

Bayesian analysis of community dynamics

Crispin Mutshinda Mwanza

Department of Mathematics and Statistics
Faculty of Science
University of Helsinki
Finland

Academic Dissertation

To be presented, with the permission of the Faculty of Science
of the University of Helsinki, for public criticism in Auditorium B 123 on May 17th 2010 at 12
o'clock noon.

Helsinki 2010

© Crispin Mutshinda Mwanza (Summary)
© Wiley-Blackwell (Chapter I)
© The authors; licensee BioMed Central Ltd. (Chapter II)
©The Royal Society (Chapter III)
©The authors (Chapters IV & V)

Author's present address:

University of Helsinki
Faculty of Science
Department of Mathematics and Statistics
P.O. Box 68
FIN-00014 Helsinki
Finland

Email: crispin.mutshindamwanza@helsinki.fi
mutshinda@gmail.com

© Crispin Mutshinda Mwanza
ISBN 978-952-92-7164-1 (paperback)
ISBN 978-952-10-6193-6 (PDF)

<http://ethesis.helsinki.fi>

Yliopisto paino
Helsinki 2010

Bayesian analysis of community dynamics

Author: Crispin Mutshinda Mwanza

Supervisor: Docent Robert B. O'Hara
Department Mathematics and Statistics
University of Helsinki
Helsinki
Finland

Current address: Biodiversity and Climate Research Centre
Frankfurt am Main
Germany

Pre-examiners: Professor Otso Ovaskainen
Mathematical Biology Group
Department of Biological and Environmental Sciences
University of Helsinki
Helsinki
Finland

Dr. Rampal S. Etienne
Community and Conservation Ecology Group,
University of Groningen,
AA Haren
The Netherlands

Opponent: Dr. Len Thomas
School of Mathematics and Statistics
University of St Andrews
St Andrews
Scotland

List of publications

This thesis is based on the following articles which are referred to in the text by their Roman numerals:

- I. Crispin M. Mutshinda, Robert B. O’Hara, & Ian. P. Woiwod (2008) Species abundance dynamics under neutral assumptions: a Bayesian approach to the controversy. *Functional Ecology* **22**: 340-347.
- II. Crispin M. Mutshinda & Robert B. O’Hara (2010) On the setting of environmental noise and the predictive performance of population dynamics models. *BMC Ecology* **10**:7.
- III. Crispin M. Mutshinda, Robert B. O’Hara, & Ian. P. Woiwod (2009) What drives community dynamics? *Proc. R. Soc. London B.* **276**: 2923-2929.
- IV. Crispin M. Mutshinda, Robert B. O’Hara, & Ian. P. Woiwod A multispecies perspective on ecological impacts of climatic forcing. Revised version submitted to the *Journal of Animal Ecology*.
- V. Crispin M. Mutshinda & Robert B. O’Hara Integrating the niche and neutral perspectives on community structure and dynamics. *Submitted for publication*.

Author’s contribution

I am fully responsible for the introduction part of this thesis. I also had the main responsibility in the design and implementation of the methods as well as the writing of the included articles (I-V).

CONTENTS

1. INTRODUCTION	8
1.1. Background	8
1.2 Aims of the thesis	10
2. METHODOLOGIES	11
2.1. Bayesian inference and the hierarchical Bayesian framework	11
2.2. Computational issues: Monte Carlo integration and MCMC methods	13
2.2.1 Monte Carlo integration	13
2.2.2. Markov chain Monte Carlo	14
2.3. Bayesian model selection	17
2.3.1 Bayes factors	17
2.3.2. The deviance information criterion	18
2.3.3. Stochastic search variable selection	19
2.4. Model assessment through posterior predictive checks	20
3. A PROBABILISTIC TEST OF THE NEUTRAL THEORY OF BIODIVERSITY	21
4. THE GENERAL FRAMEWORK	24
4.1. The process model	24
4.2. Variance partitioning	25
4.3. The sampling model	25
4.4. Incorporating environmental covariates	27
5. INTEGRATING NEUTRALITY AND NICHES	28
6. SUMMARIES OF THE ORIGINAL ARTICLES AND MAIN RESULTS	32
Article I: Species abundance dynamics under neutral assumptions: a Bayesian approach to the controversy	32
Article II: On the setting of environmental noise and the performance of population dynamical models	32
Article III: What drives community dynamics?	32
Article IV: A multispecies perspective on ecological impacts of climatic forcing	33
Article V: Integrating the niche and neutral perspectives on community structure and dynamics	33

7. CONCLUSIONS & PERSPECTIVES	----- 34
ACKNOWLEDGEMENTS	----- 35
REFERENCES	----- 35

ABSTRACT

Elucidating the mechanisms responsible for the patterns of species abundance, diversity, and distribution within and across ecological systems is a fundamental research focus in ecology. Species abundance patterns are shaped in a convoluted way by interplays between inter-/intra-specific interactions, environmental forcing, demographic stochasticity, and dispersal. Comprehensive models and suitable inferential and computational tools for teasing out these different factors are quite limited, even though such tools are critically needed to guide the implementation of management and conservation strategies, the efficacy of which rests on a realistic evaluation of the underlying mechanisms. This is even more so in the prevailing context of concerns over climate change progress and its potential impacts on ecosystems.

This thesis utilized the flexible hierarchical Bayesian modelling framework in combination with the computer intensive methods known as Markov chain Monte Carlo, to develop methodologies for identifying and evaluating the factors that control the structure and dynamics of ecological communities. These methodologies were used to analyze data from a range of taxa: macro-moths (Lepidoptera), fish, crustaceans, birds, and rodents. Environmental stochasticity emerged as the most important driver of community dynamics, followed by density dependent regulation; the influence of inter-specific interactions on community-level variances was broadly minor.

This thesis contributes to the understanding of the mechanisms underlying the structure and dynamics of ecological communities, by showing directly that environmental fluctuations rather than inter-specific competition dominate the dynamics of several systems. This finding emphasizes the need to better understand how species are affected by the environment and acknowledge species differences in their responses to environmental heterogeneity, if we are to effectively model and predict their dynamics (e.g. for management and conservation purposes). The thesis also proposes a model-based approach to integrating the niche and neutral perspectives on community structure and dynamics, making it possible for the relative importance of each category of factors to be evaluated in light of field data.

1. INTRODUCTION

1.1. Background

Understanding the mechanisms underlying the patterns of species abundance, diversity, and distribution on different spatiotemporal scales is an important quest in ecology (May 1975; Levin 1992; Ricklefs & Schluter 1993; Hubbell 2001; Enquist *et al.* 2002). Temporal fluctuations in the abundance of species typically result from the combined effects of environmental fluctuations, density-dependent regulation, demographic stochasticity, and dispersal (Lande *et al.* 2003; Wilson & Lundberg 2006). Species also exist as part of communities, so between-species interactions may be important as well (Bower 1962; Murdoch 1994; Shorrocks & Sevenster 1995; Miyashita 2001).

The community approach to biodiversity has recently enjoyed increasing attention, the main focus being on the identification of the assembly rules, along with any exogenous factors that may interact with these rules to create and maintain ecological communities. Relating to this, an on-going debate opposes the niche-assembly view and the neutral perspective on community structure and dynamics.

The niche-assembly view (Gause 1934; Hutchinson 1957; Armstrong & McGehee 1980, Tokeshi 1990) incorporates from the start the idea that species must differ in a variety of ways, including their niches in order to coexist and avoid competitive exclusion (Gause 1934; Hardin 1960). The competitive exclusion principle (CEP) (Gause 1934) asserts that no two organisms can stably coexist on exactly the same resources; one will always out-compete the other and drive it extinct. In order to underplay the competition pressure, the weakest competitor has no choice but switching to a different resource utilization pattern or a different niche (niche segregation). Niche-assembled communities can be thought of as societies primarily structured by species interactions as well as organisms' life histories, habitats, and trophic levels. According to this perspective, competing species coexist in stable equilibrium by partitioning limiting resources through niche differentiation.

The CEP has found support from mathematical models such as Lotka-Volterra's differential equations for competition (Gause 1935), and from tests using laboratory experiments (Gause 1934, 1935; Bush 1969). However, this principle has been questioned (Hutchinson 1961; Hubbell 2001; Zhou & Zhang 2008) in the face of the striking diversity of species-rich systems such as tropical forests or planktonic communities, where an apparent strong overlap of species niches does not preclude stable coexistence. This phenomenon is known as the "paradox of the plankton" after Hutchinson's (1961) paper. The niche-assembly perspective has also been criticized (Hubbell 2001) for its failure to explicitly account for dispersal limitation, the importance of which has long been recognized since MacArthur & Wilson's (1967) theory of island biogeography.

The neutral theory of biodiversity (Bell 2000; Hubbell 2001) has emerged as an alternative to the niche-assembly view. This theory maintains that, at a single trophic level, all individuals, irrespective of their species identities, can be considered to be functionally equivalent on a per capita basis with respect to their birth, death, dispersal, and speciation rates. As a consequence, biodiversity patterns are entirely attributed to random fluctuations in these demographic and evolutionary processes. The main criterion for species coexistence under the neutral perspective is random migration into the same habitable region. This explains why the

neutral view is also called “dispersal-assembly”, in contrast to niche-assembly theories which emphasize the uniqueness of all species in ecological communities.

The most documented formulation of the neutral theory of biodiversity is the unified neutral theory of biodiversity and biogeography (UNTB) introduced by Hubbell (2001). The UNTB considers community dynamics on two different spatio-temporal scales: the local or “ecological” community defined as a group of trophically similar species that actually or potentially compete for the same or similar resources in a geographical area, and the large-scale “metacommunity” which can be viewed as a regional collection of local communities, the arena where speciation occurs and the pool from which local communities are colonized (Hubbell 2001; Magurran 2005). Under the UNTB, the number of individuals comprising an ecological community is treated as a constant (the zero-sum assumption), and local community dynamics are entirely driven by stochastic demographic processes of birth, death, and immigration (from the metacommunity). On the other hand, metacommunity dynamics include an evolutionary mechanism of random speciation through which new species originate. Unlike in local communities, metacommunity species may not actually compete because of scale separation in space or time. In the absence of speciation, the random demographic processes will eventually drive all but one species extinct. Hubbell (2001) refers to the zero-sum neutral dynamics as “ecological drift”, by analogy with genetic drift (Kimura 1983), and calls the resulting relative species abundance distribution the “zero-sum multinomial” (ZSM) distribution.

Neutral community models have allegedly been effective at predicting many community and macroecological patterns including abundance, diversity, and distribution both locally and regionally (Holyoak & Loreau 2006; Hubbell 2001). The UNTB has extended previous ecological research on the quantification of biodiversity patterns, grounding their description in biologically interpretable processes. In particular, it provides sampling models for both local and metacommunity dynamics with two core parameters, the immigration rate (m) and the speciation rate (ν) respectively. It also predicts a “fundamental biodiversity number”, $\theta = 2 J_M \nu$, where J_M is the metacommunity size. The fundamental biodiversity number is related to Fishers’ α (Fisher *et al.* 1943), and represents the diversity-generating capacity of a system i.e., the expected number of new species per generation.

However, the neutral theory has been extensively falsified (e.g. Engen *et al.* 2002; McGill 2003, Etienne & Olf 2004; Maurer & McGill 2004; Poulin 2004; Chase 2005; Graves & Rahbek 2005; Turnbull *et al.* 2005; Williamson & Gaston 2005; Dornelas *et al.* 2006; Adler *et al.* 2007). Most empirical evaluations have rejected the neutral theory on the grounds that natural communities typically fluctuate more widely than neutral drift would predict. The excess variability over neutral predictions is presumably due to the range of mechanisms overlooked by the neutral perspective namely, density-dependent regulation, species differential responses to environmental fluctuations, and differences in competitive abilities across species.

An emerging consensus amongst ecologists suggests that natural communities are shaped in a convoluted way by a combination of abiotic random forcing, biotic interactions, demographic stochasticity, and dispersal (Gravel *et al.* 2006; Adler *et al.* 2007).

Comprehensive models and suitable inferential and computational tools are needed to untangle the relative contribution of these different factors beyond the limits of the prevailing dichotomist approaches of the form neutrality *versus* the niches (Turnbull *et al.* 2005) or compensatory dynamics *versus* environmental forcing (Houlahan *et al.* 2007).

There are obvious issues, both conceptual and methodological, in dealing with multispecies data. Nevertheless, recent advances in hierarchical Bayesian (HB) modelling can be utilized to set up comprehensive models that consistently integrate scientific knowledge about the study system and information from various data sources (Wikle 2003). This thesis combined the HB modelling approach with the computationally intensive methods known as Markov chain Monte Carlo, and species abundance data from long-term monitoring systems, to develop methodologies for teasing out the factors which control the structure and dynamics of ecological communities.

1.2 Aims of the thesis

The main goal of this thesis is to develop comprehensive models, along with adequate inferential and computational tools for identifying and evaluating the mechanisms underlying the structure and dynamics of ecological communities. At the outset, a discrete-time version of the unified neutral theory of biodiversity and biogeography, with the zero-sum assumption relaxed, was fitted to long-term macro-moth (Lepidoptera) community data to evaluate how well the neutral model would explain the dynamical behaviour of the study systems. The results suggested that neutrality and/or demographic stochasticity alone cannot explain the large variation observed in the focal communities. More realistic models were then developed to accommodate the mechanisms overlooked by the neutral theory. These models were utilized to analyze species abundance time series from a variety of taxa: macro-moths (Lepidoptera), fish, crustaceans, birds, and rodents. More specifically, the following aims were set, and approached through five articles referred to in the text by their Roman numerals **I-V**:

(1) Use data from long-term monitoring systems to examine the validity of the neutral theory of biodiversity (**I**)

(2) Investigate whether incorporating environmental noise in one or/and another parameter makes a difference to the performance of population dynamical models (**II**)

(3) Develop a methodology for partitioning the variation in species abundances into contributions from environmental forcing, and intra-/inter-specific interactions to see which ones dominate (**III**)

(4) Extend the scope of previously proposed population dynamical models under climatic forcing to multi-species systems (**IV**)

(5) Develop a model-based methodology for integrating the niche and neutral perspectives on community structure and dynamics (**V**)

2. METHODOLOGIES

Ecological processes are notoriously complex and the related data are usually fraught with missing data and observation errors (e.g. Clark & Bjørnstad 2004). Analyzing such processes in a sensible way requires a combination of scientific knowledge about the underlying mechanisms with information from various data sources. The Bayesian statistical approach (Bernardo & Smith 1994; Gelman *et al.* 2003) is well suited to this purpose (Anderson 1989; Clark 2005); in particular the refined hierarchical Bayesian approach (Berliner 1996; Gelman *et al.* 2003; Wikle & Hooten 2006) which provides a convenient way of representing complex phenomena as a series of simple conditional structures. However, fitting hierarchical Bayesian models to the data typically requires a resort to numerical methods such as Markov chain Monte Carlo simulation methods (Gilks *et al.* 1996) on which analyses in this thesis rely. Model adequacy is also worth checking before a model can be relied upon for inference. In this section, the Bayesian and hierarchical Bayesian model formulations are revisited, along with the related computational tools and prominent model selection and model assessment techniques in the Bayesian framework.

2.1. Bayesian inference and the hierarchical Bayesian framework

Bayesian inference is the process of using data to update the knowledge about the values of parameters of a probability model (Gelman *et al.* 2003). This is done by applying Bayes' formula

$$p(\theta | y) = \frac{p(y | \theta)p(\theta)}{p(y)} \propto p(y | \theta)p(\theta), \quad (2.1)$$

where $p(\theta)$ is the prior distribution which encodes the beliefs about the parameter, θ , before observing the data, $p(y | \theta)$ is the likelihood of the observed data, and $p(\theta | y)$ is the posterior distribution which reflects the levels of beliefs computed in light of the observed data. The normalizing constant $p(y) = \int_{\Theta} p(y | \theta) p(\theta) d\theta$ is the marginal distribution of the data, which is also known as the prior predictive distribution. The integral (2.1) is taken over the support, Θ , of the parameter θ .

Bayesian inference draws entirely on probability theory, with all conclusions arising in the form of (posterior) probability statements about model parameters or functions of them (estimation), or as yet unobserved data (prediction). Bayesian conclusions turn out to be intuitive by contrast to classical tools such as p -values or conventional confidence intervals which are often erroneously interpreted. Bayesian inference is also a learning process in the sense that, as more data are obtained, they are added to the store of information by multiplying their likelihood to the current posterior distribution.

The computational overhead of performing the high-dimensional integrals involved in Bayesian inference, in particular the computation of the normalizing constant $p(y)$ in (2.1) has long prevented the widespread of Bayesian inference or limited its applications to the computationally convenient class of conjugate priors. However, the advent of modern computer technology such as Markov chain Monte-Carlo (MCMC) (Gilks *et al.* 1996) has led to a burst of applications of Bayesian methods in many areas, including ecology and environmental sciences,

where many new and exciting modelling techniques are nowadays applied to problems that could not been dealt with previously.

An appealing feature of the Bayesian approach is the fact that it deals with uncertainty explicitly, in particular in making predictions and handling missing values and nuisance parameters. Having observed the data y , Bayesian prediction for a future observation, \tilde{y} , is based on the so-called posterior predictive distribution

$$p(\tilde{y} | y) = \int_{\Theta} p(\tilde{y} | \theta) p(\theta | y) d\theta, \quad (2.2)$$

which is nothing but the posterior distribution of \tilde{y} . Missing data are also handled in the same way, and nuisance parameters (i.e. parameters which are not of immediate interest) are dealt with by integrating them out. For example, if $\theta = (\theta_1, \theta_2)$ and interest in on θ_1 only, then θ_2 becomes a nuisance parameter, and

$$p(\theta_1 | y) = \int_{\Theta_2} p(\theta_1 | \theta_2, y) p(\theta_2 | y) d\theta_2, \quad (2.3)$$

which can be interpreted as a weighted average of the conditional density $p(\theta_1 | \theta_2, y)$, the weights being the posterior probability densities $p(\theta_2 | y)$ (O'Hara *et al.* 2002).

The Bayesian approach therefore, provides a unified framework for inference and prediction, where parameter uncertainty is naturally transferred to predictive inference.

In extending Bayesian models to more complex settings going beyond the simple likelihood-prior-posterior scheme, a hierarchical formulation is often required. Hierarchical Bayesian (HB) models (Berlinier 1996; Gelman *et al.* 2003; Wikle & Hooten 2006) are Bayesian models in which parameters in the likelihood depend on other parameters not mentioned herein, which themselves require priors that may depend on new parameters, the process coming to an end when no new parameters are introduced. The HB formulation makes it possible to entertain a much richer class of models that can better capture the scientific understanding of phenomena under investigation, by decomposing the joint distribution of a collection of random variables into a series of simple conditional models.

Berlinier (1996) and Wikle & Hooten (2006) delineated three basic stages in the HB modelling of hidden processes: the observation or data model, the process model, and the parameter model. These three model components are intended to be conditionally linked in a hierarchical structure comprising the data model at the lowest level. The observation model specifies the distribution of the data, y , given the state of the actual process of interest, x , taking into account sampling errors, and/or the fact that only a proportion of statistical units may be sampled. The parameters involved in the data model (e.g. capture probability) are denoted by φ , and referred to as “data parameters”. The process model describes the underlying dynamics and involves a number of parameters denoted by θ , and referred to as “process parameters”. Examples of process parameters include demographic rates: birth, death or immigration rates. The parameter model describes the uncertainty about both the data parameters and the process parameters.

The target distribution for inference is the joint posterior distribution, $p(x, \theta, \varphi | y)$, of the latent states of the process, the process parameters and the data parameters, which derives from Bayes' theorem as

$$p(x, \theta, \varphi | y) \propto p(y | x, \varphi) p(x | \theta) p(\theta, \varphi). \quad (2.4)$$

The corresponding posterior predictive distribution $p(\tilde{y} | y)$ is given by

$$p(\tilde{y} | y) = \int_X \int_{\Theta} \int_{\Psi} p(\tilde{y} | \varphi, \theta, x) p(\varphi, \theta, x | y) d\varphi d\theta dx. \quad (2.5)$$

It goes without saying that the increase in model complexity resulting from the HB formulation carries along further computational challenges, making the recourse to numerical methods such as Monte Carlo integration (e.g. Robert & Casella 2004) or Markov chain Monte Carlo (MCMC) simulation (Gilks *et al.* 1996) inevitable. An overview of Monte Carlo integration and MCMC methods is provided below.

2.2. Computational issues: Monte Carlo integration and MCMC methods

2.2.1 Monte Carlo integration

It is often possible to approximate high dimensional integrals with respect to a probability measure, in particular the integrals arising in Bayesian analysis, by finite sums involving random samples from the focal probability distribution. For example, let $X = (X_1, \dots, X_k)$ be a k -dimensional random variable with probability density function $f(\cdot)$, where k may be large. If a reasonable number, n , of random vectors $\{x^{(i)}\}_{i=1}^n$ can be drawn from $f(\cdot)$, then the expected value $E[X] = \int x f(x) dx$ can be approximated by the finite sum $\sum_{i=1}^n x^{(i)} / n$. In general, having a sample $\{x^{(i)}\}_{i=1}^n$ from $f(x)$, $I = \int h(x) f(x) dx$ can be approximated by $\hat{I} = \sum_{i=1}^n h(x^{(i)}) / n$. This process is known as Monte Carlo integration (e.g. Robert & Casella 2004).

The Monte Carlo estimate, \hat{I} , is valid and useful since $\hat{I} \rightarrow I$ almost surely (i.e. with probability 1) as $n \rightarrow \infty$, by the strong law of large numbers (Ross 2007). The simulation error can be evaluated by the variance $\text{Var}(\hat{I}) = \sum_{i=1}^n (h(x_i) - \hat{I})^2 / n(n-1)$, which is reminiscent of the variance of the sample mean for a normal population with unknown variance.

Monte Carlo integration is highly relevant to Bayesian inference in the sense that, if $h(\theta)$ is an interesting function of θ and $p(\theta | y)$ is the posterior distribution of θ , then the posterior mean $E[h(\theta) | y] = \int h(\theta) p(\theta | y) d\theta$ can be approximated by $\sum_{i=1}^n h(\theta^{(i)}) / n$, where $\theta^{(i)}$ is a sample from the posterior distribution $p(\theta | y)$ for $i = 1, \dots, n$. Other distributional summaries can be obtained in a similar fashion. Moreover, a $100x(1-\alpha)\%$ credible interval for θ is given by the $\alpha/2$ - and $(1-\alpha/2)$ -percentiles of $p(\theta | y)$ estimated from the simulated posterior. So, the only problem would be to obtain samples from the posterior distribution, and Markov chain Monte Carlo methods make this possible.

2.2.2. Markov chain Monte Carlo

A stochastic process $\{X_t \in S : t \in T\}$ is a collection of random variables, where t is frequently (but not necessarily) a time index. If we think of X_t as the state of the process $\{X_t\}$ at time t , and assume the following dependence condition known as the Markov property: $\Pr(X_{t+1} = i_{t+1} | X_t = i_n, \dots, X_0 = i_0) = \Pr(X_{t+1} = i_{t+1} | X_t = i_n)$, then X_t is said to be a (discrete-time) Markov chain (MC) with state-space S . The Markov property implies that the future state of the process is independent of the past, given the present. Thus, a MC wanders about the state space, remembering only where it has just been in the last time step. The transition probability from state i to state j , $\Pr(X_t = j | X_{t-1} = i)$, is denoted by p_{ij} . The transition probabilities are usually collected into a matrix, $P = \{p_{ij}\}$, called the transition matrix or transition kernel, where $p_{ij} \geq 0$ and $\sum_{j \in S} p_{ij} = 1$. Herein attention is restricted to time-homogeneous MCs, where the transition probabilities are constant over time. The probability of going from state i to state j in n time steps, $\Pr(X_n = j | X_0 = i)$, is denoted by $p_{ij}^{(n)}$, and the n -step transition matrix is $\{p_{ij}^{(n)}\}$ is denoted by $P^{(n)}$.

If u and v are two non-negative integers such that $n = u + v$, then $p_{ij}^{(n)} = \sum_{k \in S} p_{ik}^{(u)} p_{kj}^{(v)}$. This result is known as the Chapman Kolmogorov equation (Ross 2007). Stated otherwise, $P^{(n)} = P^{(u)} P^{(v)}$, meaning that the n -step transition matrix, $P^{(n)}$, is nothing but P^n .

Let $\pi_i^{(n)}$ denote the probability, $\Pr(X_n = i)$, that the MC is in state i at time n , ($\pi_i^{(n)} \geq 0$ and $\sum_i \pi_i^{(n)} = 1$). The row vector $\pi^{(n)} = \{\pi_i^{(n)}\}$ is called the distribution of the MC at time n , $\pi^{(0)}$ being the initial distribution of the MC. It turns out that $\pi_j^{(n)} = \sum_{i \in S} \{\Pr(X_n = j | X_{n-1} = i)\} \{\Pr(X_{n-1} = i)\}$ i.e., $\pi_j^{(n)} = \sum_{i \in S} \pi_i^{(n-1)} p_{ij}$. In matrix form, $\pi^{(n)} = \pi^{(n-1)} P = \dots = \pi^{(0)} P^n$. So, a MC is entirely determined by the initial distribution and the transition kernel.

The row vector π is said to be a stationary distribution of a MC with transition kernel P if the distribution of the MC is invariant with respect to P , meaning that $\pi = \pi P$ or equivalently, $\pi_j = \sum_{i \in S} \pi_i p_{i,j} \quad \forall j \in S$. In other words, π is a left eigenvector of the transition matrix with eigenvalue 1. The existence and uniqueness of the stationary distribution depend on the structure of the MC as discussed below.

A state j is said to be accessible from another state i , which is denoted by $(i \rightarrow j)$, if there exists an integer $n > 0$ such that $\Pr(X_n = j | X_0 = i) = p_{ij}^{(n)} > 0$. If $j \rightarrow i$ and $i \rightarrow j$, then i is said to intercommunicate with j , which is denoted by $i \leftrightarrow j$. Intercommunication is an equivalence relation, and intercommunicating states are the equivalence classes in this relation. If $i \leftrightarrow j$ for all i and j , then the MC is said to be irreducible.

A state i is said to have period k if any return to state i must occur in multiples of k time steps. Formally, the period of a state is defined as $k = \gcd\{n : \Pr(X_n = i | X_0 = i) > 0\}$, where \gcd stands for greatest common divisor. If the \gcd of return time to any of the states is 1, the chain is said to be aperiodic.

Let $T_i = \min\{n > 0, X_n = i | X_0 = i\}$ denote the first return time to state i . If $\Pr(T_i < \infty | X_0 = i) = 1$, then state i is said to be “recurrent” or “persistent”; otherwise it is said to be “transient.” If a state i is recurrent and $\nu_i = E[T_i | X_0 = i] < \infty$, then i is said to be “positive recurrent” or “non-null recurrent”, otherwise it is said to be “null recurrent”. A finite state MC which is irreducible, positive recurrent and aperiodic is said to be ergodic. Two interconnected states have the same classification: transient, null recurrent or positive recurrent.

The fundamental theorem of Markov chains (Motwani & Raghavan 1995) states that, if a MC is ergodic, then high powers of the transition matrix converge to the rank one matrix with all rows equal to π . i.e., an ergodic MC has a unique limiting distribution, π , that is independent of the initial state, X_0 , and depends only on the transition kernel. Thus, starting from any state i , the n th step of a run of an ergodic MC has a chance close to π_j of being in state j if n is large. A MC with transition kernel P is said to satisfy the local balance (or detailed balance) if there exists a probability measure π such that for any two states i and j ,

$$\pi_i p_{ij} = \pi_j p_{ji} \tag{2.6}$$

Local balance is a sufficient (but not necessary) condition for π to be the stationary distribution of a MC with transition kernel P (i.e. local balance implies stationarity, but not the other way around). Starting with $\pi_j p_{ji} = \pi_i p_{ij}$ and integrating (summing) the two sides with respect to i , one gets $\pi_j \sum_{i \in S} p_{ji} = \sum_i \pi_i p_{ij}$ which yields $\pi_j = \sum_{i \in S} \pi_i p_{ij}$ $\sum_{i \in S} p_{ji} = 1$. On the other hand, $\pi_j = \sum_{i \in S} \pi_i p_{ij}$ implies that π is a left eigenvector of the transition matrix associated with eigenvalue 1 i.e., π is a stationary distribution of the MC. More details on Markov chains can be found in Brémaud (2001) or Ross (2007).

The probability distributions arising in the hierarchical Bayesian framework are usually high-dimensional with unknown normalization constants. This makes it difficult to directly sample from them. As a consequence, a great deal of research has been devoted to developing algorithms that can be used to generate samples from complex probability distributions. One of the most successful methods of this kind is Markov chain Monte Carlo (MCMC) (Gilks *et al.* 1996).

The rationale of MCMC simulation methods is to construct an ergodic Markov chain that has the target distribution of interest as its stationary distribution. This MC can be initialized with any state, being guaranteed to converge to its stationary distribution. After convergence and removal of the pre-convergence part of the chain called the “burn-in”, samples from the stationary distribution of the MC can be used as draws from the target distribution, and inferences can be based on the simulated samples. The two most widely used MCMC algorithms are the Gibbs sampler and Metropolis-Hastings algorithm which are sketched below.

Let $\theta = (\theta_1, \dots, \theta_d)$ denote the d -dimensional parameter vector of interest, and $p(\theta | y)$ be the posterior distribution of θ given the data, y . The Gibbs sampler (Gelfand & Smith 1990; Casella & George 1992) proceeds by sampling each component of θ from its full conditional (the conditional distribution of the focal component with all other components held to their current values). More specifically, the Gibbs sampler proceeds as follows.

1. Pick $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})$ arbitrarily in the support of θ and set $i = 0$

2. Generate $\theta_1^{(i+1)} \sim p(\theta_1 | \theta_2^{(i)}, \dots, \theta_d^{(i)}, y)$
 $\theta_2^{(i+1)} \sim p(\theta_2 | \theta_1^{(i+1)}, \theta_3^{(i)}, \dots, \theta_d^{(i)}, y)$

.....

$\theta_d^{(i+1)} \sim p(\theta_d | \theta_1^{(i+1)}, \dots, \theta_{d-1}^{(i+1)}, y)$

3. Set $i = i + 1$, and repeat steps 2-3 until “convergence”.

Note that Gibbs sampling is only feasible when the full conditionals have well-known forms (i.e. under conditional conjugacy). Otherwise, alternative sampling schemes such as the Metropolis-Hastings (MH) (Metropolis *et al.* 1953; Hastings 1970) algorithm are required.

Let $q(\theta^{prop} | \theta)$ be a proposal kernel, where θ is the current state and θ^{prop} a proposed move from the current state θ . The MH algorithm proceeds as follows:

1. Pick arbitrarily $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})$ in the support of θ and set $i = 0$

2. Generate a proposal θ^{prop} from $q(\theta^{prop} | \theta^{(i)})$

3. Compute $a(\theta^{(i)}, \theta^{prop}) = \min \left\{ 1, \frac{p(\theta^{prop} | y)}{p(\theta^{(i)} | y)} \frac{q(\theta^{(i)} | \theta^{prop})}{q(\theta^{prop} | \theta^{(i)})} \right\}$

Draw u uniformly over the interval $(0, 1)$

If $u < a(\theta^{(i)}, \theta^{prop})$, then set $\theta^{(i+1)} = \theta^{prop}$ otherwise set $\theta^{(i+1)} = \theta^{(i)}$

4. Set $i = i + 1$, and repeat steps 2-4 until “convergence”

An important feature of the MH algorithm is the fact that the normalizing constant of the posterior density cancels out from $a(\theta^{(i)}, \theta^{prop})$ as it is involved in both the numerator and the denominator of the ratio $p(\theta^{prop} | y) / p(\theta^{(i)} | y)$. If in addition the proposal kernel q is selected to be symmetric, meaning that $q(\theta^{prop} | \theta^{(i)}) = q(\theta^{(i)} | \theta^{prop})$, then $a(\theta^{(i)}, \theta^{prop})$ takes the simpler form $a(\theta^{(i)}, \theta^{prop}) = p(\theta^{prop} | y) / p(\theta^{(i)} | y)$.

Random walk, $\theta^{prop} = \theta^{(i)} + \eta$, provides a useful example of a symmetrical proposal kernel, where $\theta^{(i)}$ is the current state and η is a zero-mean random disturbance to the current state, typically chosen to be normal. The proposal density is typically required to have no thinner tails than the target distribution. For Gaussian disturbance, Random walk proposals have the form $\theta^{prop} = \theta^{(i)} + s \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} \sim \text{MVN}(0, \Sigma)$ and s is a nonnegative constant which is often tuned to achieve an acceptance rate of between 20% and 40%. The tuning of s is

typically done by trial and error, by running the chain for different values of s and monitoring the acceptance rate.

It is important to ensure that the chains have mixed and to have a sense that further simulation will not change inferences significantly. There are a number of tools for assessing the convergence of MCMC (see e.g. Gelman *et al.* 2003 for details). Visual inspection of traceplots and autocorrelation functions is an informal approach to assessing the convergence and the mixing of MCMC.

In situations where only some components of the parameter vector can be sampled from directly, a hybrid Gibbs/Metropolis sampler may be required. This amounts to updating the parameters in blocks, where each block is altered using the Gibbs sampler or a Metropolis jump (Tierney 1994; Gelman *et al.* 2003).

In practice, MCMC can be implemented using available software such as WinBUGS (Spiegelhalter *et al.* 2003) or OpenBUGS (Thomas *et al.* 2006). These are Windows versions of the BUGS software program for Bayesian statistical analysis using MCMC, which can be used to model a wide variety of complex statistical problems.

The advent of computer intensive statistical methods such as MCMC has greatly impacted applied statistics by opening new horizons in the scale of the problems that one can deal with, thus enhancing the position of statistics in most applied fields. Model selection and model validation are also important aspects of Bayesian inference.

2.3. Bayesian model selection

Choosing amongst alternative models or scientific hypotheses is a fundamental problem faced by researchers in any scientific discipline. Model selection can be viewed as a wide scale testing problem where models rather than parameters are of interest (Robert 2001). In the Bayesian framework, the prominent model selection techniques include Bayes factors, (Kass & Raftery 1995), the deviance information criterion (Spiegelhalter *et al.* 2002), variable selection methods such as stochastic search variable selection (George McCulloch 1993), and posterior predictive model assessment (Gelman *et al.* 1996, 2003). A brief account of these methods is given below.

2.3.1 Bayes factors

Bayes factors (Kass & Raftery 1995) provide a natural means of evaluating the support provided by the data in favor of one model or scientific hypothesis against another, using rules of probability theory. Suppose that one is interested in comparing two models (or hypotheses) M_i and M_j with respective parameters θ_i and θ_j . Given that models are considered as unknowns, prior distributions are required on them as well. If $p(M)$ denotes the prior for model M , the posterior odds ratio of model M_i versus M_j is given by

$$\frac{p(M_i | y)}{p(M_j | y)} = \frac{p(y | M_i) p(M_i)}{p(y | M_j) p(M_j)}, \quad (2.7)$$

where $p(y|M_k) = \int_{\Theta_k} p(y|\theta_k, M) p(\theta_k|M) d\theta_k$. The first term on the right-hand side of equation (2.7) is called the Bayes factor (BF) of M_i versus M_j , which is denoted here by B_{ij} . In essence, B_{ij} evaluates the evidence provided by the data in favour of M_i against M_j . More explicitly, it is the amount by which the observed data change the prior odds of model M_i over M_j into posterior odds: $BF = \text{Posterior Odds} / \text{Prior Odds}$. If both hypotheses are simple and $p(M_i) = p(M_j)$, then B_{ij} boils down to the usual likelihood ratio. One says that the data provides support for M_i over M_j if $B_{ij} > 1$, and *vice-versa*. Jeffreys (1961) proposed the following scale for interpreting Bayes the factor B_{12} of H_1 against H_2 .

$BF < 1$: Negative (i.e. support for H_2); $1 \leq BF < 3$: Barely worth mentioning evidence for H_1 ; $3 \leq BF < 10$: Substantial support for H_1 ; $10 \leq BF < 100$: Strong support for H_1 ; $BF > 100$: Decisive support for H_1

It is worth emphasizing that strong support for M_i versus M_j need not translate into high posterior probability for M_i and a low one for M_j . Both models might be improbable, M_j being more unlikely than M_i . Bayes factors can be difficult to compute in practice. In particular, if $p(\theta)$ is improper, then $p(y|M)$ necessarily is, and Bayes factor is not well defined. Plummer (2008) also pointed out that Bayes factors are numerically unstable when proper, but diffuse priors are used.

2.3.2. The deviance information criterion

From a decision theoretic perspective, model selection can be cast in terms of minimizing a loss function appropriate to the decision problem at hand. A general loss function based the likelihood function is the deviance, $D(y, \theta) = -2 \log(L(y|\theta))$, where $L(y|\theta)$ denotes the likelihood function and $\log(x)$ is the natural logarithm of x . $D(y, \theta)$ is minimized as the corresponding utility function, the (log)-likelihood function, is maximized. Model selection can thus proceed by minimizing the deviance $D(\hat{\theta}) = -2 \log(L(y|\hat{\theta}))$, where $\hat{\theta}$ is the maximum likelihood estimate (MLE) of θ .

However, the introduction of extra parameters will result in an increased likelihood, regardless of whether or not those new parameters are actually relevant. As a consequence, a simple comparison of the maximum likelihoods (or alternatively the deviances) of different models will always favour the model with more parameters. Information criteria (basically penalized likelihood measures) have been developed to overcome this bias towards higher dimensional models. Their principle is to penalize the more complex models to offset any improvement in the maximum likelihood that may arise from the extra parameters.

In the classical framework, the most popular information criteria are the Akaike information criterion (AIC) $AIC = -2 \log(L(y|\hat{\theta})) + 2k$ (Akaike 1973), and the Bayesian information criterion (BIC) or Schwarz criterion $BIC = -2 \log(L(y|\hat{\theta})) + k \log(n)$ (Schwarz 1978), where k is the number of free parameters, and n is the sample size.

In the Bayesian framework, the deviance information criterion (DIC) introduced by Spiegelhalter *et al.* (2002) is widely used. The DIC is defined as

$$\text{DIC} = \bar{D} + P_D. \quad (2.8)$$

$\bar{D} = E[D(\theta | y)] = -2 \int_{\Theta} [\log(L(y | \theta))] p(\theta | y) d\theta$ is the posterior mean of the deviance which is interpreted as a measure of fit, whereas the “effective number of parameters” $P_D = \bar{D} - \hat{D}$ acts as a penalty for model complexity, where $\hat{D} = D(E[\theta | y])$ is the deviance evaluated at the posterior mean of the model parameters.

The model with the smallest DIC is favoured as this indicates the best balance between goodness of fit and model complexity. In Article II, the DIC was used, along with the mean squared error (MSE) to investigate whether including environmental noise in one or another parameter makes a difference to the fit and predictive performance of population dynamical models.

2.3.3. Stochastic search variable selection

Stochastic search variable selection (SSVS) (George McCulloch 1993) is a Bayesian procedure for selecting promising subsets of predictors in a regression set-up. Its rationale is to embed a multiple regression in a hierarchical normal mixture model, where latent indicators are used to identify the relevant predictors. To keep things simple, the sketch of SSVS given below is restricted to the multiple linear regression setting.

Let $y = (y_1, \dots, y_n)^T$ denote observations of the response variable and $x = (x_1, \dots, x_J)$ be a set of J potential predictors for the responses y_i . A linear regression of y over x has the form $y_i = \beta_0 + \sum_{j=1}^J \beta_j x_j + \varepsilon_i$, where $\varepsilon_i \sim N(0, \sigma^2)$, independently for $i = 1, \dots, n$. In order to find a subset of the predictors with significant effects, SSVS utilizes indicator variables γ_j , one for each covariate x_j ($j = 1, \dots, J$), such that γ_j takes the value one when x_j is included in the model, and is zero otherwise. Bernoulli priors, $\gamma_j \sim \text{Bern}(p_j)$, ($0 < p_j < 1$), are assigned to the indicator variables, where p_j quantifies the *a priori* expectation that x_j needs to be included into the model.

The prior distribution of β_j is defined as a mixture of two normal distributions, conditional on the indicator γ_j . That is, $\beta_j | \gamma_j \sim (1 - \gamma_j) N(0, c_1) + \gamma_j N(0, c_2)$, where the constant $c_1 > 0$ is selected to be very small, whereas $c_2 > 0$ is set to be large. With this prior specification, the effect β_j is forced to be close to zero when $\gamma_j = 0$ since the corresponding prior is a narrow “spike” clustered around zero, while being freely estimated from the data when $\gamma_j = 1$ since the corresponding “slab” component of the prior is diffuse.

A Gibbs sampling methodology is used to generate samples from the joint posterior $p(\beta, \sigma^2, \gamma | y)$. The covariates with significant effects can be identified as those with higher

inclusion probabilities. The importance of a single predictor x_j can also be evaluated through the Bayes factor,

$$\frac{\Pr(\gamma_j | y) / [1 - \Pr(\gamma_j | y)]}{p_j / (1 - p_j)},$$

which evaluates the evidence provided by the data in favour of including *versus* not including x_j into the model. This approach was used in Articles **III** and **IV** to constrain the effects of spurious inter-specific interactions to be close to zero, so that they do not affect the amount of variance attributable to different factors of fluctuation in species abundances.

2.4. Model assessment through posterior predictive checks

Model assessment implies the willingness to evaluate a statistical model in order to discover what (if anything) may be deficient about the model at hand, and find ways of remedying the problem, independently of any other model. There are many ways of assessing a model e.g., checking the distributional assumption about residuals in a regression setting. A standard method for Bayesian model assessment is posterior predictive checking (Gelman *et al.* 1996, 2003), where some function of the observed data is compared to its counterpart based on replicated data from the fitted model, and systematic discrepancies between the two taken as evidence for potential failings of the model. One approach is posterior predictive cross-validation which consists of omitting some of the observed data, and forecasting them to find out whether the omitted data are consistent with their posterior predictive distributions. The posterior predictive assessment is made in terms of posterior probability statements about the posterior predictive distribution, $p(y.rep | y)$, where y denotes the observed data and $y.rep$ the replicated data from the fitted model. This approach to model assessment was utilized in article **IV**.

3. A PROBABILISTIC TEST OF THE NEUTRAL THEORY OF BIODIVERSITY

The neutral theory of biodiversity (Bell 2000; Hubbell 2001) is one of the hottest topics in ecology right now. It refers to the underlying assumption that traits differences between trophically similar species have no impact on their relative abundances or their speciation rates. Although empirical evaluations of the neutral theory are becoming quite common, tests that elucidate the underlying mechanisms are still missing. Most of previous tests of the neutral theory have been limited to examining the consistency of empirical abundance-frequency patterns of local communities with the equilibrium predictions of the neutral theory, using snapshot data or temporally pooling several samples.

A sensible way of evaluating a dynamical model is to fit the model to temporal data and ask whether the resulting parameter values are sensible. In Article I a discrete-time version of the “unified neutral theory of biodiversity and biogeography” (UNTB) (Hubbell 2001), with the zero-sum assumption relaxed, was fitted to long-term macro-moth (Lepidoptera) community data from the Rothamsted Insect Survey (RIS) network in the UK, to see how well the neutral model would explain the changes in species abundances.

More specifically, let $N_{i,t}$ denote the abundance of species i at time t (S species in total), in the focal community assumed to be dynamically coupled to a metacommunity through immigration. Neutral dynamics were modelled by assuming that the expected number, $\lambda_{i,t}$, of individuals of species i at time t is given by

$$\lambda_{i,t} = J_{t-1} [(1 - m_t) \times C_{i,t-1} + m_t \times P_i], \quad (3.1)$$

where $J_t = \sum_{i=1}^S N_{i,t}$ is the total community size at time t , m_t is the immigration rate at time t , which is assumed to be equal across species (under neutrality, immigrants are taken randomly from the metacommunity, their relative abundances being the sole arbiter of the probability that an immigrant is of a particular species), $C_{i,t} = N_{i,t} / J_t$ is the relative abundance of species i in the community at time t , and P_i is the metacommunity relative abundance of species i , which was assumed to be constant on the ecological time-scale of the data sets considered in Article I, and speciation was not included into the model.

Since neutrality implies no selective difference between species, the fraction $(1 - m_t)$ of individuals of species i produced locally will have the same expected relative frequency as in the previous generation (i.e. $C_{i,t-1}$). Equation (3.1) is essentially a discrete-time version of the UNTB, except that the zero-sum is relaxed.

Denoting by $JP_{i,t}$ the composite variable $J_t \times P_i$ and recognizing that $J_{t-1} \times C_{i,t-1} = N_{i,t-1}$, equation (3.1) can be written as $\lambda_{i,t} = [(1 - m_t) \times N_{i,t-1} + m_t \times JP_{i,t}]$. Random drift was incorporated into the model by thinking of $\lambda_{i,t}$ as a birth rate, so that $N_{i,t}$ can be assumed to be distributed as

$$N_{i,t} \sim \text{Poisson}(\lambda_{i,t}). \quad (3.2)$$

Further, the observed abundance of species i at time t , $y_{i,t}$, was assumed to be distributed as

$$y_{i,t} \sim \text{Poisson}(N_{i,t} \times q_t), \quad (3.3)$$

where the “sampling rate”, q_t , represents the proportion of $N_{i,t}$ which has been observed. In a similar fashion to the variance effective population size (e.g. Berthier *et al.* 2002), J_t is the size of a neutral community that would show the same amount of variation as the focal community. If the neutral model is a reasonable description of the community dynamics, then the sampling rates should lie between 0 and 1. This can be tested out by relaxing the prior on the sampling rates to be e.g., $\text{Gamma}(3, 6)$ which has the same expected value and variance as the uniform distribution between 0 and 1, but with wider support.

By not restricting the parameter range, one can check whether the neutral model can produce sensible results with realistic community sizes. Under neutrality, the sampling rate corresponds to the probability of capture and should lie between 0 and 1, with the value 1 corresponding to the case where the community is completely observed. For partially observed communities like the Lepidoptera communities considered in Article I, neutral dynamics should imply sampling rates that are necessarily between 0 and 1. Consequently, the neutral model can be rejected if it predicts sampling rates that are higher than 1.

The model was fitted to the data with a Bayesian approach, and the analyses were carried out both for the full macro moth data sets and for the geometrids (*Geometridae*) only. Geometrids are a moth family whose members are known to respond in a similar way to light and are sampled particularly efficiently by the Rothamsted Insect Survey light traps (Taylor & French 1974). Figure 1 shows the 95% credible intervals of the community sizes, the sampling rates and the immigration rates for the geometrids.

These results suggest that the neutral model would need sampling rates that are higher than 1 in order to fit to the data well. This is unrealistic as it implies community sizes that are lower than the observed numbers. So, directly fitting the neutral model to the data showed that it is unrealistic.

An obvious explanation for the poor performance of the neutral model is the fact that it attributes all the dynamics to chance, ignoring many potentially important factors namely (1) environmental stochasticity which is known to affect the dynamics of populations regardless of their sizes (Lande *et al.* 2003; Chase 2005), so it is not surprising that we find it in community dynamics as well; (2) density-dependence, which has been shown to occur in a larger set of similar RIS moth data (Woiwod & Hanski 1992), and (3) competitive asymmetries across species.

The next section is devoted to the development of a more comprehensive framework designed to accommodate the mechanisms overlooked by the neutral perspective. The model is used to analyze community time series data from a variety of taxa for evaluating the relative importance of different factors.

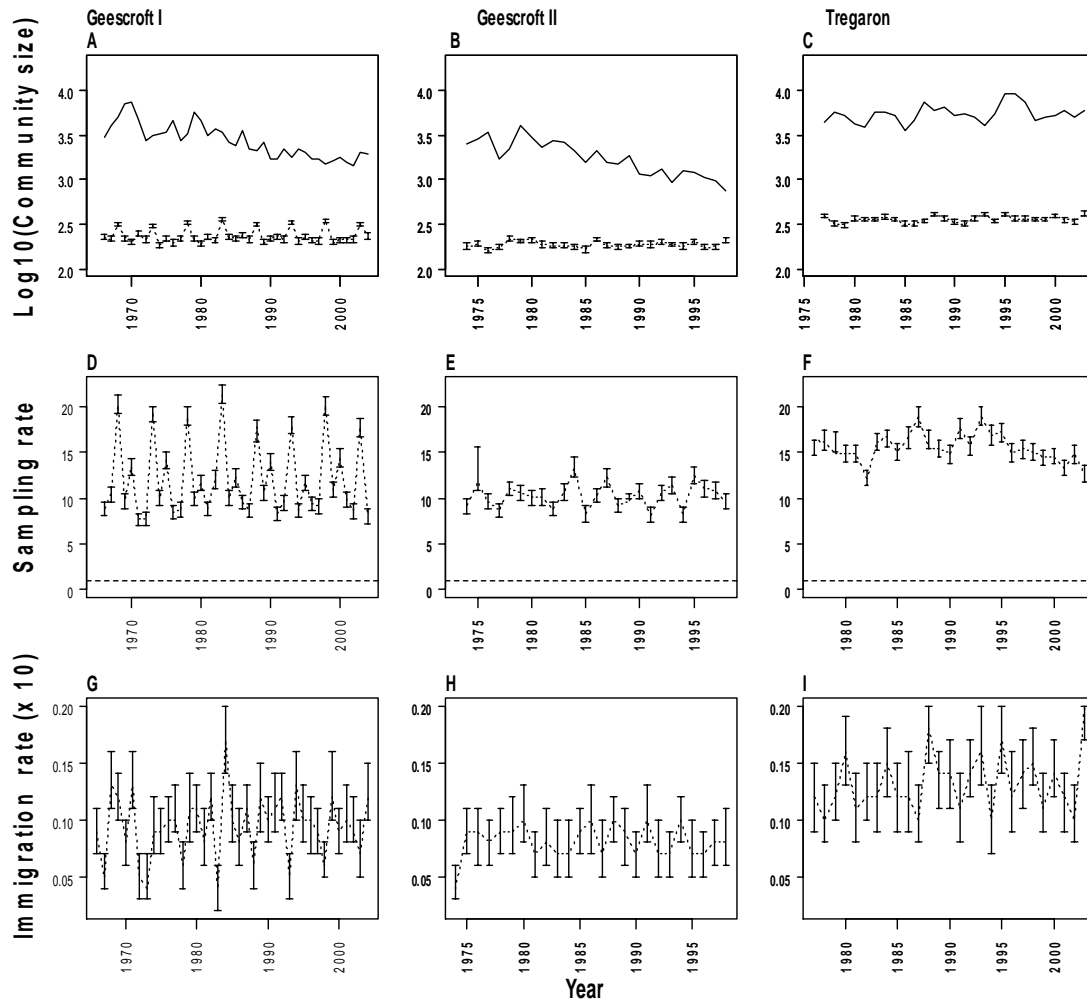


Figure 1: Posterior medians and 95% credibility sets for the estimated community sizes and the observed values (solid lines): (A–C); posterior medians and 95% credibility sets for the estimated sampling rates: (D–F) and for 10 x immigration rates: (G–I). The dashed horizontal lines in panels D - F are drawn at height 1. These results concern the geometrid species only.

4. THE GENERAL FRAMEWORK

Ecological data are often fraught with measurement errors and missing data (Clark & Bjørnstad 2004). The state-space model formulation (de Valpine & Hastings 2002; Buckland *et al.* 2004; Clark & Bjørnstad 2004; Rivot *et al.* 2004; Gimenez *et al.* 2007) provides a natural approach to integrating the process model which describes the dynamical behavior of a system, usually by conditional Markovian transition between successive states, and the observation model (or sampling model) intended to map the observed data to the states of the underlying process with regard to sampling errors. The general framework described in this section assumes Gaussian noise for both the process and the sampling model.

4.1. The process model

The underlying model assumed for species abundance dynamics is a discrete-time stochastic Gompertz model (Jacobson *et al.* 2004; Dennis *et al.* 2006). That is,

$$N_{i,t} = N_{i,t-1} \exp \left\{ r_i \left[1 - \sum_{j=1}^S \alpha_{i,j} \log N_{j,t-1} / k_i \right] + \varepsilon_{i,t} \right\}, \quad t = 2, 3, \dots, T, \quad (4.1)$$

where S is the total number of species, $N_{i,t}$ is the number of individuals of species i at time t , r_i and k_i are respectively the intrinsic growth rate and the natural logarithm of the carrying capacity of species i ; $\alpha_{i,j}$ is the coefficient of interaction between species i and j , which measures the per capita effect of species j on the growth of species i , with all intra-specific coefficients, $\alpha_{i,i}$, set to 1, and $\varepsilon_{i,t}$ is a zero-mean random disturbance affecting the dynamics of species i at time t , and assumed to be Gaussian (see below). On the natural logarithmic scale, equation (4.1) becomes

$$n_{i,t} = n_{i,t-1} + r_i \left[1 - \sum_{j=1}^S \alpha_{i,j} n_{j,t-1} / k_i \right] + \varepsilon_{i,t} \quad (t = 2, 3, \dots, T), \quad (4.2)$$

where $n_{i,t}$ denotes the natural logarithm of $N_{i,t}$. Equation (4.2) can be compactly written in matrix form as

$$\mathbf{n}_t = \mathbf{n}_{t-1} + \mathbf{R}(\mathbf{1}_S - \mathbf{A}\mathbf{n}_{t-1}) + \boldsymbol{\varepsilon}_t, \quad \text{for } t = 2, 3, \dots, T, \quad (4.3)$$

where \mathbf{n}_t denotes the S -dimensional vector of log-abundances of all species at time t , \mathbf{R} is a S -by- S diagonal matrix with $\mathbf{R}_{i,i} = r_i$, and $\mathbf{1}_S$ is the S -dimensional vector with all elements equal to 1; $A_{i,j} = \alpha_{i,j} / k_i$, and $\boldsymbol{\varepsilon}_t = (\varepsilon_{t,1}, \dots, \varepsilon_{t,S})^T$ is the vector of process disturbances affecting the community dynamics at time t , with one element by species. The vectors $\boldsymbol{\varepsilon}_t$ are set to be serially independent and multivariate normally distributed around the zero-vector, with a covariance matrix denoted by Σ_t . This covariance matrix is further decomposed as

$$\Sigma_t = \mathbf{C} + \mathbf{D}, \quad (4.4)$$

where $\mathbf{D}_t = \text{diag}(\delta_i^2 / N_t)$ and \mathbf{C} are respectively the demographic and environmental components of the underlying dynamics. Note the inverse scaling of demographic stochasticity with the population size (e.g. Sæther *et al.* 2000; Bjørnstad & Grenfell 2001), implying that demographic stochasticity can be neglected when dealing with large populations.

Correlations in species' responses to environmental fluctuations are accommodated by specifying \mathbf{C} as a full matrix with diagonal elements, $C_{i,i}$, and off-diagonal elements, $C_{i,j}$ ($i \neq j$), representing species-specific and common responses to environmental fluctuations, respectively. Consequently, the correlation between the responses of species i and j to the environmental disturbances is evaluated by $\rho_{i,j} = C_{i,j} / (C_{i,i} C_{j,j})^{1/2}$.

4.2. Variance partitioning

The temporal fluctuations in the abundance of individual species can be decomposed into relative contributions from intra-specific interactions, inter-specific interactions, and environmental forcing. More specifically, let $z_{i,t} = n_{i,t} - n_{i,t-1}$ denote the change in the log-abundance of species i from time $t-1$ to time t , $z_{i,\cdot} = \{z_{i,t}\}_{t=1}^T$ and let $v_{i,i}$ designate the (stationary) variance of $n_{i,\cdot} = \{n_{i,t}\}_{t=1}^T$. In settings where demographic stochasticity can be neglected (i.e. for large populations), $\text{var}(z_{i,\cdot})$ can be additively decomposed into the contributions $v_{i,i} (r_i / k_i)^2$, $(r_i / k_i)^2 \sum_{i \neq j} \alpha_{i,j}^2 v_{j,j}$, and $C_{i,i}$ from intra-specific interactions, inter-specific interactions, and environmental forcing, respectively. This methodology was applied in Article III to analyze community data from a range of taxa: macro-moths (Lepidoptera), fish, crustaceans, birds, and rodents.

Environmental forcing accounted for most of the temporal variation, followed by intra-specific interactions. The contribution of inter-specific interactions was found to be broadly minor. Figure 2 shows the posterior distributions of the proportion of temporal variation attributable to environmental forcing and to intra-specific interactions in the dynamics of moths, fish, and crustaceans.

4.3. The sampling model

A convenient setting for accommodating sampling error is when replicated samples are available. Assume k replicates (e.g. data from k different traps on the same site) are available, and let $y_{i,t,k}$ denote the natural logarithm of the observed number of individuals of species i at time t , from trap k . Under Gaussian observation errors, a sensible sampling model that takes into account differences in capture probabilities between species across traps is given by

$$y_{i,t,k} | n_{i,t} \sim \text{N}(n_{i,t} + b_{i,k}, \tau_i^2) \quad (t=1, 2, \dots, T), \quad (4.5)$$

where τ_i^2 is the sampling variance specific to species i . The random variable $b_{i,k} \sim N(0, \tau_b^2)$ is intended to correct for possible differences in trapping efficiency between traps with regard to species i , and is set to zero for one of the replicates (one of the traps), to force identifiability.

This sampling model was used in Article IV to analyze macro-moth (Lepidoptera) light-trapping abundance time series from the Rothamsted Insect Survey network in the UK, using replicated samples from two traps (Geescroft I and Geescroft II) on Rothamsted farm in Hertfordshire, UK (Woiwod & Harrington 1994), with all values of the random variable b set to zero for Geescroft I as discussed above.

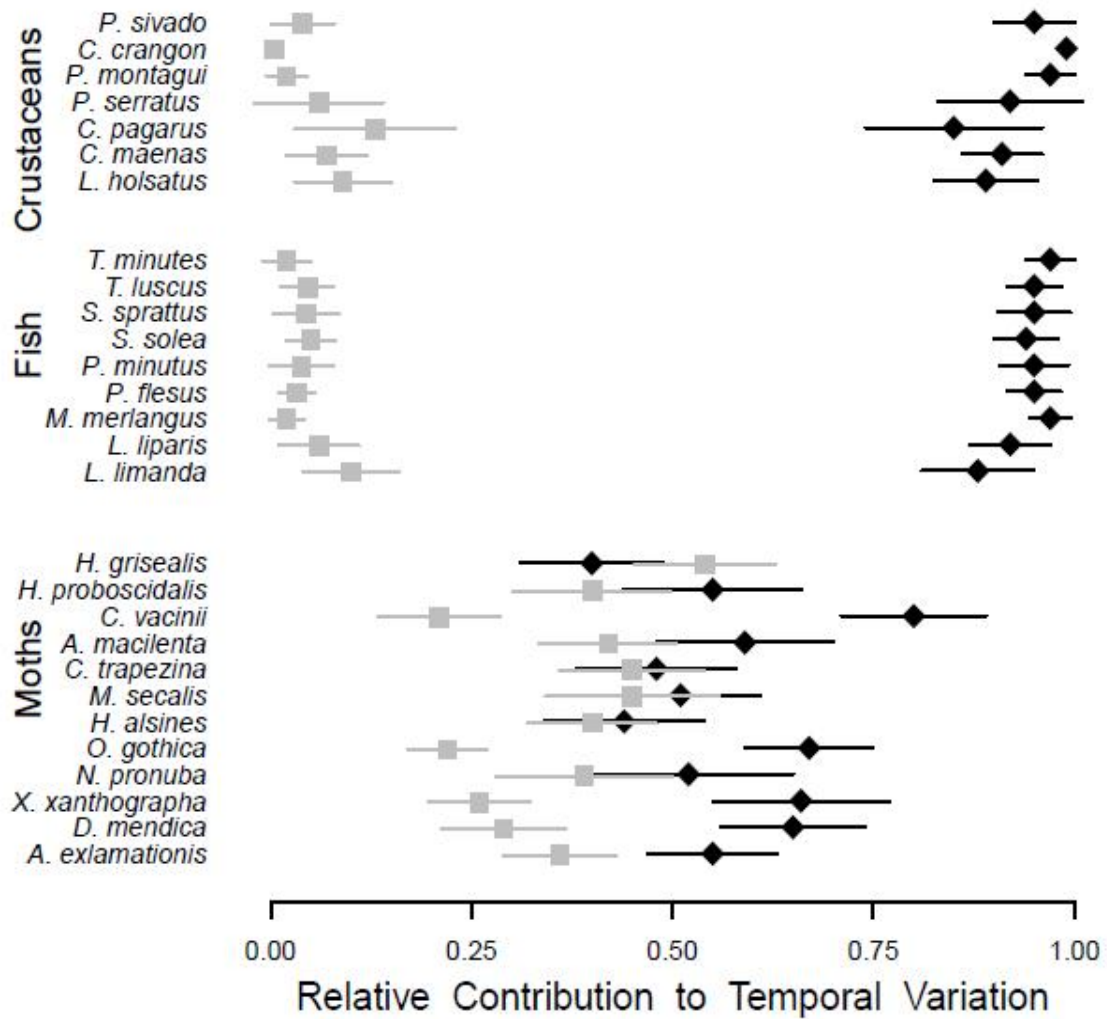


Figure 2: Error-bars (mean \pm 1SD) for the proportions of variance attributable to environmental forcing (black diamonds) and to intra-specific interactions (grey boxes) in the dynamics of individual moth, fish, and crustacean species.

4.4. Incorporating environmental covariates

The model described by equations (4.2) & (4.3) can be extended to explicitly include the effects of K environmental covariates X_1, \dots, X_K as

$$\mathbf{n}_t = \mathbf{n}_{t-1} + \mathbf{R}(\mathbf{1}_S - \mathbf{A}\mathbf{n}_{t-1}) + \boldsymbol{\beta} \mathbf{X}_t + \boldsymbol{\varepsilon}_t \quad (4.6)$$

where $\mathbf{X}_t = (X_{1,t}, \dots, X_{K,t})^\top$, $\boldsymbol{\beta}_i^\top = (\beta_{i,1}, \dots, \beta_{i,K})^\top$, and $\beta_{i,k}$ quantifies the strength of dependency of the growth rate of species i on the environmental variable \mathbf{X}_k . All environmental variables \mathbf{X}_k are assumed to be standardized to have zero-mean and unit variance.

In article **IV**, the model (4.6) was fitted to replicated light-trapping data of twelve noctuid (*Noctuidae*) macro-moth (Lepidoptera) from the Rothamsted Insect Survey network in the UK. The data used in Article **IV** came from two traps: Geescroft I and Geescroft II sampled on Rothamsted Farm in Hertfordshire UK (Woiwod & Gould 2008), and covered the period 1973-2003 for Geescroft I and 1973-1998 for Geescroft II. Two weather variables namely, mean winter (Dec-Feb) rainfall (in mm), mean winter temperature were used as environmental covariates.

The model fitting was carried out by MCMC simulation via OpenBUGS (Thomas *et al.* 2006). The model adequacy was evaluated through posterior predictive cross-validation by leaving out the last five observations for Geescroft I and using the model to forecast them and determine how well the model predictions would approximate the omitted data. The model predictions were consistent with the discarded data

Figure 3a summarizes the posterior distributions of the parameters evaluating the effects of the two weather variables on individual species. Figure 3b summarizes the posterior distributions of the variable b representing the trapping efficiency of the Geescroft II trap relative to Geescroft I for each species, with all values corresponding to Geescroft I set to zero to force identifiability. Negative values of b represent lower capture efficiency of Geescroft II relative to Geescroft I. Fig. 3c shows the relative contribution of environmental variation (not including inter-specific interactions) as well as intra- and inter-specific interactions to the total variation in the dynamics of individual species, and Fig. 3d shows the proportion of environmental variance explained by the included weather variables both individually and collectively.

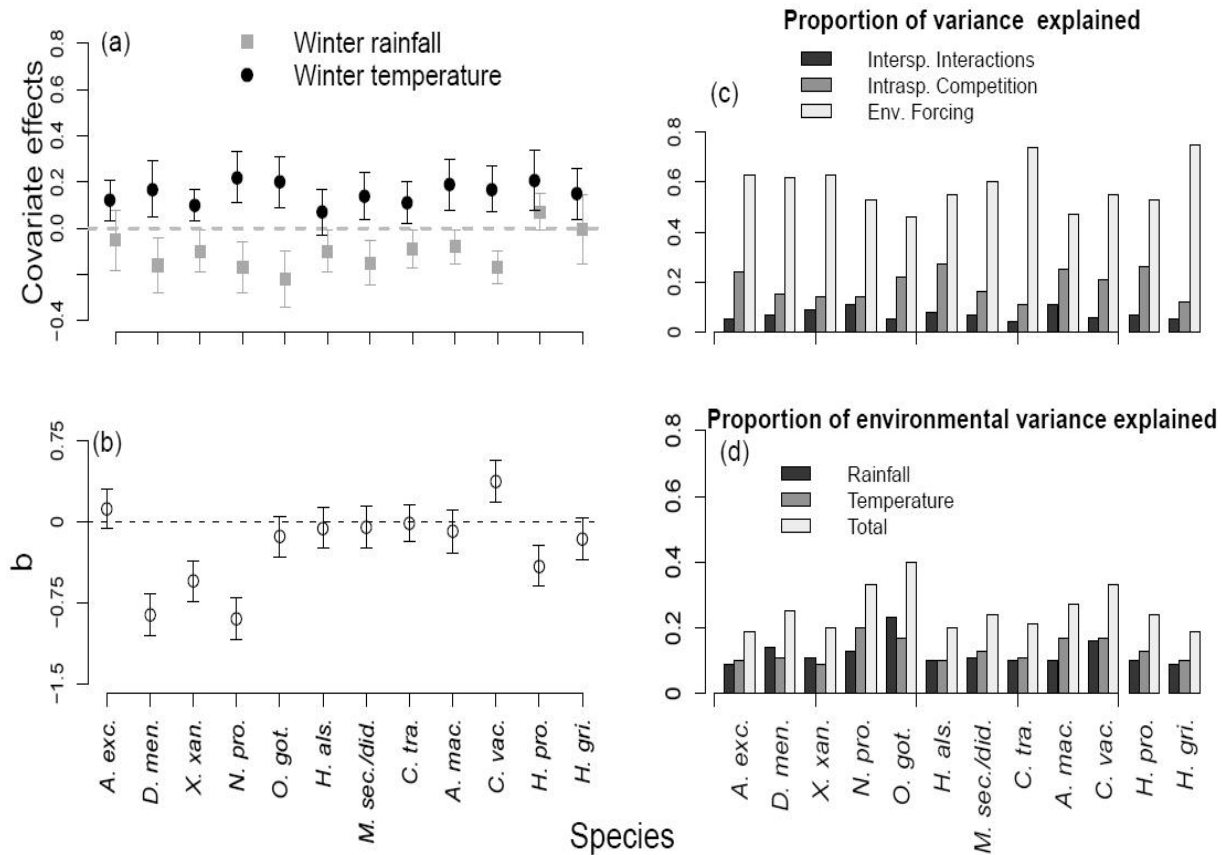


Figure 3: (a) Error-bars (mean \pm SD), representing 68% credible intervals for the effects of winter rainfall (grey squares) and winter temperature (black circles) on the growth of individual species; (b) posterior means and 68% credible intervals for the variable *b* representing the efficiency of the Geescroft II trap relative to Geescroft I for each species, with all values corresponding to Geescroft I set to zero; (c-d) barplots for the proportions of environmental variance attributable to environmental stochasticity, as well as intra-/inter-specific interactions in the dynamics of individual species (c), and for the proportion of environmental variance explained by winter rainfall and winter temperature, both individually and collectively (d).

5. INTEGRATING NEUTRALITY AND NICHES

Many ecological scholars recognize that trait-mediated and stochastic factors operate simultaneously to generate the biodiversity patterns observed in nature (Chave *et al.* 2002; Gravel *et al.* 2006; Héraut 2007). Accordingly, it has been suggested that neutrality and stringent niche segregation can be viewed as extremes of a continuum (Gravel *et al.* 2006). There is a call for comprehensive frameworks within which the relative importance of niche-mediated and neutral factors processes can be evaluated in light of field data. However, attempts to make the niche-neutral continuum operational remain sparse.

This subsection describes a model-based approach to operationalizing the niche-neutral continuum. The methodology connects the niche configuration (i.e. the way species break up the resource pool) to community dynamics through a mechanistic description of interaction coefficients, and includes demographic stochasticity to accommodate the stochastic character of species abundance dynamics. Species' niches are defined in terms of their resource utilization kernels.

Utilization of a resource with attributes x by a species i is described by a probability density function $f_i(x)$. If the resource is uniform, $f_i(x)$ describes the probability density (or probability mass function for resources with discrete states) for which resource state x an individual of species i will be found using, and is referred to as species i 's resource utilization curve (RUC).

Competition between species is assumed to result from overlaps in their RUCs, with the competition being more or less intense depending on the degree of overlaps in species niches and the niche breadths of potential competitors. This is how the niche aspect is incorporated into the model.

Several niche breadth and niche overlap have been proposed (see e.g. Colwell & Futuyma 1971; Hurlbert 1978; or Abrams 1980 for a review). Herein, the Levins' (1968) niche breadth index for species i , defined as $w_i = 1/\int f_i^2(x) dx$ is adopted. The formulae for niche breadth, niche overlap, and competition coefficient given here also apply to resources occurring in discrete states, with sums in place of integrals and probability mass functions in place of probability density functions. Following Pianka (1974), we define the overlap, O_{ij} , in the niches of species i and j as

$$O_{ij} = \int f_i(x) f_j(x) dx / \sqrt{\int f_i^2(x) dx \int f_j^2(x) dx} \quad (5.1)$$

Intuitively, O_{ij} quantifies the probability that two individuals found in a given resource state belong to species i and j relative to the geometric mean of the probabilities that they both belong to either species (the denominator of eqn. 5.1). Pianka's niche overlap index takes values between zero (no resources used in common) and one (complete overlap). In particular, the overlap of a species' niche with itself is 1, so $O_{ii} = 1$.

Along the lines of MacArthur & Levins (1967) and Simberloff (1982), the competitive effect of species j on species i , resulting from overlap in utilization of a single resource is defined as

$$\alpha_{ij} = \int f_i(x) f_j(x) dx / \int f_i^2(x) dx . \quad (5.2)$$

Intuitively, α_{ij} evaluates the probability that two individuals found in a given resource state belong to species i and j , relative to the probability that they both belong to species i . The full effect of competition from species j on species i at a given t is $\alpha_{i,j} n_{j,t}$. This completes the link between niche configuration and community dynamics.

In this setting, the competitive effect of species j on species i , α_{ij} , is related to the two species' niche characteristics through

$$\alpha_{ij} = O_{ij} \sqrt{w_i / w_j}, \quad (5.3)$$

MacArthur & Levins (1967) suggested that under Gaussian RUCs,

$$O_{ij} = \exp\left\{-\frac{(\mu_i - \mu_j)^2}{2(w_i^2 + w_j^2)}\right\}, \quad (5.4)$$

where μ_i and μ_j are the niche modes, and w_i and w_j the niche breadths (standard deviations of the RUCs) for species i and j , respectively. If in particular the niche breadths are equal

i.e. $w_i = w_j = w$, then $O_{ij} = \exp(-d_{i,j}^2 / 4w^2)$, where $d_{i,j} = |\mu_i - \mu_j|$ (Begon *et al.* 1996).

Figure 4 illustrates the connection between the niche configuration and the extent of competition experienced by a species i from another species j . The difference in species niche breadths is represented by the ratio of the niche breadth of species i to that of species j . The quantity

$$\phi = \frac{\sum_{i \neq j} b_{i,j}}{(S-1) \sum_{i=1}^S b_{i,i}},$$

where $b_{i,j} = -r_i \alpha_{i,j} / k_i$, can be used to evaluate the departure of the dynamical behaviour of a community from neutral drift, as a proxy for the community position on the niche-neutral continuum. The metric ϕ takes values between 0 and 1 with the extremes 0 and 1 corresponding to strict niche separation and neutral drift, respectively.

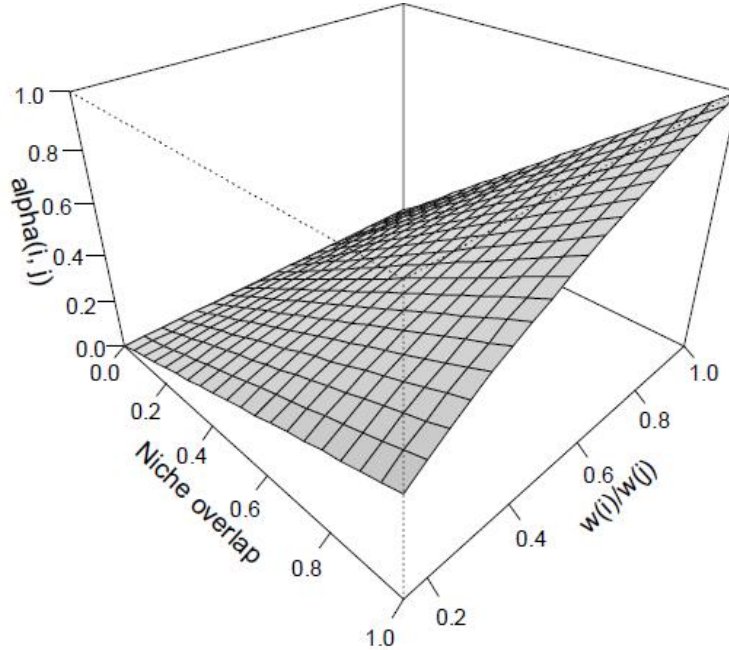


Figure 4: Illustration of the connection between niche configuration and the extent of competition experienced by species i from species j . The x-axis is the niche overlap, the y-axis represents the ratio of the niche breadths of species i to that of species j , and the z-axis is the competitive effect of species j on species i .

This methodology was used in Article V to analyze rodent web-trapping data from the Sevilleta National Wildlife Refuge in central New Mexico, USA (Figgens 2006), involving two species: *Dypodomys ordii* (Ord's Kangaroo rat) and *Dypodomys spectabilis* (Bannertail Kangaroo rat). Replicated time series of species abundances were analyzed under the process model described in equation (4.1), and the resulting competition coefficients were used to reconstruct the niche configuration along a single niche axis with continuous and normally-distributed resource states, using eqns (5.3) and (5.4).

The replicated nature of the data (the sampling webs were monitored for three consecutive nights, which provides three replicated samples) was used to estimate the capture probabilities across species over time. When dealing with count data, it might be desirable to use a sampling model that honors the discrete nature of the data. For example

$$Y_{i,t,k} | p_{i,t} \sim \text{Bin}(N_{i,t}, p_{i,t}) \quad (5.5)$$

(Royle & Dorazio 2006), where $N_{i,t}$ denotes the actual number of individuals of species i at time t , $Y_{i,t,k}$ is the observed number of individuals of species i at time t in the k th replicate, and $p_{i,t}$ represents the detection probability for species i during year t .

Figure 5 shows the niche configuration for *D. ordii* and *D. spectabilis* resulting from the competition coefficients derived from species abundances time series using equations (5.3) and – (5.4), with the niche mode and niche breadth of one of species (*D. ordii*) set to zero and 1, respectively for identifiability

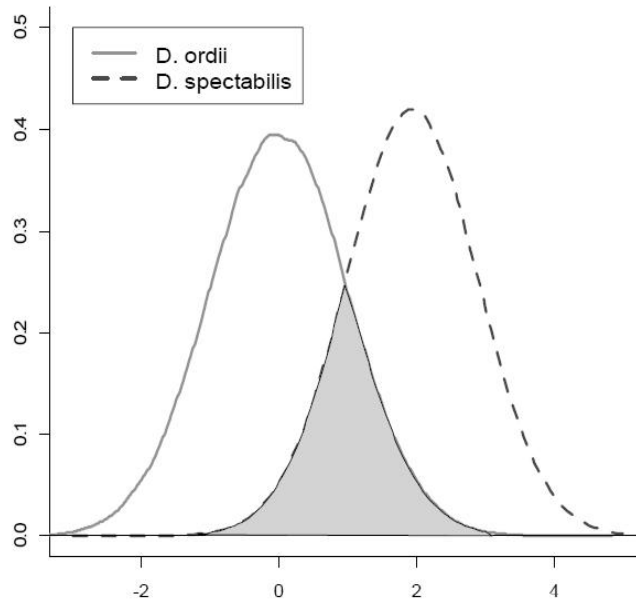


Figure 5: Niche configuration implied by the competition coefficients derived from time series of species abundances, with the niche mode and niche breadth of *D. ordii* set to 0 and 1, respectively for identifiability.

6. SUMMARIES OF THE ORIGINAL ARTICLES AND MAIN RESULTS

Article I: Species abundance dynamics under neutral assumptions: a Bayesian approach to the controversy

In article **I**, a discrete-time version of Hubbell's (2001) "unified neutral theory of biodiversity and biogeography" was developed with the assumption of a constant community size relaxed. The model was fitted with a hierarchical Bayesian approach to long-term macro-moth (Lepidoptera) light-trapping data from the Rothamsted Insect Survey network in the UK. The results suggested that the neutral model would need parameter values that are impossible in order to fit the data well. This is because the community sizes fluctuate more than expected under neutrality. The excess of variability over neutral drift owes presumably to mechanisms such as density-dependent effects, differential responses to environmental fluctuations, and potential imbalances in competitive abilities across species, which are overlooked by the neutral theory. This was one of the first times that the neutral theory was tested on temporal data, and it turned out that demographic stochasticity alone and/or neutrality cannot explain the fluctuations in the study communities.

Article II: On the setting of environmental noise and the performance of population dynamical models

Article **II** investigated whether including environmental noise in one or/and another parameter makes a difference to the behaviour of population dynamical models, with a focus on model fit and predictive performance. To do this, three population dynamical models of the Ricker type were developed, with the noise included in the growth rate (Model 1), in the carrying capacity (Model 2), or in both (Model 3). Several synthetic datasets were generated from each model. The three models were fitted to the simulated data with a Bayesian approach, and their performances in fitting to the data and in forecasting new observations were assessed. The results suggested that the way environmental noise is incorporated into a model may profoundly affect its performance. Overall, Models 1 and 3 broadly outperformed Model 2, the first having the advantage of simplicity and computational tractability. An interesting result emerging from the analyses was the broad positive correlation between mean square errors and deviance information criteria, suggesting that the latter is also informative about the predictive performance of the model.

Article III: What drives community dynamics?

Article **III** tackled one of the central questions in community ecology: what determines patterns of temporal variation in species abundance, and what is the relative importance of environmental forcing *versus* compensatory dynamics? This issue was approached through the combination of multi-species stochastic dynamics modelling with refined statistical methods and high quality data to decompose the community dynamics into the different components namely environmental stochasticity and intra-/inter-specific interactions. The methodology was implemented using data from a range of taxa including moths, fish, crustaceans, birds and

rodents. The results suggested that environmental variation is the most important driver of community dynamics across taxa. Whilst there was always evidence for some density-dependent regulation, the influence of inter-specific interactions on community-level variances remained broadly minor. This implies that community-level variances depend mainly on how species respond to environmental fluctuations. This result is in direct contrast with the view that inter-specific competition reduces community-level variability by promoting negative covariances between species abundances (Tilman & Downing 1994; Lawton & Brown 1993). We thus need to understand species-environment interactions and species differences in their responses to environmental fluctuations if we are to effectively predict their dynamics (Ives *et al.* 1999).

Article IV: A multispecies perspective on ecological impacts of climatic forcing

This article touched upon a timely research area, namely investigating the effects of weather and/or climate fluctuations on population and community dynamics. It used the state-space model formulation in the increasingly popular hierarchical Bayesian framework to extend the scope of previously proposed models of population dynamics under climatic forcing to multi-species systems. The model was designed to explicitly accommodate covariance patterns in species responses to latent environmental factors. The implementation of the model was demonstrated by fitting it to macro-moth (Lepidoptera) community data from the Rothamsted Insect Survey network in the UK, using winter rainfall and winter temperature as environmental covariates. The proportions of environmental variances and covariances attributable to the included climatic covariates in the dynamics of individual species were evaluated.

Article V: Integrating the niche and neutral perspectives on community structure and dynamics

In Article V, a model-based approach was proposed to operationalize the now widely accepted idea that there exists a continuum of community dynamical behaviours between neutrality and niches. The methodology consisted of separating out the role of trait-mediated and stochastic factors, linking the niche configuration to community dynamics through competition, and adding demographic stochasticity. This resulted in a comprehensive framework including neutrality and stringent niches separation as extreme cases. The proposed approach brings mechanisms and patterns closer together, making it possible for the importance of niche-mediated and stochastic factors to be examined in light of the data. A metric for evaluating the position of a community on the niche-neutral continuum was developed as well. The methodology was illustrated with rodent web-trapping data from the Sevilleta National Wildlife Refuge (NWR) in central New Mexico, USA. The coefficients of inter-specific competition estimated from species abundance time series were used to reconstruct the niche configuration.

7. CONCLUSIONS & PERSPECTIVES

This thesis combined the hierarchical Bayesian modelling approach with the computational intensive methods known as Markov chain Monte Carlo, to develop methodologies for identifying and evaluating the workings of community structure and dynamics.

Models of increasing complexity were developed and used to analyze data from real-world systems. The results of a preliminary analysis (**I**) suggested that the failure to acknowledge a range of potentially important mechanisms such as density-dependent regulation, species differences in their ecologies and in their responses to environmental fluctuations makes neutral models inadequate for explaining and predicting the patterns of species abundances and the structure of many real-world systems.

Article **II** was devoted to investigating whether the way environmental noise is incorporated into a population dynamical model makes a difference to the model performance, with a focus on model fit and predictive performance.

In article **III**, a comprehensive framework was developed to evaluate the relative importance of different factors namely, environmental forcing, and inter-/inter-specific interactions as drivers of community dynamics. The model was used to analyze species abundance time series from a variety of taxa: moths (Lepidoptera), fish, crustaceans, birds, and rodents. Environmental forcing was consistently found to be the most important driver of community dynamics across taxa. There was also broad evidence for population regulation through density-dependence, but the contribution of inter-specific interactions to temporal variation in species abundances was broadly weak. The model developed in article **III** was extended in Article **IV** to include environmental covariates and explicitly accommodate sampling error through replicated samples.

A model-based approach to integrating the niche and neutral perspectives on community structure and dynamics was developed in Article **V**, and a proxy for community position on the niche-neutral continuum was proposed.

This thesis contributes to elucidating the mechanisms that drive the structure and dynamics of ecological communities by demonstrating that environmental variation rather than inter-specific competition dominates the dynamics in a diverse range of communities. It is quite amazing to see that broadly similar results would arise for such groups of species as moths, fish, crustaceans, birds, and rodents which are so different from one another, with environmental forcing consistently dominating the dynamics. These findings call into question the alleged importance of negative covariance between species in stabilizing ecological communities (Tilman & Downing 1994; Lawton & Brown 1993), and emphasize the need to better understand how species are affected by the environment.

The scope of the methods presented in this thesis can be extended in different ways. For example, a spatially explicit setting (Hastings 1980, Hastings & Wolin 1989) can be included. This means adding migration into the general framework, which would result in a structured meta-community model. The communities can also be embedded in a larger food web (Ings *et al.* 2009). All these refinements can be achieved under the flexible hierarchical Bayesian framework adopted here.

ACKNOWLEDGEMENTS

I owe my deepest gratitude to my supervisor Bob O'Hara for providing me all the support and guidance I needed to carry out this Ph.D project. He has been a great source of inspiration for my growth as a researcher and a scientist.

I have been extraordinarily fortunate to collaborate with Ian Woiwod. His experience was invaluable in providing me an insight into the world of scientific research.

I would like to thank Otso Ovaskainen and Rampal Etienne for their helpful comments when reviewing this thesis.

I offer my regards to all of those who supported me in any respect during the completion of this Ph.D project; in particular Pasi, Riitta, Raili, Tarja and everyone in the administrative corridor for their willingness to help.

I am grateful to Mikko Sillanpää for extending my research options. Our weekly discussions have been very inspiring.

Many thanks to my friends Michal Wolanski, Lena, Maria, and Imoh for all the times we spent together.

Where would I be without my family? I owe much to my parents for the tremendous sacrifices they made to ensure that I had an excellent education. For this and much more, I am forever grateful. I am indebted to Ya Nelly and Tonton Leon more than they might imagine. I dedicate this thesis to my wife, Youyou Musau, and my kids, Lauren Chizanga and Brandon Munse.

This work was financially supported by the Academy of Finland (project no. 205371 to Bob O'Hara).

REFERENCES

- Abrams, P. (1980) Some comments on measuring niche overlap. *Ecology*, **61**: 44-49.
- Adler, P. B., Leiker, J. & Levine, J. M. (2009) Direct and indirect effects of climate change on a prairie plant community. *PLoS ONE* **4**: e6887.
- Adler, P. B., HilleRisLambers, J. & Levine, J. M. (2007) A niche for neutrality *Ecology Letters* **10**: 95-104.
- Akaike H (1973) Information theory and an extension of the maximum likelihood principle In: Petrov BN, Csaki F. (eds) *Proc. 2nd International symposium on information theory*. Akademiai Kiado, Budapest, 267–281.
- Alonso, D & McKane, A..J. (2004) Sampling Hubbell's neutral theory of biodiversity. *Ecology Letters* **7**: 901-910.
- Anderson, J. L. (1998) Embracing uncertainty: the interface of Bayesian statistics and cognitive psychology. *Conservation Ecology* **2**(1): 2 URL: <http://www.consecol.org/vol2/iss1/art2/>
- Armstrong, R. A. & McGehee, R. (1980) Competitive exclusion. *American Naturalist* **115**: 151-170.
- Begon, M., Mortimer, M. & Thompson, D. J. (1996) *Population ecology*, 3rd edition Blackwell Science.
- Bell, G. (2000) Distribution of Abundance in Neutral Communities. *American Naturalist* **155**: 606-617.
- Berlinier, L. M. (1996) Hierarchical Bayesian time series *Maximum Entropy and Bayesian Methods* (eds K. Hanson & R. Silver) pp. 15-22.
- Bernardo, J. M. & Smith, A. F. M. (1994) *Bayesian Theory*. Wiley, New York.

- Berthier, P., Beaumont, M. A., Cornuet, J.-M. & Luikart, G. (2002) Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: a genealogical approach *Genetics* **160**: 741–751.
- Bjørnstad, O. N., & Grenfell, B. T. (2001) Noisy clockworks: time-series analysis of population fluctuations in animals. *Science* **293**: 638-643.
- Brémaud, P. (2001) *Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer-Verlag.
- Bower, L. P. (1962) Evidence for interspecific competition in natural populations of the monarch and queen butterflies, *Danaus plexippus* and *D. gilippus berenice* in south central Florida. *Ecology* **43**: 549–552.
- Buckland, S. T., Newman, K. B., Thomas, L. and Koesters, N. B. (2004) State-space models for the dynamics of wild animal populations. *Ecological Modelling* **171**: 157-175.
- Bush, G. L. (1969) Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera, Tephritidae). *Evolution*, **23**: 237-251.
- Casella, G. & George, E. (1992) Explaining the Gibbs Sampler. *American Statistician* **46**: 167-174.
- Chase, J. M. (2005) Towards a really unified theory for metacommunities. *Funct. Ecol.* **19**: 182-186.
- Chave J, Muller-Landau, H. C. & Levin, S. A. (2002) Comparing classical community models: theoretical consequences for patterns of diversity. *Am Nat* **159**: 1-23.
- Clark, J. S. (2005) Why environmental scientists are becoming Bayesians. *Ecology Letters* **8**: 2-14.
- Clark, J. S. & Bjørnstad, O. N. (2004) Population inference from messy data: errors, missing and hidden states, and lagged responses. *Ecology* **85**: 3140-3150.
- Colwell, R. K., Futuyma, D. J. (1971) On the measurement of the niche breadth and overlap. *Ecology*, **52**: 567-576.
- de Valpine, P. & Hastings, A. (2002) Fitting population models incorporating process noise and observation error. *Ecological Monographs* **72**: 57-76.
- Dennis, B., Ponciano, J. M., Lele, S. R., Taper, M. L., and Staples, D. F. (2006) Estimating density dependence, process noise and observation error. *Ecology* **76**: 323-341.
- Dornelas, M., Connolly, S. R. & Hughes, T. P. (2006) Coral reef diversity refutes the neutral theory of biodiversity. *Nature* **440**: 80-82.
- Engen, S., Lande, R., Walla, T. & Devries, P. J. (2002) Analyzing Spatial Structure of Communities Using the Two-Dimensional Poisson Lognormal Species Abundance Model. *American Naturalist* **160**: 60-73.
- Enquist, B. J., Sanderson, J. & Weiser, M. D. (2002) Modeling macroscopic patterns in ecology. *Science* **295**: 1835–1837.
- Etienne, R. S. & Olf, H. (2004) A novel genealogical approach to neutral biodiversity theory. *Ecology Letters* **7**: 170-175.
- Fisher, R. A., Corbert, A. S. & Williams, C. B. (1943) The relationship between the number of species and the number of individuals in a random sample of an animal population, Part 2. Results obtained by means of a light-trap at Rothamsted. *J. Anim. Ecol.* **12**: 42-58.
- Friggens, M. (2006) Sevilleta small mammal population data SEV008 <http://sevilleta.unm.edu/data/>.
- Gause G. F. (1935) Experimental demonstration of Volterra's periodic oscillations in the numbers of animals. *J. Exper. Biol.* **12**: 44-48.
- Gause, G. F. (1934) *The Struggle for Existence*. Hafner Publishing Company, New York, New York, USA.
- Gelman, A., Meng, X. L. & Stern, H. (1996) Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica* **6**: 733-760.
- Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. (2003) *Bayesian Data Analysis* (2nd ed.). Chapman & Hall.
- Gelfand, A. E. & Smith, A. F. M. (1990) Sampling-based approaches to calculating marginal densities. *J. Am. Stat. Assoc.* **85**: 398-409.

- George, E. I. & McCulloch, R. E. (1993) Variable selection via Gibbs sampling. *J. Am. Stat. Assoc.* **88**: 881-889.
- Gilks, W. R., Richardson, S. & Spiegelhalter, D. J., Eds., (1996) *Markov Chain Monte Carlo in practice* Chapman and Hall, London, UK.
- Gimenez, O., Rossi, V., Choquet, R., Dehais, C., Doris, B., Varella, H. Vila, J-P. & Pradel, R. (2007) State-space modelling of data on marked individuals. *Ecol. Modelling* **206**: 431-438.
- Gravel, D., Canham, C. D., Beaudet, M., Messier, C. (2006) Reconciling niche and neutrality: the continuum hypothesis. *Ecology Letters* **9**: 399-409.
- Graves, G. R. & Rahbek, C. (2005) Source pool geometry and the assembly of continental avifaunas. *Proc. Natl Acad. Sci. USA* **102**: 7871-7876.
- Hardin, G. (1960) The Competitive Exclusion Principle. *Science* **131**: 1292-1297.
- Hastings A & Wolin, C. L. (1989) Within-patch dynamics in a metapopulation. *Ecology* **70**: 1261-1266.
- Hastings A (1980) Disturbance, coexistence, history, and competition for space. *Theor. Pop. Biol.* **18**: 363-373.
- Hastings, W. K. (1970). "Monte Carlo Sampling Methods Using Markov Chains and Their Applications". *Biometrika* **57**: 97-109.
- Héraud, B. (2007) Reconciling niche and neutrality through the Emergent Group approach *Pers. in Plant Ecol. Evol. & Syst.* **9**: 71-78.
- Holyoak, M. & Loreau, M. (2006) Reconciling empirical ecology with neutral community models *Ecology* **87**: 1370-1377.
- Houlahan, J. E., Currie, D. J., Cottenie, K., Cumming, J. S., Ernest, S. K. M., Findlay, C. S., Fuhlendorf, S. D., Gaedke, U., Legendre, P., Magnuson J. J., McArdle, B. H., Muldavin, E. H., Noble, D. Russel, R., Stevens, R. D., Willis, T. J., Woiwod, I. P., and Wondzell, S. M. (2007) Compensatory dynamics are rare in natural ecological communities. *Proc. Natl. Ac. Sci. USA* **104**: 3273-3277.
- Hubbell, S. P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press.
- Hurlbert, S. H. (1978) The measurement of the niche overlap and some relatives. *Ecology*, **59**: 67-77.
- Hutchinson G. E. (1961) The paradox of the plankton. *Am. Nat.* **95**: 137-145.
- Hutchinson, G. E. (1957) Concluding remarks. Cold Spring Harbor symposium on quantitative biology **22**: 415-425.
- Ings, T. C., Montoya J. M., Bascompte J, Bluthgen N., Brown L., Dormann C. F., Edwards F., Figueroa D., Jacob U., Jones J. I., Lauridsen R. B., Ledger M. E., Lewis H. M., Olesen J. M., van Veen F. J. F., Warren P. H., Woodward, G. (2009) Ecological networks - beyond food webs. *J. Anim. Ecol.* **78**: 253-269.
- Ives, A. R., Gross, K. & Klug, J. L. (1999) Stability and variability in competitive communities. *Science* **286**: 542-544.
- Jacobson, A. R., Provenzale, A., von Hardenberg, A., Bassano, B., and Festa-Bianchet, M. (2004) Climate forcing and density dependence in a mountain ungulate population. *Ecology* **85**: 1598-1610.
- Jeffreys, H. (1961) *Theory of Probability*. Clarendon Press, Oxford.
- Kass, R. E. and Raftery, A. (1995) Bayesian factors. *J. Am. Stat. Assoc.* **90**: 773-795.
- Kimura, M. (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, New York.
- Lande, R., Engen, S. & Saether, B. E. (2003) *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press, New York.
- Lawton, J. H. & Brown, V. K. (1993) Redundancy in ecosystems. In: Schulze, E.-D., Mooney, H. A. (Eds.), *Biodiversity and Ecosystem Function*. Springer, Berlin.
- Levin, S. (1992) The problem of pattern and scale in ecology. *Ecology* **73**: 1943-1967.
- Levins, R. (1968) *Evolution in changing environments: some theoretical explanations*. Princeton University Press, Princeton, N. J.
- Magurran, A. E. (2005) Species abundance distributions: pattern or process? *Funct. Ecol.* **19**: 177-181.

- Maurer, B. A. & McGill, B. J. (2004) Neutral and non-neutral macroecology. *Basic & Appl. Ecol.* **5**: 413–422.
- MacArthur, R. H. & Levins, R. (1967) The limiting similarity, convergence and divergence of coexisting species. *American Naturalist*, **101**: 377–385.
- MacArthur, R. H. & Wilson, E. O. (1967) *The Theory of Island Biogeography*. Monographs in Population Biology, Princeton Univ.
- May, R. M. (1975) Patterns of species abundance and diversity. In M. L. Cody and J. M. Diamond (eds.) *Ecology and Evolution of Communities*. Harvard Univ. Press, Cambridge, MA. pp.81–120.
- McGill, B. J. (2003) A test of the unified neutral theory *Nature* **422**: 881–885.
- Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. (1953) "Equations of State Calculations by Fast Computing Machines". *Journal of Chemical Physics* **21**: 1087–1092.
- Miyashita, T. (2001) Competition for a limited space in kleptoparasitic *Argyrodes* spiders revealed by field experiments. *Popul. Ecol.* **43**: 97–103.
- Murdock, W. W. (1994) Population regulation in theory and practice. *Ecology* **75**: 271–287.
- Motwani, R. & Raghavan, P. (1995) *Randomized Algorithms*. Cambridge University Press.
- Pianka, E. R. (1974) Niche overlap and diffuse competition. *Proc. Natl. Acad. Sci. USA* **71**: 2141–2145.
- Plummer M. (2008) Penalized loss functions for Bayesian model comparison. *Biostatistics*, **9**: 523–539.
- Poulin, R. (2004) Parasites and the neutral theory of biodiversity. *Ecography* **27**: 119–123.
- O'Hara, R. B., Arjas, E., Toivonen, H. & Hanski, I. (2002) Bayesian analysis of metapopulation data. *Ecology* **83**: 2408–2415.
- Ricklefs, R. E. & Schluter, D. (eds) (1993) *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*. University of Chicago Press, Chicago.
- Rivot, E., Prevost, E., Parent, E., & Bagliniere, J. L. (2004) A Bayesian state-space modelling framework for fitting a salmon stage-structured population dynamic model to multiple time series of field data. *Ecological Modelling* **179**: 463–485.
- Robert, C. P. & Casella, G. (2004) *Monte Carlo Statistical Methods* 2nd Edn. Springer-Verlag, New York.
- Robert, C. P. (2001) *The Bayesian Choice: from Decision-Theoretic Motivations to Computational Implementations*. Springer-Verlag, New York.
- Ross, S. (2007) *Introduction to Probability Models* 9th ed. Academic Press.
- Royle, J. A. & Dorazio, R. M. (2006) Hierarchical models of animal abundance and occurrence. *J. Agr. Biol. & Env. Stat.* **11**: 249–263.
- Saether, B-E., Tufto, J., Engen, S., Jerstad, K., Røstad O. W. & Skåtán, J. E. (2000) Population dynamical consequences of climate change for a small temperate songbird. *Science* **287**: 854–856.
- Schwarz, G. E. (1978) Estimating the dimension of a model. *Annals of Statistics*, **6**: 461–464.
- Shorrocks, B. & Sevenster, J. G. (1995) Explaining local species diversity. *Proc. R. Soc. Lond. B* **260**: 305–309.
- Simberloff, D. (1982) The status of competition theory in ecology. *Ann. Zool. Fennici* **19**: 241–253.
- Spiegelhalter, D., Thomas, A., Best, N. and Lunn, D. (2003) WinBugs version 1.4 User manual <http://www.mrc-bsu.cam.ac.uk/bug>.
- Spiegelhalter D. J., Best N. G., Carlin P., van Linder A. (2002) Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B* **64**: 583–639.
- Taylor, L. R. & French, R. A. (1974) Effects of light-trap design and illumination on samples of moths in an English woodland. *Bull. Entomol. Res.* **63**: 583–594.
- Thomas, A., O'Hara, R. B., Ligges, U. & Sturtz, S. (2006) Making BUGS Open. *R News* **6**: 12–17.
- Tierney, L. (1994) Markov Chains for exploring posterior distributions (with discussion) *The Annals of Statistics* **22**: 1701–1702.
- Tilman, D. & Downing, J. A. (1994). Biodiversity and stability in grasslands. *Nature* **367**: 363–365.
- Tokeshi, M. (1990). Niche apportionment or random assortment: species abundance patterns revisited. *J. Anim. Ecol.* **59**: 1129–1146.
- Turnbull, L. A., Manley, L. and Rees, M. (2005) Niches, rather than neutrality, structure a grassland pioneer guild. *Proc. Roy. Soc. Lond. B* **272**: 1357–1364.

- Wilson, W. G. & Lundberg, P. (2006) Non-neutral dynamics: empirical predictions for ecosystem function and diversity from linearized consumer–resource interactions. *Oikos* **114**: 71–83.
- Wikle, C. K. & Hooten, M. B. (2006) Hierarchical Bayesian spatio-temporal models for population spread *Applications of Computational Statistics in the Environmental Sciences : Hierarchical Bayesian and MCMC Methods* (eds J. S. Clark & A. Gelfand) pp. 145-169 Oxford University Press, Oxford.
- Wikle, C. K. (2003) Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology* **84**: 1382-1394.
- Williamson, M. & Gaston, K. J. (2005) The lognormal is not an appropriate null hypothesis for the species- abundance distribution. *J. Anim. Ecol.* **74**: 409-422.
- Woiwod, I. P. & Gould, P. J. L. (2008) Long-term moth studies at Rothamsted in *The Moths of Hertfordshire* eds. C. W. Plant (Hertfordshire National History Society) pp 31-44.
- Woiwod, I. P. & Hanski, I. (1992) Patterns of density dependence in moths and aphids. *J. Anim. Ecol.* **61**: 619–629.
- Woiwod, I. P. & Harrington, R. (1994) Flying in the face of change: the Rothamsted Insect Survey *Long-term Experiments in Agricultural and Ecological Sciences* (eds R.A. Leigh & A.E. Johnston), pp. 321–342. CAB International, Wallingford.
- Zhou, S. & Zhang, D. (2006) Neutral theory in community ecology. *Front. Biol. China* **30**: 1-8.