they are significant. So, we first remove the effects of these predictors on species distributions, and then test the significance of the MEM axes on the residuals. We use the usual forward selection procedure, plus an additional selection criterion that is designed to help with over-fitting problems that have been observed with eigenvectors like MEMs (see Gilbert and Bennett 2010).

Table 15.4 Results from RDA that include only significant MEMs

Permutation test for rda under reduced model					
Model: rda(*X*=species.matrix, *Y*=MEM.selected)					
	Df	Var	*F*	N.Perm	Pr(>*F*)
Model	61	0.084236	7.1256	199	0.005**
Residual	563	0.109107			

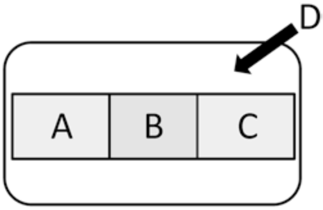


Figure 15.3 Representation of variation explained in variation partitioning. The entire area ($A+B+C+D$) represents the total variation in the dataset, and the *grey rectangles* represent variation explained by each component. Segment *A* represents the component of variation explained solely by environmental variables, while segment *C* represents the component solely explained by spatial variables, and segment *B* represents variation jointly explained by spatial and environmental variables. *D* represents unexplained variation

Step 6: Perform RDA with Species and MEM. We now do the analysis on the selected variables and determine the variation explained, as with the previous examples (Table 15.4).

Q13 How many spatial variables were in this analysis? What is the corrected R^2 for the spatial signal using MEM? Is this very different from the uncorrected R^2? Why is that? How does this compare with the corrected R^2 of the spatial polynomials?

Part 4. Variation Partitioning

As we have noted, the spatial determinants of species composition may be environmentally correlated, which can lead to problems if we draw conclusions from environmental analyses without considering spatial signals, or conversely, from spatial analyses without considering environmental factors. We therefore need a technique that will allow us to divide the signal of community change into separate environmental and spatial components, as well as the component that is shared between them. **Variation partitioning** (Borcard et al. 1992; Figure 15.3) allows different independent components of variation to be allocated. The technique works in steps using simple algebra and requires results from three separate constrained ordinations. Because you will need information from prior Exercise 6, we continue numbering our steps from the previous exercise.

Step 7: Record the Variation Explained by the RDA with Environmental Variables. The corrected R^2 result from Exercise 3 (RDA using forward-selected environmental variables) gives this portion of the variation explained. In Figure 15.3, the variation explained from this ordination is represented by $A+B$.

Step 8: Record the Variation Explained by the RDA with Spatial Variables. The corrected R^2 result from either the trend-surface analysis (Exercise 5) OR the RDA with MEM (Exercise 6, Step 6) gives this portion of variation explained. In Figure 15.3, the variation explained from this ordination is represented by $B+C$.

Step 9: Record the Variation Explained by an RDA with all Significant Spatial and Environmental Variables. This involves creating a dataset that includes the selected environmental variables from Exercise 2 and spatial variables from Exercise 5 or Exercise 6, Step 6. In Figure 15.3, the variation explained form this ordination is represented by $A+B+C$.

EXERCISE 7: Partition the Variation Explained

Variation explained solely by environmental variables is represented by component A; variation explained solely by spatial variables by component C; and shared variation by component B. All of these can be derived using algebra. Depending on the research question, ecologists may be interested in any combination of the components above. The shared spatial, shared environmental, and total explained variation are easily attained using the ordinations described above.

In order to get component A (the independent environmental signal), take the variation explained by all selected variables ($A+B+C$), and subtract the variation

Table 15.5 Results from partitioning spatial (trend-surface) and soil variables

Explanatory tables:						
$X1$: soil.selected						
$X2$: t.s.selected						
No. of explanatory tables: 2						
Total variation (SS): 120.65						
Variance: 0.19334						
No. of observations: 625						
Partition table:						
			Df	R^2	AdjR^2	Testable
$[a+b]$	=	$X1$	32	0.29269	0.25446	TRUE
$[b+c]$	=	$X2$	9	0.20996	0.1984	TRUE
$[a+b+c]$	=	$X1+X2$	41	0.34799	0.30214	TRUE
Individual Fractions						
$[a]$	=	$X1 \mid X2$	32		0.10374	TRUE
$[b]$			0		0.15072	FALSE
$[c]$	=	$X2 \mid X1$	9		0.04768	TRUE
$[d]$	=	Residuals			0.69786	FALSE

explained by the spatial and shared component: $(A+B+C)-(B+C)=A$. This is accomplished by subtracting the variation explained in Step 9 from that in Step 8. Similarly, to get component C (the independent, or "pure" spatial signal), take the variation explained by all selected variables ($A+B+C$, Step 9), and subtract the variation explained by the environmental and shared component ($A+B$, Step 7).

The partitioning of these components can also be calculated directly with a specialized function in R. For example, Table 15.5 gives the results from partitioning the environmental and trend-surface components. In this output, the fraction "[$a+b$]" or variable "$X1$", refers to the environmental plus shared variation explained ($A+B$ in the figure above). Likewise, "[$b+c$]" or variable "$X2$" refers to the spatial plus shared variation ($B+C$ in the figure above) while individual fractions refer to the "pure" environmental or spatial variation explained ("[a]" and "[c]", respectively), and the shared variation "[b]". Component "[d]", the residuals, refers to the variation that is NOT explained by the ordination.

Q14 What is the total (adjusted) variation explained by the ordination? What signal appears greater, space, or environment? How does the variation explained by either fraction compare to the shared variation, and what does this mean, in terms of interpretation of results?

Q15 Partition the variation explained by the MEM spatial predictors and the soil variables. How do these results differ from the trend-surface results above? Why do you think these techniques give different results?

The differences in signals using the trend-surface and MEM techniques illustrate an important aspect of spatial analysis: the method of representing the spatial configuration of samples or communities on a landscape can have profound influences on the results. Recall that the polynomial trend-surface model the spatial signal as curves that can be drawn on a map while the MEMs model space as a series of waves with different periodicities. The latter technique is much more flexible in terms of what is considered "spatial," making it difficult to attribute specific biological causes to the signal. Given these differences, it is important to keep in mind one's original research question when choosing a sampling configuration and an analytical technique. Similarly, the sampling design can also influence the outcome of these analyses: sample plots that are surveyed with spatial lags of kilometers are unlikely to show the same patterns as contiguous plots (see Fortin and Dale 2005 for details on this issue). Below, we clarify how these results can be interpreted and incorporated into landscape level models.

EXERCISE 8: Testing the Independent Components

Although we have already tested the significance of shared components ($A+B$, for example), we have not yet tested the independent components. The significance of the independent environment component (A) can only be tested by including the spatial component ($B+C$) as a covariate. Similarly, the independent spatial component (C) is tested by including the environmental component ($A+B$) as a covariate.

Table 15.6 Exploring the spatial structure of important soil explanatory variables

Variable	Component $A+B$ *	Component A	Component B	Percent spatial $(100*B/(A+B))$
P1	0.085	0.014	0.071	83.079
Al1	0.043	0.006	0.037	85.586
K1	0.033	0.008	0.026	76.908
Cu1	0.067	0.010	0.056	84.753
N_min1	0.042	0.005	0.037	88.407

These numbers differ from Table 15.3 because the variables are tested individually, without considering the correlations with other variables

Using the R code in the Appendix, we can see from testing both of these fractions that when using the spatial trend-surface predictors, each component is significant.

EXERCISE 9: Interpreting and Applying the Partitioning Results

Many studies in spatial ecology have stopped after step 10 and reported the variation partitioning results as evidence for environmental or biotic processes. This approach must be taken with caution. It has been shown that partitioning results can give a good indication of the processes at work; however, they are sensitive to sampling design, unmeasured (especially spatially structured) variables, the spatial model used (MEM or trend-surface), and whether correct transformations of environmental variables were used (Legendre and Legendre 1998; Gilbert and Bennett 2010).

Despite these reservations about strictly interpreting partitioning results, the separate components do contain information that is extremely useful to landscape ecologists. In particular, the component B (Figure 15.3, Table 15.5) represents the spatially structured environment that predicts species distributions. Although component B cannot be tested statistically, it can be compared with other components. A large amount of explained variation in this component indicates that some portion of species' spatial distributions are explained by spatially structured environmental variables. One method of better understanding this component is to examine how much of a variable's predictive power is spatially structured.

We do this for the five most significant variables (Table 15.6). We can see that for the tropical forest studied, the most important predictors are strongly spatially structured (Table 15.6).

The spatial structure of these variables, determined using the methods presented earlier and elsewhere in this book, can be employed in landscape models that explicitly consider how spatial processes may work in conjunction with responses to abiotic variables. The important benefit of the partitioning approach is that it has allowed us to identify the abiotic variables that are significant predictors of species distributions, and also provide an indication of how well spatially explicit landscape models could capture the effects of these predictors.

CONCLUSIONS

When contrasting spatial and environmental signals, it is important to realize that the two signals are almost always somehow intertwined. Figure 15.1 illustrates such a relationship in the BCI dataset. In fact, nearly all environmental variables have a spatial component: hard environmental boundaries such as sheer cliffs are not nearly as common as more gradual ecotones. The analyses we presented are useful tools for understanding the spatial structuring of communities due to both biotic and abiotic influences. If appropriate variables and sampling techniques are employed, these analyses can be used to gain an understanding of how abiotic and biotic influences act independently, and whether one influence tends to overshadow another. Partitioning results also offer a unique opportunity for landscape ecologists to identify variables that both structure communities and that are themselves spatially structured. Incorporating these results into landscape models can allow for quantitative estimates about the relevance of landscape models for predicting species distributions.

SYNTHESIS

Q16 Using the BCI dataset, design a research question that uses both spatial and environmental components. What variables and analyses would you test with this question? What are some potential unknowns and limitations in your analyses?

Q17 Consider a new and different dataset (either a potential hypothetical dataset or data from your own research) and answer Q16. Be sure to explain the type of data you might examine.

REFERENCES AND RECOMMENDED READINGS[1]

Blanchet FG, Legendre P, Borcard D (2008) Forward selection of explanatory variables. Ecology 89:2623–2632

Bolker BM (2004) Continuous-space models for population dynamics. In: Hanski I, Gaggioti OE (eds) *Ecology, genetics, and evolution of metapopulations*. Elsevier Science, San Diego, pp 45–69, Chapter 3

*Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. Ecology 73:1045–1055. *A classic paper outlining some of the approaches covered in this lab in more detail.*

*Borcard D, Gillet F, Legendre P (2011) *Numerical ecology with R*. Springer, New York. *A useful reference for a variety of analyses in R.*

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecol Model 196:483–493
Fortin M-J, Dale MRT (2005) *Spatial analysis: a guide for ecologists*. Cambridge University Press, Cambridge, UK
Gilbert B, Bennett JR (2010) Partitioning variation in ecological communities: do the numbers add up? J Appl Ecol 47:1071–1082
Hanski I (1991) Metapopulation dynamics: brief history and conceptual domain. Biol J Linn Soc 42:3–16
Jongman RHG, Ter Braak CJF, Van Tongeren OFR (1995) *Data analysis in community and landscape ecology*. Cambridge University Press, Cambridge, UK
Legendre P, Gallagher ED (2001) Ecologically meaningful transformations for ordination of species data. Oecologia 129:271–280
Legendre P, Legendre L (1998) *Numerical ecology*, 2nd edn. Elsevier Science, Amsterdam
*Lepš J, Šmilauer P (2003) *Multivariate Analysis of Ecological Data using CANOCO*. Cambridge University Press, Cambridge. *Chapters 1, 3 and 5 provide an excellent, readable introduction into basic ordination techniques.*
Levins R (1970) Extinction. In: Gesternhaber M (ed) *Some mathematical problems in biology*. American Mathematical Society, Providence, Rhode Island, pp 77–107
Macarthur RH, Wilson EO (1967) *The theory of island biogeography*. Princeton University Press, Princeton, NJ
Muller-Landau HC, Wright J, Calderon O, Condit RS, Hubbell SP (2008) Interspecific variation in primary seed dispersal in a tropical forest. J Ecol 96:653–667
Peres-Neto PR, Legendre P (2010) Estimating and controlling for spatial structure in the study of ecological communities. Glob Ecol Biogeogr 19:174–184
Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) *Mixed effects models and extensions in ecology with R*. Springer, New York

Part V
Ecosystem Processes and Feedbacks in Social–Ecological Landscapes

With expanding interest in social-ecological systems comes a new set of challenges. Rather than considering only biotic or environmental factors, the interactions of people and nature become paramount. These exercises explore fundamental concepts of connectivity and heterogeneity from the perspective of social-ecological landscapes, both in terrestrial and marine systems. Chapter 16 uses a straightforward spatial modelling approach in Excel to explore ways to incorporate landscape heterogeneity into ecosystem processes and services. The lab includes a well-loved "build-your-own-adventure" framework suitable for fun group projects. Chapter 17 also takes a user-friendly approach to account for heterogeneity in tropical landscapes managed for their carbon storage potential. Using the lens of carbon accounting, the role of spatial heterogeneity at different scales is assessed, exemplifying contemporary challenges of ecosystem service management at landscape scales. Chapter 18 explores spatial resilience and regime shifts in a coral reef social-ecological landscape and demonstrates the universality of spatial resilience principles (such as feedbacks) in diverse landscape (or seascape) settings. Chapter 19 examines tradeoffs among ecosystem services in an agricultural setting using a realistic modelling environment via web interface. Lastly, Chapter 20 offers a challenging perspective on social network connectivity using another marine-based example. Building on network themes from prior modules, social network connectivity is used to examine the flow of information related to fishing practices across a heterogeneous marine landscape. Taken together, this suite of exercises demonstrates the myriad ways in which landscape principles and tools are relevant to sustainability challenges in social-ecological landscapes throughout the world.

Chapter 16
Modeling Spatial Dynamics of Ecosystem Processes and Services

Sarah E. Gergel and Tara Reed

OBJECTIVES

Understanding and predicting rates of ecosystem processes (e.g., soil erosion, nutrient flux) across large heterogeneous landscapes is an enduring challenge in ecosystem and landscape ecology and underpins the knowledge base for managing ecosystem services. Many current problems in ecosystem services management (e.g., maintenance of water quality and reduction of soil erosion) occur over broad spatial scales and across ecosystem boundaries and thus are influenced by landscape pattern (Syrbe and Walz 2012). When scaling up, ecosystem ecologists and watershed hydrologists have often used fine-scale plot experiments to infer rates of ecosystem processes at broader scales (Schindler 2012). This approach can present difficulties as the results of fine-scale studies may not reflect the heterogeneity evident in a larger area (McClain et al. 2003). Because collection of ecosystem data at broad scales is often difficult and costly and many ecosystem services are difficult to measure directly, modeling is a vital tool for addressing both basic and applied questions in this realm. In this lab, you will examine several fundamental issues of modeling landscape-level ecosystem processes and services in order to:

1. Gain an appreciation for the need and challenges associated with examining ecosystem processes and associated ecosystem services at broad spatial scales;
2. Learn to conceptualize how ecosystem dynamics can be modeled at the scale of a landscape;

S.E. Gergel (✉)
University of British Columbia, Vancouver, BC, Canada
e-mail: sarah.gergel@ubc.ca

T. Reed
University of Wisconsin - Green Bay, Green Bay, WI, USA

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_16

3. Examine the implications of heterogeneity in rates of ecosystem processes; and
4. Explore a variety of important spatial assumptions that may affect spatial assessments of ecosystem services.

In Parts 1 and 2, we focus on the flux of phosphorus through an agricultural watershed using a very simple landscape model that enables one to easily incorporate and explore the impact of spatial heterogeneity on results. In Part 3, a series of synthesis questions helps you consider additional landscape ecological concepts important to understanding ecosystem services. Part 4 inspires you to build your own simple ecosystem service model using information provided for an urban landscape or even explore an entirely new situation. These exercises require a spreadsheet **ecosys.xls** that will constitute your modeling environment which can be downloaded from book web site.

NOTE: Before you proceed, save an extra backup copy of the model that you DO NOT manipulate in case you accidentally irreversibly alter the model.

INTRODUCTION

Eutrophication, or the enrichment of aquatic systems by excessive input of nutrients, constitutes the major threat to water quality (Schindler 2012; Howarth and Paerl 2008). Phosphorus (P) is often the limiting nutrient to algal productivity in freshwater systems (Carpenter 2008; Schindler 2012). As a result, phosphorus enrichment can lead to toxic algal blooms and increases in hazardous protozoa (Schindler 1977) which can threaten a variety of ecosystem goods and services provided by watersheds, including fisheries production, drinking water supplies, and recreation (Carpenter et al. 1998). Nuisance algal blooms can also reduce habitat diversity in shallow waters and deplete oxygen in bottom waters causing massive fish die-offs (Kaufman 1993). Additionally, the water may smell and taste foul and even cause skin irritation.

The most ubiquitous cause of eutrophication is non-point source pollution (USEPA 1990). **Non-point source pollution** refers to material entering aquatic systems from diffuse sources, such as runoff from agricultural fields; this is in contrast to point sources, such as a sewage treatment outflow pipe. Agricultural areas, particularly during storm events, can contribute significantly to non-point source phosphorus pollution (Omernick et al. 1991; Osborne and Wiley 1988; Correll et al. 1999). **Riparian buffer strips**, bands of uncultivated vegetation adjacent to surface waters, slow phosphorus flow and can be used to mitigate fertilization of water bodies in agricultural areas (Hoffmann et al. 2009). Wetlands also play a pivotal role in the biogeochemistry of landscapes (Verhoeven et al. 2006). Such types of biogeochemical "hotspots" (sensu McClain et al. 2003) are essential to consider when understanding nutrient fluxes in landscapes.

In this lab, we present an ecosystem model of a hypothetical agricultural landscape surrounding a canal. The canal leads to a nearby lake used by the public for swimming and boating. The model represents the flow of phosphorus from fertilized agricultural fields, through the riparian buffer strip, and then into the canal during a large storm event. While the model presents a highly simplified version of riparian and P dynamics (Hoffmann et al. 2009), it provides a useful introduction to modeling ecosystem processes at the scale of a landscape. This model is designed to address questions such as: *How much phosphorus can farmers apply to their fields without causing severe algal blooms?* as well as *At a given phosphorus application level, how much phosphorus must be retained by the buffer strip to maintain low phosphorus levels in the canal?*

Part 1. Conceptualizing Landscape-Level Ecosystem Models: Phosphorus Loading in an Agricultural Landscape

Open the file **ecosys.xls** using Excel© spreadsheet software. The spreadsheet has been configured to represent a model agricultural landscape. The brown cells on the landscape represent farmed lands. After fertilizer is applied, some P flows downhill towards the canal, represented by blue cells. The green cells represent vegetated buffer strips. The number in each cell represents the total amount of P available to leave that cell, after within cell uptake and processing is taken into account. Each cell in the model landscape represents one hectare (ha), a 10,000-m^2 area. The model approximates P flow across an agricultural landscape during a single storm event using the following simple parameters.

MODEL INPUT

Storm flow volume (m^3/ha) [G6] is the total amount of stream flow in each cell in the canal for the duration of the storm event.

Buffer absorption capacity (kg/ha) [G10] represents the ability of the buffer strip to prevent the passage of P to the next cell, expressed as the total amount of P that could be retained by the buffer cell. Riparian buffers stop the flow of P in a variety of ways, including: uptake by plants, trapping of soil to which the phosphorus is bound, and soil adsorption and immobilization. Here, we combine all the mechanisms into one equation for simplicity, representing the sum total of the buffer strips' ability to prevent P from entering the canal. In our idealized landscape, values for this parameter range from 20 to 40 kg/ha (Peterjohn and Correll 1984; Osborne and Kovacic 1993).

Amount of phosphorus applied (kg/ha) [G15] refers to the amount of P in the fertilizer applied to each individual cell in the model. This model assumes that fertilizer is evenly applied throughout the field. In practice, the amount of P applied through fertilizer is highly variable, ranging from 50 to 200 kg/ha (Nowak et al. 1996).

Transfer Coefficient [G12]. Our model assumes that farmers are not applying fertilizer in the rain and that 60% of the P in the cell runs off one pixel to another during a storm. This percentage is a simplification, in a real agricultural field the amount of P runoff would vary with vegetation, slope and especially rainfall intensity.

MODEL OUTPUT

Total phosphorus loading (kg) [G19] is the sum total of phosphorus entering the canal waters.

In-stream phosphorus concentration (mg/m^3) [G22] is the resulting concentration of phosphorus in the canal surface water after the total phosphorus is thoroughly mixed throughout the water column. The total concentration was multiplied by 1,000,000 to convert kg to mg, shown as "=(G19*10^6)" in the equation. We then divided by the number of cells in the stream (75 cells) times storm flow to find the mg/m^3, shown as "/(G6*75)." When the in-stream P concentration exceeds 75 mg/m^3 the system is at risk for algal blooms (Lathrop et al. 1998).

Exploring the Model—How Does it Work?

Agricultural Fields. The brown cells on the spreadsheet represent farmed areas. These fields slope down to the stream running down the middle of the spreadsheet. The number in each cell represents the amount of phosphorus "left over" after uptake within the cell is accounted for; that is, the amount available to leave the cell and flow downhill to the next cell.

Select cell **N4**. Note the equation for the amount of P that leaves this cell. It is composed of two parts. The first part of the equation: "(M4 + $G$15)" calculates the amount of phosphorus entering the cell. **M4** is the amount flowing in from the adjacent upstream cell. "$G$15" is the amount of fertilizer applied directly to the cell by the farmer (an input parameter you can alter). The sum of these numbers is the total amount that entered the cell. However, not all of this phosphorus flows to the adjacent downhill cell during a storm, as some is taken up by plants, adsorbed to soil, or leached into groundwater before it reaches cell **O4**. Thus, the 0.6 multiplier (or transfer coefficient), accounts for the fact that only 60% of the P that entered the cell can be washed into the next cell (i.e., 40% is taken up). This is an oversimplification. In reality, soil cannot bind an infinite amount of P. The model also assumes that flow is unidirectional, downhill towards the canal. This is another simplification. Flow is likely to be much more complex in a natural landscape.

Buffer Strips. Next examine the buffer strips (the green areas) along the banks of the canal. Select cell **T22**.

Q1 Write the formula for cell **T22** and explain in words what it means. (*NOTE:* These cells contain the Excel© equivalent of an "IF THEN" statement to prevent the program from printing negative numbers. These statements read: IF(x<y, print this if true, print this if false). For your answer, describe what the equation in the "print this if false" section means).

Drainage canal. Eventually, some P may make its way into the canal. Notice the differences in loading values for nearshore stream cells due to the variable width of the buffer at different locations.

Q2 Write the formula and explain in words how the output parameter **in-stream phosphorus concentration** is calculated.

Now, select cell **V15**. Enter a value of five into the cell. Repeat for cells **V4**, **V10**, **V11**, **V12**, and **V22**. Did the **in-stream phosphorus concentration** increase? By how much?

You just simulated several "cow patties" produced by a small herd of cows wading in the canal.

Part 2. Heterogeneity in Ecosystem Processes

In this section, you will manipulate different components of the model to gain familiarity with how it can be used to explore alternative scenarios involving spatial variation in parameters and rates.

EXERCISE 1: Phosphorus Application Rates

As with all simulation models, important simplifying assumptions have been made for this model. Notice that all agricultural areas, for example, have the same amount of P applied to each cell. In reality, the amount applied to each cell could vary for several reasons. For example, a farmer might determine that a certain area of the field needs more fertilizer than other areas due to soil type. Also, different fertilizer application techniques might result in uneven P application throughout a watershed.

Q3 Consider that two farmers live on opposite sides of the creek and simulate the effect of different farming practices on the landscape. Implement this by changing the formulas in the cells, or by summing total P runoff for different sides of the landscape under alternative P application rates.
(a) Explain your modification.
(b) What effect does this heterogeneity in fertilizer application have on the in-stream phosphorus concentration?

Q4 Another difference in P movement could be due to differences in crop type. For example, hay production requires less P than corn (Newman 1997). Describe how you would change the model to incorporate differences in crop type. What equation would you change? How would the equation be changed?

EXERCISE 2: Topographic Heterogeneity and Transfer Rates

Additional factors may cause heterogeneity in ecosystem processes (McClain et al. 2003; Hoffmann et al. 2009). Consider the importance of heterogeneity in the rates of P movement across the landscape caused by topography. Erosion of P-containing sediment is often greater in areas of steep slopes, particularly during rain events.

Change the model to account for slope differences throughout the landscape. The easiest way to do this is by changing the amount of P leaving an individual cell, thereby simulating a reduction or increase in the processing of P in that cell. Right now the processing rate is 40% of the inputs (i.e., 60% exits the cell), but this might vary depending on whether the slope is gentle or steep.

Q5 Describe the changes you made, and the effects on P loading and concentration. What other factors might you expect to influence the movement of P (other than the transport across buffers)?

EXERCISE 3: Variation in Buffer Strip Width vs. Application Rates

You probably noticed earlier that the width of the buffer strip is important in determining P loading into the canal in this model. For the sake of managing water quality in the surrounding surface waters, a land manager or farmer may be interested in the relative importance of buffer strip width versus the amount of fertilizer applied in influencing total P inputs.

Q6 *For the modeling scenario examined here*, does it appear that individual farmer behavior (i.e., application rates) or buffer width is more important in maintaining low concentrations of in-stream phosphorous? Answer in light of the constraints of the model and the range of parameters given.

Continue to manipulate the model, changing parameters at will. *Be certain that you understand all the model parameters and how all model formulas were derived.*

Part 3. Synthesis of Spatial Approaches to Ecosystem Services

Many of the same challenges you examined for understanding ecosystem processes at broad scales (e.g., heterogeneity, scaling up, terrain) are equally important when considering ecosystem services. Ecosystem services refer to the benefits humans receive from nature (Daily 1997). A wide of variety of definitions of ecosystem services exist and are vigorously debated (as in de Groot et al. 2002). Our goal in this section is to explore ways that spatial arrangement and spatial heterogeneity can impact ecosystem services (Syrbe and Walz 2012) at the scale of a broad landscape.

Each of the synthesis questions below is based around a key paper (or two) in the ecosystem services literature. Your instructor may wish to assign one question/one paper to different teams to explore in detail. Alternatively, you may wish to explore these questions (more quickly, during class) as thought exercises.

SYNTHESIS

Q7 Consider the ecosystem processes you just modeled and which ones are related to ecosystem services. How would you distinguish a service vs. a process? Is this distinction important? Why or why not? (*HINTS*: see Haines-Young and Potschin 2010 or Keeler et al. 2012)

Q8 Are there any ecosystem services for which spatial heterogeneity or spatial arrangement might NOT be important to consider? Explain your reasoning.

Q9 The primary dynamic explored in the previous modeling exercise is that of trade-offs: whereby management for one ecosystem service can negatively impact the provisioning of another. Food production affecting freshwater provisioning is a "classic" ES trade-off of great concern. Another type of interaction is a synergy whereby managing for one particular service helps augment another service. Using your knowledge of ecology, explain a few potential ecosystem service synergies (*HINTS*: see Bennett et al. 2009 or Qiu and Turner 2013).

Q10 Spatial characteristics of ecosystem services are important for a multitude of reasons and can be another way to organize or classify ecosystem services. Consider Costanza (2008) (reproduced in Table 16.1 here) which outlines five spatial characteristics potentially important to consider. Which of these five categories were already represented in the ecosys.xls model? Consider a spatial characteristic NOT represented in the model and explain how you might incorporate it.

Q11 Another important spatial consideration for ecosystem services is that of access which is influenced not only by *where* in the landscape services are produced but also by regulations, roads, as well as characteristics, abilities, and preferences of people who may wish to access various services. Some example ES might include bird-watching or harvesting wild foods (fish, berries, mushrooms, wild rice). Consider how one would model an ecosystem service with access considerations incorporated. Explain the type of spatial information you might incorporate and how you would link the new information to ecosystem processes, services, and access.

Q12 The long-term dynamics of ecosystems and the impact of landscape history have been of interest to landscape ecology for some time. It is appreciated that ignoring landscape history and/or baseline conditions can be problematic for truly understanding ecosystems. How might ignoring landscape history and prior conditions impact ecosystem services? How might incorporating landscape history improve our understanding of ES? (*HINTS*: see Tomscha and Gergel 2016; Sutherland et al. 2016; Renard et al. 2015).

Table 16.1 Categorization of ecosystem services based on spatial characteristics (adapted from Costanza 2008)

Spatial characteristics	Ecosystem service
Global (independent of proximity)	Climate regulation
	Carbon sequestration (NEP)
	Carbon storage
	Cultural/existence value
Local (depends on proximity)	Disturbance regulation/storm protection
	Waste treatment
	Pollination
	Biological control
	Habitat/refugia
Directional flow (from point of production to point of use)	Water regulation/flood protection
	Water supply
	Sediment regulation/erosion control
	Nutrient regulation
In situ (point of use)	Soil formation
	Food production/non-timber forest products
	Raw materials
User movement related (flow of people to unique natural features)	Genetic resources
	Recreation potential
	Cultural/aesthetic

Part 4. Constructing Your Own Model

Now that you have been introduced to the fundamentals of a simple landscape model and explored its parameters and possibilities, you have the basic tools to design your own landscape model. Next, you will use the same basic concept of combining cells of landscape elements (in a spreadsheet) to build your own landscape-level ecosystem model. You might also wish to incorporate your spatial understanding of ecosystem services (from Part 3) into your next model.

(*NOTE:* At this point, we switch our focus to urban landscapes, but those interested in continuing with ecosystem services in an agricultural setting, but with a more sophisticated and realistic modeling environment, are encouraged to explore Chapter 19.)

EXERCISE 4: Basic Modeling Version

Your task is to build a model to answer a specific question regarding the dynamics of P runoff in an urban landscape. Your urban environment is a city, such as Chicago or Seattle. In Excel©, you will model a city using a set of cells representing different elements of the urban environment (Table 16.2). Each element has its own level of phosphorus runoff and/or absorption. Using your imagination, create a city that

Table 16.2 Parameters for a simple spatial model of phosphorus flux through an urban watershed

Land cover type	Amount of phosphorus produced (g/20 m^2)	Phosphorus absorption capacity (g/20 m^2)	Simplified transfer coefficient (proportion)
Lawn (heavily fertilized)	30	–	0.60
Lawn (slightly fertilized)	4	–	0.60
City park (slightly fertilized)	4	–	0.60
Residential homes	70	–	1
Apartments	30	–	1
Commercial district	20	–	1
Industrial district	40	–	1
Construction site	200	–	1
Road	0	0	1
Runoff treatment wetland	–	40	0.60
Forest	–	50	0.60

contains at least a small proportion of all of the provided urban elements. Your city is adjacent to a small river which receives urban storm-water runoff.

Design and then manipulate your model specifically to answer at least *one* of the following questions:

1. City parks tend to be sinks for phosphorus although they may be slightly fertilized. What proportion of the city must be occupied by parks to maintain in-stream P levels below 75 mg/m^3? How does the spatial arrangement of the parks affect the proportion of the city that parks must occupy to maintain in-stream P levels below 75 mg/m^3?
2. What proportion of the stream must be bordered by runoff treatment wetlands in order to reduce in-stream P concentrations by 10%? By 50%? To eliminate phosphorus input altogether? What proportion of the stream must be bordered by treatment wetlands to maintain in-stream P levels below 75 mg/m^3?
3. Keeping the total area occupied by housing constant, what effect does varying the proportions of residential housing in apartments vs. homes (e.g., 30/70, 50/50, 90/10) have on P runoff to the stream? What proportions would you recommend to maintain in-stream P levels below 75 mg/m^3?
4. Consider your urban landscape from the perspective of one (or more) terrestrial ecosystem services provided by urban trees and vegetation (Escobedo et al. 2011). For example, urban parks are important for a variety of recreational purposes, greenspace has been linked to human health outcomes and well-being, and urban vegetation affects a variety of wildlife species in positive and negative ways. Redesign the provided urban model to address one or more of these terrestrial ecosystem services.

Model Parameters. You are provided with the following parameter estimates (Table 16.2) and in the spreadsheet. Notice, however, that the resolution of the runoff and absorption estimates is different than for the model you examined in Parts 1 and 2. You will probably want to adjust the scale of your model from the 1-ha resolution used in the agricultural model as city lot sizes are rarely that large. Here, we have provided the model parameters in units of g/20 m^2. For an urban landscape, 20 m^2 cells roughly approximate the minimum size (or spatial grain) of the landscape elements you will model. Building on this cell size, you could combine 1 residential housing pixel with 1 lawn pixel to represent one residence.

To incorporate both urban and agricultural areas in your landscape you can use values from Part 1 but will need to do some conversions (remember 1 ha = 10,000 m^2). You may adjust the grain size further as appropriate for your model and the questions you are trying to address, but be sure to choose an appropriate grain size for your model, and adjust the runoff and absorption capacity values accordingly. Lastly, you can assume that all processes that contribute to phosphorus runoff and/or absorption have been taken into account with the parameters given.

Transfer coefficients. In addition to the absorption capacity of a land cover type, the amount of P transferred to the next cell may also be diminished by a transfer coefficient. This reflects that some land-cover types are less permeable to runoff than others such that more runoff moves from one cell to the next. In the agricultural model, we used a transfer coefficient of 0.6, meaning that only 60% of the P in a cell was available to move out to the next cell. In this section, only wetland areas and forests have values for absorption capacity. We have, however, included transfer coefficients to account for soil permeability in lawns and parks, which we examine next.

Building Your Model. Switch to the second page of the spreadsheet file by clicking on the tab labeled **Urban Landscape** at the bottom of the spreadsheet. Again, here are all the elements with which to build your urban landscape, identical to those in Table 16.2. Click on the cells in the **Equations** column to view the equations, which incorporate transfer coefficients in some cases, for different land-cover types. The cells in the example column can be cut and pasted into the spreadsheet to build your urban landscape.

(*NOTE:* These equations represent P flow only from left to right. Unless you want to rewrite some of the equations to represent flow in the opposite direction, place your river, stream, or canal on the right-hand boundary of your landscape. Be sure to examine each cell to see which other cells are referenced).

Construct your model in the same general form as the model in Part 1. For simplicity, you may assume that flow is unidirectional, downhill towards the canal. Thus, as before, the phosphorus values in each cell represent the amount leaving that cell. This includes the runoff entering from the adjacent upstream cell plus or minus the runoff/absorption estimate for that land cover type, and in some cases, a transfer coefficient. *Remember that the number in each cell should represent the total phosphorus available to leave the cell, after any within cell uptake or processing or reduction due to the transfer coefficient.* The concentration in the water can be calculated by the amount of P flowing into the stream cell multiplied by the total volume of water that flowed through the stream during the storm event. When your model construction and manipulation are finished, complete the write-up portion of the lab.

EXERCISE 5: Advanced Modeling Version

Your task is to model any landscape-level ecosystem process of your choosing. You will use the basic concept of landscape element blocks in Excel©, but you are free to design those elements using your own knowledge, experience, and imagination. As in the basic version (above), *your model must be designed to answer a clearly defined question* (or set of related questions), but you will choose the question yourself. Be sure that you have a clear understanding of the underlying assumptions of your model throughout the building process, and be able to state those assumptions clearly.

Be sure to explicitly determine the appropriate grain size of your model. Also consider whether the values in each cell represent amount *entering* or *leaving* a given cell. If you have more than 1 day to complete this assignment, we recommend that you spend some time researching the literature and use realistic parameters to construct your model. Keep in mind that you must be able to manipulate your model to address your initial question. When the model and manipulation are finished, complete the write-up that follows.

Modeling Hints

1. Consider using the **Format**, then **Cell**, then **Patterns** commands on your spreadsheet's pull-down menu to assign different colors identifying different landscape elements.
2. Learn how to use the **$** symbol when cutting and pasting. For example, if you wanted to copy a formula "**=$F$6+5**" from one cell to a cell in the next column over, the **$F** preserves the column reference, while **$6** preserves the row reference; thus, the formula would remain **=$F$6+5** when copied and pasted. Otherwise, the formula typed as "**=F6+5**" becomes **=G6+5** when copied one cell to the right or becomes **=F7+5** when copied to the cell below.

WRITE-UP

Include the following sections in your report:

1. **Introduction**
   (a) State the question(s) your model addresses.
   (b) Provide some context for why this question is important.

2. **Description of Model**
   (a) State the underlying assumptions of your model.
   (b) Describe your model. Define the spatial and temporal scale of your model. (For the advanced version, list and explain all model parameters).

3. **Simulations and Results**
   (a) Clearly describe each "simulation experiment" with the model and summarize the results.
   (b) Answer the question(s) your model was designed to address.

4. **Discussion**
   (a) What are the implications of heterogeneity in rates of ecosystem processes in your model scenario?
   (b) Within the realm of the ecosystem process that you have modeled, what are the limitations of your model? Why?
   (c) What additions/modifications would you make to your model to address the limitations listed above?
   (d) When would considering the spatial arrangement of landscape elements or the role of landscape heterogeneity not matter to your results?
   (e) When would sampling at broad scales not be important?
   (f) How would a longer temporal scale effect your results?

5. **Literature Cited** (not included in page limits)
6. **Appendix** (not included in page limits)
   (a) If required, a copy of the answers to the exploratory questions posed in Parts 1 and 2 of this chapter
   (b) Print out of the Excel© file containing *YOUR* model

Your instructor will determine page lengths depending on the amount of time you have to complete your assignment. Consider giving oral presentations of your results.

REFERENCES AND RECOMMENDED READINGS[1]

*Bennett EM, Peterson GD, Gordon LJ (2009) Understanding relationships among multiple ecosystem services. Ecol Lett 12(12):1394–1404. *A carefully considered framework for examining interactions among ecosystem services, with examples from a variety of landscape types.*

Carpenter SR (2008) Phosphorus control is critical to mitigating eutrophication. Proc Natl Acad Sci 105(32):11039–11040

Carpenter SR, Caraco NF, Correll D et al (1998) Non-point pollution of surface waters with phosphorus and nitrogen. Ecol Appl 8(3):559–568

Correll DL, Jordan TE, Weller DE (1999) Effects of precipitation and air temperature on phosphorus fluxes from Rhode River watersheds. J Environ Qual 28:144–154

*Costanza R (2008) Ecosystem services: multiple classification systems are needed. Biol Conserv 141:350–352. *A practical, brief overview of why and spatial characteristics of ecosystem services matter.*

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Daily G (1997) *Nature's services: societal dependence on natural ecosystems*. Island Press, Washington, DC, p 392

de Groot R, Wilson MA, Boumans RM (2002) A typology for the classification, description and valuation of ecosystem functions, goods and services. Ecol Econ 41(3):393–408

Escobedo FJ, Kroeger T, Wagner JE (2011) Urban forests and pollution mitigation: analyzing ecosystem services and disservices. Environ Pollut 159(8):2078–2087

Haines-Young R, Potschin M (2010) The links between biodiversity, ecosystem services and human well-being. In: Raffaelli DG, Frid CLJ (eds) *Ecosystem ecology: a new synthesis*. Cambridge University Press, Cambridge, pp 110–139

Hoffmann CC, Kjaergaard C, Uusi-Kämppä J et al (2009) Phosphorus retention in riparian buffers: review of their efficiency. J Environ Qual 38(5):1942–1955

*Howarth R, Paerl HW (2008) Coastal marine eutrophication: control of both nitrogen and phosphorus is necessary. Proc Natl Acad Sci 105(49):103. *A review of eutrophication problems globally with an emphasis on nitrogen fluxes with an emphasis on terrestrial, aquatic and marine interactions, as well as contrasts between N and P.*

Kaufman L (1993) Catastrophic change in species-rich freshwater ecosystems. Bioscience 42:846–858

Keeler BL, Polasky S, Brauman KA et al (2012) Linking water quality and well-being for improved assessment and valuation of ecosystem services. Proc Natl Acad Sci 109(45):18619–18624

Lathrop RC, Carpenter SR, Stow CA et al (1998) Phosphorus loading reductions needed to control blue-green algal blooms in Lake Mendota. Can J Fish Aquat Sci 55(5):1169–1178

*Likens GE, Bormann RH (1995) *Biogeochemistry of a forested ecosystem*. Springer, New York. *The Hubbard Brook Ecosystem Study, began in 1963, has been foundational to the study of ecosystem processes at the scale of large watersheds and was been a pioneering scientific acheivement in broad-scale whole ecosystem experiments.*

*Lovett GM, Jones CG, Turner MG et al (eds) (2005) *Ecosystem function in heterogeneous landscapes*. Springer, New York, p 489. *Consider the very helpful book review by Carol Wessman describing the chapters and flow and of this exceptionally synthetic volume. Ecology, 88(3), 2007, pp. 803–804.*

*McClain ME, Boyer EW, Dent CL et al (2003) Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. Ecosystems 6(4):301–312. *Highly-cited classic paper examining spatial and temporal heterogeneity in ecosystem processes.*

Newman EI (1997) Phosphorus balance of contrasting farm systems, past and present. Can food production be sustainable? J Appl Ecol 34:1334–1347

Nowak P, Shepard S, Weiland C (1996) *Utilizing a needs assessment in water quality program implementation for the Lake Mendota watershed*. The Farm Practices Inventory (FPI) Report #2. Available via Environmental Resources Center, University of Wisconsin

Omernick JM, Abernathy AR, Male LM (1991) Stream nutrient levels and proximity of agricultural and forest land to streams: some relationships. J Soil and Water Conserv 36:227–231

Osborne LL, Kovacic DA (1993) Riparian vegetated buffer strips in water quality restoration and stream management. Freshw Biol 29:243–258

Osborne LL, Wiley MJ (1988) Empirical relationships between land use/landcover and stream water quality in an agricultural watershed. J Environ Manage 26:9–27

Peterjohn WT, Correll DL (1984) Nutrient dynamics in an agricultural watershed: observations on the role of a riparian forest. Ecology 65:1466–1475

Qiu J, Turner MG (2013) Spatial interactions among ecosystem services in an urbanizing agricultural watershed. Proc Natl Acad Sci USA 110(29):12149–12154

Renard D, Rhemtulla JM, Bennett EM (2015) Historical dynamics in ecosystem service bundles. Proc Natl Acad Sci 112(43):13411–13416

*Running SW, Nemani RR, Peterson DL et al (1989) Mapping regional forest evapotranspiration and photosynthesis by coupling satellite data with ecosystem simulation. Ecology 70:1090–1101. *Landscape-level simulation model of annual evapotranspiration and net photosynthesis in a mountainous region by some of the original leaders in developing spatial ecosystem models linked to remotely-sensed data.*

Schindler DW (1977) Evolution of phosphorus limitation in lakes. Science 195:260–262
Schindler DW (2012) The dilemma of controlling cultural eutrophication of lakes. Proc R Soc Lond B Biol Sci 283(1827):1–12
*Sklar FH, Costanza R (1991) The development of dynamic spatial models for landscape ecology: a review and prognosis. In: Turner MG, Gardner RH (eds) *Quantitative methods in landscape ecology*. Springer, New York, pp 239–288. *A classic overview of spatial modeling as approached by both the social and natural sciences.*
Sutherland I, Bennett EB, Gergel SE (2016) Recovery trends for multiple ecosystem services reveal non-linear responses and long-term tradeoffs from temperate forest harvesting. For Ecol Manage 374:61–70
*Syrbe RU, Walz U (2012) Spatial indicators for the assessment of ecosystem services: providing, benefiting and connecting areas and landscape metrics. Ecol Indic 21:80–88. *Extremely thoughtful piece exploring how ecosystem services flow across large regions. Also introduces exciting ideas about how to link landscape metrics to ecosystem service provisioning.*
*Tallis H, Polasky S (2009) Mapping and valuing ecosystem services as an approach for conservation and natural-resource management. Ann N Y Acad Sci 1162(1):265–283. *Overview of one of the earliest spatial models of ecosystem services, InVest, with some example applications. Further development of this software continues at a rapid pace.*
Tomscha SA, Gergel SE (2016) Ecosystem service trade-offs and synergies misunderstood without landscape history. Ecol Soc 21(1):43
USEPA (1990) *The quality of our nation's water*. U. S. Environmental Protection Agency 440/4-90-005. Washington, DC
USEPA (1997) *Index of watershed indicators*. U. S. Environmental Protection Agency -841-R-97-010. Washington, DC
Verhoeven JT, Arheimer B, Yin C et al (2006) Regional and global concerns over wetlands and water quality. Trends Ecol Evol 21(2):96–103

Chapter 17
Heterogeneity in Ecosystem Services: Multi-Scale Carbon Management in Tropical Forest Landscapes

Kathryn R. Kirby, Jeanine M. Rhemtulla, and Sarah E. Gergel

OBJECTIVES

Landscape management is increasingly focused on trade-offs among various ecosystem services. For example, while clearing forests may produce timber and provide land for agriculture, it also releases significant amounts of carbon to the atmosphere, influencing the global climate system. Evaluating the tradeoffs among ecosystem services is made difficult by the inherent heterogeneity of social–ecological systems at many levels of ecological (and social) organization. For example, the provisioning of ecosystem services may change with the size of organisms, the species composition of communities, and with variation in landscape pattern through time. In this chapter, we introduce common methods for estimating the amount of carbon stored in forests and explore the implications of spatial and temporal heterogeneity for carbon management at the landscape level. Assuming little prior knowledge of these issues, these exercises will enable students to:

1. Estimate standing stocks of carbon using methods appropriate to different spatial scales;
2. Explore the social–ecological implications of the relationship between tree sizes, wood density, and carbon stocks;
3. Quantitatively evaluate the relative impact of forest loss versus forest fragmentation on landscape-level carbon storage; and
4. Contrast the impact of alternative management regimes on total carbon stocks as well as flows through time.

K.R. Kirby (✉)
University of Toronto, Toronto, ON, Canada
e-mail: kate.kirby@utoronto.ca

J.M. Rhemtulla • S.E. Gergel
University of British Columbia, Vancouver, BC, Canada

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_17

In Part 1, we introduce basic methods appropriate for carbon accounting at the level of individual trees and forest stands. In Part 2, we explore why variation in landscape composition as well as landscape arrangement are both important to consider. Lastly, in Part 3, students use a simple landscape simulation model (in Excel) to explore how management activities may impact carbon storage in forests, including a consideration of nonlinearities in the amount of carbon stored over space and time. The only material needed for the exercise is access to a computer and the spreadsheet entitled **carbon.xlsx**, downloadable from the book webpage. Your instructor may wish to assign Part 1 as a "pre-class" assignment (to be completed prior to coming to class) in order to save classroom time for Parts 2 and 3.

Part 1. Estimating Carbon Stocks: From Trees to Forest Stands

Forests are "living storage units" for carbon. As plants grow, they sequester atmospheric carbon (CO_2) through photosynthesis. The sequestered carbon is incorporated into plant structures (e.g., stems and foliage) so that approximately 50% of the biomass, or dry weight, of plants is carbon. This stored carbon (referred to as **carbon stocks)** is only released back into the atmosphere after plants (or their parts) break down, usually through decomposition or sometimes fire. The **flux** of carbon through forests is thus determined by the relative rates of photosynthesis (sequestration) and decomposition (emissions). Deforestation and logging accelerate the release of carbon stored in vegetation back into the atmosphere through tree death and decomposition. Because elevated levels of atmospheric carbon are driving global climate change, scientists are increasingly interested in quantifying the impacts of forest management on forest carbon.

One method commonly used to quantify forest carbon storage combines models of tree allometry with nondestructive measurements of trees (e.g., measurements collected as part of standard forest inventories). For example, **allometric models** have been developed that relate tree diameter to total tree biomass. Tree diameter is typically measured during forest inventories as **DBH**, or diameter-at-breast-height, so-called because it is measured at a standard height of 130 cm above the ground. Applying a DBH-biomass allometric model to forest inventory data therefore produces an estimate of biomass for the inventoried forest. This can be transformed into an estimate of carbon based on the relationship between biomass and carbon (on average, 47% of the biomass of tropical trees is carbon) (Martin and Thomas 2011).

EXERCISE 1: Predicting Carbon Stocks Using Diameter at Breast Height (DBH) for a Common Neotropical Tree Species

Anacardium excelsum is a rather conspicuous tree in the forests of Panama. It is known locally as "espavé," from the Spanish phrase "es para ver" ("is to see"), a reference to its tall height and utility as a look-out point when climbed. *A. excelsum* is harvested for timber, but is also used to make furniture, boats, and dugout canoes.

Here, you will estimate the carbon stored in *Anacardium excelsum* trees of different diameters using an allometric model that relates tree DBH to biomass (Chave et al. 2005).

- Open the workbook **carbon.xlsx**, and go to the worksheet **Tree**.
- The allometric model of Chave et al. (2005) has been entered for you in the spreadsheet as an Excel function [cell **C12**]. Apply this function to *A. excelsum* trees with DBH ranging from 10–150 cm by adding an "=" sign in front of the function, then copying and pasting into the cells below.
- Multiply the resulting biomass estimates by 0.47 to convert biomass to carbon [cells **D12** to **D26**].
- Plot carbon vs. DBH.

Q1 What do you notice about the relationship between carbon stored and tree DBH (diameter at breast height)?

Q2 Considering this relationship, what type of tree or stand would be most valuable to managers aiming to maximize standing carbon stocks?

Q3 List some ecosystem services other than carbon storage provided by forests. Include at least one example each of services with ecological, economic as well as cultural value. Are any of these other ecosystem services dependent on individual trees? Do you expect these services to vary with the size and species of individual trees? Explain why or why not? (*HINT:* see Ellison et al. 2005; Manning et al. 2006 or Salick et al. 2007).

EXERCISE 2: Contrast the Carbon Stocks of a Hardwood and Softwood Tree Species

Wood density varies widely among tree species and impacts carbon storage by determining how much carbon is stored per unit of tree volume. The general allometric model you used in Exercise 1 can be applied to all species of tropical trees using average values for wood density; however, using species-specific wood density values greatly improves accuracy (Chave et al. 2005). Here, you will contrast carbon storage by two species of trees native to Central America. The first, *Dalbergia retusa* ("cocobolo" in Spanish, "rosewood" in English), is a hardwood species that is used by artisans for fine wood carving; it has a wood density of 0.86 g cm^{-3}. The second, *Ochroma pyramidale* ("balsa") is a light wood used locally to raft down rivers; it has a wood density of 0.16 g cm^{-3}.

- Go to the worksheet **Tree2**.
- The allometric model of Chave et al. (2005) has been entered for you in the spreadsheet as an Excel function. Apply this function to *D. retusa and O. pyramidale* trees with DBH ranging from 10–110 cm (add a "=" sign in front of the functions in cells **C14** and **E14**, then copy and paste into the rest of the cells in the column).

- Multiply the resulting biomass estimates by 0.47 to convert biomass to carbon [cells **D14–D24** and **F14–F24**].

Q4 Contrast the carbon stocks of a *D. retusa* and *O. pyramidale* individual of the same diameter. What is the difference in carbon storage among the two individuals? At what DBH would an *O. pyramidale* tree store the same carbon as a 30 cm DBH *D. retusa* tree?

Q5 Can you think of other groups of organisms in which the efficiency of ecosystem service provisioning differs among species? (*HINT:* for an example related to pollination, see Brittain et al. 2013).

EXERCISE 3: Scaling From Individual Trees to a Forest Plot

Barro Colorado Island, Panama, has been a site of intensive ecological research since 1923. Today, its lowland moist tropical forests are some of the best studied in the world. Beginning in 1982, and then every 5 years since 1985, a detailed inventory has been conducted of 50 ha of mature tropical forest on the island (Condit 1998; Hubbell et al. 1999, 2005). As part of the inventory, the DBH and species name of each tree exceeding 10 cm DBH are recorded. Here, you will work with inventory data from a one-hectare subplot on Barro Colorado Island, in which 416 individual trees belonging to 82 species were measured. (*NOTE:* Data from this site were also used for the exercises in Chapter 15 to explore spatial statistics). The data are contained in the worksheet **Plot**.

- The worksheet **Plot** contains a list of all of the individuals measured in the 1 ha plot, including their ID number, species name, wood density, and DBH.
- Apply the allometric model of Chave et al. (2005) to all the trees measured in the plot (the function has been entered for you in cell **F2**; copy it to the cells below). (*NOTE:* biomass is now being calculated in Mg (megagrams or metric tons) rather than kg).
- Convert biomass to carbon by multiplying by 0.47.
- Sum the carbon stored in all the trees in the 1 ha plot. Record your answer.

EXERCISE 4: Impact of Selective Logging on Forest Carbon Storage

Selective logging is typical in the tropics and differs from the clear-cut harvesting approach common in temperate coniferous forests. Trees in less diverse temperate conifer forests, which tend to be dominated by one or a few species, may have the same end-use and thus are easily processed at the same mill. However, the diverse panoply of tree species typically encountered in tropical forests may have a wide variety of end-uses that require different processing technologies. As a result, tropical loggers will typically remove only a subset of the tree species present, often targeting hardwood species that reach the largest diameters when mature. This form of "high-grading" can have important implications for ecosystem services. Here, you will compare the impacts on plot-level carbon stocks of two approaches to logging.

First, simulate selective logging of the 10% of individuals with the largest diameters.

- Using the worksheet **Plot**, sort the list of trees by DBH.
- Select the 42 trees with the largest diameters at breast height, and delete them (as though you were harvesting 10% of the largest trees in the plot).
- Record the new carbon stock total for the plot.

Second, simulate logging of a random selection of 10% of individuals in the 1 ha plot (as though you were clear-cutting 10% of individuals in the forest stand, without regard to species identity or tree size).

- Start again with the original plot of trees. (You can either undo your changes to make sure the trees you deleted are back in the plot or download the file again and save as another name).
- In column **H** in the worksheet **Plot**, use Excel's random number generator "**=RAND()"** to generate a random number for each tree in the plot.
- Select the column, Copy it, then go to **Paste Special > Paste Values** (this ensures that the random numbers will not be recalculated every time you sort the worksheet).
- Sort the worksheet by the random number column.
- Select the first 42 trees for simulated harvest (i.e., those with the highest random numbers). Delete these trees, thereby removing a random selection of 10% of trees from the plot.
- Record the new carbon stock total for the plot.

Q6 Determine the relative impact of removing 10% of the largest diameter trees as opposed to removing 10% of trees via random selection on total carbon storage. Does removal of 10% of trees reduce total carbon storage by 10%?

Q7 Which natural disturbances remove random vs. large diameter trees? How does this compare to anthropogenic disturbances? Are there differences in tropical versus temperate forests?

OPTIONAL: Field Exercise

A field exercise has been developed building on Part 1 that provides an opportunity for students to estimate carbon stocks using their own forest sampling data. For the exercise, students establish local sample plots and collect DBH measurements and species names for trees in the plots. As in Exercise 3, these data are converted to estimates of carbon stocks per unit area, allowing students to compare carbon stocks among sample plots and/or land-cover and forest types. Please see Powers and Velásquez-Runk (2016) which is part of the online supplementary materials for this book chapter.

Part 2. Impact of Forest Loss and Forest Fragmentation on Landscape-Level Carbon Stocks

EXERCISE 5: Relationships Between Land Cover and Carbon Stocks

The relationship between biomass and carbon stocks (as in Part 1) means that land-cover types with more vegetation—and, in particular, with more large trees—store more carbon above ground. The following table presents the average above-ground carbon stocks for one hectare of six common land-cover types in eastern Panama. Keep in mind that these values are averages: initially, the carbon stocks of fallow areas will be lower than the average presented in Table 17.1, and eventually they will surpass them. Later, in Part 3, we will address these assumptions by incorporating growth curves into a temporally explicit model.

Q8 Briefly consider the impacts of transitions among the different classes (e.g., converting mature forest to pasture or allowing pasture to regenerate as secondary forest). Which land-cover transitions would result in the greatest loss of carbon stocks (i.e., emissions of carbon to the atmosphere)? Which result in the greatest increase in carbon stocks (i.e., sequestration of atmospheric carbon)? (*NOTE:* Some of the changes occur over very short time periods whereas others occur over many years.)

Q9 Just as there is variation in carbon storage among tree species (Part 1), there is also variation in carbon storage among sites. For example, forests on rich soils may reach greater statures and therefore store more carbon then nearby forests on poor soils. How could this be taken into account by researchers aiming to provide general values for forest carbon for an entire region? (*HINT:* Pelletier

Table 17.1 Mean above-ground carbon stocks in 1 ha of six common land-cover classes in eastern Panama (based on Pelletier et al. 2012).

Class	Description of Forest type	Above-ground carbon (Mg C ha^{-1})
Mature forest	Old-growth and mature secondary forest	140
Old secondary forest	Secondary forest that is approx. 40 years old	129
Young secondary forest	Secondary forest that is approx. 10 years old	48
Fruit-tree agroforest	Orchard dominated by fruit trees, sometimes also containing fuelwood, timber, fiber, and medicinal tree species	50
Fallow	Crop field or pasture recently abandoned (regrowing forest less than 5 years old)	36
Pasture or annual crop field	Land used for growing annual crops or for cattle grazing	4.2

et al. 2012 explore this problem and show that it can have important impacts on landscape-level carbon estimates).

EXERCISE 6: Exploring the Effects of Forest Loss and Forest Fragmentation on Carbon Stocks

The total carbon stored in a landscape is affected not only by the amount of each land-cover type present, but also by the configuration of those land-cover types. More fragmented forests have more **edges**, which are exposed to more wind and other disturbance than are forest interiors. As a result, forest edges typically store less carbon than forest interiors (Laurance et al. 1998). The following exercise explores the impacts of these two factors (forest loss and forest fragmentation) on landscape-level carbon stocks.

You will examine three hypothetical landscapes for the exercise (Figure 17.1). In these landscapes, each pixel represents one hectare (100×100 m^2). All three landscapes contain the same proportion of mature forest and pasture (0.50 each); however, the degree of forest fragmentation increases from left to right. As a result, the number of forest edge pixels also increases from left to right.

- In the **carbon.xlsx** workbook, go to the **Landscape calculations** worksheet. Notice that the total number of pasture, mature forest edge, and mature forest interior pixels in each of the three landscapes of Figure 17.1 has been tallied for you (cell **B8** and below).

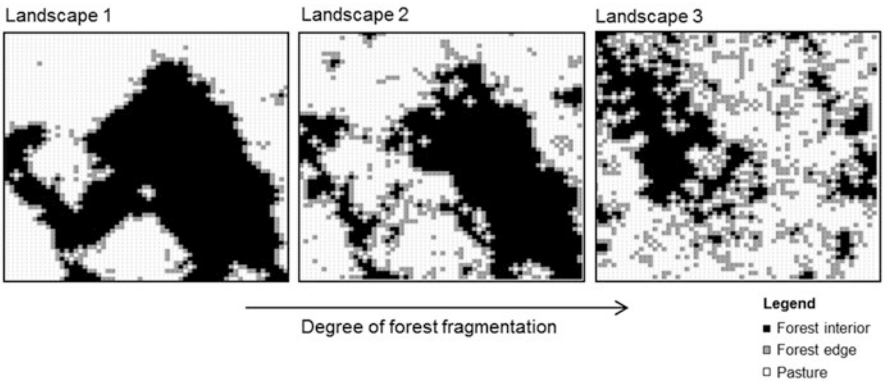


Figure 17.1 Three landscapes with identical proportions of pasture and mature forest but differing levels of forest fragmentation. Fragmentation creates forest edge (*grey pixels*), with different environmental conditions producing different ecosystem services than interior forest (*black pixels*). (*NOTE:* The landscapes were created using QRule neutral landscape software introduced in Chapter 6—with levels of clumping set at $H=0.9$, $H=0.5$, and $H=0.1$. You may wish to experiment with other configurations, and/or introduce additional land-cover types to these landscapes)

- Enter the above-ground carbon estimates for mature forest (interior) and pasture from Table 17.1 into the spreadsheet (cell **C8** and below), then multiply these values by the number of hectares of each land-cover type (column **B**) to calculate total carbon for these land-cover categories.
- To incorporate edge effects, assume that mature forest edges store 10% less carbon than mature forest interior. Reduce the mature forest carbon value in Table 17.1 by 10%, and use this as an estimate of mature forest edge carbon stocks [cells **C10, C16, C22**].
- For each landscape, sum across land-cover types to calculate total landscape-level carbon stocks in cells **D11**, **D17**, and **D23**.

Q10 What is the effect of forest fragmentation on landscape-level carbon stocks? How does the effect of fragmentation on carbon stocks compare to net losses of forest? (*HINT:* try recalculating landscape-level carbon stocks after "converting" an additional proportion of forest pixels to pasture pixels in each landscape).

Q11 Determine the sensitivity of your answer to the previous question to your assumptions regarding the carbon stocks of edges. Assume that forest edges lose carbon not just because of natural disturbances along edges (e.g., increased tree death due to wind damage) but also because of increases in human activities, such as fuel wood collection. Recalculate landscape-level carbon stocks assuming a reduction of 40% of carbon in edge forest relative to interior forest. Now what is the effect of fragmentation on landscape-level carbon stocks?

Q12 Consider the variety of ecological and abiotic differences between forest edges and forest patch interiors. Can you think of other ecosystem services that might differ between forest patch edges and interiors? Explain the differences you might expect, and why. (*HINT:* see Laurance et al. 2011).

Part 3. Impacts of Alternative Forest Management Regimes on Carbon Stocks and Fluxes Through Time

So far you have explored the effects of tree size and land-cover types on carbon stocks *at a single point in time*. However, trees sequester carbon as they grow and emit it when they die and decompose—these are *fluxes* or flows of carbon *through time*. You have already seen that large trees store disproportionately more carbon, but perhaps counterintuitively, younger trees actually have higher rates of carbon sequestration (because they are growing quickly). This has led some managers to argue that the best way to use forests to mitigate climate change is to clear mature forests and replace them with young, fast growing forests that sequester a lot of carbon. The following exercise tests this idea over a one-hundred year time period for a hypothetical tropical landscape. It uses the model contained in the **StockFlux** worksheet in the **carbon.xlsx** Excel workbook.

EXERCISE 7: Stocks vs. Flux in a Forest Landscape Through Time

In this exercise, you will compare how different forest management regimes change the stocks and fluxes of carbon in a forest landscape through time. The **StockFlux** worksheet contains a series of tables that track changes in stand age, carbon stocks, and carbon flux through time in response to forest logging and regrowth. The landscape is 10,000 ha, and the model tracks forest growth in 10-year age cohorts every 10 years over a 100-year time period. Rather than estimate carbon based on the DBH of individual trees in a stand, the model uses a carbon growth curve based on stand age. This is less accurate than estimating carbon using DBH, but is more efficient for estimating carbon over large areas.

- Open the **StockFlux** worksheet.
- Examine **Excel.Table 1** embedded within the worksheet, which shows how the forest stand age distribution of the landscape changes through time.

(*NOTE:* At the beginning of the simulation, the entire landscape (10,000 ha) is old-growth (100+ years old) **[cell K6]**, and remains this way in the absence of logging).

- Examine the graphs at the top of the sheet showing carbon stocks and fluxes over time.

Q13 Assuming no forest logging, how much above-ground carbon is stored in trees in the 10,000 ha of old-growth forest? How much carbon is sequestered and emitted through time? (*HINT:* Answers to this question also appear in embedded **Excel.Table 3** (carbon sequestered) and **Excel. Table 4** (carbon emitted) within the spreadsheet).

You can simulate logging in the model by entering the number of hectares to cut each decade in cell **B2**. Each decade, the oldest available forests will be cut, and the table will track the fate of these cut forests (which move into the youngest stand age in the following time period) and the growth of uncut forests for each 10-year period. The resulting changes in above-ground tree carbon through time are shown in **Excel.Table 2** (change in carbon stocks), **Excel.Table 3** (carbon sequestered), and **Excel.Table 4** (carbon emitted) within the worksheet.

- Log 1% of the landscape each year (10%/decade) by entering “1000” in cell **B2**.

Q14 Describe what happens to the forest stand age distribution and carbon stocks and fluxes under a 10%/decade logging regime. After 100 years, what is the mean stand age [cell **M16**], the landscape-level carbon stocks [cell **L31**], and cumulative carbon sequestered [**M44**] and emitted [**M57**]? How do these differ from the unlogged forest?

- Explore the impact of alternative forest management regimes on landscape-level carbon stocks/flux. Open the sheet **Landscape balance**, which has been linked to the sheet **StockFlux** used in the previous exercises.
- Simulate different management regimes by systematically varying the area of forest cut per decade [cell **B2** in worksheet **StockFlux**]. Record the impacts of

each simulation on: mean stand age, total carbon stocks, and cumulative carbon sequestration and emission using the table provided in worksheet **Landscape balance** (see worksheet for further instructions).

- Create two graphs: (1) total carbon stock (*Y*-axis) *vs.* mean stand age (*X*-axis), and (2) carbon sequestered and carbon emitted (both on the *Y*-axis) *vs.* mean stand age (*X*-axis)

Q15 How do carbon stocks and fluxes vary with stand age? Are the relationships linear? What type of forest has the highest carbon stocks? Highest carbon flux?

- Next, calculate the carbon balance for the landscape after 100 years of each simulated management regime in column **G** of the table in the **Landscape Balance** worksheet.

Q16 It has been suggested that replacing old-growth forest with young quickly growing forests would be a good way to help mitigate climate change. Based on your simulation results, do you agree?

In the simple model that you used in this exercise, old-growth forests reach a carbon equilibrium after 80 years of age (i.e., carbon storage ceases to increase). Recent research (e.g., Luyssaert et al. 2008) suggests that old-growth forests may in fact continue to accumulate carbon as they age. The model also assumes that when trees are cut, their entire carbon stocks are emitted. This may not be the case if, for example, timber is used for construction and so does not immediately decay.

Q17 If these two primary model assumptions discussed above are not true, would it change your answer to ***Q16***? How? (*HINT:* see Harmon et al. 1990).

CONCLUSIONS

Managing forests for multiple ecosystem services requires a careful consideration of the heterogeneity of these services across scales and through space and time. These challenges become especially obvious when trying to implement policies to enhance ecosystem service provisioning that require reliable measurements of the services. For example, international climate negotiations over the past decades have included discussions on financial incentives that would either reduce emissions from deforestation and forest degradation (e.g., the United Nation's "REDD" programme), or increase reforestation and afforestation (e.g., the Kyoto Protocol's Clean Development Mechanism). But questions such as how best to measure the gains in carbon stocks, how to ensure that deforestation isn't simply shifted in space, and what historical time period to use as a reference point, continue to be challenging. With thoughtful management, forests can play an important role in climate mitigation while contributing to local livelihoods and providing other ecosystem services.

Acknowledgments The authors thank two reviewers for their comments on earlier versions. Data from the BCI forest dynamics research project was made possible by National Science Foundation grants to Stephen P. Hubbell: DEB-0640386, DEB-0425651, DEB-0346488, DEB-0129874, DEB-00753102, DEB-9909347, DEB-9615226, DEB-9615226, DEB-9405933, DEB-9221033, DEB-9100058, DEB-8906869, DEB-8605042, DEB-8206992, DEB-7922197, support from the Center for Tropical Forest Science, the Smithsonian Tropical Research Institute, the John D. and Catherine T. MacArthur Foundation, the Mellon Foundation, the Small World Institute Fund, and numerous private individuals, and through the hard work of over 100 people from 10 countries over the past two decades. The plot project is part the Center for Tropical Forest Science, a global network of large-scale demographic tree plots.

REFERENCES AND RECOMMENDED READINGS[1]

*Balvanera P, Kremen C, Martinez-Ramos M (2005) Applying community structure analysis to ecosystem function: examples from pollination and carbon storage. Ecol Appl 15:360–375. *Presents a framework for considering how community attributes other than simple species richness affect ecosystem functioning, and how species within a community differ in their relative contributions to the function. The framework is applied to case studies of watermelon pollination by bees and carbon storage by trees. Management implications are discussed.*

Brittain C, Williams N, Kremen C et al (2013) Synergistic effects of non-Apis bees and honey bees for pollination services. Proc R Soc B 280:20122767

Bunker DE, DeClerk F, Bradford JC et al (2005) Species loss and above-ground carbon storage in a tropical forest. Science 310:1029–1031

Chave J, Andalo C, Brown S et al (2005) Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145:87–99

Condit R (1998) *Tropical forest census plots*. Springer, Berlin

*Coomes OT, Grimard F, Potvin C (2008) The fate of the tropical forest: carbon or cattle? Ecol Econ 65:207–212. *Discusses alternate land-use scenarios proposed as carbon-emission reduction strategies in the tropics, and their economic costs and risks to rural households and communities. Based on a case study from Panama.*

Ellison AM, Bank MS, Clinton BD et al (2005) Loss of foundation species: Consequences for the structure and dynamics of forested ecosystems. Front Ecol Environ 3:479–486

Harmon ME, Ferrell WK, Franklin JF (1990) Effects on carbon storage of conversion of old-growth forests to young forests. Science 247:699–702

Hubbell SP, Foster RB, O'Brien ST et al (1999) Light gap disturbances, recruitment limitation, and tree diversity in a Neotropical forest. Science 283:554–557

Hubbell SP, Condit R, Foster RB (2005) Barro colorado forest census plot data. http://ctfs.arnarb.harvard.edu/webatlas/datasets/bci

Laurance WF, Laurance SG, Delamonica P (1998) Tropical forest fragmentation and greenhouse gas emissions. For Ecol Manage 110:173–180

*Laurance WF, Camargo JLC, Luizão RCC et al (2011) The fate of Amazonian forest fragments: a 32-year investigation. Biol Conserv 144:56–67. *Synthesizes findings from a 30-year experimental study of forest fragmentation in central Amazonia. Provides an excellent introduction to what is known about the impacts of tropical forest fragmentation on forest microclimate, tree recruitment and mortality, and various other species groups, from beetles to birds.*

Luyssaert S, Schulze ED, Börner A et al (2008) Old-growth forests as global carbon sinks. Nature 455:213–215

Manning A, Fischer J, Lindenmayer D (2006) Scattered trees are keystone structures—Implications for conservation. Biol Conserv 132:311–321

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

Martin AR, Thomas SC (2011) A reassessment of carbon content in tropical trees. PLoS One 6, e23533

*Pelletier J, Kirby KR, Potvin C (2012) Consequences of carbon stock uncertainties on emissions reductions from deforestation and forest degradation in developing countries. Forest Policy Econ 24:3–11. *Explores how forest heterogeneity, different field sampling protocols, and alternative allometric models contribute to uncertainty in forest carbon stock estimates. The impacts of this uncertainty on carbon emission baselines, and thus on projections of how much a given project could contribute to carbon-emission reductions, are discussed.*

Powers J, Velásquez-Runk J (2016) Field exercise: quantifying aboveground biomass stocks in different land-use and forest types. Online supplementary material for Heterogeneity in ecosystem services: multi-scale carbon management in tropical forest landscapes (Chapter 17). In: Gergel SE, Turner MG (eds) *Learning landscape ecology: A practical guide to concepts and techniques*. Springer, New York.

Salick J, Amend A, Anderson D et al (2007) Tibetan sacred sites conserve old growth trees and cover in the eastern Himalayas. Biodivers Conserv 16:693–706

Ziter C, Bennett EM, Gonzalez A (2013) Functional diversity and management mediate aboveground carbon stocks in small forest fragments. Ecosphere 4:art85

Chapter 18
Regime Shifts and Spatial Resilience in a Coral Reef Seascape

Jennifer C. Selgrath, Garry D. Peterson, Matilda Thyresson, Magnus Nyström, and Sarah E. Gergel

OBJECTIVES

Ecosystems are shaped by natural processes such as predator–prey interactions and climate, as well as by human activities such as harvesting and pollution. Resilient ecosystems are able to absorb disturbances, but chronic stressors may reduce the capacity of an ecosystem to cope with change (Nyström et al. 2000). The ability of ecosystems to absorb disturbance and at the same time maintain their structure, processes, and function is known as **resilience** (sensu Holling 1973). Accumulated evidence from many systems (e.g., coral reefs, forests, rangelands, and shallow lakes) suggests that when pushed past a threshold (i.e., beyond their resilience), ecosystems can undergo a regime shift to an alternative state (Walker and Salt 2006; Knowlton 1992; Dublin et al. 1990; Scheffer et al. 1993; Peterson 2011). From an anthropocentric perspective these alternative states may be less desirable than the initial state depending on the ecosystem goods and services they produce (Moberg and Folke 1999). Strong feedbacks in the alternate state may also make recovery to the original state difficult, even after the original stressors are removed (Scheffer et al. 2001; Nyström et al. 2012). Human dimensions such as opportunity and

J.C. Selgrath (✉) • S.E. Gergel
University of British Columbia, Vancouver, BC, Canada
e-mail: j.selgrath@oceans.ubc.ca

G.D. Peterson • M. Thyresson • M. Nyström
Stockholm Resilience Centre, Stockholm, Sweden

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_18

governance also comprise an important aspect of resilience because they influence how sustainably resources are used (Ostrom 2009; Cinner 2009), thereby shifting the resilience threshold. The objectives of this lab are to:

1. Investigate how the ecological dynamics of a system can promote resilience or lead to regime shifts;
2. Explore how interactions between social and ecological processes can influence the state of a system; and
3. Use simple spatial modeling to investigate spatial aspects of resilience and to examine how resilience is influenced by social–ecological processes operating at different scales.

In this lab, you will investigate how social and ecological factors influence resilience across scales using simple nested models and maps within a spreadsheet (**reef.xlsx**) which can be found on the book web site. First, you will explore a model of a patch coral reef. This model will allow you to understand the ecological feedbacks that maintain reefs in a healthy coral-dominated state, in contrast to a degraded, algae-dominated state (Nyström et al. 2012). You will also learn how fishing practices affect reef resilience. Secondly, you will work with a spatial model that expands the patch reef dynamics to a series of linked reefs. This helps us understand how the spatial adjacency of multiple reefs influences resilience and spatial resilience (Nyström and Folke 2001). Third, you will explore how social and ecological factors across multiple scales interact to influence the resilience of the seascape. The lab concludes with a series of synthesis questions which allow for opportunity to think about conservation implications of the lab and spatial resilience in other systems.

NOTE: Before you begin, we recommend two things:

- View the fantastic color images of coral reef systems and small-scale fisheries in the Appendix. These images explain the organisms and fishing practices and will help you visualize the heterogeneity of this seascape;
- Print the map associated with this lab which will be used in Part 3. It can be found on the web site for this book (file called **reef_map.pdf**) and within your **reefs.xlsx** spreadsheet under the tab entitled: **seascape map PRINT THIS**.

INTRODUCTION

Biotic and abiotic factors influence how an ecosystem functions. New elements can be added (invasive species), old pieces can be lost (local extinction), and the most common components can change over time. Such changes can cause an ecosystem to shift to an alternative state. In a coral reef system, a shift from a coral-dominated to a macroalgae-dominated system provides one classic example of a regime shift (see Appendix Image Series 1). Such shifts are generally driven by human impacts, and the alternate state is characterized by changes in ecological structures and species interactions. Regime shifts can have a significant influence on societal and economic development through changes in ecosystem services provided by the system. For example, coral-dominated reefs provide protein and livelihoods to millions of people in tropical coastal zones (Burke et al. 2011). In contrast, once reefs are dominated by macro-algae, reefs are likely to provide less food and fewer types of fish to dependent communities. Consequently, social and ecological systems are intimately linked (Graham et al. 2013).

An Introduction to the Dynamics of Coral Reefs

Corals—the main architects of coral reef ecosystems—are small colonial invertebrates (3–56 mm) that form colonies and build the reefs as shelters (Appendix Image Series 2). The evolutionary success of reef-building (hermatypic) corals in forming reefs is to a large extent due to the symbiosis between the coral host (polyp) and its unicellular symbiotic microalgae (zooxanthellae). The creation of this three-dimensional framework has supported many dependent species and, over time, has made coral reefs to one of the most diverse ecosystems on Earth. Coral reefs have suffered mass extinctions throughout geologic history and the present reef ecosystems are therefore a product of only the past 45–50 Ma of evolution. The current distribution of corals is much the result of the last ice age (i.e., approximately 10,000 years ago (Kauffman and Al Fagerstrom 1993).

Herbivory and the Balance Between Corals and Macroalgae

Reef ecosystems are shaped by important ecological feedbacks (Figure 18.1; Nyström et al. 2012). Competition between corals and algae for resources such as light and space is paramount (Burkepile and Hay 2008). When in a healthy state, corals are the primary space-holders (Appendix Image Series 1a) and algae are kept under control by the constant grazing of herbivorous fish (Hughes et al. 2007). This process depends on the abundance of herbivores and the area of algae that needs to be grazed (Mumby et al. 2007). Reefs face many perturbations, such as typhoons or destructive fishing, which open up space by removing or killing corals. The loss of corals promotes algal growth and can overwhelm the grazing capacity of herbivores. This is when the regime shift occurs (Appendix Image Series 1b; Williams et al. 2001). Once macroalgae are established there is a range of feedback mechanisms that can reinforce their presence and the macroalgae-dominated state.

Coral Reefs in the Anthropocene

Despite their long history, coral reefs have suffered significant impacts from human activities (Pandolfi et al. 2003), which have resulted in worldwide loss of coral reefs (Gardner et al. 2003; Bruno and Selig 2007). The drivers causing these impacts are in many cases related to human activities, such as overfishing, pollution, and

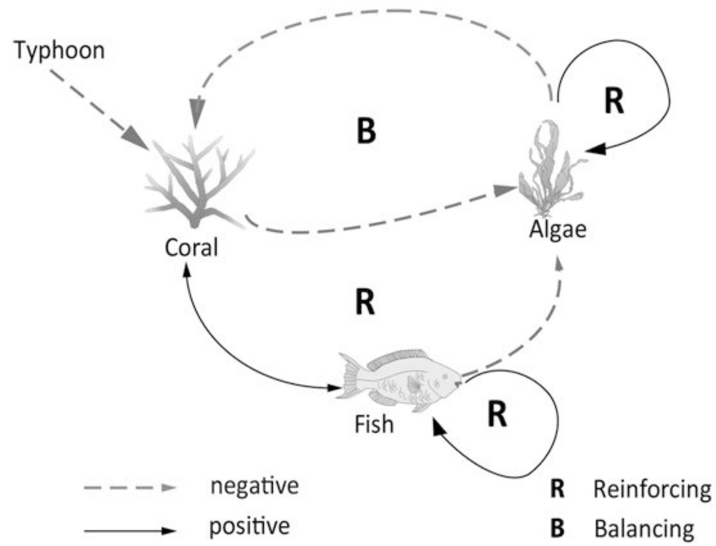


Figure 18.1 A conceptual model of the dynamics influencing whether coral or algae dominates a reef ecosystem. Note both negative (balancing) feedbacks as well as the positive (self-reinforcing) feedbacks. Images: Dieter Tracy, Tracy Saxby. IAN image library (ian.umces.edu/imagelibrary)

climate change, but they operate in tandem with natural disturbance regimes such as hurricanes and diseases (Hughes and Connell 1999; Nyström et al. 2000). Loss of resilience is making coral reefs increasingly vulnerable to these perturbations (Nyström et al. 2000). Since an algae-dominated ecosystem may no longer provide the goods and services desired by society (Moberg and Folke 1999; Burke et al. 2011), avoiding shifts to degraded states is important for societal and economic development (Nyström et al. 2012).

This lab is set in a hypothetical landscape based on the central Philippines (Figure 18.2), an area with breathtaking marine biodiversity, located in the Coral Triangle which is considered the global center of marine biodiversity. Despite their rich diversity, coral reefs in the Philippines are increasingly vulnerable to disturbance with a large human population that depends heavily on the reefs for livelihoods.

Part 1. Patch-Level Dynamics of a Reef

Model Description

We have developed a model of coral reef dynamics that explores the shift from a coral to an algal dominated state in one patch reef. The model is based on basic population dynamics where the birth rate of a population is steady, the death rate

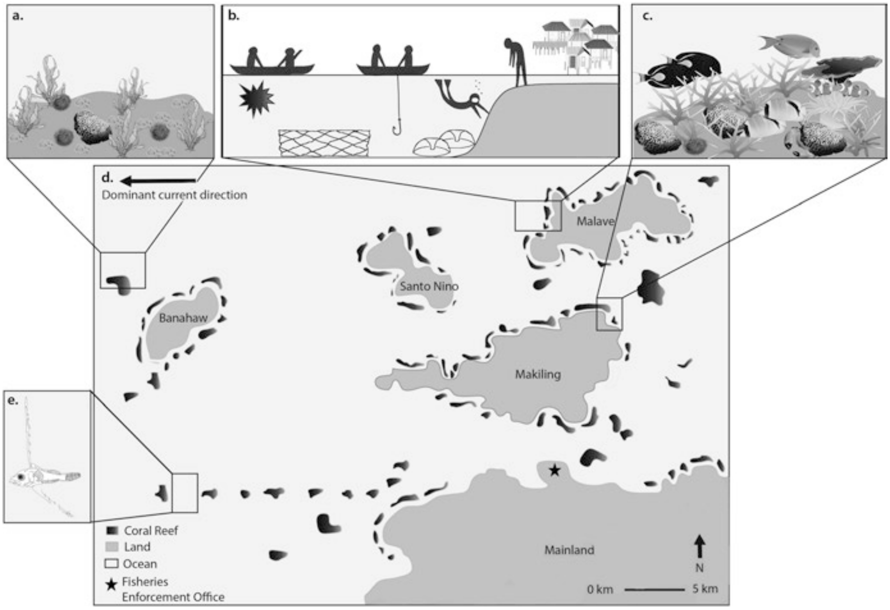


Figure 18.2 Hypothetical landscape in central Philippines with a westward ocean current. Zoom panels show different aspects of this social–ecological landscape. Panel (**a**) Degraded reef with low spatial complexity and low fish biomass. Panel (**b**) Small-scale fishers using different gears on the fringing reefs near their village. From *right* to *left*: blast fishing; nets; hand lines; traps; skin diving; gleaning. Panel (**c**) A healthy reef with high spatial complexity and high fish biomass. Panel (**d**) The marine landscape. Panel (**e**) Reef fish larvae dispersing between reefs. Fisher drawings: Danika Kleiber. Images: Dieter Tracy, Tracy Saxby. IAN image library (ian.umces.edu/imagelibrary)

increases with increasing population size, and the population size is stable, where the birth rate and the death rate intersect (Figure 18.3). The model incorporates stochastic dynamics, which are the random events that naturally occur in ecosystems such as storms and disease outbreaks. You can find this model in the **1.reef fishery** tab of the **reef.xlsx** file.

The model reef is composed of coral, algae, and a population of fish. In the model, coral growth depends upon the cover of coral, space available to colonize, and the biomass of herbivorous fish. The herbivorous fish keep algae in control and hence help to maintain high coral cover. Herbivorous fish biomass is influenced by feedbacks because fish biomass affects competition and reproduction. The maximum fish biomass is set by the amount of coral available as shelter. In a reef with low coral cover, algae will outgrow coral and dominate the reef (Figure 18.4). When fish biomass is low algae become more competitive and a high coral cover is required for corals to be self-sustaining. These dynamics mean that a reef can be dominated by either coral or algae, and that the size of the herbivorous fish population increases the resilience of a coral reef. In the model, fishing lowers the resilience of a coral reef by removing fish and coral. For the model's *Initial Conditions,* 1 represents the proportion of the potential coral cover or fish biomass for the site (i.e., a proportion of 1 = 100%).

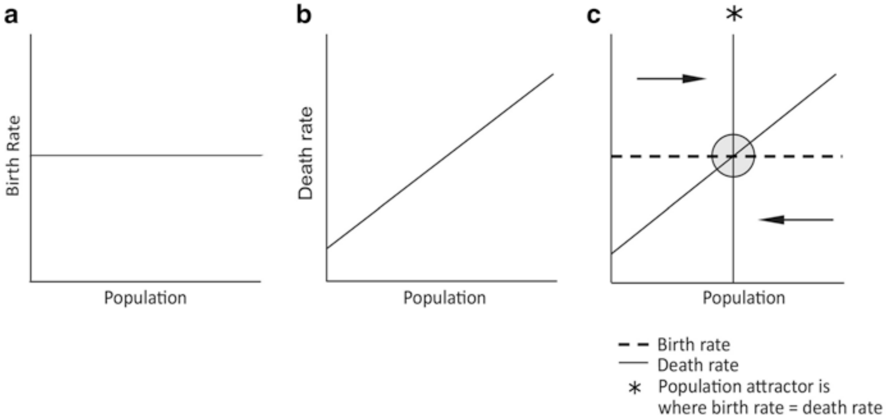


Figure 18.3 Dynamics of a population. (**a**) The birth rate is steady, (**b**) the death rate increases with population size, and (**c**) the population is stable when the birth rate equals the death rate

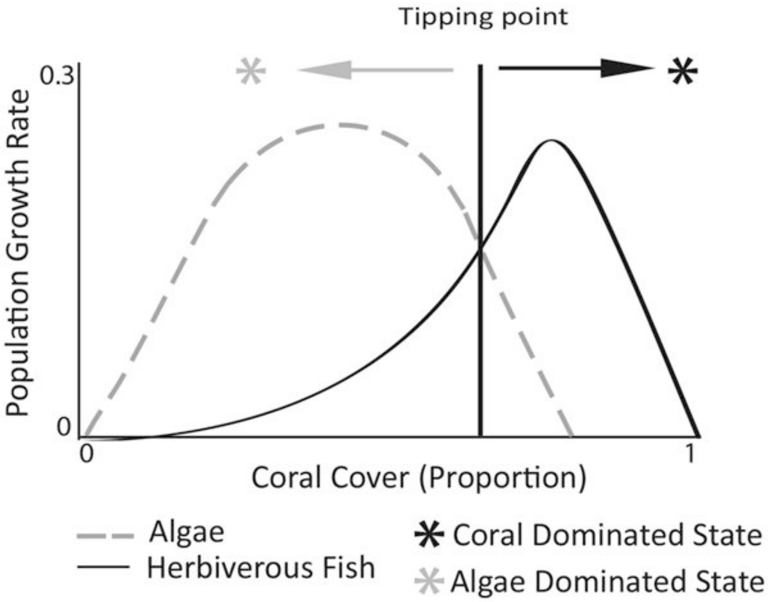


Figure 18.4 The dynamics of coral reef systems are influenced by the relationships between the growth of algae, the growth of the herbivorous fish population, and coral cover. The tipping point of a coral ecosystem exists at the balance between herbivore biomass and algal density. The system tends towards a coral-dominated state above the tipping point and towards and algal-dominated state below the tipping point

SPREADSHEET HINTS
• Download the file **reefs.xlsx** from the book web site
• Figures and Tables embedded in the reefs.xlsx file are labeled with Excel before their name (e.g., **Excel.Figure 1.1** will be in reefs.xlsx in the 1.reef fishery tab).
• Answers for all questions can be entered into reefs.xlsx. The answers you type will be added automatically to the **5.answers** tab. Hitting the update key (explained below) will update your answers.
• Update or rerun the model by pressing "F9" on a PC. • For Macs, rerun the model by pressing "command" and "=" at the same time.
• You can also rerun the model by changing the text in an unused cell and pressing Enter. On old versions of Excel for Macs "+/-" also works.

EXERCISE 1: Dynamics in a Small Coral Reef

1. Begin by using the **1.reef fishery** tab in the **reef.xlsx** file.
2. Set the initial model parameters in the **1.reef fishery** tab in the **Excel.Table 1.1 Model Parameters** table to match below (Figure 18.5).
3. To generate stochastic dynamics in the model, you can rerun the model by hitting an **update** key or set of keys. [*"***F9***"* on a PC; for a Mac use "**command**" and "**=**" simultaneously]. Each time you rerun the model, the graph in **Excel.Figure 1.1** will update. This happens because every time you rerun the model you create a new random trajectory for the reef. The trajectory is based on two things: the starting conditions and the years when typhoons occur. Typhoons are a shock to the ecosystem, which reduce resilience by suddenly removing large amounts of coral. In the graph in **Excel.Figure 1.1**, typhoons are indicated by black squares. In the same figure, the different color lines represent characteristics of corals (orange), fish (purple), algae (green), and the fishers' catch (blue).
4. Try three different **Initial Conditions** values for **Coral Cover:** 0.99; 0.65; and 0.50. Changing the initial condition value for coral initializes the system within different regimes (i.e., different dynamics of fish, algae, and coral and different responses to typhoons). Run the model **10+ times** for each initial condition and track the dynamics of the reef over time in **Excel.Figure 1.1**. Pay attention to the system's response (a) to one typhoon, and (b) to multiple typhoons that occur over short time periods.

Excel.Table 1.1 Model Parameters			
FISHING		**Initial Conditions (Proportion)**	
Yes(1)/No(0)	0	Coral Cover	0.99
Catchability (q)	0.15	Herbiverous Fish Biomass	0.98
Effort	0.99		

Figure 18.5 Start with these Model Parameter values in the **1.reef fishery** tab in your spreadsheet

Q1 Compare the trajectories of reefs with initial coral cover values of 0.99 and 0.65.

(a) What is the meaning of the two initial values for coral cover?
(b) When you change the initial coral cover, does the reef remain coral dominated or shift to algae?
(c) How many years does it take these changes to occur for fish and coral?

EXERCISE 2: Reef Fisheries

Human activities such as fishing can alter the ecological dynamics of coral reef systems. Overfishing of herbivorous fish can change the competitive balance between corals and algal (Figures 18.1 and 18.4). Different fishing gears vary in their effectiveness in capturing fish and also in the damage they cause to coral (see Appendix Image Series 3). The use of destructive gears such as **blast fishing** (using explosives) is an extremely effective way to catch many fish in a short time. However blast fishing is bad for the ecosystem and the fishery. Blast fishing kills all nearby fish and invertebrates, including species not targeted by the fishers, juveniles, and corals (Alcala and Gomez 1987). By damaging corals, blast fishing destroys the shelters and reproduction grounds of herbivorous fish (Appendix Image Series 3a; Graham et al. 2006). **Traps** are more selective in what they catch. They can cause some localized damage if they get caught in the corals, but the damage is minimal when compared to blast fishing (Appendix Image Series 3b). **Hook and Line** fishing is another commonly used gear in the region (Appendix Image Series 3c, d). However, this gear catches fewer herbivorous fish and does virtually no damage to corals. In this sense, hook and line fishing has less of an effect on the resilience of coral reefs than destructive gears such as blast fishing.

Next, you will explore the relative impact of fishing on the resilience of a small reef patch. The **Catchability (q)** parameter characterizes the efficiency of a fishing gear that is catching herbivorous fish so switching this parameter is similar to switching gear. Catchability is influenced by several factors including the efficiency of fishing gear, fish behavior, and fish biomass. In the model, catchability can range from 0 (no fish are caught) to 1 (all fish are caught). The **Effort** parameter describes the amount of time that fishers spend fishing. In this model, the effort parameter describes relative effort: 0 means no effort and 1 means fishing as much as possible.

Effects of Fishing and Catchability

- Remain on the **1.reef fishery** tab. Turn on fishing in **Excel.Table 1.1** by typing in the new **Model Parameters** shown in Figure 18.6.
- Run the model **10+ times** and track the outcomes.

Q2 What happens to the resilience of the system when fishing is added? Explore the impact of different levels of **catchability** (0.05, 0.1, 0.2) and track the outcomes.

Excel. Table 1.1 Model Parameters			
FISHING		**Initial Conditions (Proportion)**	
Yes(1)/No(0)	1	Coral Cover	0.99
Catchability (q)	0.1	Herbiveous Fish Biomass	0.98
Effort	0.99		

Figure 18.6 Use these Model Parameter values to turn on fishing in the **1.reef fishery** tab

Q3 How does the catchability of the fishing gear influence the fish population's resilience to typhoons? How could you manage fishing gear catchability to improve system resilience?

Q4 How could you manage fishing gear catchability and/or fishing effort to improve system resilience?

Part 2. Spatial Dynamics of Reefs

The resilience of an ecosystem is influenced by anthropogenic impacts. However, impacts from threats such as fishing can be camouflaged by sharing of resilience among spatially separated areas, such as neighboring reefs (Nyström and Folke 2001; McCook et al. 2009). One of the reasons that reefs share resilience is because of the movement of individual fish and invertebrates among reefs during their larval dispersal phase (Kinlan and Gaines 2003). During this phase most marine organisms, including corals, fish, and algae, travel from their birth site to recruit (i.e., settle) at a new location. This process is made possible by ocean currents. In this sense, spatial exchange of biodiversity provides some insurance against disturbances (Loreau et al. 2003).

EXERCISE 3: Linked Reefs

You will now work with a spatially explicit model that is found on the **2.spatial reefs** tab of **reefs.xlsx**. The spatial reefs model represents the dynamics for 10 reefs that are spatially linked. In this model, larval fish (Figure 18.2e) and algae disperse among reefs, but reefs at the edges receive fewer recruits than central reefs because they only have one neighbor.

1. Open the **2.spatial reefs** tab in **reefs.xlsx**.
2. Examine **Excel.Figure 2.1**. This figure shows the average value of coral, fish, algae, and catch for the 10 linked reefs.
3. Examine **Excel.Figure 2.2**. The reefs are numbered based on their distance to the village (1 = close, 10 = far). The arrows indicate that larvae from the reefs travel in both directions. There is connectivity between adjacent reefs.
4. Examine **Excel.Figures 2.3–2.5**. These figures show the spatial dynamics of coral cover, fish biomass, and catches, respectively. From top to bottom each figure shows the value at reefs 1–10. Thus, the top of the graphs are the reefs

Excel.Table 2.1 Model Parameters						
Fishing		**Catchability**		**Initial Conditions**		
Fishing: Yes(1)/No(0)	0	q (Hook & line)	0.15	Site	Coral0	Fish0
Mobile Fishing:Yes (1)/No (0)	0	q (Blast fishing)	0.6	1	0.99	0.59
Location (if not mobile)	1			2	0.99	0.85
Gear		**Coral Damage**		3	0.99	0.94
Hook & Line(0); Blast fishing (1)	0	Hook & Line	0.01	4	0.99	0.98
Effort	0.99	Blast	0.1	5	0.99	0.98
				6	0.99	0.98
Max Catch	0			7	0.99	0.98
				8	0.99	0.94
				9	0.99	0.85
				10	0.99	0.59

Figure 18.7 In Excel.Table 2.1, set these Model Parameters in the **2.spatial reefs** tab

closest to the fishing village while the bottom of the graph indicates the reefs furthest from the fishing village. From left to right, the figure shows the progression of time from Year 1 to Year 100.

5. In the **Excel.Table 2.1,** the **Initial Conditions** indicate the relative value of coral and fish at each reef (reefs 1–10) and in the first year (time = 0). The values for Coral0 and Fish0 equal the proportional cover or abundance of corals and fishes in the first year at each reef. In the first year, coral cover is high at all reefs (Coral0 = 0.99) while fish are highest in the middle reefs (Fish0 = 0.98). We will change the fishing dynamics and see how this affects the coral, fish, and catch.
6. Confirm that the **Excel.Table 2.1 Model Parameters** match those in Figure 18.7. When fishing = 0, *no other fishing parameters are turned on because fishing is not running in the model.*
7. Run the model **10+ times** with the initial conditions described in Table 18.3. Look at the spatial and temporal dynamics in **Excel.Table 2.1** and **Excel.Figures 2.3–2.5**. Notice how coral and fish respond to typhoons.
8. In **Excel.Table 2.1,** change **Site 1** and **Site 5** to algae dominated by setting **Coral0 = 0.2**.
9. Run the model **10+ times** with these new parameter values. Track the spatial and temporal dynamics.

Q5 How do algae-dominated reefs affect the resilience of their neighboring reefs to typhoons?

EXERCISE 4: Heterogeneous Fishing and Fisher Mobility

Continuing to use the model in the **2.spatial reefs** tab, we will now examine the influence of humans on resilience by exploring how heterogeneous fishing across a reef interacts with the spatial dynamics we explored above. In the model, when:

- Mobile Fishing: Yes(1)/No(0) = 0, fishing is not mobile and is restricted to shore (Location 1).

- When Mobile Fishing: Yes(1)/No(0) = 1, fishing is mobile and fishers can fish anywhere, targeting sites with the most fish.

In the model, we represent two types of fishing gear. **Hook & line** has moderate catchability and is not destructive. **Blast fishing** has high catchability and is very destructive.

1. Set the **Excel.Table 2.1 Model Parameters** to match Figure 18.8. Return **Initial Conditions** to **Coral0 = 0.99** for all sites.
2. Run the model with fishers only fishing near land (Location 1 is the reef that is adjacent to the fishing community). Run the model **10+ times** for:
   - Hook and line fishing (set **Gear = 0** and Mobile Fishing = 0)
   - Blast fishing (set **Gear = 1** and Mobile Fishing = 0)

Q6 How does the location of stationary fishing (i.e., only targeting the closest reef) influence the dynamics of the fish and coral at that reef and at the neighboring reefs?

3. In **Excel.Table 2.1** turn on mobile fishing by changing the cell **Mobile Fishing: Yes(1)/No(0) = 1**). This allows fishers to target any reef. Run the model **10+ times** for:
   - Hook and line fishing (**Gear = 0** and Mobile Fishing = 1)
   - Blast fishing (**Gear = 1** and Mobile Fishing = 1)

Q7 How does the resilience of the ecosystem change when the fishers are able to target all of the reefs?

Q8 How would you manage fishing on a network of reefs differently from an isolated reef to make the fishery more sustainable? Is it possible to make the fishery sustainable while reducing the possibility of a regime shift?

Excel.Table 2.1 Model Parameters						
Fishing		Catchability		Initial Conditions		
Fishing: Yes(1)/ No(0)	1	q (Hook & line)	0.15	Site	Coralo	Fisho
Mobile Fishing:Yes (1)/ No (0)	0	q (Blast fishing)	0.6	1	0.99	0.59
Location (if not mobile)	1			2	0.99	0.85
Gear		Coral Damage		3	0.99	0.94
Hook & Line(0); Blast fishing (1)	0	Hook & Line	0.01	4	0.99	0.98
Effort	0.99	Blast	0.1	5	0.99	0.98
				6	0.99	0.98
Max Catch	0			7	0.99	0.98
				8	0.99	0.94
				9	0.99	0.85
				10	0.99	0.59

Figure 18.8 In Excel.Table 2.1, set these Model Parameters in the **2.spatial reefs** tab

Part 3. Linking Social–Ecological Landscapes Across Scales

In the previous section, we considered how coral cover, fish biomass, disturbance, and fishing gears interact to influence the probability of a regime shift. We looked at an individual reef and at the interaction of connected reefs across the landscape.

Here, we increase the complexity of the system and examine the potential for regime shifts to occur under social–ecological conditions operating at different scales. We will take a more in-depth look at how the biological processes that operate inside individual reefs interact with the biophysical and social processes that occur across a seascape. While the calculations in the lab are simplified from the dynamics found on reefs, they provide a conceptual outline of many processes that influence reef resilience. We will consider three social factors: human population size, livelihood availability, and enforcement of fishing regulations (as a proxy for community support of sustainable fishing).

- Look at the map shown in Figure 18.2. This seascape is based on a region of the Philippines.
- Look at the printed **seascape map** or the **4.seascape map** tab in **reefs.xlsx**. The **seascape map** (**Excel.Figure 4.1**) is a raster (grid) version of the map on Figure 18.2, which you will use to do further calculations in the lab. *NOTE*: We recommend that you print the larger version of this map found in the **6.seascape map PRINT THIS** tab in the spreadsheet or **reef_map.pdf**.

Livelihoods on Islands

The four islands in this ecosystem (Figure 18.2; seascape map) are surrounded by fringing coral reefs (reefs adjacent to the island). Islanders focus their fishing on the fringing reef adjacent to their home island. Since the human population has been increasing, there are more fishers than the reefs can support. On a large island, other livelihoods such as farming or construction work are available. However, on small islands most livelihoods depend on extracting resources from the ocean and from the nearby fringe reefs, even though catches have been declining. Some fishers have responded to declining catches by turning towards destructive gears.

Q9 Based on what you learned in Part 1 and Part 2, how might each of the social factors listed below influence the resilience of a coral reef system?
(a) Human population size
(b) Availability of alternative livelihoods
(c) Community support for sustainable fishing and enforcement

- Open the **3.seascape** tab in **reefs.xlsx**. (*NOTE:* You will scroll down through the 3.seascape tab as you go along, but do not need to see the entire worksheet at one time).
- Type your answers in the boxes provided in the **3.seascape** tab. Your answers in the blue columns will be automatically added to the **4.seascape map** tab and to the **5.answers** tab.

EXERCISE 5: Patch-Scale Influence of Structural Complexity, Reef Size, and Island Size

The total number of fish on the island are influenced by the structural complexity of the coral and the total reef area (Lingo and Szedlmayer 2006; Graham and Nash 2013). In the social realm, the area of an island can correspond to the percentage of adult men on an island who work as fishers (Selgrath, unpub data).

- In **Excel.Table 3.1,** look at the relationship between reef structural complexity and herbivore density. Notice how the density of herbivorous fish changes as structural complexity goes up or down.

Calculation 1: Reef Area

Using the **seascape map**, calculate the total area for the fringing reef (i.e., patch size) associated with each island. **Each cell on the map is 1 km × 1 km. (Excel. Table 3.1)**. *(NOTE:* Some of the calculations have been done for you to save time, but be sure to look at all of the answers. The columns where you will enter answers are indicated in green).

Calculation 2: Herbivores on Entire Reef

Based on the herbivore densities of each island (**Excel.Table 3.1**), calculate the total number of herbivores found on each island's fringing reef, using the following equation. Enter your results in **Excel.Table 3.1**.

Number of Herbivores = Herbivore Density × Reef Area

Q10 Based on relative herbivore abundance, you'll make a hypothesis about which islands have higher resilience, which we will compare with the outcomes at the end of the lab:

(a) Which island's coral reefs seem to have coral states that are resilient and which seem to not be resilient?

(b) List two reasons why you made this selection.

- In **Excel.Table 3.2,** the area of each island and the number of fishers has been calculated for you. Notice how the number of fishers is similar on some islands with different population sizes.

  Number of Fishers = % fishers on island × adult male population

- In **Excel.Table 3.3**, look at the relationship between island size and the % of men who are involved in fishing.
- Fill in the number of fishers on each island (from **Excel.Table 3.2)** into **Excel. Table 3.4** to answer Question 11. (*NOTE:* Islands in Excel.Table 3.4 are ordered by size, so are not in the same order as Excel.Table 3.2)
- Refresh **Excel.Question 11 Graph**, which shows the relationship between island size and number of fishers. The Excel.Question 11 graph will also be copied into the **5.answers** tab.

Q11 Based on **Excel.Table 3.2** and the **Excel.Question 11 Graph**, how does the number of fishers vary with island size? Why do you see this pattern? Is this the relationship you expected?

Q12 How might the percentage of people dependent on fisheries influence the ability of the island communities to adapt to a changing environment?

EXERCISE 6: Landscape-Scale Factors Influencing System Dynamics

Larval recruitment and the social conditions that influence fishers' decisions about what fishing gears to use are examples of ecological and social processes occurring over broad spatial scales. Importantly, both processes may affect resilience and may vary widely across the seascape.

Ocean Currents and Connectivity

As you learned in Part 2, herbivore recruitment is influenced by self-recruitment (larvae that stay at their home reef) and external recruitment from neighboring reefs. **Nearest-neighbor distance** (here, the distance between two patches of coral) is one factor that can influence recruitment.

- Look at **Excel.Table 3.5** where the distance between each island's fringing reef and its nearest neighboring reef is calculated for you. This distance is from the fringing reef to the nearest reef in any direction.

Calculation 3: Distance to Nearest Reef
Due to a current pattern which travels from east to west, larval recruitment of herbivores from external reefs can only come from neighboring reefs that are directly eastward. Larvae born at such **source reefs** disperse with the ocean current, and recruit to **sink reefs** where they will live as adults. Calculate the distance between each island's fringing reef and the nearest reef that is also directly east. This can include the fringing reefs of other islands. Answers go in **Excel.Table 3.5**.

Calculation 4: Recruitment via Dispersal
Assume that herbivore larvae can travel up to 4 km, and only along the prevailing East → West current. Assess whether herbivore recruitment from external reefs that are larval sources occurs at each island's fringing reef. Answers go in **Excel.Table 3.5**.

- External recruitment (via dispersal) = YES if:
  A source reef is ≤ 4 km from an island's fringing reef; **and**
  The source reef is directly up-current from the fringing reef (i.e., it is eastward)

- External recruitment (via dispersal) = NO if:
  The fringing reef > 4 km from the nearest neighboring reef that is directly up-current (i.e., it is not eastward)

Q13 How does the distance between reefs interact with current patterns to influence recruitment (via dispersal)? For Banahaw Island, explain how you might predict such external recruitment patterns differently if you did or did not have knowledge about currents.

EXERCISE 7: Seascape Co-Management, Enforcement, and Fishing Gear

In 1998, the Philippine Fisheries Code prohibited the use of most destructive gears, but the use of these practices continues. To reduce destructive fishing, some fishing communities, NGOs, and municipal governments collaborated to hire boat-based fisheries enforcement officers. This co-management model of enforcement led to more successfully managed nearshore islands, but did not have a significant effect on outlying islands. Enforcement is limited by the cost of fuel because officers are not able to afford the gas to travel to distant islands. Limited enforcement and a culture that is tolerant of destructive fishing mean that destructive fishing practices persist in these outlying areas (Marcus et al. 2007; **Excel.Table 3.6**).

Calculation 5: Gear Usage
Based on their distance to the enforcement office, what fishing gears do communities use? Enter the names of the gears used by communities in **Excel.Table 3.5.** Information to answer this question can be found in **Excel.Table 3.5** and **Excel.Table 3.6**.

EXERCISE 8: Regime Shifts and Cross-Scale Interactions

As you learned in the previous section of the lab, social–ecological factors can interact across scales to influence the resilience of these linked systems. Here, we will explore the interaction of factors operating at patch and landscape spatial scales (i.e., cross-scale dynamics).

Calculation 6: Herbivores Caught per Year per Island
Based on results from **Calculation 5**, each island's annual number of fisher catch can be estimated from **Excel.Table 3.6**. Calculate the total number of herbivores caught per year for each island. Answers go in **Excel.Table 3.7** and will be automatically added to the **4.seascape map** tables.

$$\text{Herbivores Caught per Year per Island} = \text{Number of Fishers on island} \times \text{Annual catch per fisher}$$

- Examine **Excel.Table 3.7**, where the number of **Herbivores Remaining** after 1 year is the initial number of herbivores less the fish that were caught and natural mortality.

The total number of new recruits (juvenile fish) for each fringing reef is influenced by larval supplies. Recruitment is a combination of **self-recruiting** individuals (i.e., those that stay at their home reef, which is influenced by local fish populations) and individuals that recruit from up-current reefs (i.e., via dispersal; see **Excel.Table 3.7, Recruitment to Reef**). The estimated recruitment has been calculated for you and was calculated using the formula below. Only reefs with neighbors ≤ 4 km receive external recruitment.

- If Neighbors < 4 km: self-recruitment + recruitment from up-current reefs
  Recruitment = (0.25 × remaining fish population) + (700,000 × (1/distance to nearest eastward reef))
- If Neighbors > 4 km: self-recruitment only
  Recruitment = (0.25 × remaining fish population)

Calculation 7: Final Number of Herbivores on Reef

In **Excel.Table 3.7**, estimate the final number of herbivores at the end of the year after adjusting for recruitment for each island's fringing reef.

Final Number of Herbivores = Number of Herbivores Remaining + Recruitment

Calculation 8: Difference Between Final and Original Number of Herbivores

In **Excel.Table 3.7**, calculate the difference between initial and final herbivore densities on each reef after a year of fishing.

Difference = Final Number of Herbivores − Original Number of Herbivores (from Excel.Tables 3.7 and 3.1, respectively)

Calculation 9: Number of Herbivores on Reef with High Structural Complexity

In **Excel.Table 3.7**, calculate the theoretical number of herbivores expected at each reef if the reef had high structural complexity. The number of herbivores found on reefs with high quality habitat is 148,000 per km^2.

Number of Herbivores on Healthy Reef = 148,000 per km^2 × Reef Area (from Excel.Table 3.1)

Fishing decreases fish densities directly through removal and indirectly through habitat destruction. Thus, a complex reef can sustain a high density of reef fish while a fished reef that has experienced habitat destruction can support fewer fish.

Calculation 10: Percent of Herbivores Present on Reef

Determine what percent of the total possible herbivores on healthy reefs is present at each reef after a year of fishing? Place your answers in **Excel.Table 3.7**. (*NOTE*: Enter as a percentage).

$$\text{Percent of Herbivores} = \frac{\text{Final Number of Herbivores (Calculation 7)}}{\text{Number of Herbivores on Healthy Reef (Calculation 9)}}$$

The likelihood of a coral reef to shift to an algae state when hit by a disturbance (e.g., a typhoon) is affected by the abundance of herbivorous fish.

Calculation 11: Resilient Reef
Based on the percent of herbivores present on the reef from **Excel.Table 3.7**, use **Excel.Table 3.8** to predict if each island's fringing reef is resilient and hence likely to remain in a coral dominated state. Here, reefs are considered resilient if herbivore populations are more than 40% of the herbivore population size expected if the reef had high habitat structure.

$$\text{Resilient Reef} = \text{Is Calculation } 10 > 40\%\text{? (Yes/No)}$$

- Open the **4.seascape map** tab in **reefs.xlsx**. On this tab, the several answers from your calculations in the **3.seascape** tab can be found, but this time they are arranged by island. Blue indicates cells containing the answers you calculated.
- Examine how characteristics of the island vary spatially. Use the information about each island to answer Questions 14–17.

DISCUSSION QUESTIONS

Open **4.Seascape map** to answer these two questions.

Q14 Which island(s) are the most vulnerable to fishing impacts? Which island(s) are ecologically resilient (i.e., which island(s) have a combination of social–ecological factors that are keeping them from undergoing a regime shift)? How does this differ from your original predictions (based on your response to **Q10**)?

Q15 Although it may seem homogenous at first glance, this seascape is quite diverse. How can accounting for spatial variability in this or other landscapes improve our understanding of an ecosystem's spatial resilience?

SYNTHESIS

These are optional and can be assigned as homework.

Q16 We've presented a simplistic model of the relationship between enforcement and gear choice. In real situations, the use of illegal fishing gears is often influenced by complex factors such as management resources, corruption, social acceptability of illegal practices, and external actors such as migrant fishers. Discuss how one of these factors might influence the effectiveness of community efforts to manage their fisheries.

Q17 Marine Protected Areas (MPAs) are an important management tool in spatially managing fisheries. The placement of MPAs can vary based on different conservation priorities, including protecting vulnerable areas, protecting important source areas (i.e., where larvae come from), protecting areas that are the least vulnerable to climate change, and protecting areas that receive the greatest community support for protection.

(a) Considering how fishing pressure, recruitment patterns, and social situations might influence MPA placement, come up with three MPA sites for this seascape. Each MPA can include up to 5 grid cells (1 grid cell = 1 km^2).
(b) Explain where you would put the MPA on the map and describe how the placement of each of the MPAs would meet a conservation priority that you identify.
(c) You can draw your MPAs on the printed copy of the seascape.

Q18 When considering regime shifts in terrestrial environments, recruitment may be a factor, but will operate in different ways. In the case of urban forest patches, acorns are spread up to 18 km by Jays gathering acorns in distant patches and storing them in seed caches (Lundberg et al. 2008). In this way, Jays increase a system's resilience by creating greater connectivity among patches and improving the seed supply to isolated patches.

(a) Discuss similarities and differences between this terrestrial dispersal process and the coral dispersal process modeled in this lab.
(b) If you were going to make an urban protected area, how would you design the park to support this seed dispersal process?

Q19 In this lab, we have considered the resilience of a coral reef ecosystem; however, many other types of ecosystems are subject to regime shifts. Visit the regime shift database (www.regimeshifts.org) and look over other examples of regime shifts. Choose one regime shift and list:

(a) the alternate regimes
(b) the drivers behind the regime shift, and
(c) the feedbacks reinforcing the alternative states.

Q20 For the regime shift you picked in the prior question, consider the following:

(a) Over what scale(s) are drivers and feedbacks operating?
(b) Are the drivers and feedbacks are social or ecological or perhaps interlinked social–ecological?

Q21 For your chosen regime shift, use pictures (cornerstones, graphs, ball and cup diagram, etc.) and write a one page description of the ways that the drivers and feedbacks at different scales might interact to lower the resilience of one system, cause a regime shift, and then maintain the new system in a new state.

APPENDIX: CORAL REEF AND SMALL-SCALE FISHING IMAGE SERIES

Image Series 1 Two potential regimes in coral reefs: (a) shows a "healthy" coral-dominated reef with abundant herbivorous fish, habitat complexity, and high productivity, whereas (b) shows a reef dominated by macro-algae. Image Credits: Jennifer Selgrath/Project Seahorse

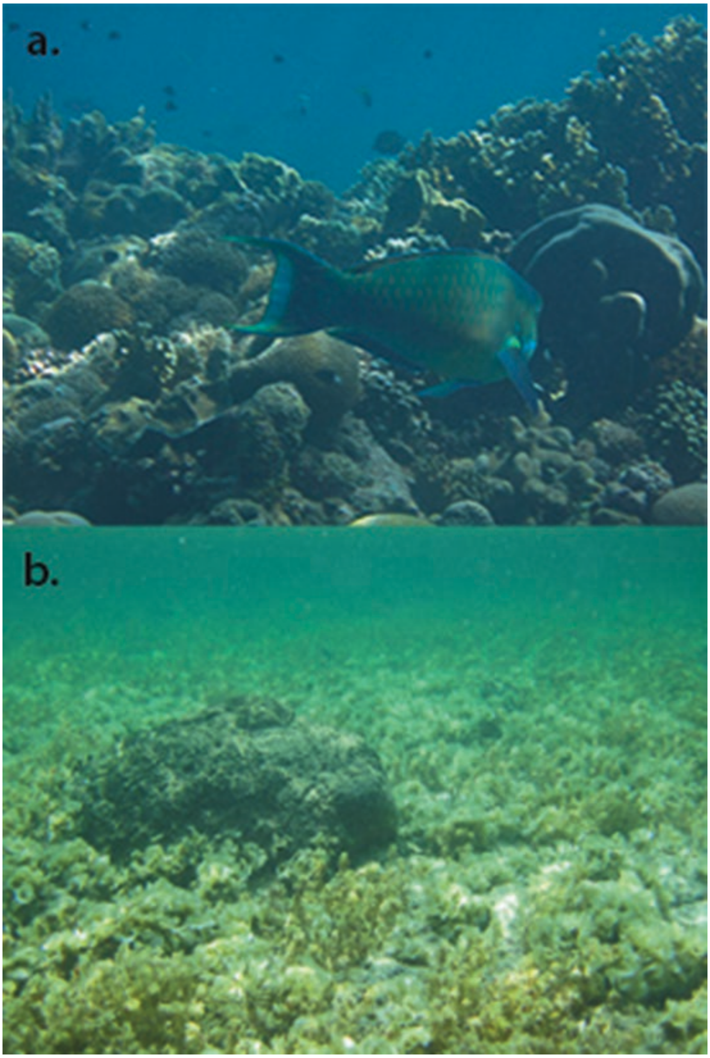
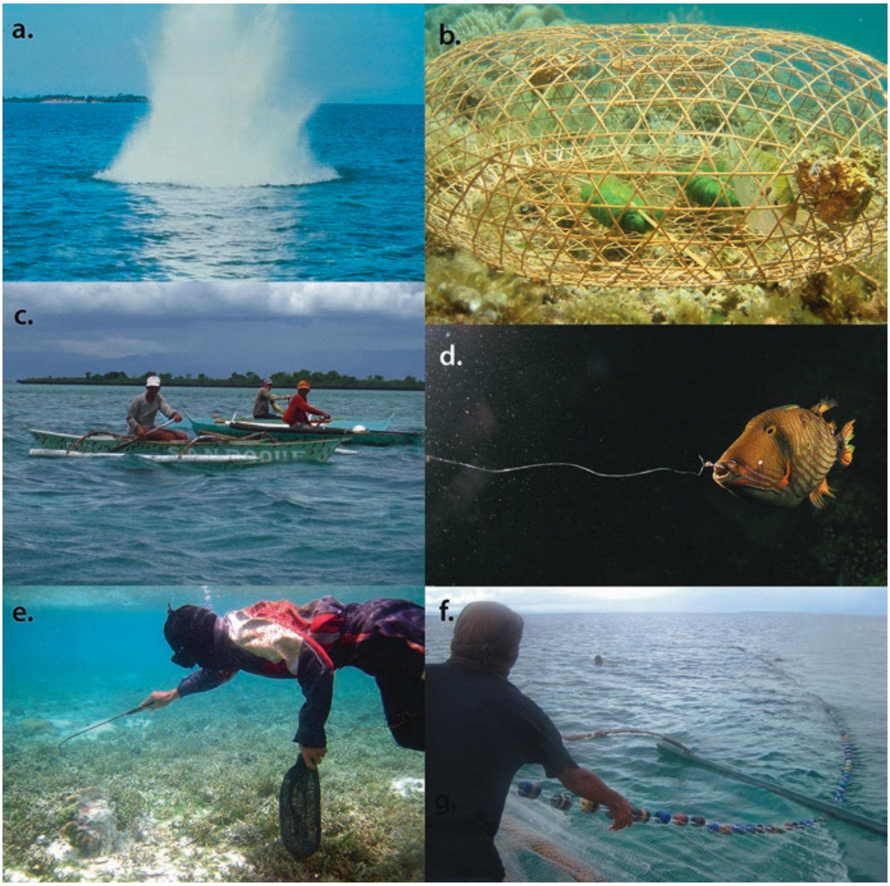


Image Series 2 Living coral polyps (a) form colonies (b) and are the building blocks of coral reefs. (c) Destruction of corals from blast fishing. Image Credits: Jennifer Selgrath/Project Seahorse

Image Series 3 Images of gears used by small-scale fishers in the Philippines: (a) blast fishing explosion; (b) fish trap with three fish inside; (c) hand line fishers paddling to their fishing grounds; (d) trigger fish caught by a hook; (e) dive fisher using a crowbar (KayKay) to pry abalone put of their hiding places in the coral reef; and (f) encircling gill net being pulled in by squid fishers. Image Credits: (a) Wolcott Henry 2005/Lynn Funkhauser; (b) Rebecca Weeks/Marine Photobank; (c, e, f) Jennifer Selgrath/Project Seahorse; (d) Lawrence Alex Wu/Marine Photobank



REFERENCES AND RECOMMENDED READINGS[1]

Alcala AC, Gomez ED (1987) Dynamiting coral reefs for fish: a resource destructive fishing method. In: Salvat B (ed) *Human impacts on coral reefs: facts and recommendations*. Antenne Museum E.P.H.E, French Polynesia

Bruno JF, Selig ER (2007) Regional decline of coral cover in the indo-pacific: timing, extent, and subregional comparisons. PLoS One 2:1–8

Burke L, Reytar K, Spalding MD et al (2011) *Reefs at risk revisited*. World Resources Institute, Washington, DC

Burkepile DE, Hay ME (2008) Herbivore species richness and deeding complementarity affect community structure and function on a coral reef. Proc Natl Acad Sci USA 105:16201–16206

Cinner JE, McClanahan TR, Daw TM et al (2009) Linking social and ecological systems to sustain coral reef fisheries. Curr Biol 19:206–212

Dublin HT, Sinclair ARE, Mcglade J (1990) Elephants and fire as causes of multiple stable states in the Serengeti-Mara woodlands. J Anim Ecol 59:1147–1164

Gardner TA, Côté IM, Gill JA et al (2003) Long-term region-wide declines in Caribbean corals. Science 301:958–960

Graham NAJ, Nash KL (2013) The importance of structural complexity in coral reef ecosystems. Coral Reefs 32:315–326

Graham NAJ, Wilson SK, Jennings S et al (2006) Dynamic fragility of oceanic coral reef ecosystems. Proc Natl Acad Sci USA 103:8425–8429

Graham NAJ, Bellwood DR, Cinner JE et al (2013) Managing resilience to reverse phase shifts in coral reefs. Front Ecol Environ 11:541–548

Holling CS (1973) Resilience and stability of ecological systems. Ann Rev Ecol Evol Syst 4:1–23

Hughes TP, Connell JH (1999) Multiple stressors on coral reefs: a long-term perspective. Limnol oceanogr 44:932–940

Hughes TP, Rodrigues MJ, Bellwood DR et al (2007) Phase shifts, herbivory, and the resilience of coral reefs to climate change. Curr Biol 17:360–365

Kauffman EG, Al Fagerstrom J (1993) The Phanerozoic evolution of reef diversity. In: Schluter D, Ricklefs RE (eds) *Species diversity in ecological communities: historical and geological perspectives*. Universtiy of Chicago Press, Chicago

Kinlan BP, Gaines SD (2003) Propagule dispersal in marine and terrestrial environments: a community perspective. Ecology 84:2007–2020

Knowlton N (1992) Thresholds and multiple stable states in coral reef community dynamics. Am Zool 32:674–682

Lingo ME, Szedlmayer ST (2006) The influence of habitat complexity on reef fish communities in the northeastern Gulf of Mexico. Environ Biol Fishes 76:71–80

Loreau M, Mouquet N, Holt RD (2003) Meta-ecosystems: a theoretical framework for a spatial ecosystem ecology. Ecol Lett 6:673–679

Lundberg J, Andersson E, Cleary G et al (2008) linkages beyond borders: targeting spatial processes in fragmented urban landscapes. Landsc Ecol 23:717–726

Marcus JE, Samoilys MA, Meeuwig JJ et al (2007) Benthic status of near-shore fishing grounds in the central philippines and associated seahorse densities. Mar pollut bull 54:1483–1494

*McClanahan TR, Graham NAJ, MacNeil MA et al (2011) Critical thresholds and tangible targets for ecosystem-based management of coral reef fisheries. Proc Natl Acad Sci USA 108:17230–33. *Quantifies resilience by comparing fish biomass to metrics characterizing the state of the ecosystem. This paper also considers the fisheries and ecosystem management strategies that correspond to the various ecosystem states.*

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

McCook LJ, Almany GR, Berumen ML et al (2009) Management under uncertainty: guide-lines for incorporating connectivity into the protection of coral reefs. Coral Reefs 28:353–366

Moberg F, Folke CS (1999) Ecological goods and services of coral reef ecosystems. Ecol Econ 29:215–233

Mumby PJ, Harborne AR, Williams J et al (2007) Trophic cascade facilitates coral recruitment in a marine reserve. Proc Natl Acad Sci USA 104:8362–8367

*Nyström M, Folke C (2001) Spatial resilience of coral reefs. Ecosystems 4:406–417. *In-depth discussion of various aspects of spatial resilience.*

Nyström M, Folke CS, Moberg F (2000) Coral reef disturbance and resilience in a human-dominated environment. Trends Ecol Evol 15:413–417

*Nyström M, Graham NAJ, Lokrantz J et al (2008) Capturing the cornerstones of coral reef resilience: linking theory to practice. Coral Reefs 27:795–809. *Discusses approaches to operationalizing (quantifying) resilience.*

*Nyström M, Norström AV, Blenckner T et al (2012) Confronting feedbacks of degraded marine ecosystems. Ecosystems 15:695–710. *Explores how human actions influence ecological feedbacks in six marine ecosystems.*

*Olds A, Pitt KA, Maxwell PS et al (2012) Synergistic effects of reserves and connectivity on ecological resilience. J Appl Ecol 49:1195–1203. *Quantifies the influence of spatial location and protected status on reef fish.*

Ostrom E (2009) A general framework for analyzing sustainability of social-ecological systems. Science 325:419–422

Pandolfi JM, Bradbury RH, Sala E et al (2003) Global trajectories of the long-term decline of coral reef ecosystems. Science 301:955–958

Peterson GD (2011) Estimating resilience across landscapes. Ecol Soc 6:17–28

Scheffer M, Hosper SH, Meijer ML et al (1993) Alternative equilibria in shallow lakes. Trends Ecol Evol 8:275–279

Scheffer M, Carpenter SR, Foley JA et al (2001) Catastrophic shifts in ecosystems. Nature 413:591–596

Walker BH, Salt D (2006) Resilience thinking: sustaining ecosystems and people in a changing world. Island Press, Washington, DC, Coral Reefs

Williams ID, Polunin NVC, Hendrick VJ (2001) Limits to grazing by herbivorous fishes and the impact of low coral cover on macroalgal abundance on a coral reef in Belize. Mar Ecol Prog Ser 222:187–196

Chapter 19
Understanding Land-Use Feedbacks and Ecosystem Service Trade-Offs in Agriculture

Lisa A. Schulte and John C. Tyndall

OBJECTIVES

Globally, a burgeoning human population and rise of a middle class are placing greater demands on our planet to produce ecosystem services than ever before. This pressure is particularly acute for current and future demands placed on agricultural land use. Farmers and other land managers are expected to satisfy existing commodity markets (e.g., corn, soybean, forage, livestock), produce options for emerging agricultural markets (e.g., biomass for bioenergy), while protecting water quality, biodiversity, and recreational opportunities. Our global citizenry needs to understand how ecosystems function in relation to the services desired and how land-use choices impact this functionality. Challenges to developing this understanding are multifold, and include complexities in landscape ecological functionality; time lags and spatial mismatches in how land-use decisions manifest ecosystem services; and cumulative impacts of multiple decision makers acting independently. Furthermore, ecosystem services are not always needed or desired by the people who own the lands that produce ecosystem services. As such, the science associated with different ecosystem services, how they interact, and how people value them is not always well understood.

In this lab, you will work with an interactive online tool "People in Ecosystems Watershed Integration," or PEWI, to evaluate feedbacks and trade-offs between agricultural land uses and environmental factors, including patterns in topography, soil, and rainfall in a realistic spatial simulation environment. Additional economic and social survey data will be used to determine how landowners and society value

L.A. Schulte (✉) • J.C. Tyndall
Iowa State University, Ames, IA, USA
e-mail: lschulte@iastate.edu

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_19

ecosystem services derived from different land-use configurations. This lab is designed to help students:

1. Understand the impacts of common agricultural land uses on the delivery of ecosystem services;
2. Explore how land use interacts with spatial and temporal factors to influence the level of ecosystem services produced;
3. Interpret, anticipate, and manage ecosystem and human system feedbacks;
4. Visualize and simulate coproduction and trade-offs among specific ecosystems services; and
5. Examine how individuals and societies value ecosystem services differently.

Working independently, students will design an agricultural watershed within PEWI to balance delivery of ecosystem services, including conventional crop and biomass production, clean water, soil carbon sequestration, and habitat for biodiversity. In Part 1, you will familiarize yourself with the levels of ecosystem services associated with different land uses and create your ideal watershed. Class discussion associated with this section will help you understand how different people value and trade off ecosystem services differently. In Part 2, you will design the watershed to produce specific, predetermined ecosystem service outcomes. Finally, in Part 3, you will use additional data to determine the economic value of the ecosystem services produced. You'll need a computer with either Google Chrome or Mozilla Firefox installed as a web browser and an internet connection to complete the lab. Because the PEWI model will be updated with new scientific findings and new modules will be developed over time, please continue with the lab online, starting here:

https://www.nrem.iastate.edu/pewi/feedbacks-and-tradeoffs

Chapter 20
Social Networks: Uncovering Social–Ecological (Mis)matches in Heterogeneous Marine Landscapes

Örjan Bodin and Beatrice I. Crona

OBJECTIVES

Ecological and socioeconomic processes often operate over different spatial and temporal scales. This can lead to increased risks of resource misuse and overexploitation if management is not well aligned with ecological processes operating in the landscape. One important way to ensure better alignment of social and ecological processes is through improved communication among relevant stakeholders operating at different scales and/or localities. Thus, understanding the structure and function of social networks is an important aspect of disentangling outcomes where different stakeholders come together to deal with natural resource dilemmas (Hahn et al. 2006; Olsson et al. 2006; Bodin and Crona 2009; Bodin and Prell 2011). For example, active successful networking of a few key actors at the onset of a resource management initiative was important for building trust and buy-in from local farmers (Hahn et al. 2006; Olsson et al. 2006). Elsewhere, external connections were key to why some rural communities were more successful in initiating economic development; a few key individuals with enough education and skills had contacts with donors and agencies outside the village. These ties to external actors with resources were crucial in differentiating successful outcomes in otherwise very similar rural Indian communities (Krishna 2002). In resource-dependent communities, particularly in the developing world, a lack of formal institutions or enforcement of regulations often means that resource users resort to informal social networks for coordinating resource use. To understand if and how social networks influence resource management, it is important to analyze both the patterns of communication but also how these patterns relate to key ecological processes in the landscape.

Ö. Bodin (✉) • B.I. Crona
Stockholm Resilience Centre, Stockholm University, Stockholm, Sweden
e-mail: orjan.bodin@stockholmresilience.su.se

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4_20

In this lab, students explore a social network of small-scale resource users targeting multiple species in a heterogeneous landscape. The study system is an artisanal fishery in a rural, coastal fishing community in East Africa. Students will learn to:

1. Analyze important characteristics of a social network of small-scale fishers in an East African coastal village;
2. Investigate the extent to which different personal characteristics (attributes) coincide with patterns of social relationships;
3. Analyze how patterns of social relations among resource users can be tied to the geographic distribution of resource extraction;
4. Discuss possible implications of their network results with regard to social and ecological (mis)matches; and
5. Gain exposure to commonly used software within the field of social network analysis.

To accomplish the above objectives, this lab is divided into three exercises. In Exercises 1 and 2, students learn to analyze real social network data in conjunction with information on personal attributes of fishers, such as occupation, ethnicity, age, and education. In Exercise 3, these social networks are examined relative to maps of different fishing localities throughout the landscape. All network-related analyses will use NetDraw (available as a free trial version when downloading the software program Ucinet). Prior familiarity with graph theory is assumed; thus, Chapter 12 in this book is a prerequisite for this lab. For those less familiar with small-scale fishing communities, Chapter 18 can also be a very helpful complement. All the necessary files (and links to software) are accessible from the book website.

INTRODUCTION

The ways in which natural resources are extracted (and potentially misused) by societies is a result of multiple socioeconomic processes (e.g., economics, poverty, culture, and tradition) as well as other structures (e.g., institutions that guide resource use, access to roads and markets). Another important factor affecting resource use and extraction is the amount and quality of information and knowledge available to resource users and other stakeholders. For example, a fisher unaware of the phenomenon of climate change and how it affects coral reefs cannot incorporate this consideration into his decisions about how much to fish or which gears to use to avoid further damaging the corals.

One can categorize the knowledge and information about the environment for resource extraction into two different but somewhat overlapping categories: (1) knowledge on how to maximize harvest while minimizing effort and (2) knowledge on how to extract resources in accordance with the natural limits posed by an ecosystem. The first category of knowledge, in a fishing context, would correspond to questions such as: "*Where are the fish?*"; "*What gears should I use to catch the fish?*"; and "*How do I deploy those gears in the most efficient way?*" The second category is more concerned with how extraction can be done without negatively affecting future use of the resource. In a fishing context, this could correspond to understanding fish stock regrowth limitations (e.g., maximum sustainable yields), how different fish stocks migrate between different localities during different life stages, and how the targeted fish species interact with or depend on other species and the physical environment. Knowledge of these processes reduces the likelihood of overexploitation. Yet it is important to acknowledge that even with good knowledge of a resource, overexploitation can and will often occur due to socioeconomic dilemmas like strong economic incentives and acute poverty (see Ostrom 1990). In such situations, the need to feed a hungry family will most often override any long-term concern for the biodiversity or resource sustainability (Barrett et al. 2011). In conclusion, both types of knowledge are important in achieving sustainable natural resource use. Thus, ways to create a better understanding of different social structures and processes that create, maintain, and distribute information and knowledge in a natural resource management setting are important.

Informal Social Networks as a Conduit for Resource Users' Communication

What can a social network perspective add to our understanding of resource governance issues? To answer this, let us first briefly define what we mean by network analysis. A **network** can be thought of as a set of **nodes (actors)** and their **ties (relations)**. Network analysis is thus the study of social relations among a set of actors. A pair of actors that have relations can be said to share a social tie. Whereas mainstream social science is concerned with attributes of individual actors (e.g., income, age, sex), network analysis is concerned with attributes of pairs of individuals and the relations between them. These relations can be categorized into kinship (such as brother of, father of), social roles (boss of, friend of), affective relations (such as likes, dislikes), and actions (talks to, or attacks). This relational approach can add several important aspects to our understanding of resource management.

First, in any resource governance/management setting people collaborate and interact. Who is included or excluded from deliberation processes or decision-making can be important for management outcomes. Such patterns can be uncovered using a network approach. For example, in situations where users share resources, such as a fish stock, it is important that they are all willing and able to agree on and abide by common rules limiting resource extraction (Ostrom 1990, 2005). From a network perspective, one could argue this is more easily achieved when users are socially well connected, as opposed to in isolated groups without much communication. A well-connected group is more likely to agree on what rules need to be developed. Social connectivity could also make it easier to monitor fellow resource users to report or sanction rule-breakers.

Second, many resources constantly flow between people in any social setting, be it information, knowledge, capital, new ideas, etc. All these resources impact how people behave and choices they make and as such will influence resource management outcomes. However, the flow of resources between people is seldom homogenously distributed among all actors. Some people will share more resources with others, and some less. Some actors will have many ties, and thus access to many resources, while others will have few.

Imagine a class of high school students. A few students tend to be extra popular and have lots of friends. The vast majority of students may have a somewhat smaller number of friends, while often classes tend to have one or two students who do not socialize with anyone in the class. Translating this into network terms, we could say that the popular students would be hyperlinked, and the students without class friends would be considered isolates. Now imagine you arrived new to this class and wanted to quickly build up relations. By befriending a popular student, and getting invited to an upcoming party, your exposure to new people has increased very rapidly. If, on the other hand, you first become acquainted with one of the "isolated" students, you would most likely not be invited to the party. Within the context of the class, this student is unlikely to introduce you to any further friends. A central person, such as a popular student, has a certain amount of power to broker contacts, and depending on the social atmosphere in the class, acceptance or rejection by such a central student could greatly affect your future social network in the class. Actors playing such roles as **hubs** and "gatekeepers" exist in many networks and can be important in facilitating or impeding flow of resources throughout the network. In this lab, we are particularly interested in how the heterogeneous distribution of relations affects flow of information and knowledge among different actors in a network and how this could impact the two different types of knowledge outlined at the start.

By now, it should start to become apparent that mapping and analyzing the patterns of social relations among a set of resource users is useful in assessing and interpreting information flows and knowledge generating processes in a natural resource management context. In addition to the connections among actors, the characteristics (or attributes) of each actor are also important. This information can be vital in trying to understand why certain subgroups appear, if central actors tend to share some common feature or skill, and what this could mean for the study system.

Social Networks and Ecological Processes in a Rural Fishing Village in East Africa

With our short introduction to basic social network concepts in mind, let us now turn briefly to the type of networks in focus in this lab. Here, you will examine a social network of small-scale fishers who share information and knowledge about the state of the natural environment (the marine system) as well as extractive fishing practices. The fisher are all residents of a rural fishing village located on the Kenyan coast (see Crona and Bodin 2006 for a full description). The village has approximately 200 households and an estimated 1000 inhabitants. The surrounding area has approximately 5 km^2 of mangroves with mudflats and seagrass meadows in the shallows and fringing coral reefs outside the lagoon. The use of resources in the village is centered on fishing, and in this lab a social network consisting of all 85 households (self-identified fishing households) is in focus. The study area represents a spatially heterogeneous landscape where many different species are harvested by many actors.

The high levels of heterogeneity and complexity in this system poses many management challenges, including: (a) many different groups of fishers are actively using the resource; (b) enforcement of regulations is weak due to limited governmental financial resources as well as difficulty in monitoring fishing occurring both day and night (making it difficult to predict when fishers will land their catch at the beach and thus assess their catch); (c) contrary to many developed-world industrial fisheries, this is a multispecies fishery, thus its management requires knowledge of not just one, but many different fish stocks (which are composed of species which also compete and prey upon each other). These challenges are similar to ones found in many terrestrial systems, such as small-scale agricultural landscapes or forests owned and managed by multiple different actors. Thus, the social network approach used in this lab is relevant to many other types of social–ecological landscapes.

EXERCISES

EXERCISE 1: Visualizing the Social Network

The dataset used in this lab consists of a set of social relations among small-scale fishers whose relations are used to exchange information and knowledge on issues related to the natural environment. Communication can occur in different ways. Some fishers work together on a boat may spend long hours at sea with ample time to share ideas and knowledge. Others may sit together after returning from the sea, discussing issues while mending their nets or enjoying a cup of tea at the local shop. However communication occurs, these types of informal social relations form the basis for understanding how knowledge and information flows through the community of fisher.

In this exercise, we will visually analyze the social network of these fishers. Each respondent is assumed to be the head of a fishing household, and thus no two fishers in the dataset are from the same household. The network data are in a file format used by the software program **Ucinet** (Borgatti et al. 2002) which is one of the most commonly used software packages for analyzing social networks. In this exercise, we use the helper application **NetDraw**, which accompanies Ucinet, to visually present the network in different ways.

Spring Embedding Network Visualization

There are many different techniques to visually present networks. A commonly used method is the **spring embedding** technique, which is a layout algorithm where each tie is treated as a spring which pulls actors towards each other (and the absence of a tie acts as a repelling spring pushing actors apart). All attracting and repelling forces of the ties are considered together and the nodes (actors) arranged accordingly. Other visualization methods often build on this simple technique. In using the spring embedding technique, actors who are very central (with ties to many other actors) tend to be arranged in the middle, whereas less connected actors end up in the periphery of the plot. Subgroups (characterized by the fact that subgroup members have more ties among themselves than with others in the network) are arranged as clusters in the visualization of the network.

In the data for this lab, all ties are **binary** (present or not), **undirected** (we do not consider who in the pair had named the other, as long as one actor has named another we consider there to be a tie), and **unweighted** (we do not try to estimate tie strength, such as by asking actors how often they interact). Social network analysis often *does* include directed and weighted ties and use of such data will have implications for the interpretation of the results. While we do not delve any deeper into to this here, for more advanced analysis we refer readers to SNA text books such as Wasserman and Faust (1994). To get started, follow these steps:

- Start **NetDraw**.
- Open the network datafile **Fisher.##h** using the pulldown menu **File** then **Open** then **Ucinet dataset** then **Network**.
- Choose **Layout** then **Graph theoretic layout**. A pop-up window labelled Spring Embedding will appear and click **OK**.
- Study the visual representations of the social network, answer, and/or reflect on the following questions.

Q1 To what extent is the community of fishers in contact with each other?

Q2 What is the relative connectivity of the fishers in the network? Why are some more central (i.e., more connected) than others?

Q3 Based on the patterns of relations observed here, to what degree do you think the community of fishers would able to come together and agree on common measures to regulate fish extraction?

Visualize Attributes of Fishers

Not all fishers in the village are the same as they typically fish at their particular favorite fishing grounds, specialize in different gears, and are embedded within different socioeconomic contexts. As a result, one might expect differences in knowledge and experiences among fishers. However, this does not mean that one fisher's knowledge is not helpful to another. In fact, the usefulness of others' experiences is particularly salient in the context of complex ecosystem dynamics. For example, accounting for multiple species interactions and migration might be crucial to successful fishing. Exchange of information and knowledge among different fishers would, at least in theory, provide for better opportunities to acquire a better and more holistic understanding of the underlying ecosystem.

NetDraw can be used to visually present the different attribute values of the fisher (the nodes). This is a powerful way to get a first impression of how and if some attributes coincide with structural features of the network. In order to explore this further:

- Open the attribute datafile **FisherAttributes.##h** and make sure it is opened as an **attribute** datafile.
- Use NetDraw's drawing abilities to show the attributes in different colors, sizes, and shapes according to the attribute values. Use the pulldown menu items **Properties** then **Nodes** then **Symbols**
- Consider the meaning of attributes in Table 20.1.
- Study the visual representations of the social network, answer, and/or reflect on the following questions:

Table 20.1 Description of fisher attributes

Attribute	Description	Values
Gear	What gear is the fisher's primarily fishing gear/method?	7 Middle man 11 Gill net 12 Spear gun 13 Hand line 14 Deep sea 15 Seine net
Religion	What religion?	0 Unknown 1 Islam 2 Christianity
Lived in village	How long has the household been resident in the village?	0 Unknown 1 0–5 years 2 5–10 years 3 10–20 years 4 >20 years
Number of child	How many children in the household?	
House type	What type of roof of the household's house (an indication of how wealthy the household is)	0 Unknown 1 Mud and thatch 2 Cement and thatch 3 Cement and iron plates
Age	How old is the fisher	
Tribe	To which tribe does the fisher belong	
Outside relative	How many close relatives outside the village does he/she have?	0 No relatives 1 Only relatives that are local (i.e., close to the village) 2 At least one relative from Tanzania

NOTE: The term "middleman" refers to the profession "fish monger" (not a gear type). Such individuals do not fish but are traders who buy fish at a landing site. They were included in the analysis as many have previously been active as fishers and thus are very knowledgeable and interact directly with the fishers on a daily basis.

Q4 What attributes are shared among the most connected fishers?

Q5 How are fishers with similar attributes connected in the network? Do they cluster together or are they dispersed throughout the network?

Q6 Can you think of any implications—in terms of fishers' understanding of ecosystem dynamics—arising from either of these patterns (clustered vs. dispersed)?

EXERCISE 2: Analyze the Fisher Network

In this exercise, we will start to formally analyze some structural characteristics of the network. We will start looking for highly central individuals and then explore how different subgroups can be identified.

Identify Central Actors

In most networks, there will be a smaller number of actors who are significantly more connected than others (captured by their **degree centrality**, Figure 20.1 Panel A). In a communication network of ecological knowledge, an actor with a high degree centrality could be influential since many people turn to him/her to access information about the natural environment. Another form of centrality is **betweenness centrality** which is the extent to which an actor indirectly connects other actors in the network. It indicates the potential of the actor to act as a channel for flow of information as well as other resources such as ideas and disease. Actors with high betweenness can be crucial in bringing different subgroups socially closer to each other by acting as bridges. The dark grey actors in Figure 20.1 Panel B have higher betweenness centrality than the others in the subgroups. Some basic but fundamental analyses of node centrality can be done using NetDraw (and even more centrality analyses are available using Ucinet):

- In NetDraw, centrality measures can be found in the menu **Analysis** then **Centrality measures**. Select both **degree** and **betweenness centrality**. NetDraw will then create new node attributes using the centrality scores for each node in the network.
- Use this new attribute data to draw the size of the nodes in accordance with their degree centrality scores. Use the menu item **Properties** then **Nodes** then **Symbols** then **Size** then **Attribute based** and select the **degree centrality** attribute.
- Other attributes can be simultaneously visualized using a different node color. For example, set each node's color based on its **gear type** attribute value.
- Repeat the above steps procedure for (Freeman's) **betweenness centrality** before reflecting on the questions below.

Q7 What attributes are shared among the most connected fishers (i.e., the ones with the highest degree centrality)?

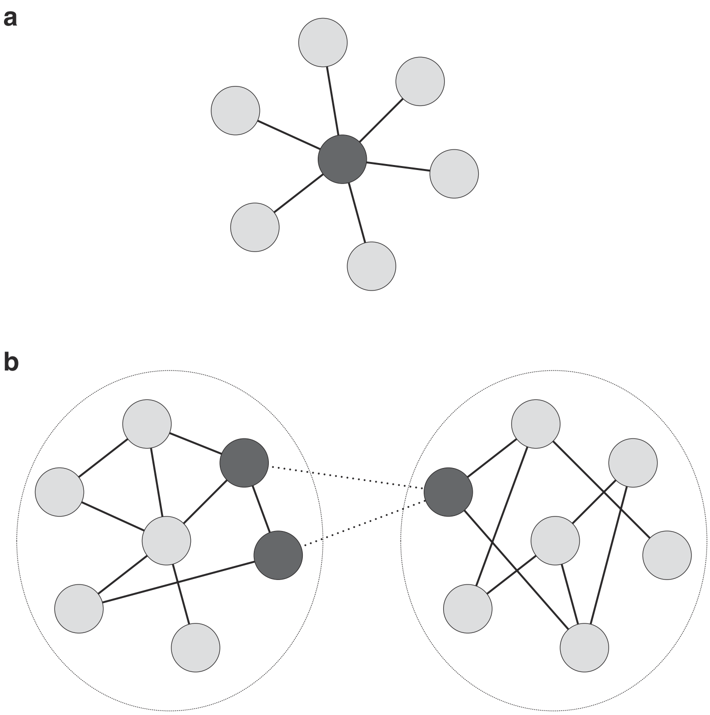


Figure 20.1 Illustration of network characteristics. In *Panel A*, the network is highly centralized with an actor in the center with a higher degree centrality (i.e., higher number of links) than the others. In *Panel B*, the network is composed of two distinct subgroups (surrounded by circles). These subgroups are interconnected through links (*dotted lines*) connecting actors (in *dark grey*) who serve as bridges between the subgroups

Q8 Who are the fishers with the highest betweenness centrality?

Q9 Why do you think fishers with the highest degree and betweenness centrality could be important for the community's ability to manage their fish stocks?

Identify Subgroups

In many real-world social networks, actors tend to be clumped into subgroups (Figure 20.1 Panel B). A few actors in subgroups may have ties to members of other subgroups thus connecting the network as a whole (Figure 20.1 Panel B, dotted lines). Specific subgroups appear for different reasons related to the type of network under investigation. For example, in a friendship network they could represent cliques of close friends. In a network of fishers, it could represent a group of fishers using similar gears.

Numerous different analytical methods are available to identify different subgroups in networks. In this exercise, we will focus on **relationally** defined subgroups. Relationally defined subgroups are defined based on subgroup members being socially tied in similar ways and is in contrast to subgroups defined by some common **attribute** value (like being members of the same tribal group).

NetDraw can be used to identify subgroups (and even more methods are available using Ucinet). Here, we use a clustering algorithm which divides the network into partitions consisting of different subgroups according to a **"Q" score** (aka **Modularity**, see Girvan and Newman 2002). Higher values of Q indicate a network composed of more distinguishable subgroups.

- Use the menu **Analysis** then **Subgroups**. Select a high value for the maximum number of clusters (e.g., select 30).
- Visualize subgroup membership using different colors (as in the previous exercise).
- Choose the partition with the highest Q-value. Notice how the Q-value increases as the number of groups increase up to a certain point, and then typically declines when the number of groups further increase.
- Layout the network grouped by the subgroup membership attribute (**Layout** then **Group by Attribute** then **Categorical Attribute** menu item), and then try visualizing some other node attribute using different colors (such as type of gear).
- Consider if some of the attributes seem to coincide with the subgroup partitioning and the answer the following questions.

Q10 Can you find any distinguishable subgroups in the network?

Q11 What attributes distinguish these subgroups?

Q12 What do you think could be the cause of the patterns you've just observed? How could it affect governance of the resource?

EXERCISE 3: Compare Social Networks and Ecological Data

In this final exercise, the social network will be used to (qualitatively) analyze how the network and the ecological processes in the study area "match up," and explore possible consequences for natural resource governance. Our goal is to determine if there is a good "fit" between the flow of information among the fisher and ecological flows between different localities and species therein.

Analyze the Spatial Heterogeneity of the Seascape

The fish species targeted in the seascape represent a fairly heterogeneous mix of species at different trophic levels (grazers, benthivores, predators) which also exhibit different spatial distributions. Figure 20.2 illustrates the distribution of some major targeted species. Different types of fishing gear are typically used at different locations (Figure 20.2), and Table 20.2 describes these fishing gears along with the

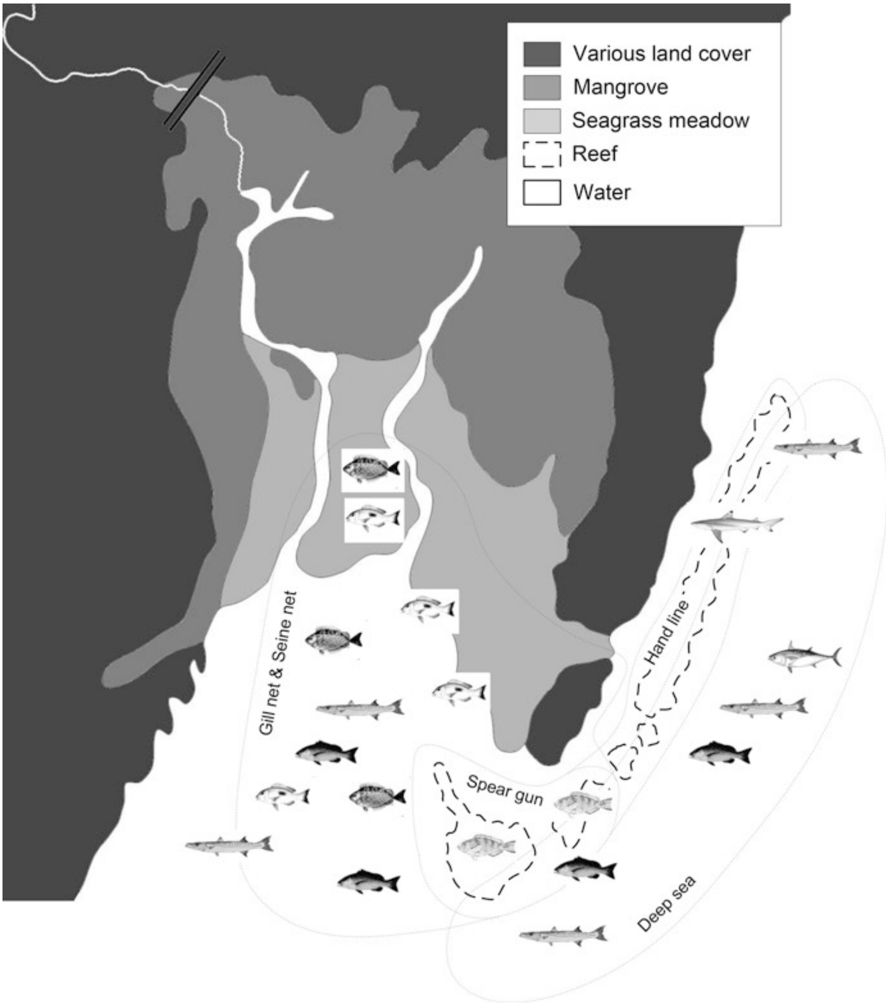


Figure 20.2 Seascape map depicting fishing grounds according to type of fishing gear used. The localities used to catch targeted fish species are also identified. For printing, this image is available from the book website as **seascape.jpg**

major species targeted by each gear. Images and key characteristics of the major targeted species are also provided in Table 3 (from the book website).

- Obtain a printout of Figure 20.2 (or **seascape.jpg**) as well as Table 3. Both are available from the book website.
- Outline the spatial distribution of all the fish species *individually*, by circling the areas where they are found.
- Identify areas of overlap between these areas and the gear-defined fishing grounds (shown as dotted lines).

Table 20.2 Description of fishing gear and some species targeted by these gears

Fishing gear	Description
Gill net	A net which is commonly used as a set net, i.e., it is not dragged actively in the water. Mesh size can vary.
	Main species targeted: *rabbitfish, emperor, snapper, barracuda*
Spear gun	A contraption (often home made) which, when deployed, releases a spear. Mainly used for larger individuals of varying species, but primarily used on the reef.
	Main species targeted: *parrotfish, snapper*
Hand line	A line, usually with one hook, often dragged behind a canoe but also larger vessels. Hooks are often bated with smaller fish or squid depending on species targeted.
	Main species targeted: *barracuda, sharks, tuna-like species (e.g., kawkawa)*
Deep sea (purse seine)	A large net, usually of medium mesh size. It is deployed using two vessels. Divers in the water identify a school of fish and the vessels circle the school. Divers dive down and tie off (close) the bottom of the net which is then dragged onto the boat.
	Main species targeted: *mainly semi-pelagic species such as kawakawa, but also snappers, barracudas*
Seine net	Usually, a fairly small-meshed net of varying size. Can be very large. It is deployed out in the open water in the lagoon. A number of people proceed to haul the net onto the beach, while fishers in the water assist by making sure the net does not get caught on the bottom. The net tends to scrape the bottom substrate (much like a trawl) and catch species and individuals of all sizes, and is therefore considered a destructive gear.
	Main species targeted: *rabbitfish, sea grass parrotfish (not the same as S. ghobban found on the reef), as well as juveniles of snappers, barracudas, emperors, and semi-pelagics like kawakawa.*

Additional information and images of fish species can be found in Table 3 an online resource

- Species being targeted by one or several types of fishing gears can be distinguished (also use information in Table 3).

Q13 Which species are targeted using multiple gears, and which species are targeted by only one specific gear?

Q14 What resource management implications might arise when different fishers using different gears target the same species?

Alignment of Social Networks and Ecological Processes

Next, we consider how well the social network and the heterogeneous seascape "match up." This is a group exercise so these questions should be discussed in small groups. Each group will then present their insights to the rest of the class. One initial way to consider how well social networks among resource users and ecological patterns/processes in land- and seascapes align is to simply overlay and visualize them

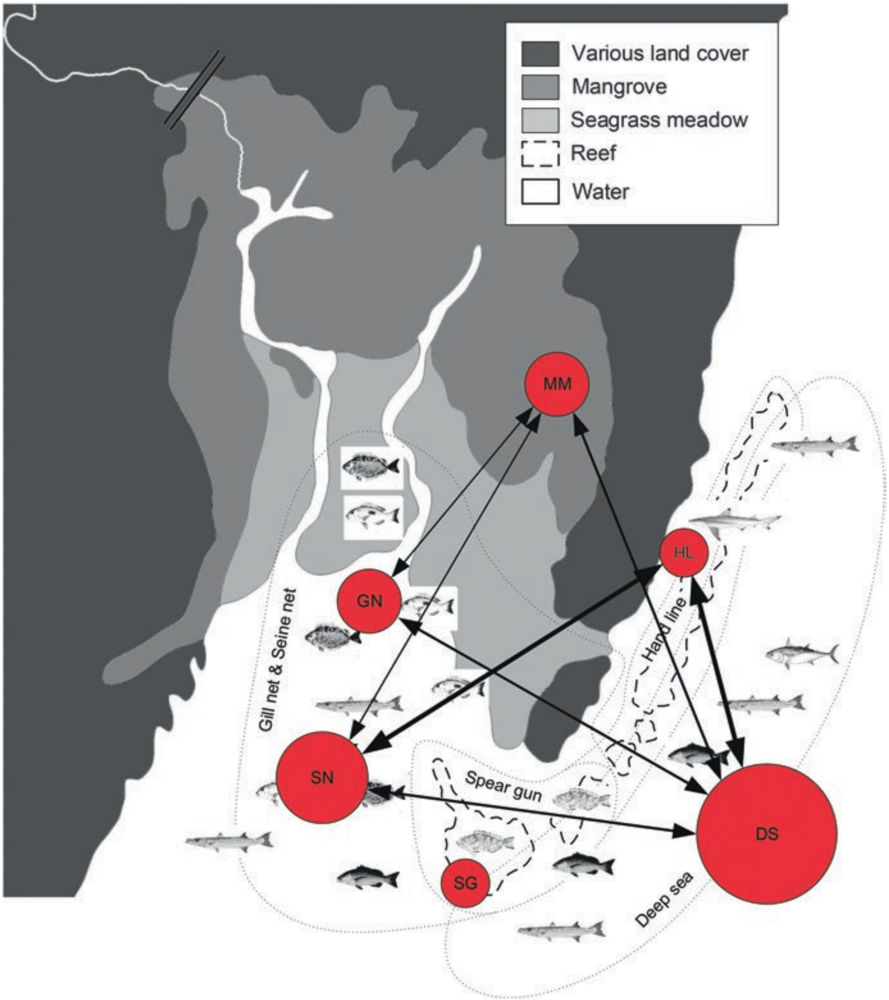


Figure 20.3 The social network of fishers relative to the seascape and fishing grounds. Nodes correspond to subgroups defined by gear type used (and node size is proportional to size of subgroup). The figure helps depict the extent of communication among subgroups of fisher using different gears who may target different (and sometimes similar) fish species in different locales throughout the seascape. *DS* = Deep sea, *GN* = Gill net, *SN* = Seine net, *HL* = Hand line, *SG* = Spear gun, *MM* = Middle men

across the landscape. A compressed simplified version of the complete social network has been drawn atop the seascape (Figure 20.3) showing subgroups based on gear type. (*NOTE:* if interested, see additional material from the book website for how to do this). Previous research (Crona and Bodin 2006) has shown that fishing gear type often correlates with *relationally* defined subgroup of fishers (as you may also have discovered in the second part of Exercise 2). Because of this identified correlation, it makes sense to divide the network into gear-defined subgroups. In this new network, each node represents a gear type subgroup and any ties represent subgroup interaction.

Use Figure 20.3, along with your answers and reflections from the first part of Exercise 3 to consider the concordance between different social and ecological processes and its implications:

Q15 Which gear-defined subgroups are communicating more than others, and with whom?

Q16 Are some fishers more (or less) decoupled from the other groups? Who would gain (or lose) most from increased knowledge and information exchange?

Q17 Consider the shared knowledge of fish species across different scales and localities. Which gear-defined groups are likely to have similar knowledge? Take into account subgroup communication as well as whether subgroups are targeting the same species.

Q18 Can you imagine any potential conflicts between different fishers targeting the same species (at similar or different localities)? Or between groups targeting interdependent species (e.g., predator and prey)?

Q19 To what extent might potential conflicts between subgroups of fishers (such as targeting the same species), coincide with social relations? What might the implications of such overlap (or lack thereof) be for conflict resolution?

SYNTHESIS

Let's take a step back and broaden our perspective and consider the following:

Q20 Consider the social and the ecological parts of the coupled social–ecological system (SES) we studied here—do they "match up" or align? Are there any apparent mismatches?

Q21 To what extent are the issues and themes of the lab specific to the context of small-scale fisheries vs. other heterogeneous social–ecological landscapes such as small-scale agricultural systems? Watershed management? Urban systems, forests, or parks and nature preserves? Choose another type of social–ecological landscape and explain.

Q22 Consider the shared knowledge of fish species operating over different scales and localities and think particularly about the kind of knowledge outlined in the introduction where focus was on understanding ecosystem processes. Which subgroup would you judge as having the most advanced understanding of ecosystem processes? Why?

Q23 Consider if some fishers increased their fishing efforts towards specific target species and specific localities. Can you identify any particularly vulnerable species and/or localities?

CONCLUSIONS

In this lab, we explored how the communication networks of resource users align with species distribution patterns to explore how patterns of social communication match ecological processes. We used a fishery example, but any human-dominated landscape could be analyzed in a similar fashion. The comparison of social networks and ecological processes was, in this lab, qualitative. More quantitative approaches are possible (examples include Janssen et al. 2006; Cumming et al. 2010; Johnson and Griffith 2010; Bodin and Tengö 2012; Guerrero et al. 2015; Treml et al. 2015; Bodin et al. 2014).

Resource extraction behavior and knowledge generation are inherently social processes affected by social embeddedness (Johannes 2002). How knowledge of ecological processes is distributed throughout community networks, and how this knowledge is translated into institutions that regulate extraction, are both impacted by social processes such as influence and diffusion, and are crucial for understanding spatial mismatches. The dilemma of the commons (Hardin 1968; Ostrom 1990), which can result in overexploitation and resource depletion, can be overcome through social collaboration and development of extractive norms (Ostrom 1990, 2005). A prerequisite for such collaboration is some basic communication and knowledge sharing to forge a collective understanding of the status of the system to be managed. Hence, communication between groups targeting similar species and operating in overlapping locations is of vital importance to avoid resource degradation and enhance the capacity for sustainable management.

REFERENCES AND RECOMMENDED READINGS[1]

Barrett CB, Travis AJ, Dasgupta P (2011) On biodiversity conservation and poverty traps. Proc Natl Acad Sci USA 108(34):13907–13912

*Bodin Ö, Crona BI (2009) The role of social networks in natural resource governance: what relational patterns make a difference? Glob Environ Chang 19:366–374. *A review paper that gives an overview of how and why social network analysis might be useful in studying natural resource management.*

Bodin Ö, Crona B, Thyresson M, Golz A-L, Tengö M (2014) Conservation Success as a Function of Good Alignment of Social and Ecological Structures and Processes. Conserv. Biol. 28, 1371–1379. doi:10.1111/cobi.12306. *A comparative case study, where one of the two cases is*

[1] *NOTE:* An asterisk preceding the entry indicates that it is a suggested reading.

the Kenyan fishery in this lab, using the social-ecological network approach presented in Bodin & Tengö 2012.

*Bodin Ö, Tengö M (2012) Disentangling intangible social-ecological systems. Glob Environ Chang 22:430–439. *Presents a theoretical and methodological framework for analyzing coupled social- and ecological networks. An analytical approach on how social-ecological systems can be described and analyzed as social-ecological networks.*

Bodin Ö, Prell C (eds) (2011) *Social networks and natural resource management: uncovering the social fabric of environmental governance*. Cambridge University Press, Cambridge

Borgatti SP, Everett MG, Freeman LC (2002) *Ucinet for windows: software for social network analysis*. Analytic Technologies, Harvard

*Crona BI, Bodin Ö (2006) WHAT you know is WHO you know?—communication patterns among resource extractors as a prerequisite for co-management. Ecol Soc 11:7. *In this study the basic context and some findings relating to the ecological knowledge of the studied fishermen community are presented.*

Cumming GS, Bodin Ö, Ernstson H et al (2010) Network analysis in conservation biogeography: challenges and opportunities. Divers Distrib 16:414–425

Girvan M, Newman MEJ (2002) Community structure in social and biological networks. Proc Natl Acad Sci USA 99:7821–7826

Guerrero AM, Bodin Ö, McAllister RRJ, Wilson KA (2015) Achieving social-ecological fit through bottom-up collaborative governance: an empirical investigation. Ecol. Soc. 20, 41. doi:10.5751/ES-08035-200441. *Investigates if a bottom-up collaborative conservation initiative has led to adequate social-ecological alignment ("fit").*

*Hahn T, Olsson P, Folke C et al (2006) Trust-building, knowledge generation and organizational innovations: the role of a bridging organization for adaptive comanagement of a wetland landscape around Kristianstad, Sweden. Hum Ecol 34:573–592. *Includes a variety of compelling arguments on why social networks are important in natural resource governance.*

Hardin G (1968) The tradgedy of the commons. Science 162:1243–1248

Janssen MA, Bodin Ö, Anderies JM et al (2006) A network perspective on the resilience of social-ecological systems. Ecol Soc 11:15

Johannes RE (2002) The renaissance of community-based marine resource management in Oceania. Annu Rev Ecol Syst 33:317–340

Johnson JC, Griffith DC (2010) Linking human and natural systems: social networks, environment, and ecology. In: Vaccaro I, Smith EA, Aswani S (eds) *Environmental social sciences: methods and research design*. Cambridge University Press, Cambridge, pp 212–237

Krishna A (2002) *Active social capital. Tracing the roots of development and democracy*. Columbia University Press, New York

Olsson P, Gunderson LH, Carpenter SR et al (2006) Shooting the rapids: navigating transitions to adaptive governance of social-ecological systems. Ecol Soc 11:18

Ostrom E (1990) *Governing the commons: the evolution of institutions for collective action*. Cambridge University Press, Cambridge

Ostrom E (2005) Understanding institutional diversity. Princeton University Press, Bloomington

Treml EA, Fidelman PIJ, Kininmonth S, Ekstrom JA, Bodin Ö (2015) Analyzing the (mis)fit between the institutional and ecological networks of the Indo-West Pacific. Glob. Environ. Chang. 31, 263–271. doi:10.1016/j.gloenvcha.2015.01.012

Wasserman S, Faust K (1994) *Social network analysis—methods and applications*. Cambridge University Press, Cambridge

Index

S.E. Gergel, M.G. Turner (eds.), *Learning Landscape Ecology*,
DOI 10.1007/978-1-4939-6374-4