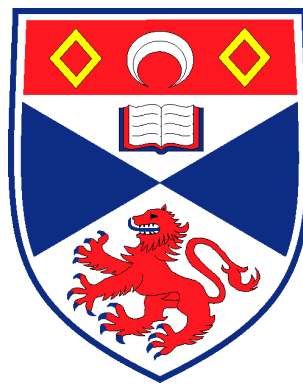# Quantifying biodiversity trends

in time and space

Angelika C. Studeny

Thesis submitted for the degree of
DOCTOR OF PHILOSOPHY
in the School of Mathematics and Statistics
UNIVERSITY OF ST ANDREWS

October 2012

Declaration

I, Angelika Caroline Studeny, hereby certify that this thesis, which is approximately 50,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in January 2008 and as a candidate for the degree of Doctor of Philosophy in Statistics in January 2008; the higher study for which this is a record was carried out in the University of St Andrews between 2008 and 2012.

date: _____ signature of candidate: _____

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in Statistics in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

date: _____ signature of supervisor: _____

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. We have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

Embargo on both all or part of printed copy and electronic copy for the same fixed period of 2 years on the following ground:

publication would preclude future publication.

date: _____ signature of candidate: _____

date: _____ signature of supervisor: _____

Abstract

The global loss of biodiversity calls for robust large-scale diversity assessment. Biological diversity is a multi-faceted concept; defined as the 'variety of life', answering questions such as 'How much is there?' or more precisely 'Have we succeeded in reducing the rate of its decline?' is not straightforward. While various aspects of biodiversity give rise to numerous ways of quantification, we focus on temporal (and spatial) trends and their changes in species diversity.

Traditional diversity indices summarise information contained in the species abundance distribution, i.e. each species' proportional contribution to total abundance. Estimated from data, these indices can be biased if variation in detection probability is ignored. We discuss differences between diversity indices and demonstrate possible adjustments for detectability.

Additionally, most indices focus on the most abundant species in ecological communities. We introduce a new set of diversity measures, based on a family of goodness-of-fit statistics. A function of a free parameter, this family allows us to vary the sensitivity of these measures to dominance and rarity of species. Their performance is studied by assessing temporal trends in diversity for five communities of British breeding birds based on 14 years of survey data, where they are applied alongside the current headline index, a geometric mean of relative abundances. Revealing the contributions of both rare and common species to biodiversity trends, these 'goodness-of-fit' measures provide novel insights into how ecological communities change over time.

Biodiversity is not only subject to temporal changes, but it also varies across space. We take first steps towards estimating spatial diversity trends. Finally, processes maintaining biodiversity act locally, at specific spatial scales. Contrary to abundance-based summary statistics, spatial characteristics of ecological communities may distinguish these processes. We suggest a generalisation to a spatial summary, the cross-pair overlap distribution, to render it more flexible to spatial scale.

Acknowledgements

Friendship isn't a big thing
— it's a million little things.

When I set out on this journey, I did not know how much it would have in store for me. I am thankful for everything I have learned, about a new subject but equally much about myself, with the help of many good travelling companions and teachers. Since my first week in St Andrews, I have felt immensely privileged to be able to work in such a welcoming and inspiring environment and I would like to thank everyone at the Centre for Research into Ecological and Environmental Modelling as well as those outside of the Observatory who have made this experience unique.

I would like to express my special gratitude to my supervisors, Professor Steve Buckland, Professor Anne Magurran and Dr. Janine Illian, for their confidence in letting a mathematician get her hands on real data, for sharing their vast knowledge with me and for their patience. I greatly appreciated their constant willingness to discuss my work, their guidance and encouragement and valuable comments on two paper drafts as well as, finally, my thesis. In particular, I would like to thank Steve for his incredible fast feedback on various drafts, for extending my funding for six more months after three years and for supporting my attendance of various conferences and workshops. I am immensely grateful to Anne for inviting me to the Kavli discussion meeting. The opportunity to meet and discuss the challenges of biodiversity research with so many experts in the field was a unique and inspiring experience for me. Many thanks belong to Janine for fruitful discussions which brought me back on track more than once, and for organising two research retreats, one of which was the birth place of the weighted xPODs, and beyond that for her friendship.

I wish to convey many thanks to Dr. Stuart Newson from the British Trust for Ornithology (BTO) for his help in getting me to work on the BBS data, for sharing his knowledge about birds with me and for helpful comments on two manuscripts. My special thanks is extended to the BTO for granting me access to their data. My gratitude goes to all the BTO staff and the volunteer observers who collect and process the data. Their work is invaluable.

The research for this thesis was funded by the National Centre for Statistical Ecology (NCSE)through the University of St Andrews for three years for which I am most appreciative.

A big 'thank you' belongs to Calum Brown for many interesting and inspiring discussions that eventually led us to the weighted xPODs, for carrying out the simulations to investigate their use in practice and for sharing his R code with me, in particular that for generating Figs. 6.2 and 6.3. I greatly enjoyed our work together.

I want to express my sincere thanks to Dr. Finn Lindgren and Dr. Daniel Simpson for helping me get to grips with INLA. I am particularly grateful to Finn for his kind permission to use Fig. 5.3.

I would like to thank Phil Harrison for providing me with the the coordinates for the UK map and for helpful discussions and comments.

My work would not have been possible without the 'good spirits' that make things work smoothly, Rhona Rodgers and Phil LeFeuvre. Many thanks for their kind help.

A big thank you to all the degrees of freedom for filling my life with music and the joy that comes with it.

From all my heart, I wish to convey my special thanks to Peter Jupp for so many things: for proof-reading my thesis and for all the encouragement over the last months and for sharing his statistical knowledge with me, but beyond that for his friendship, for my cultural and linguistic education, and amongst other adventures, for taking me to the performance of the Magic Flute in Perth.

Thank you, merci, xiexie does not really do justice to the gratitude I feel towards Yuan Yuan, Marjolaine Caillat and Vanessa Cave for their friendship and their support. I would not have been able to do this without them. I am forever grateful to Marjolaine and Yuan for letting me stay numerous times, in particular to Yuan for food and shelter over the weekends for the last months as well as for all the walks to keep me sane. Un bisou pour Elliot pour m'avoir amené des rayons du soleil au bureau.

Another, equally big, thank you belongs to my wonderful housemates, Elizabeth Hutchin and Sunil Bellur, for being my family in Scotland. I cannot thank Elizabeth enough for 4 years of friendship, of sharing house and food, and for making me laugh. Many thanks to Sunny for the many many lifts which made my life so much easier, and for making me dinner so many times.

Finally, I want to thank my sister, Sylvia Garny, for encouraging me to be adventurous four years ago. It was one of the best experiences in my life. A thank you is in order for all the letters and postcards of encouragement as well as the help with LaTeX tables and bibtex files that she provided.

Un gros merci et bisou to my partner, Olivier Gras, for his patience and for

being there for me, despite the distance.

My deep and lifelong gratitude belongs to my parents, Wilfried and Gisela Studeny, for their love, support and encouragement over the years. I would not be who I am without them.

For my grandfather,
who taught me the love of nature.

Für meinen Opa,
der mich die Natur lieben lehrte.

Abandon the urge to simplify everything, to look for formulas and easy answers, and begin to think multidimensionally, to glory in the mystery and paradoxes of life, not to be dismayed by the multitude of causes and consequences that are inherent in each experience – to appreciate the fact that life is complex. — M. Scott Peck

Contents

List of Figures

List of Tables

Introduction

Biodiversity has become a prime topic of discussion in various fields, ranging from ecology through conservation to socioeconomics: inherently an ecological concept, its importance has been emphasised by both theoretical and applied ecologists (Pielou, 1975; May, 1988; Gaston & Spicer, 2004; Magurran, 2004; Magurran & McGill, 2011). In particular, it is considered essential to ecosystem functioning (Naeem *et al.*, 1994; Hooper *et al.*, 2005). Beyond the walls of academic research institutes, it has recently received growing attention in terms of conservation concern (natural resource management, sustainability) as well as from a socioeconomic and political perspective (biodiversity as an asset for a general quality of life) (Brechin *et al.*, 2002; Adams & Hutton, 2007).

The assessment of biodiversity has always been an important subject of ecological research (Magurran, 2004). In particular, the variation in species and their abundances has been recognised early on, as commented by Darwin (1859)

> '*It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent on each other in so complex a manner ...*' — The origin of species, p.489.

However, it was not until the middle of the last century that a quantitative framework of biodiversity was developed (Fisher *et al.*, 1943; MacArthur, 1960; Pielou, 1969). Amongst others, Fisher (1943) laid the foundation by expressing the distribution of individuals across species in mathematical terms. Further models have since been suggested, trying to incorporate more realistic biological mechanisms (Sugihara, 1980; Tokeshi, 1990, 1996). The variation within an assemblage has been of interest, not only to biology, but to other subjects, too. C.E. Shannon derived the expected information contained in a string of characters or numbers, based on the frequency of their occurrence (Shannon, 1948). 'Shannon's index' was soon adopted by ecologists to measure the diver-

sity of a community (Pielou, 1966a) and a variety of alternative diversity indices has been proposed since (Simpson, 1949; Pielou, 1969; McIntosh, 1967; Hurlbert, 1971; Heip, 1974; Hill, 1973; Patil & Taillie, 1982; Smith & Wilson, 1996). The resulting 'jungle' of measures (Ricotta, 2005) has led some ecologists to criticise diversity as a 'non-concept' and doubt its usefulness (Hurlbert, 1971). However, biodiversity as a concept has persisted, and with it the need for its quantification. Instead of giving up in face of the complexity of the concept, ecologists continued to work towards a more nuanced methodological framework by acknowledging its many facets (Magurran & McGill, 2011). The focus on species diversity and numbers of species has widened to other aspects such as genetic diversity (Culver *et al.*, 2011), phylogenetic diversity (Pavoine *et al.*, 2005; Vellend *et al.*, 2011), taxonomic diversity (Warwick & Clarke, 1995) and functional diversity (Tilman, 2001; Dìaz & Carbido, 2001). In addition, limitations of the information contained in the species abundance distribution have been discussed (McGill *et al.*, 2007) and alternatives are now sought. Promising advances have been made by taking into account spatial information (Condit *et al.*, 2000; Harte *et al.*, 2005; Brown *et al.*, 2011).

'Biodiversity' gained significance on a wider public level in 1992 at the Rio Earth Summit, when 150 government representatives signed the Convention on Biological Diversity (CBD) with the aim to 'halt the loss of biodiversity' by 2010 (CBD, 1992; Dobson, 2005). With the initial deadline for the target come and gone, it is difficult to assess whether and in what ways progress towards the 2010 target has been made, not least because of the difficulty of quantifying it. The general consensus is that, apart from a few exceptions, we can consider the target as failed (Butchart *et al.*, 2010). The original objective has now been extended in the more extensive catalogue of the 20 'Aichi' biodiversity targets, to be achieved by 2020 (Scholes *et al.*, 2012).

However, the international agreement on a biodiversity target sparked a discussion about the way diversity is assessed — monitored as well as quantified (Dobson, 2005; Walpole *et al.*, 2009). While policy makers have focussed on suitable headline indices (Walpole *et al.*, 2009; van Strien *et al.*, 2012), scientists have searched for a comprehensive methodological framework to assess diversity and changes in diversity on large temporal, spatial and spatio-temporal scales (Magurran, 2011; Magurran & Dornelas, 2010). Compared to previous studies, the perspective to diversity assessment changed from a focus on site-specific diversity to an interest in following changes across larger temporal and spatial scales. In particular, the following key points have been discussed

- Instead of an absolute (one point in time or space) evaluation of diversity of a community, relative assessment has become more important. This has led to an increase in popularity of the geometric mean, especially with policy makers (Loh *et al.*, 2005). The aim is to monitor **long-term trends** in biodiversity and to identify changes in the rate of change in trends. A slowing or reversing of the rate at which biodiversity currently declines, would indicate that conservation efforts are effective (Magurran *et al.*, 2010; Magurran, 2011).

- Probability of detection varies between individuals from different species; however, most existing methods do not account for this variation (Yoccoz *et al.*, 2001; Buckland *et al.*, 2011a) which may lead to biased results. While this has been acknowledged as a short-coming, only a few large-scale monitoring programmes and studies have tried to actually incorporate **detectability**.

- Traditional diversity measures are 'non-spatial' in that they are calculated based on the species abundance distribution which does not contain information on **spatial diversity** (McGill *et al.*, 2007). While these non-spatial summary statistics can be compared across space, additional information could be gained by including spatial characteristics explicitly in diversity assessment (spatial diversity measures).

The research for this thesis has been conducted with these points in mind. In particular, it seeks to contribute to setting appropriate methods in place which can be used to assess progress towards international biodiversity targets without compromising the complexity of the concept itself. It introduces novel methodology for diversity quantification, and evaluates these methods in practical applications. In particular, it looks at assessing temporal trends in diversity and identifying turning points in these trends, i.e. points that indicate an increase or a decrease in the 'rate of loss of biodiversity'.

In detail, this thesis is structured as follows: chapter 1 sets the stage by reviewing ways to define biodiversity and its various measurable aspects. It clarifies which of these aspects are considered in this thesis, and thus provides the context for the work presented here. Problems arising with the assessment of biodiversity are discussed, in particular those that concern monitoring diversity and those that stem from describing it in quantitative terms. With regard to the latter, an overview of existing methods of quantification is given. Chapter 2 explores the idea of using goodness-of-fit statistics as measures of diversity,

more precisely of evenness. This is extended to a one-parameter index family based on goodness-of-fit statistics, which has been introduced and studied in statistics by Cressie & Read (1984). We discuss its properties in theory and practice, where the latter is achieved through applications to simulated data as well as a first small example using real data from Scottish farmland birds. Issues in the context of imperfect detection of species as well as that of variation in detectability between individuals from different species are looked at in chapter 3, with particular attention to the effects on the methods introduced in chapter 2. A comprehensive analysis of temporal trends amongst British breeding birds in chapter 4 puts the 'goodness-of-fit approach' into practice. Employing these novel evenness measures alongside a traditional geometric mean index provides new insights into how ecological communities change through time. Finally, chapters 5 and 6 look at assessment of diversity in space, although from two different perspectives. Chapter 5 makes a first step towards mapping local diversity throughout a region and discusses possibilities to assess compositional turnover across large spatial scales. The aim of this chapter is to set an example that sparks future work; we use a new algorithmic method to fit a Poisson model to the observed counts which takes account of the spatial autocorrelation structure. Chapter 6 on the other hand uses the information contained in the spatial locations of observations to draw inference on underlying processes promoting the coexistence of species. Its methodological background lies in spatial point process theory which provides a range of summary statistics for multi-type patterns in space. We propose a generalisation of one of these summary statistics that renders evaluation more flexible to spatial scale. Again, the use of this extension is illustrated in simulations as well as with data, collected at a plant diversity 'hotspot' in South-Western Australia. In contrast to the rest of this thesis which focusses on large-scale assessment of diversity, this chapter gives an example of identifying processes that operate on very local scales. We conclude with a final discussion of the results obtained, their limitations and an outlook to future work.

Because of its complexity, biodiversity will never be easy to quantify, but the development of new approaches as well as methods that allow us to integrate information across space and time in more flexible ways will hopefully provide a more realistic insight and foster deeper understanding, and in the long run increase awareness and care for our most vital resource — the 'variety of life' (Gaston, 2000) as it (still) surrounds us.

Chapter 1

Biological diversity and its measurement –
a multidimensional problem

We should preserve every scrap of biodiversity as priceless while we learn to use it and come to understand what it means to humanity. – E. O. Wilson

It is that range of biodiversity that we must care for - the whole thing - rather than just one or two stars. – David Attenborough

When we read quotes like the above, the importance of biological diversity (or biodiversity) and the necessity to protect it, seem to be beyond question. But what is the 'range of biodiversity', what is 'the whole thing' and how can we be sure to capture 'every scrap' of it? Any answer to these questions relies first of all on a systematic and comprehensive assessment of biodiversity. This starts with defining exactly what it is that we want to conserve (and why) and is quickly followed by how we want to go about monitoring biodiversity as well as our progress in 'caring' for it, and last but not least how we can describe it in quantitative terms (Yoccoz *et al.*, 2001). Biodiversity is a central concept in ecology — and one of the broadest. Hence, its assessment and measurement are anything but straightforward tasks (Hurlbert, 1971; Magurran, 2004). The richness of the concept is reflected in the various approaches to capturing it (Magurran & McGill, 2011). This chapter reviews the current understanding of biodiversity and discusses major challenges that come with the task of its assessment, in particular its quantification and the design of monitoring programmes.

1.1   What is biological diversity?

Before setting up a methodological framework for its assessment, we need to clarify what the term *biological diversity* or *biodiversity* [1] refers to and which of its aspects are considered in the present work.

The list of definitions is long and they differ greatly in the extent of their range (DeLong, 1996; Magurran, 2004). While some authors narrow biodiversity down to the number of species, i.e. *species richness* only, others include further components. For example, Hubbell (2001) considers biodiversity as *'synonymous with species richness and relative abundance in space and time'* (p.3).

At the other end of the spectrum, we find descriptions which try to capture the full complexity of biological diversity as *'...the variability among living organisms from all sources including,* inter alia*, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems'* (CBD (1992), Art. 2; for more details see Appendix A).

Given by the Convention on Biological Diversity (CBD) in 1992, this definition can be considered equally relevant as a scientific statement as well as the basis for political actions like the 2010 target. Widely cited (Harper & Hawksworth, 1995; Magurran, 2004; Gaston & Spicer, 2004) and mirrored by similar definitions from other authors (Noss, 1990), it followed Norse *et al.* (1986) in its recognition of three main organisational levels of biodiversity — genetic, species and ecosystems (see Fig. 1.1). The term 'biological diversity' directly translates to 'variability ... from all sources' or as Gaston (2000) puts it, the 'variety of life' itself. This can be specified as the genetic variation within a species and how it evolves with time or the diversity of an assemblage formed by different species (usually on the same trophic level). Neither the evolution of species nor that of communities is independent of the surrounding environment. Diversity of a whole ecosystem refers to the complexity and the interdependence of species and the environment they live in (Harper & Hawksworth, 1995; Rosenzweig, 1995; Magurran & Dornelas, 2010).

However, DeLong (1996) criticises definitions like the CBD's for their general inclusion of ecosystems, as they comprise biotic as well as abiotic components, whereas the term *bio*-diversity should only be used for the biotic parts to avoid confusion:

---

[1]We follow Magurran (2004) in using 'biological diversity' and 'biodiversity' interchangeably.

FIGURE 1.1 – Example of the three organisational levels of biodiversity. (a) shows the phylogenetic relationships among *Corydoradinae*, a group of freshwater catfishes (genetic diversity), (b) pictures species diversity in a marine evironment and (c) gives a schematic overview of a foodweb in the marine ecosystem.

Sources: (a) reprinted by permission from Macmillan Publishers Ltd: Nature (Alexandrou *et al.*, 2011), copyright 2011; (b) http://www.publicdomainpictures.net and (c) courtesy U.S. National Oceanic and Atmospheric Administration (NOAA).

'Biodiversity <u>is</u> *a state or attribute of a site or area and specifically refers to the variety within and among living organisms, assemblages of living organisms, biotic communities, and biotic processes, [...].'*

He furthermore stresses the importance of a clear distinction between the classification (i.e. semantic and structural definition) of biodiversity as a state or attribute (which we aim to assess) and the operational aspects of its measurement.

'Biodiversity <u>can be measured</u> *in terms of genetic diversity and the identity and number of different types of species, assemblages of species, biotic communities, and biotic processes, and the amount (e.g., abundance, biomass, cover, rate) and structure of each. It can be observed and measured at any spatial scale ranging from microsites and habitat patches to the entire biosphere'* (DeLong (1996), p.745).

Although this definition might read as merely a more elaborate rewording of the CBD's statement above, it carries some subtle differences which are relevant to this thesis. While biodiversity is fully acknowledged as a complex concept in ecology, it points to the fact that the broadness of the concept renders a unitary methodological framework for its measurement and assessment difficult, if not impossible. In fact, methods of quantification depend on the organisational level we are looking at, as well as what aspect of diversity we are interested in. In contrast to definitions that are very much tied to number of species and/or species proportions, it explicitly includes more recent approaches looking at diversity of a community in terms of genetic differences (phylogenetic diversity, e.g. (Pavoine *et al.*, 2005); taxonomic distinctness (Warwick & Clarke, 1995)) as well as the contribution of the different species towards ecosystem maintainance (functional diversity, e.g. (Tilman, 2001)). While this definition explicitly excludes inorganic components of an ecosystem as direct contributors to its biodiversity, one should not forget that they can be prerequisite for a species existence and as such for biodiversity.

With respect to quantification of biodiversity, a complex definition such as DeLong (1996) seems to increase the difficulty of the task (Hurlbert, 1971). But biodiversity assessment must acknowledge that there is not a unique measure to capture every aspect of it (Magurran & McGill, 2011). If we ground our methods on a definition which focuses on maybe one or two of the measurable components

only, as in Hubbell's definition at the beginning of this section, we seemingly decrease the complexity of our task and escape the necessity of justifying why we are concentrating on these aspects. Reasons might be as trivial and legitimate [!] as the lack of resources or accessibility within a monitoring scheme or just the current boundaries of our knowledge, but ultimately we always have to aim for the *'whole range of biodiversity'* as David Attenborough put it. A definition should mirror this objective by reflecting biodiversity in its entire complexity. It then allows for the possibility of gathering information on different levels to get a fragmented, possibly quite incoherent, but finally more complete picture than we would achieve by only counting the number of species (for example Magurran & McGill (2011)).

DeLong's definition also explicitly includes a possible variety of 'currencies' which can be used for biodiversity measurement. The question of the 'currency' of diversity has gained increasing interest in ecology (Chiarucci *et al.*, 1999; Morlon *et al.*, 2009) So far, the focus of biodiversity quantification has clearly been on individual counts. Not only have doubts been raised about this always being the best choice (**?**), but the combined use of different currencies can also be applied to gain information from the extent of agreement between them (Warwick & Clarke, 1994; Henderson & Magurran, 2010).

The currency in which biodiversity is measured is not the only attribute that is open to choice. DeLong (1996) makes the important point that biodiversity, and therefore its quantification, ranges across different spatial scales, from local sites over regions to a national and even global dimension. The spatial aspect is perhaps a minor issue when quantifying biodiversity of single sites (i.e. a single, selected location in space), but it cannot be ignored when assessment is for larger regions and entire countries, as intended in the CBD's 2010 target (Pereira & Cooper, 2006). It is mainly the latter that we consider in this thesis and hence we are concerned with diversity 'in space' (Rosenzweig, 1995). Scaling properties of unifying concepts like biodiversity are reckoned to enable inference on global ecological patterns (Storch *et al.*, 2007).

However, space does not enter biodiversity assessment solely in terms of scale. Recently, space and spatial patterns have received growing attention in ecology in their own right (McGill & Collins, 2003; McGill, 2011). It has been emphasised that assemblages of species do not carry information only on their abundances, but also on the spatial relations between individuals (Wiegand *et al.*, 2007; Brown *et al.*, 2011; Rajala & Illian, 2012). The latter promise to provide statistical methods for investigating ecological communities and processes that are shaping them. They can, for example, be used to test for the
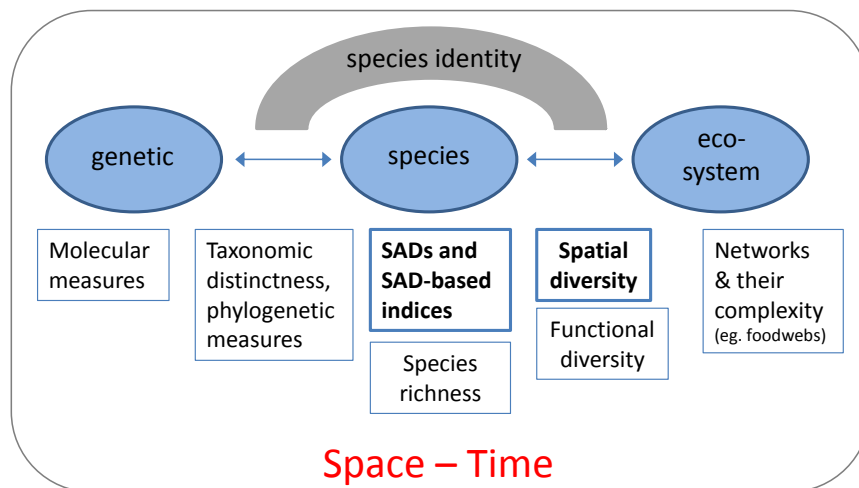
FIGURE 1.2 – Overview of the three organisational levels of biodiversity and various ways to quantify them. The complexity of biodiversity as a concept cannot be represented by a single method of quantification. Depending on which component is of interest to a study, a range of methods have been suggested. Biodiversity is independent neither of time nor space since populations evolve and structures of communities and ecosystems change. On the background of this framework, we look at diversity quantification based on species abundance distributions (SADs) and spatial diversity characteristics of ecological communities.

absence of interactions between species or to evaluate conjectured shapes of the species abundance distribution (SAD). This adds a new perspective and an additional (measurable) component to the biodiversity concept, namely *spatial diversity*, and might lead a way out of the traditionally SAD-focussed view of biodiversity research (McGill *et al.*, 2007).

Based on this discussion, this thesis views biodiversity as a multi-dimensional, unifying concept spanning three major levels of organisation which relate to different subject fields in biology. These are genetic diversity, diversity of species assemblages and of entire ecosystems. This is summarised in Fig. 1.2. Measurable components of biodiversity are various, can be found within any of these levels and can be quantified in different currencies and along different spatial scales. On the genetic level, molecular measures are used to determine the genetic distance between individuals of the same species (distance to the most recent common ancestor), but also to time speciation events (Nichols, 2001). However, they can also be applied to identify species as dissimilar and thus measure the diversity of a species assemblage based on their phylogeny (Pavoine *et al.*, 2005). Traditionally, biodiversity has been very much associated with species diversity (see for example Hubbell's (2001) definition of biodiversity).

Quantification on this level typically considers species at the same trophic levels; it aims at determining the number of species (species richness, May (1988); Gotelli & Colwell (2001)) or is based on the species abundance distribution, i.e. the proportion of abundance allocated to each species in the assemblage (McGill *et al.*, 2007). Both the problem of quantifying species richness reliably and summarising information contained in the species abundance distribution has led to the development of a wide range of statistical methodology. More recently, species diversity has been extended to take into account a species' contribution to the maintenance of ecosystem functions (Tilman, 2001; Dìaz & Carbido, 2001). At the most complex level, quantification of biodiversity could aim at the interactions and processes driving an ecosystem as a whole. Obviously, the diversity of a whole ecosystem cannot be captured by a single measure. Rather, measures of all aspects (genetic, taxonomic, SAD, functional, spatial) have to be combined to characterise ecosystems (Magurran & McGill, 2011). In particular, spatial diversity characteristics play an important role here, as they aim to reveal the processes which drive ecosytems, primarily species interactions. To get a picture of an entire ecosystem and its diversity across all trophic levels, network approaches can be applied (Dunne *et al.*, 2002; McCann, 2007).

Species identities are by definition relevant to genetic and phylogenetic diversity measures; they likewise are important when we look at specific species' functional roles within an ecosystem or interactions between individual species. On the contrary, the classical SAD-approaches do not take into account species identities, i.e. they always compare ordered species proportions (McGill *et al.*, 2007; Magurran & Henderson, 2010).

Throughout this work, our perspective is that of the CBD's 2010 target, i.e. biodiversity is considered as a measureable state of a large region (typically a whole country), that can vary both in time and space. From a methodological point of view, this thesis concentrates on the species level of diversity. While we acknowledge the significance of estimating species richness (chapter 3.2) and the progress that has been made towards overcoming the problem of undetected species (Mao & Colwell, 2005; Chao *et al.*, 2009), our focus is on measures based on the species abundance distribution (SAD) here. In particular, we discuss methods based on species proportions that allow us to separate trends in rare species from those in common species. While these methods, which we introduce in chapter 2, are based conceptually on individual counts, they can to some extent be applied to other currencies of biodiversity. Diversity on the species level has traditionally been defined as either information or variation (Maurer & McGill, 2011), as we will see section 1.3.1. Later, we extend this

view by defining and exploring spatial diversity characteristics (see chapter 6).

1.2   General problems in diversity assessment

1.2.1   Multidimensionality

Even if we limit our view to the species level, there is no unique way to quantify diversity of species' assemblages (Magurran, 2004). The most immediate (and in most cases first to be looked at) is the number of species in the assemblage, or *species richness* (May, 1988). Though often almost equated with species diversity, species richness as a measure of diversity is not without pitfalls: It is independent of neither scale nor sampling effort (Rosenzweig, 1995) and estimates of richness tend to be sensitive to sampling effects and survey design (Gotelli & Colwell (2001, 2011), see also section 1.2.3 and chapter 3). But the number of species is not the only possible characteristic of a community that can be used to quantify (species) diversity. As Hubbell's definition of biodiversity in the previous section states, not only the presence of species but also their abundances add to the variability within an assemblage.

Any sensible form of assessment and measurement of biodiversity (species diversity) should register changes in any of the following components, (Buckland *et al.*, 2005)

- the number of species

- the total abundance

- the distribution of individual abundances across species.

The last is referred to as a change in *evenness* (Pielou, 1969). The closer this distribution is to a uniform distribution, the more diverse we usually view the community it represents. Because any of these aspects (richness, abundance, evenness) characterise and contribute to overall (species) diversity of an assemblage, the choice of a diversity 'measure' is preceded by the decision to focus on a specific characteristic. While this depends on scientific interests or aims of a study, any measure of diversity should behave sensibly with regards to changes in the three dimensions: a change in any of species richness, total abundance or evenness while the remaining two components stay constant should result in an equivalent change in overall diversity (Buckland *et al.*, 2005).

The multidimensional nature becomes most apparent when we consider evenness. Typically it has been seen as independent of the number of species and

is has been assumed that diversity can be decomposed into a richness and an evenness factor (Heip, 1974; Gosselin, 2006; Jost, 2007). In consequence, it has been considered crucial that any measure of evenness should be insensitive to changes in number of species (Smith & Wilson, 1996). Only recently has it been acknowledged that evenness and species richness cannot be regarded as independent components of diversity (Jost, 2010). Jost (2010) supports his argument by mathematical derivation of evenness factors which depend on species richness. Based on the same derivation, he then suggests a (multiplicative) partitioning of species richness into independent diversity and evenness factors. Diversity is in this case seen as the 'effective number of species' (Hill, 1973; Jost, 2006), a concept which will be discussed further in section 1.3.2 below. While mathematically sound, this is hard to interpret in an intuitive and ecologically sensible way. That evenness and richness cannot be measured independently, is evident from the fact that a change in numbers of species leads to a change in the dimension of the uniform distribution that corresponds to complete evenness. Thus evenness can be regarded as complementary to species richness; it contains additional information about a assemblage of species, but it is a relative concept rather than an absolute one: changes in richness will alter the reference point of what is regarded as complete evenness (see also chapter 2).

1.2.2 Sampling aspect, the concept of a community and units of measurement

The different components of species diversity describe characteristics of an ecological community; a diversity measure should reflect the actual properties ('truth') of the community under consideration. However, this is not straightforward. Firstly, ecologists do not agree upon the concept of a 'community' and what defines it; some go as far as to doubt the usefulness of the concept as a whole when it comes to diversity assessment (Hurlbert, 1971; Smith & Wilson, 1996). Often, the diversity of a community is equated to that of a sample from it (Maurer & McGill, 2011): for example, Smith & Wilson (1996) abandon the idea of *'a sample from a community'* as *'unrealistic'* and *'prefer to see the quadrat or sample as a small, fully censused piece of biotic space'*. This might at the most hold for a narrow, site-focused approach to diversity where sites are sufficiently small to sample everything, but is impossible to achieve for many taxa (Lawton *et al.*, 1998; Longino *et al.*, 2002) and at larger spatial scales (Buckland *et al.*, 2011a). In general, an approach like this ignores the problem of detectability (see following section and chapter 3). While it is al-

ready doubtful that this is realistic for single sites, it surely cannot be upheld if we are interested in biodiversity quantification of a large region – the spatial scale we are interested in when assessing the international biodiversity targets. Rather, we have to assure that sites and the community of interest are chosen as representative of the entire biota of a region or a country (Buckland *et al.* (2011a, In prep); and see also section 1.4).

Maurer & McGill (2011) provide a more precise, stochastic (metapopulation) framework, where the species abundance distribution that represents the community in question is derived as the expectation over a finite number of realisations from a multinomial (equilibrium) distribution. This expectation itself is considered an 'ecological sample' as ecological processes act as a 'filter' for a species' presence at any given time. However, this ecological sample is usually not directly accessible and any of its characteristics, such as, for example, diversity, have to be estimated based on 'empirical samples' from the community.

Even if the delimiters of the community of interest are clear, we rarely are in a position to account for every individual belonging to it. Thus, the 'true' species abundance distribution, the 'true' diversity and any of its aspects are not directly accessible, but we have to rely on samples from the species assemblage. Whether or not the concept of a community is agreed on, it has to be kept in mind that any quantity derived from a sample is an *estimate*. Hence, we will usually measure diversity via an estimator or a summary *statistic*. As such we are interested in the statistical properties of the summary statistics we use, in particular their bias and precision (e.g. Hellmann & Fowler (1999) for species richness). While much effort has gone into developing robust statistical methodology for richness estimators (Gotelli & Colwell, 2011), the same would be needed for estimation of biodiversity 'headline' indices. Ideally, we would want estimators of diversity to show little bias and high precision. This is included in the criteria for a reasonable diversity measure by Buckland *et al.* (2005).

The fact that we rely on samples to estimate diversity of an assemblage also influences what we might consider the appropriate 'currency' of diversity. Commonly, individuals are the sample units and hence count data are the basis for the estimation. However, sometimes it might be impossible to sample single individuals (such as in grassland plants, Tilman *et al.* (2006)) or individuals might differ very much in some aspect (for example size in fish communities, Henderson & Magurran (2010)). It is then often more appropriate to consider other units of measurements, such as biomass, instead of individual counts (**?**).

1.2.3   Detectability and rarity of species

As we rely on samples to estimate diversity, detectability becomes an issue (Buckland *et al.*, 2011a). 'Detectability' refers to the probability with which an object of interest is recorded during the observation process. In terms of biodiversity assessment, this is an obvious problem when considering species richness (Gotelli & Colwell, 2001, 2011); the major concern about richness as a measure of diversity is due to the fact that we do not expect to see all the species that form the community (Boulinier *et al.*, 1998). Extending the survey area or increasing sampling effort will almost certainly reveal additional species. However as long as we do not detect any individuals of a species, the species remains undetected. Completely undetected species pose a problem for which various parametric and non-parametric methods have been proposed, but that still is not readily dealt with (see chapter 3), in particular for speciose communities where most species are rare, such as tropical arthropods (Longino *et al.*, 2002). It is necessary to distinguish this form of detectability (of a species) from that of an individual.

The latter has so far not received much attention (Yoccoz *et al.*, 2001; Buckland *et al.*, 2011a). When estimated from samples, diversity measures are usually calculated based on the observed individual counts – without any attention to the fact that the individuals in the sample from different species might differ in their detection probability: observed counts from species A in a sample could be lower than those from species B because there are actually fewer individuals from species A in the community *or* because individuals from species A are less likely to be observed. If diversity is estimated from the recorded individuals without taking differences in detection probabilities into account, we treat the sample as if we were certain that species A has indeed less individuals. However, appropriate survey design can provide additional information that allows us to estimate detection probabilities and to derive abundance estimates for each species (see 1.4 below). Diversity measures can then be based on the abundance estimates instead of the counts (MacKenzie & Kendall (2002); Buckland *et al.* (2011a); and see chapters 3 for further discussion and 4 for an example).

This second issue of individual detectability is bound to affect all measures which are based on species abundances. As it depends solely on the occurrence of a species, species richness as a measure of diversity does not need to be adjusted for individual detection probability, but suffers from the more difficult problem of species detectability (MacKenzie *et al.*, 2003). In particular in highly speciose communities, the number of species observed will not have reached an

asymptote, even after thorough sampling (Longino *et al.*, 2002).

Rarity of a species is problematic in terms of both species and individual detectability. The encounter probability for a rare species in the sampling process is small and it is likely to remain undetected (MacKenzie *et al.*, 2005). Even if detected, observations will be sparse and hence reliable estimation of abundance hindered. Often detection probabilities cannot be estimated directly, but have to be 'borrowed' from a similar common species. Typically, rare species reduce precision of the diversity assessment drastically, and for some diversity measures, like the geometric mean, they pose a problem for the calculation of the measure. There are arguments for not including rare species in diversity assessment, as they cause various problems but in general do not contribute much in terms of biomass or with regards to functions of an ecosystem, on the other hand they are usually the ones with the highest conservation concern (Gaston, 2008).

1.2.4   Temporal trends

While the composition of communities has never been seen as static, existing biodiversity measures and their use have long focussed on determining site-specific diversity at a single point in time or used space as a proxy (Magurran, 2011; Magurran & Dornelas, 2010). In the face of objectives like the international biodiversity targets, monitoring diversity trends on a long-term basis becomes essential (Dobson (2005); Pereira & Cooper (2006); for an example see Fig. 1.3).

Based on either count data (or other currencies) or estimates of abundance or biomass, the time series of any diversity measure is likely to show a change between two times, as it is driven by short-term fluctuations (see Fig.1.3 (a)). By smoothing these out, long-term trends can be extracted (Fewster *et al.* (2000); and see chapter 4). However, extensive time series of data are needed to see long term temporal effects (for example of climate change) on diversity (Magurran *et al.*, 2010; Magurran & Dornelas, 2010). In the face of the current decline in biodiversity, following the trend curve alone is not sufficient. The international 2020 biodiversity targets call for an improvement of 'the status of biodiversity by safeguarding ecosystems, species and genetic diversity' preceding the former objective to 'halt the rate of loss of biodiversity' (CBD, 1992). If we want to assess progress with respect to this target, we are interested in determining whether or not there has been a change to the current rate of biodiversity decline

(Buckland *et al.*, 2005). This information can be accessed through the second derivative of the trend curve. Points where there is an indication for an actual change in the rate of change can be identified (see chapter 4). This is usually done numerically and does not require an extensive time series (Fewster *et al.*, 2000)

While diversity indices summarise information on dominance structure of the underlying community, they are insensitive to species identities (Magurran & Henderson, 2010). Hence, a diversity index can stay exactly the same, while the composition of a community changes, as long as there is no change in the distribution of species proportions (abundance or biomass). This can be mended by following species turnover in time in addition to trends in diversity, for example using rank abundance clocks (Collins *et al.*, 2008). However, like trends in diversity, species composition at single sites will vary between points in time, especially over short periods of time, more due to random fluctuations than to a real change in the underlying community. On the other hand, if we are interested mainly in regional diversity, the species catalogue over a large spatial area (such as a region) will be fairly stable over long periods, while turnover happens across space. Existing turnover measures are very much based on a site-specific approach, looking at either changes within a single site or differences between pairs of sites (β-diversity) (Jost (2007); Tuomisto (2010); see also discussion in the next section). This cannot readily be extended to cover turnover between a large number of sites across a whole region (see chapter 5).

1.2.5 Spatial components of diversity

While temporal trends in biodiversity have not received much attention until recently, there has long been an interest in its spatial patterns (Pielou, 1969; Whittaker, 1972; Cody, 1975; Gering *et al.*, 2003). This is reflected in classical partitioning of diversity into the following spatial components (Whittaker, 1972; Cody, 1975):

- α – (average) diversity within a site

- β – diversity or variation between sites

- γ – diversity pooled over all sites

The discussion about the nature of the partition is almost as old as the diversity concept itself (Whittaker, 1972; Lande, 1996) and still ongoing (Jost, 2007). While some favour an additive decomposition of γ-diversity into α- and
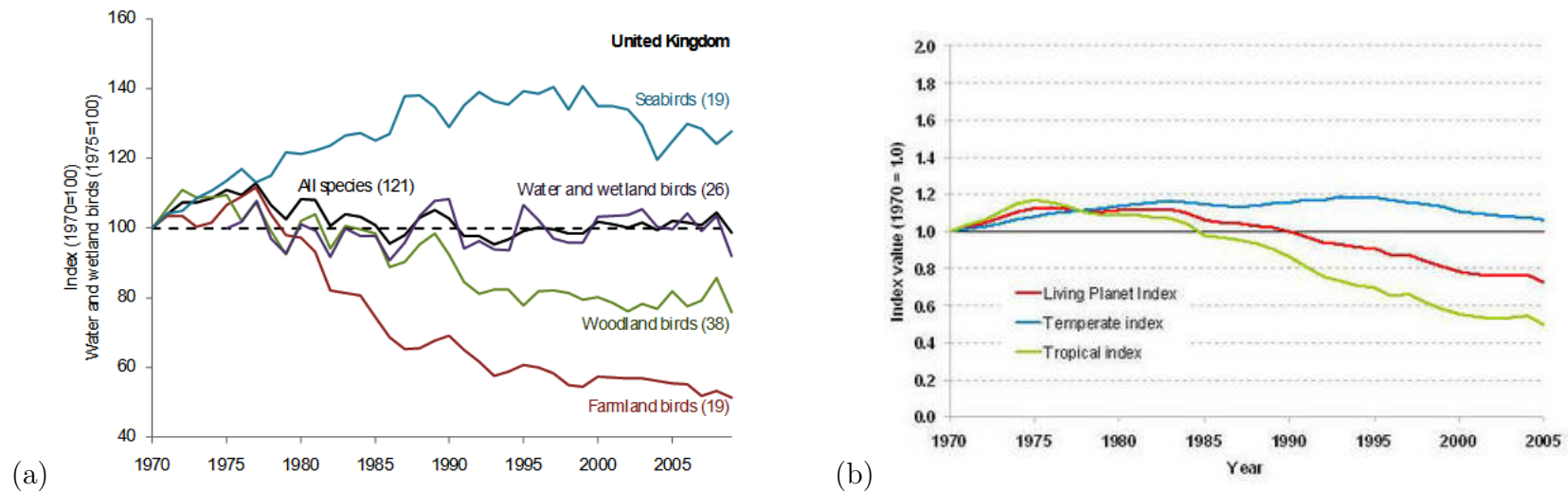
FIGURE 1.3 – Examples of long-term trends in diversity. (a) The UK Wildbird Indicator (BTO, RSPB, BirdLife, DEFRA) is part of a set of indices trying to measure sustainable policies and ecosystem health throughout the UK; (b) the Living Planet Index (WWF, ZSL, UNEP) tries to capture trend in global biodiversity by summarising population trends of more than 2,500 species of fish, amphibians, reptiles, birds and mammals.

Source: (a) courtesy of the British Trust for Ornithology (http://www.bto.org/science/monitoring/developing-bird-indicators); (b) reprinted with kind permission from Loh *et al.* (2010)

β-components similar to an analysis of variance (Lande, 1996; Gering *et al.*, 2003; Maurer & McGill, 2011), others argue for a multiplicative relationship (Whittaker, 1972; Jost, 2007). Preference for either one or the other also depends on the choice of diversity measure. However, the paradigm of diversity partitioning and its discussion is closely linked to a site-based approach to diversity and hence loses its significance when diversity is considered at a regional level; this holds in particular for β-diversity or species turnover between sites.

While the concept of α-diversity is not as relevant on a large (regional) spatial scale and regional survey schemes often do not provide enough data to estimate diversity reliably at each site, β-diversity as species turnover across space (and time) would be of interest (Buckland *et al.*, In prep), but suffers from scale dependence and the fact that it is only meaningful for a few single sites. This will be discussed further in chapter 5, along with first potential steps towards alternative approaches. Only γ-diversity can be readily estimated on a regional level (see chapter 4). In consequence, this thesis mainly considers γ-diversity in application as it offers the only immediately available way to assess regional diversity trends. Nonetheless, the measures developed in chapter 2 could also be calculated at the site-level given sufficient data.

Any of α-, β- and γ-diversity components are based entirely on the species abundance distribution, i.e. only the frequency of a species, not its spatial relationship to other species. Although the partitioning implicitly contains variation across space in the species abundance distribution, this is not taking spatial information explicitly into account.

In contrast, spatial diversity looks at the spatial positions of species with respect to each other. This can be informative in terms of processes that are relevant for the maintenance of ecosystems, but that are not accessible through the species abundance distribution (Pielou, 1969; McGill, 2011). Spatial statistics and spatial point pattern analysis have recently been applied to access such information about the spatial composition of ecological communities (Perry *et al.*, 2006; Law *et al.*, 2009) and have been shown to provide a valuable instrument to investigate interactions. More generally, this promises valuable insights to various niche-based processes, including competition, attraction and repulsion between species (Brown *et al.*, 2011). This is in particular relevant to studies of plant diversity, as the spatial position of plants does not undergo rapid changes and is fairly easily determined. However, spatial point process analysis has also been successfully used to analysis seasonal animal movement patterns, such as for example muskoxen herds in Greenland (Illian & Hendrichsen, 2010). In chapter 6 we extend this kind of approach to 'spatial diversity' to cover flexible

spatial scales and study its properties and advantages.

1.3 Existing methods to quantify species diversity

This section was written by me. An edited version has been incorporated into Buckland *et al.* (In prep), following a request by the editors of Environmental and Ecological Statistics to incorporate into that paper a review of methods for quantifying biodiversity trends.

1.3.1 Diversity indices

Whether biodiversity is considered a fundamental characteristic of our planet or a socio-economic asset, there is by now little question that diversity assessment is needed in order to maintain it in face of its global loss (Pereira & Cooper, 2006). Reliable quantification is essential for any form of assessment (Balmford *et al.*, 2003; Dobson, 2005) and the question of how to quantify biodiversity is as old as its introduction as one of the main concepts in ecology (Fisher *et al.*, 1943; Pielou, 1969). Methods that are used to measure diversity are not inherent to ecology, but span various disciplines such as information theory (Shannon, 1948; Rényi, 1965), statistics (Fisher *et al.*, 1943), population genetics (Simpson, 1949; Hubbell, 2001), physics (Tsallis, 1988) and economics (Cowell, 1980; Hoffmann, 2008).

   The multidimensionality of diversity as a concept prevents us from measuring it in a single way and hence a variety of quantities have been suggested. As mentioned above, some approaches concentrate on species richness alone and attempt to determine the number of species in a community beyond the number of species observed in samples (Chao, 1984; Chao & Lee, 1992; Shen *et al.*, 2003; Chao *et al.*, 2009). More complex quantities combine information on several aspects, the latter are usually referred to as diversity indices (Magurran, 2004). Here, we use the terms 'diversity index' and 'diversity measure' interchangeably. (However, we point out that these measures are in general not measures in the strict mathematical sense, i.e. as defined by measure theory (see, for example, Doob (1994), p.17/18).)

   It is beyond the scope of this work to go through a complete collection of the numerous diversity indices that have been suggested (Hurlbert, 1971; Tóthmérész, 1995; Magurran, 2004; Maurer & McGill, 2011). Instead we review

those indices and methods that are relevant to this thesis, either because they are used in applications (chapter 4) or because they serve as a basis for comparison (chapter 2). This section also introduces general notation used in the following chapters.

Our primary interest is in quantifying diversity on a large spatial scale and to follow changes in diversity over time, hence we consider a *region* of interest with a total number $S > 1$ of species present (usually assumed on the same trophic level). This list of species is assumed fixed for the time period we are interested in. (We return to this assumption in chapter 3.) All individuals belonging to these species form what we consider the community of interest. The *species abundance distribution* of this community is given by the vector

$$\mathbf{N}_t = (N_1^t, \ldots, N_S^t) \tag{1.1}$$

where N_i^t is the number of individuals of species i in the region of interest, $i = 1, \ldots, S$, at time t. Here and in the following, we drop the index t if we are looking at a single time. The majority of diversity indices are defined based on the species proportions given by

$$\mathbf{p}_t = (p_1^t, \ldots, p_S^t) = \left(\frac{N_1^t}{N^t}, \ldots, \frac{N_S^t}{N^t} \right) \tag{1.2}$$

where $N^t = \sum_{i=1}^{S} N_i^t$ is the total abundance of the community at time t. For a more precise mathematical definition, see Maurer & McGill (2011) who consider $\mathbf{N}$ as the expected abundances for a random realisation from a metapopulation in equilibrium.

Traditionally, diversity indices are calculated from the species proportions p_i of individual counts; however, sometimes it might not be possible or biologically appropriate to take individuals as units of measurements. More recently, diversity quantification has also been considered based on other currencies than individual counts, for example biomass (Henderson & Magurran, 2010).

As discussed in section 1.2.2 above, the quantities in equations (1.1) and (1.2) are in general not directly observable; hence diversity indices (or any other summary characteristic of the species abundance distribution) must be estimated based on samples of the community of interest. A sample will be denoted by $\mathbf{n}_t = (n_1^t, \ldots, n_S^t)$ and $n = \sum_{i=1}^{S} n_i^t$ will denote its sample size. Note that due to the sampling effects discussed above, zeros are possible: n_i is zero if species i is not recorded in the sample. In this context, and to stress

the sampling aspect of diversity assessment, we also speak of diversity indices as 'summary statistics' and their corresponding estimators. Indices and their estimators differ in their sensitivity to zeros in a sample (see chapter 3).

Maurer & McGill (2011) distinguish two main conceptual bases for the classical diversity indices used in ecology: a variance-based view and an information-theoretic derivation. The two most commonly used indices — Simpson's index (Simpson, 1949) and Shannon's entropy (Shannon, 1948) — are also typical representatives of each group: Simpson's index is calculated as

$$D_t = \sum_{i=1}^{S} (p_i^t)^2, \tag{1.3}$$

which is the probability that two randomly drawn individuals (from the ecological community) belong to the same species. This probability is lower the more diverse the community is, and hence D measures homogeneity rather than diversity. Thus, usually transformations $1 - D$, $1/D$ and $-\log D$ are used as indices. The probability $1 - D$ that two individuals do not belong to the same species can be interpreted as the within-species contribution to the total variation contained in $\mathbf{p}$ (Maurer & McGill, 2011)

$$\mathbb{V}[\mathbf{p}] = 1 - \sum_{i=1}^{S} p_i^2 - 2 \sum_{i<k} p_i p_k. \tag{1.4}$$

Maurer & McGill (2011) furthermore identify this within-species variance as a probabilistic measure of species richness, while the between-species part (covariance between p_i and p_k) accounts for evenness.

On the other hand, Shannon's entropy

$$H_t = - \sum_{i=1}^{S} p_i^t \log p_i^t \tag{1.5}$$

was defined in information theory to quantify the amount of uncertainty contained in a message (Shannon, 1948; Rényi, 1965). The message is seen as a random string of units of information drawn from an underlying 'alphabet' which is described by the frequencies for each letter. It has been adopted in ecology as a diversity measure (Pielou, 1966b). In the context of an ecological community, the set of species is considered the alphabet with frequencies given by the species abundance distribution $\mathbf{p}$. Under some assumptions like additivity, it can be shown that $-\log p_i$ is a sensible way of quantifying the

uncertainty in whether or not a randomly drawn individual belongs to species i. The uncertainty is highest if all species are equally frequent.

While interesting from a statistical point of view, Maurer & McGill (2011)'s interpretation of $1 - D$ as a 'probabilistic measure' of species richness lacks biological meaning. A different concept which links Simpson's index in the form $\frac{1}{D}$ to Shannon's entropy is that of 'effective number of species' (Hill, 1973; Jost, 2006). However, it suffers likewise from the fact that as a theoretical (mathematical) framework it does not readily offer a biological interpretation. The latter aproach is discussed further in section 1.3.2.

Simpson's as well as Shannon's indices combine information on both the number of species (observed in a sample) and evenness (Magurran, 2004; Maurer & McGill, 2011). Species that are not contained in a sample (but are considered part of the species catalogue $1, \ldots, S$, such that $n_i = 0$) make no contribution to either index. Indeed, both indices tend to neglect the contribution from rare species, however Simpson's index is slightly more driven by the dominant species (this will become evident in the section 1.3.2). Because of their insensitivity to rare species, Simpson's and Shannon's indices usually show good precision (Buckland *et al.* (2011b) and see chapter 4). As they are solely based on species proportions **p**, they do not register a decline if all species' abundances go down at the same rate. This is one of their major drawbacks when it comes to determining the current loss of biodiversity (Buckland *et al.*, 2005, 2011b).

An alternative measure is the geometric mean of relative abundances (Buckland *et al.*, 2011b; van Strien *et al.*, 2012). In contrast to the usual convention, relative abundances here refer to a species' abundance *relative* to its abundance in a baseline year t_0, i.e. quotients $N_i^t/N_i^{t_0}$ for $i \in \{1, \ldots, S\}$ at times t. The geometric mean is defined as

$$G_t = \exp\left(\frac{1}{S} \sum_{i=1}^{S} \log \frac{N_i^t}{N_i^{t_0}}\right). \tag{1.6}$$

It meets many requirements of a headline index for biodiversity (Buckland *et al.*, 2005, 2011b), in particular it picks up a decline if all species are declining at the same rate. The geometric mean respects the multiplicative nature of the relative abundances – if one species' abundance doubles relative to the baseline while another species halves its abundance, the relative abundances are 2 and 0.5, respectively. In additive terms, the first species' increase would dominate over the second species' decline. However, relative abundances are quotients and hence work multiplicatively, i.e. the increase of species one equals

the decrease of species two (2 is the reciprocal of 0.5). In particular, compared to the arithmetic mean, which works on an additive scale, the geometric mean is more appropriate when relative abundances are considered (Buckland *et al.*, 2005). It summarises *within* species trends. This enables the user to combine data from different surveys, as for example it is currently done in the Living Planet index (Loh *et al.*, 2005) (see Fig. 1.3 (b)). Traditionally seen as a pure measure of trends in species' abundance (Gregory *et al.*, 2005; Buckland *et al.*, 2005), it has recently been pointed out that the geometric mean reflects evenness, too. This is due to the fact that it places equal weights onto rare and common species when the mean is taken (Buckland *et al.*, 2011b). Because of these advantages, the geometric mean is an appropriate headline index and has been adopted by policy makers as a diagnostic for the assessement of the CBD's biodiversity targets, on an international level (Living Planet Index) as well as the national level (for example the UK Wildbird Indicator). It is also the measure of choice for many ornithologists (Gregory & van Strien, 2010). However, the geometric mean index has some disadvantages, too: Because it takes rare species as much into account as common ones, its precision is usually lower than that of Shannon's or Simpson's index (see chapter 4). While the latter simply ignore zeros in a sample, the geometric mean cannot be evaluated in this case. Species with missing observations at any point during a survey period usually have to be excluded from the set list of species.

The conceptual difference between the geometric mean index and traditional measures like Shannon's or Simpson's is that of a relative assessment versus an absolute one. While Simpson's or Shannon's index can be evaluated for one single time, the geometric mean is applicable only to a time series, as it requires a baseline for the individual abundance indices.

Yet another group of indices is based on the idea of comparing curves or areas — an approach mainly used to assess evenness. The best known is the *Lorenz* curve (Lorenz, 1905). A popular method to assess evenness of distributions in economics, it has been successfully applied to ecological problems and gained popularity (Tóthmérész, 1995; Rousseau *et al.*, 1999). In principal a cumulative distribution function, it plots the cumulated species proportions against the rank of a species (in percentage, by increasing order). In case of a completely even species abundance distribution, the abundances of $x\%$ of the species equal $x\%$ of the total abundance and hence the Lorenz curve is the bisecting (straight) line through the origin of the unit square. Any derivation from evenness results in a concave line with the same start and end point. This curve can be used to compare different distributions (independent from

any difference in numbers of species). If their respective Lorenz curves do not intersect, one is necessarily closer to the straight line and hence more even than the other. However, this only defines a partial order since two distributions with intersecting Lorenz curves cannot be compared. The *Gini*-coefficient which is calculated based on the Lorenz curve, tries to overcome this (Morgan, 1962). It is based on the idea that the larger the area between the straight line representing complete evenness and the Lorenz curve, the more uneven is the corresponding distribution. Hence the fraction of this area to the triangular straight line defines a measure of unevenness (the Gini-coefficient). A similar approach is used by Camargo (1995) and Bulla (1994). Instead of considering the cumulative distribution function, they look directly at the area under the species abundance distribution and compare it to the one expected under perfect evenness (for the same number of species) to calculate an index of evenness.

So far, we have seen indices as summary statistics of the (true) species abundance distribution. In reality, $\mathbf{N}$ and $\mathbf{p}$ are typically not known and hence we are looking at estimators of diversity rather than the indices themselves (Lande, 1996). It is common to get an estimate by replacing the species abundances N_i by the observed individual counts n_i in formula (1.3), (1.5) and (1.6) above. However, as we will see in chapter 3 this leads to biased results in most cases. More generally, indices based on $\mathbf{N}$ can be estimated by replacing N_i with an estimate $\hat{N}_i$ ('plug-in estimator'). (This can be the individual count if a less biased estimate is not available.) Apart from reducing bias, estimation of N_i (based on observed n_i) sometimes also permits us to overcome problems caused by species with missing observations at some time points during the survey period: if we are able to fit a temporal or spatio-temporal model to the observed data, we can estimated diversity based on the model predictions instead of the counts (Gotelli *et al.*, 2010; Buckland *et al.*, 2011b). In chapters 4 and 5 we apply this to improve estimates of diversity of British breeding birds.

1.3.2   Parametric index families

As discussed in the last section, diversity indices typically combine information on different aspects of biodiversity (richness, evenness, abundance) in one single number (scalar measure). This was seen by many as insufficient to capture the multidimensionality of the concept (Hurlbert, 1971; Hill, 1973; Patil & Taillie, 1982). In addition, the absolute value of a scalar diversity number carries little meaning; comparability is limited and is usually only comparability with respect

to the same measure and under similar conditions (Tóthmérész, 1995; Liu *et al.*, 2007).

The problem with diversity quantification is that there is no unique way to determine which of two species abundance distributions is the more 'diverse' — mathematically, there is no natural order for elements in $\mathbb{R}^n$ when $n \geq 2$. Diversity indices seemingly solve this issue by assigning a single number (scalar) to each species abundance distribution and hence the natural order in $\mathbb{R}$ becomes available. But this orders the SAD only according to the aspect that the index primarily captures, and a change in the choice of the index may well reverse the order.

To achieve greater generality, diversity indices have been combined into parametric families which describe the diversity measure as a function of a free parameter (Hill, 1973; Patil & Taillie, 1982; Ricotta, 2003; Leinster & Cobbold, 2012). This approach is again strongly linked with information theory (Ricotta & Avena, 2002). Hill (1973) introduced a transformation of Rényi's generalised entropy to ecology, which is commonly used. Other parametric index families exist, for example Patil & Taillie (1982), but many of them are transformations of Hill's numbers and are typically outperformed by them (Tóthmérész, 1995).

Hill's (1973) one-parameter family of 'diversity numbers' J is given by

$$J_t(\lambda) = \left(\sum_{i=1}^{S} (p_i^t)^{\lambda} \right)^{\frac{1}{1-\lambda}}, \tag{1.7}$$

with the restriction $\lambda \geq 0$. (This is equivalent to the expontential of Rényi's generalised entropy.)

The fundamental idea of a parametric index family is that the free parameter, here λ, enables the user to consider not only one measure at a time but to evaluate and compare several simultaneously. And indeed, for specific values of λ, Hill's family (1.7) corresponds to several classical diversity measures, namely species richness $J_t(0) = S$, the exponential of Shannon's entropy $J_t(1) = e^{H_t}$, and the 'inverse' Simpson index $J_t(2) = 1/D_t$ (Hill, 1973). However, all existing families still combine information on both richness and evenness, where richness is typically equated to the number of observed species (observed species richness) if J_t is evaluated for a sample. Typically, the choice of the parameter value decides about the weight put on rare and common species. Values close to 0 give equal weight to all species and usually $\lambda = 0$ correponds to the number of species observed (or a transformation thereof). Values greater than one shift more and more focus towards dominant species; in the limit $\lambda \to \infty$ Hill's family

equals the Berger-Parker-Index (Berger & Parker, 1970) which is the reciprocal of the most abundant species' proportion (in a sample).

Hill (1973) gives an interpretation of $J(\lambda)$ as the 'effective number of species'. For $\lambda = 0$, the index is the actual number of species, so that every species contributes equally to the count. As $\lambda \to \infty$, fewer and fewer species are counted 'effectively' as the free parameter shifts more weight from the least abundant species to the dominant species. If species are very similar in their abundances (close to evenness), there is less change in weight and effectively all species are counted independently of the value of λ. (This independence of λ with increasing evenness will be exploited in the following section where we discuss diversity profiles). This also provides means of assessing evenness of a community of interest by looking at quotients $J(a)/J(b)$ for some $a, b > 0, a \neq b$.

Jost (2006) recently gave the concept of effective number of species a more precise meaning: If we calculate a diversity index for an arbitrary community, then in theory we can find a completely even community for which the diversity index would give the same value. (Mathematically, we can say that the diversity index defines an equivalence relation on the 'set' of all possible finite communities). The number of species in this completely even community (which is unique) is then defined as the effective number of species.

While the geometric mean of relative abundances (1.6) is different from measures based on species proportions, it could also be generalised in a parametric form and put into context to other means, such as, for example, the arithmetic mean or the harmonic mean. The *generalised mean* of relative abundances is defined by

$$M_t(\lambda) = \left(\frac{1}{S} \sum_{i=1}^{S} \left(\frac{N_i^t}{N_i^{t_0}} \right)^{\lambda} \right)^{\frac{1}{\lambda}}, \tag{1.8}$$

where now λ can be any non-zero real number. In the limit $\lim_{\lambda \to 0} M_t(\lambda) = G_t$. For $\lambda = -1$ and $\lambda = 1$ we get the harmonic and the arithmetic mean, respectively. Furthermore, $\lim_{\lambda \to \infty} M_t(\lambda) = \max \left(N_1^t/N_1^{t_0}, \ldots, N_S^t/N_S^{t_0} \right)$ and $\lim_{\lambda \to -\infty} M_t(\lambda) = \min \left(N_1^t/N_1^{t_0}, \ldots, N_S^t/N_S^{t_0} \right)$, similar to the Berger-Parker-index being the limit of Hill's numbers. Even more general formulations exist; the weighted general mean $M_t(\lambda) = \left(\frac{1}{w} \sum_{i=1}^{S} \left(w_i \frac{N_i^t}{N_i^{t_0}} \right)^{\lambda} \right)^{\frac{1}{\lambda}}$, where $w = \sum_{i=1}^{S} w_i$ are weights given to the each species, gives in particular a weighted geometric mean $G_t = \exp \left(\frac{1}{w} \sum_{i=1}^{S} \log \left(\frac{N_i^t}{N_i^{t_0}} \right)^{w_i} \right)$. A related generalisation is that of the
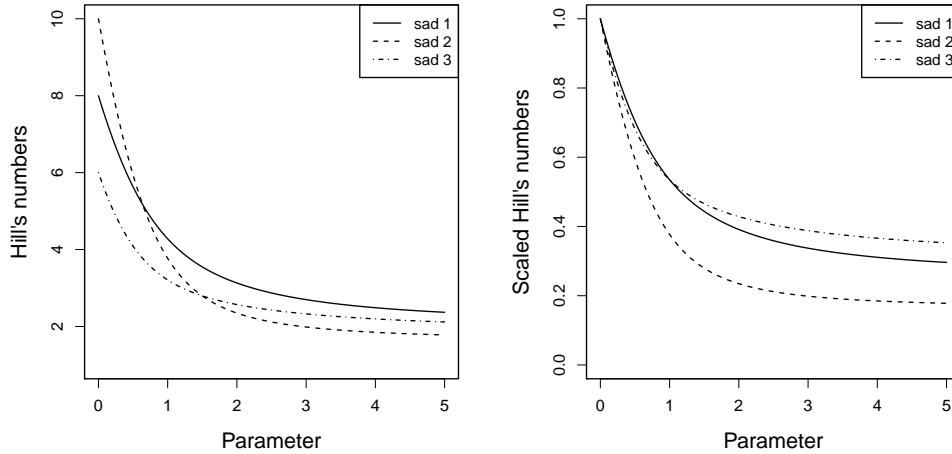
FIGURE 1.4 – Example of diversity profiles for 3 different species abundance distributions – sad 1 = (1,1,1,2,5,7,13,30), sad 2 = (1,1,1,1,2,4,5,7,12,58) and sad 3 = (1,1,1,7,15,30). The left-hand side shows Hill's untransformed diversity numbers, the right-hand side the scaled version $J(a)/J(0)$ (i.e. scaled by number of species). Even in the scaled version, comparability is limited due to the differences in species richness.

quasi-arithmetic or generalised f-mean

$$M_t(f) = f^{-1}\left[\left(f\left(\frac{N_1^t}{N_1^{t_0}}\right) + \ldots + f\left(\frac{N_S^t}{N_S^{t_0}}\right)\right)/S\right], \qquad (1.9)$$

with f being an arbitrary continuous and injective function into the real numbers $\mathbb{R}$. Again, we can find the three standard means, the harmonic (for $f(x) = \frac{1}{x}$, the arithmetic (for $f(x) = ax + b, \quad a \neq 0$) and finally the geometric mean (for $f(x) = \log x$).

1.3.3 Diversity profiling

At first sight, it is not evident what we gain from the description of diversity indices as a function of a free parameter. However, this becomes clearer when we look at the graph of this function, known as *diversity profile*, i.e. the plot of the values of the diversity measures in the family against parameter values λ (Hill, 1973; Patil & Taillie, 1982; Tóthmérész, 1995).

Diversity profiles plot indices against parameter values and provide a graphic display of the whole family. Thus we can access more information than that contained in a single index value. This provides us with a 'partial order': if the diversity profiles for two species abundance distributions do not intersect, one of them can be identified as the more diverse without any ambiguity. However, as $\lambda = 0$ typically corresponds to species richness, diversity profiles for two communities with different numbers of species cannot be easily compared. This

is usually met by dividing the whole family by $J_t(0) = S$ to scale the initial point to one (Heip, 1974; Pielou, 1975). (In fact, often the number of species is equated with those that are observed in a sample, i.e. $J_t(0) = S_{\text{obs}}$.) Comparability is still compromised, though, as the underlying reference point (S) differs for the two communities.

As a community gets more even, the values of a parametric family become increasingly independent of the value of λ (Hill, 1973). This is represented in the corresponding profile being almost flat. Hence, the curvature of the profile is informative in terms of a community's evenness (see chapter 2). An example of a diversity profile is shown in Fig. 1.4.

1.4   Biodiversity monitoring & survey design

With the CBD's 2010 target, a new objective for biodiversity assessment was set. The international agreement as well as its extension in the 2011-2020 Aichi targets focusses on biodiversity across large spatial scales, namely entire regions and countries. This has implications on the design of monitoring programmes, in particular the way data are collected, i.e. survey design (Dobson, 2005; Buckland *et al.*, In prep).

Traditionally, biodiversity studies have often been limited in their spatial extent, concentrating on few single monitoring sites, and diversity indices were calculated directly from the observed species records (Boulinier *et al.*, 1998; Yoccoz *et al.*, 2001). This implicitly assumes that a sample is representative of the community of interest and that all individuals have an equal probability of being detected. Already debatable for single site monitoring in general, this approach is clearly bound to fail when we want to consider the biodiversity of a whole region or country. On a small spatial scale and by looking only at the local community, it is possible to get an exhaustive sample if survey plots are sufficiently small (such that we can assume a more or less complete census within the plot) and if there are sufficiently many plots (to ensure most of the species within the local community are recorded). However, even on a local scale this is limited to mostly larger, well known vertebrate species (Landres *et al.*, 1988). For more cryptic and smaller organisms and very heterogeneous environments, as for example microbes (Hughes *et al.*, 2001) or tropical arthropod communities (Longino *et al.*, 2002), it will be neither clear which species belong to the local community nor possible to get a full census; nevertheless, rarefaction techniques have been applied with some success in both cases. If extended to a regional, national or even global scale it is obvious that exhaustive sampling

is not feasible and that we have to rely on a selection of sites to draw inference for the whole region. Likewise biodiversity assessment should be aimed at the entire biota of the region or country. This raises the question of representativeness for both sites and species sampled (Buckland *et al.*, 2011a, In prep). A representative sample of sites can be ensured by a carefully designed random sampling scheme (Magurran *et al.*, 2010). For example, the UK Breeding Bird Survey (Riseley *et al.*, 2011) follows a stratified random sampling design (see chapter 4). Random allocation of survey sites makes estimation of precision in the chosen summary statistics possible, while stratification can eliminate or at least reduce bias stemming from differences in sampling effort or heterogeneity from structural changes in the environment across the region.

To determine a representative list of species is less straightforward. A catalogue of a country's entire biota will in most cases remain incomplete; especially in biodiversity hotspots which are extremely species-rich, while at the same time the majority of diversity hotspots are located in developing countries where resources for monitoring are low. The species catalogue will be affected by seasonal changes as well as longterm turnover caused by natural or anthropogenic changes to the environment (Magurran *et al.*, 2010). In addition, a survey method that might be adequate for sampling some species, can be insufficient for others. Due to its specific design, the UK Breeding Bird Survey, for example, leaves nocturnal and coastal species undersampled. Regional biodiversity monitoring therefore relies on a chosen set of taxa from the entire biota that will necessarily be reduced. (For further discussion see chapter 3.) The concept of indicator species provides one possiblity to direct this choice (Lawler *et al.*, 2003; Pereira & Cooper, 2006). The task of monitoring can then be further facilitated by considering only a random subset of the set list; this is particularly indicated where resources for monitoring programmes are limited. However, there is the risk that species which might be less easily sampled, will be underrepresented in such lists (Landres *et al.*, 1988; Lawton *et al.*, 1998).

Nevertheless, even after the decision on a set of species to monitor, in general these species still differ in their probability of detection (see 1.2.3 above and chapters 3 and 4). If not accounted for, this leads to bias in the estimation of diversity (Yoccoz *et al.*, 2001; Buckland *et al.*, 2011a). When planning a survey to monitor regional biodiversity, attention should therefore be paid to basing it on methods that allow for explicit estimation of detection probabilities and hence a correction of the diversity indices. Distance sampling (Buckland *et al.*, 2001, 2004a), as for example applied in the UK Breeding Bird Survey, is one option, and is useful for species that are easily detected and identified at least

in the proximity of the observer. It will be discussed in more detail in chapter 4. Other alternatives are mark-recapture methods which can be used where individuals can be trapped, and in rare cases removal and other catch methods (Borchers *et al.*, 2002).

Rare species pose a problem to any kind of sampling scheme that is not exhaustive. They will be missing from samples most of the time and seldom provide sufficient data when they are observed, for drawing solid statistical inference (Cunningham & Lindenmayer, 2005). Estimation of detection probabilities is often out of reach and precision of diversity measures that are very sensitive to rare species is substantially reduced (Buckland *et al.*, 2011b). Focal surveys which collect additional information on species that tend to be underrecorded in a nationwide monitoring programme can supplement biodiversity assessment and help to reduce this problem, at least partially. However, the question of survey compatibility arises. This is a wider issue for large scale biodiversity monitoring in general, as we require surveys with a broad temporal and spatial extent in order to assess the 2010 target and its follow-up targets (Magurran *et al.*, 2010). Long time series of data on a national level from an adequate sampling design, however, are sparse. Sampling schemes like the UK Breeding Bird Survey are still the exception and were only recently established. As long as such monitoring programmes are not set in place more widely and cover a range of taxa, the only option is a sensible combination of available data from different surveys and probably different survey methods. But not all diversity measures readily allow for such an integration; the geometric mean has a strong advantage here in that it can easily combine different surveys (Buckland *et al.*, 2011b). This puts it in favour as a headline index on a global scale, such as the Planet of Life index (Loh *et al.*, 2005).

1.5   Chapter summary

The intent of this chapter was to review particular issues that arise when we want to assess biodiversity. With the international biodiversity targets in mind, we consider mainly monitoring and assessment over large spatial regions and trends in their diversity in time. Assessment includes in particular measurement. We have discussed biodiversity as a multidimensional concept and its various measurable aspects. Concentrating on species diversity, different methods of quantifications have been reviewed. Diversity indices usually have to be estimated based on survey data. While good survey design is important

to reduce bias in these estimates, sampling ecological communities comes with specific problems that should not be overlooked. Notably, rarity of species will lead to sampling zeros which complicate diversity assessment. We have focussed on problems that play a role on large spatial scales and over longer periods of time and will return to an example how to address some of them in chapter 4. This chapter has set the ground for the following in that it has summarised general problems that any method of biodiversity quantification faces (multidimensionality, sampling aspects, detectability) and those that are particular to the international biodiversity targets (temporal trends, spatial aspects, survey design). We will return to some of them in the chapters to come, be it to evaluate methods that are developed in the course of this work or to suggest ways to overcome them partly.

Chapter 2

Goodness-of-fit measures of evenness: a new tool for exploring changes in community structure

Biodiversity on the species level summarises the structure within an ecological assemblage, taking into account its size (species richness) as well as dominance and rarity of species (see section 1.2.1 in chapter 1). As a multivariate concept, biodiversity cannot be well-represented by a single measure (as discussed in sections 1.3.1 and 1.3.2 in the previous chapter). Any scalar index typically combines information on different characteristics of the assemblage. However, as discussed in section 1.3.3, diversity profiles summarise the multivariate nature of multi-species datasets, and allow a more nuanced interpretation of biodiversity trends than unitary metrics.

Although several families of diversity indices have been suggested (see section 1.3 and Tóthmérész (1995) for an overview), families of evenness measures have rarely been considered and none has found wide application in ecological studies (Ricotta, 2003). However, quantification of equality and inequality has a long tradition in information theory as well as economics (Kullback, 1968; Hoffmann, 2008). Although derived from a statistical perspective here, the methods considered in this chapter are closely related to inequality measures used by economists (Cowell, 1980).

In this chapter, we investigate goodness-of-fit statistics as measures of diversity. Based on the knowledge that an ecological community is never completely even, this approach uses this departure from perfect evenness as a novel and

insightful way of measuring diversity (Studeny *et al.*, 2011).

After motivating this idea in terms of the two most commonly used goodness-of-fit statistics, we give a generalisation in the form of a one-parameter family of statistics and discuss its theoretical properties.

We study how this approach relates to existing ecological indices: it provides a generalisation in that it contains transformations of the classical measures, Simpson's and Shannon's index. Consequently, we want to examine how it is connected to other generalisations of these indices. In this context, Hill's diversity numbers (Hill, 1973) have recently regained interest in ecology (Jost, 2006, 2010; Jost *et al.*, 2011; Leinster & Cobbold, 2012) and are used here as the main reference. However, connections are not restricted to the ecological context — we discuss a similar approach that is well established in economics.

As with diversity profiling (see 1.3.3), we plot this measure of departure as a function of a free parameter, to generate 'evenness profiles'. These profiles allow us to separate changes due to dominant species from those due to rare species, and to relate these patterns to shifts in overall diversity. This separation of the influence of dominance and rarity on overall diversity enables the user to uncover changes in diversity that would be masked in other methods. In this context, we also explore their behaviour with regard to ecological criteria for evenness indices (Smith & Wilson, 1996).

In a first application, we evaluate this goodness-of-fit based method in terms of predicted community structure (following Tokeshi's niche models) and present an example assessing temporal trends in diversity of British farmland birds. We conclude that it is an informative and tractable parametric approach for quantifying evenness. It provides novel insights into community structure, revealing the contributions of both rare and common species to biodiversity trends. This will be used in chapter 4 for an extensive study of biodiversity trends in British breeding birds.

Major parts of the work presented in this chapter have been published in Studeny *et al.* (2011).

2.1   General idea: Perfect evenness as the yardstick

No real ecological community is perfectly even; the contrary is actually true (Rabinowitz *et al.*, 1986). Species abundance distributions are so remarkably similar in their shape showing few very abundant species and many rare ones

that this 'hollow curve' shape is assumed to follow a universal pattern (McGill *et al.*, 2007).

This has been exploited to measure evenness contained in the species abundance distribution: we can take the uniform distribution that represents perfect evenness as a 'null model' (Gotelli & Graves, 1996) and quantify its *divergence* from the species abundance distribution of the assemblage of interest. Divergence measures have been widely used in information theory (Kullback, 1968; Rao, 1982; Ricotta & Avena, 2002), where they have been developed into a very general framework (Karagrigoriou & Mattheou, 2010). While the focus is on the discrepancy from the null model there, the similarity of two distributions could be assessed equivalently. As discussed briefly in chapter 1.3, the latter approach is used by Camargo (1995) and Bulla (1994) to derive a scalar measure by comparing the area under the species abundance distribution to that under a uniform distribution.

In the following, we adopt an approach to quantifying divergence based on goodness-of-fit statistics which is inherently sample-based, but provides an estimate of the true divergence. Given a sample, we can ask ourselves how likely it is that this sample came from a completely even species abundance distribution. Goodness-of-fit statistics offer a natural way of doing so. Considering the two most commonly used goodness-of-fit statistics, the likelihood ratio G and Pearson's X^2, we look at

$$G \;=\; 2\sum_{i=1}^{S} n_i \log\left(\frac{n_i S}{n}\right), \tag{2.1}$$

$$X^2 \;=\; \sum_{i=1}^{S} \frac{(n_i - n/S)^2}{n/S}. \tag{2.2}$$

Instead of the standard comparison to a χ^2-distribution, we use the values of these statistics as an estimate of the degree of divergence of the true species abundance distribution from perfect evenness. More precisely, take $\sum_{i=1}^{S} p_i \log\left(p_i/p_i^*\right)$ and $\sum_{i=1}^{S} p_i\left[(p_i/p_i^*) - 1\right]$ as divergence measures between $\mathbf{p}$ and $\mathbf{p}^*$ (the former is the well-known Kullback-Leibler divergence, Kullback (1968)). Then $G/2n$ and X^2/n provide asymptotically unbiased maximum likelihood estimators of these divergences, respectively (as will be shown in a more general context in section 2.3 below). In fact, these two statistics are transformations of the most prominent classical diversity measures, Shannon's index and Simpson's index (see section 2.3).

This idea can be generalised; Cressie & Read (1984) introduce and study a one-parameter family of goodness-of-fit statistics which incorporates both G and X^2. This family enables us to extend these (scalar) goodness-of-fit 'measures' of evenness to a parametric approach in the next section.

2.2   A family of goodness-of-fit statistics

Introducing a free parameter λ, T. Read and N. Cressie derived the following parametric form for a generalised goodness-of-fit statistic (Cressie & Read, 1984; Read & Cressie, 1988)

$$I_{\mathbf{n}}(\lambda) = \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{S} n_i \left[\left(\frac{n_i}{n/S} \right)^{\lambda} - 1 \right]. \tag{2.3}$$

By changing values for λ in this expression, we switch between different goodness-of-fit statistics. Although equation (2.3) does not define $I_{\mathbf{n}}(\lambda)$ for $\lambda = -1$ and $\lambda = 0$, limits $\lambda \to -1$ and $\lambda \to 0$ can be taken. Parameter values $\lambda = 0$ and $\lambda = 1$ give, as special cases, the statistics G and X^2, respectively. (For details and proofs see section 2.3 below).

Analogously to G and Pearson's X^2, when divided by $2n$ this family of goodness-of-fit statistics provides an estimator of a measure of divergence between the true species abundance distribution $\mathbf{p}$ and the perfectly even distribution $\mathbf{p}^*$ (Read & Cressie, 1988). This divergence measure quantifies the departure of $\mathbf{p}$ from evenness using a parametric form

$$I_{\mathbf{p}}(\lambda) = \frac{1}{\lambda(\lambda+1)} \sum_{i=1}^{S} p_i \left[\left(\frac{p_i}{p_i^*} \right)^{\lambda} - 1 \right]; \tag{2.4}$$

and we have

$$\hat{I}_{\mathbf{p}}(\lambda) = \frac{1}{2n} I_{\mathbf{n}}(\lambda). \tag{2.5}$$

Read & Cressie (1988) focus on the statistical aspects of the family of goodness-of-fit statistics in equation (2.3) (distributional properties, asymptotics, their statistcal power amongst others). In particular, they discuss its connection to information theoretic divergence measures, e.g. the Kullback-Leibler-divergence. They refer to the family of divergence measures given by (2.4) as the 'power-divergence family'.

By varying the parameter λ, we change the weights given to the terms in

the sum and hence their contributions to the index; in ecological applications, this corresponds to the influence of different species on overall diversity. By recording the changes in index values with λ, we can infer information about evenness. If the data exhibit perfect evenness, the estimator is independent of λ. The greater the departure from evenness, the more pronounced the changes in index values with λ are. This is an important feature which enables us to build evenness profiles in section 2.6.1 below.

2.3   Mathematical and statistical properties of the family of goodness-of-fit statistics

In section 1.2.1, desirable properties of a diversity index have been discussed in the light of the multidimensional nature of the concept. This included statistical properties as well as how an index reacts to changes in the different components of diversity, i.e. species richness, overall abundance and evenness. As diversity assessment is generally based on samples from the community of interest, statistical properties of a diversity estimator are crucial (Lande, 1996; Buckland *et al.*, 2005).

In this section, we will focus on the family of goodness-of-fit measures given by equation (2.3) as an estimator of the family of divergences (2.4), and establish its mathematical and statistical properties. Its behaviour as a sensible measure of diversity is discussed further in section 2.6.2, where we exploit it in terms of diversity profiling. (This section is mostly based on Read & Cressie (1988).)

2.3.1   Continuity at $\lambda = -1$ and $\lambda = 0$

The first point to note when looking at expressions (2.3) and (2.4) is that neither of them is defined at $\lambda = -1$ and $\lambda = 0$. This limitation is overcome by taking limits

$$
\begin{aligned}
\lim_{\lambda \to 0} I_{\mathbf{p}}(\lambda) &= \lim_{\lambda \to 0} \frac{1}{\lambda(\lambda+1)} \sum_{i=1}^{S} p_i \left[\left(\frac{p_i}{p_i^*} \right)^{\lambda} - 1 \right] \\
&= \sum_{i=1}^{S} p_i \lim_{\lambda \to 0} \frac{1}{\lambda} \left[\left(\frac{p_i}{p_i^*} \right)^{\lambda} - 1 \right] \\
&= \sum_{i=1}^{S} p_i \log \left(\frac{p_i}{p_i^*} \right)
\end{aligned}
$$

and

$$
\begin{aligned}
\lim_{\lambda \to -1} I_{\mathbf{p}}(\lambda) &= \lim_{\lambda \to -1} \frac{1}{\lambda(\lambda+1)} \sum_{i=1}^{S} p_i \left[\left(\frac{p_i}{p_i^*} \right)^{\lambda} - 1 \right] \\
&= \lim_{\lambda \to 0} \frac{1}{(\lambda-1)\lambda} \sum_{i=1}^{S} p_i \left[\left(\frac{p_i}{p_i^*} \right)^{\lambda-1} - 1 \right] \\
&= \lim_{\lambda \to 0} \frac{1}{(\lambda-1)\lambda} \sum_{i=1}^{S} p_i^* \left[\left(\frac{p_i}{p_i^*} \right)^{\lambda} - 1 \right] \\
&= -\sum_{i=1}^{S} p_i^* \log \left(\frac{p_i}{p_i^*} \right),
\end{aligned}
$$

where in each case we used

$$
\lim_{h \to 0} \frac{1}{h} \left(x^h - 1 \right) = \frac{\mathrm{d}}{\mathrm{d}h} x^h \bigg|_{h=0} = \frac{\mathrm{d}}{\mathrm{d}h} \exp(h \log x) \bigg|_{h=0} = \log x.
$$

Analogous results can be derived for the family of estimators $I_{\mathbf{n}}(\lambda)$ given by equation (2.3) by replacing p_i with its standard estimator n_i/n in the equations above.

However, one has to be careful if samples with zeros are involved. While we can still derive finite results for $\lambda = 0$ because $\lim_{x \to 0} x \log x = 0$, the limit for $\lambda = -1$ no longer exists. There are two possible solutions to this problem, namely either by considering only parameter values strictly above -1 or by adding a small quantity to the sample proportions for all species in order to regain non-zero values. Both approaches will be discussed explicitly in section 2.6.1 and studied in applications in section 2.7.

2.3.2 G and X^2 as part of the family $I_{\mathbf{n}}(\lambda)$ of goodness-of-fit statistics

As mentioned, the classic goodness-of-fit statistics — Pearson's X^2 and the likelihood ratio G — are part of the family $I_{\mathbf{n}}(\lambda)$ for $\lambda = 0$ and $\lambda = 1$, respectively:

$$
\begin{aligned}
\lim_{\lambda \to 0} I_{\mathbf{n}}(\lambda) &= 2 \sum_{i=1}^{S} n_i \log \left(\frac{n_i}{n/S} \right) = 2\,G, \\
I_{\mathbf{n}}(1) &= \sum_{i=1}^{S} n_i \left(\frac{n_i}{n/S} - 1 \right) = \sum_{i=1}^{S} \frac{(n_i - n/S)^2}{n/S} = X^2.
\end{aligned}
$$

We will see below that these two prominent goodness-of-fit statistics are

indeed connected to the two most common diversity indices — Shannon's index (Shannon, 1948) and Simpson's index (Simpson, 1949).

2.3.3 Maximum likelihood estimation and asymptotics of the family of goodness-of-fit measures

Under the assumption of multinomial sampling, the statistical theory of the family of goodness-of-fit statistics (2.3) is well studied (Cressie & Read, 1984; Read & Cressie, 1988). In particular, their asymptotic behaviour is known. Based on aymptotic normality of the components of the sampling vector, the following results were proved by Read & Cressie (1988).

If $(n_1, \ldots, n_S)$ is the realisation of a multinomial $\mathcal{M}(n, \mathbf{p}^*)$ from the perfectly even species abundance distribution $\mathbf{p}^*$ and assuming large enough sample size n, all members of the family are equivalent with regard to their asymptotic distribution, which is χ^2_{S-1}. In particular we have

$$\mathbb{E}_{\mathbf{p}=\mathbf{p}^*}\left[I_{\mathbf{n}}(\lambda)\right] \approx S - 1, \tag{2.6}$$

$$\mathbb{V}_{\mathbf{p}=\mathbf{p}^*}\left[I_{\mathbf{n}}(\lambda)\right] \approx 2\,(S - 1)\,. \tag{2.7}$$

Typically, the focus of any goodness-of-fit statistic is on the equality of the distributions compared. However, here we know that in reality the species abundance distribution will never be perfectly even, therefore we are more interested in the distributional properties when $\mathbf{p} \neq \mathbf{p}^*$. In this case, the distribution of the statistics in the goodness-of-fit family is not independent of λ, but can be approximated by a normal distribution, where the mean and variance are given by

$$\mathbb{E}_{\mathbf{p}\neq\mathbf{p}^*}\left[I_{\mathbf{n}}(\lambda)\right] \approx \frac{2n}{\lambda(\lambda+1)} \sum_{i=1}^{S} p_i\left[(p_i S)^\lambda - 1\right],$$

$$\mathbb{V}_{\mathbf{p}\neq\mathbf{p}^*}\left[I_{\mathbf{n}}(\lambda)\right] \approx \frac{4n}{\lambda^2}\left[\sum_{i=1}^{S}(p_i S)^{2\lambda} p_i - \left[\sum_{i=1}^{S}(p_i S)^\lambda p_i\right]^2\right].$$

Hence, for large n we have

$$\mathbb{E}_{\mathbf{p}\neq\mathbf{p}^*}\left[\frac{1}{2n}I_{\mathbf{n}}(\lambda)\right] \longrightarrow I_{\mathbf{p}}(\lambda).$$

As $\hat{p}_i = n_i/n$ is the usual maximum likelihood estimator, the estimator $\frac{1}{2n}I_{\mathbf{n}}(\lambda)$ is a maximum likelihood estimator (invariance property of the maxi-

mum likelihood estimator). Because of its asymptotic behaviour, it is consistent and asymptotically unbiased (in the setting of multinomial sampling). Later, we explore the consequences of dropping the assumption of a multinomial sampling distribution for a setting which is more realistic in terms of ecological applications (see chapter 3).

2.3.4　The goodness-of-fit measures in a simplex setting

Discrete probability distributions in S dimensions can be interpreted as points in the $(S-1)$-simplex $\Delta = \{\mathbf{p} \in \mathbb{R}^S | \sum_i p_i = 1\}$ (see Box 2.1 for details). In this mathematical framework the search for an evenness measure translates to the definition of an appropriate metric $d : \Delta \times \Delta \to \mathbb{R}$ on the simplex to quantify the distance of an arbitrary point in the simplex to the barycentre which corresponds to the point $\mathbf{p}^*$. To be a metric in a strict mathematical sense, d has to have certain properties (see Box 2.1), one of which is symmetry, i.e. $d(\mathbf{p}, \mathbf{p}^*) = d(\mathbf{p}^*, \mathbf{p})$. For the symmetry property to hold for arbitrary $\mathbf{p} \in \Delta$ we need

$$\sum_{i=1}^S p_i \left[\left(\frac{p_i}{p_i^*} \right) - 1 \right] = \sum_{i=1}^S p_i^* \left[\left(\frac{p_i^*}{p_i} \right) - 1 \right]$$

$$\Leftrightarrow \sum_{i=1}^S \left[\frac{p_i^{\lambda+1}}{p_i^{*\,\lambda}} - p_i - \frac{p_i^{*\,\lambda+1}}{p_i^{\lambda}} + p_i^* \right] = 0 \Leftrightarrow \sum_{i=1}^S \left[p_i^{2\lambda+1} - p_i^{*\,2\lambda+1} \right] = 0.$$

As $\mathbf{p}^*$ is fixed, this holds if and only if $p_i^{2\lambda+1} = p_i^{*\,2\lambda+1} = 1 \Leftrightarrow \lambda = -1/2$.

Consequently, in general the members of the family of goodness-of-fit measures do not define a metric between $\mathbf{p}$ and the barycentre $\mathbf{p}^*$ in a strict mathematical sense. Only for $\lambda = -1/2$ is the symmetry property fulfilled. Indeed, by applying the square root transformation $\mathbf{x} := (\sqrt{p_1}, \ldots, \sqrt{p_S})$ and $\mathbf{y} := (\sqrt{p_1^*}, \ldots, \sqrt{p_S^*})$, we see that

$$0 \le \|\mathbf{x} - \mathbf{y}\|^2 = 2 - 2 \sum_{i=1}^S \sqrt{p_i p_i^*} = 2 \left(1 - \sum_{i=1}^S \sqrt{p_i p_i^*} \right) = \frac{1}{2} I_{\mathbf{p}} \left(-1/2 \right).$$

Hence, $I_{\mathbf{p}}(-1/2)$ inherits its properties as a metric from the Euclidean distance in $\mathbb{R}^S$ and we conclude: $I_{\mathbf{p}}(\lambda)$ defines a metric on the $(S-1)$-simplex if and only if $\lambda = -1/2$. Its square root is called *Matusita* distance.

Box 2.1: Excursion into metric spaces — simplices and norms

Indices for species diversity are commonly based on the species propor-tions $\mathbf{p}$. While diversity indices can be derived as variance or as infor-mation contained in $\mathbf{p}$, we can also use a geometric setting to describe $\mathbf{p}$ and its diversity. This is of particular interest for determining evenness within a community or differences between communities.

$$\Delta_{S-1} = \left\{ (p_i)_{1 \leq i \leq S} \in \mathbb{R}^S \,\big|\, \sum_{i=1}^{S} p_i = 1 \right\}$$

is the collection of all possible vectors $\mathbf{p}$ (species abundance distribu-tions) corresponding to communities of S species. Together, the vectors form a subspace of $\mathbb{R}^S$ which is called the $(S-1)$-*simplex*. Hence, we can find any community with S species represented by a point in the sim-plex. In particular, the centre of gravity (*'barycentre'*) of the simplex corresponds to the completely even community $\mathbf{p}^*$.

Some evenness measures are based on the idea of 'divergence' between the communities represented by $\mathbf{p}$ and $\mathbf{p}^*$. This divergence is reflected in the distance between the two points in the simplex. Hence, an evenness measure can be defined by an appropriate metric (or distance function) on the simplex. Mathematically, a metric is defined as a function $d : \Delta_{S-1} \times \Delta_{S-1} \to \mathbb{R}_0^+$ that fulfils the following axioms:

- $d(p_1, p_2) = 0 \leftrightarrow x = y$ (identity of indiscernibles)
- $d(p_1, p_2) = d(p_2, p_1)$ (symmetry)
- $d(p_1, p_3) \leq d(p_1, p_2) + d(p_2, p_3)$ (triangle inequality).

These correspond to our intuitive way to think about distance between two points. For example, a common metric can be derived from the Euclidean norm on $\mathbb{R}^S$: $d(p_1, p_2) = \|p_1 - p_2\| = \sqrt{\sum_{i=1}^{S} (p_{i_1} - p_{i_2})^2}$.

In this setting, Simpson's index $D = \sum_{i=1}^{S} p_i^2$ gives the squared Euclidean length of the vector $\mathbf{p}$. It is not a divergence measure since it is does not establish the distance between $\mathbf{p}$ and a reference point within the simplex. However, it does account for evenness indirectly, in that the closer $\mathbf{p}$ is to $\mathbf{p}^*$, the smaller its length is. In reality, we estimate diversity based on sample(s). The vector of species proportions in the sample can again be located in the simplex; however, it might lie in a subsimplex if not all species are observed. It is immediately clear that determining diversity of the sample and equating the result with the evenness of the underlying community generally leads to bias.

Figure showing the 3-simplex.
We can locate any community with 4 species on this simplex through its species abundance distribution. The endpoints correspond to the most uneven communities; the barycentre (red) marks the completely even community.

Although goodness-of-fit statistics do not define a metric on the simplex in general (see main text), they can give an idea of the 'distance' to the reference point (here, the perfectly even $\mathbf{p}^*$): if a sample is not from the completely even distribution, but a different $\mathbf{p}$, the value of the goodness-of-fit statistics is likely to be larger, the further away $\mathbf{p}$ is of $\mathbf{p}^*$. However, the simplex representation has another feature — it keeps track of species identities while these are neglected by diversity indices as summaries of $\mathbf{p}$ in general. For example, $p_1 = (0.2, 0.6, 0, 0.2)$ and $p_2 = (0.2, 0.2, 0.6, 0)$ are not distinguished by any of the common diversity indices despite being clearly different points in the simplex (see figure).

In fact, this is not an issue as long as we are only interested in quantifying evenness as the divergence from the (fixed) point $\mathbf{p}^*$; the situation where we want to swap $\mathbf{p}$ and $\mathbf{p}^*$ does not arise. However, since the identification of ecological communities with points in the simplex allows us to keep track of species identities (see Box 2.1), we might use the divergence between two arbitrary points as a measure of their similarity in terms of species composition. In this case, we are interested in this measure being, for example, symmetric.

If we are more interested in following changes in a community through time (temporal turnover), this could be represented by a point moving through the simplex as a state process. When we survey the community, we get an observation based on the underlying state at the time and some sampling distribution.

2.4   Connection to Shannon's and Simpson's index and other possible transformations

As indicated in section 2.1 above, the two most prominent members of the family of goodness-of-fit statistics given by equation (2.3) in 2.2, the likelihood ratio G and Pearson's X^2 are related to traditional diversity indices, namely Shannon's entropy $H = -\sum p_i \log p_i$ and $D = \sum p_i^2$, which is the basis for any

version of Simpson's index. More precisely, we can express G and X^2 as linear transformations

$$
\begin{aligned}
G &= 2n\left[\sum_{i=1}^{S}\frac{n_i}{n}\log\left(\frac{n_i}{n}\right)+\log S\right]=2n\left(\log S-\hat{H}\right)=2n\left(H^*-\hat{H}\right),\\
X^2 &= n\left[\sum_{i=1}^{S}\left(\frac{n_i}{n}\right)^2 S-1\right]=n\left(\hat{D}S-1\right)=n\left(\frac{\hat{D}}{D^*}-1\right)
\end{aligned}
$$

of the sample-based estimates $\hat{H}$ and $\hat{D}$ of Shannon's and Simpson's indices. $H^*=\log S$, $D^*=1/S$ are the index values if the species abundance distribution is completely even ($\mathbf{p}=\mathbf{p}^*$). (Note that the estimator $\hat{D}$ is not corrected for sampling bias here (Lande, 1996).)

From the equations above, we see that the goodness-of-fit statistics measure the difference between the value of the respective diversity index and the value expected under complete evenness. While G does this in additive terms, X^2 is based on a ratio. In terms of interpretability and familiarity with these indices, it would be desirable to find a linear transformation of $I_\mathbf{p}(\lambda)$ which directly gives Shannon's index H for $\lambda=0$ and a version of Simpson's index for $\lambda=1$. However, this is infeasible because of the different nature (additive/multiplicative) of the relationship of $I_\mathbf{p}(\lambda)$ to H and D, respectively. If Simpson's index is considered to be preferable to Shannon's index, we can choose the transformation

$$
\tilde{I}_\mathbf{p}(\lambda)=\log\frac{S}{(\lambda+1)I_\mathbf{p}(\lambda)+\lambda},\tag{2.8}
$$

which gives

$$
\tilde{I}_\mathbf{p}(1)=\log S-\log(2I_\mathbf{p}(1)+1)=\log S-\log SD=-\log D
$$

and

$$
\tilde{I}_\mathbf{p}(0)=\log S-\log(\lim_{\lambda\to 0}I_\mathbf{p}(\lambda))=\log S-\log(\log S-H).
$$

This yields a commonly used version of Simpson's index for $\lambda=1$, but obscures the relationship to H. By taking the exponential of $\tilde{I}_\mathbf{p}(\lambda)$, we still get $1/D$ as a commonly used version of Simpson's index, while the expression for H is closer to the original.

On a different matter, low values of $I_\mathbf{p}(\lambda)$ correspond to high evenness; this might be perceived as counter-intuitive in an evenness measure. A simple transformation that orients the measures in the family in the same direction as standard diversity indices, is by taking the reciprocal $1/I_\mathbf{p}(\lambda)$. However, this is

still not restricted in its range. For reasons of standardization and comparison it might be desirable to have an evenness measure that ranges from 0 (if all except one p_i equal to 0) to 1 (for perfect evenness). A transformation that can be used, should an application require a measure of diversity with more traditional range of support, is given by

$$\rho_{\mathbf{p}}(\lambda) = 1 - \frac{I_{\mathbf{p}}(\lambda)}{\max I_{\mathbf{p}}(\lambda)}, \quad \lambda > -1, \tag{2.9}$$

where $\max I_{\mathbf{p}}(\lambda) = 1/[\lambda(\lambda+1)](S^\lambda - 1)$, i.e. the value of $I_{\mathbf{p}}(\lambda)$ for the least even abundance vector $\mathbf{p} = (1, 0, \ldots, 0)$. However, applications show that changes in both $1/I_{\mathbf{p}}(\lambda)$ and $\rho_{\mathbf{p}}(\lambda)$ with varying λ are not as easily interpreted as in the untransformed index family (see section 2.7 below). In addition, $\rho_{\mathbf{p}}(\lambda)$ is only defined for $\lambda > -1$ and differentiates well between abundance distributions only for $-1 < \lambda < 1$.

A different approach would be to use the expectation under the null model of perfect evenness, $\mathbb{E}_{\mathbf{p}=\mathbf{p}^*}[I_{\mathbf{n}}(\lambda)] \approx S - 1$, as a benchmark for comparison and look at the ratio

$$\gamma_{\mathbf{n}}(\lambda) \equiv \frac{S-1}{I_{\mathbf{n}}(\lambda)} \tag{2.10}$$

for a sample $\mathbf{n} = (n_1, \ldots, n_S)$ as a measure of divergence from the expectation under perfect evenness. However, $\gamma_{\mathbf{n}}(\lambda)$ is entirely sample-based as the standardisation uses the mean of the sample distribution of $I_{\mathbf{n}}(\lambda)$ under evenness; in this, it is not meaningful for the divergence family. Moreover, even under perfect evenness, the expression (2.10) does usually not evaluate to 1. Only if we had (many) repeated samples and considered their average, the latter would approach 1, but note that in general

$$\mathbb{E}_{\mathbf{p}=\mathbf{p}*}\left[\frac{S-1}{I_{\mathbf{n}}(\lambda)}\right] \geq \frac{S-1}{\mathbb{E}[I_{\mathbf{n}}(\lambda)]} = 1$$

by Jensen's inequality and the strict convexity of $\phi(x) = \frac{1}{x}$; indeed the sample mean can be bigger than 1, even for large sample sizes (for more detail see Appendix E).

In addition, we need to estimate S in order to calculate $\gamma_{\mathbf{n}}(\lambda)$. While this is true for the untransformed family of goodness-of-fit measures, too, (and will be discussed in more detail in chapter 3), $\gamma_{\mathbf{n}}(\lambda)$ is particularly sensitive to underestimation of S, while overestimation needs to be severe in order to affect the measure substantially (see Fig. E.1(d) in Appendix E). In view of this, an index based on $\gamma_{\mathbf{n}}(\lambda)$ does not seem to have a strong foundation. For this

reason, we only consider the transformations $1/I_{\mathbf{p}}(\lambda)$ or $\rho_{\mathbf{p}}(\lambda)$ from now on.

2.5 Connection to other index families

There are numerous measures for diversity, some of which give rise to evenness indices (Smith & Wilson, 1996). If we want to compare our approach with existing (parametric) methods, we are faced with a number of options to choose from (see chapter 1.3.2 and Tóthmérész (1995)). For parametric diversity indices it has been shown that Hill's numbers not only perform best in terms of distinguishing species abundance distributions, but also that most other parametric approaches can be derived from them (Tóthmérész, 1995). More recently, a new transformation of them, termed 'inequality factors' was introduced (Jost, 2010). This set of numbers claims to respect the interdependence and correctly separate the effect of richness from evenness. We show that our approach is closely connected to both these families of diversity measures, before we then compare their performance when they are applied to distinguish between abundance models with different degrees of evenness.

2.5.1 Hill's diversity numbers

Hill's (1973) one-parameter family of '*diversity numbers*' $J(\alpha) = \left(\sum p_i^\alpha\right)^{1/(1-\alpha)}$, $\alpha > 0$ was introduced in chapter 1.3.2 and was shown to contain several classical (scalar) diversity measures. We can rewrite the index family $I_{\mathbf{p}}(\lambda)$ in terms of Hill's diversity numbers as

$$I_{\mathbf{p}}(\lambda) = \frac{1}{\lambda(\lambda+1)} \left[\left(\frac{J(0)}{J(\lambda+1)} \right)^\lambda - 1 \right]. \tag{2.11}$$

Hence the parametric family considered here is a transformation of quotients of Hill's numbers. Hill himself suggested the use of quotients $J(a)/J(b)$ of two diversity numbers to quantify evenness - giving a value of 1 when the underlying distribution is completely even (Hill, 1973). In theory, any two real numbers a, b can be chosen. Commonly, quotients involving $J(0) = S$ are used to render the measure independent of species richness (Heip, 1974; Pielou, 1969; Camargo, 1995; Smith & Wilson, 1996). Nonetheless, Hill argues against the use of $J(0) = S$ because of the difficulty of determining S. Since our measure involves quotients based on $J(0)$, we need to justify why it is sensible to base an evenness measure on the number of species, which is usually unknown. Before we do so in section 2.6.3, we first discuss certain quotients of Hill's numbers

that have recently been emphasised in ecology (Jost, 2010) as well as a family
of inequality indices in economics that is similar to the divergence measures.

2.5.2   Inequality factors

Traditional evenness measures have typically used division by S with the aim of
achieving independence from species richness. Recently, Jost (2010) argued that
richness and evenness cannot be independent and instead proposes factorising
richness into a diversity component and an evenness component where diversity
is quantified via Hill's diversity number $J(a)$. These evenness components are
called *'inequality factors'* and are defined by the quotients $J(0)/J(a)$ — which,
given equation (2.11) above, moves them close to the index family considered
here. Jost's approach is closely connected to the interpretation of Hill's diversity
numbers as the 'effective number of species' (see chapter 1). A weighted count,
this number is always smaller than or equal to the true species richness, as rare
species contribute less and less to it with increasing parameter value. The more
uneven a community is, the faster the effective number of species decreases
as the free parameter increases. Jost proposes quantifying this discrepancy
between the true and the effective numbers of species in multiplicative terms,
introducing an inequality factor

$$S = J(\alpha) \times Q_{\mathrm{ineq}}(\alpha), \quad \alpha \geq 0. \tag{2.12}$$

Since $S = J(0)$, these inequality factors are given by $Q_{\mathrm{ineq}}(\alpha) = J(0)/J(\alpha)$
(and can be transformed to equality factors by taking the reciprocal $Q_{\mathrm{eq}} = J(\alpha)/J(0)$).

Jost discusses different transformations of these evenness components. If
equality factors are used to compare communities, they suffer from the drawback
that their range depends on the underlying species richness (as the effective
number of species $J(\alpha)$ is bound by S). Hence they can give a misleading
picture if communities are compared that differ greatly in numbers of species. To
overcome this problem, Jost applies the linear transformation $(x-x_{\mathrm{min}})/(x_{\mathrm{max}}-x_{\mathrm{min}})$ to derive relative versions of Q_{ineq} and Q_{eq} which measure evenness relative
to the minimum and maximum possible for a given number of species. To
preserve the complementarity of equality and inequality at the same time, he
favours applying this transformation to the logarithm of the equality factors.

This gives relative logarithmic evenness

$$\frac{\ln Q_{\text{eq}} + \ln S}{\ln S} = \frac{\ln J(\alpha)}{\ln S}, \tag{2.13}$$

which for $\alpha = 1$ is the known evenness measure proposed by Pielou (1969).

2.5.3   Generalised entropy inequality

The issue of measuring 'evenness' is not only relevant in ecology. One of the transformations discussed by Jost (2010) in connection with the inequality factors described in the previous paragraph is closely related to the goodness-of-fit based approach: Cowell (1980) introduces a family of measures, known as *generalized entropy inequality*, which is a well established measure in economics to quantify the degree of evenness in the distribution of wealth across households (Jenkins, 2009). Mathematically, it is close to the power-divergence family in (2.4) under reparametrisation $\lambda = \nu - 1$ (Cowell, 1980). More precisely, it is essentially equation (2.4) without the factor p_i outside the square brackets.

An important and well discussed problem in economics, the numerical summary of the distribution of wealth or shares across participants of a market (firms, households) has received much attention (Bruckmann, 1969; Studeny, 1973; Cowell, 1980; Hoffmann, 2008). Indeed, there are parallels between the quantification of evenness as an aspect of biodiversity and the quantification of the distribution of assets or market shares across shareholders in economics. The species forming the community of interest and their species abundance distribution which describes their proportion within the whole assemblage can be seen as analogous to the relative share of total wealth by a number of firms or households. However, there is an important difference. In economics, the total number of shareholders as well as their proportional shares are in general known exactly. In contrast, ecological data is almost always a sample from some underlying *unknown* population. Uncertainty persists with regard to the true number of species (species richness) as well as detectability of individuals from different species during the sampling process. This will be discussed in more detail in chapter 3.

2.6   Analysing ecological communities based on a graphical representation

2.6.1   Evenness profiles and their properties

As functions of the free parameter λ, families such as $I_{\mathbf{p}}(\lambda)$ allow us to plot index values over the range of that parameter. Thus they provide a graphical display of information contained in the species abundance distribution in addition to a quantification of diversity. We show in this section that this is highly useful and more intuitive for conveying multidimensional information than deriving either a single scalar metric or a range of of them. The graphical equivalent of a parametric index family has been termed a 'profile' (Patil & Taillie (1982); see also section 1.3.3 in chapter 1).

Profile plots of existing index families are usually restricted to the positive parameter range. For most, the parameter value of zero corresponds to the number of species (or a transformation thereof). In this case, by simply counting the number of species, we assign equal weight to all species. Plots are usually a monotonically decreasing function of the parameter, and the metrics tend to be increasingly driven by the dominance of the most abundant species. Negative parameter values would put increasingly greater weight on the rarest species, but would at the same time reverse the order when comparing distributions with the same number of species but different levels of evenness. This violates what Ricotta (2003) refers to as 'consistency with the intrinsic order' (as given by the Lorenz curve, see chapter 1.3.1). Because this consistency is necessary for defining at least a partial order between species abundance distributions, it is a desirable property when profile plots are used for comparison in ecological applications. Hence, negative parameter values have so far been excluded in any of the existing parametric index families. Nevertheless, they do provide valid index values and contain information on community structure. We will see that profile plots that are based on our parametric approach, and which we call *'evenness profiles'* in the following, do not suffer from this change in order and we can indeed explore the whole parameter range.

As with diversity ordering, we derive these evenness profiles by plotting the family of divergence measures $I_{\mathbf{p}}(\lambda)$ for a given species abundance distribution $\mathbf{p}$ as a function of λ. This profile can be used to analyse community structure and compare species abundance distributions. The resulting graph is always continuous and U-shaped. In this it differs from existing approaches, which would be sigmoid were negative parameter values included; it is this sigmoid
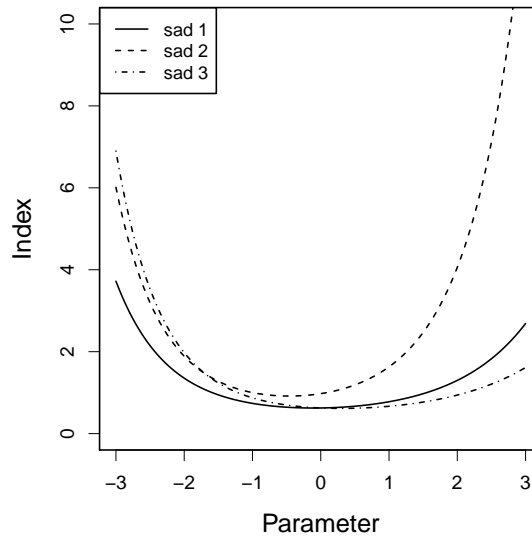
FIGURE 2.1 – Example of evenness profiles for three species abundance distributions – sad 1 = (1,1,1,2,5,7,13,30), sad 2 = (1,1,1,1,2,4,5,7,12,58) and sad 3 = c(1,1,1,7,15,30) (the same have been used as a toy example in the corresponding figure in chapter 1.3.3). A greater degree of curvature corresponds to increased unevenness. The profiles provide more nuanced information than standard diversity profiles: sad 1 and sad 3 are similar for positive parameter values, hence with respect to dominance of species; for negative parameter values, i.e. when rare species are considered, sad 1 is closer to sad 2.

shape which causes the ordering to reverse. Typically, we are interested in diversity metrics as relative rather than absolute measures, to compare species abundance distributions through either space or time. The curvature of our evenness profiles provides further information to aid such comparisons. The closer a species abundance distribution is to perfect evenness, the less the index values depend on λ until in the limiting case of perfect evenness, all members of the family are equal (corresponding to a horizontal line). Hence, the degree of curvature of the profile plot reflects the degree of unevenness in the species abundance distribution.

In traditional diversity profiling, the focus shifts between the two components, evenness and species richness, as the parameter varies. By contrast, for these evenness profiles, the focus switches between the two opposite ends, rarity and dominance, of one component (evenness) (see Fig. 2.1). Although not completely independent of the number of species S (which must be fixed in advance), an evenness profile is equally dependent on S everywhere, allowing comparison of profile plots.

In terms of a sample, the members of the family $I_{\mathbf{n}}(\lambda)$ vary in their sensitivity to more extreme ratios of observed to expected frequencies per class (species) for different values of λ. Large positive λ put more weight on large ratios of

observed frequencies to those expected under perfect evenness, thus detecting unevenness due to pronounced dominance of a species. At the other end, large negative λ highlight observations that lie below their expected levels under perfect evenness, i.e. on rare species. Thus, these profiles are able to distinguish between a community with high dominance of a species and one where the unevenness lies mainly within the rare species, while the common species have similar abundance (for an example see sections 2.7 and 2.8 below). This is not detected by existing approaches. For a graphical assessment of community structure, the range of $-3 \leq \lambda \leq 3$ usually provides sufficient information.

2.6.2   (Ecological) criteria for an evenness index

We have discussed properties of the family of evenness measures based on a goodness-of-fit approach from a purely mathematical and statistical point of view in section 2.3. However, our interest is in applying this approach in the form of evenness profiles to ecological data. Thus we also need to ask about their meaningfulness in an ecological context.

In addition to statistical properties, Buckland *et al.* (2005) stated criteria for the general behaviour of a diversity index. With a headline index for diversity in mind, their criteria refer to changes in all three components, species richness, evenness and overall abundance, which should be adequately reflected by the index: a change in any one component while the other two remain constant should lead to an analogous change in the diversity measure. By concept, the measures in the index family $I_{\mathbf{p}}(\lambda)$ are evenness measures, and, as such, less adequate as a headline index (however they can be applied as a useful supplement to a headline index, see chapter 4). Relevant properties with regard to the criteria by Buckland *et al.* (2005) are as follows. Since it is based on relative abundances the family of indices $I_{\mathbf{p}}(\lambda)$ is not affected by changes in individual abundances or the overall abundance, as long as number of species and the degree of evenness remains the same. (In this, they show the same drawback as the related Shannon's and Simpson's indices.) For the untransformed index, a decrease in evenness leads to an increase in the measure (for any parameter value λ). The interdependence with species richness has already been mentioned briefly and its problematic nature will be discussed separately in more detail below.

An earlier study by Smith & Wilson (1996) concentrated explicitly on evenness indices. In a detailed analysis of existing measures and their own suggested evenness index, they established a comprehensive list of 'requirements' and 'fea-

Requirements	$I_{\mathbf{p}}$	$1/I_{\mathbf{p}}$	$\rho_{\mathbf{p}}$
Invariant under replication of the SAD *(independence of species richness)*	✓	✓	⊘
Decreasing when abundance of rare species is reduced	—	✓	✓
Decreasing after addition of a rare species	—	✓	✓
Invariant when SAD multiplied by a constant *(unaffected by units used)*	✓	✓	✓

Features	$I_{\mathbf{p}}$	$1/I_{\mathbf{p}}$	$\rho_{\mathbf{p}}$
Maximal at perfect evenness	—	✓	✓
Maximum value equals 1	—	—	✓
Minimal for any number of species when SAD most unequal	—	✓	✓
Value close to min (< 0.5) when community as maximally uneven	—	—	⊘
Minimum value equals 0	—	✓	✓
Minimum value attainable with any no of species	—	—	✓
Value in the middle of scale for intermediate levels of evenness	—	—	—
Reasonable response to changes in an SAD *that systematically increase evenness*	✓	—	✓
Symmetry with regard to minor and abundant species	—	—	—
Lower value for skewed distributions	—	—	—

TABLE 2.1 – Behaviour of the evenness index $I_{\mathbf{p}}(\lambda)$ and its transformations $1/I_{\mathbf{p}}(\lambda)$ and $\rho_{\mathbf{p}}(\lambda)$ with respect to the criteria postulated by Smith & Wilson (1996). '⊘' indicates that the respective feature is only just not met, whereas '—' means that the index fails to fulfil this criterion completely. Except for the first requirement, all other criteria were evaluated for $\lambda = -0.5$.

tures'. Requirements are properties that these authors consider as essential for any evenness measure in ecology, while features are desirable, but not necessary. Using the same toy examples as in the original article, we checked our evenness profiles for these criteria. Table 2.1 provides a summary of the results (R code and more detailed results are given in Appendix B). The untransformed goodness-of-fit measures $I_{\mathbf{p}}(\lambda)$ fail the majority of these criteria, mostly because of their lack of standardisation (see 2.4 above). Taking the reciprocal of $I_{\mathbf{p}}(\lambda)$ assures that all requirements are met. However, only three of the features are fulfilled. Due to its standardisation the transformed family of measures $\rho_{\mathbf{p}}(\lambda)$ does much better, but fails what Smith & Wilson (1996) refer to as independence from species richness. In the context of their study, this means that multiples of the same SAD should result in the same value. Hence the index

should not distinguish between

$$v_1^1 = (1497, 1, 1, 1)$$
$$v_1^2 = (1497, 1, 1, 1, 1497, 1, 1, 1)$$
$$v_1^3 = (1497, 1, 1, 1, 1497, 1, 1, 1, 1497, 1, 1, 1)$$
etc.

(example taken from Smith & Wilson (1996)). $\rho_{\mathbf{p}}(\lambda)$ shows a stronger 'dependence' on S in this sense if the SAD is very uneven. However, for a large number of species (roughly $S > 50$) this becomes negligible. Although other authors also point out the importance of the invariance under this kind of replication (Hill, 1973; Routledge, 1983; Jost, 2010), it is debatable whether independence in this sense is a reasonable requirement. Evenness reflects the distribution of the total number of individuals over the number of species. Replicates of a species abundance vector do not only increase the number of species but change the overall abundance as well. In general, this changes the SAD expected under perfect evenness (and hence the null model). The choice of an evenness measure then also involves the decision whether or not we want this change to be reflected in the quantification. (This is similar to the differences between Jost's concepts of absolute and relative evenness.)

Given the multidimensionality of biodiversity in general and the interdependence of species richness and evenness in particular, condensing it into an index (or even a family of indices) cannot avoid a loss of information. Thus the choice of the index needs to address the specific goal of the biodiversity assessment. As our focus is on regional trends in biodiversity, i.e. a fairly large spatial scale, where numbers of species are generally big, we can assume that the goodness-of-fit indices are largely unaffected by variation in species richness.

2.6.3 Interdependence of evenness and species richness

Traditionally, evenness metrics were sought that are largely independent of species richness, since evenness and richness were viewed as two separate characteristics of the species abundance distribution. More recently, it has been recognized that their relationship is more complex and that they cannot be regarded as independent components of diversity (Gosselin, 2006; Jost, 2010); adding or reducing the number of species, without making any adjustments to the abundances of the remaining ones, will in general change the degree of evenness. For example, if species i disappears from the species pool, i.e.

$n_i = 0$, traditional scalar and existing parametric diversity and evenness metrics will register the reduction both in numbers of species and in evenness at the same time. This makes it difficult to distinguish between the two effects. Our approach gives the option of either retaining the same species richness S or resetting it to $S - 1$. In the first case, evenness is then evaluated with respect to S species (species i making a negative contribution) or in the latter case, species i is not taken into account. This allows us to choose which we consider to be the most appropriate, at the cost of having to specify the assumed species pool rather than let it be chosen by the measure.

A potential problem in this context lies in the fact that it is generally difficult to estimate the number of species S without bias or high uncertainty (Lande, 1996; Gotelli & Colwell, 2001, 2011) . This issue is further discussed below when we look at the problem of unobserved species in samples. However, any approach that is - like ours - based on $J(0) = S$ should either justify why this is sensible given the difficulty of determining the number of species, or look for ways of avoiding this problem. Hill's solution was to consider quotients with $J(2)$ in the nominator instead of $J(0)$. This could be done for the method suggested here, too, but would result in breaking the connection to goodness-of-fit statistics which provides a natural interpretation of the measure and a direct sample based estimate. Furthermore, quotients $J(a)/J(2)$ are difficult to interpret in most cases and show ambiguous behaviour (Peet, 1974; Jost, 2010). They do not increase monotonically with evenness anymore and hence can show similar values for highly even and for uneven species abundance distributions (see scenario 1 in section 2.7). In applied studies, the problem is usually avoided by assessing evenness of a sample (i.e. taking into account only observed species) rather than drawing inference on the species abundance distribution. In this case, the measures do not reflect a characteristic of the population but of the sample itself (Smith & Wilson, 1996; Hill, 1997). Parametric indices on the other hand have been defined and studied theoretically for given (and fully known) species abundance distributions. Both approaches are on their own unsatisfactory, the first because we are interested in characteristics of the population, not the sample, and the latter because we have to rely on samples to draw inferences on these characteristics.

Instead, we assume here that the number of species S in a study area is fixed (and known) at least over the period of time we are interested in. We think it is plausible to base an evenness measure on S for two reasons. First, theoretical studies (e.g. Bulla (1994)) are often implicitly built on the assumption of a fixed number of species. These studies provide valuable insight in the general

behaviour and properties of diversity indices. By stating this assumption on S explicitly, we set a defined framework in which our analyses and conclusions hold. Second, in the context of monitoring temporal change in biodiversity within a site or region, it is self-consistent to define a list of potential species, and to estimate species proportions as zero for those species not recorded at a given time. Disappearance of species will then be reflected by a change in evenness. If unanticipated species are recorded in later surveys, the list of potential species may be revised, and the entire sequence of biodiversity measures updated.

The comparison of two completely independent communities on the other hand is more complex. Because of the multidimensionality of diversity as a concept, the question of which of two or more such communities is the more diverse will rarely have a single answer. Yet, if we can assume either that their species richness is comparable or that more or less the same catalogue of species holds across communities (this can always be achieved artificially by pooling species), our family of evenness measures offers an instrument of comparison. As it is sensitive to unobserved or missing species, it is to some extent able to pick up differences between these communities. However, it remains a diversity measure and does not account for species identities, i.e. turnover between communities (as has been pointed out in more detail in section 2.3.4 above).

2.7   Applications I: Tokeshi's niche models (simulation study)

After discussing its various properties as a measure of biodiversity as well as its connection to other approaches, we next want to evaluate the use of the goodness-of-fit based index family to analyse ecological assemblages in practice. This will first be done by looking at predictions from well-known models of community composition and followed by an example of actual data that gives an outlook on later chapters.

Box 2.2: Dominance decay, power-fraction and dominance preemption — Tokeshi's models explained

Niche models in ecology try to explain the coexistence of species despite the fact that resources and suitable habitat are limited. Tokeshi (1990, 1996) introduced several stochastic models to describe how species sequentially join a community and by doing so, they take over a certain share of the available niche space. By assumption, this share is proportional to a species' abundance. Hence, the total niche space is always shared between all species present; an arriving species changes the distribution of niche space and therefore the species abundance distribution. In all Tokeshi models invaded niches are divided at random. However, the models differ in the way the next niche to be occupied is chosen. If we represent the available niche space by a line of length 1 and assume three species present, where $\mathbf{p} = (1/2, 1/3, 1/6)$. Then the following gives a schematic description of the three model classes (dominance decay, power-fraction and dominance preemption):
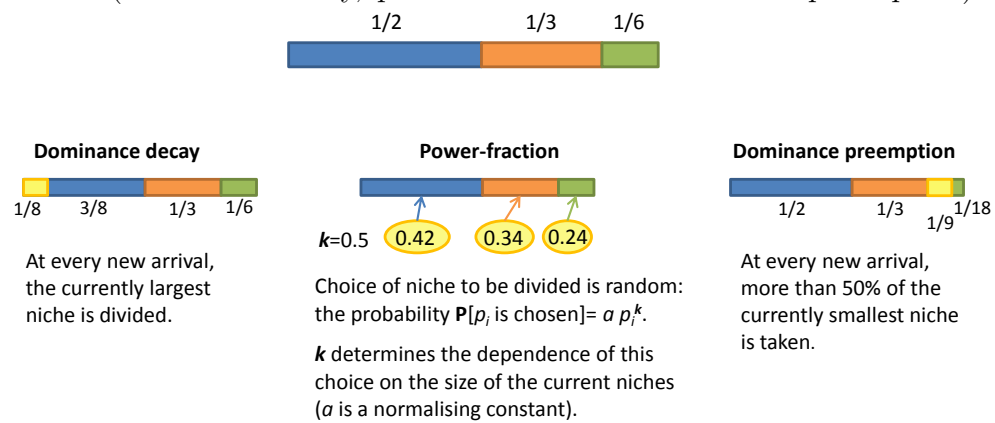
The different stochastic mechanisms lead to different degrees of evenness; the table below gives a summary.

Model	Specifications	Degree of evenness
Dominance decay	Newly colonising species take random part of largest current niche	high
Power-fraction	Niche is chosen randomly, depending on its size (degree of dependence regulated via a model parameter k); random break point; $k = 0$ (random fraction) completely random niche choice, $k = 1$ (MacArthur fraction) largest niche has the highest possible probability of being chosen	variable between $k = 0$ (low evenness) and $k = 1$ (high evenness)
Dominance preemption	Arriving species take more than 50% of the remaining niche space	low

2.7.1 Tokeshi's niche models

Niche partitioning tries to describe how available space and resources in an ecosystem are partitioned between the species in an assemblage (Marquet *et al.*, 2003). Tokeshi (1990, 1996) introduced a group of stochastic models (dominance preemption, power fraction, dominance decay) to predict species abundance distributions based on niche apportionment (for details see Box 2.2). These models have been shown to provide a good fit to a range of real data (e.g. Fesl (2002)). Meanwhile, the discussion of niche models in ecology has evolved, and is concerned with the complex processes governing community structure (Kelly *et al.*, 2006). Nevertheless, Tokeshi's models allow to simulate realistic species abundance distribution with varying degrees of evenness. As such, they provide means of testing the behaviour of the evenness profiles and compare them to the other approaches discussed in section 2.5.

Tokeshi's models can be distinguished as three different model classes (see also Box 2.2). In the dominance preemption model niche space is partitioned by species consecutively taking more than half of the remaining niche space, thus being dominant over all the following species. This leads to a highly uneven species abundance distribution. The power-fraction model on the other hand allows systematic variation of the degree of evenness of the species abundance distribution derived from it by changing the values of the model parameter k within the range from 0 to 1. An increase in k corresponds to a more even allocation of niche space. While for $k = 0$ niche space is divided completely randomly (random fraction), the model for $k = 1$ (MacArthur fraction) chooses niche space that is to be partitioned with a probability proportional to the size of the niche. Since larger niches are divided with higher probability, this leads to a more even species abundance distribution. In the model with the highest degree of evenness – the dominance decay model – the largest current niche is always chosen to be partitioned further.

Originally introduced as a way to describe niche apportionment and the dynamics that structures ecological communities, Tokeshi's models can also be taken as a static description of the expected species abundance distribution as the outcome of niche partitioning with a fixed number of species (Magurran (2004), p.47). By looking at species abundance distributions simulated from these models, we can illustrate the properties and the performance of the goodness-of-fit index family when assessing evenness. In addition to the evenness profiles, we consider the quotients of Hill's numbers as a measure of evenness and Jost's logarithmic relative evenness for comparison.
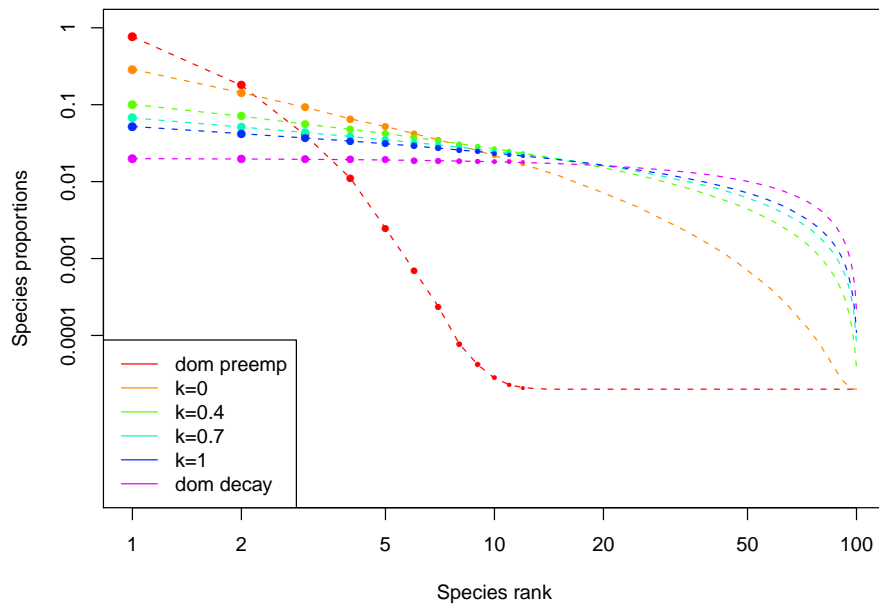
FIGURE 2.2 – Rank-abundance-plots for six species abundance distributions with varying degrees of evenness, generated through simulations of Tokehi's models of niche apportionment. The dots mark the 12 most abundant species.

2.7.2   Specifications of the simulations

The following simulations of Tokeshi's models were implemented in the statistical language R (R core development team (2011); code details can be found in Appendix C). For each model, 500 random realisations were generated by a discrete algorithm where available niche space was fixed by setting the total abundance $N = 50000$. The expected species abundance distributions were then derived by averaging over these and are shown as rank abundance plots in Fig. 2.2. We consider the following scenarios:

1. In scenario 1, we construct evenness profiles based on the true species abundance distributions for all the Tokeshi models and compare them to other parametric evenness measures. The number of species $S = 100$ is fixed and for the power fraction model values $k \in \{0, 0.4, 0.7, 1\}$ are chosen for the model parameter.

2. Scenario 2 explores the effects of changes in species richness on the different profile plots. We do so by looking at the two extremes in terms of evenness, the dominance decay and the dominance preemption model. In the setting of the Tokeshi models, changes in species richness can be envisaged as invading species while total abundance represents total available niche space (and is kept constant here).
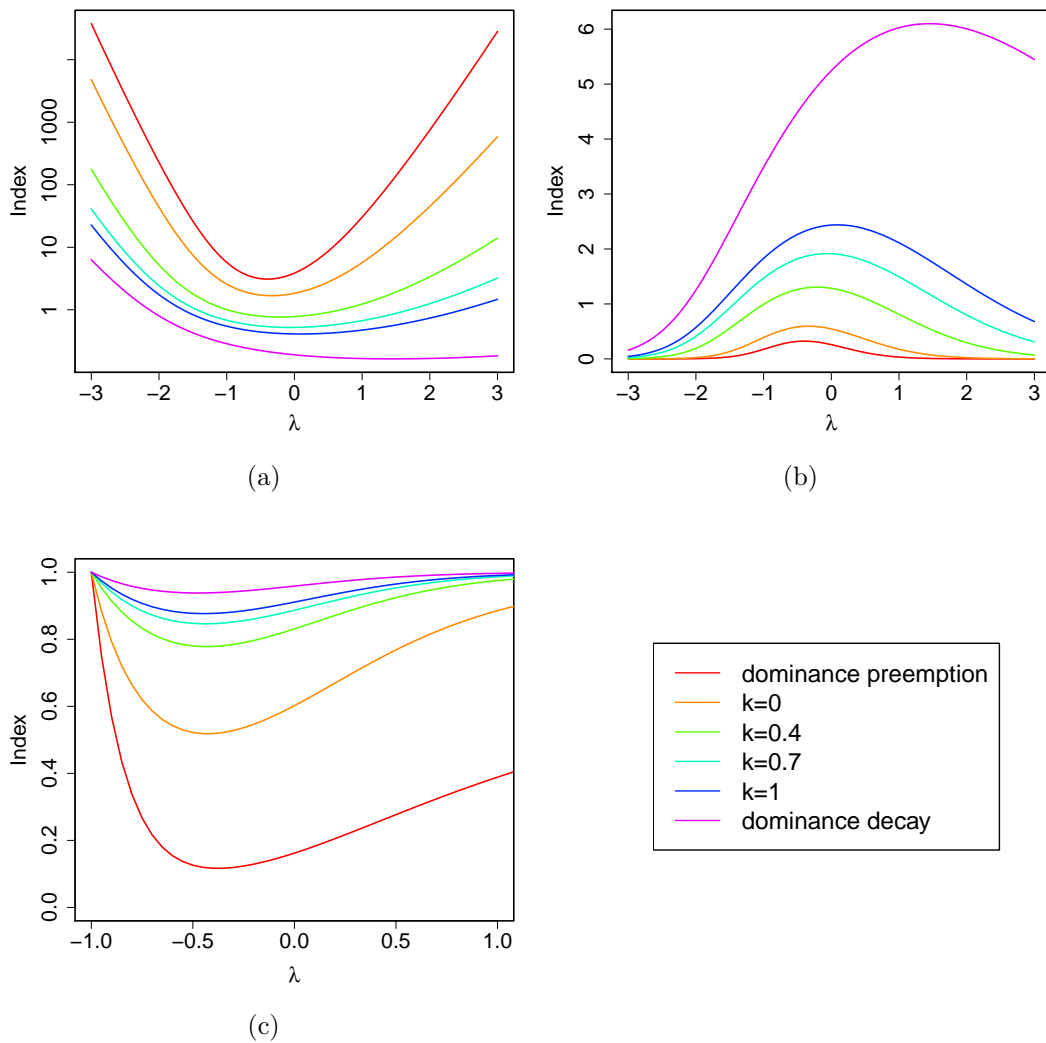
FIGURE 2.3 – Goodness-of-fit based evenness profiles and transformations for a range of Tokeshi's models. (a) shows untransformed evenness profiles $I_{\mathbf{p}}(\lambda)$ as a function of λ, (b) profiles based on the reciprocal $1/I_{\mathbf{p}}(\lambda)$ and (c) a transformation that not only conforms the ordering of the profiles from low to high values of evenness, but also standardises their range to the interval $[0, 1]$.

Both scenarios evaluate evenness profiles based on the community structure predicted by the Tokeshi models for known N, S and k to study their general behaviour. In chapter 3 we will discuss problems that arise specifically because diversity assessment has to be based on information contained in samples. In this context we return to the Tokeshi models to study the effects of sampling, in particular how we can deal with unobserved species (scenario 3).

2.7.3   Scenario 1: Ordering of the Tokeshi models

Evenness profiles put Tokeshi's models in the correct order (see Fig. 2.3). Curvature decreases corresponding to higher degrees of evenness. For the dominance decay model, only the negative parameter range enables us to detect departure from evenness, as evenness is reduced by the rare species while common species are highly homogeneous in numbers of individuals. Overall, the similar curvature in the negative parameter range reveals that it is the degree of dominance of the abundant species that distinguishes between Tokeshi's models, rather than differences in occurrence of rare species. (This confirms them as realistic ecological models, as they follow the universal hollow curve (Magurran, 2004; McGill *et al.*, 2007).)

By default, high values in the evenness profiles are associated with departure from evenness. This might be seen as counter-intuitive in ecological applications, where traditional diversity indices and index families usually give high values at high evenness. Possible transformations were discussed in section 2.4 and are shown in Fig. 2.3 as an alternative. However, neither of them preserves the distances between the curves and the possibility of directly interpreting the curvature of the original evenness profile is lost. This has to be kept in mind when drawing conclusions based on the transformed indices. In particular, the transformation $1/I_{\mathbf{p}}(\lambda)$ no longer displays evenness as a nearly horizontal line; the lower the index values in the original index, reflecting high evenness, the higher they are under this transformation, with no upper limit, while distances between profiles representing uneven distributions decrease as the index values approach zero. Especially, if we are looking at changes in a community over time, the same change can be reflected in quite different amounts of increase or decrease depending on whether we look at positive or negative parameter values. For example, consider a community that changes from following Tokeshi's dominance decay model to Tokeshi's power fraction model with $k = 1$ over time. If we assess this change in evenness by looking at, say, $\lambda = -2$ and $\lambda = 2$, we will see a much larger decrease in $1/I_{\mathbf{p}}(\lambda)$ for the latter. While we should be careful about comparing the amount of decrease for the two parameter values in this case, we can certainly state that there is an overall decrease in evenness. Likewise, if there were no visible change for a positive parameter value while there is a decrease for the negative parameter range, we would be able to conclude a negative trend for rare species while evenness remains unchanged for the dominant species. This has to be kept in mind when these profiles are applied to real data in chapter 4.
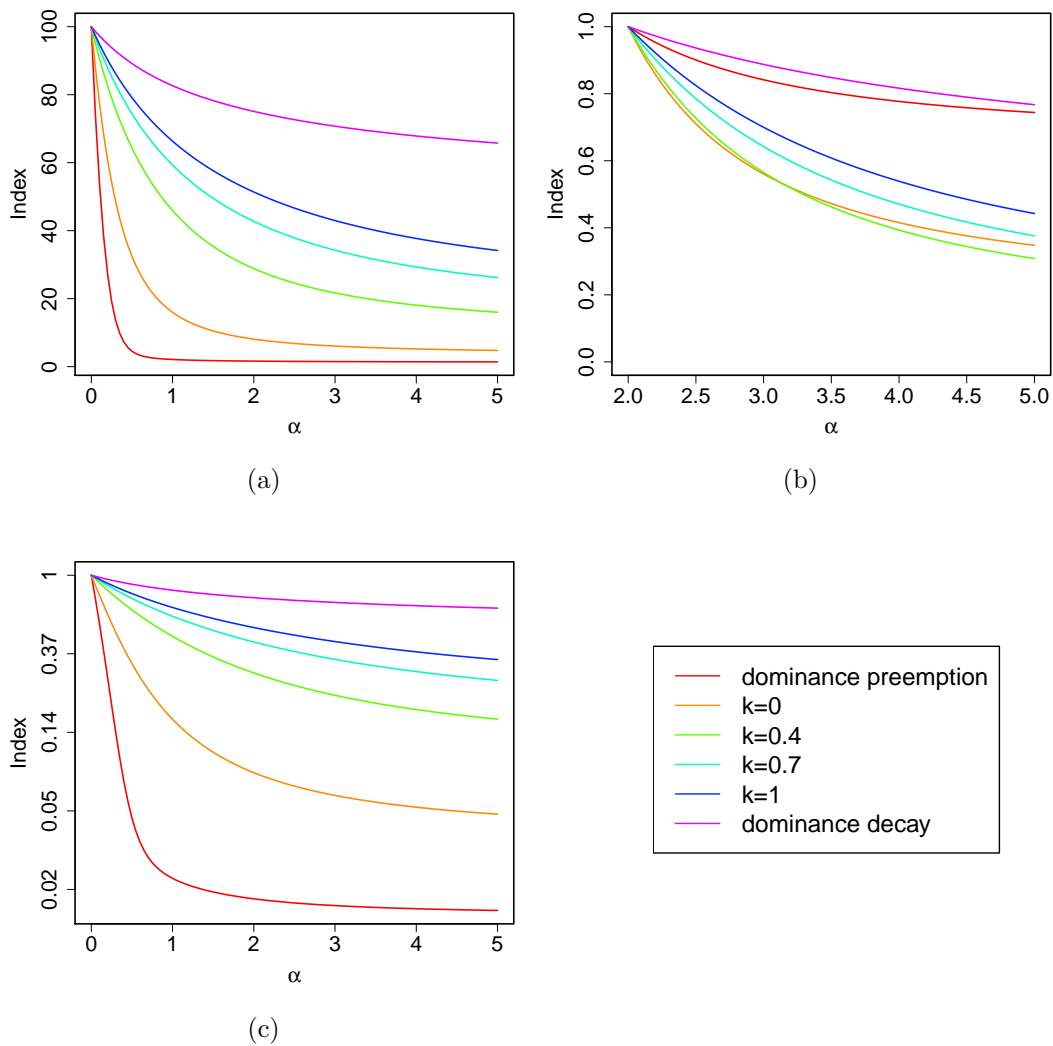
FIGURE 2.4 – Profile plots for Tokeshi's models based on quotients of Hill's diversity numbers, showing (a) Hill's untransformed diversity numbers, (b) quotients $(J(\alpha)/J(2))^2$ and (c) Jost's evenness quotients $(J(\alpha)/J(0))$ (on a logarithmic scale). Hill's evenness quotients fail to distinguish correctly between Tokeshi's models. All approaches are restricted to positive parameter values $0 \le \alpha \le \infty$.

The third transformation suggested in section 2.4 (equation (2.8)), leads to a profile plot that is no longer continuous, but that resembles a hyperbola, with an asymptote between 0 and 1 (see plot in Appendix). The position of the asymptote gives a rough idea of the evenness of the underlying community; the closer it is to 0 the higher the evenness. However, the latter is again a scalar (one-dimensional) description.

For comparison, profiles based on Hill's diversity numbers are shown in Fig. 2.4, with the usual restriction to positive parameters. All transformations suggested to account for evenness are scaled to the range of $[0, 1]$ with 1 corresponding to perfect evenness. Because only positive parameter values are usu-

ally considered, a quotient close to 1 can either be the result of the underlying species abundance distribution being truly close to evenness or simply due to homogeneity in abundance within the more dominant species, while the tail of rare species is neglected. Plots based on $J(a)/J(2)$ (suggested by Hill) show the aforementioned lack of consistency; for higher parameter values, the highly even dominance decay model is placed closest to the most uneven dominance preemption model.

On the other hand, logarithmic versions of $J(a)/J(0)$, as suggested by Jost (2010), do well in differentiating between the different Tokeshi models. With respect to the dominance structure, they are as effective as the evenness profiles. However, since they are restricted to the positive parameter range, they do not reveal that the models are similar with respect to the tail of rare species. In particular, high evenness of the dominance decay model is picked up, but the remaining unevenness due to rare species is undetected.

2.7.4 Scenario 2: Effect of changing species richness

While some argue that quantification of evenness should not be affected by changes in species richness (Smith & Wilson, 1996), it is quite obvious in the conceptual setting of Tokeshi's niche models that a change in the number of species (envisaged as additional species invading the available niche space) is in general expected to change community structure and hence potentially evenness. Additional species reduce the niche sizes expected under a completely even apportionment. In the dominance preemption model this means that the deviation of the size of the larger niches (abundant species) from complete evenness becomes bigger as species invade niche space. On the other hand, niches taken by the incoming species get smaller and smaller, but so do the differences in niche sizes for (those) rare species. Hence we expect to see a more even distribution of niches for the least abundant species. Quite the opposite is to be expected for the dominance decay model. As additional species always invade the largest niche, no single species becomes dominant at any time; we expect to see a fairly flat profile and dominance structure should largely remain unchanged. In particular, from the profile plot it is clear that at $S = 50$ we can expect a nearly even distribution of niche space and no dominant species. However, this means that invasion adds rare species to the assemblage, because further species either take a smaller portion of a niche or leave one existing species with only a small part of their original niche. This generates at least one rare species and, because it is almost certain that the random split of the

QUANTIFYING BIODIVERSITY TRENDS IN TIME AND SPACE

Angelika C. Studeny

**A Thesis Submitted for the Degree of PhD
at the
University of St Andrews**
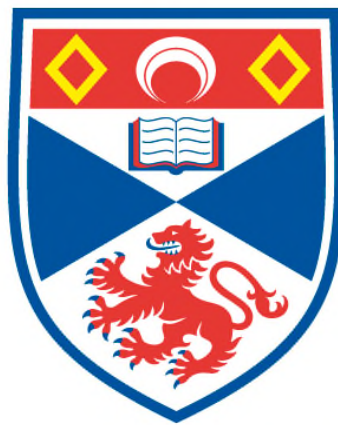
2012
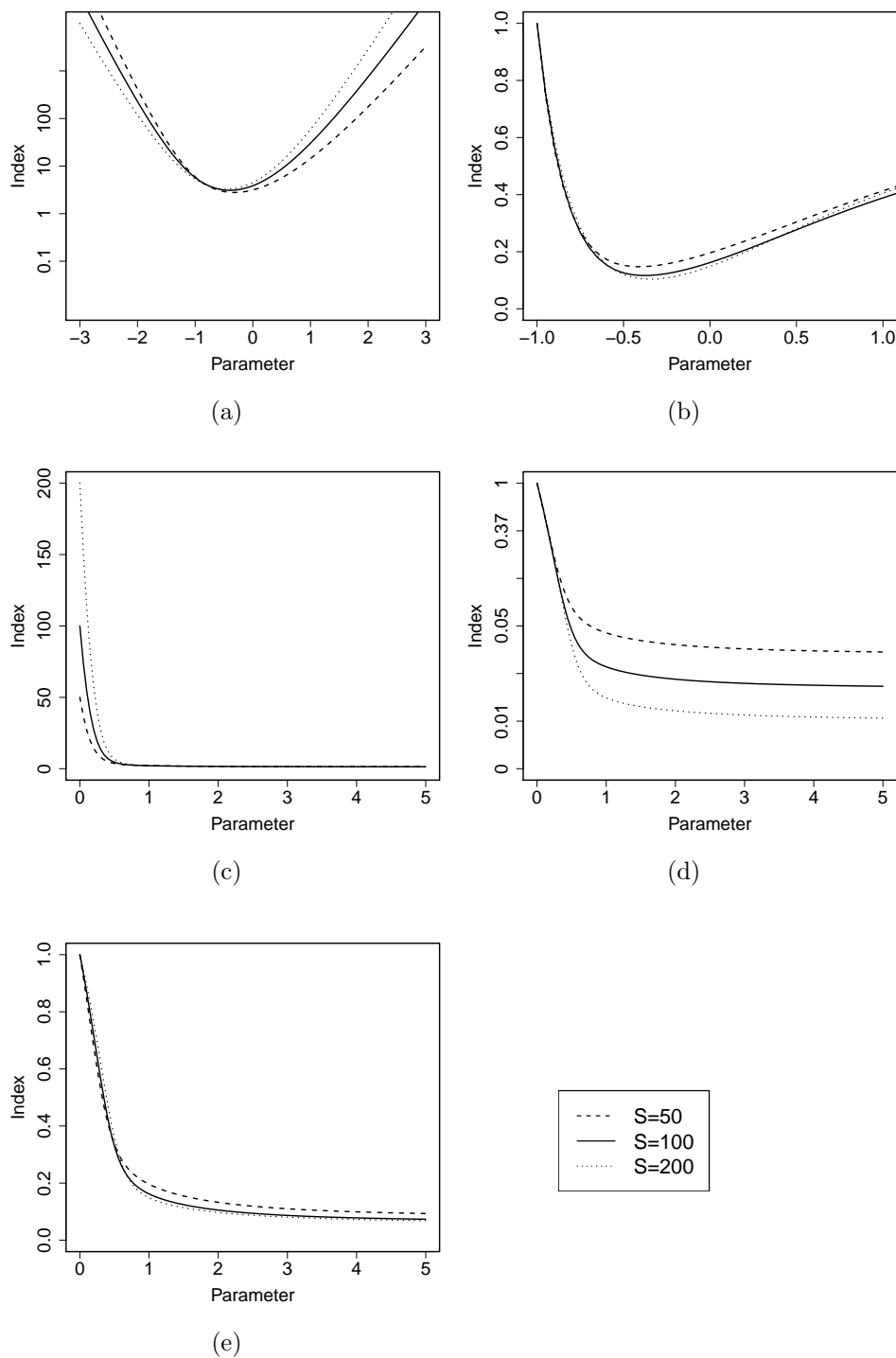
FIGURE 2.5 – Profile plots for varying number of species in the *dominance preemption model*. Additional species invading the niche space lead to changes in community structure in Tokeshi's models (see main text for details). Different profile plots register these changes differently — (a) untransformed evenness profiles and (b) their standardised version, (c) Hill's untransformed numbers and two of Jost's (2010) transformations, (d) equality factors (on a logarithmic scale) and (e) logarithmic relative evenness.
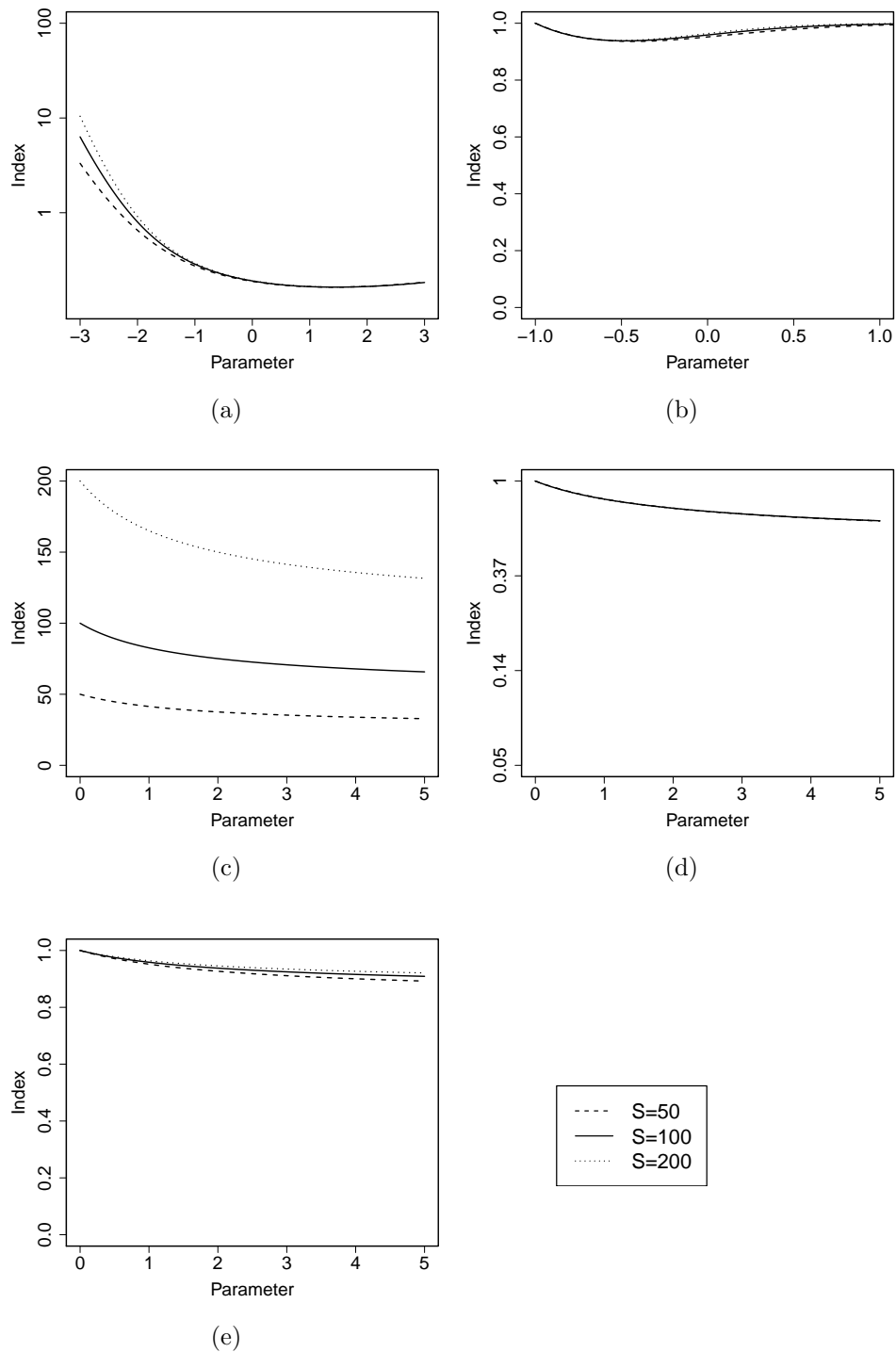
FIGURE 2.6 – Profile plots for varying number of species in the *dominance decay model*. Additional species invading the niche space lead to changes in community structure in Tokeshi's models (see main text for details). Different profile plots register these changes differently — (a) untransformed evenness profiles and (b) their standardised version, (c) Hill's untransformed numbers and two of Jost's (2010) transformations, (d) equality factors (on a logarithmic scale) and (e) logarithmic relative evenness.

niche is uneven, it increases unevenness amongst the rare species. This is visible in the untransformed evenness profiles. However, the standardised version does not display these effects well. Hill's diversity numbers $J(\alpha)$ separate the two models when species richness is gradually increased. Nevertheless, it is not evident whether the difference that is visible in the plots is genuinely caused by a difference in evenness or whether it is due only to the varying number of species (i.e. the difference in the intercept for each curve). This holds in particular for the dominance preemption model, where differences are only visible very close to the origin. When plotted on a logarithmic axis, Jost's equality factors separate the dominance preemption models well, similarly to the evenness profiles for the positive parameter range. As they only take into account the dominance structure, no effect is visible for the equality factors for the dominance decay model. Relative logarithmic evenness measures proposed by Jost (2010) for comparing communities independent of their differences in numbers of species do not pick up a change in evenness with changes in species richness (as we would expect of them).

2.8   Applications II: Assessing trends in evenness over time in British farmland birds (Common Bird Census data)

As discussed in the previous section, focusing solely on the more dominant species and neglecting unevenness caused by the rare species may give a false impression of homogeneity. The following example shows that this is particularly true for assessing trends in diversity over time if the proportions of dominant species remain fairly stable, while there is a change in some of the rare species. The data come from the UK Common Bird Census (CBC) (Marchant *et al.*, 1990). This long-term survey of breeding birds organized by the British Trust for Ornithology was run over 38 years (1962 to 2000). Over this time period, changes in agricultural practice had major effects on farmland birds in the UK (Siriwardena *et al.*, 1998). The survey protocol follows a territory mapping approach. Fewster *et al.* (2000) used generalized additive modelling to smooth the time series of territory counts for a set of 13 farmland species during years 1962 to 1995, which yields predicted counts for every species for each site in each year, whether or not a site was surveyed in a given year.

To look at time trends in evenness of these farmland birds, we plot a profile over the parameter range $[-5, 5]$ for every year of the survey based on these smoothed estimates of individual counts. The results are compared with diversity trends estimated by taking the geometric mean of the 13 species-specific relative abundance estimates obtained by Buckland *et al.* (2005).

We see that the intermediate parameter range, which covers the transformations of Shannon's and Simpson's indices, does not display any substantial change in evenness over time, aside from an increase (reduced biodiversity) from the late 1970s to the late 1980s (see Fig. 2.7). This is in accordance with Buckland *et al.* (2005), who remarked that these classical measures register a decline in diversity after 1975, but show no trend in the first half of the survey. Our 3D surface plot not only confirms this, but makes it apparent that this is true for all measures in the power divergence family which concentrate on dominance of a species. The extended parameter range, however, reveals further changes in evenness. For more clarity, profiles for selected parameter values are plotted in addition to the surface plot in Fig. 2.7. For non-negative $k = 0, 1, 2, 3, 4$, they look fairly similar over time, displaying the same qualitative information as Shannon's or Simpson's index. Profiles based on negative parameter values on the other hand show a decrease in the corresponding indices (i.e. increased evenness) within the group of rare species during the early years, followed by an increase (i.e. decreased biodiversity) from the late 1970s, which continued to the end of the survey period. Buckland *et al.* (2005) also detected the increasing trend in biodiversity in the early years, using the geometric mean of relative abundances and attributed this to a recovery after the severe winter of 1962/63. While a geometric mean considered the community as a whole, our analysis reveals that this affected primarily the rarer species.

Overall, this leads to the conclusion that biodiversity in British farmland birds increased during the 1960s and early 1970s, followed by a decline from 1975 to the late 1980s. Further decline occurred amongst the rare species, but not amongst the abundant species, until the end (1995) of the survey period. Traditional indices do not reflect the changes amongst the rarer species. A more detailed statistical analysis would also include precision of the evenness profiles and establish points in time which show a significant change in trend. This can be achieved by considering bootstrap resamples of survey sites as in Buckland *et al.* (2005).
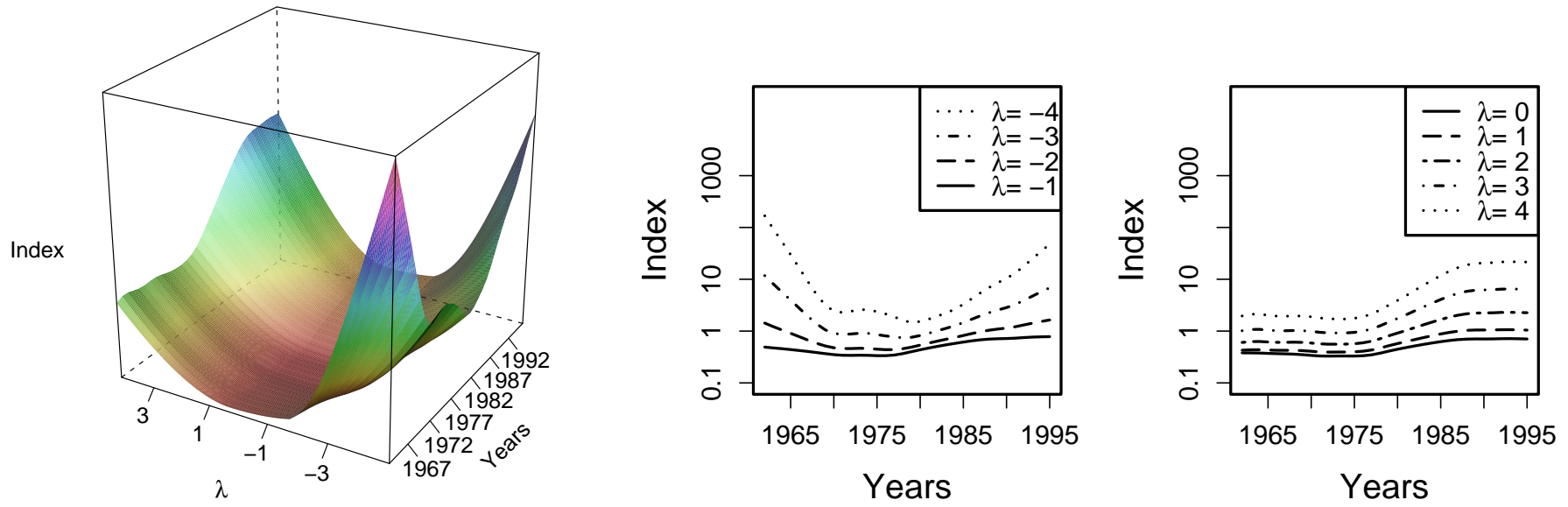
FIGURE 2.7 – The 3D profile surface (left) for the CBC data 1962 - 1995 gives a summary of changes in the farmland bird community along a continuous gradient from rare (negative λ) to dominant species (positive λ). Temporal trends in evenness are extracted as cross-sections for certain values of λ (middle and right).

2.9   Conclusions: performance of evenness profiles

Following the idea of diversity profiling of plotting parametric index families as functions of their parameter, we constructed evenness profiles. The parameter range of traditional diversity profiles is restricted to the positive values, and hence concentrates mostly on the abundant species. In contrast, evenness profiles can sensibly be evaluated for both positive and negative parameter values. These profile plots do not only give index values, but display information on evenness via their curvature. A horizontal line would correspond to perfect evenness. This feature is especially valuable for comparison.

Simulated species abundance distributions from Tokeshi's (1990, 1996) models allowed us to examine the performance of evenness profiles based on our approach using simulated data. The main advantage of this method was illustrated by the example of time trends in British farmland bird diversity, to reveal that changes beyond a general decline during the late 1970s and mid-1980s are due mainly to fluctuations in the abundance of rare species. With the full parameter range at hand, these evenness profiles allow exploration in both directions, focusing on dominant species for positive and on rare species for negative parameter values. Thus they detect that Tokeshi's dominance decay model, despite its highly even allocation of niche space, still contains some unevenness — as every natural community would do. In the example of British farmland birds, classical indices, which are represented by the positive parameter range of the evenness profiles, exhibit changes only in the second half of the survey period following changes in agricultural practices. The effects of recovery from a harsh winter in 1962/63 are, however, visible only along the negative parameter range of the profile plots.

2.10   Discussion

The Convention on Biological Diversity's 2010 and 2020 targets are a response to large-scale loss of biodiversity (Butchart *et al.*, 2010). They reinforced the need for reliable methods to assess change in ecological communities. Loss in biodiversity often affects rare species first, while the dominance structure of a community might change only gradually (Gotelli *et al.*, 2010). However, most methods are limited in their ability to detect change amongst rare species (Colwell & Coddington, 1994).

We have proposed a parametric approach to quantify evenness which includes explicit information on rare species. In particular, it allowed us to separate changes in rare species from those in common species for British farmland birds. The method suggested is based on a correspondence between divergence measures and a family of goodness-of-fit statistics (Read & Cressie, 1988). In general, a high degree of evenness or uniformity of the species abundance distribution is equated with high biodiversity. The idea of this method is that, while ecological communities will never be perfectly even, evenness can serve as a 'null model' (Gotelli & Graves, 1996) and we can measure the departure of the species abundance distribution from perfect evenness to gain insight into the structure of a community. Goodness-of-fit statistics provide us with genuine insight into the properties of the community while quantifying this departure on a sample level. By combining these statistics in a single family, we capture more information on the species abundance distribution (as a multivariate object) than can a single scalar index. We have shown that these measures are closely connected to Hill's numbers, which have been the subject of recent interest as diversity metrics (Jost, 2006; Jost *et al.*, 2011; Jost, 2010; Leinster & Cobbold, 2012).

Ultimately, there is no unique answer to the question of how diversity is best assessed. The method developed here allows us to concentrate on the evenness aspect of the species abundance distribution. The evenness profiles extend the range of visible information and display contents of the abundance distribution differently from existing index families. Thus they offer a new perspective which can lead to further insights. Their explicit focus on rare species may be of special interest when examining the impact of anthropogenic disturbance on diversity (Dornelas *et al.*, 2009; Dornelas, 2010). As with any other method, the results may be biased if evenness profiles are drawn without regard to the sampling scheme (Colwell & Coddington, 1994; Gotelli & Colwell, 2001, 2011). Being sensitive to the number of species that are assumed to be observable, evenness profiles depend on additional information and a deliberate decision on the species that are included in the analysis. While for some taxa, like birds, a complete list of species and therefore exact knowledge of the population of interest is more easily achieved, there are many taxonomic groups for which we will never be able to compile a (nearly) complete list, such as tropical insects (Longino *et al.*, 2002). Evenness profiles allow us to draw inference even if we are not certain about the actual number of species. Provided we have a rough estimate of the number of species, we can include rare, unobserved species by extending the sample vector. This should reduce bias in inference. Qualitatively,

information on evenness is contained in the curvature of the profiles. Curvature seems to be less sensitive to the assumed number of species than the evenness profiles themselves are. Hence, even if we disregard the additional information contained in the plots because of potential bias, curvature should still give us an accurate qualitative result.

2.11   Chapter summmary

In this chapter we have introduced the idea of quantifying evenness based on a goodness-of-fit approach. An existing one-parameter family of goodness-of-fit statistics can be used to derive an entire set of evenness measures. A closely related 'power-divergence' index family that has been defined in information theory and is used in a similar form by economists to quantify inequality in distributions. We have discussed mathematical and statistical properties and explained how this approach provides a graphical representation of an ecological community. The latter is termed an 'evenness profile', analogous to general diversity profiling techniques (see chapter 1). We have explored connections to other index families used in ecology and economy and discussed properties which are particularly relevant in an ecological setting. Finally, we have studied the behaviour of these profiles in practice, by applying them to simulated data as well as to actual data from British farmland birds.

Chapter 3

Detectability: reducing bias in diversity assessment

Any form of diversity assessment is faced with the problem that individuals of some species will be underrepresented or even missing from samples (May (1988); Magurran (2004); Gotelli & Colwell (2011); see also chapters 1.2.3 and 2.6.3). In the previous chapter, we have already briefly discussed issues in connection with the number of species S, which is in general unknown. This chapter sets out to explore imperfect detection of both species and individuals and the consequences for diversity assessment in more detail. In particular, our discussion brings up subtle differences if evenness is quantified by a member of the family of goodness-of-fit statistics and if a divergence measure is used.

In chapter 1 we have distinguished between the problem of missing species (species detectability) and that of variation in detection probability of individuals across different species (individual detectability). The first part of this chapter is concerned with the uncertainty about the true number of species when the diversity of an ecological community is assessed, and options to deal with this problem. By looking at another simulation from Tokeshi's models we investigate these options in practice as well as the effects of misspecification on the inference from profile plots. Given that the focus of this thesis is on large-scale assessment of biodiversity trends, we discuss the issue of species being missed when data is collected from this perspective.

A further issue arises because detection probabilities for individuals are likely to differ across species , even for observed species (Yoccoz *et al.*, 2001; MacKenzie & Kendall, 2002). The second part of this chapter highlights how diversity estimation can be biased when these differences in individual detectability are not taken into account (Buckland *et al.*, 2011a). In particular, we can expect

evenness measures to suffer from this bias. Hence, we explicitly investigate consequences for the goodness-of-fit based measures if these are adjusted for detectability. Again, a distinction is made between divergence measures and measures based on goodness-of-fit statistics.

Sections 3.2.3 and 3.2.4 are published in Studeny *et al.* (2011); in particular, the simulation study using Tokeshi's models can be found there, including the results for the empirical ϵ-correction. The discussion on model-based approaches has been added to this thesis. Parts of section 3.3 are published in Buckland *et al.* (2011a) and Buckland *et al.* (2011b); in both cases, I contributed the data analysis including the description of methods and a summary of the results. In particular, the example shown in Box 3.2 has been presented in Buckland *et al.* (2011a).

3.1   Two forms of detectability

As discussed briefly in chapter 1.2.3, when sampling ecological communities we usually encounter two different, but related issues of detectability — that of individuals and that of species.

The former stems from the fact that surveys are unlikely to 'detect' individuals from different species with the same probability (Yoccoz *et al.*, 2001; MacKenzie & Kendall, 2002; Buckland *et al.*, 2011a). A specific survey method will be more adapted to one species' behaviour than to that of another (Southwood & Henderson, 2000); it will work better in a certain environment or be unable to sample sites entirely randomly (e.g. due to limited access). Hence, it detects individuals from some species more easily than others (Magurran & Henderson, 2003; Hutchens & DePerno, 2009). This is in the following referred to as 'individual detectability' (Buckland *et al.*, 2011a). Strictly speaking, it does not only concern differences between species, but detection by a certain survey method can also differ between individuals from the same species, for example from different age classes or sexes (due to differences in size or behaviour); examples are fishing gear that might miss juveniles (Koeller *et al.*, 1986) or different vocalising behaviour in male and female birds when detection is by audio cues (Poesel *et al.*, 2004).

A different issue arises when no individual of a species is detected. Referred to as the 'detectability of a species' by Buckland *et al.* (2011a), this results in unobserved species in a data set. These pose a problem to diversity assessment because we can rarely determine why a species is missing from a sample (Magurran & Henderson, 2003; Mao & Colwell, 2005) and hence whether or not

it should be included in the assessment and at what point (if temporal trends in diversity are considered). Absences can have various causes: the species may have disappeared from the species pool; it may be temporarily absent from the region at the time of the survey (e.g. migrants); it may be present but absent from the sampled plots (more likely for rare species); it may be present on at least one sample plot, but evade detection (more likely for rare and elusive species). While in the latter two cases we surely would want to include the unobserved species in any form of diversity assessment (Gotelli *et al.*, 2010), this is debatable for the first two cases. Are we interested in a momentary state of biodiversity or long-term monitoring? We probably do not want to take truly absent species into account for a 'snapshot' of biodiversity, but might be interested in including such species when analysing changes over time and/or space. Climate change effects can be expected to add to this problem since species are likely to shift their ranges, and possibly natural habitats (Parmesan, 2006; Buisson *et al.*, 2008).

In the following, we discuss the effects of both forms of detectability on diversity assessment and, where possible, potential ways to deal with problems stemming from these issues. Our focus is on the goodness-of-fit measures introduced in the previous chapter.

3.2 Incomplete species detectability and its consequences

3.2.1 Unobserved species and sampling effects

We rely on observations at sampling locations within the study area to estimate evenness profiles based on $I_{\mathbf{p}}(\lambda)$, which then allow us to draw inference on the community. However, as pointed out above, species may be missing from samples (Colwell & Coddington, 1994; Yoccoz *et al.*, 2001; Gotelli & Colwell, 2001; Shen *et al.*, 2003; MacKenzie *et al.*, 2005).

Commonly used diversity measures are in general insensitive to unobserved species, i.e. they only quantify the diversity of the sample itself (Hill, 1973, 1997). This holds, for example, for measures contained in Hill's family of diversity numbers $J(\alpha)$. Only $J(0) = S$ would in theory be an exception if we explicitly decided to count missing species; the latter is rarely done in practice. However, see Chao & Shen (2003) for adjusting estimation of Shannon's index for unobserved species.

In contrast, the approach to evenness considered in chapter 2 is sensitive to unobserved species in a sample. Goodness-of-fit statistics treat zeros in the sample vector $\mathbf{n}$ as observations that are smaller than what would be expected for an even community; hence they increase departure from evenness. However, the contribution of zeros to the measure is evaluated differently whether we take a goodness-of-fit based view (i.e. use $I_{\mathbf{n}}(\lambda)$ to quantify evenness) or estimate the divergence $\hat{I}_{\mathbf{p}}(\lambda) = 1/2n I_{\mathbf{n}}(\lambda)$. Rewriting expressions (2.4) and (2.3) in chapter 2 slightly, we have

$$I_{\mathbf{p}}(\lambda) \;=\; \frac{1}{\lambda(\lambda+1)} \sum_{i=1}^{S} \left(p_i - \frac{1}{S} \right) Z_{\lambda}(\cdot)$$

$$I_{\mathbf{n}}(\lambda) \;=\; \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{S} \left(n_i - \frac{n}{S} \right) Z_{\lambda}(\cdot),$$

where $Z_{\lambda}(x) = x^{\lambda} + x^{\lambda-1} + \ldots + x + 1$ is a polynomial of order λ, evaluated at $x = \frac{p_i}{1/S}$ and $x = \frac{n_i}{n/S}$, respectively. We see that in the case of the power divergence $I_{\mathbf{p}}(\lambda)$ a zero $p_i = 0$ reduces the sum by a constant term $-1/S$, whereas this negative contribution depends on the total abundance of the remaining species if we look at the goodness-of-fit based $I_{\mathbf{n}}(\lambda)$. Hence, while making a contribution to divergence, missing species only lead to a translation from a measure based on a sum over species proportions (such as the measures in Hill's family). This puts divergence measures closer to classical diversity measures (Read & Cressie, 1988). While this makes no difference in relative terms, if we evaluate evenness by a goodness-of-fit statistic as an absolute measure (i.e. not divided by $2n$), the negative contribution by zeros is bigger, the larger the total abundance of the sample is (or as we will see later, the total estimated abundance). Quantifying evenness by either a divergence based measure or by a goodness-of-fit statistic thus involves an explicit decision about how we want to treat zeros.

Often, there will be a valid reason to consider missing species as part of the community (MacKenzie *et al.*, 2003). The goodness-of-fit method then allows us to include these and thus goes beyond a simple assessment of characteristics of a sample. In particular, this enables us to register cycles in diversity patterns due to migration or the change in evenness, if a species genuinely disappears from the population. (For an example see Box 3.1.) However, given the sensitivity of the goodness-of-fit approach to zeros, a careful and educated decision about the assumed species catalogue is needed in order to avoid biased results.

Even if we know the exact number of species, zeros in the vector of sample proportions $(n_1/n, \ldots, n_S/n)$ cause problems: $I_{\mathbf{n}}(\lambda)$ is no longer continuous at

$\lambda = -1$ if $n_i = 0$ for some i. Consequently, the evenness profiles estimated by $I_{\mathbf{n}}(\lambda)$ have an asymptote at $\lambda = -1$ which cannot be interpreted. There are several ways to overcome this problem and derive informative evenness profiles, at least partially. The standard approach for traditional diversity measures is to discard zeros and calculate diversity indices or construct diversity profiles solely based on the species observed in the sample (Hill, 1973, 1997). As discussed above, this accounts for the properties of the sample, but not those of the community.

If we want to retain full diversity profiles, a common solution is to add small quantities ϵ to achieve non-zero values for all n_i. Alternatively, we can restrict the analysis to positive parameter values, taking into account zeros, but without the explicit focus on rare species provided by the negative parameter values. The latter will be a good solution if sampling intensity is low, since it discounts the part of the sampled distribution that carries higher uncertainty. Finally, it might be possible to fit a model to the sample data predicting the underlying community structure. Instead of estimating the evenness profile directly from the sample, we can then do so based on the predictions from the model. We return to simulations from Tokeshi's models in 3.2.3 below, to study and discuss these approaches in application.

Box 3.1: Sensitivity of diversity measures to unobserved species

Consider the following toy example of a community where the most abundant species shows a seasonal migration pattern (see Fig.), while all other species' abundances stay constant. The initial abundance vector is $\mathbf{N}_0 = (376, 145, 43, 9,8,7,3,3,3,3)$.
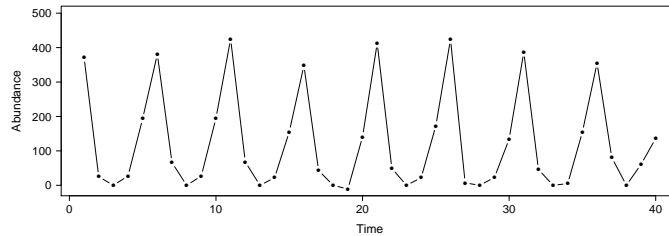
We evaluate Hill's diversity number at $\alpha = 2$ (which results in the inverse of Simpson index), the family for divergence measures $I_{\mathbf{p}}(\lambda)$ at $\lambda = 1$ (which gives a different transformation of Simpson's index) and the goodness-of-fit statistic $I_{\mathbf{N}}(1)$. Both Hill's $J(2)$ and the divergence measure $I_{\mathbf{p}}(1)$ show almost the same values at points where the most abundant species is present and dominant, compared to points when it is absent. The goodness-of-fit based measures on the other hand follow the the migration pattern of the dominant species; they drop when the migrating species is most abundant, reflecting the low evenness at these points.
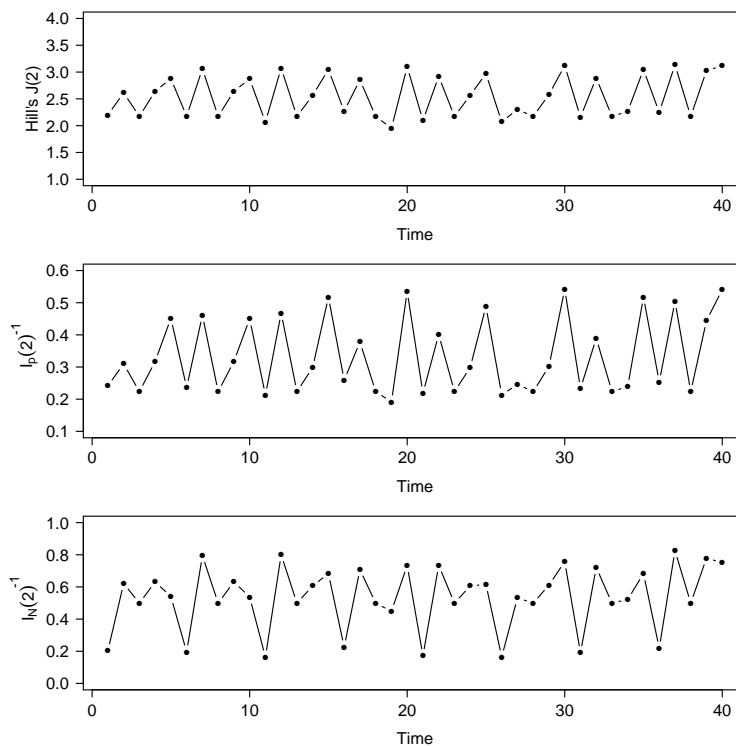
[The plots for $I_{\mathbf{p}}(\lambda)$ and $I_{\mathbf{N}}(1)$ are based on the inverse (for direct comparability with traditional indices) where the y-axis has been scaled appropriately.]

3.2.2   Determining S

Species detectability obviously is a prime issue when species richness is considered the main measure of choice for biodiversity (Magurran, 2004; Gotelli & Colwell, 2001, 2011). Because the number of species also indirectly influences what is taken to be the reference point for evenness, evenness measures are not unaffected by determining the 'true' number of species S (Brose *et al.* (2003); Magurran (2004); and see section 2.6.3 above). In both cases, we require an answer to the question "What is 'S'?". However, different approaches will be appropriate depending on the aim of diversity assessment. In general, the 'true' species richness of an ecological community will be 'elusive' (May, 1988; Gotelli & Colwell, 2001): first, rare species are likely to be missing from samples (Longino *et al.*, 2002; Mao & Colwell, 2005) and second, species richness is neither scale-invariant nor insensitive to sampling effort (Brose *et al.*, 2003) — we expect to find more species in larger survey areas (species-area relationship, Rosenzweig (1995)) and when we increase sampling effort (species accumulation curves, Colwell & Coddington (1994)). Hence, the observed number of species provides only a lower bound on the true number of species.

By now, a range of methods have been developed to adjust the observed S for missing species (Bunge & Fitzpatrick, 1993; Gotelli & Colwell, 2011). They can broadly be distinguished into three groups: parametric approaches fit a model to either the rank-abundance plot (i.e. the ordered SAD) or to the species accumulation curve. In the first case, the number of missing species from the sample can, for example, be 'unveiled' by the tail of the fitted distribution (Preston, 1948), whereas in the second approach S is determined by extrapolation of the curve until an asymptote is reached (Colwell & Coddington, 1994). Alternatively, various non-parametric estimators have been introduced (Chao, 1984; Chao & Lee, 1992); the latter have advantages over parametric methods (Gotelli & Colwell, 2011) and statistical methodology has been thoroughly studied and improved to reduce bias (Chao *et al.*, 1993; Lee & Chao, 1994; Chao *et al.*, 2009; Lopez *et al.*, 2012). The majority of these estimators are based on the idea that the ratio of 'singletons' (species observed exactly once in a sample) to 'doubletons' can be used to estimate the number of unobserved species (Good, 1953; Chao, 1984).

Species richness estimators share the assumption of other types of diversity assessment that individuals are sampled randomly (Magurran, 2004). As we will see below in 3.3, this assumption is almost always violated because of variability in individual detection probabilities across species. Variation in detectability

across species is likely to affect estimates of S (Boulinier *et al.*, 1998; Nichols *et al.*, 1998). One way to avoid this issue of individual detectability, would be to make sample plots small enough that a complete census of each plot is possible. However, at the same time this generally means that on a single plot many species will in fact be missing and consequently, we need a big number of small plots throughout the survey area to assure that most of the species present are actually observed (Buckland *et al.*, 2011a, In prep). This might not always be feasible (Lawton *et al.*, 1998)

Although species richness is probably the most intuitive measure of diversity, the difficulties of its estimation and the fact that it accounts only for one aspect of species diversity do not make it a preferred headline index for monitoring diversity on a large scale (Buckland *et al.*, 2005, In prep). If our aim is long-term monitoring on a large spatial scale (as with regards to the international biodiversity targets), then on one hand we are interested in the entire biota of a region, on the other hand species richness in a large region (such as a country or larger) will probably not change much on the typical time scale on which data is collected, except for short-term fluctuations. Extinction rates could in some way serve as a proxy to assess whether the 'rate of loss of biodiversity' has been 'halted', but determining extinction caused by anthropogenic factors besides naturally occurring extinction as well as immigration events, and estimating extinction rates reliably, suffers from even greater uncertainty (Balmford *et al.*, 2003).

However, a complete species inventory is often not necessary for establishing trends in a headline index (such as the Living Planet Index, for example). It is sufficient to determine a set of 'indicator species', an appropriate catalogue of species that are representative of the entire community and can reflect the status of the whole ecosystem (Landres *et al.*, 1988; Pearson, 1994). Hence it might be more important to determine functional groups (Tilman *et al.*, 1997) or distinguish between different habitat types (Boulinier *et al.*, 1998; Newson *et al.*, 2009). Although the contribution to ecosystem functioning of rare species is often disputed (Lawton, 1994; Lyons & Schwartz, 2001), in terms of representativeness as well as conservation concern, at least some of the rare species should be included in such an inventory (Lawler *et al.*, 2003). On the downside, rare species increase uncertainty in those indices that are sensitive to them (such as the geometric mean of relative abundances or the goodness-of-fit measures for negative λ). If an index permits integration of data from different surveys, additional species-specific surveys can be conducted to target rare species and reduce the uncertainty (MacKenzie *et al.*, 2005). This is done,

for example, for the Living Planet index (Loh *et al.*, 2005).

Another issue of interest in long-term monitoring is change in species composition (turnover) (Magurran & Henderson, 2003; Magurran, 2011; Magurran *et al.*, 2010). Because of the relative stability of species richness on large spatial scales, presence-only based assessment of turnover is unlikely to provide enough information to reveal changes. Again, monitoring the estimated abundances of a set of indicator species might be a better solution.

Species composition can change across space as well as over time. In the first case, we are comparing communities at different locations (or more precisely, different expressed communities stemming from the same assumed species pool), while in the latter, changes in a single community are followed through time. A combination of both is possible and of particular interest (spatio-temporal trends); however it is more complicated to analyse and untangle effects on both scales (Levin, 2000). Again, an appropriate inventory of species is needed as a basis for comparison. Because of the sensitivity of (observed) species richness to scale (area of sample plots) and sampling effort, changes in the observed number of species between communities (across time or space) are likely to occur (Rosenzweig, 1995; He & Condit, 2007). As with any form of diversity assessment, but maybe more obviously here, assessment should hence not be based on the observed number of species without any adjustment. Rarefaction (Sanders, 1968; Simberloff, 1972; Gotelli & Colwell, 2001; Colwell *et al.*, 2004; Gotelli & Colwell, 2011) is a well established technique that 'standardises' species diversity by interpolation to a common (and hence comparable) number of individuals or samples. However, this 'down-sized' comparison reflects mostly the more common species as rare species affect mainly the upper end of the rarefaction curve. Rarefaction curves also average across species (observed) in order to determine the expected number of species. As long as we are purely interested in differences in numbers of species or even the 'effective number of species', this will not matter. Detecting changes in species composition on the other hand relies on keeping track of species identities in some way. Like other forms of diversity assessment, long-term and large-scale monitoring of compositional changes can be based on a set of indicator species as the basis for any evaluation of differences in species composition in the sense of identifying principal changes. Even if we make sure that plot sizes are standardised and sampling effort constant throughout the survey area (or use a stratified design where this is not possible), any measure of spatial turnover depends on the scale at which we choose to evaluate these measures (Gering & Crist, 2002; Mac Nally *et al.*, 2004; Gaston *et al.*, 2007).
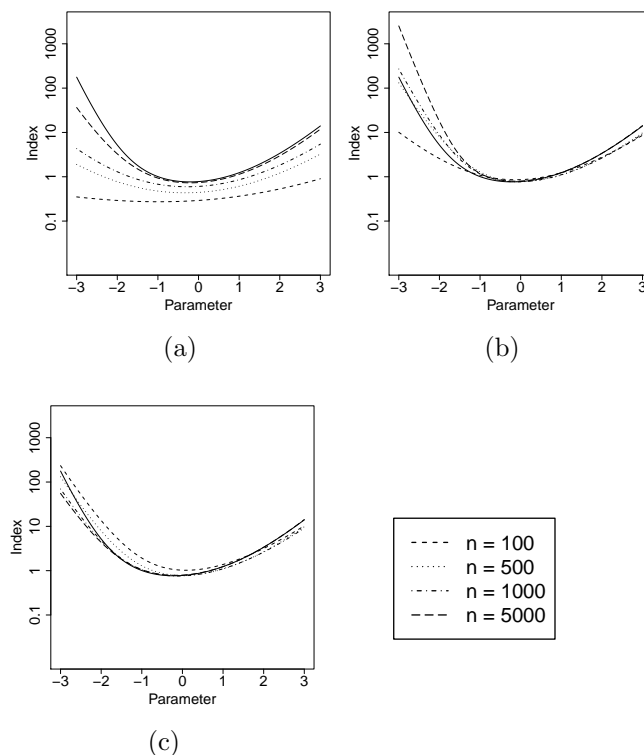
FIGURE 3.1 – Goodness-of-fit based evenness profiles for 4 samples with increasing sample size ($n = 100, 500, 1000, 5000$) from Tokeshi's power fraction model with $k = 0.4$. (a) only taking into account observed species (b) correcting for all unobserved species by adding the same $\epsilon = 0.1$, independent of sample size and (c) adjusting the ϵ correction according to sample size. The solid line shows the profile for the underlying Tokeshi model.

3.2.3   Tokeshi's models revisited (scenario 3)

To study the effects of sampling and the problem of unobserved species when diversity is estimated, we once more use simulated data derived from Tokeshi's models of niche apportionment (see chapter 2.7 and Studeny *et al.* (2011)). After the two scenarios studied in the previous chapter, a third scenario is now looked at. To investigate sampling effects, diversity profiles for one sample are studied for each of three power fraction models. Values 0, 0.4 and 1 are chosen for the model parameter k to cover a range of uneven, intermediate and even distributions. Different ways of dealing with unobserved species in a sample are applied.

After acquiring a sample from the community of interest, the data analyst first needs to decide whether or not he or she wants to take unobserved species into account. If counts equal to zero are included in the evaluation of evenness profiles, a small ϵ has to be added in order to plot the profile for $\lambda \leq -1$ and hence a second question is the appropriate size of this small quantity added.
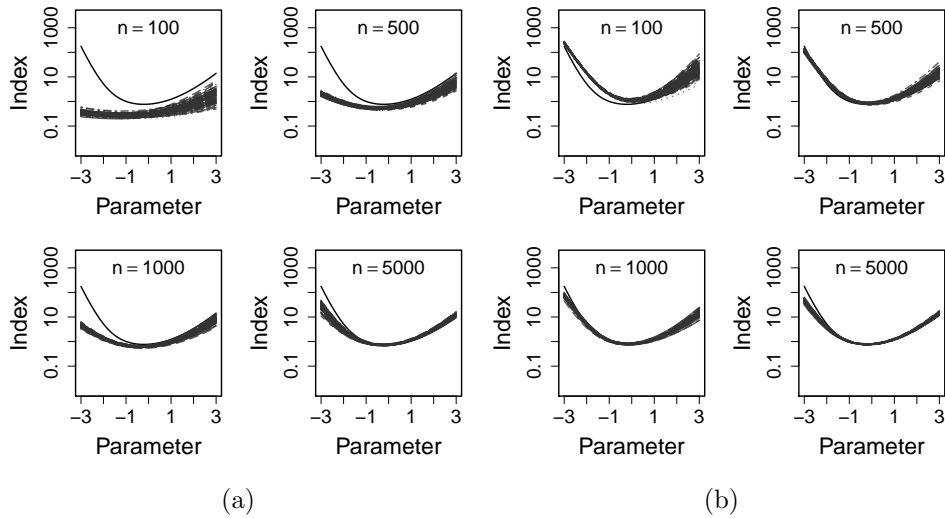
FIGURE 3.2 – Sampling variation in goodness-of-fit based evenness profiles for 4 samples with increasing sample size ($n = 100, 500, 1000, 5000$) from Tokeshi's power fraction model with $k = 0.4$ (a) only taking into account observed species and (b) with variable ϵ correction.

Figs. 3.1 and 3.2 show profiles based on samples of increasing size taken from a Tokeshi power fraction model with $k = 0.4$ and their sampling variation, if zeros are omitted or corrected for. Apart from the convergence for larger sample size (as discussed in the section on asymptotics of the goodness-of-fit measures towards the 'true' divergence), we can see that the correction by ϵ can reduce bias drastically, especially for smaller samples. However, the performance of the correction is affected by the choice of the size of ϵ. What is a 'good' value depends on the size of the fraction of the population sampled, as well as the degree of unevenness of the community.

Once we decide to correct a sample for missing species, the question about the (in general) unknown number of species S arises. The number of absences, added to the sample vector **n** as zeros, is determined by making an assumption about S. We could, for example, use one of the existing species richness estimators or rarefaction methods for an 'educated guess'. Naturally, it is of interest to what extent deliberate underestimation (by taking only the number of species observed in the sample) versus overestimation of S affects our inference on the true community structure. To investigate this, we plot sample profiles for a range of assumed values of S along with the true profile (Fig. 3.3) for a series of Tokeshi models with variable evenness. This is again compared to other diversity profiling methods, namely Hill's quotients $J(a)/J(2)$ and relative logarithmic evenness suggested by Jost (2010). Where a higher number of
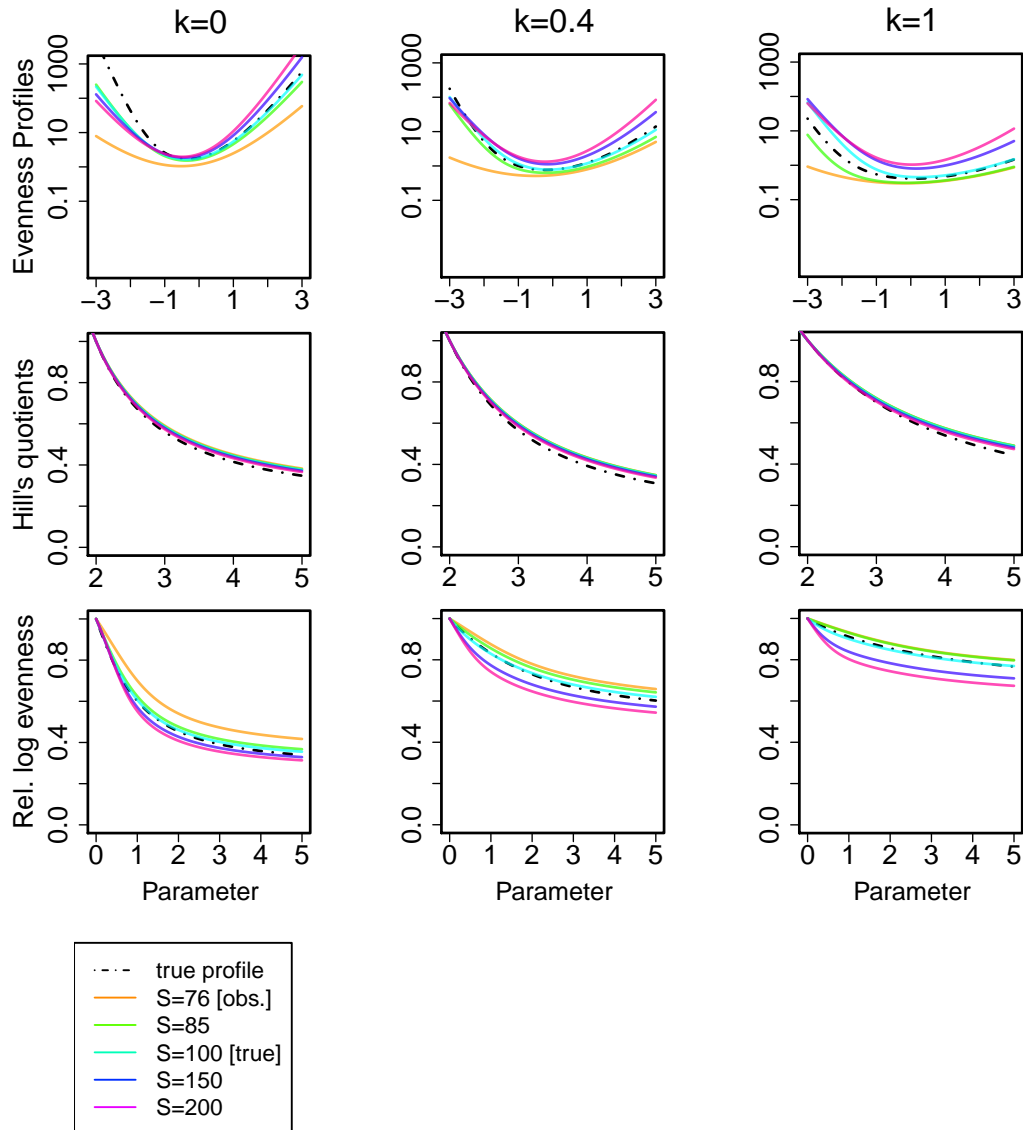
FIGURE 3.3 – Estimation of evenness and other diversity profiles based on one sample of a series of Tokeshi's power-fraction models ($k = 0, 0.4, 1$). Estimation is based on a range of assumed values for the number of species S. If only observed species are taken into account, evenness profiles as well as relative logarithmic evenness are markedly biased.

species was assumed than was observed, a small ϵ is added to species counts in order to achieve a full profile plot including negative values for λ.

For all profile plots in Fig. 3.3, we see that in general we do better by taking unobserved species into account. This is relevant if we are interested in the contribution of rare species (or changes to them), which can be detected by looking at the part of the evenness profile corresponding to negative parameter values. Provided we do not grossly underestimate the number of missing species, the picture we get for rare species is quite robust to mis-specification of our

sample. There is a slight drawback: while the curves are largely similar, the bias from the true profile depends on the value of ϵ. Here, the same ϵ correction was applied to all profiles. However, we see that it does not work equally well for all Tokeshi models. For the positive parameter range, the sample evenness profiles are not much affected by the choice of ϵ as the dominance structure is not changed by adding ϵ. If we assess evenness by looking at the curvature rather than the absolute values of the evenness profiles in Fig. 3.3, the dependence on ϵ is in general less of an issue. By contrast, if we omit unobserved species, profiles are almost flat, giving a false indication of evenness of the community. This holds no matter which evenness measure we choose. Even if we are only interested in the (true) dominance structure of a community, hence focussing on positive parameter values, disregarding unobserved species leaves a bias. Bias in the estimated profiles increases slowly as our estimate of S moves further from the true number of species, in either direction. The exception is Hill's evenness quotient based on $J(2)$ which appears invariant to the assumed S. However, bias and precision of the latter are highly sensitive to the choice of ϵ (see Appendix F). We also need to keep in mind that Hill's evenness quotients generally fail to display the difference in the models correctly and can place species abundance distributions with high evenness close to those with low evenness. This is also evident in Fig. 3.3 with little difference in the quotients for the power-fraction models with $k = 0$ and $k = 0.4$.

The difficulty of choosing an appropriate ϵ is certainly inconvenient. An approach that provides an alternative in some cases is model-based estimation of the evenness profiles. Instead of a direct estimation from samples, the latter can be used to fit a model provided we have repeated samples and a species is not missing from all of them. Evenness profiles can then be derived from the species abundances predicted by the fitted model. Nevertheless, in general this does not free us from the need to make an assumption about S. To study whether or not a model-based approach might have an advantage over the ϵ correction, we look at the sample from the Tokeshi power-fraction model with $k = 0$. This was the example that showed the most remaining bias for negative λ values when the evenness profiles were plotted after correction by ϵ (Fig. 3.3). Again we deliberately under- and overestimate S, while fitting a Tokeshi model to the sample by importance sampling of k (Doucet *et al.*, 2001). Importance sampling belongs to the set of Monte Carlo simulation methods as they are frequently used in Bayesian analysis; we consider here a basic importance sampling procedure where a random sample of 500 values from a uniform (prior) distribution on the unit interval $[0, 1]$ is generated for k and a Tokeshi model is simulated for each of

FIGURE 3.4 – Model-based estimation of the evenness profile for Tokeshi's power fraction model for $k = 0$. While under- and overestimating the number of species S, the model parameter k was estimated from one sample ($n = 500$) and the evenness profile plotted for the estimated model. Bias along the negative parameter range essentially vanishes.
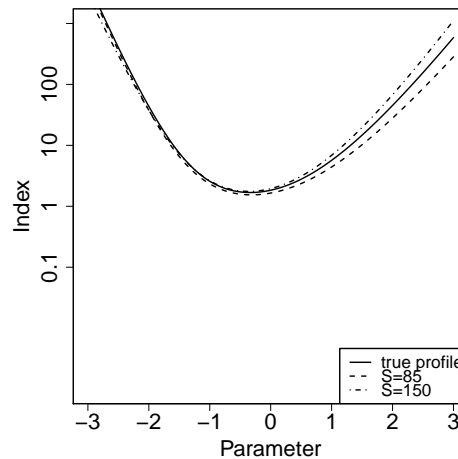
these values; the importance sampling weights are determined by the probability that the observed sample is a realisation of the corresponding Tokeshi model and the posterior distribution is simulated by a weighted resampling of the original sample for k; we derive the estimated k as the mean of this resample. After simulating the Tokeshi model with this estimated k, we compare the evenness profiles for the fitted model to the original Tokeshi model the sample has been taken from.

Indeed, bias is visibly reduced for both under- and overestimates of S, in particular for the negative parameter range of the profiles (see Fig. 3.4). This could be extended to a joint estimation of k and S by the same importance sampling procedure (assuming independence of k and S the joint density would be the product density). However, there is little or no gain in bias reduction compared to setting S in advance (results not shown). A model-based estimation was essentially used in the example for British farmland birds in chapter 2.8 above where a generalized additive model was used to get predicted counts in every survey location (Fewster *et al.*, 2000). We will further exploit a model-based approach when analysing temporal and spatial trends in diversity of British breeding birds in the next chapters.

3.2.4   Assessing large-scale temporal and spatial trends in diversity

A particular question of interest is the decline or increase in diversity over time (Magurran & Dornelas, 2010). Monitoring and conservation programs often focus on rare species (Lyons & Schwartz, 2001). Our example of the data from the Common Bird Census in chapter 2 showed that the proposed evenness profiles are able to separate changes that are mostly within the rare species from

those that affect common species. We will see a more detailed example later, in chapter 4.

As discussed in the previous sections, the goodness-of-fit based evenness measures allow us to include species that are temporarily absent or undetected (Studeny *et al.*, 2011). Valid quantitative and qualitative conclusions on time trends can be drawn as long as we can establish a representative catalogue of species which we assume form the community over the period of time we are interested in. The loss of biodiversity should a species truly disappear from the community would be registered as long as the species is included in the assumed catalogue of species in the community. If it is found necessary to revise this list as more data become available, we can re-evaluate the index family over the entire time period, to improve our knowledge of the dominance and rarity characteristics of the community (Magurran & Henderson, 2011). Headline indices used to monitor national and international biodiversity trends are often evaluated for a chosen set of indicator species, which is assumed to be representative of the entire biota under consideration (see discussion in section 3.2.2 above). However, in order to make use of the evenness profiles introduced in chapter 2 and in order to gain a realistic picture of the state of biodiversity, this list should not only be based on the most common species. Trends for rare species are often different from those in common species (Lawler *et al.*, 2003) and it is one advantage of the goodness-of-fit measures that rare species can be included in the analysis — as long as missing observations can be appropriately dealt with. We will return to the specific difficulties of including less common species in biodiversity assessment when analysing temporal trends in diversity of British breeding birds in the next chapter.

Changes in counts of individuals, in particular of rare species, are driven by stochasticity and hence undergo short-term fluctuations. In addition to getting predictions for missing observations, a model of either observed counts or abundance indices can in this case also be used to smooth out these short-term fluctuations before analysing long-term trends in a diversity statistic. There are parametric as well as non-parametric options for scatterplot smoothers, including generalized additive models (Hastie & Tibshirani, 1990; Wood, 2006, 2008), kernel regression methods (Bowman & Azzalini, 1997) and latent Gaussian models (Rue *et al.*, 2009).

If we are interested in temporal trends alone, appropriate randomisation of survey sites across the region of interest, assures representativeness of sites and eliminates bias stemming from spatial trends in data (Thompson, 1992; Buckland *et al.*, 2011a). (Sometimes additional sampling techniques, such as

stratification, have to be applied to account for inhomogeneity of the data across space.) However, in the context of large-scale biodiversity assessment there is also an interest in analysing spatial patterns in diversity explicitly (e.g. identify 'hotspots' of biodiversity and of change in biodiversity, Rodrigues *et al.* (2000)). This is not only relevant to objectives aimed at conservation and protection of biodiversity, such as the 2010-20 targets. Revealing spatio-temporal trends in $\alpha-$diversity as well as changes in species composition across space and time simultaneously (spatio-temporal turnover) could also help to track effects of climate change (Hannah *et al.*, 2002). Similarly to smoothing temporal trends, spatial modelling of observed species' counts can be applied to predict diversity in continuous space by a smooth surface. Such a model-based approach is able to take account of spatial, and even spatio-temporal autocorrelation in the data and can incorporate known and unknown sources of variation (Cameletti *et al.*, 2012; Lindgren *et al.*, 2011). While the former can be thought of as the standard terms in a linear model, the latter can be described by a random structure in space where only distributional assumptions are made. A detailed spatial analysis of this kind is beyond the scope of this thesis, but we attempt a 'peek behind the curtain' in this direction with a first basic analysis in chapter 5.

3.3   Individual detectability and diversity assessment

3.3.1   Bias in diversity indices from differences in individual detectability

While ecologists are very much aware of the problems arising from species that are completely missed in the observation process (discussions about the definition of the 'community', species richness estimators, etc.), the issue of differences in individual detectabilities has received less attention (Yoccoz *et al.*, 2001; Buckland *et al.*, 2011a). Indeed, diversity measures are often calculated under the implicit assumption that the data are in fact a random sample of the community (Magurran (2004), p.136), that consequently every individual in the community has the same probability of being detected and that the species' proportions $p_i = \frac{N_i}{N}$ are correctly reflected in the observed counts. However, as discussed in section 3.1 above, survey methods can rarely guarantee equal detection probabilities for all individuals, except if we have census data or sample plots are so small that every individual is detected. The latter is in general not feasible and hence ignoring this kind of variability in the detection process when estimating diversity indices based on the data, is expected to lead to biased results (Yoccoz *et al.*, 2001; Buckland *et al.*, 2011a).

Extending the notation introduced in chapter 1, let P_i be the probability of detecting an individual of species i. (As above, a further superscript t will indicate points in time where temporal trends in diversity indices are considered.). Then $\mathbb{E}(n_i) = N_i P_i$ is the expected number of detected individuals from species i. Expected species proportions in the sample are then given by $q_i = \frac{\mathbb{E}(n_i)}{\sum_i \mathbb{E}(n_i)}$ and commonly estimated by the sample proportions $\hat{q}_i = \frac{n_i}{n}$.

To derive the diversity of the ecological community the sample is taken from, a diversity index would have to be calculated based on the true species proportions $p_i = \frac{N_i}{N}$. Since these are in general not directly observable, diversity indices have to be estimated, ideally based on an estimator $\hat{p}_i = \frac{\hat{N}_i}{\hat{N}}$.

However, in practice indices are often estimated using the sample proportions $\hat{q}_i$, i.e. we estimate the diversity of the expected sample abundance distribution **q** rather than the true underlying species abundance distribution **p**. Any diversity index that accounts at least partly for evenness, usually measures variation in species proportions in some sense.

Box 3.2: Adjusting diversity indices for detectability

Variability in individual detection probabilities across species can bias diversity assessment.

Standard diversity indices (Shannon's $\hat{H} = -\sum_{i=1}^{S} \hat{p}_i \log \hat{p}_i$, Simpson's $1 - \hat{D} = 1 - \left(\sum_{i=1}^{S} \hat{p}_i{}^2\right)$ and a geometric mean of relative abundances G) were evaluated for a set of 20 Scottish farmland birds. To account for detectability, the estimator $\hat{p}_i = \hat{N}_i / \hat{N}$ was used where $\hat{N}_i = n_i / \hat{P}_i$ is the estimated abundance and $\hat{P}_i$ the estimated probability of detection for an individual of species i. The thus adjusted diversity indices are labelled with an 'a'. Where the indices are based on the sample proportions $\hat{p}_i = n_i / n$, this is indicated by a subscript 'c'. [Figure taken from Buckland *et al.* (2011a)]
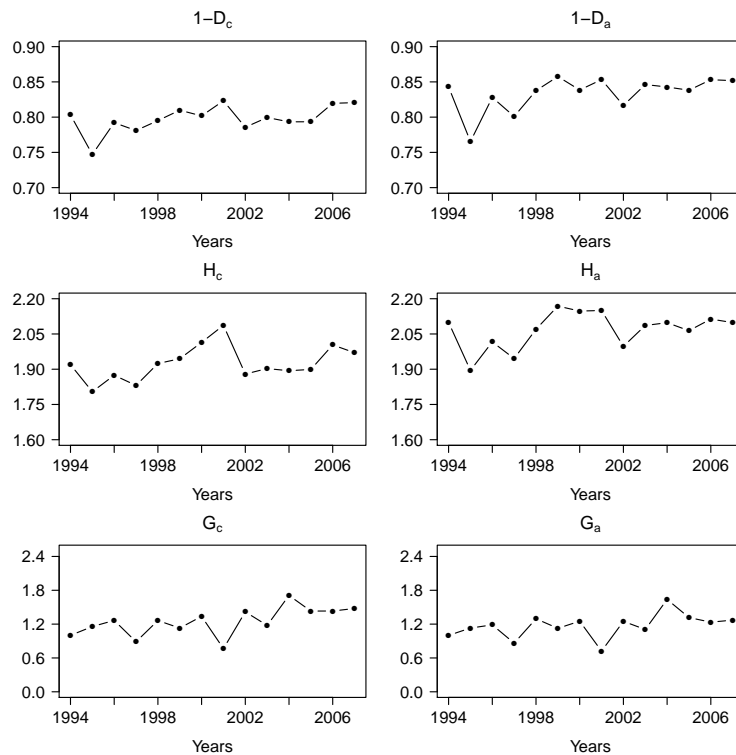
Indices that are based on species proportions are biased if their calculation is directly based on species counts; because the detection probabilities have been estimated by pooling data across all survey years 1994-2007 for a large part of the species and do not change much over time where they have been estimated for each year, we do not see any differences in the geometric means (see text for further explanation). However, when testing for trend in the indices by fitting a linear model, evidence for an increase in the geometric mean G_c became non-significant when we adjusted counts for detection (for detail see Buckland *et al.* (2011a)).

To study effects of calculating diversity based on the sampling proportions $\mathbf{q}$, Buckland *et al.* (2011a) compared variation in $\mathbf{q}$ and $\mathbf{p}$ by deriving

$$\operatorname*{var}_i(\log \mathbb{E}(n_i)) = \operatorname*{var}_i(\log N_i) + \operatorname*{var}_i(\log P_i) + 2\operatorname*{cov}_i(\log N_i, \log P_i).$$

We can see that in general variation in $\mathbf{q}$ is not the same as that of the true species abundance distribution. If we can assume detection probabilities to be independent from species' abundances (i.e. $\operatorname{cov}(\log N_i, \log P_i) = 0$), variation in the $\mathbb{E}(n_i)$ is the same as that in the N_i if and only if $\operatorname{var}(\log P_i) = 0$. This is equivalent to all P_i being constant, and hence all species being equally detectable. In general, variation in $\mathbb{E}(n_i)$ will be higher than in the N_i, meaning that we underestimate evenness when estimating diversity based on $\hat{q}_i$. Unless we make the (probably unrealistic) assumption that abundance and detectability are independent, evenness can be under- or overestimated (depending on the direction and size of the correlation between $\mathbf{N}$ and $\mathbf{P}$). Only if $\operatorname*{var}_i(\log P_i) + 2\operatorname*{cov}_i(\log N_i, \log \pi_i) = 0$, is the variation of the two distributions the same. This corresponds to $\operatorname*{cov}_i(\log N_i, \log P_i) = -0.5\operatorname*{var}_i(\log P_i)$. If there is high variability in detection probabilities between species, this can only occur if the correlation between abundances and detectabilities is strongly negative.

Box 3.2 illustrates this using Scottish farmland breeding birds as an example. In the next chapter, this is extended to an analysis of diversity trends in British breeding birds across the UK and different habitat groups. Here as later, we account for detectability by estimating diversity indices based on $\hat{p}_i = \frac{\hat{N}_i}{\hat{N}}$. Appropriate survey design is needed, though, in order to estimate detection probabilities (see chapter 1.2.3).

In this context, the geometric mean has an advantage over measures based on species proportions (Buckland *et al.*, 2011a,b). Since it is a relative measure summarising abundances relative to each species abundance at a baseline point in time (see chapter 1.3.1), it is unaffected by variability in detection probabilities between species as long as these remain constant in time: if $P_i^{t_0}$ is the detection probability of species i at the reference point and $P_i^t = P_i^{t_0}$, it cancels when the geometric mean is estimated

$$
\begin{aligned}
\hat{G}_t &= \exp\left(\frac{1}{S}\sum_{i=1}^{S}\log\frac{\hat{N}_i^t}{\hat{N}_i^{t_0}}\right) = \exp\left(\frac{1}{S}\sum_{i=1}^{S}\log\frac{\hat{\mathbb{E}}(n_i^t)/P_i^t}{\hat{\mathbb{E}}(n_i^{t_0})/P_i^{t_0}}\right) \\
&= \exp\left(\frac{1}{S}\sum_{i=1}^{S}\log\frac{\hat{\mathbb{E}}(n_i^t)}{\hat{\mathbb{E}}(n_i^{t_0})}\right) = \exp\left(\frac{1}{S}\sum_{i=1}^{S}\log\frac{n_i^t}{n_i^{t_0}}\right).
\end{aligned}
$$

However, if there is a trend in detection probability over time, the geometric mean will be affected more severely than absolute diversity indices.

3.3.2   Effect of differences in detection probabilities on the goodness-of-fit measures

Based on this discussion, we conclude that diversity measures should be adjusted for individual detectability, in particular if they are quantifying evenness. As the goodness-of-fit measures introduced in the previous chapter are conceptually evenness measures and at the same time sample based because of their connection to goodness-of-fit statistics, we will now study the effect of explicitly including detection probabilities into their estimation.

As stated above, an essential assumption in the goodness-of-fit framework is that the sample vector $\mathbf{n}$ comes from a multinomial $\mathcal{M}(n, \mathbf{p})$ distribution. This assumption implies that the detection of individuals of different species is equally likely — analogous to what has been discussed for other diversity measures above. In reality this is rarely the case. What is the consequence for the goodness-of-fit approach to diversity quantification if we drop the assumption of the multinomial model and take into account variable individual detection probabilities?

Assume that data are collected in a way that provides information on these detection probabilities and allows us to derive estimates $\hat{P}_i$ (see chapter 1.4 on survey design; for further details on appropriate survey and estimation methods refer to Buckland *et al.* (2011a)). Consider observations of different species to be sampled independently and according to a product binomial model, i.e. $\mathbf{n} \sim \bigotimes_{i=1}^{S} B(N_i, P_i)$ (Fewster & Jupp, 2009). (Note that we can retrieve a multinomial model by conditioning on the sample size $n = \sum_{i=1}^{S} n_i$.) Consider the family of power divergences (2.4) in this setting. Instead of using the sample proportions $\mathbf{q}$ directly to estimate $\mathbf{p}$, we can use adjusted estimates $\hat{p}_i = \frac{\hat{N}_i}{\hat{N}} = \frac{n_i/\hat{P}_i}{\sum n_j/\hat{P}_j}$. Plugging this into the family of power-divergences (equation 2.3 in chapter 2.2), we derive the following family of estimators

$$\hat{I}_{\hat{\mathbf{p}}}(\lambda) = \frac{1}{\lambda(\lambda + 1)} \sum_{i=1}^{S} \frac{\hat{N}_i}{\hat{N}} \left[\left(\frac{\hat{N}_i}{\hat{N}/S} \right)^{\lambda} - 1 \right]. \tag{3.1}$$

Alternatively, from the point of view of goodness-of-fit of the sample $\mathbf{n}$ to a completely even distribution, we could look at the departure of the observations from what we expect to see under perfect evenness. With the product binomial

model we have $\mathbb{E}_{\mathbf{p}=\mathbf{p}^*}[n_i] = P_i N/S$ if the sample came from a perfectly even community. This changes Cressie and Read's family of goodness-of-fit statistics to

$$I_{\mathbf{n}}(\lambda) = \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{S} n_i \left[\left(\frac{n_i}{NP_i/S} \right)^\lambda - 1 \right]. \qquad (3.2)$$

Since N, P are unknown, we need to use the estimated detection probabilities $\hat{P}_i$ and the estimated total abundance $\hat{N} = \sum \frac{n_i}{\hat{P}_i}$. Comparing the estimators of (3.1) and (3.2), we see that

$$
\begin{aligned}
\hat{I}_{\mathbf{n}}(\lambda) &= \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{S} n_i \left[\left(\frac{n_i}{\hat{N}\hat{P}_i/S} \right)^\lambda - 1 \right] \\
&= \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{S} \hat{P}_i \hat{N}_i \left[\left(\frac{\hat{N}_i}{\hat{N}/S} \right)^\lambda - 1 \right].
\end{aligned}
$$

Hence, if variable detection probabilities between species are taken into account, the estimators based on the family of goodness-of-fit statistics are no longer equivalent to the plug-in estimators for the family of power divergences, but give a 'weighted' version of $2\hat{N}\hat{I}_{\hat{\mathbf{p}}}(\lambda)$ where the weights are equal to the detection probabilities P_i (but note that the vector $\mathbf{P} = (P_1, \ldots, P_S)$ does in general not sum to one here).

In applications where we can derive abundance estimates adjusted for detection probability, this leaves us with the decision of which approach to take. We have seen at the beginning of this chapter, the goodness-of-fit measures react differently from the divergence family to zeros in a sample. As discussed in 3.2.1 above, if we have reason to include unobserved species, evaluating evenness profiles based on the goodness-of-fit family can be more informative than based on the estimated divergences. However, we showed here that there is a conceptual difference if we incorporate detectability into these measures. The goodness-of-fit statistics no longer provide a direct estimate for the divergences. While they still have their sample-based interpretation, this only indirectly provides information about the underlying species abundance distribution. The divergences on the other hand are based on the concept of a 'distance' between two distributions (the true underlying species abundance distribution and the hypothetical even distribution) and can be evaluated from a sample by the plug-in estimator $\hat{I}_{\hat{\mathbf{p}}}(\lambda)$. This follows a standard statistical framework. (With some

effort, we could potentially develop a similar framework for the goodness-of-fit setting by employing a metapopulation framework in which we can interpret the community under consideration as a realisation, and hence a sample, of the metapopulation.) For the application to diversity trends in the next chapter, we decided to use the estimator $\hat{I}_{\hat{\mathbf{p}}}(\lambda)$ rather than the purely sample-based evaluation of the goodness-of-fit statistics.

3.4 Conclusions

The goodness-of-fit approach is sensitive to zeros in the species proportions. This allows us to include species in the diversity assessment even if they are not detected in our samples. As with any other approach, but perhaps more obviously here, careful consideration of the assumed list of species to include in an analysis is required. In reality the reasons for missing species are manifold, and we might want to distinguish between a true absence, a temporary absence, and rarity or difficulty of detection of a species that is present (Gotelli *et al.*, 2010). Assessing underlying evenness based on a sample from the Tokeshi models showed that omitting unobserved species gives a false impression of higher evenness for both rare and common species, for any of the evenness metrics considered.

However, as long as we do not greatly underestimate the true number of species, mis-specification has little effect on inference for rare species. Analysis based on the curvature of the evenness profile is even less sensitive to the number of species. Some bias for common species remains; this is shared by alternative approaches such as Jost's evenness factors and their logarithmic transformations. To draw evenness profiles that extend to unobserved species, some small quantity ϵ must be added to counts. Some caution is needed when this quantity is chosen, as the negative parameter range is sensitive to this choice. Thus, if the information about rare species is not essential for the objectives of a study or if uncertainty about rare species is too high, it is advisable to plot profiles for the positive parameter range only. In this case, dominance is still displayed relative to all species and not only the sampled ones.

Where the data are such that a statistical model can be fitted, diversity assessment can be based on the predictions from the model instead of the observations. This is a good alternative to overcome the problem of missing observations and avoids the decision about an appropriate ϵ correction. Although it does not provide a solution to the problem of the unknown S, simulations from the Tokeshi models showed that misspecification seems to be even less of

a problem when evenness profiles are evaluated based on predicted values from the fitted model.

In addition to the problem of missing species, differences in detection probabilities between individuals that are actually observed lead to bias in the diversity assessment if they are not accounted for. This affects evenness measures in particular, as it is estimation of the variation in the species abundance distribution which is prone to bias when measures are calculated solely from the species proportions in a sample. Although addressable with appropriate survey design, correcting for individual detectability has received less attention than species being missed entirely in the sampling process. We saw that the direct connection between the divergence of the (true) species abundance distribution $\mathbf{p}$ from evenness given by the power-divergence family $I_{\mathbf{p}}(\lambda)$ and the family of goodness-of-fit statistics that is defined for a sample $\mathbf{n}$ no longer holds, if both are corrected for detectability. Instead the estimators $\hat{I}_{\mathbf{n}}(\lambda)$ are now a weighted version of $I_{\mathbf{p}}(\lambda)$.

3.5 Chapter summary

This chapter highlighted detectability issues of both species and individuals and how they affect diversity assessment. Individual detectability can be handled with appropriate survey design and statistical methods and where this is possible, diversity estimates should be based on abundance estimates taking into account detection probability rather than observed counts to avoid bias. The discussion brought up subtle, but important differences between the sample-based goodness-of-fit measures and the power divergences which are calculated based on the species proportions, both in terms of undetected species and variation in detection of individuals. While both approaches are sensitive to zeros in species abundances, the divergences only register them in a translation and are hence closer to traditional diversity index families. Quantifying evenness through a goodness-of-fit statistic is more complex in the way zeros are treated; it takes into account the commonness and the rarity of the species that is missing as well as that of the other species. Thus it can be more appropriate where we think that the contribution of a species, even unobserved in a sample, should not be neglected. If this is the case, we would also choose the detectability adjusted version of the family of goodness-of-fit measures over the divergence family. By taking into account variation in detection probabilities we lose the direct connection between the goodness-of-fit statistics and the divergences and hence cannot take the former to estimate the latter anymore.

Chapter 4

Fine-tuning the assessment of large-scale temporal trends in biodiversity using the example of British breeding birds

This chapter has been prepared for publication in Studeny *et al.* (In prep) and is currently under revision for Journal of Applied Ecology.

4.1 Introduction

We are faced with an unprecedented decline in biodiversity at a time when the pressure on the Earth's ecosystems is growing (Butchart *et al.*, 2010). International responses to the Convention on Biological Diversity (CBD) demand large-scale assessments of biodiversity trends (de Heer *et al.*, 2005; Pereira & Cooper, 2006; Jones, 2011) rather than the local-scale evaluations that have typically been pursued in the past. As such, data must be representative on at least a national scale. These data in turn require analyses that will minimise potential bias, extract long-term trends, and determine whether the rate of biodiversity loss has been reduced (Magurran *et al.*, 2010; Magurran & Dornelas, 2010; Buckland *et al.*, In prep).

Diversity assessment is not straightforward as 'biological diversity' is a concept with a wide meaning in ecology (Harper & Hawksworth, 1995; Gaston & Spicer, 2004; Magurran, 2004). This is reflected in the CBD's biodiversity targets which are primarily directed towards ecosystem and regional diversity. However since diversity is typically quantified at the level of species (Magur-

ran, 2004; Maurer & McGill, 2011), groups of species must be chosen that are representative of the state of a nation's ecosystem. Common breeding birds are seen as good indicators of ecosystem health (Gregory *et al.*, 2003, 2005; Gregory & van Strien, 2010): they are widespread across different habitat types and extensive monitoring programmes exist in many countries.

The UK Breeding Bird Survey (BBS) is an annual survey that has been running since 1994 (Riseley *et al.*, 2011). Adopting a nationwide randomised stratified sampling design, it meets the requirements of the CBD's biodiversity plan and provides high quality data for statistical analysis which is tailored towards national monitoring. The UK's Wild Bird Indicator, which integrates data from the BBS and other UK bird surveys, has been adopted by the UK government to form part of a group of headline indicators for sustainability (Gregory *et al.*, 2005; DEFRA, 2011). It is based on a geometric mean of relative abundances of species, i.e. it summarises within-species trends in abundance relative to each species' abundance in a baseline year (see chapter 1.3.1). It also reflects the evenness component of species diversity (Buckland *et al.*, 2011b). The geometric mean has several advantages over more traditional measures of species diversity like Shannon's or Simpson's indices (Buckland *et al.*, 2011b), and is therefore preferred as a headline index (Buckland *et al.*, 2005; Lamb *et al.*, 2009; van Strien *et al.*, 2012; Renwick *et al.*, 2012). However, the summary it provides can conceal finer details in diversity trends (as any scalar measure does).

Recently, a family of indices based on goodness-of-fit statistics has been introduced in an ecological context as a diversity measure (Studeny *et al.* (2011), and see chapter 2). Similar methodology has been applied in economics to assess inequality of wealth (Cowell, 1980). As derived in chapter 2, this family is a generalisation of traditional evenness measures based on Shannon's and Simpson's diversity index. A free parameter controls the sensitivity of the measures in this family towards either rare or more common species. Thus they allow us to tailor diversity assessment with respect to different degrees of rarity within the chosen group of species. This provides a tool for analysing diversity trends in detail and offers additional information on the structure of ecological communities and their changes, which would not be revealed by the geometric mean measure alone.

One source of potential bias in estimating diversity is due to the neglect of variation in detectability for individuals from different species and over time (Yoccoz *et al.* (2001), and see chapter 3). However, diversity measures are easily corrected if they are based on estimated abundances which take account

of detection probabilities (Buckland *et al.*, 2011a). This assumes that survey methods can provide information on detectability. Distance sampling along line transects, as used by the BBS, is one way of ensuring this. Other options include point transect sampling and versions of mark-recapture techniques (Borchers *et al.*, 2002).

As pressures on species are often habitat specific, the ability to quantify trends in diversity at the habitat level is essential for effective conservation management (Newson *et al.*, 2009). Farmland birds, for instance, experienced a marked decline in the UK in the 1980s (Siriwardena *et al.*, 1998). However, an overall decline does not mean that all categories of species have been equally affected. Recent studies of woodland birds, for example, have shown that specialist and generalist birds respond differently to change, the latter showing some strong increases in abundance (Vickery *et al.*, 2004; Davey *et al.*, 2012), while bird species associated with human habitats such as *Passer domesticus* (House Sparrow), *Sturnus vulgaris* (Common Starling) and *Apus apus* (Common Swift) are the focus of growing concern (Baillie *et al.*, 2010). Conservation managers therefore need to be able to pinpoint trends for both rare and common species, and to place these in the context of the overall trends seen across the different habitat groups.

In this chapter, we evaluate large-scale trends in biodiversity across five major habitat types (farmland, grassland, urban species, wetland and woodland) using UK breeding birds. Application of the goodness-of-fit based evenness measures in conjunction with a geometric mean allows us to separate trends in less common species from those in abundant species. This provides an example of how these methods can be used to obtain a robust and more informative assessment of temporal trends in diversity. Given the role birds play as indicator species, this analysis has wide relevance for ecosystem management.

4.2 Material and Methods

4.2.1 The data

We analyse data from the British Breeding Bird Survey (BBS), which has been conducted yearly since 1994. It is organised by the British Trust for Ornithology (BTO) and carried out by volunteer observers (Greenwood *et al.*, 1995; Newson *et al.*, 2005, 2008). We use data for years 1994 to 2008, except for 2001, access to many survey sites was restricted due to an outbreak of foot and mouth disease.

The BBS is based on a stratified random design where sampling units are

1 km squares. These are allocated randomly within strata which are based on regions corresponding closely to UK counties. The sampling rate in each stratum is proportional to the number of available volunteers. Observers visit their survey square twice a year, once in April or early May, and once in late May or June. Visits are a minimum of four weeks apart. In this analysis we have focussed on data from the first visit to minimise the possibility of including juvenile birds around later in the season, except for late breeding birds, such as summer migrants (see list of species in Appendix G). In their allocated square, the volunteers walk two parallel transect lines of 1 km each while recording every bird seen or heard and assigning the observed bird to one of four categories ($0 - 25$ m from the line, $25 - 100$ m, > 100 m, and flying over). In accordance with Newson *et al.* (2008), we consider data from the first two categories only.

We began by considering the entire suite of bird species recorded through this scheme with the exception of nocturnal species as they are not well covered by the survey design. However, since the geometric mean cannot be calculated if a species' index of abundance equals zero, and precision of estimates is often poor if rarely-recorded taxa are included, very rare species were excluded from the analysis. This constraint applies to all studies that use the geometric mean to assess diversity, including those that adopt the Living Planet Index (Loh *et al.*, 2005; Buckland *et al.*, 2005, 2011b). Thus only data for species which were sufficiently widespread (observed at more than 10 sites overall) and abundant (a minimum of 15 records on average per year) were used, to allow estimation of a year effect in detection probabilities. In addition, two wetland species, *Recurvirostra avosetta* (Pied Avocet) and *Limosa limosa* (Black-tailed Godwit) had to be excluded despite being classified as sufficiently common, because of zero abundance indices in several years. Occurrence of some grassland and wetland species correlates negatively with sampling effort. Both habitat groups also comprise waders for which large counts of non-breeding individuals have been reported at a small number of sites (Field & Gregory, 1999). These would not be representative in terms of average bird density. Suspecting records of flocks behind unusual high numbers of observed birds, we decided to omit these from the analysis as outliers. A total of nine records across all years were identified as such outliers and excluded, along with a square falling in the Abbotsbury Swannery which we considered as not representative for *Cygnus olor* (Mute Swan) records.

In addition to extreme records in these two habitat groups, a few of the bootstrap resamples did not contain records for some species on the original list. This leads to asymmetric and wide confidence intervals for some diversity

estimates. As we felt that those few species unduly compromised precision, potentially caused by a misfit between the survey design and their heterogeneous distribution, these were excluded from the results presented here. This only affected grassland and wetland species. These removals left a total of 98 species across all habitats (a list can be found in Appendix G, where excluded species are also listed).

4.2.2 Habitat classification

These 98 species were considered as belonging to one of six 'communities', where the community is defined by a species' main habitat here. Using habitat information recorded along with species counts, a classification method based on Jacobs' preference index (Jacobs, 1974; Newson *et al.*, 2008) was applied to assign each species to one community (coastal, farmland, grassland, near human habitation, wetland, woodland) according to their primary habitat use. However, we excluded coastal species from the analysis as they are not adequately surveyed by BBS methods.

4.2.3 The analysis

Diversity measures

The data analysis seeks to identify long-term trends in biodiversity of British breeding birds and to determine points in time at which the rate of change in trend changes. Diversity indices differ in the degree to which they are sensitive to more dominant or rarer species. The existing UK Wild Bird Indicator is based on a geometric mean of relative abundances (i.e. abundance of each species relative to the abundance of that species in a baseline year, see chapter 1.3.1). Based on a list of S species, we calculate the geometric mean from the estimated abundances $\hat{N}_{ij}$ for each species i in each year j as

$$\hat{G}_j = \exp\left(\frac{1}{S}\sum_{i=1}^{S}\log\frac{\hat{N}_{ij}}{\hat{N}_{i1}}\right).$$

The geometric mean meets many of the requirements of a headline index (Loh *et al.*, 2005; Buckland *et al.*, 2005, 2011b). It summarises species-specific trends in abundance as well as evenness and gives equal weight to rare and more common species. It has been adopted by policy makers as a headline index, for example in the Living Planet Index and the aforementioned UK Wild Bird Indicator.

More traditional diversity measures are often based on species proportions N_{ij}/N_j, where $N_j = \sum_i N_{ij}$, and focus on dominance. The most prominent are Shannon's and Simpson's indices (Shannon, 1948; Simpson, 1949). While there are arguments in favour of Simpson's index for its sampling behaviour, both indices reflect largely the same properties, i.e. a combination of species richness and evenness. Based on species proportions, they remain unchanged if all species decline at the same rate (Buckland *et al.*, 2005). Shannon's index is slightly less focussed on the dominant species than Simpson's (Hill, 1973). A parametric family of goodness-of-fit based measures was proposed by Studeny *et al.* (2011) for quantifying evenness (see chapter 2). A free parameter λ in this family allows us to weight our biodiversity measure towards either rare or dominant species. It is given by

$$\hat{I}_j(\lambda) = \frac{1}{\lambda(\lambda+1)} \sum_{i=1}^{S} \frac{\hat{N}_{ij}}{\hat{N}_j} \left[\left(\frac{\hat{N}_{ij}}{\hat{N}_j/S} \right)^{\lambda} - 1 \right]. \tag{4.1}$$

As discussed in detail in chapter 2, it corresponds to a family of goodness-of-fit statistics (Read & Cressie, 1988) as well as a family of inequality measures used in economics (Cowell, 1980), and includes as special cases linear transformations of both Shannon's index and the log-version of Simpson's index for parameter values $\lambda = 0$ and $\lambda = 1$, respectively. Low values of the measures in this family correspond to high evenness and vice versa. We consider here its reciprocal $1/\hat{I}_j(\lambda)$ for easier comparison with other measures.

Like the geometric mean index, the goodness-of-fit measures for negative λ cannot be computed if an annual abundance estimate for any species is zero. Either such species must be removed from the species list or the missing abundance estimate must be replaced by a predicted (non-zero) value from a model, for example using generalized additive models (see below).

Correcting for differences in detectability

Measures based on species proportions, like the goodness-of-fit based indices, are biased when they are calculated from count data assuming equal detection probabilities independent of species identity (Yoccoz *et al.*, 2001; Buckland *et al.*, 2011a). By contrast, the geometric mean of counts is unchanged by species differences in detectability of individuals, but is biased if there is a trend in detectability with time (see details in chapter 3). Hence, where possible, detectability should be taken into account explicitly when estimating diversity. The BBS follows a basic distance sampling protocol and therefore we can apply

standard methods to estimate detection probabilities for each species (Buckland *et al.*, 2001). In standard line transect sampling, all animals on the line are assumed to be detected, and probability of detection is assumed to drop with distance from the line. This decrease can be described by a model for the detection function. For each species, we fit a half-normal model to the binomial count data corresponding to numbers of birds counted within 25 m of the line and between 25 m and 100 m of the line. To allow for trends in detection probabilities, year can be incorporated as either a continuous covariate or as a factor in the scale parameter (Marques & Buckland, 2003). For each species, we fit a detection function that is assumed independent of year, together with one in which year was a continuous covariate. For those species recorded at more than 10 sites in every year, we also fit a model with year as a factor. We then select the model with the smallest AIC. The following total UK abundance estimate for each species takes the original survey stratification into account:

$$\hat{N}_{ij} = \sum_r \frac{A_r}{m_{jr}a} \left[\sum_s \sum_k \frac{1}{\hat{P}_{ijksr}} \right]$$

where $\hat{P}_{ijksr}$ is the estimated detection probability of the kth detected bird of species i in year j at site s in region (stratum) r. Within a plot, we have two strips each of length 1 km and half-width 100 m, giving $a = 0.4$ km^2 as the survey area covered per plot, m_{jr} is the number of plots visited in year j in region r, and A_r is the size (i.e. total number of available squares) of region r.

Estimating long-term trends

The point estimates of abundance typically show variation over time as they are largely driven by short-term fluctuations. To establish underlying long-term trends, we smooth the yearly fluctuations applying a scatterplot smoother (Hastie & Tibshirani, 1990). We follow Fewster *et al.* (2000) and use generalized additive models (GAM). The GAM was fitted to mean counts for each species in each year. The mean was calculated as an average of the region (stratum) means, weighted by the size of the region (to take account of stratification). An offset term was included for the detectability conversion to UK density estimates

$$\mathbb{E}[\bar{n}_{ij}] = \exp \left(\log(a\hat{P}_{ij}) + f(\text{year}) \right), \tag{4.2}$$

where again $a = 0.4$ km^2 is the survey area covered per plot and $f(\cdot)$ is a smooth function. From the fitted values, the predicted abundance is calculated as

$$\frac{\mathbb{E}[\bar{n}_{ij}] \cdot \sum_r A_r}{a\hat{P}_{ij}}. \tag{4.3}$$

One choice of error distribution for a mean of counts would be a Gamma distribution, but it does not allow us to smooth over years where no birds were observed and the mean count is zero. Tweedie distributions provide a flexible way to model the mean-variance-relationship of over- or underdispersed data through a parameter ν (Jørgensen, 1997). The Poisson ($\nu = 1$) and Gamma ($\nu = 2$) distributions are special cases. By choosing a Tweedie error distribution with $\nu = 1.9$, we are reasonably close to a Gamma distribution, but can include years with missing observations in the smoothing procedure. The generalized additive model was fitted by thin-plate spline regression using the mgcv library in R (R core development team, 2011) and was given an upper limit of three degrees of freedom (df) where the actual df is determined by in-built cross-validation (Wood, 2006, 2008).

Quantifying precision

For a randomized survey design, precision of biodiversity measures can be quantified by a nonparametric bootstrap, using sites as the resampling unit (Fewster *et al.*, 2000). To take the original survey stratification into account, we sampled visited sites within each region with replacement to get the same number of sampled squares in each region as for the original sample. This is repeated (here 999 times), and 95% confidence limits for the annual diversity index are derived by the percentile method (Buckland, 1984). As it is based on a relative measure, precision of the geometric mean can be low (Buckland *et al.*, 2011b). By definition, the index is unity with zero variance in the baseline year, while confidence intervals become wider and less useful over time. Precision for a subsequent year is driven by the variance in the baseline year (here 1994, the first year of the BBS survey) as well as the subsequent year. Low effort in the early years of the survey increases uncertainty for the entire time series.

Changes in temporal trends

Fewster *et al.* (2000) and Buckland *et al.* (2005) successfully used numerical evaluation of the second derivative in combination with the non-parametric bootstrap described above to determine years in which there is a change in

(a) Negative changes   (b) Positive changes
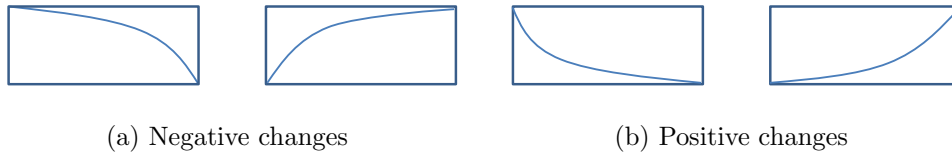
FIGURE 4.1 – Schematic plots of changes in trends. Changes can be (a) negative: increasing rate of decline or decreasing upwards trend or (b) positive: slowing rate of decline or increasing upwards trend. These changes (in slope) can be identified by looking at the 2nd derivative of the trend curves; in (a) curves are concave downwards (negative 2nd derivative), in (b) they are concave upwards (positive 2nd derivative)

the rate of change of diversity. We apply the same method to identify likely points of an accelerated or slowed loss of biodiversity for each habitat group in the BBS data set. A change for the better (either a slowed decrease or an accelerated increase in diversity) is indicated by the confidence interval for the second derivative lying entirely above zero, while a negative change (accelerated decrease or slowed increase in diversity) is reflected in an interval spanning only negative values (see Fig. 4.1). As Buckland *et al.* (2011b) note, these results are independent of the choice of baseline year for the geometric mean, and confidence interval length does not increase with increasing length of the time series.

4.3 Results

4.3.1 Overall results

Figs. 4.2 and 4.3 show trend curves for the geometric mean index and a series of members of the goodness-of-fit measures for the five habitat groups. Trends vary between habitat groups as does precision of the diversity estimates. Rare species increase uncertainty in the estimates. Hence, indices which are more sensitive towards rare species, like the geometric mean index and the goodness-of-fit measures for negative λ, show lower precision and have less power to detect trend change points in general. If species with no records in one or more survey years are not excluded when the geometric mean or the goodness-of-fit measure for negative λ is calculated, both indices show higher variation. (The smoothing ensures that zero abundance estimates do not prevent calculation of the indices, provided the smoothed estimates are all non-zero.) However, only the grassland and wetland habitat groups were adversely affected by this.
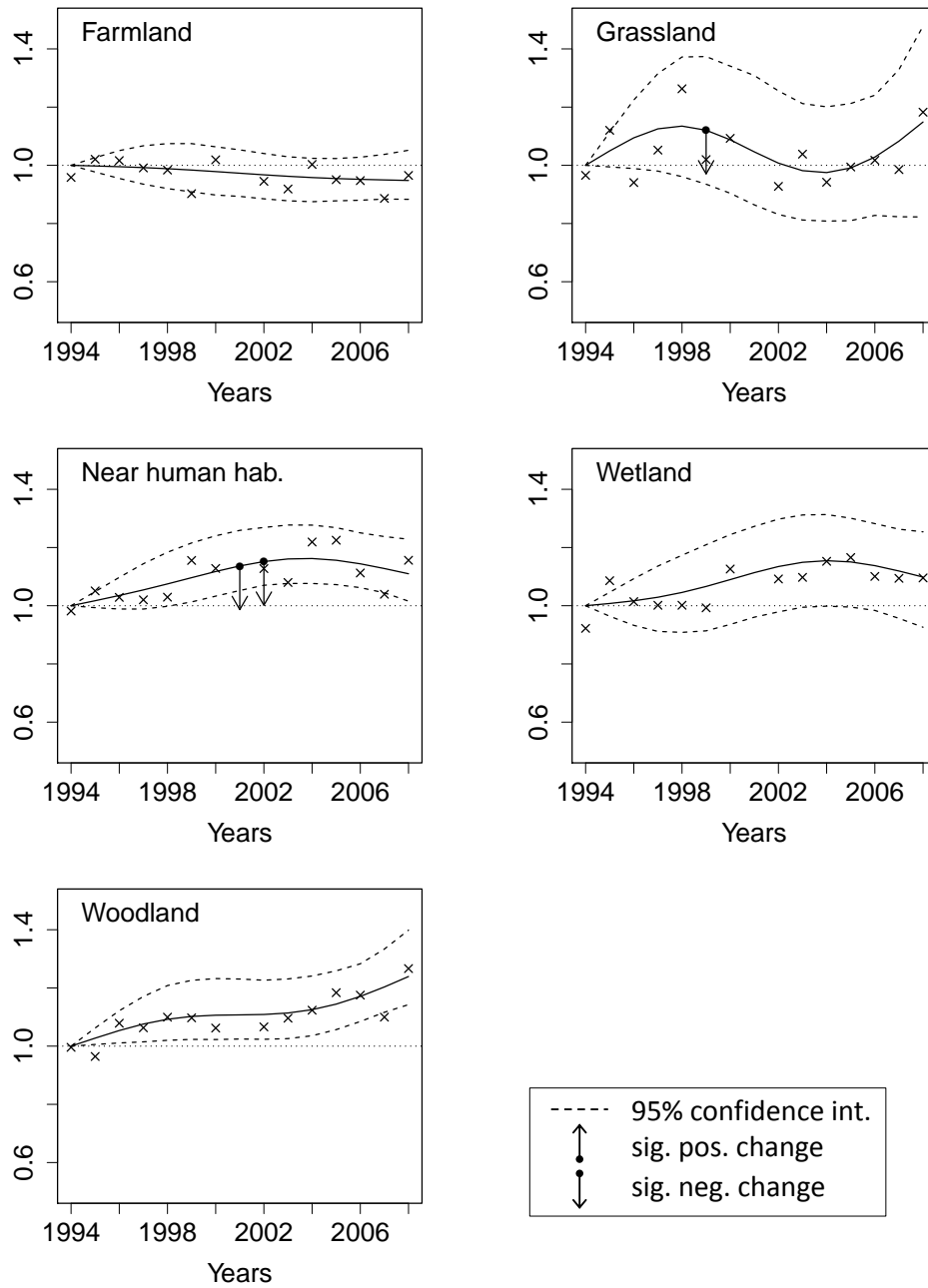
FIGURE 4.2 – Geometric mean indices of diversity for the UK breeding bird survey data 1994 - 2008. The panels show trends in diversity for five different communities of breeding birds. Birds have been assigned to one community, according to their primary habitat (farmland, grassland, near human habitation, wetland, woodland).

4.3.2   Habitat-specific trends

Farmland birds

Neither the geometric mean nor the goodness-of-fit equivalents for Shannon's or Simpson's indices give any indication of an increase or a decline in trend over the survey period. This suggests that the strong decline in the 70s and 80s has been halted, but not reversed. However, the goodness-of-fit based measures for negative parameter values show a more nuanced and less reassuring picture. The index for $\lambda = -1$ results in a continuing decline in evenness, i.e. abundances for rare species lie increasingly far below the mean abundance across species. Looking at second derivatives, there is no indication for a significant change in trend for this habitat group.

Urban species

For urban species, all indices show roughly the same pattern but to different extents. The geometric mean shows a significant increase in diversity between 1994 and the early 2000s. The second derivative suggests a change for the worse in 2001/02. The goodness-of-fit indices confirm this, but show a stronger upwards trend from the beginning of the survey. This suggests that the pattern of trend for urban species is primarily driven by trends in evenness rather than in abundance. As for the geometric mean, for positive parameter values of λ, the second derivative indicates that this increase slowed down between 2000 and 2003 with little change thereafter. For the more common species in this habitat group, diversity in 2008 (the last year considered here) is well above the value for 1994. The picture is again less positive if we look at negative parameter values for the goodness-of-fit measures and hence less common species. The initial upward trend is less pronounced, while this trend is slowed significantly around 2002 reversing the trend and leaving diversity just below the level of 1994 in 2008. This provides further clarification of why the geometric mean, which gives equal weight to rare and common species, shows less of an upward trend in the first half of the period.

Woodland birds

The geometric mean indicates an increase in diversity over the time period, with the value in 2008 being about 25% ([14%-40%]) higher than at the beginning of the survey period. No significant change in the second derivative is evident in any year. However, the goodness-of-fit based measures give little indication for

a positive trend. Positive changes in trend are picked up for the more dominant species (positive parameter values) in 2002 and for the following four years in the case of the transformed Shannon's and Simpson's indices. However, taking the confidence intervals into consideration, there is little change. While there is no significant change in the second derivative for negative λ, the trend curve suggests a decline in diversity for the less common species.

Grassland birds

Neither the geometric mean index nor the goodness-of-fit based measures indicate a substantial change in diversity for this group over the time period. Potentially due to low sampling effort at the beginning of the survey, precision for the geometric mean is low and trend estimates hence of little use. There is an indication for a change for the worse in the second derivative in 1999. The goodness-of-fit based measures draw a similar picture. For positive λ, the second derivative indicates a negative change around 1998. For the more common species, further positive changes are picked up: the first in 2002 reversing the previous negative direction, and a second in 2005 ($\lambda = 2$), indicating an improvement in the rate of change of diversity. Nevertheless, the confidence intervals for the diversity estimates do not suggest a significant change in diversity between 1994 and 2008. For less common species only, there is a hint of a positive trend in the goodness-of-fit measure, although we cannot be confident of this conclusion given the wide confidence interval.

Wetland birds

Estimated precision for all diversity indices is low for this group, in particular for those that are sensitive to rare species (geometric mean index, goodness-of-fit measure for negative λ). The proportion of rare species is high in this group, as well as species which are inhomogeneously distributed across the UK. In addition, as for the grassland group, precision for the geometric mean is likely to be affected by the lower sampling effort at the beginning of the survey. Neither the geometric mean nor the goodness-of-fit measures indicate a significant trend. There are no significant changes registered by the second derivative for any of the indices.
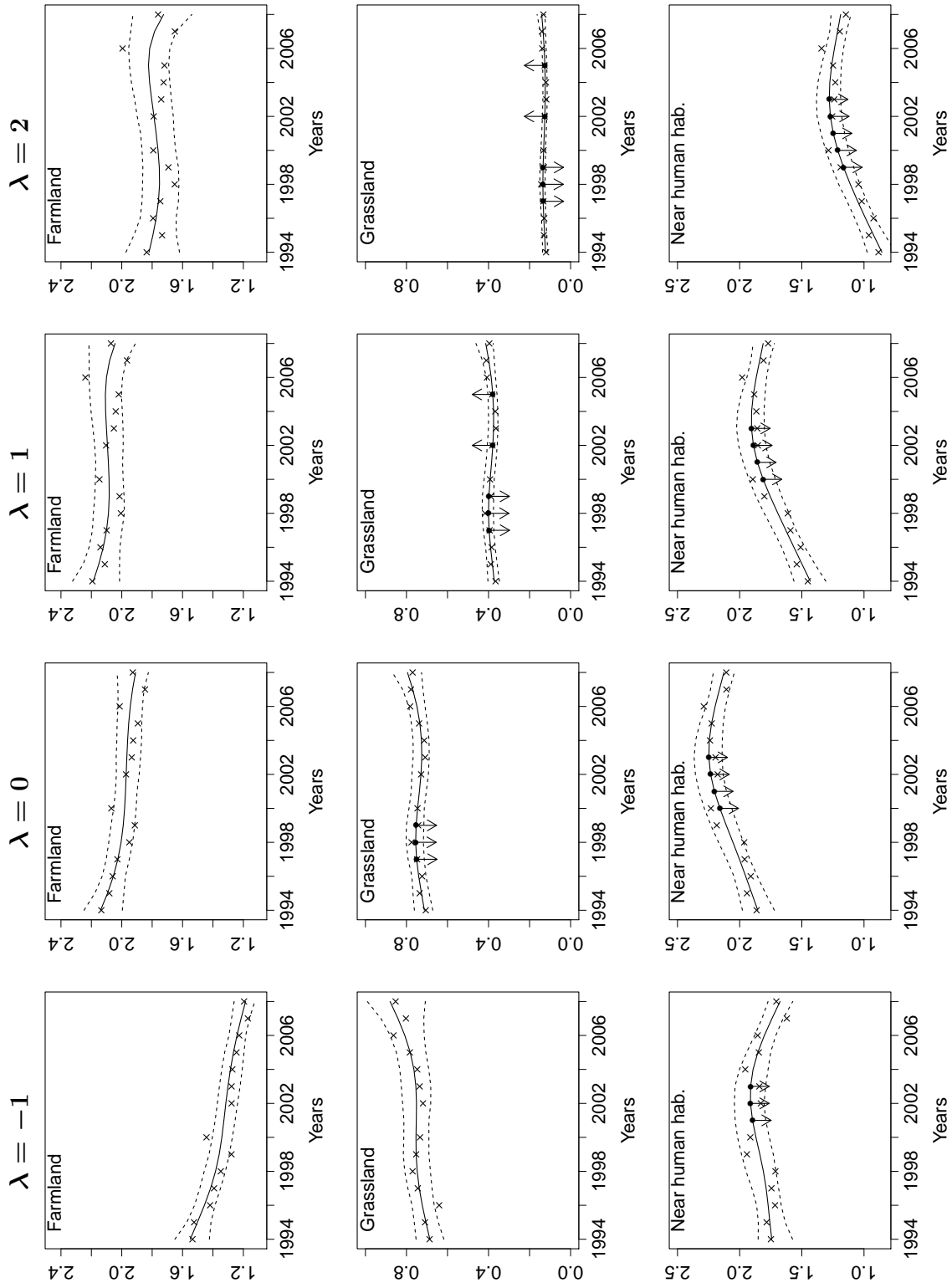
FIGURE 4.3 – Goodness-of-fit based diversity measures for the UK breeding bird survey data 1994 - 2008. [Figure continues on next page]
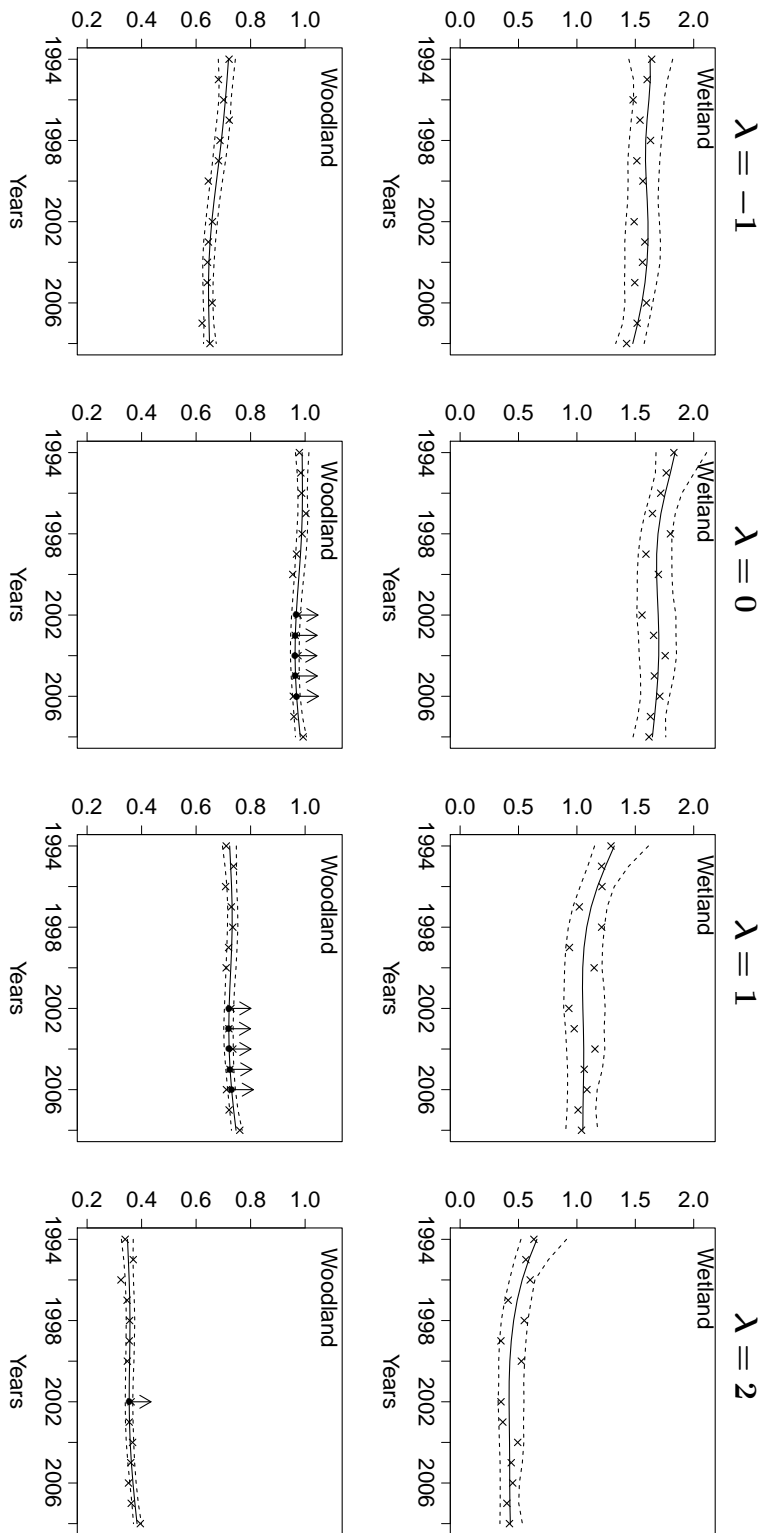
FIGURE 4.3 – Goodness-of-fit based diversity measures for the UK breeding bird survey data 1994 - 2008. The panels show the reciprocal $1/I_{\hat{\mathbf{p}}}^t(\lambda)$ where species proportions are based on abundance estimates $\hat{p}_i^t = \hat{N}_i^t/\hat{N}^t$. Varying λ changes the sensitivity of the measure towards rare (negative λ) and common species (positive λ). Changes in the rate of change (2$^{\text{nd}}$ derivatives) are determined based on the untransformed $I_{\hat{\mathbf{p}}}^t(\lambda)$. [Legend is the same as in Fig. 4.2]

4.4 Discussion

National and international biodiversity conservation plans rely on monitoring programmes as well as methods to determine large-scale biodiversity trends and assess whether the rate of loss of biodiversity is successfully slowed down, halted or even reversed (Dobson, 2005; Walpole *et al.*, 2009).

Here, we present a comprehensive study of diversity trends in British breeding birds classified by their primary habitat use. As indicator species of ecosystem and environmental health, birds are currently contributing to the national headline index to evaluate progress with regards to the international biodiversity targets. This headline index is based on a geometric mean of relative species abundances. We complemented the geometric mean index by measures that are based on goodness-of-fit statistics and that generalise classic evenness indices. The danger of a single headline index is that it concentrates on either selective aspects of the biodiversity concept or on the most abundant species. Importantly, the choice of index can have a strong influence on whether and even what kind of trend (positive or negative) is identified. The ability of the goodness-of-fit based index family to separate effects for rare and common species sheds light on why different indices pick up contrasting trends.

In addition, a positive change in the geometric mean might not be representative of all species. When a parametric measure is chosen which allows us to shift focus between rare and dominant species, trends for rare species can be looked at explicitly as long as the species are still common enough to be included in an analysis. This latter group appears to be the 'losers' in terms of diversity trends. Positive trends in our results are mostly associated with the more abundant species. The UK Wild Bird Indicator, confirmed by the geometric mean we calculated, suggests that the negative trends could have stabilised for farmland birds and even have reversed for woodland species. However, looking at the goodness-of-fit index for $\lambda = -1$, we see that the negative trend in fact continued for the less common species in both groups. This can be confirmed by single species trends in abundance for some of these birds (see individual trend curves in Appendix G). This indicates a weakness in monitoring programmes; general surveys cover the most abundant species well and are supplemented by single species surveys for the rare, endangered species (not included here) that are the focus of conservation plans. Those species that fall between the two categories (e.g. willow tit, wood warbler) are not monitored well by any scheme. This is worrying and emphasizes the necessity for a more nuanced assessment.

However, less common species pose a problem from a statistical point of view.

Their low numbers result in fewer data, and greater short-term fluctuations in abundance estimates. Uncertainty in the estimated detection probabilities for such species is also high. As diversity indices are summary statistics across a set of species, rarer species introduce uncertainty to the diversity estimates if they are included. The geometric mean, which gives equal weight to rare and common species, is especially affected, while traditional indices concentrate on the most dominant species and hence show higher precision. The goodness-of-fit measures of diversity clearly show this — confidence intervals tend to be wider for the negative parameter range. Hence there is a trade-off between the inclusion of as many species as possible and precision of the diversity estimates. Results could potentially be improved by conducting targeted surveys for some of the rare species and combining them with the general results from surveys such as the BBS, as is already done for a number of species, especially raptors and very rare species. However, the design for such surveys has to be carefully considered for the results to be included in a statistical analysis. Particularly rare species might at the same time not be homogeneously distributed. On the other hand, data collected in pristine locations might assure sampling success but will not be representative on a larger scale.

In most ecosystems the majority of species are rare while there are only few highly abundant species (Rabinowitz *et al.*, 1986; McGill *et al.*, 2007). But in terms of ecosystem function, the contribution of rare species is debated (Lawton, 1994; Lyons & Schwartz, 2001; Smith & Knapp, 2003). Yet there are examples of rare species being crucial in maintaining ecosystem functions (Lyons & Schwartz, 2001). In light of this, monitoring schemes and differentiated diversity assessment should pay attention to rare and less common species.

In this study, we consider long-term diversity trends based on abundance estimates for the whole of the UK (γ-diversity). However the time period that is covered by the BBS is not yet very long. Substantial changes in trends are not expected to occur over the course of 14 years. Nevertheless, our results already reveal important issues for large-scale monitoring. As biodiversity schemes are to be extended over the coming years according to the CBD's action plan and coordinated on an international level, this study provides relevant information at an early stage. The modelling approach presented here could be extended to identify 'hotspots' of biodiversity change, both spatially and temporally. For example, similar to the GAMs considered here, a modelling approach could be used to smooth across both space and time and thus provide predicted abundance for all grid squares across the UK. Based on this, the total predicted abundance and hence (γ-)diversity can be calculated at different spatial scales.

If, however, biodiversity indices are determined at the 1 km square level and then averaged at the chosen spatial resolution, we get a measure of α-diversity. Comparing these γ- and α-diversity estimates, we can potentially identify areas where the amount of spatial turnover (β-diversity) is changing through time. Habitat and climate information could also be usefully incorporated into such a model to identify how these covariates affect the spatial and temporal turnover in biodiversity. For instance, using a related approach, Davey *et al.* (2012) found increasing homogenization in breeding birds linked to the warming climate in Britain.

4.5   Chapter summary

This chapter investigated temporal trends in species diversity on a large spatial scale (across the UK). At the same time, this analysis illustrated in application what has been discussed in previous chapters, in particular the goodness-of-fit based measures.

In order to reliably assess diversity trends at this scale, appropriate and sufficient data is needed. Here, we analysed data from the British Breeding Bird Survey; birds are one of the few taxa for which national monitoring schemes are established according to a randomised survey design. This guarantees representative data at the national level. We examined trends separately for five major bird communities, which were defined by their primary habitat use (farmland, grassland, near human habitation, wetland, woodland).

The geometric mean of relative abundances of breeding birds underpins a headline index currently used to monitor biodiversity, ecosystem health and sustainable practices (UK Wild Bird Indicator). Diversity measures based on goodness-of-fit statistics, which have been introduced in chapter 2 offer a novel way of separating trends in dominant species from those in rarer ones. This makes them an ecologically informative complement to a headline index. In this chapter, we estimated diversity trends using both a geometric mean and goodness-of-fit based measures.

Bias from variation in detectability between individuals from different species was reduced by using abundance estimates instead of recorded counts of birds. We applied a scatterplot smoother to point estimates of abundances to separate long-term trends from short-term fluctuations and determine significant changes in diversity trends.

The results show that diversity trends vary amongst habitat types. However, the nuances in trends within the different habitats, particularly with regard to

changes in evenness, are masked when the geometric mean is adopted as the sole measure of diversity. Analysing diversity using goodness-of-fit based measures, highlights differences in trends between common and rare species. In particular, it shows that species that are scarce, but not yet rare enough to be the focus of conservation action, may be the 'losers' in diversity action plans.

Chapter 5

Spatial variation in diversity and species turnover across a region - British breeding birds (part II)

Up to now, this thesis has looked at diversity assessment in a largely aspatial way. While the previous chapter showed an approach to diversity across an entire region instead of single, selected sites, this was done by considering what is traditionally referred to as γ-diversity, i.e. diversity is assessed across all sample locations. It does not take into account spatial variability in diversity. Here, we are concerned with the concepts of α- and β-diversity (Whittaker, 1972; Cody, 1975) and we investigate how they can be interpreted in continuous space. The character of this chapter is different from the other parts of this thesis in that it does not aim for a comprehensive analysis; instead we set out to demonstrate and discuss first steps towards diversity assessment in space which will hopefully prepare the ground for future development of more advanced methods.

As discussed in chapter 1, the traditional view of partitioning diversity into α-, β- and γ-components (Whittaker, 1972; Cody, 1975; Lande, 1996; Jost, 2007) is not readily transferred to a regional setting where we are interested in making inference on diversity across the entire space instead of establishing site-specific diversity and variability among a finite (typically small) number of observed sites only.

α-diversity, i.e. the average site-specific diversity, may in theory be easily calculated for every observed location and could then be smoothed across space to extrapolate onto the whole region. However, in a large scale survey single sites will generally not provide enough data on every species for reliable diversity estimation, even for species that are easily monitored. In section 5.3, we use

a model-based approach to estimate α-diversity which overcomes this issue by replacing observations with expected counts from a model.

Based on the standard partitioning, spatial variation in diversity is typically thought of as β-diversity, the *between-sites* diversity or difference in species composition (Tuomisto, 2010; Jost *et al.*, 2011). However, extending the concept of β-diversity across a whole region is even less evident than for α-diversity (see discussion in chapter 1). Along with looking at the usual definition of β-diversity by either an additive or multiplicative decomposition of γ-diversity, we will discuss potential ways of accounting for β-diversity more adequately in continuous space in section 5.5. Similarity measures can identify locations with similar species composition (Jost *et al.*, 2011); however, calculating similarity between many pairs of sites quickly becomes computationally challenging for large regions.

5.1   Motivation

In the previous chapter, we saw an example of large-scale biodiversity monitoring; in particular, how changes in temporal trends can be assessed. Based on total UK abundance estimates for breeding birds, we were able to estimate γ-diversity. Our approach took into account potential variation in detection probabilities across species and, by grouping birds according to their habitat preferences, we could analyse habitat-specific trends in diversity.

While this (or a similar) kind of assessment is essential when we seek to monitor changes in diversity and assess progress made towards its conservation, large-scale monitoring should ideally go further (Buckland *et al.*, In prep). Besides the global picture of γ-diversity, it is also informative to estimate diversity 'locally' and how it changes across space (Ter Steege *et al.*, 2003); and ultimately, how this spatially explicit diversity changes over time (Magurran *et al.*, 2010; Magurran & Dornelas, 2010). Traditionally, this is captured by α-diversity: in a site-specific survey of several plots, α-diversity represents the average diversity of a plot. Generalising this to regional monitoring, the aim is to determine diversity at each survey site and to use this information to derive expected α-diversity throughout the whole region. On a large-scale, only a small part of all possible locations is visited during a survey. In addition, stochastic fluctuations at single sites are likely to be substantial. Hence, similar to temporal trends, estimating diversity trends locally and across space should not be based directly on the observed counts (Gotelli *et al.*, 2010; Buckland *et al.*, 2011a). Instead, we adopt a model-based approach here to derive pre-

dicted density of each species and calculate diversity indices based on the model predictions. As with γ-diversity, efforts should be made to adjust the local α-diversity estimate for detectability. However, even for well monitored and easily observed species like birds, counts are likely to be zero for a number of sites unless a species is very common and heterogeneously distributed across the whole region. As a consequence, a single site usually does not provide enough data to fit a 'local' detection function for the majority of species. Ideally, this could be integrated in a modelling approach, but appropriate methodology has yet to be developed; we do not pursue this here, but use the detection probabilities based on the data pooled across the UK derived in the previous chapter (see table in Appendix G).

5.2 The data

Once more, we look at the data from the British breeding bird survey (Riseley *et al.* (2011); for more details on the survey design see also chapter 4.2). The spatial distribution of bird species is likely to depend in large parts on the availability of suitable habitat (Gregory & Baillie, 1998; Benton *et al.*, 2003; Renwick *et al.*, 2012). Hence any spatial analysis should ideally include information on habitat through suitable covariates. However, data sources covering information on the local environment across the entire UK are sparse. In theory, the BBS data themselves contain habitat information collected by the observers (Riseley *et al.*, 2011; Renwick *et al.*, 2012). We found this habitat information unsuitable to provide covariates for two reasons; first, they evidently cover only visited squares and hence cannot be used for predictions for unobserved sites and second, their very detailed description makes it difficult to find an appropriate (qualitative or quantitative) summary for the whole square. Using satellite image and digital cartography, UK-wide land cover data have been collected by the Centre for Ecology and Hydrology (CEH) in connection with the Countryside Survey (Morton *et al.*, 2011); on a resolution of 1 km^2 on the ordnance survey grid, each grid square is assigned one of ten aggregated land cover categories (see Table 5.1). As this data set provides local habitat information collected on the same grid and with the same spatial resolution as the BBS, it seems a suitable choice as a covariate describing local habitat suitability.

Since the BBS has been launched, two such land cover maps have been compiled by CEH, in 2000 and 2007, respectively. Technical improvements for the 2007 land cover map assure a continuous spatial coverage, while the 2000 map is based on 100 km $\times$ 100 km tiles which are not straightforward to

TABLE 5.1 – Land cover classes and their frequencies across the UK aggregated at the 1 km square level.

Land cover	%area	Land cover	%area
arable	30.5	coniferous woodland	6.1
improved grassland	26.0	coastal	2.9
mountain/heath/bog	16.2	broadleaf woodland	2.5
semi-natural grassland	10.2	saltwater	0.03
built-up areas/gardens	6.1	freshwater	0.004
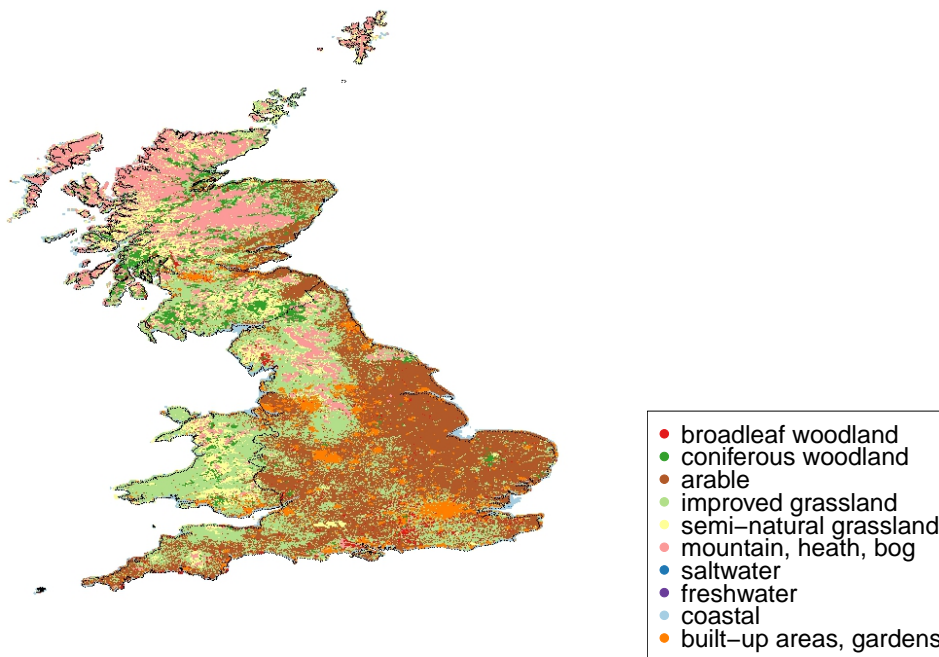
FIGURE 5.1 – Land cover map of the UK in 2007 based on data collected by the Centre for Ecology & Hydrology in connection with the Countryside Survey partnership.

combine; this can lead to inconsistencies at the boundaries of the tiles. Direct comparability between the 2000 and the 2007 map is thus limited (Morton *et al.* (2011), p.84). For this reason, we decided to only use the BBS data for 2007 in combination with the land cover map for the analysis. We further excluded all the squares falling into the land cover categories 'freshwater' and 'saltwater' because these landclasses are underrepresented in comparison to the rest (see Table 5.1). Therefore they do not provide sufficient data for reliable estimation of model parameters.

Given that the aim of this chapter is more of an outlook at what could be done than a comprehensive analysis, we further focussed on the species whose habitat has been classified as 'within or near human habitation'. Considering

computational effort, the decision to use this specific community was led partly by the fact that it consists of a comparatively small number of species while all except one of its species are sufficiently common to guarantee enough available data and avoid further deletion of species within the group (16 species classified as common, consistent with the analysis in chapter 4). Furthermore, conservational concerns for this particular community are growing (Baillie *et al.*, 2010).

5.3 The modelling approach: a hierarchical model in a Bayesian setting

We now propose a basic spatial model for the density of a species (as number of individuals per km^2) across the UK. This model is fitted to each of the 16 bird species in the 'near human habitation' category. The choice of modelling approach is directed by a recently developed fitting algorithm based on **i**ntegrated **n**ested **L**aplace **a**pproximation (INLA) (Rue & Martino, 2007; Rue *et al.*, 2009). It has been shown to be powerful, in terms of both fast and accurate computation and applicability to a wide range of spatial and non-spatial models. It is implemented in the `R`-library `R-Inla` (www.r-inla.org, Martino & Rue (2010)). Although we keep the model deliberately simple here, this restriction is solely due to the limited scope of this thesis. The methods used to fit the model come with great flexibility and can handle highly complex spatio-temporal models (Lindgren *et al.*, 2011; Simpson *et al.*, 2011a; Cameletti *et al.*, 2012). We discuss possible extensions in section 5.6 below.

Observed counts $\mathbf{y} = (y_1, \ldots, y_L)$ at locations $v_1, \ldots, v_L$[1] are modelled depending on the prevalent type of habitat given by the land cover category for the corresponding square. In addition, we expect counts to be spatially autocorrelated. This autocorrelation as well as unexplained variation in the observations are taken into account by a random spatial process $\boldsymbol{\xi}$ (a 'spatial field').

The fitting algorithm is based on a discretisation of space. More precisely, the random field is fitted to the data explicitly only in a finite number of points across space and interpolated otherwise (Lindgren *et al.*, 2011). These points are given by a triangulation (see Fig.5.2). The discrete representation of space renders computation very fast (for several reasons) while it turns out to be accurate if certain assumptions hold (Simpson *et al.*, 2011a). (Technical details are explained in more detail below.) The triangulation does not have to be

---

[1]Locations here refer to the coordinates as (easting, northing) of the lower left corner of the survey square ℓ, $\ell = 1, \ldots, L$.

FIGURE 5.2 – Triangulation of the UK, based on BBS survey squares observed in 2007. This discretisation of space is the basis of the model fitting algorithm INLA. The mesh is extended beyond the coastline to take edge effects into account.
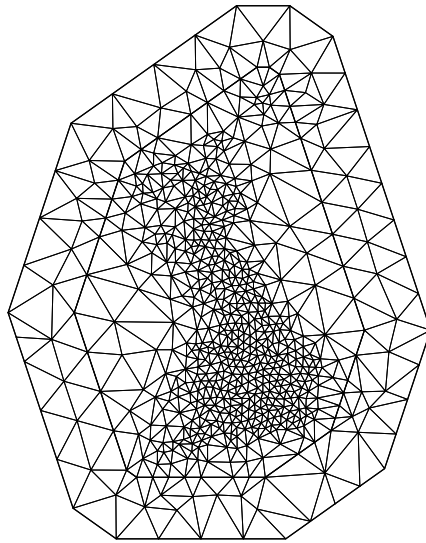
regular and thus can take into account that sampling effort differs by regions, i.e. we can place more evaluation points where we have more observations. The spatial field describes a latent, large-scale spatial trend while covariates explain local variation in the observations explicitly (Martino & Rue, 2010; Cameletti *et al.*, 2012).

We will now describe the model more formally and provide theoretical details about the fitting algorithm. In particular, the modelling is done within a Bayesian framework (Martino & Rue, 2010). The reader who is not familiar with Bayesian statistics can find some general background information in Box 5.1.

5.3.1   The model

In the following, let $V \subset \mathbb{R}^2$ denote the set containing all points on a map of the UK and let $v_1, \ldots, v_L \in V$ be all grid squares observed in 2007. We expect the abundance of a species within a square to depend on suitable habitat and hence use the land cover class ('lc') of a square as an explanatory variable. However, on top of habitat-specific variation, species densities are likely to vary in space, either due to further (unobserved) covariates or due to random fluctuations. We assume that this can be described by a stochastic process $\boldsymbol{\xi}$ in space which is also called a (latent) *random field* and which is specified by its probability distribution. In the approach taken here, the computation relies on $\boldsymbol{\xi}$ being Gaussian, i.e. a multivariate normal stochastic process in space. As a Gaussian process, $\boldsymbol{\xi}$ is fully determined by its mean and covariance matrix where the latter is given by a covariance function describing the autocorrelation structure of the

---

Box 5.1: Bayesian statistics in a nutshell

In specifying a statistical model, we make assumptions about the structure underlying the data. Usually the definition of the model includes one or more model parameters. Contrary to classical frequentist statistics, these parameter(s) are not considered fixed in a Bayesian setting, but are seen as random quantities themselves. As such they are characterised by their probability distribution.

In order to fit the model a *prior* assumption on this distribution is made; based on this and the data, the aim is to determine the *posterior* distribution

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}),$$

i.e. the probability density of $\boldsymbol{\theta}$ *given* the data. This is achieved by exploiting a basic result on conditional probabilities — Bayes' theorem — stating

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \propto \pi(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}),$$

where $\pi(\boldsymbol{\theta})$ is said to be the prior distribution and $\pi(\boldsymbol{y}|\boldsymbol{\theta})$ is the probability of observing $\boldsymbol{y}$ given $\boldsymbol{\theta}$ (this is the classical likelihood). The two sides of the equation are equal up to a normalising constant that ensures that $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ is a valid probability density function.

If several parameters are involved (as is likely in spatio-temporal settings), the joint posterior $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ is usually too complex to interpret and one is more interested in the marginal posteriors

$$\pi(\theta_j|\boldsymbol{y}) = \int \pi(\boldsymbol{\theta}|\boldsymbol{y})\mathrm{d}\boldsymbol{\theta}_{-j}.$$

The notation $\int \mathrm{d}\boldsymbol{\theta}_{-j}$ is short-hand for integration over ('integrating out') all components of $\boldsymbol{\theta}$ except the j-th. Due to the complexity of the joint posterior this integral can be derived analytically only in exceptional cases. The strength of Bayesian methods here is their powerful computational framework, commonly in the form of Markov chain Monte Carlo simulations, which allow us to derive estimates for the marginal distributions.

---

field. In general, it is not possible to define this covariance function *ad hoc*, but a parametric covariance model is chosen from a class of functions suggested in the literature (Zimmerman & Stein, 2010; Cressie & Wikle, 2011). The most commonly used covariance functions in spatial modelling belong to the *Matérn family* which, in $\mathbb{R}^2$, is given by

$$C_\nu(v, w) = \frac{1}{2^{\nu-1}\Gamma(\nu)} \left(\kappa\|v-w\|\right)^\nu K_\nu\left(\kappa\|v-w\|\right) \tag{5.1}$$

for $v \neq w \in V$, where $\nu, \kappa > 0$, Γ is the gamma function and K_ν a modified Bessel function (see for example Stein (1999), p.31). Bessel functions play a major role in physics where they are used to describe equilibrium states of fields such as electromagnetic potentials or wave propagation in cylindric or spherical coordinates (Gbur (2011), chapter 16). K_ν is exponentially decreasing (in this case with distance); it can be represented for example by an integral, but cannot be evaluated analytically. From the definition of the Matérn covariance family in (5.1), we notice that the value of $C_\nu(\cdot, \cdot)$ depends only on the (Euclidian) distance $\| \cdot \|$ between points and not their individual locations. As a consequence, $\boldsymbol{\xi}$ equipped with $C_\nu(\cdot, \cdot)$ is what is called a second-order stationary and isotropic field. Its degree of smoothness is determined by the parameter ν which is considered fixed (broadly, we can view it as analogous to the degrees of freedom of a scatterplot smoother). The scale parameter κ has to be estimated; broadly speaking, it corresponds to the range of non-zero autocorrelation.

With these preliminaries, we can now define an additive regression model for a species' density. The observed counts $y_1, \ldots, y_L$ are considered realisations from a Poisson distribution where the logarithm of its mean is given by the linear predictor

$$\boldsymbol{\lambda}(v) = \log \mathbb{E}[Y(v)] = \log(\text{offset}) + \beta_0 + \sum_{k=1}^{7} \beta_k(v)\mathbf{1}_{\{\text{lc}(v)=k\}} + \boldsymbol{\xi}(v), \qquad (5.2)$$

for $v \in V$. The offset term is the same as in model (4.2) in the previous chapter, the detection probability for an individual of species i on the square, and it converts counts into density (at location v). (We use the estimated detection probabilities derived in chapter 4 for the whole of the UK, hence the offset is constant here; this does not need to be the case in general.)

Collecting all random quantities in $\boldsymbol{\eta} = (\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\xi})$, we see that $\boldsymbol{\eta}$ is a Gaussian random field, namely

$$\boldsymbol{\eta} \sim \mathcal{N}\left(\boldsymbol{\lambda}, \sigma^2 \boldsymbol{\Sigma}\right) \qquad (5.3)$$

where the covariance is

$$\boldsymbol{\Sigma}(v, w) = \begin{cases} 1 & \text{if } \|v - w\| = 0, \\ C_\nu(v, w) & \text{otherwise.} \end{cases}$$

Note the hierachical structure of the model: We observe $\mathbf{y} = (y_1, \ldots, y_L)$ where

$$y_\ell = \exp\left(\log(\text{offset}) + \beta_0 + \sum_{k=1}^{7} \beta_k(v_\ell)\mathbf{1}_{\{\text{lc}(v_\ell)=k\}} + \boldsymbol{\xi}(v_\ell)\right)$$

as realisations of the underlying field $\boldsymbol{\eta}$, which itself is governed by the variance and scale parameter $\boldsymbol{\theta} = (\sigma^2, \kappa)$ that determine the spatial structure[2]. Hierarchical models with this kind of structure have become popular, especially in a Bayesian framework (Cressie & Wikle, 2011). Commonly, MCMC (Markov chain Monte Carlo) algorithms are used to 'update' the marginal distributions of all model parameters by repeated stochastic simulation until an equilibrium is reached ('the chain has converged') (Robert & Casella, 1999). However, in the context of spatial modelling MCMC algorithms often do poorly in terms of mixing and convergence, due to dependence between the model parameters (Rue *et al.*, 2009). While techniques have been developed to (partly) overcome these problems (Rue & Held, 2005), MCMC methods are hampered by high computational costs (Rue *et al.*, 2009; Simpson *et al.*, 2010). Recently, INLA has been proposed as a deterministic alternative to MCMC and was shown to be highly accurate while computation takes only a fraction of the time required by MCMC algorithms (Rue *et al.*, 2009; Simpson *et al.*, 2011b).

5.3.2   Model fitting with INLA

Before we discuss the restrictions of this 'magic wand', we explain the basic ideas behind the INLA algorithm. Given priors on the hyperparameters $\boldsymbol{\theta}$ and assuming a multivariate Gaussian distribution for the random field $\boldsymbol{\eta}$ given $\boldsymbol{\theta}$, our aim is to derive the posterior marginal distributions for the hyperparameters and, more importantly, all components of the random field. In mathematical terms, we look for

$$\pi(\eta_{j_1}|\mathbf{y}) = \int \pi(\eta_{j_1}|\boldsymbol{\theta}, \mathbf{y})\pi(\boldsymbol{\theta}|\mathbf{y})\mathrm{d}\boldsymbol{\theta}, \tag{5.4}$$

$$\pi(\theta_{j_2}|\mathbf{y}) = \int \pi(\boldsymbol{\theta}|\mathbf{y})\mathrm{d}\boldsymbol{\theta}_{-j_2}, \tag{5.5}$$

where $\pi(\cdot|\mathbf{y})$ is the conditional density given the observations and the subscripts j_1, j_2 refer to the components of $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$, respectively.

Hence, computational steps involved are

---

[2]These are usually referred to as 'hyperparameters', since they are not explicit parameters of the model itself, but of the spatial field.

- to derive an approximation $\tilde{\pi}(\boldsymbol{\theta}|\mathbf{y})$,

- to derive an approximation $\tilde{\pi}(\eta_{j_1}|\boldsymbol{\theta},\mathbf{y})$,

- to evaluate the integrals in (5.4) and (5.5) by replacing the conditional densities with their approximate versions and by numeric integration.

To get the first approximation, the definition of conditional probabilities gives us

$$\pi(\boldsymbol{\theta}|\mathbf{y}) = \frac{\pi(\boldsymbol{\theta},\boldsymbol{\eta}|\mathbf{y})}{\pi(\boldsymbol{\eta}|\boldsymbol{\theta},\mathbf{y})}. \tag{5.6}$$

As with any Bayesian analysis, it is first and foremost based on the application of Bayes' Theorem (see Box 5.1) which lets us calculate the joint posterior in the numerator as a product of the priors for $\boldsymbol{\theta}$, $\boldsymbol{\eta}$ and the likelihood of $\mathbf{y}$

$$\pi(\boldsymbol{\theta},\boldsymbol{\eta}|\mathbf{y}) \propto \pi(\boldsymbol{\theta})\pi(\boldsymbol{\eta}|\boldsymbol{\theta})\prod_{\ell=1}^{L}\pi(y_\ell|\boldsymbol{\eta},\boldsymbol{\theta}).$$

Replacing the denominator in (5.6) by its Laplace approximation π_G (Tierney & Kadane (1986); see Box 5.2 for details) and evaluating at the mode $\boldsymbol{\eta_0}$ of the Gaussian we get

$$\tilde{\pi}(\boldsymbol{\theta}|\mathbf{y}) \propto \left.\frac{\pi(\boldsymbol{\theta},\boldsymbol{\eta}|\mathbf{y})}{\pi_G(\boldsymbol{\eta}|\boldsymbol{\theta},\mathbf{y})}\right|_{\boldsymbol{\eta}=\boldsymbol{\eta_0}(\boldsymbol{\theta})}. \tag{5.7}$$

The approximation $\pi_G(\boldsymbol{\eta}|\boldsymbol{\theta},\mathbf{y})$ of the conditional density by a Gaussian is in general very accurate because $\boldsymbol{\eta}$ is Gaussian by assumption and the data $\mathbf{y}$ are usually 'well-behaved', i.e. conditioning on them does not lead to a radical change in distribution of $\boldsymbol{\eta}$ (Rue & Martino, 2007; Rue *et al.*, 2009). (We point out that the posterior for $\boldsymbol{\theta}$ resulting from (5.7) is generally anything but Gaussian.)

Based on this, one is tempted to derive the approximation of the marginal distribution $\tilde{\pi}(\eta_{j_1}|\boldsymbol{\theta},\mathbf{y})$ by the corresponding marginal Gaussian of $\pi_G(\boldsymbol{\eta}|\boldsymbol{\theta},\mathbf{y})$. Unfortunately, the posterior marginals are often non-symmetric and as a consequence the latter approximation is often poor; it suffers from an error in correctly locating the mode and a lack of skewness (Rue & Martino, 2007). Instead, another (simplified) Laplace approximation can be applied directly to the marginal density $\pi_G(\eta_{j_1}|\boldsymbol{\theta},\mathbf{y})$ and extracting information to correct the marginal Gaussian for location and skewness. Rue *et al.* (2009) state that this assures the correct posterior marginals for many observation models, including the Poisson which we are considering here. Finally, we can calculate the integrals in (5.4) and (5.5) by replacing them by finite (weighted) sums. (Note that if the ob-

servation model is indeed Gaussian this last step is the only approximation involved.)

In order to apply the INLA algorithm, certain assumptions have to be met that we have not explicitly mentioned so far, although two of them have been stated implicitly (Rue *et al.*, 2009). When calculating $\pi(\boldsymbol{\theta}|\mathbf{y})$ we relied on the conditional independence of the observations $y_1, \ldots, y_L$ given $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$. Crucially, for the Laplace approximation to work, $\boldsymbol{\xi}$ is assumed to be Gaussian random field, i.e. Gaussian (multivariate normal) priors have to be chosen for all components of $\boldsymbol{\xi}$ in the above. (The hyperparameters $\boldsymbol{\theta}$, which govern the spatial structure of the field here, do not underlie any distributional restrictions.)

However for the INLA algorithm to be superior in terms of computation time, the covariance matrix $\boldsymbol{\Sigma}$ is desired to be sparse[3] (rendering computational cost for the linear algebra operations cheap). More precisely, we would like the spatial field $\boldsymbol{\xi}$ to be a *Gaussian Markov random field* (GMRF) (Rue *et al.*, 2009; Simpson *et al.*, 2010, 2011a). GMRFs are characterised by certain conditional independence properties (Markov properties[4]) which can be represented in a neighbourhood structure and let the entries of the precision matrix $\mathbf{Q} = \boldsymbol{\Sigma}^{-1}$ be mostly zero[5]. The requirement on sparseness of the precision/covariance matrix is in general not given if $\boldsymbol{\Sigma}$ is defined by a complex covariance function like (5.1). This hurdle has been overcome recently by Lindgren *et al.* (2011) who exploited a direct correspondence between the Matérn covariance family and a certain stochastic partial differential equation (SPDE). Namely, a Gaussian field $\boldsymbol{\xi}$ governed by a Matérn covariance function is a solution to said SPDE. Moreover, under a discretisation of space and choosing certain (simple) basis functions, we can represent $\boldsymbol{\xi}$ by a discrete version

$$\boldsymbol{\xi}(v) = \sum_{m=1}^{M} \phi_m(v)\omega_m, \tag{5.8}$$

where M is the number of points in the discretisation, $(\phi_m)_{1 \leq m \leq M}$ the set of basis functions and $(\omega_m)_{1 \leq m \leq M}$ Gaussian distributed weights. The basis functions are chosen such that ϕ_m is equal to 1 at point m and is zero otherwise. Hence the weights ω_m actually represent the values of the field at these points (see Fig. 5.3 for illustration). The covariance matrix of the weights $\boldsymbol{\omega} = (\omega_m)_{1 \leq m \leq M}$

---

[3]This means that the precision matrix contains only a limited number of non-zero values

[4]values of the field at locations v_{ℓ_1} and v_{ℓ_2} are independent given the values at all other locations $\boldsymbol{v}_{-\ell_1\ell_2}$

[5]$Q_{v_{\ell_1},v_{\ell_2}} \neq 0 \leftrightarrow v_{\ell_1}$ is neighbour of v_{ℓ_2} (where the definition of 'neighbourhood' can be more complex than just two locations being close in space).

Box 5.2: Laplace approximation in a nutshell

This technique, first introduced by Laplace in 1774, allows us to numerically approximate functions under certain conditions by a Gaussian integral. In particular, in statistics and probability theory it is applied to derive expected values and posterior marginal distributions.

In the basic case assume a probability density function (pdf) $\pi(x)$ which is unimodal and twice differentiable and has a global maximum at the mode x_0. These properties are then inherited by the log-likelihood $L(x) = \ln \pi(x)$. We can approximate the latter by a quadratic polynomial where we match the value of the function and the first two derivates in x_0

$$L(x) \approx L(x_0) + L'(x_0)(x - x_0) + L''(x_0)(x - x_0)^2,$$

(by Taylor's theorem). As π has a global maximum at x_0, the first derivative of $L(x)$ in x_0 equals zero and hence the second term in the equation vanishes. Setting $L''(x_0) = -1/\sigma^2$, we can rewrite $\pi(x)$ based on this as

$$\pi(x) = \exp(L(x)) \approx \text{const.} \exp\left(-\frac{(x - x_0)}{2\sigma^2}\right)$$

where we recognise the density of a normal distribution $\mathcal{N}(x_0, \sigma^2)$ on the right-hand side. Thus any unimodal, twice differentiable pdf can be approximated by a Gaussian. This is useful, for example, for computational reasons. Laplace's method has been generalised to multivariate and multimodal pdfs. In particular, Tierney & Kadane (1986) developed it further for use in Bayesian analysis, reducing the error term. The figure shows the Laplace approximation (dotted line) to a beta distribution. (Obviously, if the original pdf is itself Gaussian, Laplace's method is exact.)
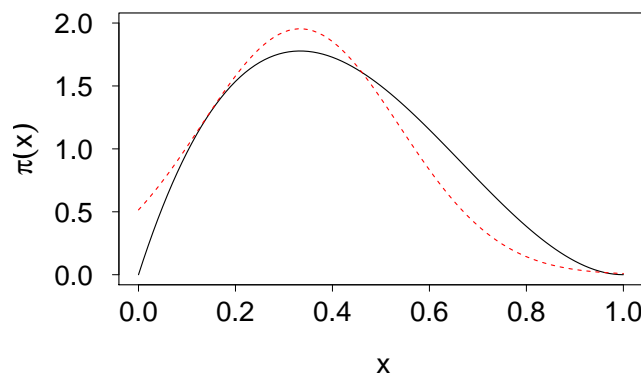
turns out to be sparse and hence they define a GMRF (Lindgren *et al.*, 2011; Simpson *et al.*, 2011a). Replacing $\boldsymbol{\xi}$ by its discrete equivalent $\boldsymbol{\omega}$ then enables us to combine this with INLA (Cameletti *et al.*, 2012; Simpson *et al.*, 2011b). To not lose too much information, we need to assure that the resolution of the
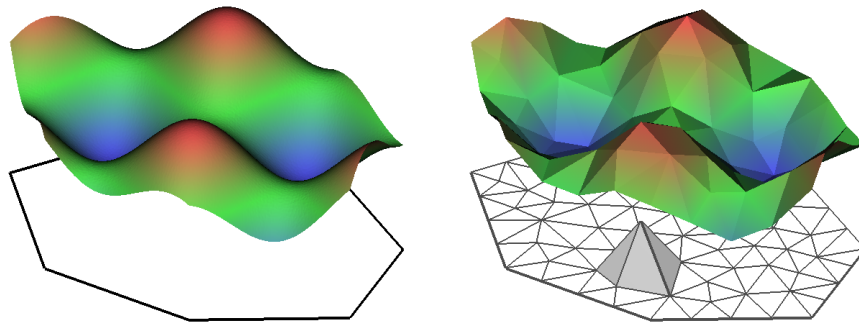
FIGURE 5.3 – Illustration of the spatial field $\boldsymbol{\xi}$ (left) and its representation by (piecewise linear) basis functions (right), as given in equation (5.8). Each basis function (example shown in grey) assumes the value of the field in a node of the triangulation and is zero otherwise. [Plot taken from Cameletti *et al.* (2012), with kind permission of Finn Lindgren.]

triangulation is fine enough to represent the underlying spatial structure. As INLA is fast, this can be done by gradually refining the resolution.

The triangulation as well as calculations combining the SPDE approach with INLA have been implemented in R and can be used within the `R-Inla` library.

5.4 First results: α-diversity maps for the British human habitation bird community

The model (5.2) was fitted independently to each species whose primary habitat has been classified as 'near human habitation' (16 species in total, see table in Appendix G). Based on the fitted model, density was estimated for each 1 km square on the ordnance survey grid. Densities are smoothed across space by the spatial field, which is continuous. As an example, the predicted density is shown for house sparrow (*Passer domesticus*) and mistle thrush (*Turdus viscivorus*) in Fig. 5.4 together with the posterior mean and posterior standard deviation for the spatial field. As expected, standard deviation is higher for the rare species, i.e. mistle thrush here. The spatial field accounts for trend *not* explained by the fixed covariate (here, land cover class of the square).

Posterior means and standard deviations for the field parameters for each species are given in Table 5.2; κ reflects the strength of spatial autocorrelation and tends to be higher for less abundant species. Plotting the posterior mean field, a more or less pronounced gradient between North and South is recognisable, except for jackdaw, spotted flycatcher and feral pigeon. (For the latter two, the prior had to be adjusted in order to achieve convergence of the model
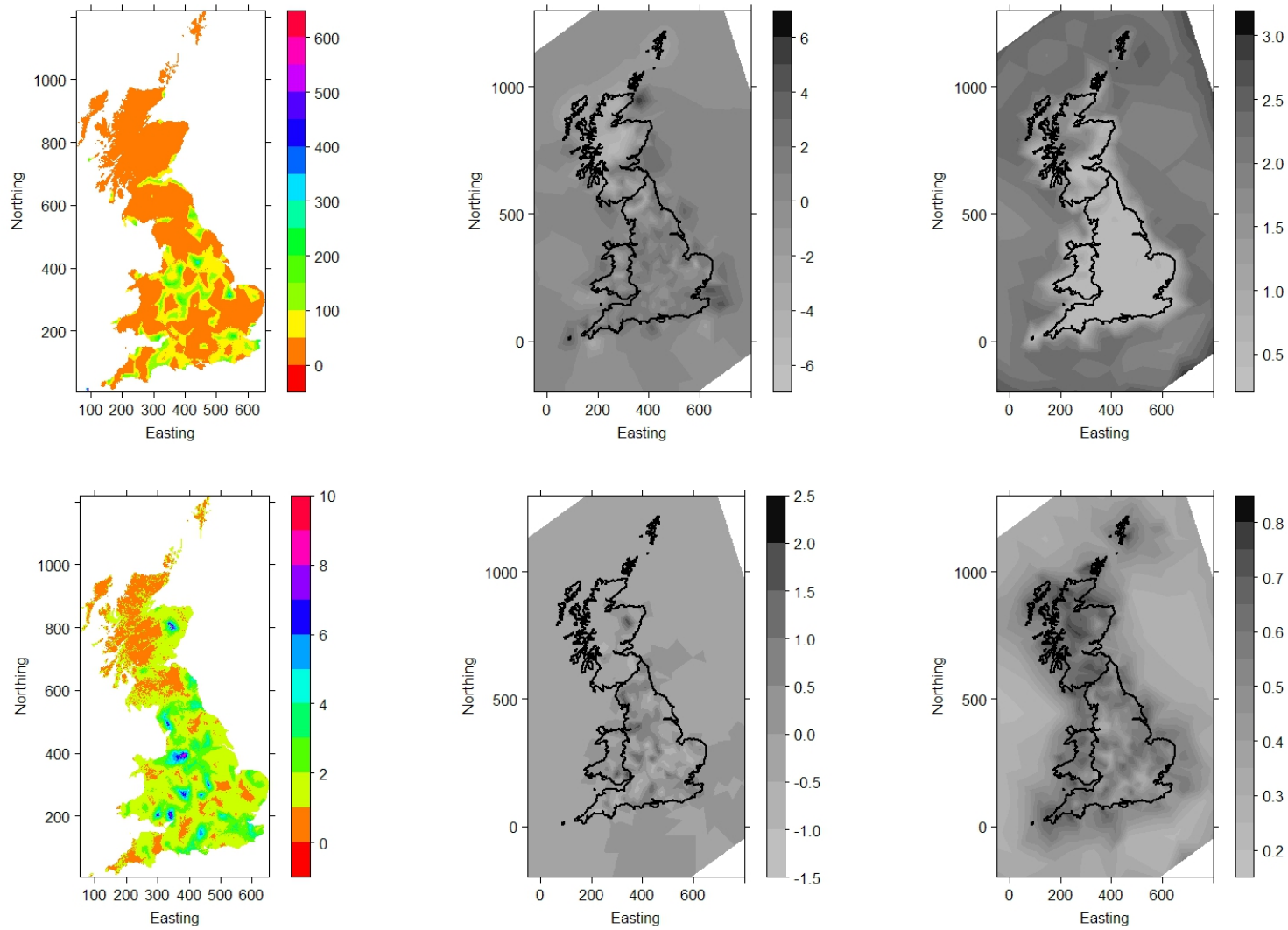
FIGURE 5.4 – Estimated densities (in colour) of house sparrow (upper row) and mistle thrush (lower row), as well as the posterior mean field (middle) and posterior standard deviation (left).

TABLE 5.2 – Posterior summary statistics of the parameters κ, σ^2 of the spatial field. Mean and standard deviation (sd) are given along with the 2.5%, 50% (median) and 97.5% quantiles of the posterior distribution. (By default, INLA returns logarithms of field parameters.)

species	mean	sd	2.5%	50%	97.5%
$\log \kappa$					
blackbird	-3.57	0.12	-3.82	-3.57	-3.33
collared dove	-3.35	0.14	-3.63	-3.34	-3.06
dunnock	-3.83	0.17	-4.18	-3.83	-3.50
feral pigeon	-3.10	0.16	-3.42	-3.10	-2.80
goldfinch	-3.36	0.17	-3.70	-3.36	-3.02
greenfinch	-3.52	0.14	-3.81	-3.52	-3.24
house martin	-1.87	0.04	-1.94	-1.87	-1.79
house sparrow	-3.45	0.13	-3.71	-3.45	-3.20
jackdaw	-3.49	0.16	-3.81	-3.48	-3.18
magpie	-3.98	0.20	-4.40	-3.97	-3.61
mistle thrush	-2.27	0.28	-2.78	-2.28	-1.68
pied wagtail	-2.78	0.22	-3.19	-2.78	-2.34
spotted flycatcher	-2.57	0.28	-3.08	-2.59	-1.99
starling	-3.74	0.16	-4.06	-3.73	-3.45
swallow	-2.83	0.15	-3.11	-2.83	-2.53
swift	-1.63	0.28	-2.12	-1.66	-1.01
$\log \sigma^2$					
blackbird	0.70	0.18	0.35	0.70	1.06
collared dove	1.39	0.18	1.02	1.39	1.75
dunnock	0.65	0.26	0.14	0.64	1.17
feral pigeon	4.12	0.18	3.77	4.12	4.47
goldfinch	0.66	0.21	0.24	0.66	1.08
greenfinch	0.98	0.20	0.59	0.98	1.39
house martin	2.62	0.10	2.41	2.62	2.81
house sparrow	1.73	0.18	1.37	1.72	2.09
jackdaw	1.81	0.22	1.38	1.81	2.26
magpie	1.64	0.33	1.03	1.63	2.32
mistle thrush	-0.25	0.27	-0.76	-0.26	0.29
pied wagtail	-0.42	0.16	-0.74	-0.42	-0.11
spotted flycatcher	0.59	0.25	0.10	0.59	1.09
starling	2.86	0.23	2.41	2.85	3.34
swallow	0.66	0.12	0.42	0.66	0.90
swift	3.24	0.39	2.57	3.21	4.10
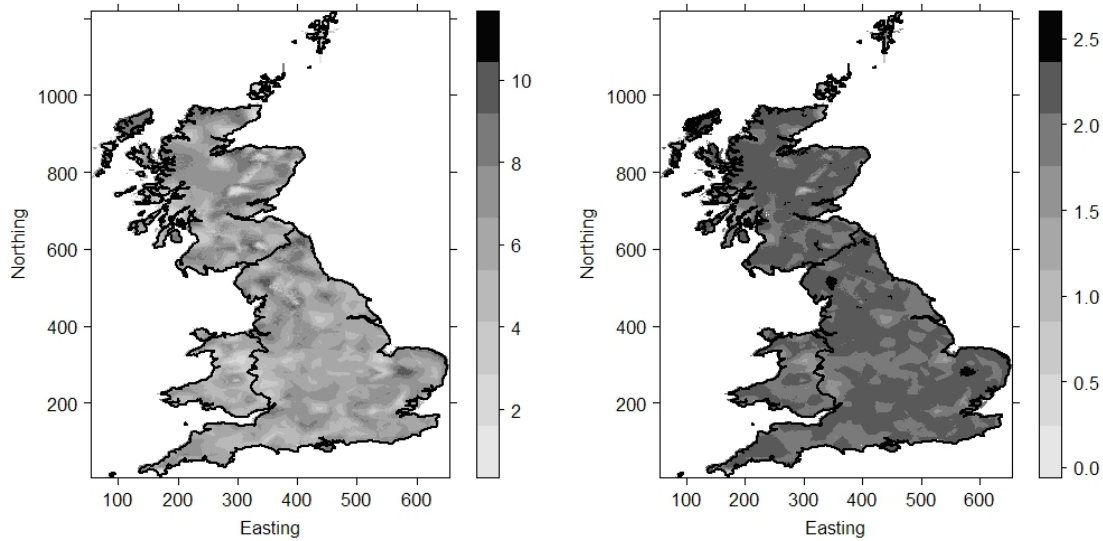
FIGURE 5.5 – α-diversity maps of UK breeding birds whose natural habitat is near human habitation. The maps show the inverse of Simpson's index $1/D$ and Shannon's index H on a resolution of 1 km squares. There is no apparent spatial trend. ($1-D$ was calculated too, with similar results which are not shown.)

fitting algorithm, hence results have to be treated with caution.) The spatial field reflects variation left after taking into account the land cover (with more mountainous, less populated areas in the North). Posterior summary statistics for coefficients of the land cover categories are presented in Table 5.3. The North–South gradient suggests that there is some remaining structure which is not explained by the covariates.

Based on the estimated single species densities we calculate α-diversity on the resolution of the grid (diversity of each 1 km square). As one would expect, Simpson's and Shannon's indices give similar results (see Fig. 5.5). No distinct spatial trend can be identified. Topographic effects are only faintly recognisable.

We also calculated a geometric mean across space where here relative abundance of a species is its abundance at a grid location relative to that in a reference square. The latter was chosen (arbitrarily) in the Southwest corner of mainland Britain (easting '134' and northing '24' on the map used here, see Fig. 5.6). Contrary to Simpson's and Shannon's index, the spatial geometric mean clearly displays a trend with numbers decreasing in areas of higher altitude (mountainous areas). This demonstrates once more the weakness of the classical indices to pick up trends (Buckland *et al.*, 2011b). While abundance within each species varies across space (leading to a North-South gradient in the geometric mean), species proportions can stay roughly the same if change in abundance is the same across species. As a consequence, neither Shannon's

TABLE 5.3 – Posterior means for the different land cover coefficients β for each species with posterior standard deviation given in brackets.

Species	Intercept β_0	Coniferous woodland	Arable	Improved grassland	Semi-natural grassland	Mountain, heath, bog	Coastal	Built-up areas, gardens
blackbird	2.93(0.27)	-0.01(0.06)	0.08(0.03)	0.07(0.03)	0.02(0.05)	0.00(0.06)	-0.19(0.07)	0.14(0.04)
collared dove	0.89(0.35)	0.3(0.15)	0.12(0.08)	0.00(0.08)	-0.56(0.15)	-0.45(0.15)	0.01(0.13)	0.15(0.08)
dunnock	2.36(0.34)	-0.32(0.11)	-0.12(0.06)	-0.19(0.06)	-0.44(0.10)	-0.32(0.10)	-0.13(0.10)	0(0.06)
feral pigeon	-3.91(1.25)	-0.71(0.20)	-0.76(0.08)	-0.77(0.08)	-1.43(0.20)	-0.49(0.15)	-0.28(0.14)	-0.94(0.08)
goldfinch	1.51(0.26)	-0.09(0.16)	0.18(0.10)	0.17(0.10)	0.17(0.13)	-0.05(0.14)	0.17(0.16)	0.39(0.10)
greenfinch	1.74(0.32)	-0.18(0.12)	0.07(0.06)	-0.03(0.06)	-0.13(0.10)	0.14(0.10)	-0.07(0.11)	0.2(0.07)
house martin	-0.9(0.28)	0.58(0.27)	0.38(0.21)	0.6(0.21)	0.04(0.29)	1.01(0.26)	-0.1(0.35)	0.15(0.23)
house sparrow	2.57(0.40)	0.27(0.08)	0.26(0.05)	0.22(0.05)	-0.04(0.08)	0.26(0.07)	0.04(0.08)	0.39(0.05)
jackdaw	0.75(0.47)	0.61(0.09)	0.05(0.07)	0.24(0.07)	0.41(0.09)	-0.04(0.09)	-0.49(0.14)	0.1(0.07)
magpie	0.11(0.64)	0.17(0.15)	0.08(0.08)	-0.02(0.07)	0.02(0.13)	0.06(0.12)	0.02(0.15)	-0.03(0.08)
mistle thrush	0.35(0.17)	0.03(0.23)	0.10(0.16)	-0.07(0.16)	-0.19(0.22)	-0.41(0.21)	-0.06(0.26)	0.16(0.17)
pied wagtail	1.34(0.17)	-0.03(0.2)	-0.10(0.15)	-0.08(0.15)	-0.01(0.19)	-0.12(0.18)	-0.94(0.3)	-0.02(0.16)
spotted flycatcher	-0.84(0.43)	-0.14(0.56)	-0.27(0.42)	0.13(0.41)	0.63(0.46)	-0.40(0.50)	-0.13(0.62)	-0.31(0.50)
starling	2.46(0.86)	0.38(0.09)	0.02(0.05)	-0.11(0.05)	-0.11(0.08)	0.25(0.07)	-0.19(0.09)	0.20(0.05)
swallow	1.43(0.19)	-0.03(0.17)	-0.02(0.12)	0.04(0.12)	-0.17(0.15)	0.15(0.14)	-0.17(0.18)	0.08(0.12)
swift	-4.10(0.56)	-0.95(1.18)	1.00(0.47)	1.20(0.47)	-6.84(11.56)	0.31(0.60)	1.99(0.60)	1.75(0.46)
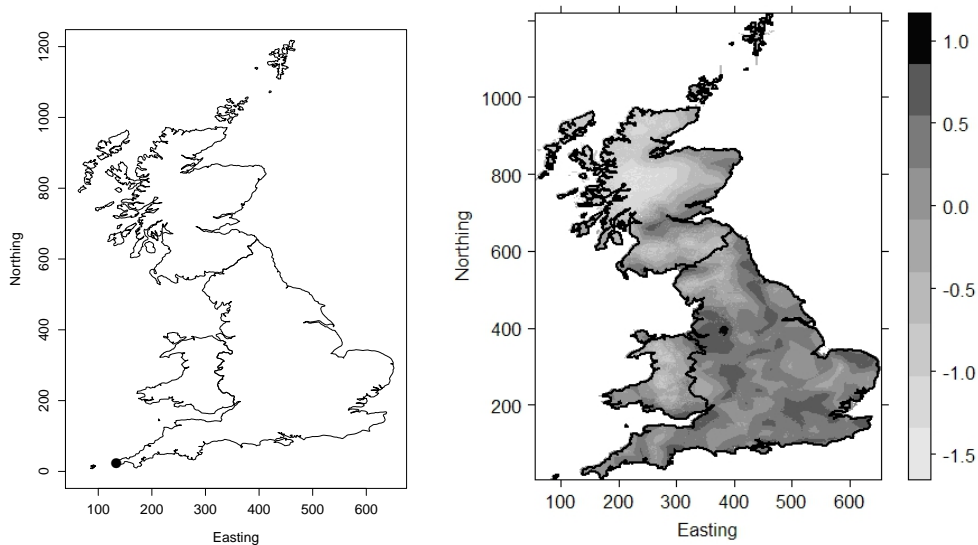
FIGURE 5.6 – Map of a spatial geometric mean index of UK breeding birds near human habitation [on a log scale]. The geometric mean has been calculated from the species abundances relative to its abundance in a base square. The location of the base square is indicated • on the outline of the UK on the left-hand side.

nor Simpson's index reveal a distinct spatial trend in diversity.

5.5 First steps towards regional β-diversity

In the previous section, we calculated local diversity (α-diversity) for the BBS data set. α-diversity is not based on individual species' identities, and hence does not account for spatial turnover in species composition. β-diversity, i.e. changes in species composition between separate sites, or along spatial and temporal gradients, has traditionally played an important part in diversity assessment (Whittaker, 1972; Vellend, 2001; Jost, 2007; Tuomisto, 2010). Especially large-scale diversity monitoring cannot neglect β-diversity as species composition is very likely to change across a country (or even larger region).

Classically there are different approaches for quantifying β-diversity (Jost *et al.* (2011); and see chapter 1): partitioning of the total (pooled) diversity (γ) into α- and β-components follows either an additive decomposition (in analogy to an analysis of variance within and between sites; Lande (1996)) or a factorisation (stressing independence between the α- and β-component; Whittaker (1972)).

These two different approaches for partitioning γ-diversity correspond to the choice of diversity index. In particular, additive partitioning relies on certain
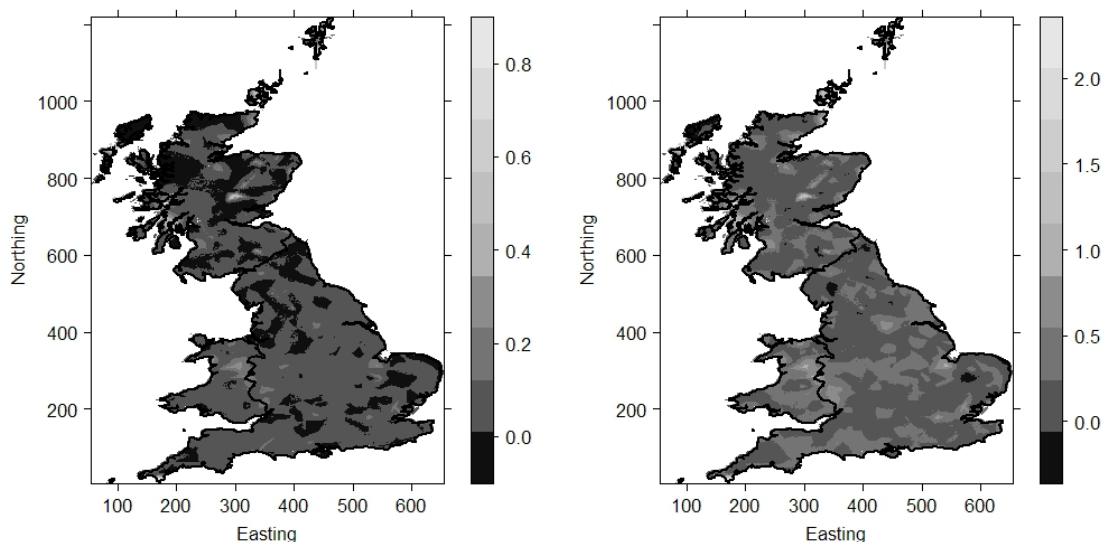
FIGURE 5.7 – β-diversity map of the UK breeding birds near human habitation, derived by an additive partitioning of γ-diversity as given by Simpson's $1-D$ (on the left) and Shannon's H (on the right).

properties of the index used to quantify diversity (concavity, Lande (1996)). The Gini-Simpson index $1-D$ as well as Shannon's index H fulfil these criteria. Hence, we can derive β-diversity maps (as shown in Fig. 5.7) from the estimated α-diversity by calculating $\gamma - \alpha$, where γ is the (constant) value of the respective index ($1 - D$ or H) from the pooled analysis of the previous chapter. The exponential of Shannon's index e^H and the inverse Simpson $1/D$ naturally go with a multiplicative decomposition (Jost, 2007, 2010) and are shown in Fig. 5.8. Neither an additive nor a multiplicative partitioning is independent of scale. Lowering spatial resolution is likely to reduce β-diversity (Mac Nally *et al.*, 2004).

The concept of partitioning diversity in α-, β- and γ-components is not applicable to the geometric mean as a relative measure; as such, it depends on a reference point. In space, the latter needs to be at the same resolution as the data. Hence it becomes meaningless as resolution increases. On the other hand, as the latter accounts for *within*-species trend across space, the map of the spatial geometric mean (Fig. 5.6) could be interpreted as a measure of β-diversity (where turnover is with respect to the base square) rather than α-diversity. It is less sensitive to spatial resolution, but not independent from the choice of the base square (in the same way as the temporal geometric mean is not independent of the baseline year, Buckland *et al.* (2005, 2011b)).

It is questionable whether extending the concept of diversity partitioning
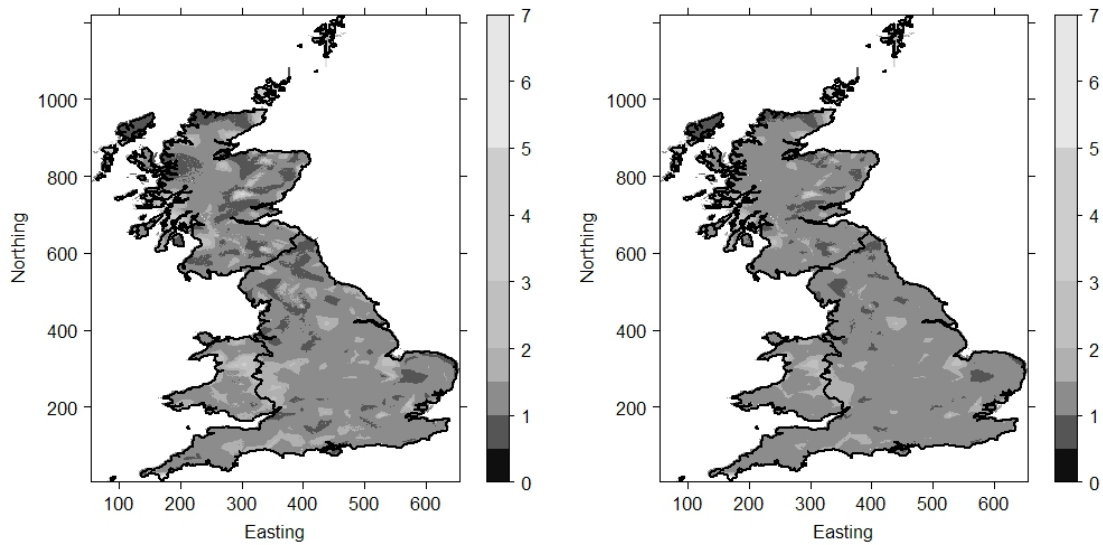
FIGURE 5.8 – β-diversity map of the UK breeding birds near human habitation, derived by a multiplicative partitioning of γ-diversity as given by the inverse Simpson $1/D$ (on the left) and the exponential of Shannon's e^H (on the right).

to continuous space in this way provides relevant information. In particular, additive decomposition is very much based on a set of discrete sites and it is not immediately clear how 'between' locations components can be interpreted in continuous space.

Independently from diversity partitioning, although often referred to as measures of β-diversity, similarity measures are used to quantity turnover between sites (Jost *et al.*, 2011). Such measures can be based purely on species' occurrences (Koleff *et al.*, 2003), but some of them can take into account species abundances and hence are not only registering change in composition, but also in the species abundance distribution (Chao *et al.*, 2006). As an attempt at an alternative way to quantify regional β-diversity, the following steps might be envisaged.

- Determine an appropriate set of n locations across the region
  This can be the original set of sampling locations or a subset. A uniform spatial coverage proportional to areas with low, medium and high α-diversity could potentially be useful. A coarser but similar triangulation to that for the model-based estimation of α-diversity could also be used here. The total number of locations will be limited by computational power.

- Quantify similarity between the chosen locations
  We derive an $n \times n$ similarity matrix by calculating a similarity measure

of choice, for example the Morisita-Horn index (Horn, 1966), for all pairs of locations based on the predicted species counts in each location (Jost *et al.*, 2011).

- Identify locations with high similarity
  Based on a distance matrix, clustering algorithms have been developed to group elements which are 'close' to each other, including visualisation techniques, such as multidimensional scaling (Everitt *et al.*, 2001; Borcard *et al.*, 2011). Using the similarity matrix as a measure of proximity, this allows us to identify locations with high similarity in species composition. In general, this results in a categorical label attached to each location. If we use multidimensional scaling (mds) and reduce the dimension down to one, we end up with an arrangement of the locations along an ordinal scale.

- Extrapolate onto the whole region
  Given the categorical label or the value resulting from mds into one-dimensional space, we need to smooth this across space. In the latter case, this could potentially be achieved assuming an underlying spatial random field as a model and using the INLA algorithm as described to fit it based on the values at the chosen set of locations. This model can only consist of the unstructured random spatial effect (i.e. the field) or contain covariates, e.g. if we think that local habitat explains local species composition. If we have categorical labels, we might be able to achieve an ordering within each category and could then follow the same procedure to get smoothed maps for each category. Alternatively, based on the clustering algorithm used, once the categories are established based on the set of chosen locations, every other location on the map could be assigned into either exactly one or several categories. Colour-coding then allows us to draw a map that shows similarities and dissimilarities in species composition across space; for more clarity this could be done separately for regions with low, medium and high α-diversity and shown in a lattice plot.

It remains to be tested whether this approach is applicable in practice. The biggest limitation is the computation of the similarity matrix, due to its dimension. If calculated between too many locations, a memory allocation problem is likely to occur (as it happened when the author of this thesis made a first attempt). However, if we choose too few locations, the extrapolation onto a

map will not be reliable.

There is a representation of ecological communities that naturally keeps track of species' identities: we can consider the vector of predicted species proportion in each spatial location as a point in the $(S-1)$-simplex (see section 2.3.4 in chapter 2). In this setting, communities with similar composition are close (in Euclidean distance). The trade-off is not only the complexity as S can be large, but also the loss of the two dimensional spatial information contained in the original location in space. From a purely mathematical point of view, the latter could be reattached to the points on the simplex, resulting in a (quantitatively) marked point pattern in $\mathbb{R}^S$ (Illian *et al.*, 2009a; Baddeley, 2010b). Unfortunately, fitting point process models is already challenging in the low dimensions of two-dimensional space, and hence applications in practice are likely to be limited; however, they might merit further investigation.

We conclude that the main problem to quantify and visualise changes in species composition across large spatial scales is the high dimensionality that comes with continuity in space. Traditional approaches tend to lose their meaning in this context, while practical alternatives are not readily available.

5.6   Discussion and future directions

Covariate data and spatio-temporal modelling

As the results shown are only a first step towards mapping diversity in space, conclusions drawn from them can be considered preliminary. Even further, we are ultimately interested in analysing spatio-temporal patterns to detect change in the different components of diversity (Magurran & Dornelas, 2010; Buckland *et al.*, In prep). The latter requires covariate data with a temporal coverage that matches the full BBS time series or at least a sufficient subset of it, which the land cover maps considered here do not provide. While we could have taken into account the map of 2000 in addition to that of 2007 and compared results for both, we decided against this. The two maps are based on different techniques of data collection and data processing and hence not directly comparable (Morton *et al.*, 2011). Based on the preliminary results presented here, it seems that land cover might not have sufficient explanatory power as a covariate on its own; the latent spatial field which accounts for variation not explained by the covariate shows a more or less pronounced North-South trend for the majority of species. This could reflect the differences in sampling efforts across the different survey regions (with substantially less coverage in the Western Scottish Highlands).

However, we would then expect this trend in the spatial field to appear in all species (unless they are so rare that even in areas with high coverage, numbers will be low). The land cover class which should account for expected counts in mountain areas to be less, does not seem to do so reliably (see estimated coefficents in Table 5.3). This needs further investigation, in particular other covariates should be included in the model to test them for their predictive power (e.g. elevation). The search for suitable covariate information becomes even more challenging if we considered spatio-temporal models. While INLA can handle this without problems (Lindgren *et al.*, 2011; Cameletti *et al.*, 2012), time series of data with adequate sample design as well as temporal and spatial coverage are sparse, if they exist at all (Magurran *et al.*, 2010; Buckland *et al.*, In prep). In the light of discussions about climate change and its effects (Thuiller *et al.*, 2005; Buisson *et al.*, 2008; Davey *et al.*, 2012), it would be of particular interest to include climate variables into a model . However, climate data are usually based on an average over at least 30 years (for example, of mean monthly temperature), and consequently require exceptionally long time series.

Model structure

Overall, a more complex model structure could be considered and compared to the results from the basic Poisson model shown here. The latent random field is also referred to as a 'structured random effect', based on terminology commonly used for mixed effect models, as it describes the spatial autocorrelation in the data. The model (5.2) could be extended by including further 'unstructured' effects (Rue *et al.*, 2009; Martino & Rue, 2010); this can be random effects, for example accounting for observer effects, or a white noise term. Given the high number of zeros in the data, a model accounting for zero-inflation is probably more adequate than a Poisson model (Zuur *et al.* (2009), chapter 11). Several options are implemented in INLA; either a zero-inflated Poisson or zero-inflated negative binomial could be chosen to model counts, where the latter can account for overdispersion along with zero-inflation. Both are further able to differentiate between two different 'types' depending on whether zeros can occur naturally at a sampling location ('true zeros') and are inflated by the 'false zeros' (type 1), or whether the latter are the only way a zero count can occur (type 0). A zero-inflated Poisson and a zero-inflated negative binomial model have been run (as type 1). With increasing complexity, run time for the INLA algorithm goes up, but still remains under 30 sec for a single species fit. Comparing the autocorrelation based on the estimated model parameters with

the theoretical one predicted by the Matérn covariance function, suggests that a negative binomial is not a good fit, but this needs verification. If several models are fitted, the INLA-library provides DIC (deviance information criterion, Spiegelhalter *et al.* (2002)) for model comparison in a Bayesian context.

Estimation of detectability

Estimating diversity throughout a region, whether locally or globally, should include information about the detectability of a species (see chapter 3). We have adjusted the estimated density for each species by an offset term in the model before calculating α-diversity; however, the offset term here was constant because we only had detection probabilities readily available which had been estimated by pooling the data across the UK and for which the detection function did not include covariate information on local habitat. Realistically, detection probability is likely to vary depending on the local environment and even conditions on the day when an observer went out (Boulinier *et al.*, 1998; Yoccoz *et al.*, 2001; Buckland *et al.*, 2011a). This should ideally be taken into account; it would be interesting to explore possibilities of modelling spatial variation in counts as well as detectability in an integrated approach. This would necessitate an extension of the hierarchical approach: a model for the detection probability of a species including a latent field accounting for spatial autocorrelation which then enters the count model. The latter can include a further latent random field along with other fixed and random effects. State-space modelling could provide a methodological framework (Buckland *et al.*, 2004b), with the true density being the underlying state and the detection probabilities governing the observations. INLA can already handle some state-space models (Ruiz-Cárdenas *et al.*, 2012), but the methods required for the outlined integrated spatial modelling approach still need to be developed.

Partitioning diversity in space

A further direction for development has already been discussed in the previous section, along with the limitations of partitioning diversity into α-, β- and γ-components in continuous space. α-diversity could have been derived in a more traditional way given the stratified design of the BBS as the average diversity of all sites visited within a stratum. β-diversity could then have been interpreted as the 'between-stratum' diversity. While differences in sampling effort between strata could probably have been adjusted by a rarefaction approach (Gotelli & Colwell, 2001), this division of space into sampling regions is an artificial

one; a different stratification is likely to change the outcome and hence these 'between-stratum' differences do not provide reliable means to make inference on compositional changes across (continuous) space. Although possible alternatives have only been sketched and remain to be studied further with respect to their applicability and usefulness, the hope is to encourage future research 'outside the trodden paths'.

New paths certainly need to be taken when we are looking at modelling and analysing biodiversity, of species and beyond, in a large-scale spatio-temporal context. This requires an integration of methods discussed in the current and the previous chapter. We saw how temporal trends in γ-diversity and their changes can be assessed on a large-spatial scale, taking into account variation in detection probability between species. In this chapter, we demonstrated the mapping of local diversity and its changes across space based on a simple model of species counts and discussed how changes in species composition, not visible in traditional α-diversity measures, could be addressed. This encounters an additional challenge in a spatio-temporal context; changes have to be followed simultaneously and points have to be identified where rates of changes in diversity as well as species turnover are accelerating or slowing, in space and time.

5.7   Chapter summary

In this chapter we have attempted first steps to diversity assessment in space. First, we mapped local species diversity (α-diversity) across space using a simple model for species density. At the same time, this also demonstrated the use of a recently developed model fitting algorithm based on integrated nested Laplace approximation. This provides means to consider more complex, spatio-temporal models; it could also be adapted to other forms of diversity. Second, we derived β-diversity estimates and discussed limitations of classical diversity partitioning when it comes to continuous space. Some suggestions towards alternative representations have been made. However, they are restricted by computational limitations due to high dimensionality of the problem: diversity is seen as a phenomenon that can be measured at any location and at any spatial scale.

Chapter 6

Spatial diversity in a 'zoom-lens': Analysing ecological communities through weighted spatial scales

Diversity in space is also the topic of this final chapter, but it will be discussed from a very different perspective. After investigating large-scale assessment of diversity across space, we are here concerned with the identification of interspecific processes that allow species to coexist and thus maintain diversity locally. Based on the assumption that interactions between species determine their spatial locations with respect to each other, the point pattern formed by the locations of individuals in space is the object of study. This is, as such, not new; Pielou (1969) already emphasised the information contained in the spatial structure of ecological communities. However, exploratory tools to analyse these spatial patterns have developed greatly and are now available in point process statistics. We introduce a generalisation of one spatial summary statistic, the cross-pair overlap distribution (Brown *et al.*, 2011), that renders it more flexible with respect to the spatial scale at which it is evaluated.

6.1   An extra dimension to diversity assessment

It has always been central to community ecology to link ecological processes to observed patterns (Watt, 1947; Bolker & Pacala, 1997; Gotelli & McCabe, 2002). With respect to biodiversity, a particular interest is in processes that shape the diversity of a community and maintain it (Pielou, 1969; Chesson,

139

2000). Several models have been suggested to describe underlying community dynamics; prominent are the discussions about neutral models which assume equivalence amongst species at the same trophic level whose coexistence is then driven by random processes (Hubbell, 2001), and niche models stressing species' adaptive responses to their local environment (Gilbert & Lechowicz, 2004). In addition, specific hypotheses, like the Janzen-Connell (Schupp, 1992) and heteromyopia effects (Murrell & Law, 2003), suggest that disadvantageous effects between conspecifics foster coexistence.

So far, we have focussed on the species abundance distribution (SAD) as the 'carrier of information' when discussing ways to quantify diversity. SADs have also been exploited as a diagnostic tool, for example to detect disturbances (Dornelas, 2010). However, they are limited in the extent of information they carry. Indeed it has been shown that they do not distinguish well between different process-based or stochastic models that describe community structure since different models may result in the same SAD (McGill *et al.*, 2007). Any of the summary statistics considered in this thesis up to this point looked at characteristics of a community which are contained in the SAD (species richness, evenness, abundance), and which are non-spatial as such. (This does not mean that we cannot look at spatial distribution or changes across space in these characteristics.) These will in the following be referred to as *first-order characteristics*.

Although every natural ecological community always is an assemblage of individuals *in space*, explicit spatial information has only recently started to be utilised to analyse community structure (McGill *et al.*, 2007; McGill, 2011). The spatial composition of a community can itself be interpreted as an aspect of its diversity: similar to the dichotomy of the non-spatial aspects 'evenness' – 'richness', spatial diversity can be characterised along the gradients 'scattering' (clustering/regularity) and 'exposure' (segregation/mingling) where high regularity and high mingling are identified with high (spatial) diversity (Pielou, 1969; Shimatani & Kubota, 2004; Rajala & Illian, 2012). These aspects are captured by what is called *second-order characteristics* in spatial point process analysis.

Ecological processes that are indistinguishable on the first-order level, might lead to different second-order characteristics. More precisely, second-order characteristics allow us to investigate the spatial positions of individuals of different species and analysing them with respect to each other. However, this can indirectly provide insight in the underlying processes, such as interindividual interactions within and between species, that determine these spatial locations

(Brown *et al.*, 2011; Rajala & Illian, 2012). In particular, hypotheses about different types of potential interactions can be tested against the null model of complete spatial randomness (CSR) (Gotelli & McGill, 2006).

Spatial point processes and spatial point pattern analysis provide powerful statistical tools to model and analyse such spatial diversity structures (Diggle, 1983; Illian *et al.*, 2009a; Baddeley, 2010a). Several second-order summary statistics exist (Diggle, 1983; Baddeley, 2010b); they are usually functions of scale and many are based on a cumulative description of the spatial pattern up to a certain scale R (usually referred to as the 'interaction radius'). A shortcoming of the latter measures is their rigidity with respect to spatial scale. Though not restricted to a specific value, the interaction radius R is typically set *before* the statistic is evaluated. After a rigorous mathematical description and a short review of second-order summary statistics in the next section, the rest of this chapter proposes a generalisation that increases the flexibility of existing measures over spatial resolution.

6.2 Spatial diversity and its assessment

Based on this discussion, our aim is to explicitly describe and analyse the spatial positions of the individuals within the community of interest (e.g. the locations of trees or animals). Given an assemblage of several species, we want to investigate the point pattern formed by the individuals' locations, while taking into account their species identities at the same time (see Fig. 6.1 for illustration). (Again, if we are considering diversity patterns across large regions, an adequate sampling design should assure representativeness of sampling sites. Note that here we want randomness of sites across space, while considering all individuals and their locations as fixed at the time of the survey.)

6.2.1 Describing multi-species assemblages in space

Following standard point process theory, this can formally be described as a multi-type (marked) point process

$$M = \{[(x_n, y_n); m(x_n, y_n)] : n \in \mathbb{N}\}. \tag{6.1}$$

Each point (x_n, y_n) refers to an individual's (random) location and has its species' affiliation attached to it. I.e. if $N = \{(x_n, y_n) : n \in \mathbf{N}\}$ is the random

(a)   (b)
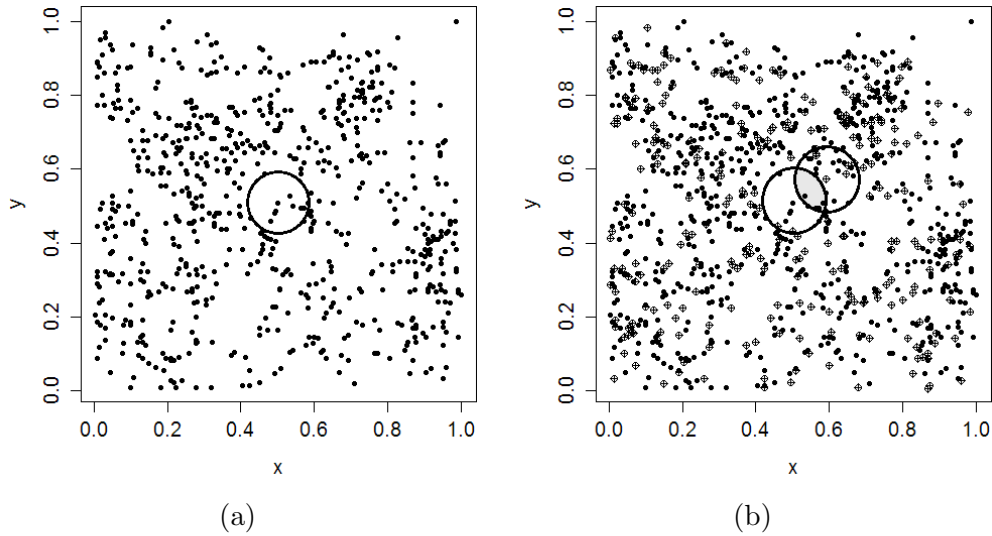
FIGURE 6.1 – The spatial pattern formed by the locations of the individuals can provide information on (a) intra- and (b) interspecific interactions. Spatial point process analysis provides summary statistics describing the degree of clumping or regularity in patterns like the ones shown based on the distances between points (compared to a random pattern); e.g. the pair correlation function g measures the probability of finding other points in a small neighbourhood around locations (indicated by the circle). It has been extended to investigate relations between patterns of different types (b) by looking at the intersection of neighbourhoods.

set of (unmarked) points on a window of unit area $(x_{\cdot}, y_{\cdot} \in [0,1])$[1], we have a mapping

$$m : N \longrightarrow \{1, \ldots, S\}, \tag{6.2}$$

which can be completely random or driven by a probability distribution taking into account associations between certain types. (For more detail on marked point processes see Illian *et al.* (2009a).)

Note that here — as above — we assume that the species catalogue and hence the number of types $m(\cdot) \in \{1, ..., S\}$ is discrete and finite. In general, the set of types can be finite or infinite and need not be discrete. (But note that the marks, although described by integers, are qualitative rather than quantitative here. Point process statistics can in general deal with both qualitative and quantitative marks, however different statistics are used.) We may consider only points of a certain type; this is denoted by $M_i = \{(x_n, y_n) : m(x_n, y_n) = i, n \in \mathbb{N}\}$. The observations in the survey area(s) can be regarded as one or several realisations of the process (6.1). The statistical analysis of such point

---

[1]Without loss of generality, we can always assume locations to be represented as points in a unit square; if necessary after appropriate rescaling

pattern is based on different summary statistics.

Here, we are interested in second-order summary statistics, which allow us to investigate relationships between points of different types (i.e. individuals of different species). Although these statistics may be defined for the general case, the point pattern is usually assumed to be stationary and isotropic. This allows us to simplify the common summary statistics to functions of only the distance between points, independent of location and is especially convenient for comparison with complete independence of types. Stationarity refers to the property that the random process giving rise to (6.1) is invariant under spatial translation (M has the same probability distribution as $M^z = \{[(x_n + z_1, y_n + z_2); m(x_n, y_n)] : n \in \mathbb{N}\}$, the point pattern that results from a translation of M by $z = (z_1, z_2)$ while the marks remain unchanged). Isotropy is defined analogously for invariance under rotations around the origin. Stationarity and isotropy might not be given, in particular under environmental heterogeneity; we will discuss limitations in section 6.6 below.

6.2.2   Analysing spatial relations within and between species

Traditional second-order summary statistics look at pairs of types and their independence from each other. One of the most commonly used and the most intuitive to interpret is the *cross-pair correlation function* $g_{ij}(r)$, which is a multi-type generalisation of the univariate pair correlation function for unmarked point processes (Illian *et al.*, 2009a; Baddeley, 2010b). Given the point pattern (6.1), the joint probability of finding a point of type i in a small circle U of area $\mathrm{d}u$ and a point of type j in a circle V of area $\mathrm{d}v$, where the distance between the centres of U and V is r, can be expressed as $\rho_{ij}(r)\mathrm{d}u\mathrm{d}v$ (ρ_{ij} is the 'product (probability) density' of the subprocesses M_i and M_j). The cross-pair correlation function is then given as

$$g_{ij}(r) \propto \frac{\rho_{ij}(r)}{\lambda_i \lambda_j} \quad \text{for } r \geq 0, \tag{6.3}$$

where λ_i is the intensity[2] of the point pattern M_i (and analogously for λ_j). Hence, g_{ij} is a standardised version of the probability given through ρ_{ij}. g_{ij} allows us to investigate if locations of species i are independent of those of species j and hence, if there is an indication of interaction between the two species; if the two subpatterns M_i and M_j are independent of each other, then

---

[2]Biologists and other applied scientists commonly call this the density of the points. However, as 'density' is the standard term in probability theory referring to the distribution, spatial statisticians use 'intensity' to avoid confusion.

their joint density $\rho_{ij} = \rho_i \rho_j \propto \lambda_i \lambda_j$ and we get

$$g_{ij}(r) \equiv 1 \quad \text{for } i \neq j^3, \tag{6.4}$$

indicating no interaction between the two species. Moreover, we can derive a direction if there is an indication for interaction (attraction vs. repulsion). More specifically, $g_{ij} > 1$ if points of type i and j are interspersed, thus suggesting positive interaction between the two species. On the other hand, $g_{ij} < 1$ indicates segregation between points and hence potential repulsion. Note that, while equation (6.4) indicates independence between M_i and M_j, it does not provide any information on the degree of randomness within either M_i or M_j (these would have to be studied through g_i and g_j, respectively).

While we concentrate on the cross-pair correlation function here, two other commonly used second-order statistics are closely related to it. The K_{ij}-function (the bivariate equivalent to Ripley's K for unmarked point processes, Ripley (1977); Baddeley (2010b)) is the expected number of points of type j in a circle of radius r where the centre is a typical point of type i scaled by the intensity λ_j. Its relationship to g_{ij} is similar to that of a cumulative probability distribution function to its density function

$$g_{ij}(r) = \frac{1}{2\pi r} \frac{\mathrm{d}}{\mathrm{d}r} K_{ij}(r) . \tag{6.5}$$

For several reasons, point process statisticians prefer to use a square root transformation of K_{ij},

$$L_{ij}(R) = \sqrt{\frac{K_{ij}(R)}{\pi}} , \tag{6.6}$$

(Illian *et al.* (2009a), p. 217). In particular, this stabilises fluctuations in K_{ij} (of both mean and variance) with increasing R. (While $g_{ij}(r)$ considers local behaviour around points that are distance r apart, K_{ij} and L_{ij} accumulate information up to scale R.)

6.2.3   A community level summary of spatial structure

An alternative bivariate cumulative second-order summary statistic, which has recently been introduced, is based on the logarithm of g_{ij},

$$A_{ij} = \int_0^R \log g_{ij}(r) \mathrm{d}r \tag{6.7}$$

---

[3] $g_{ij}(r) = g_i(r)$ for $i = j$ (where $g_i(r)$ is the univariate pair correlation function for M_i)

(Brown *et al.*, 2011). Although at first sight similar in its idea to K_{ij} because of the integral, A_{ij} has the advantage that it is equal to zero (independent of scale R) if M_i and M_j are independent (while K_{ij} increases in R). The logarithmic transformation also has a balancing effect on the values of g_{ij}, which are originally bounded from below (by zero) but not from above.

Any of these summary statistics allow us to investigate bivariate relationships between types of points. However, this is a clear limitation in any realistic ecological situation of interest. Communities are typically multivariate, and while g_{ij} (as well as K_{ij}, L_{ij}) could be applied to all possible pairs of species, this becomes quickly unmanageable to compare, in particular if we think of speciose assemblages like tropical rain forests (Condit *et al.*, 2002). Based on equation (6.7), the *cross-pair overlap distribution* (xPOD) has recently been suggested to provide a second order summary across all species in the community (Brown *et al.*, 2011). For fixed R, it considers the distribution of A_{ij} across all pairs of types in the form of a histogram. Conclusions on the presence of interactions between species can be drawn by comparing the mean and the standard deviation of the histogram to that of a process without such interactions. (The latter is determined by the chosen null model, this can be complete spatial randomness, in which case the expected mean would be zero, but other choices are possible.) Brown *et al.* (2011) show that these cross-pair overlap distributions distinguish well between simulations from a neutral model and a range of niche models.

6.3   A radius-weighted approach to spatial diversity

As mentioned in the previous section, cumulative second-order summary statistics depend on the scale R at which they are evaluated. This is evident for A_{ij} (but holds equally for K_{ij} or L_{ij}). For fixed R, A_{ij} can well equal zero, simply because opposed effects on finer scales cancel each other in the integral. As long as we restrict ourselves to only pairwise comparison between types i and j, this is not a problem. We can simply plot the value $A_{ij}(R)$ against R and hence look at all possible scales. However, this form of multi-scale evaluation is not readily transferred to the multiple species setting; we would be left with a multitude of pairwise comparisons along all scales. For speciose communities, these pairwise comparisons are neither feasible in terms of computation time, nor easily interpreted across all species.

A different interpretation of equation (6.7) enables us to derive a generalisation of the cross-pair overlap distribution that is more flexible with respect to scale. More precisely, we can identify expression (6.7) as the (rescaled) expectation of $\log g$ over the spatial scale r where equal weights are placed on $[0, R]$,

$$\frac{1}{R}\, A_{ij} = \int_0^1 \log g_{ij}(r) f(r) \mathrm{d}r, \tag{6.8}$$

and $f(r) = \frac{1}{R}\mathbf{1}_{[0,\,R]}$ is the uniform distribution. By choosing a different probability density function (pdf) f we can introduce non-uniform (and hence less rigid) weights on the scale r. Because of its natural interpretation as weights, an obvious choice if the observation window is set to $[0, 1] \times [0, 1]$ is

$$\tilde{A}_{ij} = \mathbb{E}[\log g_{ij}] = \int_0^1 \log g_{ij}(r) B(r; \alpha, \beta) \mathrm{d}r, \tag{6.9}$$

where $B(r; \alpha, \beta) = B_{\alpha,\beta}(r)$ denotes a beta distribution. The parameters of the beta distribution determine the focus on certain spatial scales (local neighbourhood, intermediate distance, far distance, or combinations of these) while considering the whole point pattern (instead of the sharp cut-off at R by the uniform distribution). Actually, the latter is included in the beta distribution as a special case — parameters $\alpha = 1$ and $\beta = 1$ correspond to a uniform distribution on $[0, 1]$.

However, any other pdf for which the expectation in (6.9) exists can be used instead. Alternative choices are the uniform distribution, which works for any scale, or any discrete probability (for which the integral becomes a sum); a truncated Normal, Gamma or more general Tweedie distributions (Jørgensen, 1997) can be used where the observation window can not be scaled to $[0, 1]$. The crucial point is that the generalisation to an expectation along weighted spatial scales (where the weights are determined by the chosen pdf) offers greater flexibility and no longer requires the evaluation radius R to be set in advance. In addition, it allows us to consider information contained in the entire observation window while the focus on a specific scale can be regulated via the parameters of the chosen distribution.

6.4   Example: comparison of simulated point patterns

To demonstrate the performance of the radius-weighted xPOD given in equation (6.9) we compare two very different, simulated point patterns on a unit square (see Fig. 6.2). In both cases, 15 different types of points were generated. For the first pattern, these were the outcome of 15 independent Poisson processes. This realisation from a random Poisson point process provides the xPOD for the usual reference point of complete spatial randomness and hence absence of any interactions on any scale. The highly structured point pattern consists of regular seed points which define the rings and random multi-type clusters around each seed point.

For the application of the weighted xPOD a beta distribution is chosen as in (6.9) where the parameters are set to (1) $\alpha_1 = 1, \beta_1 = 3$, (2) $\alpha_2 = 3, \beta_2 = 3$ and (3) $\alpha_3 = 3, \beta_3 = 1$. This corresponds to zooming in on local, intermediate, and large-scale behaviour. We expect the differences between the patterns to produce divergent xPODs at different spatial scales. In their original (unweighted) version, the xPODs of the two point patterns are indistinguishable (Fig.6.2) despite the fact that they describe very different spatial patterns, as the xPOD cannot express scale-specific behaviour. The xPOD of the structured pattern is highly sensitive to the distance between the concentric rings of point processes. For illustration, it has been chosen here to generate an xPOD similar to that under complete spatial randomness.

When the radius-weighted version is applied, the xPOD for the Poisson pattern remains virtually the same, independent of the chosen weighting – as we expect given the self-similarity of the point process across all scales. For the structured point process, on the other hand, the change in structure with scale is now clearly visible (Fig.6.3): At small radii, marks (or 'species') occur together and consistently overlap more than they would if the entire pattern was random. Hence, when weighted towards very local behaviour, the xPOD is centered around positive values. At medium radii this behaviour changes. Marks can now be wholly separated from one another by the empty areas between rings. This leads to a change in sign when the xPOD is focussed on this scale. At large radii, neighbouring rings in the pattern are encountered, and so the values in the distribution become positive again. Their range is greater than at small radii, however, as the scale is now so large that the spatial overlap of marks is partly determined by the position of points relative to the edge of the window.
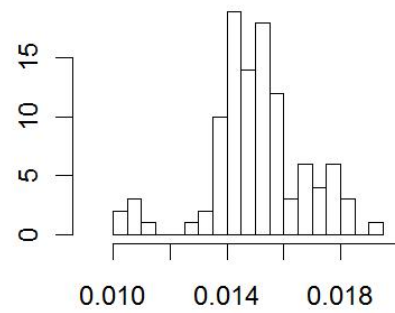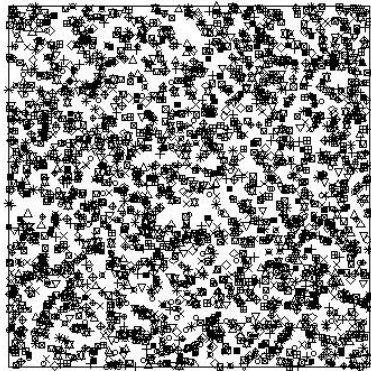
(a) Random point pattern
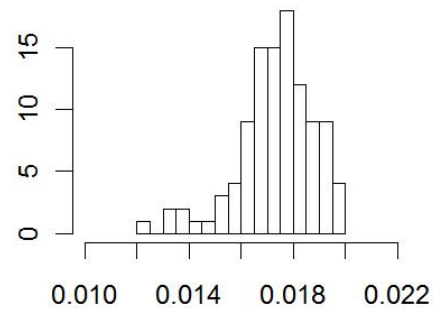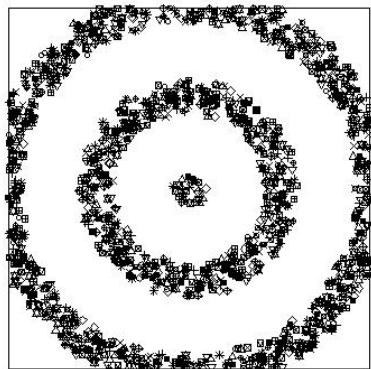
(b) Structured point pattern

FIGURE 6.2 – Two multi-type point patterns and their unweighted xPODs. The panels on the left show two simulated marked point pattern with 15 different types; (a) is a superposition of 15 individual Poisson processes, while for (b) rings were generated by seed points and multi-type clusters around them. Despite the different structure, the unweighted xPODs for the two point patterns on the left look very similar.

(a) $Beta(1,3)$ (b) $Beta(3,3)$ (c) $Beta(3,1)$
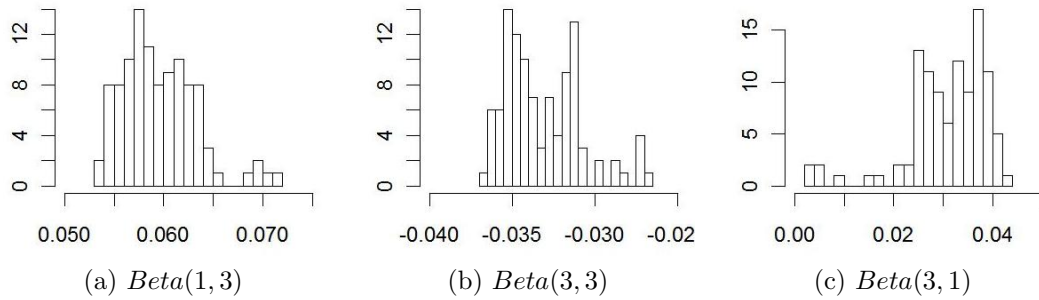
FIGURE 6.3 – Weighted xPODs for the structured point pattern. Focussing the xPOD on different scales from local (a) over intermediate (b) to large distances (c) reveals scale-specific behaviour of species' co-occurrence. (Note the different values on the x-axis.)

Our calculations are corrected for the absence of points beyond the window, and so this truly reflects the random nature of the point processes within it.

6.5   Application: a 'hotspot' of plant biodiversity

Interactions between species are likely to play on important role in maintaining a community's biodiversity (Tilman, 1994). In the following, we apply weighted xPODs to investigate a highly diverse, ancient plant community in south-western Australia. We look at a biodiversity hotspot characterised by an immense richness in species (Myers *et al.*, 2000). The coexistence of so many species is all the more astonishing because the resources in the study area are naturally poor (Armstrong, 1991; Orians & Milewski, 2007).

This community was previously studied by Illian *et al.* (2009b) who modelled the spatial point pattern taking into account 24 of its species. Crucially, their modelling approach incorporated information on the typical 'zone of influence' for the species under consideration. The range of these interaction radii varies significantly (for some examples see Illian *et al.* (2009b), Table 1). A Bayesian approach allowed Illian *et al.* (2009b) to incorporate this in the species-by-species analysis. However, it is not clear on what scale a community level summary should best be evaluated. The weighted xPOD provides means to consider all species at the same time despite their different interaction radii and 'zoom in' on certain ranges. Thus we expect to gain insight into scale dependent behaviour on the community level. In contrast to the large scale diversity patterns studied in the previous chapters, the following example is
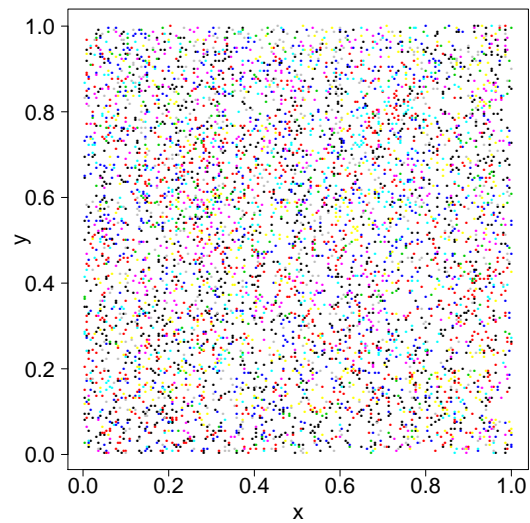
FIGURE 6.4 – Spatial locations of 18 Banksia woodland species on a plot of 22m by 22m in Western Australia.

concerned with processes acting on very small spatial scales.

6.5.1   The data

The data come from a survey site in Cataby in the South-West Australia, a region that is considered a hotspot of biodiversity (Myers *et al.*, 2000). They have been described in detail in Armstrong (1991). A full census of a 22 m by 22 m plot was carried out which, despite the relatively small survey area, revealed a total of 67 species at 6,378 individual plant locations. The majority of these species are endemic to south-western Australia. The community has formed over a substantial amount of time, with some species growing in the same location for hundreds to thousands of years. It consists of various small evergreen, shrub-like plants in low Banksia woodland that undergoes regular bushfire outbreaks (approximately every 10 years). All species have adapted to the occurrence of fires through one of two strategies: *'seeders'* are destroyed by the fire which at the same time initiates the release of seeds stored since the last bush fire (serotiny) and helps the germination, so that these plants regenerate quickly; *'sprouters'* burn down except for the plant stem which is protected by 'lignotubers', buds in the root crown preserving nutrients which enable the plant to sprout in the absence of photosynthesis. The high species richness on such a small plot is astonishing given the low levels of nutrients and water of the sandy soil characteristic for the area (Armstrong, 1991). One is inclined to assume that this leads to increased competition for the limited resources and hence inhibition between species, which is indeed the case (Richardson *et al.*, 1995). However, positive interactions can occur where soil fungi allow certain seeder species to

TABLE 6.1 – List of the 18 species considered in the analysis of the Australian plant community along with their regeneration strategy after bush fire incidence. (*) indicates a resprouter that produces some seed for regeneration, where the main form of regeneration is from the root stem.

ID	species	abundance	regeneration strategy
1	Alexgeorgea nitens	977	resprouter
2	Andersonia heterophylla	686	seeder
3	Bossieae eriocarpa	103	resprouter
4	Conospermum crassinervium	266	seeder
5	Conostylis candicans	149	resprouter
6	Dasypopgon bromeliifolius	167	resprouter
7	Eremaea asterocarpa	207	resprouter
8	Hibbertia hypericoides	148	resprouter
9	Hibbertia sp.	134	resprouter
10	Jacksonia floribunda	124	resprouter
11	Chordifex sinuosus	154	resprouter
12	Leucopogon conostephioides	657	seeder
13	Leucopogon striatus	261	seeder
14	Lomandra sp.	304	resprouter
15	Lyginia barbata	299	resprouter
16	Melaleuca scabra	377	resprouter(*)
17	Phlebocarya philifolia	207	resprouter
18	Scholtzia involucrata	170	resprouter

extract nutrients if they are in close proximity to certain sprouter species (Illian *et al.*, 2009b). Thus in particular the interactions between seeders and sprouters are of interest. Given the consistently poor soil conditions throughout and the comparatively small size of the plot, we follow Illian *et al.* (2009b) and assume homogeneity of the local environment. In consequence, any pattern detected is the result of conspecific or interspecific interactions rather than driven by heterogeneous environmental conditions. To guarantee sufficient data we consider only species that have been observed in at least 100 locations across the plot (18 species in total of which 4 are seeders). Table 6.1 gives an overview. This is a subset of the species considered previously in Illian *et al.* (2009b).

6.5.2 Identifying scale-dependent mechanisms of spatial diversity

To investigate interspecific patterns in the plant community, we evaluate weighted xPODs where a beta distribution is chosen for the weights. We inform our choice of parameters for the beta distribution by the radii given in Illian *et al.* (2009b)
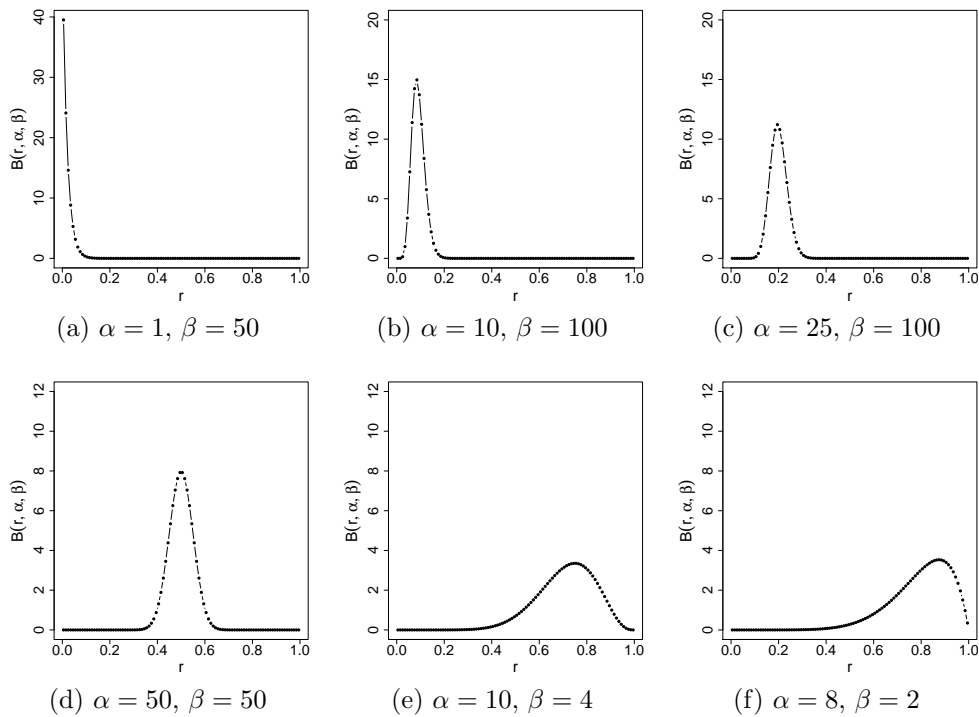
FIGURE 6.5 – Beta distributions $B(r; \alpha, \beta)$ as used for the weighted xPODs for the Banksia woodland community. The values were chosen so that they provide a nuanced evaluation on a local scale (up to 2 m, around 2 m, around 4 m) and a coarser resolution at higher spatial scales. The empirical interaction radii for all resprouting species considered here were less than 4m.

as the zone of influence for the different resprouter species. In contrast to the simulated pattern in the previous section, interspecific interactions operate on a very local scale here (with values of empirically derived interaction radii between 0.1 m and 4 m depending on species type). Based on this information, we chose the parameter values α, β for the beta distribution that allow for a spatial resolution fine enough to zoom in and differentiate spatial patterns on this local scale (see Fig. 6.5). (A coarser spatial resolution had been considered at the beginning, but was immediately recognised as not sensitive enough to the local effects.)

Contrary to Illian *et al.* (2009b) who built their model on one-directional interactions of resprouting plants to seeders, we do not make this assumption here. In fact, the xPOD as a community level summary statistic does not provide us with information on the direction of the interaction between the pairs of species. The weighted xPODs for all 18 species are shown in Fig. 6.6. There is a striking difference for the xPOD weighted for very local patterns (< 2 m), with much wider spread, very little concentration of values and a mean clearly below 0. A further effect seems to appear in the middle range while the

xPODs focussing on distances just under and around 4 m (the upper boundary of the empirical interaction radii) and those for the far distances look similar. However, the latter is less left-skewed and shows a slightly wider spread.

Comparing this with the xPODs evaluated for the resprouter (Fig. 6.8) and the seeder species (Fig. 6.9) separately, we can see that these effects are mainly due to the resprouter species. This is partly because of the low numbers of seeders in the set of species, which consequently do not contribute as much. However, it might also reflect that the resprouters, regrowing from their rootstock after a fire, have been in the same spatial location since the plant assemblage started to form.

Brown *et al.* (2011) carried out an extensive simulation study; they simulated both neutral and several niche models, including effects such as Janzen-Connell and heteromyopia. Looking at point patterns generated as the outcome of these simulations and evaluating (unweighted) xPODs for them, they investigated the ability of the xPOD to distinguish between these models. Comparing our results from the weighted xPODs with their results, the local pattern visible in the xPOD might be explained by niche or temporary niche effects. Due to the low nutrient levels, the system has evolved very slowly and over a long time in which the resprouters established stable niches. Some temporal variability might be introduced by the seeders, which regenerate from seed periodically after each bush fire incidence. Niche effects have been discussed more generally as a potential mechanism behind species coexistence in species rich areas (Tokeshi, 1996).

When the focus is on radii below 2.2 m, the mean of the xPOD is negative ($\mu = -0.15$). This indicates less overlap at very local scales than we would expect from spatial independence between the subpatterns. Hence on average, species tend to be more spatially segregated at the very local scale, which corresponds to the empirical zone of interaction for more than half the species. This suggests that competition and niche effects are the main driver at this scale. However, some species also overlap more than expected if they were spatially independent, indicating some positive interactions (Vilà & Sardans, 1999; Illian *et al.*, 2009b).

At small radii, seeder-resprouter cross-pair overlap can assume positive or negative values depending on the pairing (frequencies increase over the whole range of the xPOD when seeders are added to the resprouters). This confirms the assumption by Illian *et al.* (2009b) of a negligible impact of the seeders on the spatial overlap of the resprouters. It also is in accordance with their model-based result: interactions between species can be positive as well as negative.

(a) $\alpha = 1$, $\beta = 50$ (b) $\alpha = 10$, $\beta = 100$ (c) $\alpha = 25$, $\beta = 100$

(d) $\alpha = 50$, $\beta = 50$ (e) $\alpha = 10$, $\beta = 4$ (f) $\alpha = 8$, $\beta = 2$
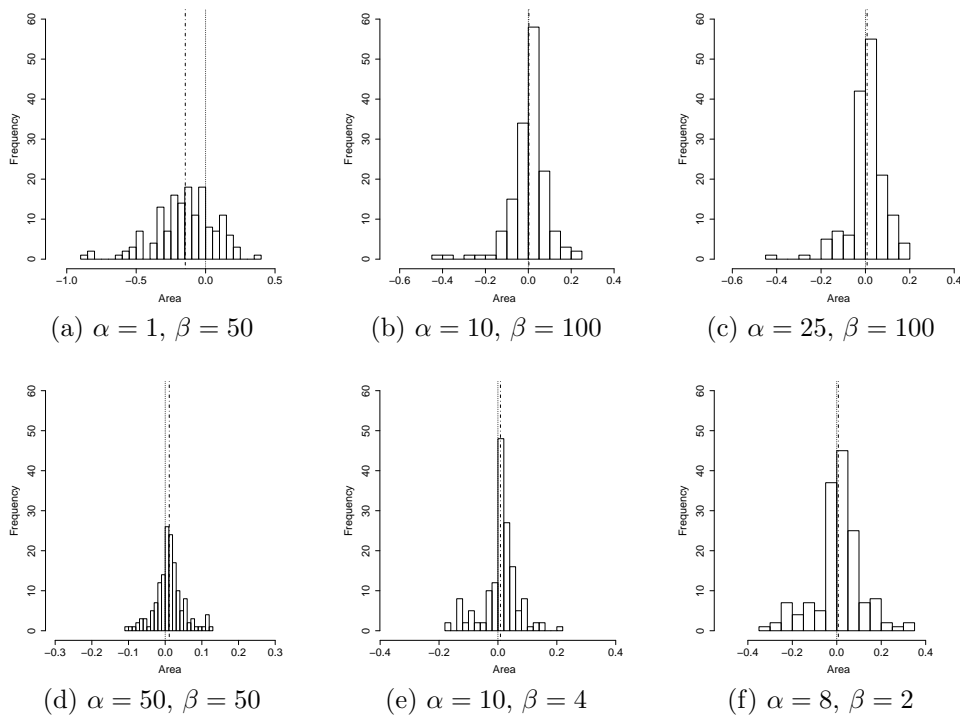
FIGURE 6.6 – Radius-weighted xPODs for 18 species of the Banksia genus on a high diversity plot in Western Australia

There are several theoretical concepts in ecology explaining coexistence of species through negative density dependence in conspecifics (Wright, 2002), such as the Janzen-Connell effect (Schupp, 1992; Wright, 2002) and heteromyopia effects (Murrell & Law, 2003). Fig. 6.7 strongly suggests that such effects occur within this plant community. Investigating average nearest neighbour distances between conspecifics as well as for pairs of species reveals that nearest neighbours tend to be of a different species rather than the same (see Fig. 6.7). This indicates conspecific competition to be greater than interspecific competition.

This might also be reflected in the mid-range effect. Based again on a comparison with simulations by Brown *et al.* (2011) for the unweighted xPODs, the reduced variance of the overlap distribution could be read as an indication of a Janzen-Connell effect, i.e. a stronger negative density effect between conspecifics than interspecific competition. It is not immediately clear why this effect should occur at this scale (around 10 m). There has been an argument that negative density dependence among more abundant species occurs at larger spatial scales (Wright, 2002), such that this pattern could be an indication of a second, 'large-scale' spatial trend. However, we have to keep in mind that the larger the scale gets the less informative the data will be because of the small area covered by the plot.

6.6 Discussion

Processes that allow species to coexist and the identification of such 'drivers' of species diversity have long been of central interest to ecologists (Watt, 1947; Chesson, 2000; Wright, 2002). Theoretical concepts such as the spatial segregation hypothesis or competition–colonisation trade-off have been conjured to explain coexistence of species despite the fact that they naturally have to compete for resources (Bolker & Pacala, 1997). Crucially, these processes can be assumed to shape the *spatial* structure of a community (Tilman, 1994). Astonishingly, statistics that are commonly in theoretical point pattern analysis as exploratory tools have not been exploited in ecology until recently to infer process from pattern (Wiegand & Moloney, 2004; Perry *et al.*, 2006; Law *et al.*, 2009).

However, processes can operate on very different spatial scales and hence summary statistics should be able to react flexibly to scale. Here, we introduce an extension to a community level spatial diversity metric, the cross-pair overlap distribution, along weighted spatial scales. The use of this generalisation is demonstrated in two different examples. Introducing weights enables us to see a clear difference in the spatial overlap for simulated point patterns that were indistinguishable for the unweighted xPOD despite their very different spatial structure. For the highly diverse Australian plant community on the other hand, application of weighted xPODs allows us to 'zoom in' on different spatial scales from local to distant neighbourhoods. As a consequence, processes operating on different spatial scales became visible. Maybe not surprisingly, niche effects appear dominant at very local scales (which is at the same time the empirically derived zones of influence for most of the species in the community). More interestingly, an additional effect is picked up at mid-distances. Here, we could only offer a preliminary interpretation. Given the coexistence of so many species in such a small area, it is not unreasonable to assume that effects fostering negative density dependence between conspecifics could offer an explanation (Wright, 2002). However, further investigation would be necessary to confirm this. Whether the mid-scale pattern is caused by an ecologically relevant process or is a purely stochastic effect, it is unlikely that it would have been uncovered by a traditional fixed scale approach and can thus be seen at least as a proof of concept.

Although other choices are possible, the beta distribution can in general be expected to work well for the weighted xPOD. It has a natural interpretation as weights, truncation or rescaling are not necessary and its expectation, and hence

the integral in equation 6.9, exist. Moreover, a very flexible shape governed by its two parameters allows the user to adapt the evaluation scale, like a 'zoom-lens', to match the scale of their data. In the example of the Australian plant community, we saw the focus of the lens, i.e. the weights, needed to be set in such a way that we could zoom in on the 'natural' (small) scale of the underlying processes, to make them visible. In particular, niche processes here act on a very local scale and only when weights are concentrated there, do they become apparent. Because of a much clearer segregation between the rings in the simulated pattern, shifting weights along a coarser scale was sufficient, and a higher resolution would not have revealed more. Hence, we recommend using some prior biological knowledge, where possible, as a base for biologically reasonable setting of the 'lens' in order to derive meaningful results. Although we did not show any kind of sensitivity analysis here, we can conclude from our experiences so far that the exact choice of parameter values for the beta distribution is not essential as long as they allow a focus on the appropriate scale. If there is no prior knowledge at all about the system under consideration, we recommend systematic variation of the parameters to inspect coarser and finer spatial resolutions.

Along with other basic spatial summary statistics, the pair correlation function and with it the xPOD are based on the assumption that the point process generating the pattern is stationary and isotropic. Because of the homogeneous environmental conditions, we could be confident of this assumption holding in the example of the Australian plant community. However, it is not likely to be the case in general. Baddeley *et al.* (2000) introduced an inhomogeneous K-function as an exploratory tool allowing spatial variation in intensity. Interestingly, their approach is also based on a weighting, in this case by the local density of points. Nevertheless, distinguishing inherent environmental conditions affecting the system from intra- and interspecific processes is difficult, if not impossible. Ecological communities are in reality shaped by the interplay of both and hence they are in general not separable. This is similar to the non-separability of trend and autocorrelation in time series analysis (Baddeley, 2010a). Similar patterns, such as clustering, can be the outcome of positive interactions, but can look identical to patches of favourable soil conditions, at least if we have only a single observation (Bartlett, 1964).

Weighted spatial scales such as have been considered here for the xPODs can in theory be applied to other cumulative spatial summary statistics. Future work could determine whether this is as useful as in the case of spatial overlap and in particular if a combination with the inhomogeneous K-function (or a

transformation of it) could provide informative results.

6.7 Chapter summary

The focus of this final chapter has been different in that it dealt with local processes maintaining community diversity rather than overall trends across large spatial scales. While observable and quantifiable in space, (species) diversity in the traditional sense does not contain spatial information itself. In contrast, this chapter built on a recently introduced concept of 'spatial diversity'. By explicitly analysing the positions of individuals of different species and their spatial structure, we learn how spatially diverse (interspersed in space) a community is. As such, this might not justify an interpretation as an additional aspect of the diversity concept; however, it is on this level that interactions between species and processes maintaining coexistence of species and hence biodiversity manifest. We introduced a generalisation of second-order point process summary statistics that allowed us to extract information on the spatial structure of a community more flexibly than with existing methods.
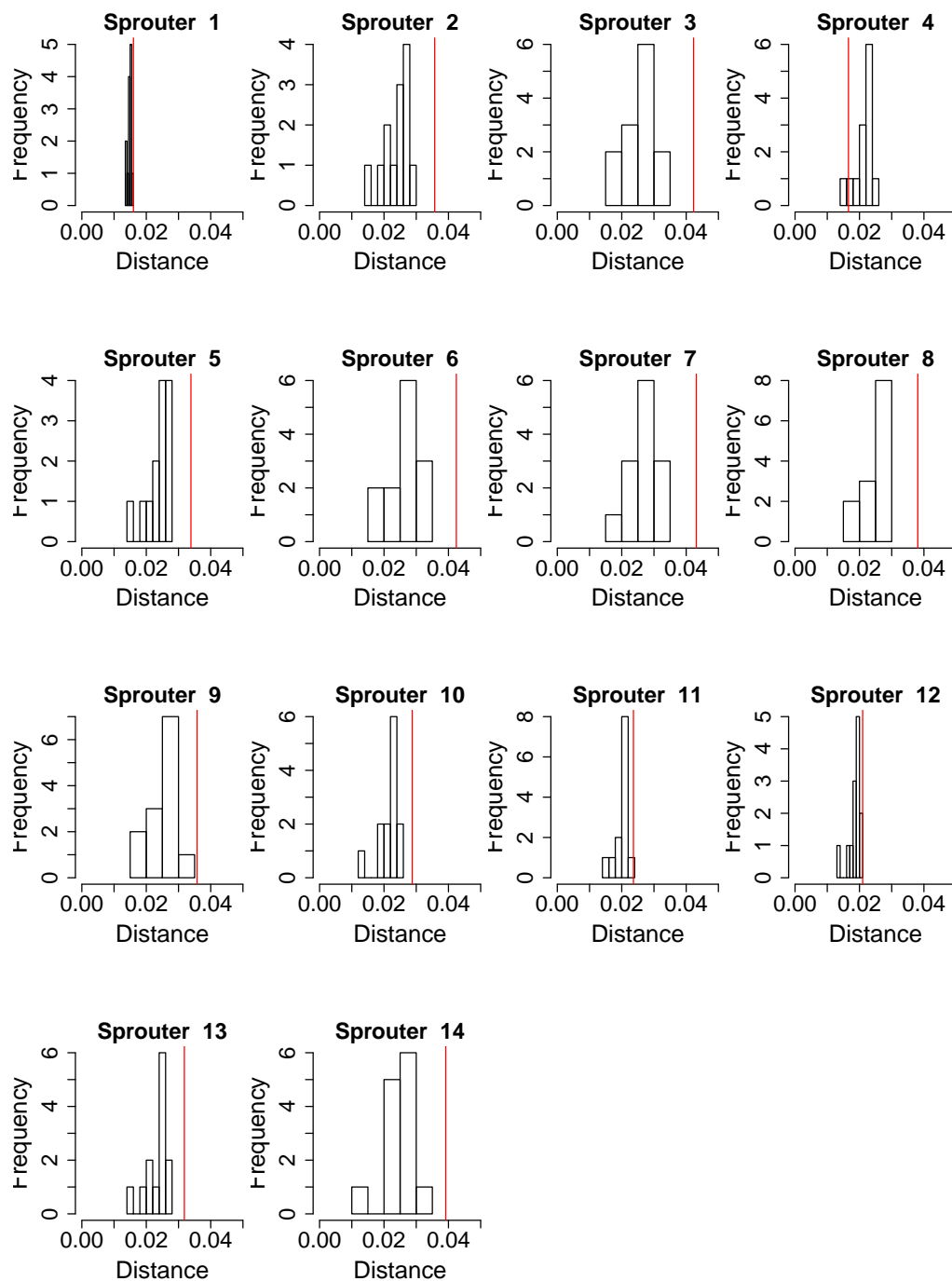
FIGURE 6.7 – Average interspecific and conspecific nearest neighbour distances for the resprouter species in the Australian plant community. For each resprouter species, the plot shows the average distance to the nearest individual of the same species (red) and a histogram for the average nearest neighbour distance of the resprouter to each of the other resprouter species.
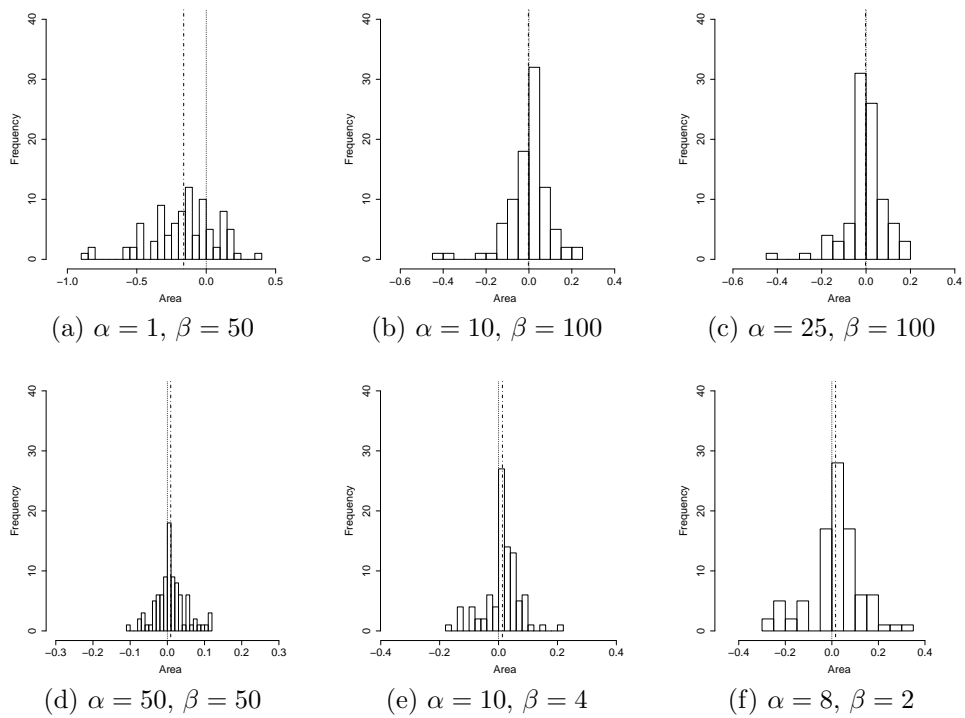
FIGURE 6.8 – Radius-weighted xPODs for 14 resprouter species of the Banksia genus on a high diversity plot in Western Australia
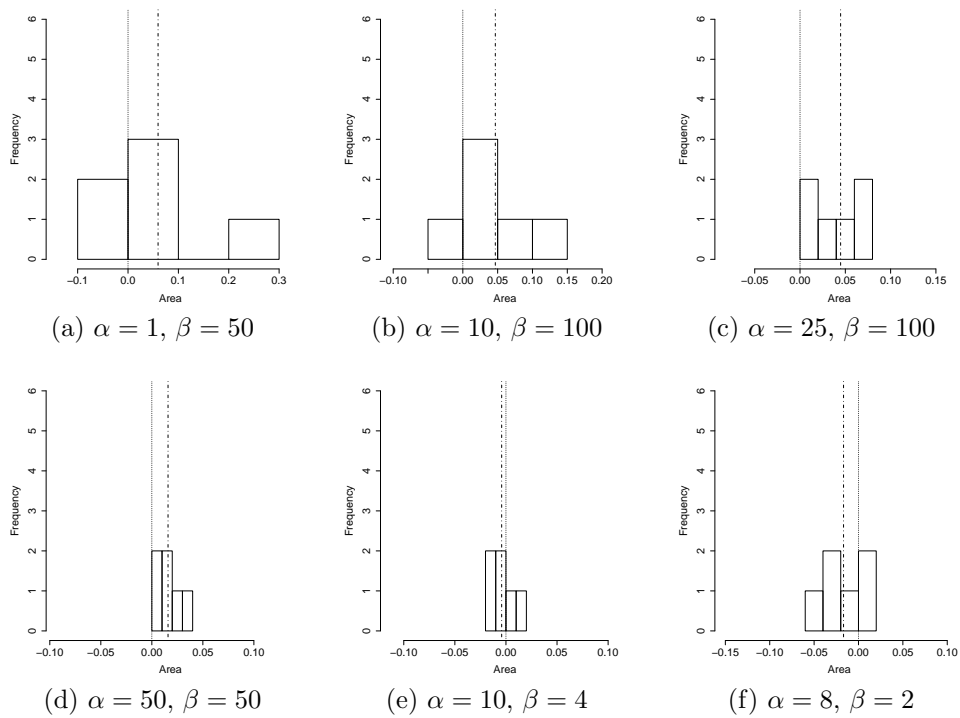
FIGURE 6.9 – Radius-weighted xPODs for 4 seeder species of the Banksia genus on a high diversity plot in Western Australia

Synthesis and future work

The debate about biodiversity and its assessment is as lively and urgent as never before, within the academic community and beyond (Magurran & McGill, 2011). The (sobering) realisation that humanity has (mostly over the last 100 years, Chapin III *et al.* (2000)) contributed to an unprecedented loss of biodiversity set the course at the turn of the century: the Millenium Goals (United Nations, 2000) contain, amongst others, the aim set by the Convention on Biological Diversity (CBD, 1992) to reverse this negative trend on a global level. The achievement of these goals cannot be successful unless we have a solid monitoring framework in place including objective, comprehensive and integrative quantitative methods that can capture progress and failure of efforts undertaken (Dobson, 2005).

Against this background, this thesis investigated methods to assess biodiversity, with a focus on large-scale, continuous monitoring (in both time and space). Herein, we concentrated on quantitative aspects of biodiversity monitoring, where the aim was to contribute to further development of appropriate methodology.

In the following review, we summarise our results with respect to six 'keystone' themes. These were identified in chapter 1 as issues of particular relevance and provided the frame for the research carried out and documented in this thesis.

1. Biodiversity as a concept is foremost characterised by its **multidimensionality**. Diversity quantification not only involves a decision on which of the components of biodiversity are of particular interest to a study, but is also faced with the dilemma of choosing between condensing information into a summary statistic and the loss of information caused by reducing its dimension by doing so. Chapter 1 positioned the research of this thesis within the plurality of aspects ranging from genetic diversity to the complexity of entire ecosystems. This emphasised that species diversity, which has been considered here, while certainly important, can only be

one of the many pieces in the puzzle (Magurran & McGill, 2011). In the context of species diversity, we reviewed parametric approaches that combine several indices and hence capture more of the multidimensionality. In particular, the information they contain can be represented by profile plots. We fully support and embrace recent developments to 'revive' earlier works on diversity profiling (Hill, 1973; Patil & Taillie, 1982; Ricotta, 2003; Jost, 2010; Leinster & Cobbold, 2012): in chapter 2 we suggested a one-parameter family of evenness measures based on goodness-of-fit statistics. We investigated this family with regard to its theoretical properties (chapters 2 and 3) as well as in applications to data (chapters 2 and 4). Analogous to other diversity profiling techniques, we derived evenness profiles. We established connections between this approach and existing ones and compared their performance. The goodness-of-fit measures overcome restrictions of other methods with regard to the range of parameter values for which they are evaluated. This provided tools to investigate diversity more flexibly in terms of commonness and rarity of species in a community, and makes them an ecologically informative complement to headline indices of diversity, such as the geometric mean index (chapter 4).

2. Diversity assessment is likely to be biased if variation in detection probability across species is not taken into account (Yoccoz *et al.*, 2001; Buckland *et al.*, 2011a). We discussed two issues of **detectability**, that of individuals and that of species (see section 1.2.3 and chapter 3). While the latter refers to a species being completely undetected by a survey, the former concerns the variability in detection probabilities between individuals of different species in the observation process.

We focussed on individual detectability; in chapter 3 we demonstrated the bias in diversity summary statistics that results from ignoring variation in detectability. Given appropriate survey design, we have statistical methods in place to adjust diversity estimates for detectability (as discussed in chapter 3.3 and demonstrated in chapter 4). Based on this, we argued that diversity monitoring (survey and analysis) should aim to incorporate information on detection probabilities.

Furthermore, chapter 3 investigated effects of both species and individual detectability on the goodness-of-fit measures introduced in chapter 2. This revealed important differences between the family of goodness-of-fit statistics and the family of divergence metrics related to it, and indicates that some care needs to be taken when these measures are chosen for

diversity assessment. As with other aspects of diversity quantification, one approach is not superior as such; divergence measures are closer to traditional indices in conception as well as linked with indices used in economics, which makes them more easily interpretable and perhaps more universal. Conceptually, goodness-of-fit measures are interpreted purely from a sample-based perspective. To give them a meaning on the level of the community, about which we are drawing inference, a metapopulation framework has to be adopted. On one hand, this might be perceived as less intuitive, on the other, this might be more realistic since, with the exception of a complete census of a closed population, any diversity assessment will always be sample-based.

3. Closely related to the detectability issue, difficulties arise because of **rare species**, which are more likely to be missed in the sampling process. This is notably an issue for large-scale monitoring programmes, where only a fraction of all possible locations is usually surveyed. Rare species can thus be expected to contribute substantially to zeros in the data (see discussion in chapter 3). Although zeros do not necessarily have to be caused by rare species alone, the consequences of missing rare species is of special interest in connection with the goodness-of-fit approach suggested in this thesis, since this method allows us to focus on rare species in particular.

   Missing observations in general and possible ways to deal with them have been discussed in chapter 3. If we have indication of a species being unobserved, although present, this zero should be included in the evaluation. Our results in chapter 3 showed that the goodness-of-fit based measures take this into account differently from both traditional indices and the divergence measures that they are linked to. Hence, they might be more appropriate where missing observations are caused by a certain behaviour of a species, such as seasonal migration (see example in chapter 3).

   Zeros can be included without problems if the goodness-of-fit family is evaluated for positive parameters. For a full profile, however, a small quantity has to be added to achieve positive (but close to zero) values to enable evaluation. On the downside, the bias in the measures for rare species (corresponding to the part of the profile corresponding to negative parameter values) depends on the choice of this small quantity. In the wider picture, this reflects the increase in uncertainty connected to rare species overall: when we first included them in the analysis presented in chapter 4, confidence intervals became basically meaningless for those

indices that are sensitive to rare species (the geometric mean and the goodness-of-fit measures for negative parameter values). This might be seen as a disadvantage of these measures in comparison to traditional indices, such as Shannon's and Simpson's. However, precision of the latter is only higher because they do not give much weight to rare species in the first place.

Sometimes a model-based approach can be used to overcome this problem (Gotelli *et al.*, 2010); by replacing observed (zero) counts by predictions from a fitted model we can get more reliable results, as shown for simulated data in chapter 3. We also adopted this approach in the analysis of data from the British Breeding Bird Survey in chapter 4 to reduce the effect of short-term fluctuations and to overcome missing observations for some rare species in single years. However, fitting a model relies on having sufficient data. In some cases this can be achieved by pooling (see analysis in 4) or additional surveys that target problematic species to gain more information. Where model-fitting is no longer possible, analysis might sometimes be carried out based on the assumption that rare species behave similarly to more common ones (for example, in terms of detection probability). In general, rare species and the corresponding lack of data will compromise precision if they are included in an analysis of data from a multispecies survey. Hence, the rarest species might have to be excluded (as can be seen in the analysis presented in chapters 4 and 6). This is not unambiguous: while there are arguments that rare species do not contribute much in terms of ecosystem functions and that they will not have much influence on trends in diversity (Smith & Knapp, 2003), in particular on a large scale, we know that the majority of species within an assemblage is likely to be rare (Rabinowitz *et al.*, 1986; Magurran & Henderson, 2003); this holds in particular for highly diverse tropical rainforests which are not easily monitored (Longino *et al.* (2002); see also section on Limitations below). In addition, studies show that at least some rare species play a crucial role within ecosystems (Lyons & Schwartz, 2001) and that trends in common species might not automatically reflect the same trend in rare species (see results in chapter 4 as well as Buckland *et al.* (2011b)).

4. Monitoring biodiversity, locally or globally, needs to be able to identify **(long-term) temporal trends** (Magurran *et al.*, 2010; Magurran, 2011). This has implications for the design of monitoring programmes (as discussed in chapter 1.4) as well as data analysis (see chapter 4). Sufficiently

long time series of data are a prerequisite for the analysis of trends in bio-diversity over time, but will often be diffcult to obtain (Magurran *et al.*, 2010). To set an example in a 'best case scenario', we thus relied on data for birds, a taxonomic group that is easily and hence well-monitored, at least in some parts of the world, for the analysis presented in chapter 4. Although considered as indicator species, by no means do we claim that their trends are universal (see section on Limitations below). However, the geometric mean, which has been emphasised as a good headline index by Buckland *et al.* (2011b) and was used in our analysis to quantify trends in diversity for British breeding birds, has an advantage in that it allows the user to combine information across different surveys (see for example the Living Planet Index, Loh *et al.* (2005)). Furthermore, the goodness-of-fit approach introduced in chapter 2 allows us to separate trends in rare species, as long as we have sufficient data for a reliable analysis, from trends in common species. Applying this to the BBS data in chapter 4, we could reveal that monitoring might be at the expense of those species that are rare, but not (yet) of conservation concern. Although derived for a specific taxonomic group here, we believe that this result might be relevant to other taxa and monitoring schemes.

The Living Planet Index, for example, averages over a large number of species by pooling data from different sources (Loh *et al.*, 2005, 2010). Where such large sample sizes cannot be obtained, diversity assessment needs to establish long-term trends by separating it from short-term (sto-chastic) fluctuations. The analysis in chapter 4 achieved this by applying generalized additive models as smoothers. It is only from these long-term trends that we can reliably locate **trend change points**. This enables us to tell whether or not the rate of change (i.e. the loss in biodiversity) has changed. Chapter 4 demonstrated how statistical inference on this can be based on the investigation of the second derivative (Fewster *et al.*, 2000; Buckland *et al.*, 2005). This provides an intuitive and easily imple-mented approach; in addition it is independent from the choice of baseline, where a relative diversity measure is considered. Change point analysis is an alternative statistical tool, although it might be more appropriate for considering rapid or sudden changes on smaller spatial scales (Baker & King, 2010; Thomson *et al.*, 2010).

5. Similar to trends over time, assessment of diversity **in space** should cap-ture continuous spatial trends if it is aimed at the global biodiversity

targets (Yoccoz *et al.*, 2001). Analysis across regions can be based on pooled data, as in chapter 4, or look at local variation in diversity in continuous space (chapter 5). The former provides an overall picture on diversity and will be the main interest when headline indices are considered. However, the latter is relevant as there will be local variation in diversity (McGill, 2011) which is concealed by a pooled analysis. Chapter 5 discussed the limitations of the classical partitioning of diversity in γ- (pooled), α- (local) and β-diversity (between spatial locations), if we are interested in inference across continuous space. We demonstrated how local diversity in bird species that live close to human habitation can be mapped and spatial trends identified. Crucially, in continuous space autocorrelation has to be taken into account. This is only possible in a model-based approach; we used a recently developed model fitting algorithm which can deal with complex autocorrelation structures. Moreover, it can be extended to spatio-temporal observations.

Space is different from time in that it is not linear. As a consequence, scale becomes an issue; this has notably been discussed for β-diversity (Gaston *et al.*, 2007). Ultimately, methods should be such that they allow for a flexible up- and downscaling. As far as our analysis in chapter 5 could reach, we expect this to fast become computationally intensive. Flexibility of scaling is the main topic in chapter 6, which looked at information contained in the spatial structure of ecological communities. However, the focus was on processes operating at a local (small) scale here, in contrast to the large-scale assessment considered in the previous chapters. In the latter, scale is determined by the level at which we aggregate data *prior* to analysis, while the former applies summary statistics that are, by definition, a function of scale.

6. Although both look at diversity in space, chapters 5 and 6 are different, conceptually and with regard to spatial scale. In the first case, the data locations were outcomes of a sampling process; it is important that this sampling process is random in order to be representative of the surveyed area, in this case the UK (see discussion on survey design in chapter 1). The observations at these random locations were then used to estimate and map diversity across a whole region (chapters 4 and 5). In contrast, the methodological framework introduced in chapter 6 is based on the assumption that the observed locations are outcomes of an underlying process and hence inference is based on the locations themselves. Pielou,

as early as 1969, considered this *spatial pattern* formed by the locations of individuals from different species to be informative for the processes underlying the structure of ecological communities. Indeed, Pielou (1969) discusses this *before* diversity assessment based on the species abundance distribution (SAD) in her book. Nevertheless, SADs and the questions of modelling them as well as of summarising information contained in them have become dominant for quantification of biodiversity (Magurran, 2004; McGill *et al.*, 2007), but are not unquestioned (McGill *et al.*, 2007; Brown *et al.*, 2011). The comparison of the two different perspectives given to diversity in space by this thesis demonstrated that both are important. If our aim is a summary of the state of a country's biodiversity and the analysis of overall trends in time and space as well as points where change in species diversity is happening, we are likely to refer to one or several summary statistics based on the SAD. Although 'monitoring biodiversity' tends to imply following only the reaction to conservation efforts, understanding the processes driving biodiversity is equally important. The latter operate on local scales (Kerr *et al.*, 2002), and chapter 6 was concerned with how they can be analysed without having to decide on the scale of evaluation prior to the analysis.

Limitations

More than anything else, monitoring biodiversity is a discipline of choices. Any form of diversity assessment, in particular on large spatial scales, will be restricted deliberately by the choices that we make, but also by what is attainable.

The multidimensionality of biodiversity (see 1) always requires a choice of which aspect is to be investigated in detail. In this thesis, we focussed on species diversity and its quantifaction, while other aspects could only be mentioned. While there are more recent, equally important advances with respect to other aspects of biodiversity (genetic, phylogenetic, taxonomic, functional diversity measures as well as network approaches to food webs and other structures of ecosystems), species diversity is (still) the aspect of biodiversity that researchers and policy makers alike will turn to, at least in the first instance. Reasons are manifold; species diversity has a long tradition (Fisher *et al.*, 1943), appears as intuitive (May, 1988) and comes with a well equipped toolbox of summary statistics (Magurran, 2004), which promises to provide accessible headline indices for large-scale monitoring (Buckland *et al.*, 2005). This makes it even more important to understand how these indices behave if they are evaluated across large regions, such as entire countries, and what their limitations are.

The work presented here set out to contribute to a better understanding in this sense.

Buckland *et al.* (2005) already highlighted short-comings of the classical ecological indices, Simpson's and Shannon's, and argued for the use of a geometric mean as a headline index. However, the work presented here showed that none of these summary statistics is able to differentiate between trends in rare and those in common species. While the goodness-of-fit approach can in principle overcome this, we saw that missing observations and uncertainty about the number of species that are considered can pose a problem. While zeros can be included to some extent, they increase uncertainty. Too little data or only geographically restricted data will not provide sufficient amount of information to draw inference on a global scale.

This connects to the part of diversity assessment which is not down to choice, but to the availability of data. If we want to get a reliable picture of diversity across large spatial scales, it should be based on the best data that are available at this scale and that do not surpass the limits of our resources. Nevertheless, this is likely to leave many taxonomic groups underrepresented. Here, we looked at birds, and in general vertebrate species tend to be covered better by surveys and in terms of data quality, while they only make up approximately five percent of all known species (Landres *et al.*, 1988). While we need these good quality data to get reliable results, we should be aware that they might not simply transfer to other groups for which we have less or no data. To some extent, this is unavoidable, but care should be taken that it is not how 'cute' a taxonomic group is, or how familiar it is, or how strong the lobby for it is, that decides over objective reasons why diversity assessment should be based on this group and not another. However, to overcome this tendency, it does not suffice to convince the public, policy makers or conservationists of the importance of those species groups which have been considered less, but first and foremost we need to develop an appropriate methodological framework for them, in terms of survey design as well as statistical analysis, on a global scale. This might for some prove infeasible, at least for the time being (Lawton *et al.*, 1998; Longino *et al.*, 2002).

In many cases, on a large scale, data will be most easily collected on the species level, but clearly not in all. Where individuals are not readily sampled or are, for example, different in size, other methods and different 'currencies' to those that were considered in this thesis might be more appropriate. The methods suggested here are limited in terms of transferring them onto other settings; by concept, goodness-of-fit statistics rely on counts and hence should

be applied to individual observations. However, the divergence approach that is connected to them is based on proportions and can hence be applied to biomass data, for example.

Similarly, the adjustment for individual detectability that we used here, might not be applicable to some species. It relies on a specific survey design (distance sampling) as well as sufficient data. While issues with the latter have been discussed, the required survey design might not be appropriate or not feasible for some species. Given the results presented here, detectability or catchability should always be considered as potential sources of bias. Re-sampling techniques, mark-recapture, or potentially genetic methods such as bar coding, offer alternatives where distances to either individuals or groups of individuals cannot reasonably be determined.

In conclusion, while we should not defer large scale assessment of biodiversity until we have the 'best method' (which does not exist), we should at the same time continue working towards well designed surveys, a variety of data from different taxa as well as the integration of methods.

To be continued ...

Apart from the aim to work towards a better integration of methods used for quantification of biological diversity in general, the research presented here stimulates further investigation and provides room for extensions.

Within the scope of this thesis, the application of goodness-of-fit statistics and of the divergence measures related to them to data could not be exhaustive. The method promises new insights, in particular for separating trends in rare and common species. As we have discussed, this could be crucial for a reliable assessment of biodiversity, at least in some communities. Although it covered different habitat types, our analysis only looked at one taxonomic group. Further study is needed in terms of its applicability and usefulness for other ecological assemblages in general. While our work revealed important differences between the goodness-of-fit statistics as measures and the divergence measures and gave a first indication as to when either might be appropriate, this clearly merits additional investigation. In particular, we need a better understanding and ways to differentiate between what causes zeros in the data; modelling approaches could potentially reduce bias and increase precision here, as discussed in chapter 3. We therefore expect that a closer linkage between model-based approaches and diversity quantification might help capture the complex structure of ecological communities and improve inference, in particular for less detectable

species (Gotelli *et al.*, 2010). Powerful model fitting algorithms, such as INLA, are now available and continue to be improved; we hope that the application we presented in chapter 5 demonstrates convincingly their usefulness for ecological questions and encourages non-statisticians to exploit theses methods.

The work in this thesis strongly advocates multidimensional approaches to diversity assessment as provided by parametric index families. Recently, a generalisation of Hill's numbers has been proposed (Leinster & Cobbold, 2012), which can include information on species similarities in addition to their dominance or scarcity. Similarity can be flexibly defined here, for example by phylogeny or through certain traits. This acknowledges that the loss of a species which is unique in its functional or phylogenetic position within an ecosystem might potentially outweigh that of a member of a highly redundant group of species. It also opens the possibility of combining information on different aspects of biodiversity, such as species diversity and phylogenetic or functional diversity, within the same measuring approach. Since the goodness-of-fit approach is in essence a transformation of Hill's numbers, it might be possible to extend it in a similar way. Although this promises a step towards integration of methods, we should not forget that by increasing the information that is incorporated in a measure, results will potentially become less easily interpretable. In particular, disentangling the effects of scarcity and similarity, either in terms of phylogeny or ecosystem function, could prove to be challenging. As weighting for similarities is within an assemblage in Leinster and Cobbold's approach, the comparison of different assemblages could also be hindered, if they are not nested (similar to problems for the goodness-of-fit approach as discussed in chapter 2).

While this provides means to account for species' similarities and hence redundancies in the SAD, none of the SAD-based approaches, including Hill's numbers and its transformations as well as the geometric mean or generalised means, allow us to keep track of species' identities. The latter are crucial if we want to consider turnover or changes in composition, either within the same assemblage or by comparison of two or more assemblages (in time or space). Equally important to the assessment of change points in γ- or α-diversity, determining rates of change in turnover might reveal 'hotspots' of rapid change in (local) community composition. If correlated with climate data, this could point at shifts of species ranges as they are happening, and provide valuable insight and tests for predictions from climate change models. However, as discussed in chapter 5, it is not evident how to extend the concept of β-diversity or measures accounting for compositional similarity to continuous space. Here, we only made a suggestion whose applicability needs to be put to the test; in

particular, we anticipated computational problems. The approach outlined in chapter 5 is also quite 'ad hoc' and does not leave much flexibility with respect to spatial scale, for example, other than repeating it for a different resolution of the data. A more rigorous, model-based framework would certainly be superior.

We demonstrated how trends in diversity across large spatial scales can be assessed, separately in time and space (chapters 4 and 5), by smoothing local fluctuations. The next step would be to combine this into a spatio-temporal modelling approach which takes into account temporal and spatial autocorrelation in their full interdependence. The INLA algorithm in combination with the SPDE theory has already proved capable of handling complex spatio-temporal environmental systems (Cameletti *et al.*, 2012; Lindgren *et al.*, 2011). We can envisage a similar application to the BBS data and ultimately for other taxonomic groups, provided data with a sufficient spatial coverage are available.

While large scale assessment is indispensable for following effects of global actions to reduce the loss of biodiversity, we need to continue to invest equal efforts in understanding the mechanics that maintain diversity and coexistence of species. Here, analysing information contained in the spatial structure has been shown to provide promising results. We have extended one spatial summary statistic, the cross-pair overlap distribution, in a way that renders it more flexible to spatial scale and enables the user to adjust its resolution, similar to a 'zoom-lens', to the ecological community of interest. This approach provides tools for explanatory analysis and might also be used to inform modelling approaches; it could be equally applied to other second-order summary characteristics which are a function of scale. Future work could involve a comparison of this approach in combination with a range of summary statistics. The flexibility was achieved by introducing variable weights on spatial scale. While this allows us to focus on certain distances, we could also see the possibility of weighting with respect to other aspects of spatial structure, including inherent characteristics such as mingling. Finally, summary statistics have been suggested in spatial point process theory that can incorporate heterogeneity in environmental conditions (Baddeley *et al.*, 2000). It would be of interest to investigate if and how the latter might be combined with weighted spatial scales.

References

Adams, W.M., & Hutton, J. 2007. People, parks and poverty: political ecology and biodiversity conservation. *Conservation and Society*, **5**, 147–183.

Alexandrou, M.A., Oliveira, C., Maillard, M., McGill, R.A.R., Newton, J., Creer, S., & Taylor, M.I. 2011. Competition and phylogeny determine community structure in Müllerian co-mimics. *Nature*, **469**, 84–88.

Armstrong, P. 1991. *Species patterning in the heath vegetation of the Northern Sandplain*. Honours Thesis, University of Western Australia.

Baddeley, A. 2010a. Modeling strategies. Pages 339–370. In: Gelfand, A.E., Diggle, P.J., Fuentes, M., & Guttorp, P. (eds), *Handbook of Spatial Statistics*. Chapman & Hall, CRC Press, Boca Raton.

Baddeley, A. 2010b. Multivariate and marked point processes. Pages 371–402. In: Gelfand, A.E., Diggle, P.J., Fuentes, M., & Guttorp, P. (eds), *Handbook of spatial statistics*. Chapman & Hall, CRC Press, Boca Raton.

Baddeley, A., Møller, J., & Waagepetersen, R. 2000. Non- and semiparametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, **54**, 329–350.

Baillie, S.R., Marchant, J.H., Leech, D.I., Renwick, A.R., Joyes, A.C., Noble, D.G., C.Barimore, Conway, G.J., Downie, I.S., Risely, K., & Robinson, R.A. 2010. *Breeding Birds in the Wider Countryside: their conservation status 2010*. BTO Research Report No 565, British Trust for Ornithology, Thetford.

Baker, M.E., & King, R.S. 2010. A new method for detecting and interpreting biodiversity and community thresholds. *Methods in Ecology and Evolution*, **1**, 25–37.

Balmford, A., Green, R.E., & Jenkins, M. 2003. Measuring the changing state of nature. *Trends in Ecoloy and Evolution*, **18**, 326–330.

Bartlett, M.S. 1964. A note on spatial pattern. *Biometrics*, **20**, 891–892.

Benton, T.G., Vickery, J.A., & Wilson, J.D. 2003. Farmland biodiversity: is habitat heterogeneity the key? *Trends in Ecology and Evolution*, **18**, 182–188.

Berger, W.H., & Parker, F.L. 1970. Diversity of planktonic foraminifera in deep-sea sediments. *Science*, **12**, 1345–1347.

Bolker, B., & Pacala, S.W. 1997. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theoretical Population Biology*, **52**, 179–197.

Borcard, D., Gillet, F., & Legendre, P. 2011. *Numerical ecology with R*. Springer, New York.

Borchers, D.L., Buckland, S.T., & Zucchini, W. 2002. *Estimating animal abundance: closed populations*. Springer, London.

Boulinier, T., Nichols, J.D., Sauer, J.R., Hines, J.E., & Pollock, K.H. 1998. Estimating species richness: the importance of heterogeneity in species detectability. *Ecology*, **79**, 1018–1028.

Bowman, A.W., & Azzalini, A. 1997. *Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations*. Oxford University Press, New York.

Brechin, S.R., Wilshusen, P.R., Fortwangler, C.L., & West, P.C. 2002. Beyond the square wheel: toward a more comprehensive understanding of biodiversity conservation as a social and political process. *Society and Natural Resources*, **15**, 41–64.

Brose, U., Martinez, N.D., & Williams, R.J. 2003. Estimating species richness: sensitivity to sample coverage and insensitivity to spatial patterns. *Ecology*, **84**, 2364–2377.

Brown, C., Law, R., Illian, J.B., & Burslem, D. 2011. Linking ecological processes with spatial and non-spatial patterns in plant communities. *Journal of Ecology*, **99**, 1402–1414.

Bruckmann, G. 1969. Einige Bemerkungen zur statistischen Messung der Konzentration. *Metrika*, **14**, 183–213.

Buckland, S.T. 1984. Monte Carlo confidence intervals. *Biometrics*, **40**, 811–817.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., & Thomas, L. 2001. *Introduction to Distance Sampling.* Chapman & Hall, London.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L, Borchers, D.L., & Thomas, L. 2004a. *Advanced Distance Sampling.* Oxford University Press, Oxford.

Buckland, S.T., Newman, K.B., Thomas, L., & Koesters, N.B. 2004b. State-space models for the dynamics of wild animal populations. *Ecological Modelling*, **171**, 157–175.

Buckland, S.T., Magurran, A.E., Green, R.E., & Fewster, R.M. 2005. Monitoring change in biodiversity through composite indices. *Philosophical Transactions of the Royal Society B*, **360**, 243–254.

Buckland, S.T., Studeny, A., Magurran, A.E., & Newson, S.E. 2011a. Biodiversity monitoring: the relevance of detectability. Pages 25–36. In: Magurran, A.E., & McGill, B.J. (eds), *Biological diversity: Frontiers in Measurement and Assessment.* Oxford University Press.

Buckland, S.T., Studeny, A.C., Magurran, A.E., Illian, J.B., & Newson, S.E. 2011b. The geometric mean of relative abundance indices: a biodiversity measure with a difference. *Ecosphere*, **2**(9), art.100.

Buckland, S.T., Baillie, S.R., Dick, J.McP., Elston, D.A., Magurran, A.E., Scott, E.M., Smith, R.I., Somerfield, P.J., Studeny, A.C., & Watt, A. In prep. How should regional biodiversity be monitored?

Buisson, L., Thuiller, W., Lek, S., Lim, P., & Grenouillet, G. 2008. Climate change hastens the turnover of stream fish assemblages. *Global Change Biology*, **14**, 2232–2248.

Bulla, L. 1994. An index of evenness and its associated measure. *OIKOS*, **70**, 167–171.

Bunge, J., & Fitzpatrick, M. 1993. Estimating the number of species: a review. *Journal of the American Statistical Association*, **88**, 364–373.

Butchart, S.H.M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J.P.W., Almond, R.E.A., Baillie, J.E.M., Bornhard, B., Brown, C., Bruno, J., Carpenter, K.E., Carr, G.M., Chanson, J., Chenery, A.M., J., Csirke, Davidson,

N.C., Dentener, F., Foster, M., Galli, A., Galloway, J.N., Genovesi, P., Gregory, R.D., Hockings, M., Kapos, V., Lamarque, J.-F., Leverington, F., Loh, J., McGeoch, M.A., McRae, L., Minasyan, A., Hernández Morcillo, M., Oldfield, T.E.E., Pauly, D., Quader, S., Revenga, C., Sauer, J.R., Skolnik, B., Spear, D., Stanwell-Smith, D., Stuart, S.N., Symes, A., Tierney, M., Tyrell, T.D., Vié, J.-C., & Watson, R. 2010. Global Biodiversity: Indicators of Recent Decline. *Science*, **328**, 1164–1168.

Camargo, J. 1995. On measuring species evenness and other associated parameters of community structure. *OIKOS*, **74**, 538–542.

Cameletti, M., Lindgren, F., Simpson, D., & Rue, H. 2012. Spatio-temporal modelling of particulate matter concentration through the SPDE approach. *AStA Advances in Statistical Analysis*. Online preview DOI 10.1007/s10182-012-0196-3.

CBD. 1992. *The Convention on Biological Diversity*. http://www.cbd.int.

Chao, A. 1984. Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statitics*, **11**, 265–270.

Chao, A., & Lee, S.-M. 1992. Estimating the number of classes via sample coverage. *Journal of the American Statistical Association*, **87**, 210–217.

Chao, A., & Shen, T.-J. 2003. Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample. *Environmental and Ecological Statistics*, **10**, 429–443.

Chao, A., Ma, M.-C., & Yang, M.C.K. 1993. Stopping rules and estimation for recapture debugging and unequal failure rates. *Biometrika*, **80**, 193–201.

Chao, A., Chazdon, R.L., R.K.Colwell, & Shen, T.-J. 2006. Abundance-based similarity indices and their estimation when there are unseen species in samples. *Biometrics*, **62**, 361–371.

Chao, A., Colwell, R.K., Lin, C.-W., & Gotelli, N.J. 2009. Sufficient sampling for asymptotic minimum species richness estimators. *Ecology*, **90**, 1125–1133.

Chapin III, F.S., Zavaleta, E.S., Eviner, V.T., Naylor, R.L., Vitousek, P.M., Reynolds, H.L., Hooper, D.U., Lavorel, S., Sala, O.E., Hobbie, S.E., Mack, M.C., & Díaz, S. 2000. Consequences of changing biodiversity. *Nature*, **405**, 234–242.

Chesson, P. 2000. Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, **31**, 343–366.

Chiarucci, A., Wilson, J.B., Anderson, B.J., & De Dominicis, V. 1999. Cover versus biomass as an estimate of species abundance: does it make a difference to the conclusions. *Journal of Vegetation Science*, **10**, 35–42.

Cody, M.L. 1975. Towards a theory of continental species diversity: bird distributions over mediterranean habitat. Pages 214–257. In: Cody, M.L., & Diamond, J.M. (eds), *Ecology and evolution of communities*. Harvard University Press, Cambridge, Massachusetts.

Collins, S.L., Suding, K.N., Cleland, E.E., Batty, M., Pennings, S.C., Gross, K.L., Grace, J.B., Gough, L., Forgione, J.E., & Clark, C.M. 2008. Rank clocks and plant community dynamics. *Ecology*, **89**, 3534–3541.

Colwell, R.K., & Coddington, J.A. 1994. Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society B*, **345**, 101–118.

Colwell, R.K., Mao, C.X., & Chang, J. 2004. Interpolating, extrapolating, and comparing incidence-based species accumulation curves. *Ecology*, **85**, 2717–2727.

Condit, R., Ashton, P.S., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., Hubbell, S.P., Foster, R.B., Itoh, A., LaFrankie, J.V., Lee, H. Seng, Losos, E., Manokaran, N., Sukumar, R., & Yamakura, T. 2000. Spatial patterns in the distribution of tropical tree species. *Science*, **288**, 1414–1418.

Condit, R., Pitman, N., Leigh Jr., E.G., Chave, J., Terborgh, J., Foster, R.B., nez, P. Nú Aguilar, S., Valencia, R., Villa, G., Muller-Landau, H.C., Losos, E., & Hubbell, S.P. 2002. Beta-diversity in tropical forest trees. *Science*, **25**, 666–669.

Cowell, F.A. 1980. Generalized entropy and the measurement of distributional change. *European Economic Review*, **13**, 147–159.

Cressie, N., & Wikle, C.K. 2011. *Statistics for spatio-temporal data*. John Wiley & Sons, Hoboken, New Jersey.

Cressie, N.A.C., & Read, T.R.C. 1984. Multinomial Goodness-of-Fit Tests. *Journal of the Royal Statistical Society B*, **46**, 440–464.

Culver, M., Fitak, R., & Herrmann, H.-W. 2011. Genetic methods for biodiversity assessment. Pages 208–218. In: Magurran, A.E., & McGill, B.J. (eds), *Biological diversity: Frontiers in Measurement and Assessment.* Oxford University Press.

Cunningham, R.B., & Lindenmayer, D.B. 2005. Modeling count data of rare species: some statistical issues. *Ecology*, **86**, 1135–1142.

Darwin, C. 1859. *On the origin of species by means of natural selection.* John Murray, London.

Davey, C.M., Chamberlain, D.E., Newson, S.E., Noble, D.G., & Johnston, A. 2012. Rise of the generalists: evidence for climate driven homogenization in avian communities. *Global Ecology and Biogeography*, **21**, 568–578.

de Heer, M., Kapos, V., & ten Brink, B.J.E. 2005. Biodiversity trends in Europe: development and testing of a species trend indicator for evaluating progress towards the 2010 target. *Transactions of the Royal Society B*, **360**(February), 297–308.

DEFRA. 2011. *UK biodiversity indicators.* Statistical Release, Department for Environment, Food and Rural Affairs, London.

DeLong, D.C. 1996. Defining biodiversity. *Wildlife Society Bulletin*, **24**, 738–749.

Dìaz, S., & Carbido, M. 2001. Vive la difference: plant functional diversity matters to ecosystem processes. *Trends in Ecology and Evolution*, **16**, 646–655.

Diggle, P.J. 1983. *Statistical analysis of spatial point patterns.* Academic Press, London.

Dobson, A. 2005. Monitoring global rates of biodiversity change: challenges that arise in meeting the Convention on Biological Diversity (CBD) 2010 goals. *Philosophical Transactions of the Royal Society B*, **360**, 229–241.

Doob, J.L. 1994. *Measure theory.* Springer.

Dornelas, M. 2010. Disturbance and change in biodiversity. *Philosophical Transactions of the Royal Society B*, **365**, 3715–3727.

Dornelas, M., Moonen, A.C., Magurran, A.E., & Bàrben, P. 2009. Species abundance distributions reveal environmental heterogeneity in modified landscapes. *Journal of Applied Ecology*, **46**, 666–672.

Doucet, A., Freitas, N. De, & Gordon, N. 2001. *Sequential Monte Carlo methods in practice*. Springer, New York.

Dunne, J.A., Williams, R.J., & Martinez, N.D. 2002. Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, **5**, 558–567.

Everitt, B.S., Landau, S., & Leese, M. 2001. *Cluster Analysis*. Arnold, London.

Fesl, C. 2002. Niche-oriented species-abundance models: different approaches of their application to larval chironomid (Diptera) assemblages in a large river. *Journal of Animal Ecology*, **71**, 1085–1094.

Fewster, R.M., & Jupp, P.E. 2009. Inference on population size in binomial detectability models. *Biometrika*, **96**, 805–820.

Fewster, R.M., Buckland, S.T., Siriwardena, G.M., Baillie, S.R., & Wilson, J.D. 2000. Analysis of population trends for farmland birds using generalized additive models. *Ecology*, **81**, 1970–1984.

Field, R.H., & Gregory, R.D. 1999. *Measuring population changes from the Breeding Bird Survey*. BTO Research Report No 217, British Trust for Ornithology, Thetford.

Fisher, R.A., Corbet, A.S., & Williams, C.B. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology*, **12**, 42–58.

Gaston, K. J. 2000. Global patterns in biodiversity. *Nature*, **405**, 220–227.

Gaston, K. J. 2008. Biodiversity and extinction: the importance of being common. *Progress in Physical Geography*, **32**, 73–79.

Gaston, Kevin J., & Spicer, John I. 2004. *Biodiversity: An Introduction*. Blackwell Publishing, Oxford.

Gaston, K.J., Evans, K.I., & Lennon, J.J. 2007. The scaling of spatial turnover: pruning the thicket. Pages 181–222. In: Storch, D., Marquet, P.A., & Brown, J.H. (eds), *Scaling biodiversity*. Cambridge University Press, Cambridge, UK.

Gbur, G.J. 2011. *Mathematical methods for optical physics and engineering.* Cambrigde University Press, Cambridge, UK.

Gering, J.C., Crist, T.O., & Veech, J.A. 2003. Additive partitioning of species diversity across multiple spatial scales: implications for regional conservation of biodiversity. *Conservation Biology*, **17**, 488–499.

Gering, Jon C., & Crist, Thomas O. 2002. The alpha-beta-regional relationship: providing new insights into local-regional patterns of species richness and scale dependence of diversity components. *Ecology Letters*, **5**, 433–444.

Gilbert, B., & Lechowicz, M.J. 2004. Neutrality, niches, and dispersal in a temperate forest understory. *Proceedings of the National Academy of Science*, **101**, 7651–7656.

Good, I.J. 1953. The population frequencies of species and the estimation of population parameters. *Biometrika*, **40**, 237–264.

Gosselin, F. 2006. An assessment of the dependence of evenness indices on species richness. *Journal of Theoretical Biology*, **242**, 591–597.

Gotelli, N.J., & Colwell, R.K. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, **4**, 379–391.

Gotelli, N.J., & Colwell, R.K. 2011. Estimating species richness. Pages 39–54. In: Magurran, A.E., & McGill, B.J. (eds), *Biological diversity: Frontiers in Measurement and Assessment.* Oxford University Press.

Gotelli, N.J., & Graves, G.R. 1996. *Null models in ecology.* Smithonian Instiute Press, Washington.

Gotelli, N.J., & McCabe, D.J. 2002. Species co-occurence: a meta-analysis of J.M. Diamond's assembly rules model. *Ecology*, **83**, 2091–2096.

Gotelli, N.J., & McGill, B.J. 2006. Null versus neutral models: what's the difference? *Ecography*, **29**, 793–800.

Gotelli, N.J., Dorazio, R.M., Ellison, A.M., & Grossman, G.D. 2010. Detecting temporal trends in species assemblages with bootstrapping procedures and hierarchical models. *Philosophical Transactions of the Royal Society B*, **365**, 3621–3631.

Greenwood, J.J.D., Gregory, R.D., Peach, W.J., & Fuller, R.J. 1995. Some new approaches to conservation monitoring of British breeding birds. *Ibis*, **137**, 16–28.

Gregory, R.D., & Baillie, S.R. 1998. Large-scale habitat use of some declining British birds. *Journal of Applied Ecology*, **35**, 785–799.

Gregory, R.D., & van Strien, A. 2010. Wild bird indicators: using composite population trends of birds as measures of environmental health. *Ornithological Science*, **9**, 3–22.

Gregory, R.D., Noble, D.G., Field, R.H., Marchant, J.H., Raven, M., & D.W. 2003. Using birds as indicators of biodiversity. *Ornis Hungarica*, **12-13**, 11–24.

Gregory, R.D., van Strien, A., Vorisek, P., Meyling, A.W. Gmelig, Noble, D.G., Foppen, R.P.B., & Gibbons, D.W. 2005. Developing indicators for European birds. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**, 269–288.

Hannah, L., Midgley, G.F., Lovejoy, T., Bond, W.J., Bush, M., Lovett, J.C., Scott, D., & Woodward, F.I. 2002. Conservation of biodiversity in a changing climate. *Conservation Biology*, **16**, 264–268.

Harper, J.L., & Hawksworth, D.L. 1995. *Biodiversity: Measurement and Estimation.* Chapman & Hall, CRC Press, London. Chap. Preface.

Harte, J., Conlisk, E., Ostling, Annette, Green, J.L., & Smith, A.B. 2005. A theory of spatial structure in ecological communities at multiple spatial scales. *Ecological Monographs*, **75**, 179–197.

Hastie, T.J., & Tibshirani, R.J. 1990. *Generalized Additive Models.* Chapman & Hall, CRC Press, London.

He, F., & Condit, R. 2007. The distribution of species: occupancy, scale and rarity. Pages 32–50. In: Storch, D., Marquet, P.A., & Brown, J.H. (eds), *Scaling biodiversity.* Cambridge University Press, New York.

Heip, C. 1974. A new index measuring evenness. *Journal of the Marine Biological Association of the United Kingdom*, **54**, 555–557.

Hellmann, J.J., & Fowler, G.W. 1999. Bias, precision, and accuracy of four measures of species richness. *Ecological Applications*, **9**, 824–834.

Henderson, P.A., & Magurran, A.E. 2010. Linking species abundance distributions in numerical abundance and biomass through simple assumptions about community structure. *Proceedings of the Royal Society B*, **277**, 1561–1570.

Hill, M.O. 1973. Diversity and Evenness: A Unifying Notation and Its Consequences. *Ecology*, **54**, 427–432.

Hill, M.O. 1997. An evenness statistic based on abundance-weighted variance of species proportions. *OIKOS*, **79**, 413–416.

Hoffmann, Sönke. 2008. *Generalized distribution based diversity measurement: survey and unification.* FEMM working paper No 23, Otto-von-Guericke Universität Magdeburg.

Hooper, D.U., III, F.S. Chapin, Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setälä, H., Symstad, A.J., Vandermeer, J., & Wardle, D.A. 2005. Effects of biodiversity on ecosystem functioning: a concensus of current knowledge. *Ecological Monographs*, **75**, 3–35.

Horn, H.S. 1966. Measurement of 'overlap' in comparative ecological studies. *The American Naturalist*, **100**, 419–424.

Hubbell, S.P. 2001. *The unified neutral theory of biodiversity and biogeography.* Princeton University Press, Princeton, New Jersey.

Hughes, J.B., Hellmann, J.J., Ricketts, T.H., & Bohannan, B.J.M. 2001. Counting the uncountable: statistical approaches to estimating microbial diversity. *Applied and Evironmental Microbiology*, **67**, 4399–4406.

Hurlbert, S.H. 1971. The Nonconcept of Species Diversity: A Critique and Alternative Parameters. *Ecology*, **42**, 577–586.

Hutchens, S.J., & DePerno, C.S. 2009. Efficacy of sampling techniques for determining species richness estimates of reptiles and amphibians. *Wildlife Biology*, **15**, 113–122.

Illian, J.B., & Hendrichsen, D.K. 2010. Gibbs point process models with mixed effects. *Environmetrics*, **21**, 341–353.

Illian, J.B., Penttinen, A., H., H. Stoyan, & Stoyan, D. 2009a. *Statistical Analysis and Modelling of Spatial Point Patterns.* John Wiley & Sons, New York.

Illian, J.B., Moller, J., & Waagepetersen, R.P. 2009b. Hierarchical spatial point process analysis for a plant community with high biodiversity. *Environmental and Ecological Statistics*, **16**, 389–405.

Jacobs, J. 1974. Quantitative measurement of food selection. *Oecologia*, **14**, 413–417.

Jenkins, S.P. 2009. The measurement of economic inequality. Pages 40–67. In: Salverda, W., Nolan, B., & Smeeding, T.M. (eds), *The Oxford handbook of economic inequality*. Oxford University Press, Oxford.

Jones, J.P.G. 2011. Monitoring species abundance and distribution at the landscape scale. *Journal of Applied Ecology*, **48**, 9–13.

Jørgensen, B. 1997. *The Theory of Dispersion Models*. Chapman & Hall, London.

Jost, L. 2006. Entropy and Diversity. *OIKOS*, **113**, 363–375.

Jost, L. 2007. Partitioning diversity into independent alpha and beta components. *Ecology*, **88**, 2427–2439.

Jost, L. 2010. The relation between evenness and diversity. *Diversity*, **2**, 207–232.

Jost, L., Chao, A., & Chazdon, R.L. 2011. Compositional similarity and β-diversity. Pages 66–84. In: Magurran, A.E., & McGill, B.J. (eds), *Biological diversity: Frontiers in Measurement and Assessment*. Oxford University Press.

Karagrigoriou, A., & Mattheou, K. 2010. On Distributional Properties and Goodness-of-Fit Tests for Generalized Measures of Divergence. *Communications in Statistics – Theory and Methods*, **39**, 472–482.

Kelly, C.K., Bowler, M.G., Pybus, O., & Harvey, P.H. 2006. Phylogeny, niches, and relative abundance in natural communities. *Ecology*, **89**, 962–970.

Kerr, B., Riley, M.A., Feldmann, M.W., & Bohannan, B.J.M. 2002. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature*, **418**, 171–174.

Koeller, P.A., Hurley, P.C.F., Perley, P., & Neilson, J.D. 1986. Juvenile fish surveys on the Scotian Shelf: implications for the year-class size assessment. *ICES Journal of Marine Science*, **1**, 59–76.

Koleff, P., Gaston, K.J., & Lennon, J.J. 2003. Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, **72**, 367–382.

Kullback, S. 1968. *Information Theory and Statistics.* Dover Publications, New York.

Lamb, E.G., Bayne, E., Holloway, G., Schieck, J., Boutin, S., Herbers, J., & Haughland, D.L. 2009. Indices for monitoring biodiversity change: are some more effective than others? *Ecological Indicators*, **9**, 432–444.

Lande, R. 1996. Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos*, **76**, 5–13.

Landres, P.B., Verner, J., & Thomas, J.W. 1988. Ecological uses of vertebrate indicator species: a critique. *Conservation Biology*, **2**, 316–328.

Law, R., Illian, J.B., Burslem, D., Gratzer, G., Gunatilleke, S., & Gunatilleke, N. 2009. Ecological information from spatial patterns of plants: insights from point process theory. *Journal of Ecology*, **97**, 616–628.

Lawler, J.J., White, D., Sifneos, J.C., & Master, L.L. 2003. Rare species and the use of indicator groups for conservation planning. *Conservation Biology*, **3**, 875–882.

Lawton, J.H. 1994. What do species do in ecosystems? *Oikos*, **71**, 367–374.

Lawton, J.H., Bignell, D.E., Bolton, B., Bloemers, G.F., Eggleton, P., Hammond, P.M., Hodda, M., Holt, R.D., Larsen, T.B., Mawdsley, N.A., Stork, N.E., Srivastava, D.S., & Watt, A.D. 1998. Biodiversity inventories, indicator taxa and effects of habitat modification in tropical rain forest. *Nature*, **391**, 72–76.

Lee, M.S.Y., & Chao, A. 1994. Estimating population size via sample coverage for closed capture-recapture models. *Biometrics*, **50**, 88–97.

Leinster, T., & Cobbold, C. 2012. Measuring diversity: the importance of species similarity. *Ecology*, **93**, 477–489.

Levin, S.A. 2000. Multiple scales and the maintenance of biodiversity. *Ecosystems*, **3**, 498–506.

Lindgren, F., Rue, H., & Lindström, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: The SPDE approach. *Journal of the Royal Statistical Society, Series B*, **73**, 423 – 498.

Liu, C., Whittaker, R.J., Ma, K., & Malcolm, J.R. 2007. Unifying and distinguishing diversity ordering methods for comparing communities. *Population Ecology*, **49**, 89–100.

Loh, J., Green, R.E., Lamoreux, J., Jenkins, M., Kapos, V., & Randers, J. 2005. The Living Planet Index: using species population time series to track trends in biodiversity. *Philosophical Transactions of the Royal Society B*, **360**, 289–295.

Loh, J., Collen, B., McRae, L., Deinet, S., De Palma, A., Manley, R., & Baillie, J.E.M. 2010. *Monitoring biodiversity - the Living Planet Index.* Technical report. WWF, ZSL and GFN.

Longino, J.T., Coddington, J., & Colwell, R.K. 2002. The ant fauna of tropical rain forest: estimating species richness three different ways. *Ecology*, **83**, 689–702.

Lopez, L.C.S., de Aguiar Fracasso, M.P., Mesquita, D.O., Palma, A.R.T., & Riul, P. 2012. The relationship between percentage of singletons and sampling effort: a new approach to recude the bias of richness estimation. *Ecological Indicators*, **14**, 164–169.

Lorenz, M.O. 1905. Methods of measuring concentration of wealth. *American Statistical Association*, **70**, 209–219.

Lyons, K.G., & Schwartz, M.W. 2001. Rare species loss alters ecosystem function - invasion resistance. *Ecology Letters*, **4**, 358–365.

Mac Nally, Ralph, Fleishman, Erica, Bulluck, Lesley P., & Betrus, Christopher J. 2004. Comparative influence of spatial scale on beta diversity within regional assemblages of birds and butterflies. *Journal of Biogeography*, **31**, 917–929.

MacArthur, R. 1960. On the relative abundance of species. *The American Naturalist*, **94**, 25–36.

MacKenzie, D.I., & Kendall, W.L. 2002. How should detection probability be incorporated into estimates of relative abundance. *Ecology*, **83**, 2387–2393.

MacKenzie, D.I., Nichols, J.D., Hines, J.E., Knutson, M.G., & Franklin, A.B. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology*, **84**, 2200–2207.

MacKenzie, D.I., Nichols, J.D., Sutton, N., Kawanishi, K., & Bailey, L.L. 2005. Improving inferences in population studies of rare species that are detected imperfectly. *Ecology*, **86**, 1101–1113.

Magurran, A. E. 2004. *Measuring Biologial Diversity*. Blackwell Publishing, Oxford.

Magurran, A. E., & McGill, B.J. (eds). 2011. *Biological diversity: Frontiers in Measurement and Assessment*. Oxford University Press.

Magurran, A.E. 2011. Measuring biological diversity in time (and space). Pages 85–95. In: Magurran, A.E., & McGill, B.J. (eds), *Biological diversity: Frontiers in Measurement and Assessment*. Oxford University Press, Oxford.

Magurran, A.E., & Dornelas, M. 2010. Biological diversity in a changing world. *Philosophical Transactions of the Royal Society B*, **365**, 3593–3597.

Magurran, A.E., & Henderson, P.A. 2003. Explaining the excess of rare species in natural species abundance distributions. *Nature*, **422**, 714–716.

Magurran, A.E., & Henderson, P.A. 2010. Temporal turnover and the maintenance of diversity in ecological assemblages. *Philosophical Transactions of the Royal Society B*, **365**, 3611–3620.

Magurran, A.E., & Henderson, P.A. 2011. Commoness and rarity. Pages 97–104. In: Magurran, A.E., & McGill, B.J. (eds), *Biological diversity: Frontiers in Measurement and Assessment*. Oxford University Press.

Magurran, A.E., Baillie, S.R., Buckland, S.T., Dick, J.McP., Elston, D., Scott, E.M., Smith, R.I., Somerfield, P., & Watt, A.D. 2010. Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in Ecology and Evolution*, **25**, 574–582.

Mao, C.X., & Colwell, R.K. 2005. Estimation of species richness: mixture models, the role of rare species, and inferential challenges. *Ecology*, **86**, 1143–1153.

Marchant, J.H., Hudson, R., Carter, S.P., & Whittington, S.P. 1990. *Population trends in British breeding birds*. British Trust for Ornithology/Nature Conservancy Council, Tring.

Marques, F.F.C., & Buckland, S.T. 2003. Incorporating covariates into standard line transect analyses. *Biometrics*, **59**, 924–935.

Marquet, P.A., Keymer, J.E., & Cofré, H. 2003. Breaking the stick in space: of niche models, metacommunities and patterns in the relative abundance of species. Pages 64–84. In: Blackburn, T.M., & Gaston, K.J. (eds), *Macroecology: concepts and consequences.* Blackwell Science Ltd for the British Ecological Society.

Martino, S., & Rue, H. 2010. Case studies in Bayesian computation using `INLA`. Pages 99–114. In: Mantovan, P., & Secchi, P. (eds), *Complex data modeling and computationally intensive statistical methods.* Springer, Italia.

Maurer, B.A., & McGill, B.J. 2011. Measurement of species diversity. Pages 55–65. In: Magurran, A.E., & McGill, B.J. (eds), *Biological diversity: Frontiers in Measurement and Assessment.* Oxford University Press.

May, R.M. 1988. How many species are there on earth? *Science*, **241**, 1441–1449.

McCann, K. 2007. Protecting biostructure. *Nature*, **446**, 29.

McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., Dornelas, M., Enquist, B. J., Green, J. L., He, F., Hurlbert, A. H., Magurran, A. E., Marquet, P. A., Maurer, B. A., Ostling, A., Soykan, C. U., Ugland, K. I., & White, E. P. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters*, **10**, 995–1015.

McGill, B.J. 2011. Measuring the spatial structure of biodiversity. Pages 152–171. In: Magurran, A.E., & McGill, B.J. (eds), *Biological diversity: Frontiers in Measurement and Assessment.* Oxford University Press.

McGill, B.J., & Collins, C. 2003. A unified theory for macroecology based on spatial patterns of abundance. *Evolutionary Ecology Research*, **5**, 469–492.

McIntosh, R. P. 1967. An index of diversity and the relation of certain concepts of diversity. *Ecology*, **48**, 392–404.

Morgan, J. 1962. The anatomy of income distribution. *The Review of Economics and Statistics*, **44**, 270–283.

Morlon, H., White, E.P., Etienne, R.S., Green, J.L., Ostling, A., Alonso, D., Enquist, B.J., He, F., Hurlbert, A., Magurran, A.E., Maurer, B.A., McGill, B.J., Olff, H., Storch, D., & Zillio, T. 2009. Taking species abundance distributions beyond individuals. *Ecology Letters*, **12**, 488–501.

Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R., & Simpson, I.C. 2011. *Final report for LCM2007 - the new UK land cover map.* Technical report. Centre for Ecology and Hydrology.

Murrell, D.J., & Law, R. 2003. Heteromyopia and the spatial coexistence of similar competitors. *Ecology Letters*, **6**, 48–59.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., de Fonseca, G.A.B., & Kent, J. 2000. Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Naeem, S., Thompson, L.J., Lawler, S.P., Lawton, J.H., & Woodfin, R.M. 1994. Declining biodiversity can alter the performance of ecosystems. *Nature*, **368**, 734–737.

Newson, S.E., Woodburn, R., Noble, D.G., & Baillie, S.R. 2005. Evaluating the Breeding Bird Survey for producing national population size and density estimates. *Bird Study*, **52**, 42–54.

Newson, S.E., Evans, K.L., Noble, D.G., Greenwood, J.J.D., & Gaston, K.J. 2008. Use of distance sampling to improve estimates of national population sizes for common and widespread breeding birds in the UK. *Journal of Applied Ecology*, **45**, 1330–1338.

Newson, S.E., Ockendon, N., Joys, A., Noble, D.G., & Baillie, S.R. 2009. Comparison of habitat-specific trends in abundance of breeding birds in the UK. *Bird Study*, **56**, 233–243.

Nichols, J.D., Boulinier, T., Hines, J.E., Pollock, K.H., & Sauer, J.R. 1998. Inference methods for spatial variation in species richness and community composition when not all species are detected. *Conservation Biology*, **12**, 1390–1398.

Nichols, R. 2001. Gene tree and species tree are not the same. *Trends in Ecology and Evolution*, **16**, 358–364.

Norse, E.A., Rosenbaum, K.L., Wilcove, D.S., Romme, W.H., Johnston, D.W., & Stout, M.L. 1986. *Conserving biological diversity in our national forests.* The Wilderness Society, Washington, DC.

Noss, R.F. 1990. Indicators for monitoring biodiversity: a hierarchical approach. *Conservation Biology*, **4**, 355–364.

Orians, G.H., & Milewski, A.V. 2007. Ecology of Australia: the effects of nutrient-poor soils and intense fires. *Biological Reviews*, **82**, 393–423.

Parmesan, C. 2006. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution and Systematics*, **37**, 637–669.

Patil, G.P., & Taillie, C. 1982. Diversity as a concept and its measurement. *J. Am. Stat. Assoc.*, **77**, 548–561.

Pavoine, S., Ollier, S., & Dufour, A.-B. 2005. Is the originality of species measurable? *Ecology Letters*, **8**, 579–586.

Pearson, D.L. 1994. Selecting indicator taxa for quantitive assessment of biodiversity. *Philosophical Transactions of the Royal Society B*, **345**, 75–79.

Peet, R.K. 1974. The measurement of species diversity. *Annual Review of Ecology and Systematics*, **5**, 285–307.

Pereira, H., & Cooper, H.D. 2006. Towards the global monitoring of biodiversity change. *Trends in Ecology & Evolution*, **21**, 123–129.

Perry, G.L.W., Miller, B.P., & Enright, N.J. 2006. A comparison of methods for the statistical analysis of spatial point patterns in plant ecology. *Plant Ecology*, **187**, 59–82.

Pielou, E.C. 1966a. The measurement of diversity in different types of biological collections. *Journal of Theoretical Biology*, **13**, 131–144.

Pielou, E.C. 1966b. Shannon's formula as a measure of specific diversity: its use and misuse. *The American Naturalist*, **100**, 463–465.

Pielou, E.C. 1969. *An Introduction to Mathematical Ecology*. John Wiley & Sons, New York.

Pielou, E.C. 1975. *Ecological Diversity*. John Wiley & Sons, New York.

Poesel, A., Debaelsteen, T., & Pedersen, S.B. 2004. Dawn song of male blue tits as a predictor of competitiveness in midmorning singing interaction. *Acta Ethologica*, **6**, 65–71.

Preston, F.W. 1948. The Commonness, and rarity, of species. *Ecology*, **29**, 254–283.

R core development team. 2011. *R: A language and environment for statistical computing.* R foundation for statistical computing, Vienna, Austria. http://www.R-project.org.

Rabinowitz, D., Cairns, S., & Dillon, T. 1986. Seven forms of rarity and their frequency in the flora of the British Isles. Pages 182–204. In: Soulé, M.E. (ed), *Consevation biology. The science of scarcity and diversity.* Sinauer, Sunderland, Massachusetts.

Rajala, T.A., & Illian, J.B. 2012. A family of spatial biodiversity measures based on graphs. *Environmental and Ecological Statistics.* Online preview DOI 10.1007/s10651-012-0200-9.

Rao, C.R. 1982. Diversity and Dissimilarity Coefficients: A Unified Approach. *Theoretical Population Biology*, **21**, 24–48.

Read, T.R.C., & Cressie, N.A.C. 1988. *Goodness-of-Fit Statistics for Discrete Multivariate Data.* Springer.

Renwick, A.R., Johnston, A., Joys, A., Newson, S.E., Noble, D.G., & Pearce-Higgins, J.W. 2012. Composite bird indicators robust variation in species selection and habitat specificity. *Ecological Indicators*, **18**, 200–207.

Rényi, A. 1965. On the foundations of information theory. *Review of the International Statistical Institute*, **33**, 1–14.

Richardson, D.M., Cowling, R.M., Laymont, B.B., & van Hensbergen, H.J. 1995. Coexistence of *Banksia* species in southwestern Australia: the role of regional and local processes. *Journal of Vegetation Science*, **6**, 329–342.

Ricotta, C. 2003. On parametric evenness measures. *Journal of Theoretical Ecology*, **222**, 189–197.

Ricotta, C. 2005. Through the jungle of biological diversity. *Acta biotheoretica*, **53**, 29–38.

Ricotta, C., & Avena, G.C. 2002. On the information-theoretical meaning of Hill's parametric evenness. *Acta Biotheoretica*, **50**, 63–71.

Ripley, B.D. 1977. Modelling spatial patterns. *Journal of the Royal Statistical Society B*, **39**, 172–212.

Riseley, K., Renwick, A.R., Dadam, D., Eaton, M.A., Johnston, A., Baillie, S.R., Musgrove, A.J., & Noble, D.G. 2011. *The breeding bird survey 2010.* BTO Research Report No 597, British Trust for Ornithology, Thetford.

Robert, C. P., & Casella, G. 1999. *Monte Carlo Statistical Methods.* Springer, New York.

Rodrigues, A.S.L, Gregory, R.D., & Gaston, K.J. 2000. Robustness of reserve selection procedures under temporal species turnover. *Proceedings of the Royal Society B*, **267**, 49–55.

Rosenzweig, Michael L. 1995. *Species Diversity in Space and Time.* Cambridge University Press.

Rousseau, R., Van Hecke, P., Nijssen, D., & Bogaert, J. 1999. The relationship between diversity profiles, evenness and species richness based on partial ordering. *Environmental and Ecological Statistics*, **6**, 211–223.

Routledge, R.D. 1983. Evenness indices: are any admissible. *Oikos*, **40**, 149–151.

Rue, H., & Held, L. 2005. *Gaussian Markov random fields: theory and applications.* Chapman & Hall, CRC Press, London.

Rue, H., & Martino, S. 2007. Approximate Bayesian inference for hierarchical Gaussian Markov random field models. *Journal of Statistical Planning and Inference*, **137**, 3177–3192.

Rue, H., Martino, S., & Chopin, N. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society B*, **71**, 319–392.

Ruiz-Cárdenas, R., Krainski, E.T., & Rue, H. 2012. Direct fitting of dynamic models using integrated nested Laplace approximations - INLA. *Computational Statistics and Data Analysis*, **56**, 1808–1828.

Sanders, H. 1968. Marine benthic diversity: a comparative study. *The Amercian Naturalist*, **102**, 243–282.

Scholes, R.J., Walters, M., Turak, E., Saarenmaa, H., Heip, C.H.R., Tuama, É.Ó., Faith, D.P., Mooney, H.A., Ferrier, S., Jongman, R.H.G., Harrison, I.J., Yahara, T., Pereira, H.M., Larigauderie, A., & Geller, G. 2012. Building a global observing system for biodiversity. *Current Opinion in Environmental Sustainability*, **4**, 139–146.

Schupp, E.W. 1992. The Janzen-Connell model for tropical tree diversity: population implications and the importance of spatial scale. *The American Naturalist*, **140**, 526–530.

Shannon, C.E. 1948. A mathematical theory of information. *Bell System Technical Journal*, **27**, 379–423.

Shen, T.-J., Chao, A., & Lin, C.F. 2003. Predicting the number of new species in further taxonomic sampling. *Ecology*, **84**, 789–804.

Shimatani, K., & Kubota, Y. 2004. Quantitative assessment of multispecies spatial pattern with high species diversity. *Ecological Research*, **19**, 149–163.

Simberloff, D. 1972. Properties of the rarefaction diversity measurement. *The Amercian Naturalist*, **106**, 414–418.

Simpson, D., Lindgren, F., & Rue, H. 2010. *In order to make spatial statistics computationally feasible, we need to forget about the covariance function.* Technical report. Norwegian University of Science and Technology, Trondheim, Norway.

Simpson, D., Lindgren, F., & Rue, H. 2011a. *Think continuous: Markovian Gaussian models in spatial statistics.* Technical report. Norwegian University of Science and Technology, Trondheim, Norway.

Simpson, D., Lindgren, F., & Rue, H. 2011b. *Fast approximate inference with INLA: the past, the present and the future.* Technical report. Norwegian University of Science and Technology, Trondheim, Norway.

Simpson, E.H. 1949. Measurement of diversity. *Nature*, **163**, 688.

Siriwardena, G.M., Baillie, S.R., Buckland, S.T., Fewster, R.M., Marchant, J.H., & Wilson, J.D. 1998. Trends in the abundance of farmland birds: a quantitative comparison of smoothed Common Birds Census indices. *Journal of Applied Ecology*, **35**, 24–43.

Smith, B., & Wilson, J.B. 1996. A consumer's guide to evenness indices. *OIKOS*, **76**, 70–82.

Smith, M. D., & Knapp, A. K. 2003. Dominant species maintain ecosystem function with non-random species loss. *Ecology Letters*, **6**, 509–517.

Southwood, R., & Henderson, P.A. 2000. *Ecological Methods.* Blackwell Science, Oxford.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., & Van der Linde, A. 2002. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society B*, **64**, 583–616.

Stein, M.L. 1999. *Interpolation of spatial data: some theory for kriging.* Springer, New York.

Storch, D., Marquet, P.A., & Brown, J.H. 2007. Introduction: scaling biodiversity — what is the problem? Pages 1–11. In: Storch, D., Marquet, P.A., & Brown, J.H. (eds), *Scaling biodiversity.* Cambridge University Press, New York.

Studeny, A.C., Buckland, S.T., Illian, J.B., Johnston, A., & Magurran, A.E. 2011. Goodness-of-fit measures of evenness: a new tool for exploring changes in community structure. *Ecosphere*, **2**(2), art.15.

Studeny, A.C., Buckland, S.T., Harrison, P.J., Illian, J.B., Magurran, A.E., & Newson, S.E. In prep. Fine-tuning the assessment of large-scale temporal trends in biodiversity using the example of British breeding birds.

Studeny, W. 1973. *Bemerkungen zur dynamischen statistischen Konzentrationsanalyse.* Ph.D. thesis, Staatswirtschaftliche Fakultät, Ludwig-Maximillians-Universität, München.

Sugihara, G. 1980. Minimal community structure: and explanation of species abundance patterns. *The American Naturalist*, **116**, 770–787.

Ter Steege, H., Pitman, N., Sabatier, D., Castellanos, H., Van Der Hout, P., Daly, D.C., Silveira, M., Phillips, O., Vaquez, R., Van Andel, T., Duivenvoorden, J., Adalardo de Oliveira, A., Ek, R., Lilwah, R., Thomas, R., Van Essen, J., Baider, C., Maas, P., Mori, S., Terborgh, J., Nunez Vargas, P., Mogollon, H., & Morawetz, W. 2003. A spatial model of tree α-diversity and tree density for the Amazon. *Biodiversity and Conservation*, **12**, 2255–2277.

Thompson, S.K. 1992. *Sampling.* John Wiley & Sons, Hoboken, New Jersey.

Thomson, J.R., Kimmerer, W.J., Brown, L.R., Newman, K.B., Mac Nally, R., Bennett, W.A., Feyrer, F., & Fleishman, E. 2010. Bayesian change point analysis of abundance trends for pelagic fishes in the upper San Francisco Estuary. *Ecological Applications*, **20**, 1431–1448.

Thuiller, W., Lavorel, S., Araújo, M.B., Sykes, M.T., & Prentice, I.C. 2005. Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences*, **102**, 8245–8250.

Tierney, L., & Kadane, J.B. 1986. Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, **81**, 82–86.

Tilman, D. 1994. Competition and biodiversity in spatially structured habitats. *Ecology*, **75**, 2–16.

Tilman, D. 2001. Functional diversity. Pages 109–120. In: Levin, S.A. (ed), *Encyclopedia of Biodiversity*, vol. III. Academic Press, San Diego.

Tilman, D., Knops, J., Wedin, D., Reich, P., & Siemann, E. 1997. The influence of functional diversity and composition on ecosystem processes. *Science*, **29**, 1300–1302.

Tilman, D., Reich, P.B., & Knops, J.M.H. 2006. Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature*, **441**, 626–632.

Tokeshi, M. 1990. Niche apportionment or random assortment: species abundance patterns revisited. *Journal of Animal Ecology*, **59**, 1129–1146.

Tokeshi, M. 1996. Power fraction: a new explanation of relative abundance patterns in species-rich assemblages. *OIKOS*, **75**, 543–550.

Tóthmérész, B. 1995. Comparison of different methods for diversity ordering. *Journal of Vegetation Science*, **6**, 283–290.

Tsallis, C. 1988. Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, **52**, 479–487.

Tuomisto, H. 2010. A diversity of beta diversities: straightening up a concept gond awry. Part 1: Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, **33**, 2–22.

United Nations. 2000. *UN Millenium Declaration*. United Nations, New York. http://www.un.org/millennium/declaration/ares552e.pdf.

van Strien, A.J., Soldaat, L.L., & Gregory, R.D. 2012. Desirable mathematical properties of indicators for biodiversity change. *Ecological Indicators*, **14**, 202–208.

Vellend, M. 2001. Do commonly used indices of β-diversity measure species turnover. *Journal of Vegetation Science*, **12**, 545–552.

Vellend, M., Cornwell, W.K., Magnuson-Ford, K., & Moors, A.Ø. 2011. Measuring phylogenetic biodiversity. Pages 194–207. In: Magurran, A.E., & McGill, B.J. (eds), *Biological diversity: Frontiers in Measurement and Assessment.* Oxford University Press.

Vickery, J.A., Bradbury, R.B., Henderson, I.G., Eaton, M.A., & Grice, P.V. 2004. The role of agri-environment schemes and farm management practices in reversing the decline of farmland birds in England. *Biological Conservation*, **119**, 19–39.

Vilà, M., & Sardans, J. 1999. Plant competition in mediterranean-type vegetation. *Journal of Vegetation Science*, **10**, 281–294.

Walpole, M., Almond, R.E.A., Besancon, C., Butchart, S.H.M., Campbell-Lendrum, D., Carr, G.M., Collen, B., Colette, L., Davidson, N.C., Dulloo, E., Fazel, A.M., Galloway, J.N., Gill, M., Goverse, T., Hockings, M., Leaman, D.J., Morgan, D.H.W., Revenga, C., Rickwood, C.J., Schutyser, F., Simons, S., Stattersfield, A.J., Tyrell, T.D., Vie, J.-C., & Zimsky, M. 2009. Tracking progress toward the 2010 biodiversity target and beyond. *Science*, **325**, 1503–1504.

Warwick, R.M., & Clarke, K.R. 1994. Relearning the ABC: taxonomic changes and abundance/biomass relationships in disturbed benthic communities. *Marine Biology*, **118**, 739–744.

Warwick, R.M., & Clarke, K.R. 1995. New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. *Marine Ecology Progress Series*, **129**, 301–305.

Watt, A.S. 1947. Pattern and process in the plant community. *Journal of Ecology*, **35**, 1–22.

Whittaker, R.H. 1972. Evolution and measurement of species diversity. *Taxon*, **12**, 250–260.

Wiegand, T., & Moloney, K.A. 2004. Rings, circles, and null-models for point pattern analysis in ecology. *Oikos*, **104**, 209–229.

Wiegand, T., Gunatilleke, S., & Gunatilleke, N. 2007. Species associations in a heterogeneous Sri Lankan dipterocarp forest. *The American Naturalist*, **170**, E77–D95.

Wood, S.N. 2006. *Generalized additive models: an Introduction with R*. Chapman & Hall, CRC Press, London.

Wood, S.N. 2008. Fast stable direct fitting and smoothness selection for generalized additive models. *Journal of the Royal Statistical Society B*, **70**, 495–518.

Wright, J.S. 2002. Plant divesity in tropical forests: a review of mechanisms of species coexistence. *Oecologica*, **130**, 1–14.

Yoccoz, N. G., Nichols, J.D., & Boulinier, Th. 2001. Monitoring of biological diversity in space and time. *Trends in Ecology and Evolution*, **16**, 446–453.

Zimmerman, D.L., & Stein, M. 2010. Classical geostatistical methods. Pages 29–44. In: Gelfand, A.E., Diggle, P.J., Fuentes, M., & Guttorp, P. (eds), *Handbook of Spatial Statistics*. Chapman & Hall, CRC Press, London.

Zuur, A.F., Ieno, E.N., Walker, N., Saveliev, A.A., & Smith, G.M. 2009. *Mixed effects models and extensions in ecology with R*. Springer, New York. Chapter 11.

Appendices

Appendix A

The Convention on Biological Diversity's (CBD) biodiversity targets

The following are relevant excerpts from the Convention on Biological Diversity (CBD, 1992), in particular the original formulation of the 2010 target as well as the AICHI targets that have now replaced it.

From the preamble of the Convention on Biological Diversity

> The Contracting Parties,
>
> [...] Conscious also of the **importance of biological diversity** for evolution and for maintaining life sustaining systems of the biosphere,
> [...] Concerned that biological diversity is being significantly reduced by certain human activities,
> **Aware of the general lack of information and knowledge regarding biological diversity** and of the urgent **need to develop scientific, technical and institutional capacities** to provide the basic understanding upon which to **plan and implement appropriate measures**,
> Noting that it is vital to **anticipate, prevent and attack the causes of significant reduction or loss** of biological diversity at source,
> Noting also that where there is a threat of significant reduction or loss of biological diversity, lack of full scientific certainty should not be used as a reason for postponing measures to avoid or minimize such a threat,
> [...] Stressing the importance of [...] **international, regional and global cooperation** among States and intergovernmental organizations and the non-governmental sector for the conservation of biological diver-

sity and the sustainable use of its components,

[...] Determined to conserve and sustainably use biological diversity for the benefit of present and future generations,

Have agreed as follows [...]

The CBD's definition of biological diversity

Article 2. Use of Terms

[...] 'Biological diversity' means the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems. [...] 'Ecosystem' means a dynamic complex of plant, animal and micro-organism communities and their non-living environment interacting as a functional unit. [...]

The 2010 target

Parties commit themselves to a more effective and coherent implementation of the [...] objectives of the Convention, to achieve by 2010 a **significant reduction of the current rate of biodiversity loss at the global, regional and national level** as a contribution to poverty alleviation and to the benefit of all life on earth.

In particular,

Goal 3. National biodiversity strategies and action plans and the integration of biodiversity concerns into relevant sectors serve as an effective framework for the implementation of the objectives of the Convention. [...]
3.1 Every Party has **effective national strategies, plans and programmes in place to provide a national framework** for implementing the three objectives of the Convention and to set clear national priorities [...]

The strategic plan for biodiversity 2011-2020, including the Aichi biodiversity targets

Strategic Goal A: Address the underlying causes of biodiversity loss by mainstreaming biodiversity across government and society

Target 1

By 2020, at the latest, people are aware of the values of biodiversity and the steps they can take to conserve and use it sustainably.

Target 2

By 2020, at the latest, biodiversity values have been integrated into national and local development and poverty reduction strategies and planning processes and are being incorporated into national accounting, as appropriate, and reporting systems.

Target 3

By 2020, at the latest, incentives, including subsidies, harmful to biodiversity are eliminated, phased out or reformed in order to minimize or avoid negative impacts, and positive incentives for the conservation and sustainable use of biodiversity are developed and applied, consistent and in harmony with the Convention and other relevant international obligations, taking into account national socio economic conditions.

Target 4

By 2020, at the latest, Governments, business and stakeholders at all levels have taken steps to achieve or have implemented plans for sustainable production and consumption and have kept the impacts of use of natural resources well within safe ecological limits.

Strategic Goal B: Reduce the direct pressures on biodiversity and promote sustainable use

Target 5

By 2020, the rate of loss of all natural habitats, including forests, is at least halved and where feasible brought close to zero, and degradation and fragmentation is significantly reduced.

Target 6

By 2020 all fish and invertebrate stocks and aquatic plants are managed and harvested sustainably, legally and applying ecosystem based approaches, so that overfishing is avoided, recovery plans and measures are in place for all depleted species, fisheries have no significant adverse impacts on threatened species and vulnerable ecosystems and the impacts of fisheries on stocks, species and ecosystems are within safe ecological limits.

Target 7

By 2020 areas under agriculture, aquaculture and forestry are managed sustainably, ensuring conservation of biodiversity.

Target 8

By 2020, pollution, including from excess nutrients, has been brought to

levels that are not detrimental to ecosystem function and biodiversity.

Target 9

By 2020, invasive alien species and pathways are identified and prioritized, priority species are controlled or eradicated, and measures are in place to manage pathways to prevent their introduction and establishment.

Target 10

By 2015, the multiple anthropogenic pressures on coral reefs, and other vulnerable ecosystems impacted by climate change or ocean acidification are minimized, so as to maintain their integrity and functioning.

Strategic Goal C: To improve the status of biodiversity by safeguarding ecosystems, species and genetic diversity

Target 11

By 2020, at least 17 per cent of terrestrial and inland water, and 10 per cent of coastal and marine areas, especially areas of particular importance for biodiversity and ecosystem services, are conserved through effectively and equitably managed, ecologically representative and well connected systems of protected areas and other effective area-based conservation measures, and integrated into the wider landscapes and seascapes.

Target 12

By 2020 the extinction of known threatened species has been prevented and their conservation status, particularly of those most in decline, has been improved and sustained.

Target 13

By 2020, the genetic diversity of cultivated plants and farmed and domesticated animals and of wild relatives, including other socio-economically as well as culturally valuable species, is maintained, and strategies have been developed and implemented for minimizing genetic erosion and safeguarding their genetic diversity.

Strategic Goal D: Enhance the benefits to all from biodiversity and ecosystem services

Target 14

By 2020, ecosystems that provide essential services, including services related to water, and contribute to health, livelihoods and well-being, are restored and safeguarded, taking into account the needs of women, indigenous and local communities, and the poor and vulnerable.

Target 15

By 2020, ecosystem resilience and the contribution of biodiversity to carbon stocks has been enhanced, through conservation and restoration, including restoration of at least 15 per cent of degraded ecosystems, thereby contributing to climate change mitigation and adaptation and to combating desertification.

Target 16

By 2015, the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization is in force and operational, consistent with national legislation.

Strategic Goal E: Enhance implementation through participatory planning, knowledge management and capacity building

Target 17

By 2015 each Party has developed, adopted as a policy instrument, and has commenced implementing an effective, participatory and updated national biodiversity strategy and action plan.

Target 18

By 2020, the traditional knowledge, innovations and practices of indigenous and local communities relevant for the conservation and sustainable use of biodiversity, and their customary use of biological resources, are respected, subject to national legislation and relevant international obligations, and fully integrated and reflected in the implementation of the Convention with the full and effective participation of indigenous and local communities, at all relevant levels.

Target 19

By 2020, knowledge, the science base and technologies relating to biodiversity, its values, functioning, status and trends, and the consequences of its loss, are improved, widely shared and transferred, and applied.

Target 20

By 2020, at the latest, the mobilization of financial resources for effectively implementing the Strategic Plan for Biodiversity 2011-2020 from all sources, and in accordance with the consolidated and agreed process in the Strategy for Resource Mobilization, should increase substantially from the current levels. This target will be subject to changes contingent to resource needs assessments to be developed and reported by Parties.

Appendix B

Evaluating the goodness-of-fit based measures with respect to the criteria by Smith & Wilson (1996)

Requirement 1: Independence of species richness

The evenness index should be invariant if multiples of a species abundance vector are considered. We use the same examples as in Smith & Wilson (1996), namely

```
> sad_1 <- c(1497, 1, 1, 1)
> sad_2 <- c(800, 400, 200, 100)
> sad_3 <- c(378, 376, 374, 372)
> models <- as.matrix(rbind(sad_1, sad_2, sad_3), nrow = 3)
> S <- ncol(models)
> abd <- models
> rel.abd <- models/rowSums(models).
```

The family of goodness-of-fit based measures is calculatd for $-3 < \lambda < 3$ for the SADs and for triples of the original SADs (see Fig B.1).

```
> lambda <- seq(-3, 3, by = 0.05)
> GoF <- t(apply(rel.abd, 1, gof, lambda = lambda, S = S))

> triples <- t(apply(models, 1, rep, times = 3))
> S <- ncol(triples)
> abd <- triples
```

```
> rel.abd <- triples/rowSums(triples)
> GoF_triples <- t(apply(rel.abd, 1, gof, lambda = lambda, S = S))
```
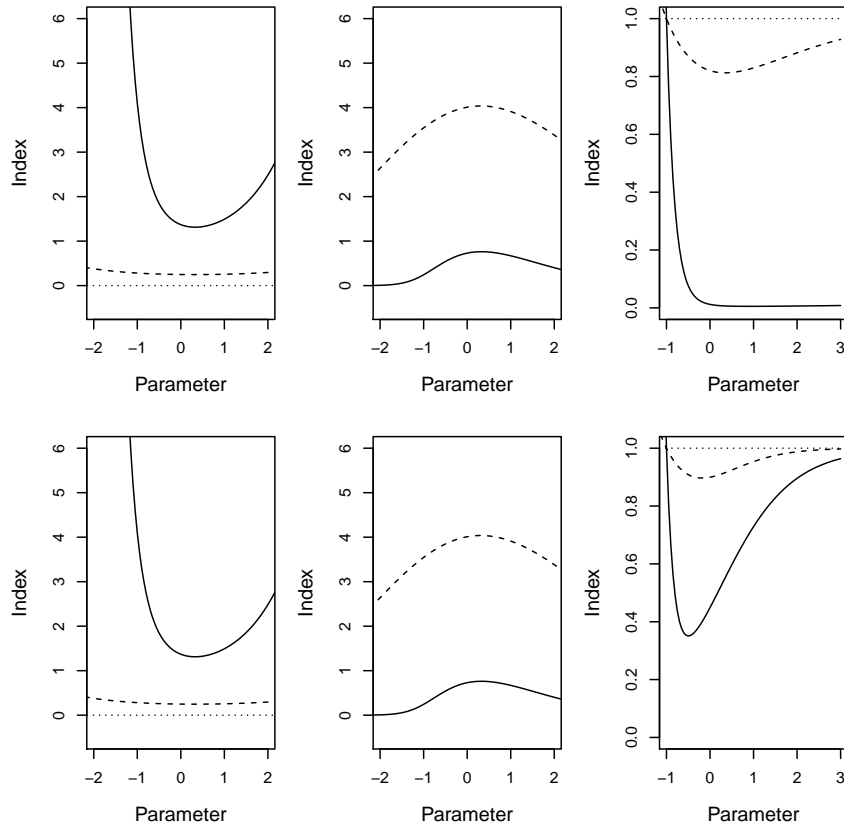

FIGURE B.1 – Smith and Wilson's first requirement for evenness measures — independence of species richness. The measure should not change if the species abundance vector is merely repeated. The upper row shows evenness profiles for the untransformed goodness-of-fit measures $I_{\mathbf{p}}(\lambda)$ for the three examples of species abundance vectors in Smith & Wilson (1996), the reciprocal $1/I_{\mathbf{p}}(\lambda)$ and the standardised transformation $\rho_{\mathbf{p}}(\lambda)$ (from left to right). The lower row shows the evenness profiles, evaluated for triples of the original abundance vectors.

Further repetitions of the original SADs are considered (up to a 40-fold increase in the original number of species $S = 4$) and for each repetition the goodness-of-fit measure for $\lambda = -0.5$ is calculated (the λ value was chosen because it results in a metric on the simplex). Fig. B.2 shows the untransformed goodness-of-fit index as well as the two transformations plotted against number of species. Complete independence of S results in a horizontal line.

```
> lambda <- -0.5
> models <- as.matrix(rbind(sad_1, sad_2, sad_3), nrow = 4)
> times <- c(1, 2, 3, 5, 10, 20, 40)
> species <- c(4, 8, 12, 20, 40, 80, 160)
```

```
> GoF <- matrix(0, nrow = nrow(models), ncol = length(species))
> rho <- matrix(0, nrow = nrow(models), ncol = length(species))
> maxi <- 1/(lambda * (1 + lambda)) * (species^lambda - 1)
> for (i in 1:length(species)) {
+     models <- t(apply(models, 1, rep, times = times[i]))
+     abd <- models
+     rel.abd <- abd/rowSums(models)
+     GoF[, i] <- t(apply(rel.abd, 1, gof, lambda = lambda, S = species[i]))
+     rho[, i] <- 1 - GoF[, i]/maxi[i]
+     models <- as.matrix(rbind(sad_1, sad_2, sad_3), nrow = 4)
+ }
```
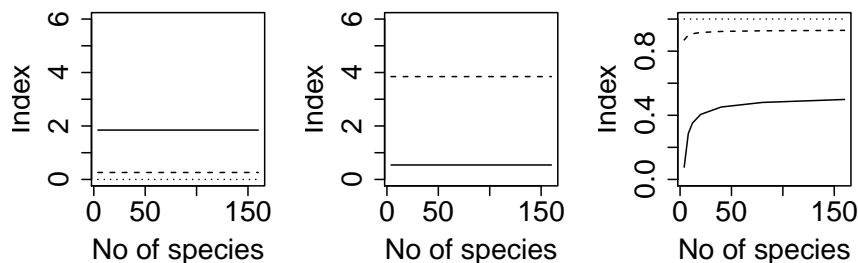

FIGURE B.2 – Smith and Wilson's first requirement for evenness measures, for systematic increases in species richness. Evaluated for $\lambda = -0.5$, the untransformed goodness-of-fit measures $I_{\mathbf{p}}(\lambda)$, its reciprocal and the standardised $\rho_{\mathbf{p}}(\lambda)$ (from left to right) are plotted for the three example SADs against an increase in species richness. S is increased twofold, threefold, fivefold, tenfold, twentyfold and fortyfold by the corresponding repetitions of the three examples of species abundance vectors.

Requirement 2: Loss of abundance of a rare species

If the rarest species decreases in abundance, the evenness index should drop.

```
> sad_1 <- c(80, 40, 20, 10, 1)
> sad_2 <- c(80, 40, 20, 10, 0.5)
> models <- as.matrix(rbind(sad_1, sad_2), nrow = 2)
> abd <- models
> rel.abd <- models/rowSums(models)
> S <- 5
> maxi <- 1/(lambda * (1 + lambda)) * (S^lambda - 1)
> GoF <- t(apply(rel.abd, 1, gof, lambda = lambda, S = S))
> rho <- 1 - GoF/maxi
> GoF
```

```
        sad_1      sad_2
[1,] 0.520288 0.5572212

> 1/GoF

        sad_1      sad_2
[1,] 1.922012 1.794619

> rho

         sad_1      sad_2
[1,] 0.7646975 0.7479943
```

Requirement 3: Addition of a rare species

If a rare species is added to the assemblage, the evenness index should decrease.

```
> sad_1 <- c(80, 40, 20, 10)
> sad_2 <- c(80, 40, 20, 10, 0.5)
> rel.abd_1 <- sad_1/sum(sad_1)
> rel.abd_2 <- sad_2/sum(sad_2)
> maxi_1 <- 1/(lambda * (1 + lambda)) * (4^lambda - 1)
> GoF_1 <- gof(rel.abd_1, lambda = lambda, S = 4)
> rho_1 <- 1 - GoF_1/maxi_1
> maxi_2 <- 1/(lambda * (1 + lambda)) * (5^lambda - 1)
> GoF_2 <- gof(rel.abd_2, lambda = lambda, S = 5)
> rho_2 <- 1 - GoF_2/maxi_2
> GoF_1

[1] 0.2599164

> GoF_2

[1] 0.5572212

> 1/GoF_1

[1] 3.84739

> 1/GoF_2

[1] 1.794619

> rho_1

[1] 0.8700418

> rho_2

[1] 0.7479943
```

Requirement 4: Invariance under multiplication of the SAD by a constant

The evenness index should not change when abundances of all species are increased or decreased by the same (constant) factor. This is interpreted as the measure being unaffected by a change in units used.

```
> sad_1 <- c(1, 2, 3)
> sad_2 <- c(100, 200, 300)
> models <- as.matrix(rbind(sad_1, sad_2), nrow = 2)
> abd <- models
> rel.abd <- models/rowSums(models)
> maxi <- 1/(lambda * (1 + lambda)) * (3^lambda - 1)
> GoF <- t(apply(rel.abd, 1, gof, lambda = lambda, S = 3))
> rho <- 1 - GoF/maxi
> GoF

         sad_1      sad_2
[1,] 0.09086446 0.09086446

> 1/GoF

      sad_1   sad_2
[1,] 11.0054 11.0054

> rho

         sad_1     sad_2
[1,] 0.9462531 0.9462531
```

Features 5 and 6: Maximality at perfect evenness

The evenness should attain its maximum at perfect evenness. The maximum value should preferably be 1.

```
> sad <- c(375, 375, 375, 375)
> rel.abd <- sad/sum(sad)
> maxi <- 1/(lambda * (1 + lambda)) * (4^lambda - 1)
> GoF <- gof(rel.abd, lambda = lambda, S = 4)
> rho <- 1 - GoF/maxi
> GoF
```

```
[1] 0

> 1/GoF

[1] -Inf

> rho

[1] 1
```

Features 7 and 9: Minimality if SAD is as unequal as possible

The evenness index should attain its minimum for any number of species if an SAD is as unequal as possible. This minimum should preferably be 0.

```
> sad_1 <- c(999, 1)
> sad_2 <- c(900, 100)
> sad_3 <- c(800, 200)
> sad_4 <- c(700, 300)
> sad_5 <- c(600, 400)
> sad_6 <- c(500, 500)
> models <- as.matrix(rbind(sad_1, sad_2, sad_3, sad_4, sad_5,
+     sad_6), nrow = 6)
> S <- 2
> rel.abd <- models/rowSums(models)
> maxi <- 1/(lambda * (1 + lambda)) * (S^lambda - 1)
> GoF <- t(apply(rel.abd, 1, gof, lambda = lambda, S = S))
> rho <- 1 - GoF/maxi
> GoF

        sad_1     sad_2     sad_3      sad_4      sad_5 sad_6
[1,] 1.083545 0.4222912 0.2052668 0.08437475 0.02025539     0

> 1/GoF

         sad_1    sad_2    sad_3    sad_4    sad_5 sad_6
[1,] 0.9228968 2.368034 4.871708 11.85189 49.36958  -Inf

> rho

          sad_1     sad_2     sad_3     sad_4     sad_5 sad_6
[1,] 0.07513673 0.6395519 0.8247938 0.9279816 0.9827109     1
```

Feature 8: Minimality if SAD is as unequal as possible

The evenness index should be close to its minimum for an SAD that is as unequal as we are likely to observe. ('close to its minimum' being less than 0.05)

```
> sad_1 <- c(1497, 1, 1, 1)
> S <- 4
> rel.abd <- sad_1/sum(sad_1)
> maxi <- 1/(lambda * (1 + lambda)) * (S^lambda - 1)
> GoF <- gof(rel.abd, lambda = lambda, S = S)
> rho <- 1 - GoF/maxi
> GoF

[1] 1.847082

> 1/GoF

[1] 0.5413946

> rho

[1] 0.07645917
```

Feature 10: Minimum attainable for any number of species

As one species' abundance goes to infinity, the species abundance vector in the limit, $(\infty, 1)$, should result in the evenness index being 0. (For numerical calculation, the abundance of the first species is set to a very high value in the following example.)

```
> sad <- c(1e+09, 1)
> S <- 2
> rel.abd <- sad/sum(sad)
> maxi <- 1/(lambda * (1 + lambda)) * (S^lambda - 1)
> GoF <- gof(rel.abd, lambda = lambda, S = S)
> rho <- 1 - GoF/maxi
> GoF

[1] 1.171483

> 1/GoF

[1] 0.8536186

> rho

[1] 7.634293e-05
```

Feature 11: Intermediate value for neither even nor uneven communities

SADs that are intuitively thought of as being neither particularly even nor uneven should lead to intermediate index values.

```
> sad_1 <- c(600, 450, 300, 150)
> sad_2 <- c(800, 400, 200, 100)
> models <- as.matrix(rbind(sad_1, sad_2), nrow = 2)
> abd <- models
> rel.abd <- models/rowSums(models)
> S <- 4
> maxi <- 1/(lambda * (1 + lambda)) * (S^lambda - 1)
> GoF <- t(apply(rel.abd, 1, gof, lambda = lambda, S = S))
> rho <- 1 - GoF/maxi
> GoF

          sad_1     sad_2
[1,] 0.1127611 0.2599164

> 1/GoF

        sad_1   sad_2
[1,] 8.868307 3.84739

> rho

          sad_1     sad_2
[1,] 0.9436195 0.8700418
```

Feature 12: Reasonable response to systematic changes

For a series of SADs that intuitively change in their degree of evenness, an evenness index should reasonably reflect this change. (Smith & Wilson (1996) consider a continuous response of an index as reasonable if it results in a convex curve when evaluated for the following sequence of abundance vectors (see Fig. B.3); other opinions exist.)

```
> sad_1 <- c(999, 1)
> sad_2 <- c(900, 100)
> sad_3 <- c(800, 200)
```

```
> sad_4 <- c(700, 300)
> sad_5 <- c(600, 400)
> sad_6 <- c(500, 500)
> sad_7 <- c(400, 600)
> sad_8 <- c(300, 700)
> sad_9 <- c(200, 800)
> sad_10 <- c(100, 900)
> sad_11 <- c(1, 999)
> models <- as.matrix(rbind(sad_1, sad_2, sad_3, sad_4, sad_5,
+     sad_6, sad_7, sad_8, sad_9, sad_10, sad_11), nrow = 11)
> rel.abd <- models/rowSums(models)
> S <- 2
> maxi <- 1/(lambda * (1 + lambda)) * (S^lambda - 1)
> GoF <- t(apply(rel.abd, 1, gof, lambda = lambda, S = S))
> rho <- 1 - GoF/maxi
```
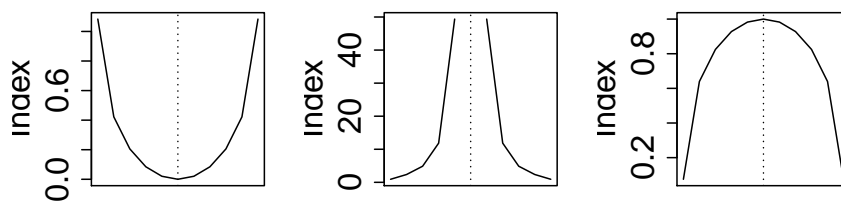

FIGURE B.3 – Smith and Wilson's feature for evenness measures. An evenness index should change continuously for a symmetric increase and decrease in evenness, resulting in a horseshoe-shaped curve

Feature 13: Symmetry with regards to dominant and rare species

An evenness index should not favour an assemblage with a certain number of dominant species and one rare species to an assemblage with the same number of species, but only one dominant species.

```
> sad_1 <- c(1000, 1000, 1000, 1)
> sad_2 <- c(1000, 1, 1, 1)
> rel.abd_1 <- sad_1/sum(sad_1)
> rel.abd_2 <- sad_2/sum(sad_2)
> maxi_1 <- 1/(lambda * (1 + lambda)) * (4^lambda - 1)
> GoF_1 <- gof(rel.abd_1, lambda = lambda, S = 4)
> rho_1 <- 1 - GoF_1/maxi_1
```

```
> maxi_2 <- 1/(lambda * (1 + lambda)) * (4^lambda - 1)
> GoF_2 <- gof(rel.abd_2, lambda = lambda, S = 4)
> rho_2 <- 1 - GoF_2/maxi_2
> GoF_1

[1] 0.4999668

> GoF_2

[1] 1.813541

> 1/GoF_1

[1] 2.000133

> 1/GoF_2

[1] 0.5514076

> rho_1

[1] 0.7500166

> rho_2

[1] 0.09322971
```

Feature 14: Skewness of SADs

An evenness index should give a lower value for assemblages with an excess in
either dominant or rare species than it does for an assemblage that is completely
balanced with regards to dominant and rare species.

```
> sad_1 <- c(1000, 1, 1, 1, 1, 1)
> sad_2 <- c(1000, 1000, 1000, 1, 1, 1)
> sad_3 <- c(1000, 1000, 1000, 1000, 1000, 1)
> S <- 6
> models <- as.matrix(rbind(sad_1, sad_2, sad_3), nrow = 3)
> abd <- models
> rel.abd <- models/rowSums(models)
> maxi <- 1/(lambda * (1 + lambda)) * (S^lambda - 1)
> GoF <- t(apply(rel.abd, 1, gof, lambda = lambda, S = S))
> rho <- 1 - GoF/maxi
> GoF
```

```
           sad_1     sad_2      sad_3
[1,] 2.113518 1.083588 0.3257897

> 1/GoF

           sad_1    sad_2     sad_3
[1,] 0.4731447 0.92286 3.069465

> rho

           sad_1     sad_2      sad_3
[1,] 0.1070925 0.5422117 0.8623622
```

Appendix C

R functions

C.1   Simulations of Tokeshi's models

The following R functions provide discrete algorithms for generating realisations of Tokeshi's niche models with a fixed number of species S and total abundance N. The latter determines the total available niche space. The functions have to be given the model specifications (S, N and k for the power fraction model) as well as one initial value b (the first niche division point); b can also be generated randomly.

Sampling function for the power fraction model

```
> powerfrac <- function(b, S, N, k) {
+     while (length(b) < S - 1) {
+         p <- c(1, b, N)
+         L <- length(p)
+         dis <- p[2:L] - p[1:(L - 1)]
+         if (k == 0) {
+             ind <- which(dis == 1)
+             if (length(ind) > 0) {
+                 dis[ind] <- 0
+                 dis[-ind] <- 1
+             }
+             else {
+                 dis <- rep(1, times = length(dis))
+             }
+             v <- sample(c(b, N), 1, prob = dis)
+         }
```

```
+           else {
+               dis <- (dis - 1)^k
+               v <- sample(c(b, N), 1, prob = dis)
+           }
+           index <- which(p == v)
+           a1 <- p[index - 1] + 1
+           a2 <- p[index] - 1
+           if (a1 == a2) {
+               b.new <- a1
+           }
+           else {
+               b.new <- sample(c(a1:a2), 1)
+           }
+           b <- sort(c(b, b.new))
+       }
+       res <- sort(c(b, N) - c(0, b), decreasing = T)
+       return(res)
+ }
```

Sampling function for the dominance preemption model

The function which gives a discrete equivalent for the dominance preemption model. Because of this discretisation, the resulting species abundance vector is slightly less uneven than expected in the case of the original model (which describes total available niche space as the continuous interval $[0, 1]$).

```
> dompreem <- function(b, S, N) {
+     if (b > N - S) {
+         return("initial b too big")
+     }
+     else {
+         while (length(b) < S - 1) {
+             if (N - max(b) > S - length(b) + 1) {
+                 k <- runif(1, 0.5, 1)
+                 b.new <- floor(k * (N - S + length(b) - max(b))) +
+                     max(b)
+                 b <- c(b, b.new)
+                 res <- sort(c(b, N) - c(0, b), decreasing = T)
+             }
```

```
+            else {
+                b.rem <- rep(1, times = S - length(b))
+                b <- c(b, b.rem)
+                res <- b
+            }
+        }
+        return(res)
+    }
+ }
```

Sampling function for the dominance decay model

```
> domdecay <- function(b, S, N) {
+     res <- c(b, N) - c(0, b)
+     while (length(b) < S - 1) {
+         d <- c(0, b, N)
+         ind <- which(res == max(res))
+         if (length(ind) == 1) {
+             b.new <- sample((d[ind] + 1):(d[ind + 1] - 1), 1)
+             b <- sort(c(b, b.new), decreasing = F)
+             res <- c(b, N) - c(0, b)
+         }
+         else {
+             ind <- sample(ind, 1)
+             b.new <- sample((d[ind] + 1):(d[ind + 1] - 1), 1)
+             b <- sort(c(b, b.new), decreasing = F)
+             res <- c(b, N) - c(0, b)
+         }
+     }
+     res <- sort(res, decreasing = T)
+     return(res)
+ }
```

C.2 R functions for the various diversity index families

Goodness-of-fit based measures

```
> gof <- function(lambda, S, rel.abd) {
+     res <- rep(0, times = length(lambda))
+     for (i in 1:length(lambda)) {
+         if (lambda[i] == -1) {
+             res[i] <- -1/S * sum(log(rel.abd)) - log(S)
+         }
+         else {
+             if (lambda[i] == 0) {
+                 res[i] <- log(S) + sum(rel.abd[rel.abd != 0] *
+                   log(rel.abd[rel.abd != 0]))
+             }
+             else {
+                 res[i] <- 1/(lambda[i] * (lambda[i] + 1)) * sum(rel.abd *
+                   ((rel.abd * S)^lambda[i] - 1))
+             }
+         }
+     }
+     return(res)
+ }
```

Hill's diversity numbers

```
> hill <- function(beta, rel.abd) {
+     res <- rep(0, times = length(beta))
+     for (i in 1:length(beta)) {
+         if (beta[i] == 1) {
+             res[i] <- exp(-sum(rel.abd * log(rel.abd)))
+         }
+         else {
+             res[i] <- (sum(rel.abd^beta[i]))^(1/(1 - beta[i]))
+         }
+     }
+     return(res)
+ }
```

The generalised mean

```
> gen.mean <- function(alpha, x) {
+     res <- rep(0, times = length(alpha))
+     for (i in 1:length(alpha)) {
+         if (alpha[i] == 0) {
+             res[i] <- exp(mean(log(x)))
+         }
+         else {
+             res[i] <- (mean(x^alpha[i]))^(1/alpha[i])
+         }
+     }
+     return(res)
+ }
```

Generalised entropy

```
> gen.entropy <- function(beta, rel.abd) {
+     res <- rep(0, times = length(beta))
+     ind <- which(beta == 0)
+     for (i in 1:length(beta)) {
+         if (beta[i] == 1) {
+             res[i] <- exp(-sum(rel.abd * log(rel.abd)))
+         }
+         else {
+             res[i] <- (sum(rel.abd^beta[i]))^(1/(1 - beta[i]))
+         }
+     }
+     return(((res/res[ind])^(1 - beta) - 1)/(-beta * (1 - beta)))
+ }
```

Appendix D

Another transformation of $I_{\mathbf{p}}(\lambda)$

The following figure shows the transformation of the goodness-of-fit based evenness measures $I_{\mathbf{p}}(\lambda)$ which gives Simpson's index $-\log D$ for $\lambda = 1$. Here, this transformation is applied to order Tokeshi's models (scenario 1). The profile plots are no longer continuous, however, the position of the asymptote gives some indication of evenness of the underlying species proportions.
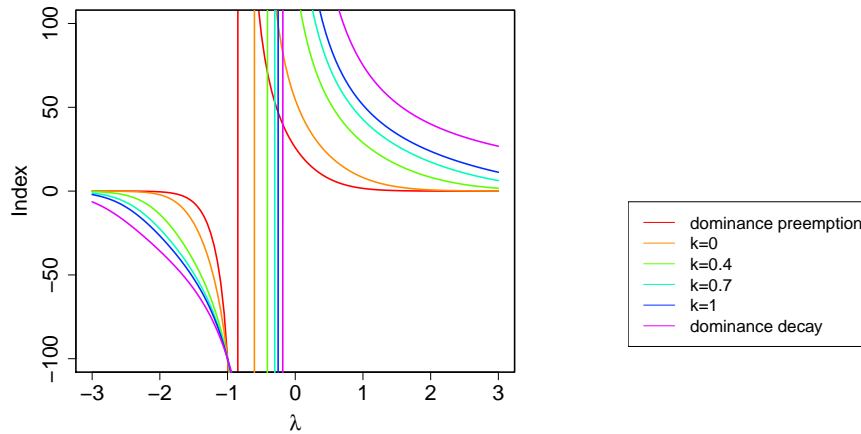
Appendix E

Sampling properties of $\gamma_{\mathbf{n}}(\lambda)$

In order to investigate the properties of the transformation

$$\gamma_{\mathbf{n}}(\lambda) = (S - 1)/I_{\mathbf{n}}(\lambda) \qquad \text{(E.1)}$$

(see equation 2.10), we evaluate the statistic for samples from a completely even distribution and a vector of species proportions of increasing unevenness (where $S{=}100$ in all cases). We compare their sample means on the background of the sampling distribution under perfect evenness (a rescaled χ^2_{99}).

Each of the following panels shows a histogram of the sampling distribution of $\gamma_{\mathbf{n}}(\lambda)$ under perfect evenness in the background. To derive this distribution, 1000 random draws were generated from a χ^2_{99} and standardised by $S{-}1$ for each of the four plots. For the two panels in the upper row, 100 random samples were simulated from the uniform distribution $(1/100, \ldots, 1/100)$, and the sample mean over the corresponding 100 values of $\gamma_{\mathbf{n}}(\lambda)$ was calculated, where in (a) the size of each sample was $n = 50$ and in (b) $n = 200$. Panel (c) shows sample means over samples from uneven species abundance distribution (again 100 samples were randomly generated, where sample size $n = 50$). The species abundance distributions are

$$
\begin{aligned}
p_1 &= (2/150, \ldots, 2/150, 1/150, \ldots, 1/150) \\
p_2 &= (0.6, 0.3, 0.2, 0.1, \ldots, 0.1, 0.006, \ldots, 0.006) \\
p_3 &= (0.15, 0.07, 0.07, 0.06, 0.5, 0.4, 0.4, 0.3, 0.2, 0.1, 0.007, \ldots, 0.007, \\
&\qquad\quad 0.006, \ldots, 0.006, 0.005, \ldots, 0.005, 0.004, \ldots, 0.004, \\
&\qquad\quad 0.003, \ldots, 0.003, 0.002, \ldots, 0.002, 0.001, \ldots, 0.001)
\end{aligned}
$$

where in p_1 half the species are exactly twice as abundant, 50 species' proportions are equal to 0.006 in p_2, and the frequencies of the species proportions in the tail of p_3 are 30 times 0.007, 10 times 0.006, 10 times 0.005, 20 times 0.004, 5 times 0.003, 5 times 0.002, 10 times 0.001.

Finally, (d) illustrates the effects of under- and overestimation of S, where again a sample mean over 100 samples (of the completely even SAD) of size $n = 50$ is calculated, roughly 40 species are observed in each sample.
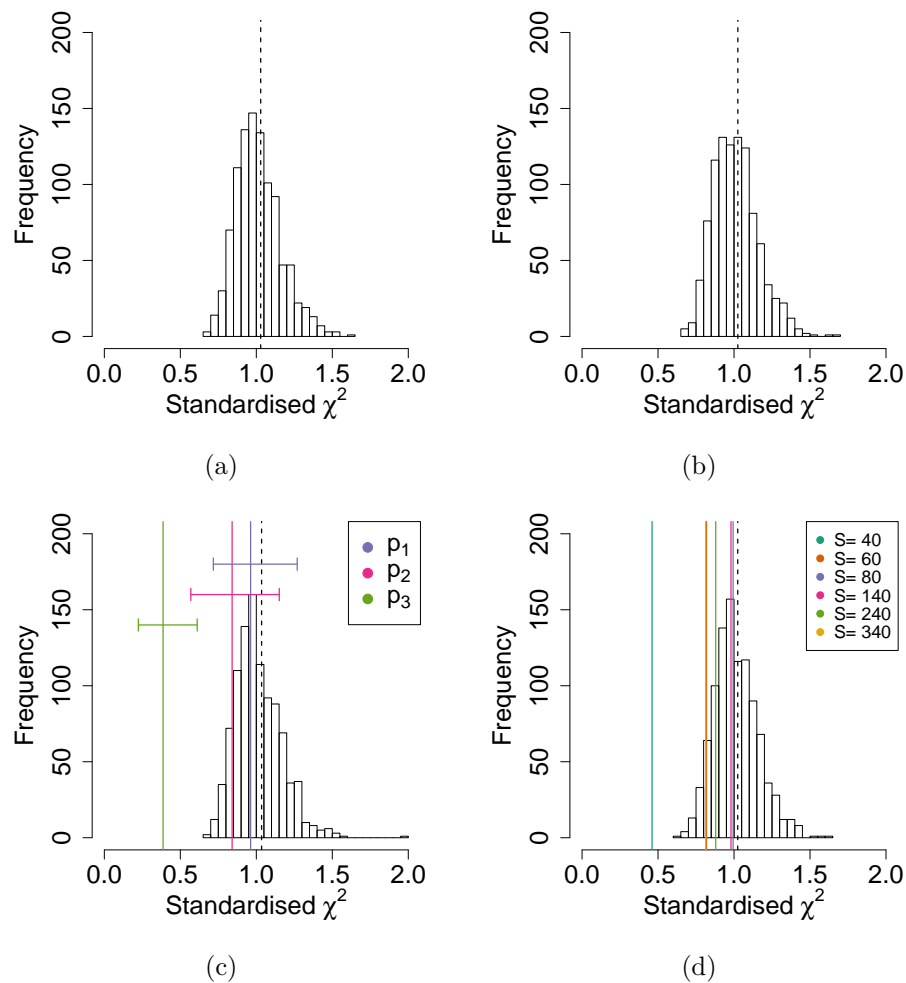
FIGURE E.1 – Sampling properties of the transformation $\gamma_{\mathbf{n}}(\lambda)$ of the goodness-of-fit measure family. $\gamma_{\mathbf{n}}(\lambda)$ is entirely sample-based and does not have a corresponding divergence measure any more. The figure shows the sample mean of this summary statistic (over 100 samples) of different size ((a) and (b)), where samples are taken from uneven species abundance distributions (c) and evaluated for under- and overestimates of species richness (d).

Appendix F

Sensitivity to the choice of ϵ
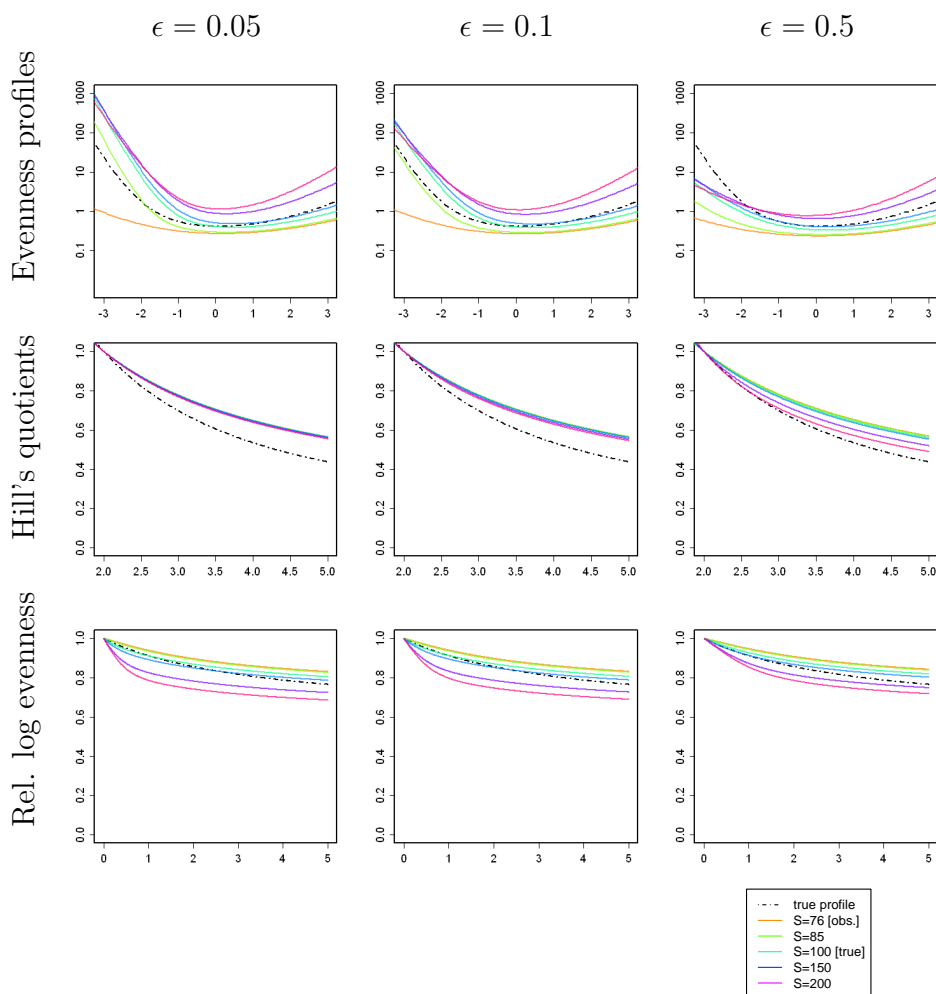
The figure shows the bias in estimated evenness profiles depending on the choice of ϵ correction for zeros in the observations. Different ϵ-corrections are applied to the sample profiles for scenario 3 for the power fraction model with $k = 1$. Analogous to Fig. 3.3 in chapter 3, the number of species is deliberately under-

and overestimated when the sample is corrected for missing species. The true profile which is shown as a black dotted line is based on the species abundances that are expected from the Tokeshi power fraction model ($k = 1$).

Appendix G

Supplementary information for the BBS analysis

The following provides additional information about the bird species included in the analysis in chapter 4.

The following table lists all species by primary habitat groups and specifies whether or not they are included in the analysis. Only common species were included in the final analysis; 'L' indicates that records from the second ('late') visit were used. Superscript ∘ for a common species indicates its exclusion because of its abundance index being zero in at least one year; the asterix ∗ indicates wetland and grassland species which were additionally excluded as they frequently appeared without any records for some years in the bootstrap (but not the original analysis). For the common species, we also list average estimated detection probability π_{est}(across years) and which model was chosen for the scale parameter of the half-normal detection function (by AIC). The two most common and the two rarest (included) species are highlighted (with respect to the smoothed abundance index N_{est} in 1994 and 2008, where abundance estimates are given in 10^6).

In addition, trends in individual species' relative abundances (relative to 1994, the first year of the survey) are shown.

Species	Scientific name	Visit	π_{est}	Model	N_{est} (1994)	N_{est} (2008)	Rare species (excluded)	
Farmland								
Red-legged Partridge	*Alectoris rufa*		0.40	-	0.365	0.552	Quail	*Coturnix coturnix*
Grey Partridge	*Perdix perdix*		0.36	linear	0.279	0.153	Red Kite	*Milvus milvus*
Pheasant	*Phasianus colchicus*		0.50	linear	1.260	1.799	Montagu's Harrier	*Circus pygargus*
Kestrel	*Falco tinnunculus*		0.47	-	**0.071**	**0.084**	Corncrake	*Crex crex*
Lapwing	*Vanellus vanellus*		0.57	linear	0.658	0.647	Stone-curlew	*Burhinus oedicnemus*
Stock dove	*Columba oenas*		0.49	factor	0.217	0.197	Brambling	*Fringilla montifringilla*
Turtle dove	*Streptopelia turtur*	L	0.50	-	0.112	**0.025**	Cirl Bunting	*Emberiza cirlus*
Rook	*Corvus frugilegus*		0.54	factor	**2.300**	**3.267**		
Carrion crow	*Corvus corone*		0.51	factor	1.791	2.638		
Hooded crow	*Corvus cornix*		0.49	factor	**0.079**	0.138		
Skylark	*Alauda arvensis*		0.54	linear	**2.866**	**2.948**		
Whitethroat	*Sylvia communis*	L	0.31	factor	1.576	1.968		
Tree Sparrow	*Passer montanus*		0.27	-	0.340	0.495		
Yellow Wagtail	*Motacilla flava flavissima*	L	0.38	factor	0.186	0.084		
Linnet	*Carduelis cannabina*		0.31	factor	2.230	2.035		
Yellowhammer	*Emberiza citrinella*		0.31	factor	2.206	2.208		
Corn Bunting	*Emberiza calandra*	L	0.41	factor	0.195	0.104		
Grassland								
Red Grouse	*Lagopus lagpus*		0.41	linear	0.250	0.211	Ptarmigan	*Lagopus muta*
Golden Plover	*Pluvialis apricaria*	L	0.50	linear	0.172	0.102	Black Grouse	*Tetrao tretrix*
Dunlin	*Calidris alpina*	L	0.44	linear	0.074	0.109	Hen Harrier	*Circus cyaneus*
Curlew	*Numenius arquata*		0.53	linear	0.381	0.249	Golden Eagle	*Aquila chrysaetos*
Redshank	*Tringa totanus*		0.53	linear	0.0980	0.137	Merlin	*Falco columbarius*

Raven	*Corvus corax*		0.53	-	**0.023**	**0.055**	Peregrine	*Falco peregrinus*	
Dartford Warbler *	*Sylvia undata*		(0.39)	(linear)	(0.003)	(0.011)	Dotterel	*Charadrius morinellus*	
Ring Ouzel *	*Turdus torquatus*	(L)	(0.44)	(linear)	(0.038)	(0.021)	Whimbrel	*Numenius phaeopus*	
Whinchat	*Saxicola rubetra*		0.45	-	0.200	0.136	Green Sandpiper	*Tringa ochropus*	
Stonechat	*Saxicola torquatus*		0.45	-	**0.053**	**0.027**	Greenshank	*Tringa nebularia*	
Wheatear	*Oenanthe oenanthe*	L	0.40	factor	**0.517**	**0.461**	Short-eared Owl	*Asio flameus*	
Meadow pipit	*Anthus pratensis*		0.39	factor	**6.805**	**5.414**	Woodlark	*Lullula arborea*	
Twite *	*Carduelis flavirostris*	(L)	(0.35)	(-)	(0.070)	(0.055)	Snow Bunting	*Plectrophenax nivalis*	

Human habitation

Feral Pigeon	*Columba livia*		0.35	factor	1.046	1.338	Black Redstart	*Phoenicurus ochruros*	
Collared Dove	*Streptopelia decaocto*		0.37	factor	1.269	1.874			
Swift	*Apus apus*	L	0.45	factor	**0.338**	**0.145**			
Magpie	*Pica pica*		0.44	factor	1.149	1.349			
Jackdaw	*Corvus monedula*		0.47	factor	1.837	3.193			
Swallow	*Hirundo rustica*	L	0.35	factor	1.652	2.258			
House Martin	*Delichon urbicum*	L	0.34	factor	0.920	0.768			
Starling	*Sturnus vulgaris*		0.39	factor	7.590	5.987			
Blackbird	*Turdus merula*		0.35	factor	**8.103**	**11.588**			
Mistle Thrush	*Turdus viscivorus*		0.45	linear	0.476	0.429			
Spotted Flycatcher	*Muscicapa striata*	L	0.29	linear	**0.293**	**0.242**			
Dunnock	*Prunella modularis*		0.29	factor	3.202	4.245			
House Sparrow	*Passer domesticus*		0.26	factor	**12.167**	**10.571**			
Pied Wagtail	*Motacilla alba*		0.34	factor	0.864	1.057			
Greenfinch	*Carduelis chloris*		0.33	-	2.635	3.448			
Goldfinch	*Carduelis carduelis*		0.28	factor	1.409	2.548			

Wetland								
Mute Swan	*Cygnus olor*		0.46	factor	0.129	0.135	Whooper swan	*Cygnus cygnus*
Greylag Goose	*Anser anser*		0.56	-	0.050	0.302	Egyptian goose	*Alopochen aegyptiaca*
Canada Goose	*Branta canadensis*		0.51	factor	0.170	0.267	Wigeon	*Anas penelope*
Shelduck	*Tadorna tadorna*		0.55	linear	0.115	0.088	Pintail	*Anas acuta*
Gadwall *	*Anas strepera*		(0.50)	(-)	(0.012)	(0.016)	Garganey	*Anas querquedula*
Teal *	*Anas crecca*		(0.47)	(linear)	(0.041)	(0.019)	Shoveler	*Anas clypeata*
Mallard	*Anas platyrhynchos*		0.38	factor	1.135	1.587	Common Scoter	*Melanitta nigra*
Pochard *	*Aythya ferina*		(0.52)	(linear)	(0.008)	(0.011)	Goldeneye	*Bucephala clangula*
Tufted duck	*Aythya fuligula*		0.54	factor	0.128	0.178	Red-throated Diver	*Gavia stellata*
Goosander *	*Megus merganser*		(0.38)	(linear)	(0.019)	(0.020)	Black-throated Diver	*Gavia arctica*
Ruddy duck *	*Oxyura jamaicensis*		(1.00)	(uniform)	(0.003)	(0.002)	Bittern	*Botaurus stellaris*
Cormorant	*Phalacrocorax carbo*		0.51	-	**0.010**	**0.019**	Little Egret	*Egretta garzetta*
Grey Heron	*Ardea cinerea*		0.44	linear	0.055	0.083	Slavonian Grebe	*Podiceps auritus*
Little Grebe	*Tachybaptus ruficollis*		0.49	-	0.019	0.022	Black-necked Grebe	*Podiceps nigricollis*
Great Crested Grebe	*Podiceps cristatus*		0.52	-	0.034	0.026	Marsh Harrier	*Circus aeruginosus*
Moorhen	*Gallinula chloropus*		0.36	linear	0.351	0.345	Osprey	*Pandion haliaetus*
Coot	*Fulia atra*		0.45	factor	0.140	0.207	Water Rail	*Rallus aquaticus*
Oystercatcher	*Haematopus ostralegus*		0.54	linear	**0.488**	0.425	Spotted Crake	*Porzana porzana*
Avocet°	*Recurvirostra avosetta*		(1.00)	(uniform)	(0)	(0.006)	Little Ringed Plover	*Charadrius dubius*
Snipe	*Gallinage gallinago*		0.44	-	0.150	0.155	Ruff	*Philomachus pugnax*
Black-tailed Godwit°	*Limosa limosa*		(1.00)	(uniform)	(0)	(0.027)	Bearded Tit	*Panurus biarmicus*
Common Sandpiper	*Actitis hypoleucos*	L	0.36	-	0.162	0.092	Savi's Warbler	*Locustella luscinioides*
Cuckoo	*Cuculus canorus*		0.54	-	0.102	0.039	Marsh Warbler	*Acrocephalus palustris*
Kingfisher	*Alcedo atthis*		0.29	-	**0.015**	**0.008**	Mandarin Duck	*Aix galericulata*
Sand Martin *	*Riparia riparia*	(L)	(0.42)	(linear)	(0.019)	(0.014)		

Cetti's Warbler *	*Cettia cetti*		(0.47)	(-)	(0.010)	(0.017)			
Lesser Whitethroat	*Sylvia curruca*	L	0.31	linear	0.151	0.093			
Grasshopper Warbler	*Locustella naevia*	L	0.49	-	0.024	0.034			
Sedge Warbler	*Acrocephalus schoenobaenus*	L	0.34	factor	0.476	**0.517**			
Reed Warbler	*Acrocephalus scirpaceus*	L	0.31	factor	0.202	0.230			
Dipper	*Cinclus cinclus*		0.27	-	0.058	0.059			
Grey Wagtail	*Motacilla cinerea*		0.28	-	0.082	0.129			
Reed Bunting	*Emberiza schoeniclus*		0.35	factor	**0.519**	**0.525**			

Woodland

Sparrowhawk	*Accipiter nisus*		0.39	-	**0.024**	0.030	Capercaillie	*Tetrao urogallus*	
Buzzard	*Buteo buteo*		0.45	-	0.102	0.180	Golden Pheasant	*Chrysolophus pictus*	
Woodpigeon	*Columba palumbus*		0.41	factor	6.379	**11.091**	Lady Amherst's Ph.	*Chrysolophus amherstiae*	
Green Woodpecker	*Picus viridis*		0.54	-	0.074	0.149	Honey Buzzard	*Pernis apivorus*	
Great Spotted Woodpecker	*Denddrocopos major*		0.47	-	0.108	0.361	Goshawk	*Accipiter gentilis*	
Jay	*Garrulus glandarius*		0.40	-	0.199	0.244	Hobby	*Falco subbuteo*	
Goldcrest	*Regulus regulus*		0.28	factor	1.063	1.570	Lesser Sp. Woodp.	*Dendrocopos minor*	
Blue Tit	*Cyanistes caeruleus*		0.26	factor	**8.110**	10.918	Golden Oriole	*Oriolus oriolus*	
Great Tit	*Parus Major*		0.33	factor	3.082	6.270	Firecrest	*Regulus ignicapilla*	
Coal Tit	*Periparus ater*		0.33	factor	1.112	1.572	Crested tit	*Lophophanes cristatus*	
Willow Tit	*Poecile montana*		0.27	-	0.053	**0.022**	Fieldfare	*Turdus pilaris*	
Marsh Tit	*Poecile palustris*		0.28	linear	0.086	0.075	Redwing	*Turdus iliacus*	
Long-tailed Tit	*Aegithalos caudatus*		0.24	linear	0.824	1.256	Nightingale	*Luscinia megarhynchos*	
Wood Warbler	*Phylloscopus sibilatrix*	L	0.47	-	0.096	**0.025**	Scottish Crossbill	*Loxia scotica*	
Chiffchaff	*Phylloscopus collybita*	L	0.47	linear	0.648	1.090	Hawfinch	*Coccothraustes coccothraustes*	
Willow Warbler	*Phylloscopus trochilus*	L	0.43	-	2.724	2.636			

| Blackcap | *Sylvia atricapilla* | L | 0.39 | linear | 0.710 | 1.320 |
Garden Warbler	*Sylvia borin*	L	0.37	-	0.267	0.243
Nuthatch	*Sitta europaea*		0.42	linear	0.135	0.253
Treecreeper	*Certhia familiaris*		0.26	-	0.220	0.271
Wren	*Troglodytes troglodytes*		0.38	factor	6.531	9.381
Song Thrush	*Turdus philomelos*		0.46	linear	1.301	1.661
Robin	*Erithacus rubecula*		0.32	factor	6.369	9.212
Redstart	*Phoenicurus phoenicurus*	L	0.38	-	0.165	0.155
Pied Flycatcher	*Ficedula hypoleuca*	L	0.37	linear	0.074	0.019
Tree Pipit	*Anthus trivialis*	L	0.46	linear	0.155	0.151
Chaffinch	*Fringilla coelebs*		0.34	factor	**11.177**	**14.436**
Siskin	*Carduelis spinus*	L	0.37	-	0.300	0.508
Lesser Redpoll	*Carduelis cabaret*	L	0.38	factor	0.124	0.231
Common Crossbill	*Loxia curvirostra*		0.43	-	**0.051**	0.097
Bullfinch	*Pyrrhula pyrrhula*		0.26	-	0.474	0.430
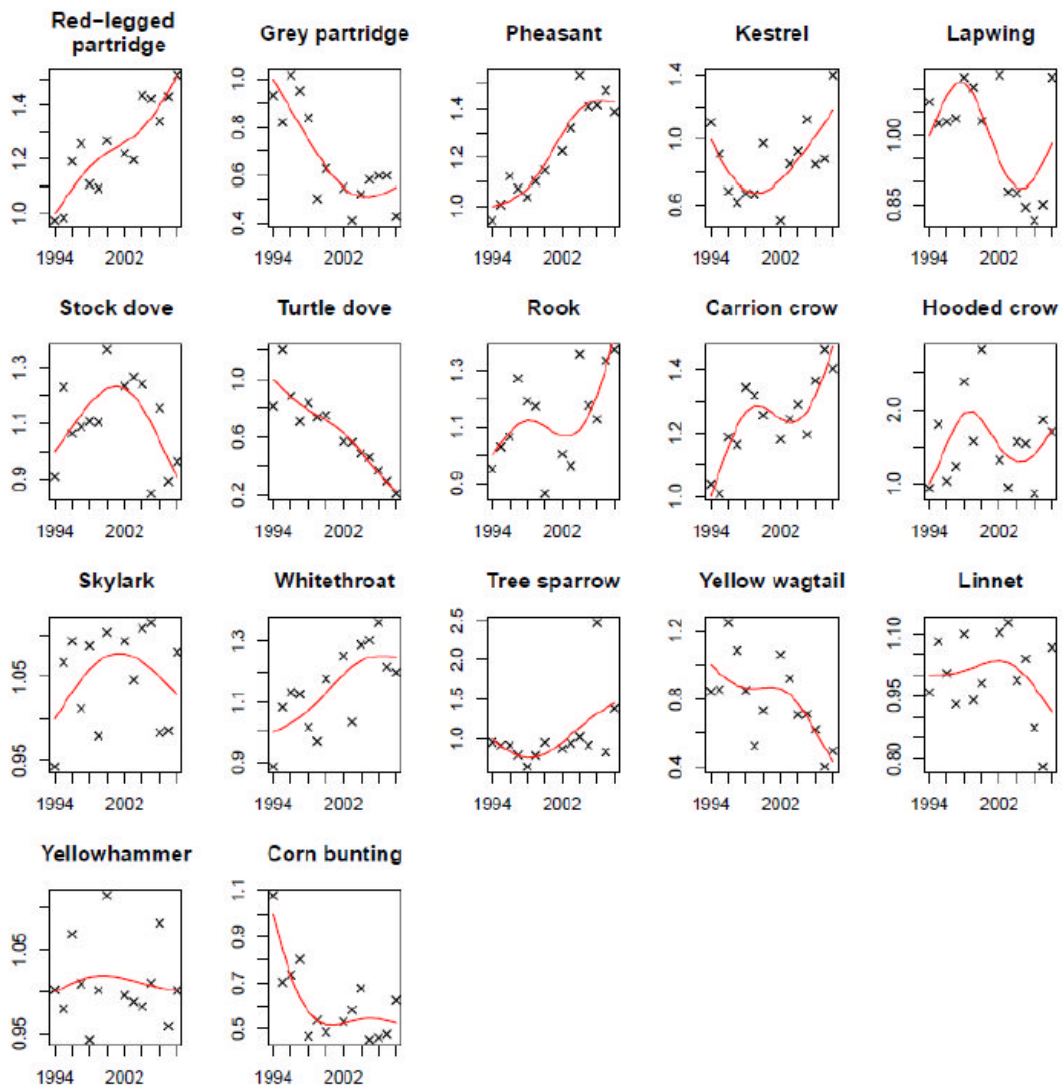
FIGURE G.1 – Single species trends for the farmland bird community. Relative abundance indices are shown.
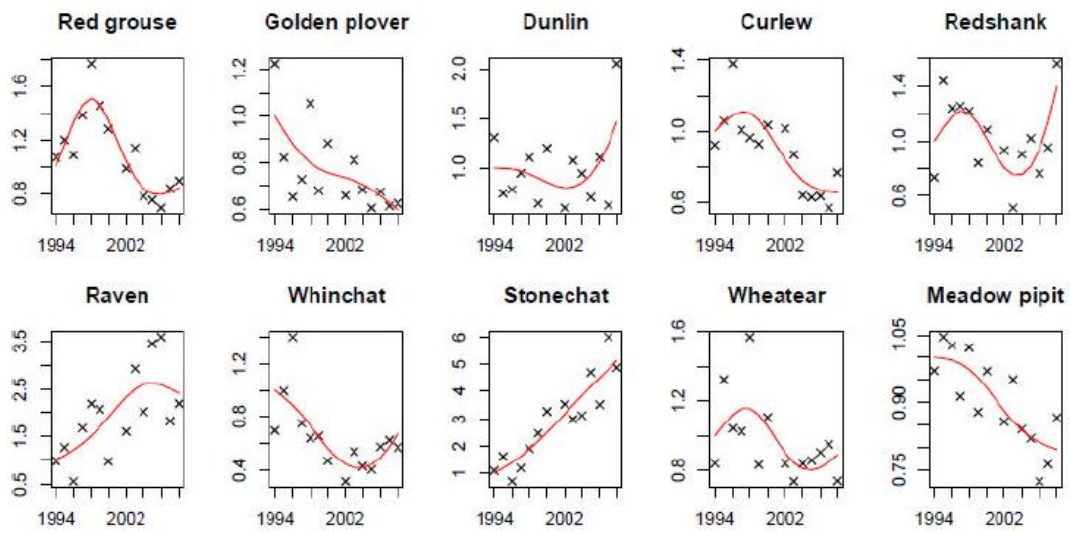
FIGURE G.2 – Single species trends for the grassland bird community. Relative abundance indices are shown.
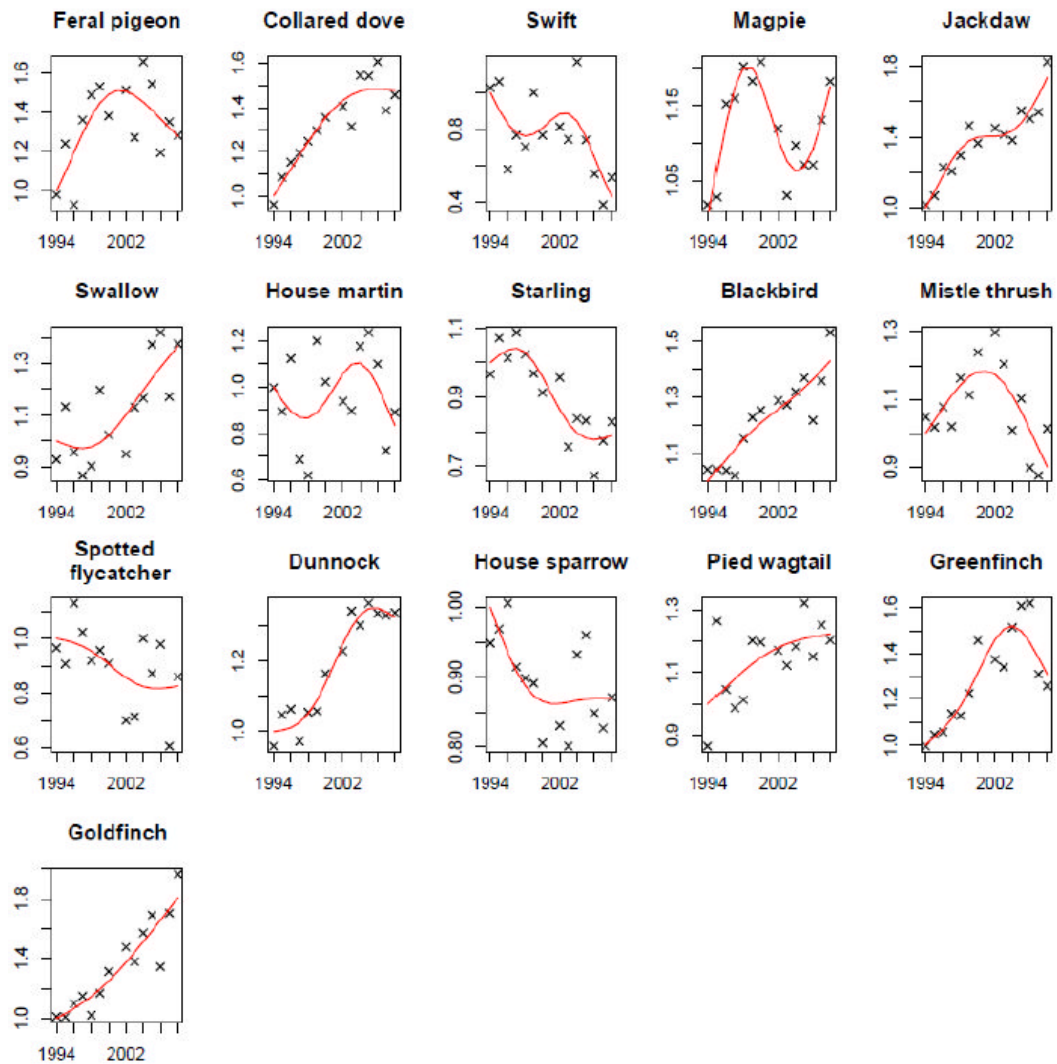
Figure G.3 – Single species trends for the near human habitation bird community. Relative abundance indices are shown.
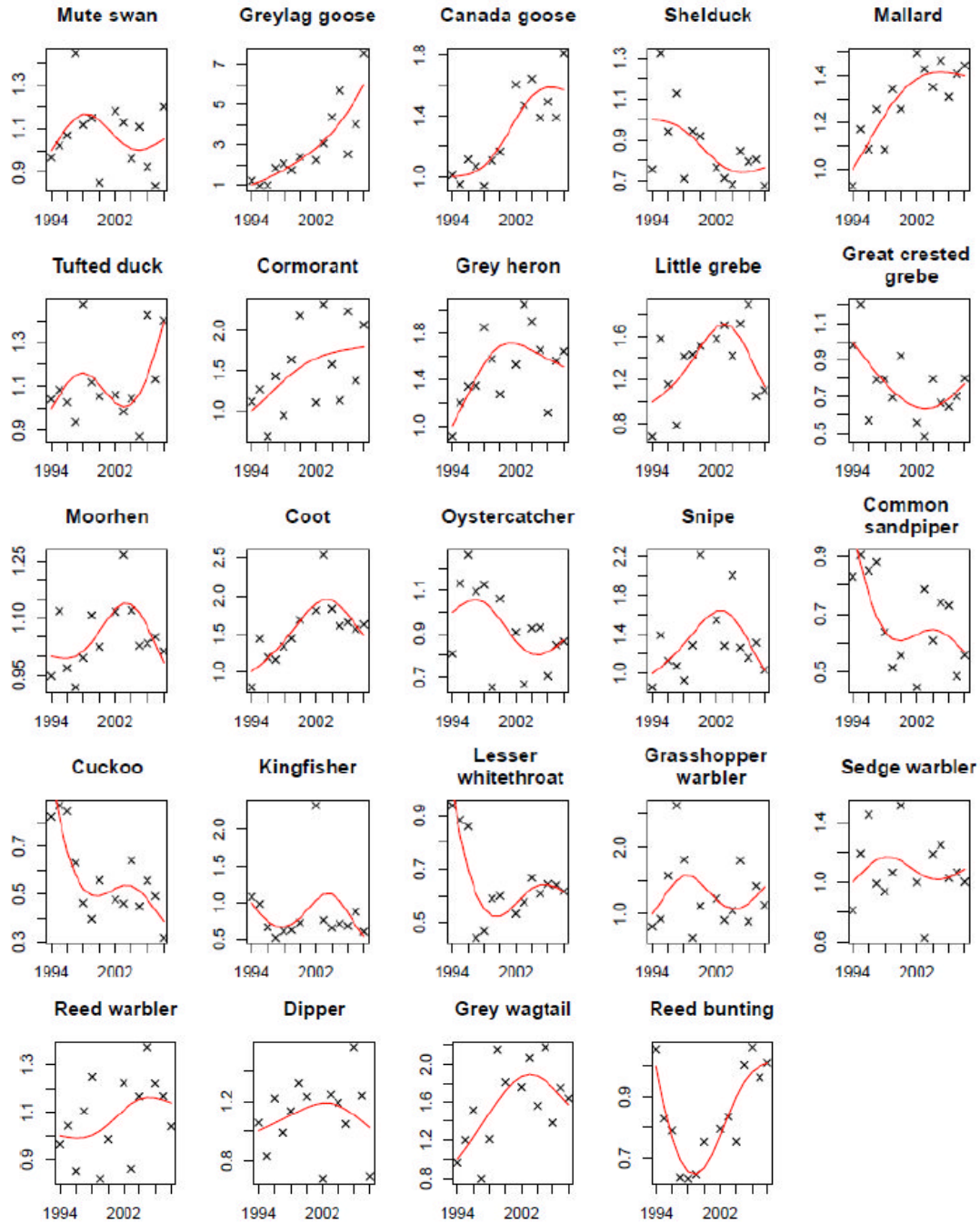
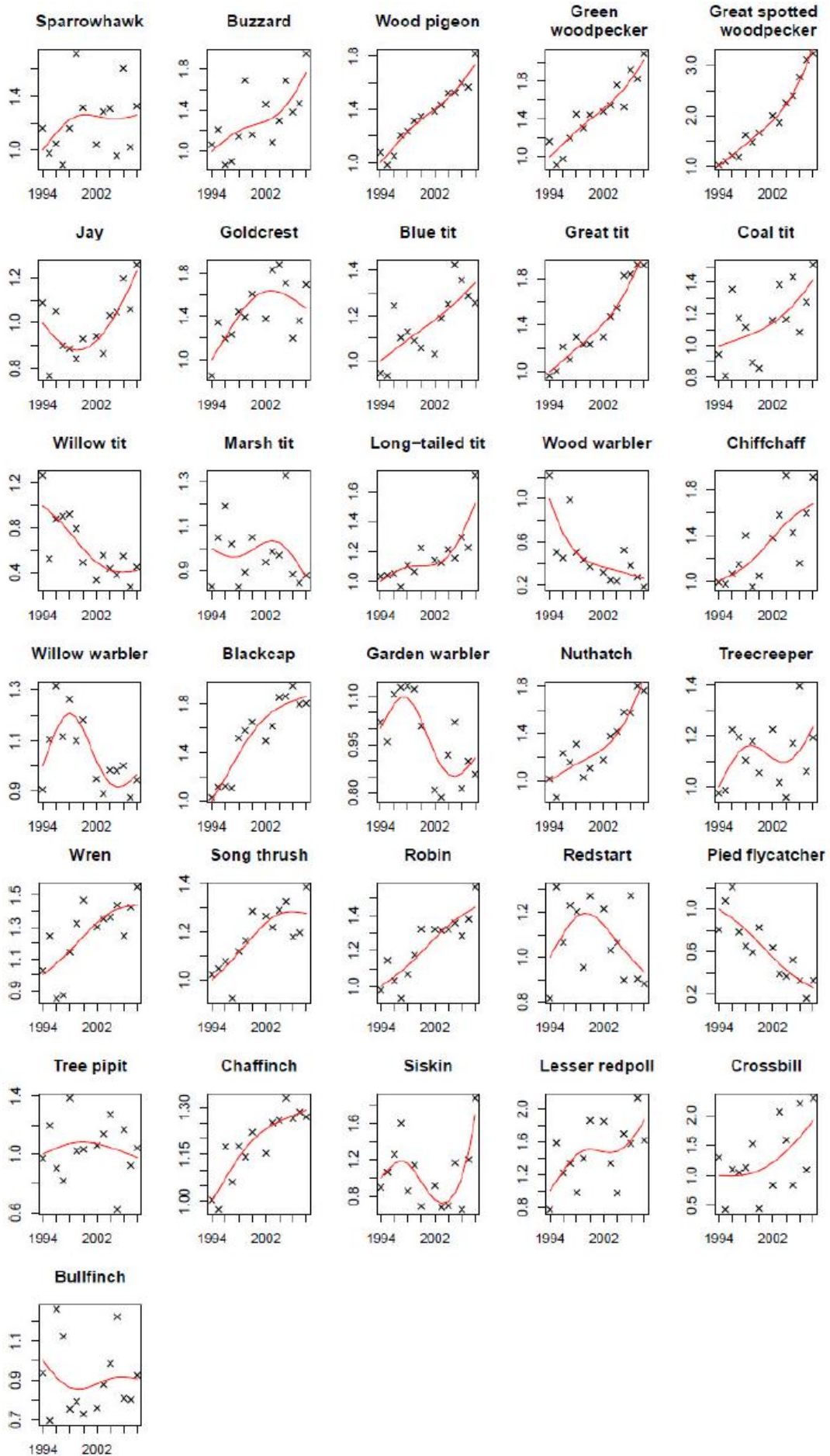FIGURE G.4 – Single species trends for the wetland bird community. Relative abundance indices are shown.

FIGURE G.5 – Single species trends for the woodland bird community. Relative abundance indices are shown.