

Groups and Niches in Food-Web Structure: Bayesian Methods and Consequences for Robustness

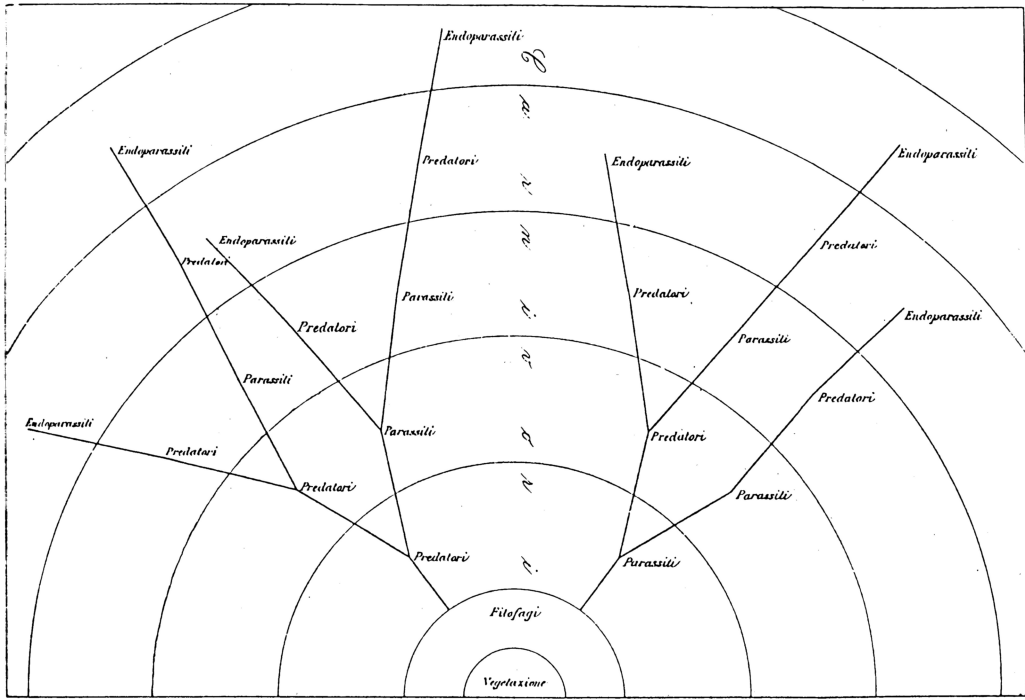
by

Edward Barnet Baskerville

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Ecology and Evolutionary Biology and Scientific Computing)
in the University of Michigan
2013

Doctoral Committee:

Professor Mercedes Pascual, Chair
Associate Professor Aaron A. King
Research Professor Rick L. Riolo
Professor John H. Vandermeer



The earliest known food web network diagram, Lorenzo Camerano, 1880.

Copyright © 2013 Edward Barnet Baskerville.

Except where otherwise noted, this work is provided under a
Creative Commons Attribution-ShareAlike Unported license.

<http://creativecommons.org/licenses/by-sa/3.0/>



Acknowledgments

My long and meandering path to a degree in ecology has been inspired and aided by many teachers, mentors, colleagues, and friends. First and foremost, I want to thank my committee: Aaron King, Rick Riolo, John Vandermeer, and Mercedes Pascual, my advisor. Aaron's thinking is simultaneously careful and expansive, mathematically abstract and biologically intuitive, and I have learned something new from every conversation we have had. Rick's course on complex systems more than a decade ago led me into ecology in the first place, and much of how I think owes to his early and ongoing mentorship; he is a treasure for the complex systems community. Having first witnessed John's passionate devotion to coffee agroecosystems in a talk many years ago, as a graduate student I have learned much from his knowledge of the interactions of actual organisms as well as his appreciation for simple, elegant models of complex phenomena.

At my first meeting with Mercedes some ten years ago, I was amazed by her brilliant mind and her boundless energy and enthusiasm for ideas, and I continue to be amazed. She manages to constantly generate flurries of fascinating thoughts, hold onto them, and transform them into inspired research. Without her guidance my love for ecology would never have grown to what it is, and I could not have asked for a better advisor. To her I give the greatest thanks.

Chapters 3 and 4 result largely from a collaboration with Andy Dobson, who, equipped with deep knowledge of both natural history and ecological theory, has taught me everything I know about the Serengeti. The whole dissertation owes much to the work of Stefano Allesina, who has been a wonderful mentor about food webs and beyond. The use of Bayesian approaches came directly from numerous conversations with Trevor Bedford, stemming from his expertise in phy-

logenetic inference. Mike Anderson provided valuable insight, interpretation, and spatial data on Serengeti plant communities. I also thank Brian Sedio for an interesting side collaboration on plant-insect interactions.

The Department of Energy Computational Science Graduate Fellowship and the EEB department provided core financial support, not to mention two wonderful communities of people. Kind and patient administrative support has come from many people at CSGF, EEB, and Scientific Computing, especially Jeana Gingery, Jane Sullivan, Sonja Botes, Gail Kuhnlein, and Pam Derry.

I owe a great debt to many teachers along the way. Long before graduate school, in Kent Daniels' fourth and fifth grade classes at Madrona Elementary School I first encountered computer programming, food webs, and environmental studies through weeklong trips down the Columbia River and around the Olympic Peninsula. Irene Barinoff introduced my high school calculus class to the Lotka-Volterra equations and to fractal geometry, and Jabe Blumenthal provided a taste of probability theory and differential equations; Stephen Bullock and Charles Doering reanimated those topics for me in later years in the two best math classes I took at Michigan. John Jamison has been greatest mentor about all things: relativity, history, music, education, complexity, life. Roger Nisbet and Leeza Pachepsky helped guide me toward ecology in my pre-Ph.D. years of wandering. Additionally, many people have made the other side of teaching both a joy and the perfect intellectual adjunct to research: Katie Wharton and the students of Biology 171, Fall 2008; and Deborah Goldberg, Aaron King, and the students of EEB 481, Fall 2012.

The Pascual lab has provided a wonderful intellectual environment thanks to a friendly, smart cast of characters over the years: David Alonso, Katia Koelle, Manojit Roy, Sarah Cobey, Diego Ruiz Moreno, Yael Artzy, Karina Laneri, Richard Zinck, Bobby Reiner, Molly Rorick, Tony Golubski, Pamela Martinez, Bob Woods, Andrew Strayer, Mauricio Santos, Daniel Zinder, Alex Livne, Stefano, Trevor, and Andres Baeza, my fellow traveler from start to finish. The rest of the extended EEB and CSCS communities, including CSAAW, home to many fascinating discussions of complexity, and BioBev, where EEB ends the week, have made academic life a pleasure. I also thank

many other Ann Arbor friends: Silvio Medoro, who provides a pizza-oriented home for ecologists and musicians alike; all the musicians; the poker night crew; Mark Haynes, wave expert and fellow cello; and Beth Percha, physics buddy, cookie-baking housemate, and brilliant thinker.

I would like to thank my family for their endless love and their infectious curiosity about the world: Gramma, Dad, Mom, and Mia, who are always too far away, and Lizzy, my fellow chronic Michigan transplant. Lastly, I thank Robin, who I am lucky to have in my life.

Table of Contents

Acknowledgments	ii
List of Figures	vii
List of Tables	viii
Chapter 1 Introduction	1
1.1 The food web: a network view into ecosystem complexity	1
1.2 Patterns in food-web structure: models and methods	2
1.3 Food-web structure, stability, and robustness	5
1.4 Overview of the dissertation	7
Chapter 2 Bayesian Inference for Food-Web Models	12
2.1 Probabilistic models for food webs	13
2.2 Bayesian inference	14
2.3 The Bayesian group model	15
2.4 Other food-web models	23
2.5 Markov-chain Monte Carlo sampling	25
2.6 Model selection via marginal likelihood	29
2.7 The mc3kit software library	33
2.8 Lessons learned	35
Chapter 3 Spatial Guilds in the Serengeti Food Web	42
3.1 Introduction	42
3.2 Results	45
3.3 Discussion	57
3.4 Methods	62
Chapter 4 Groups and Niches in Food-Web Structure	74
4.1 Introduction	74
4.2 Methods	77
4.3 Results	84
4.4 Discussion	90

Chapter 5 Food-Web Robustness and the Dual Nature of Compartments	99
5.1 Introduction	99
5.2 Methods	103
5.3 Results	110
5.4 Discussion	114
Chapter 6 Conclusion	123
6.1 Models of food-web structure: where to go from here	123
6.2 Compartments, stability, robustness, and spatial coupling	126
6.3 Future directions for the food web	127
6.4 Theory is dead, long live theory!	130

List of Figures

Figure 3.1	Posterior distributions of aggregation and group-count parameters	47
Figure 3.2	Posterior distributions of link density parameters	48
Figure 3.3	Distribution of link probability parameters	49
Figure 3.4	Affinity matrix	50
Figure 3.5	The Serengeti food web	53
Figure 3.6	Adjacency matrix ordered by groups	54
Figure 3.7	Network layout of groups	55
Figure 4.1	Group counts	85
Figure 4.2	Plant groups.	86
Figure 4.3	Herbivore groups.	87
Figure 4.4	Adjacency matrix by groups	88
Figure 4.5	Niche space	89
Figure 5.1	Group structures for robustness experiment	107
Figure 5.2	Error against model fit	111
Figure 5.3	Predicted vs. empirical extinction area	112
Figure 5.4	Extinction area by group structure	113
Figure 5.5	Extinction area vs. groupiness	115
Figure 5.6	Effect of intervality on robustness	116

List of Tables

Table 3.1	Marginal likelihood estimates and Bayes factors	46
Table 3.2	Groups identified in the Serengeti food web	51
Table 3.3	Link densities between groups	56
Table 4.1	Models	78
Table 4.2	Model-selection results	85
Table 5.1	Food webs for robustness comparison	105
Table 5.2	Models for robustness comparison	106

Chapter 1

Introduction

1.1 The food web: a network view into ecosystem complexity

In a prescient 1880 paper, the Italian zoologist and former art student Lorenzo Camerano likened the transmission of fluctuations in biological populations to the propagation of sound waves [5, 6]. This paper, apparently inspired by the pioneering work on sound and music by Heinrich von Helmholtz, included the first known diagram of a food web network, consisting of feeding links originating with plants at the center and branching out to herbivores, predators, parasites, predators of parasites, and on in a series of seven concentric circles. Camerano described a *laissez-faire* vision of nature in which a natural equilibrium among species emerges out of the Darwinian struggle for existence. If one species unexpectedly increased in abundance, the system would serve to restore its abundance to the equilibrium state. Evolution acts as Antonio Stradivari building an ecological violin, and when bow is put to string, it sings rather than snaps. Though Camerano's vision encompassed all of nature, it would be another forty years before Lotka and Volterra invented the mathematics to describe a single predator-prey pair [15, 33]. Today ecologists are still struggling to untangle the networks that Camerano saw as essentially harmonious, but which we increasingly know as sometimes robust, sometimes fragile, and certainly more complex than a violin.

Later on, Charles Elton helped establish the view of ecosystems as networks of feeding links, which he called food cycles, as foundational for modern ecology [9]. Camerano described the

natural world in terms of the population fluctuations of individual species, and this perspective, along with the population-dynamic legacy of Lotka and Volterra, have led to a long discourse on how these fluctuations interact in complex networks of interconnections [16, 10, 23, 18, 25]. It also naturally leads to the compilation of increasingly large lists of species and their feeding interactions, which lend themselves to analysis using mathematical tools. A parallel school of thought developed in American limnology from the seminal and tragically brief work of Raymond Lindeman, who described an integrated view of ecosystems as the flow and dissipation of energy through a food web [14]. The approaches that developed tend to emphasize the organization the flow of energy and nutrients in ecosystems at a large scale of organization [21], at the expense the representation of individual species. The focus of this dissertation is on the structure of networks of feeding links among species, with population dynamics and energy and nutrient flows assumed implicitly but not represented explicitly, with the goal of better establishing a basis for the more difficult problem of how structure and dynamics interact and emerge from each other.

The sections in this introduction describe approaches for looking at patterns in food-web structure (1.2); survey how structural patterns have been thought to influence aspects of stability and robustness in food webs (1.3); and introduce the investigations pursued for this dissertation (1.4).

1.2 Patterns in food-web structure: models and methods

The complexity of food webs provides many important questions for ecological science. How do food webs form, persist, collapse, reassemble? How do web-scale processes interact with the evolutionary trajectories of a food web's constituent species? One important step toward understanding these difficult questions is to identify the patterns present in natural food webs, and thus provide an empirical basis for these investigations. The emergence of food-web data sets over time [7, 27, 22, 12] has prompted the development of analysis and modeling approaches to characterize the structure of food webs.

Over the last several decades, the emphasis of studies on empirical food-web networks has shifted

from a search for general “laws” of food-web structure to approaches based on probabilistic models that can be fit to data. In the 1980s, Cohen and colleagues studied a collection of 113 empirical food webs and inferred the existence of a number of interesting patterns, including a low degree of omnivory, a linear scaling between the number of species and the number of links, and the presence of high degree of “intervality” in empirical food webs, meaning that species can be ordered in such a way that the diets of each species is largely contiguous within the ordering; that is, there is a one-dimensional “niche space” [7]. With the notable exception of the one-dimensional niche space, many of the identified patterns were found to result from poor sampling of data or other problems; however, in highly resolved data sets, the pattern that the number of links scales with the square of the number of species is plausible [24].

The use of probabilistic models for food-web structure also began with same research group, and this approach has evolved into a flexible, general way to model patterns in food webs and quantitatively test hypotheses. In their original model, the cascade model, species are imagined to be ordered, and the diet of the first species is restricted to be empty (thus representing a primary producer); the diet of the second species can only contain the first species; the diet of the second species can only contain the first two; and so on, so that the adjacency matrix of the food web is upper triangular, with each possible link being equally likely. Although the possible diet of a species under this model is nested within the possible diet of a species higher in the ordering, this model does not impose the contiguous diets they observed in real webs. The niche model of Williams and Martinez realized that requirement, along with the tendency for higher-order species to feed more widely, while relaxing the constraint that possible diets be perfectly nested [34]. Food webs generated using the niche model displayed surprising correspondence to empirical webs when measured in terms of properties such as trophic similarity, degree of omnivory, and the variability of diet with among species, improving substantially on the cascade model.

The cascade and niche models were originally used to generate food webs and compare their properties with those of empirically observed webs. In fact, the niche model cannot exactly repro-

duce the link pattern of any observed webs, because empirical data sets inevitably include some “gaps” in the nearly-contiguous diets of constituent species. The work of Allesina, Alonso, and Pascual developed the minimum potential niche model, allowing gaps in diets, and showed that it was possible to fit this model to data [1]. The likelihood-based approach also enabled the use of model-selection techniques to compare across different food-web models [2].

A parallel conversation about food-web structure has asked whether food webs are compartmentalized [26], inspired largely from the theoretical idea that compartmentalization should stabilize population dynamics [18]. Until recently, the compartmentalization hypothesis had not been tested explicitly using probabilistic models. Instead, measures of network modularity have been used to indicate that food webs are significantly compartmentalized with respect to certain random-graph models [13, 28], an approach that has recently come under strong criticism [32].

Another approach to subdividing food webs into groups is realized by the group model, under which species in the same group have identical probabilities of connecting to species in other groups [2]. The model thus tends to identify trophic guilds of species, which have statistically similar (but not identical) sets of predators and prey. When fit to empirical webs, the group model performs quite well compared with the minimum potential niche model in large food webs, indicating that this model does a reasonably good job at summarizing the overall shape of a food web using sets of trophically similar species. However, improved niche models have been shown to outperform the group model [35, 32].

The last several years of fitting probabilistic models to food webs has produced a more nuanced view of food-web structure. The one dominant pattern that remains standing from early analysis is that the contiguity of diets in a low-dimensional niche space is the main recognizable feature in food-web structure, although in larger food webs two-dimensional orderings give a better model fit than one dimension alone [35]. The precise origin of this diet contiguity is still unclear. Body size has long been hypothesized as the dominant axis of niche space, but models based on body size alone perform poorly [35, 29]. Recent work based on evolutionary models indicates that low-

dimensional diet contiguity can in fact arise from a very high-dimensional trait space controlling feeding relationships, a counterintuitive result that may resolve the contradiction that body size is a poor predictor of food-web structure [30].

1.3 Food-web structure, stability, and robustness

Elton and other thinkers of the time gravitated toward the idea that the complexity of nature provides a balance, that an abundance of species, and an abundance of interconnections among species provides a mechanism to offset the fluctuations that arise [10]. MacArthur quantified this notion by proposing that the “stability” of a food web could be measured as the freedom of energy flow up a food web, measured using Shannon entropy [16, 31]. But Elton’s argument and MacArthur’s definition both depended upon the assumption that many links and many species do, in fact, generate stability, an idea that was beautiful but untested.

Elton cites a multitude of empirical observations that simple ecological systems are far more susceptible to dynamic instability than complex ones. Tropical rainforests never seem to suffer from insect outbreaks the way temperate forests do, and simple agricultural systems are more susceptible than the complex natural systems they replace. The more complex the ecosystem, he surmised, the more stable. Elton acknowledges the likely incompleteness of this analysis, but appeals to a plausible, intuitive explanation: if every species has plenty of natural enemies, no species will be able to escape top-down control.

The assumption that complexity alone could account for stability remained unchallenged until Robert May first brought a dynamical systems approach to the issue [17]. He found that, in randomly constructed community matrices, more interactions between species lead to a lower chance of stability, and that this effect is more pronounced in webs containing more species. The implication of May’s work is not that nature is miraculously filled with large, unstable systems, but that complexity alone, in the sense of many species and many links between species, cannot explain the persistence and stability of the large systems that do exist. Thus May’s work focused ecologists on

the question of how large, complex systems that display some level of persistence and resilience can, in fact, emerge.

The question of how food web structure and dynamics interact has driven a number of theoretical studies, most focused on the effect of imposed structural constraints on different stability measures. Empirical analysis show that food webs contain a small number of strong links embedded in a network of primarily weak links, and this pattern has been shown to influence stability in some theoretical models [19]. [20] show that “long loops” of interactions in empirical soil webs tend to be dominated by weak links, a pattern hypothesized to offset the potential for loops to lead to instability. Allometric scaling of bioenergetic flow has also been shown to stabilize food web dynamics [4]. However, regardless of interaction strength, the predominance of predator-prey (+/-) interactions alone produces structures much more stable than the ones examined by May, which allowed for -/- and +/+ responses around equilibrium: although increasing the number of links and species does reduce stability, it does so much more slowly [3]. As a result, the idea that a large food webs can be stable now seems much less mysterious.

That said, a better understanding of how food webs will respond to perturbations of various kinds remains an important problem, particularly in an age of increasing extinctions caused by humans. Furthermore, when it comes to extinctions, understanding the equilibrium state prior to the initial extinction is not enough; we must also understand how the removal of species will cascade through the system [8]. Extinction cascades are mediated both through the structure of food webs in a very direct way, since if a species loses its prey it will go extinct, as well as in a more complex way due via the dynamical responses of constituent populations that may cause extinctions not necessarily implied by the structure alone.

As the topology of food web networks mediates these extinction cascades, a better understanding of that topology should help improve our understanding of those cascades. For example, a conventional wisdom has emerged that compartmentalization is an important force stabilizing food webs, because theoretical investigations have shown that compartmentalization can stabilize dynamical

models, and compartmentalization has been detected as “significant” compared to random-graph models. However, upon further inspection, the model implied by these methods is a very poor description of food-web structure. Still, because the presence of niche or group structure also implies a kind of “compartmentalization” in the food web [11], the general idea is not rendered invalid, but it must be reexamined using better descriptions of structure.

1.4 Overview of the dissertation

This dissertation includes three main ecological investigations, supported by one main development in methodology.

Chapter 2 covers the use of the Bayesian statistical framework for probabilistic food-web models. The Bayesian framework, which has become increasingly popular for statistics in general and for ecology in particular, has a main advantage of including uncertainty in model parameters as part of the model formulation. It also allows for flexible hierarchical structures in models, a feature vital to the group-niche model of Chapter 4. The Bayesian framework also enables model selection to be more easily framed in straightforward probabilistic terms. However, complex Bayesian models also require the use of computationally intensive Markov-chain Monte Carlo methods. In order to perform inference on the models developed for this dissertation, I developed a new software package, `mc3kit`, also described in the chapter.

Chapter 3 is a study of a plant-mammal food web from the Serengeti ecosystem of Tanzania. The analysis is based on a Bayesian version of the group model, and represents the first use of the full Bayesian approach for probabilistic food-web models.

Chapter 4 extends the analysis of the Serengeti web with a probabilistic food-web model that combines group structure—the identification of trophically similar species—with the explicit representation of a niche space, made possible through the use of a Bayesian hierarchical formulation. This group-niche model thus allows the simultaneous inference of similarities among groups of species whose diets are constrained by low-dimensional niche space.

Chapter 5 employs improvements in modeling food-web structure to explore the structural robustness of food webs, first by showing, unsurprisingly, that better models of structure better capture the robustness of food webs, and second by looking at how the allocation of links in group structures—in other words, compartmentalization—affects the susceptibility of food webs to bottom-up extinction cascades.

Chapter 6 sums up, and discusses future directions for studies of food-web structure, robustness, and stability.

References

- [1] S. Allesina, D. Alonso, and M. Pascual. “A general model for food web structure”. In: *Science* 320.5876 (2008), pp. 658–661. DOI: 10.1126/science.1156269.
- [2] S. Allesina and M. Pascual. “Food web models: a plea for groups”. In: *Ecology Letters* 12.7 (2009), pp. 652–662. DOI: 10.1111/j.1461-0248.2009.01321.x.
- [3] S. Allesina and S. Tang. “Stability criteria for complex ecosystems”. In: *Nature* 483.7388 (2012), pp. 205–208. DOI: 10.1038/nature10832.
- [4] U. Brose, R. J. Williams, and N. D. Martinez. “Allometric scaling enhances stability in complex food webs”. In: *Ecology Letters* 9.11 (2006), pp. 1228–1236. DOI: 10.1111/j.1461-0248.2006.00978.x.
- [5] L. Camerano. “Dell’equilibrio dei viventi merce la reciproca distruzione”. In: *Accademia delle Scienze di Torino* 15 (1880), pp. 393–414.
- [6] J. E. Cohen. “Lorenzo Camerano’s contribution to early food web theory”. In: *Frontiers in mathematical biology*. Springer, 1994, pp. 351–359.
- [7] J. E. Cohen, F. Briand, and C. M. Newman. *Community Food Webs: Data and Theory*. Biomathematics. Springer, 1990. ISBN: 0387511296.

- [8] J. A. Dunne, R. J. Williams, and N. D. Martinez. “Network structure and biodiversity loss in food webs: robustness increases with connectance”. In: *Ecology Letters* 5.4 (2002), pp. 558–567. DOI: 10.1046/j.1461-0248.2002.00354.x.
- [9] C. S. Elton. *Animal Ecology*. Text-Books of Animal Biology. London: Sidgwick & Jackson, Ltd., 1927.
- [10] C. S. Elton. *The Ecology of Invasions by Animals and Plants*. 1958.
- [11] R. Guimera, D. Stouffer, M. Sales-Pardo, E. Leicht, M. Newman, and L. Amaral. “Origin of compartmentalization in food webs”. In: *Ecology* 91.10 (2010), pp. 2941–2951. DOI: 10.1890/09-1175.1.
- [12] U. Jacob, A. Thierry, U. Brose, W. E. Arntz, S. Berg, T. Brey, I. Fetzer, T. Jonsson, K. Mintenbeck, and C. Mollmann. “The Role of Body Size in Complex Food Webs: A Cold Case”. In: *The Role of Body Size in Multispecies Systems*. Vol. 45. Advances In Ecological Research. Amsterdam: Elsevier, 2011, pp. 181–223.
- [13] A. E. Krause, K. A. Frank, D. M. Mason, R. E. Ulanowicz, and W. W. Taylor. “Compartments revealed in food-web structure”. In: *Nature* 426.6964 (2003), pp. 282–285. DOI: 10.1038/nature02115.
- [14] R. L. Lindeman. “The trophic-dynamic aspect of ecology”. In: *Ecology* 23.4 (1942), pp. 399–417. DOI: 10.2307/1930126.
- [15] A. J. Lotka. *Elements of physical biology*. Williams & Wilkins Baltimore, 1925.
- [16] R. MacArthur. “Fluctuations of animal populations and a measure of community stability”. In: *Ecology* 36.3 (1955), pp. 533–536. DOI: 10.2307/1929601.
- [17] R. M. May. *Stability and Complexity in Model Ecosystems*. Princeton University Press, 1973.
- [18] R. M. May. “Will a large complex system be stable?” In: *Nature* 238 (1972), pp. 413–414. DOI: 10.1038/238413a0.

- [19] K. McCann, A. Hastings, and G. R. Huxel. “Weak trophic interactions and the balance of nature”. In: *Nature* 395.6704 (1998), pp. 794–798. DOI: 10.1038/27427.
- [20] A.-M. Neutel, J. A. Heesterbeek, and P. C. de Ruiter. “Stability in real food webs: weak links in long loops”. In: *Science* 296.5570 (2002), pp. 1120–1123. DOI: 10.1126/science.1068326.
- [21] H. T. Odum. “Systems Ecology; an introduction”. In: (1983).
- [22] S. C. Opitz. *Trophic interactions in Caribbean coral reefs*. 1085. The WorldFish Center, 1996.
- [23] R. T. Paine. “A note on trophic complexity and community stability”. In: *The American Naturalist* 103.929 (1969), pp. 91–93. DOI: 10.2307/2459472.
- [24] M. Pascual and J. A. Dunne. *Ecological networks: linking structure to dynamics in food webs*. Oxford University Press, 2005.
- [25] S. L. Pimm. “The complexity and stability of ecosystems”. In: *Nature* 307.5949 (1984), pp. 321–326. DOI: 10.1038/307321a0.
- [26] S. L. Pimm and J. H. Lawton. “Are food webs divided into compartments?” In: *Journal of Animal Ecology* (1980), pp. 879–898. DOI: 10.2307/4233.
- [27] G. A. Polis and D. R. Strong. “Food web complexity and community dynamics”. In: *American Naturalist* (1996), pp. 813–846. DOI: 10.2307/2463091.
- [28] E. L. Rezende, E. M. Albert, M. A. Fortuna, and J. Bascompte. “Compartments in a marine food web associated with phylogeny, body mass, and habitat structure”. In: *Ecology Letters* 12.8 (2009), pp. 779–788. DOI: 10.1111/j.1461-0248.2009.01327.x.
- [29] R. P. Rohr, H. Scherer, P. Kehrl, C. Mazza, and L.-F. Bersier. “Modeling food webs: exploring unexplained structure using latent traits”. In: *The American Naturalist* 176.2 (2010), pp. 170–177. DOI: 10.1086/653667.

- [30] A. G. Rossberg, Å. Brännström, and U. Dieckmann. “Food-web structure in low-and high-dimensional trophic niche spaces”. In: *Journal of The Royal Society Interface* 7.53 (2010), pp. 1735–1743. DOI: 10.1098/rsif.2010.0111.
- [31] C. E. Shannon. “A mathematical theory of communication”. In: (1948).
- [32] P. P. Stanciczenko, M. J. Smith, and S. Allesina. “Selecting food web models: the good, the bad, and the ugly”. In: (*in prep.*) (2013).
- [33] V. Volterra. “Variations and fluctuations of the number of individuals in animal species living together”. In: *J. Cons. Int. Explor. Mer* 3.1 (1928), pp. 3–51.
- [34] R. J. Williams and N. D. Martinez. “Simple rules yield complex food webs”. In: *Nature* 404.6774 (2000), pp. 180–183. DOI: 10.1038/35004572.
- [35] R. J. Williams and D. W. Purves. “The probabilistic niche model reveals substantial variation in the niche structure of empirical food webs”. In: *Ecology* 92.9 (2011), pp. 1849–1857. DOI: 10.1890/11-0200.1.

Chapter 2

Bayesian Inference for Food-Web Models

The purpose of this chapter is to introduce Bayesian and empirical Bayes approaches for modeling food webs as a methodological foundation for the models used in Chapters 3 and 4, in a mathematically informal and mildly didactic treatment. A working knowledge of probability theory is assumed. The general Bayesian material and the description of the group model is adapted from the supplement for the paper that comprises Chapter 3 [4].

Section 2.1 introduces the use of probabilistic generative models to describe food-web structure, with a focus on the group model [2]. Section 2.2 introduces Bayesian inference in general, and Section 2.3 applies the Bayesian framework to the group model. A brief discussion of other kinds of models, including models that predict structure from trait data, makes up Section 2.4. Section 2.5 describes Markov-chain Monte Carlo (MCMC) methods, and Section 2.6 covers model selection via marginal likelihood and maximum marginal likelihood. A large side-effect of this work on food-web models was the development of a general-purpose software library for MCMC inference of statistical models, `mc3kit`, which is briefly described in Section 2.7. Finally, Section 2.8 provides some reflections on the challenges of using Bayesian and empirical Bayes methods and Markov-chain Monte Carlo.

2.1 Probabilistic models for food webs

Probabilistic modeling serves a tool for formalizing hypotheses about food-web structure and uncovering meaningful regularities in food-web data. For most of this chapter, a food web is assumed to be an observed list of who eats whom in an ecosystem, interpreted as data in the form of a binary matrix. We start with the basic question: assuming a particular model of food web structure, what is the probability of observing the data? This probability is referred to as the *likelihood* of observing the data, given model parameters. In a maximum-likelihood framework, the mechanical part of the inference process is to find the set of model parameters that makes the likelihood as great as possible, with the interpretation that this represents the best point estimate of the underlying process.

We begin with the group model [2], also known as a stochastic blockmodel [20], which in the context of food webs was originally treated in a maximum-likelihood framework. Conceptually, this model encodes the simple hypothesis that species can be divided into groups, and species in the same group have statistically similar behavior: they tend to consume species in certain groups and tend to be consumed by species in certain groups. Specifically, the probability that a species belonging to group i is eaten by a species belonging to group j is given by p_{ij} , and conversely, the probability of a link being absent is $(1 - p_{ij})$. If there are K groups, then a matrix \mathbf{P} of K^2 link probabilities is required to completely describe the relationships among all groups. The likelihood for the whole network is the product over all pairs of species of the probability of a link being present (if present) or absent (if absent). The assignment of species to groups is also an unobserved parameter in this model, which adds a layer of difficulty to parameter estimation. For example, in a network of 100 species, there are approximately 5×10^{115} different ways to partition the network into groups (see Methods). That is, if you had a computer that could process 10^{80} partitions (as many partitions as there are atoms in the universe) every femtosecond (10^{-15} s), it would take 1.5×10^{13} years to process them all, or about one thousand times the estimated time since the big bang.

The group model allows for a more flexible definition of groups than the usual approaches to

network clustering, which find groups that have large numbers of internal connections and relatively few connections between groups [15]. Because each p_{ij} parameter may take any value between 0 and 1, good model fits may result from other relationships, such as high link density between groups and low link density within groups, and may accommodate different relationships in different parts of the network. In general, the best-fitting partitions will try to maximize or minimize the number of links within specific groups and between specific pairs of groups.

2.2 Bayesian inference

In a Bayesian framework, rules of probability are taken to govern both the data and model parameters. Rather than finding the set of parameter values that maximize the likelihood, the goal becomes to estimate a probability distribution over parameters based on observed data. In this way, we can directly quantify the uncertainty in our parameters in terms of probabilities. This permits questions such as: given the model and the data, what is the probability that a parameter lies in a particular range? The name “Bayesian” comes from Bayes’ rule [5], a basic result in conditional probabilities. Bayes’ rule lets us calculate a *posterior distribution*, $\Pr(\theta|D)$, over model parameters θ , representing our (uncertain) knowledge of parameters once we have observed the data D . This calculation requires a *prior distribution* $\Pr(\theta)$, which represents assumptions prior to observing the data, possibly informed by previous results. If we are dealing entirely with discrete probability distributions, Bayes’ rule takes its most intuitive form:

$$\Pr(\theta|D) = \frac{\Pr(\theta)\Pr(D|\theta)}{\Pr(D)}. \quad (2.1)$$

The numerator of the right-hand side is the probability of producing the data from the given parameters: the *prior* probability of those parameters, $\Pr(\theta)$, times the probability of producing the data given those parameters, $\Pr(D|\theta)$, also known as the likelihood. The denominator is the *marginal*

probability of observing the data unconditional on the particular parameter values at play, which is simply the sum of the probabilities of all the different ways of producing the data using all possible parameter values, $\Pr(D) = \sum_{\theta} \Pr(\theta)\Pr(D|\theta)$. In other words, in order to calculate the posterior probability of parameters θ , we add up all the different ways of producing the data weighted by their probability, and then calculate what fraction of that probability came from parameters θ . From here, we will write these quantities in more general notation, suitable for a mix of discrete and continuous probability distributions:

$$f(\theta|D) = \frac{f(\theta)f(D|\theta)}{\int_{\theta} f(\theta)f(D|\theta) d\theta}, \quad (2.2)$$

where the integral sign represents a multiple integral over discrete and continuous parameters.

The specification of a prior can be both a useful tool and a point of arbitrariness, and, for the purpose of this dissertation, is simply viewed as part of the assumptions encoded in the model structure. The empirical Bayes approach provides one way to avoid arbitrary priors; see Section 2.6.2 for details.

2.3 The Bayesian group model

In the group model, a network of N nodes is partitioned into K groups. The groups to which a potential prey and to which a potential predator belong completely determine the probability that a feeding relationship exists between them. The assignment of species to groups is given by the vector $\mathbf{G} = (g_1, \dots, g_n)$, with $g_i \in \{1, \dots, K\}$. We refer to this assignment as a set ‘partition,’ in keeping with standard mathematical terminology. The probability that a species assigned to group i is consumed by a species assigned to group j is equal to p_{ij} . This gives a matrix \mathbf{P} of K^2 probabilities, containing the probabilities of observing directed links between members of each pair of groups, and within members of each group.

If we take \mathbf{A} to be the directed adjacency matrix of a network, with entries a_{ij} equal to 1 if a

link exists from node i to node j , 0 otherwise, then the probability of the network being generated by partition \mathbf{G} and link probabilities \mathbf{P} is given by

$$f(\mathbf{A}|\mathbf{G}, \mathbf{P}) = \prod_{i=1}^K \prod_{j=1}^K p_{ij}^{Y_{ij}} (1 - p_{ij})^{Z_{ij}}, \quad (2.3)$$

where Y_{ij} and Z_{ij} are the number of 1-entries and 0-entries in the submatrix of \mathbf{A} containing entries from rows r satisfying $g_r = i$ and columns c satisfying $g_c = j$.

In the simplest case, all nodes are assigned to the same group, and the likelihood simplifies to

$$f(\mathbf{A}|p) = p^Y (1 - p)^Z \quad (2.4)$$

where Y and Z are the total number of 1-entries and 0-entries in \mathbf{A} .

In the Bayesian framework, models must include both a formulation of the likelihood and also a prior distribution over model parameters. With the group model, this means defining a prior distribution over both link probabilities and arrangements into groups (set partitions). In general, priors may incorporate informed knowledge about the system, but in this case we simply use them to encode different variants of the same basic model. We use two distributions for partitions and two distributions for link probabilities, which are combined to form four different model variants.

We want to infer the posterior distribution over partitions and parameters,

$$f(\mathbf{G}, \mathbf{P}|\mathbf{A}) \propto f(\mathbf{G}, \mathbf{P})f(\mathbf{A}|\mathbf{G}, \mathbf{P}). \quad (2.5)$$

This requires specifying a prior distribution over partitions \mathbf{G} and link probabilities \mathbf{P} . We consider two priors over \mathbf{G} and two priors over \mathbf{P} .

2.3.1 Link probability prior

Two possible distributions for elements p_{ij} of the group-pair link probability matrix \mathbf{P} are (1) a uniform distribution between 0 and 1, and (2) a beta distribution with shape parameters α and β , which can in turn be governed by a prior distribution, such as an exponential distribution with mean 1, or fixed according to maximum marginal likelihood, covered in Section 2.6. Alternative (2), with exponential priors, is thus a hierarchical model, with exponential *hyperpriors* for α and β governing the beta prior for elements of \mathbf{P} .

Mathematically, the uniform prior over $[0, 1]$ can be written:

$$f(p_{ij}) = 1 \quad 0 \leq p_{ij} \leq 1. \quad (2.6)$$

The beta prior has the form:

$$f(p_{ij}|\alpha, \beta) = \frac{1}{B(\alpha, \beta)} p_{ij}^{\alpha-1} (1 - p_{ij})^{\beta-1}, \quad (2.7)$$

where $B(\alpha, \beta)$ is the beta function,

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1 - t)^{\beta-1} dt. \quad (2.8)$$

The parameters α and β control the shape of the distribution, which may be convex, concave, or skewed toward 0 or 1. When $\alpha = \beta = 1$, the beta prior becomes a uniform distribution.

The beta prior distribution is said to be *conjugate* to the likelihood, because the conditional posterior distribution of the parameters is also beta-distributed. This feature enables analytical integration over probability parameters, conditional on a partition (grouping).

Finally, α and β may be given hyperpriors, for example an exponential distribution with mean

1:

$$f(\alpha) = e^{-\alpha} \quad \alpha \geq 0, \quad (2.9)$$

$$f(\beta) = e^{-\beta} \quad \beta \geq 0. \quad (2.10)$$

2.3.2 Partition prior

The assignment of species into groups, or partition, is also considered a parameter of the model, and, in the Bayesian framework, is also modeled as a probability distribution. That is, the model allows uncertainty in group assignments as well as link probabilities. As a result, we must specify a prior distribution over partitions as well.

A number of choices for prior distributions are possible. In Chapter 3, we consider (1) a uniform distribution and (2) a distribution generated by the Dirichlet process, sometimes referred to as the “Chinese restaurant process” [12]. Alternative (2) is controlled by an aggregation parameter χ that is in turn drawn from an exponential distribution with mean 1. The uniform distribution assigns equal prior probability to each possible partition, irrespective of the number of groups. Because there are far more ways to partition the network at an intermediate, but relatively high, number of groups, the uniform prior implicitly biases the model toward that number. For example, for 100 nodes, the *a priori* expectation is $K = 28$ groups. In contrast, the hierarchically structured Dirichlet process prior provides flexibility via an aggregation parameter, χ . When χ is large, partitions tend to have many small groups; when χ is small, partitions tend to have few groups, with a skewed group size distribution.

In Chapter 4, we use a related alternative to the Dirichlet process, the Dirichlet-categorical distribution [16], in which assignments to groups are assumed to be governed by a categorical distribution with weighted categories, and the vector of weights for each category is governed by a Dirichlet distribution. (The name overlap is unfortunate, since the Dirichlet distribution and the Dirichlet process are different but related.) The Dirichlet process can introduce some bias in the

relative size of groups, and the Dirichlet-categorical distribution, of which the Dirichlet process can be seen as a limiting case, is more flexible in this regard.

Simpler choices that avoid the prior bias toward a moderate number of groups are also possible, such as:

1. A fixed number of groups K , with a uniform prior over groups. Different values of K are then considered to be different models, and can be compared via model selection (Section 2.6).
2. A monotonically decreasing prior distribution over the number of groups K , with a uniform distribution conditional on K . This is a weighted variant of (1). The decreasing distribution over K will be controlled by some top-level parameter, which will either need a prior itself or need to be fixed using maximum marginal likelihood 2.6.2.

Uniform partition prior

The simplest prior over partitions assigns equal probability to each possible assignment of nodes into groups. For a network of N nodes, the number of possible partitions is given by the N th Bell number,

$$\mathcal{B}(N) = \sum_{K=1}^N \mathcal{S}_2(N, K), \quad (2.11)$$

where $\mathcal{S}_2(N, K)$ is the Stirling number of the second kind, the number of ways to partition N objects into exactly K groups,

$$\mathcal{S}_2(N, K) = \frac{1}{K!} \sum_{j=0}^K (-1)^{K-j} \binom{K}{j} j^N. \quad (2.12)$$

Therefore, the prior probability of a particular partition is uniform across all possible partitions

$$f(\mathbf{G}) = \frac{1}{\mathcal{B}(N)}, \quad (2.13)$$

and the prior probability of having exactly K groups is

$$f(K) = \frac{\mathcal{S}_2(N, K)}{\mathcal{B}(N)}. \quad (2.14)$$

Dirichlet process prior for partitions

The Dirichlet process is frequently used as a Bayesian prior for assignment into groups [12, 21, 38]. Consider a restaurant with an infinitely large number of infinitely large tables, all initially empty. The first patron sits alone at the first table, and subsequent patrons may either sit at an occupied table or a new table. They choose occupied tables with weight equal to the number of current occupants, or a new table with weight equal to an aggregation parameter χ . For example, the second patron will sit at the same table as the first patron with probability $1/(1 + \chi)$. In fact, because the process is *exchangeable*, the probability of any pair of patrons sitting at the table is also $1/(1 + \chi)$. If χ is small, there will tend to be a small number of occupied tables and a skewed distribution of table sizes; if χ is large, there will be a larger number of tables occupied by few patrons.

Interpreting tables of patrons as groups of nodes, under the Dirichlet process the prior probability of a particular partition \mathbf{G} is

$$f(\mathbf{G}|\chi) = \chi^K \frac{\prod_{j=1}^K (\eta_j - 1)!}{\prod_{i=1}^N (\chi + i - 1)}, \quad (2.15)$$

where N is the number of nodes in the network, K is the number of groups in the partition, and η_j is the number of nodes in group j . The prior probability of K groups is

$$f(K|\chi) = \frac{|\mathcal{S}_1(N, K)|\chi^K}{\prod_{i=1}^N (\chi + i - 1)}, \quad (2.16)$$

where $\mathcal{S}_1(N, K)$ is a Stirling number of the first kind, equal to the coefficients on x_K in the expansion $x(x - 1)(x - 2) \dots (x - K + 1)$.

In Chapter 3, we give χ an exponential hyperprior distribution with mean 1, a basically arbitrary choice that means that large values are considered a priori unlikely, and that, in the restaurant metaphor, at the prior mean value a new patron is as likely to sit at an empty table as to sit at table with one person present:

$$f(\chi) = e^{-\chi} \quad \chi \geq 0. \quad (2.17)$$

2.3.3 Compartmental modification

In Chapter 3, we also consider a compartmental version of the group model that requires between-group link probability parameters to be lower than corresponding within-group parameters. A simple way to fix this requirement is by defining a new parameter q_{ij} between 0 and 1, and redefining values of p_{ij} so that

$$p_{ij} = \begin{cases} q_{ij} & i = j \\ q_{ij} \min(q_{ii}, q_{jj}) & i \neq j \end{cases} \quad (2.18)$$

This way, all between-group parameters are restricted to be less than the within-group parameters for both groups.

2.3.4 Consensus partitions

For the models presented here, the posterior distribution is estimated via Markov-chain Monte Carlo sampling (Section 2.5), which in the case of group models includes samples of many different partitions of species into groups. As these partitions are potentially all distinct from each other, but represent similar tendencies of species to be grouped together, it is useful to try to summarize the information contained in all the samples in a more compact form. One approach is to construct a pairwise “affinity matrix” for species in the food web, with entries equal to the posterior probability that two species are in the same group. A visual representation of this matrix can illuminate the

group structure, and, for the sake of interpretation, a consensus partition can then be constructed from this matrix using a simple clustering algorithm.

The affinity matrix \mathbf{M} is the posterior probability that two nodes are in the same group and 0 otherwise, that is,

$$\mathbf{M} = \sum_{\mathbf{G}} P(\mathbf{G}|\mathbf{A})\mathbf{M}_{\mathbf{G}}, \quad (2.19)$$

where an entry $\mathbf{M}_{\mathbf{G}}$ is 1 if nodes i and j are in the same group, that is,

$$\mathbf{M}_{\mathbf{G},ij} = \delta_{\mathbf{G}_i, \mathbf{G}_j}, \quad (2.20)$$

where δ is the Kronecker delta and \mathbf{G} is the assignment vector for the partition. This matrix is estimated from MCMC output as the fraction of MCMC samples in which the corresponding species are in the same group:

$$\hat{\mathbf{M}} = \frac{1}{N} \sum_{i=1}^N \mathbf{M}_{\mathbf{G}_i}. \quad (2.21)$$

A consensus partition is formed by applying a hierarchical clustering algorithm to the affinity matrix estimate $\hat{\mathbf{M}}$, and then cutting the dendrogram at some number of groups K , forming a consensus partition with assignment vector \mathbf{G}_K and affinity matrix \mathbf{M}_K . In Chapter 3, the goodness of fit of a consensus partition is measured as the correlation between $\hat{\mathbf{M}}$ and \mathbf{M}_K ; in Chapter 4, simple mean-square-error was used. The best consensus partition is thus identified using the value of K that gives the highest correlation.

A number of different clustering methods can be used to generate a dendrogram by treating $\mathbf{1} - \hat{\mathbf{M}}$ as a distance matrix. In Chapter 3, we used the average-linkage clustering algorithm [34], which generated the lowest error among all the methods in SciPy [25]; in Chapter 4, we used Ward's algorithm [37], for the sake of generating a more parsimonious clustering; its error was still very

close in magnitude to other methods. Both algorithms are available in R [30], SciPy [25], and other commonly used software packages.

2.4 Other food-web models

One of the main advantages of the Bayesian framework, particularly in conjunction with Markov-chain Monte Carlo sampling, is the flexibility it allows in model construction. For example, the work in Chapter 4 is based on a model that uses group structure, as in the group model, to constrain prior distributions for parameters that control the feeding relationships among individual species: species parameters are drawn from group parameters, which are in turn drawn from top-level parameters. Conceptually, structure can be detected in data at multiple scales of organization, so that the details of food-web structure are represented as refinements of higher-level patterns.

All the models considered here are based on describing the probability, p_{ij} , that a species j feeds on a species i , in terms of some set of relationships between the two species. In a random-graph model, p_{ij} is simply a constant value p for all species. In the group model, p_{ij} is defined for each pair of groups. In the niche-based models employed in Chapter 4, the probability of a link is a function of the niche-space location n_i of species i relative to the feeding center c_j and radius r_j of species j , but in addition, each of the individual species parameters are drawn from prior distributions for their group; and group assignment is also inferred.

The work here is based on generative models, which define a probability distribution over food webs without any knowledge of empirical species traits, but discriminative models, which relate the probability distribution over food webs directly to empirical traits, are also possible. Examples in the literature, which have employed Markov-chain Monte Carlo sampling but not the full Bayesian framework, include the latent trait model, which combines inferred feeding tendencies for each species with body size as a predictor [32].

A Bayesian discriminative model was used to relate mechanical and chemical defense traits of plant species of the genus *Psychotria* with their risk of being eaten by different insects, in a collabora-

tion that emerged out of this work [33]. In addition to linking trait data to feeding data, that study is an example of using non-binary link data, specifically, counts of *Psychotria* DNA detected in the guts of different insect specimens. For each insect species j , the counts of different plant species was modeled as a multinomial distribution. The probability that an observation in the gut of species j was from plant species i was modeled as

$$p_{ij} = \frac{e^{q_{ij}}}{\sum_{i'=1}^m e^{q_{i'j}}}, \quad (2.22)$$

where m is the number of plant species and

$$q_{ij} = \sum_{k=1}^n \beta_{jk} x_{ik} \quad (2.23)$$

is a linear combination of plant traits x_{ik} weighted by coefficients β_{jk} that reflect the tolerance of insect j to trait k , with n equal to number of traits. We originally formulated that model to allow inference of groups of insects with similar responses to be detected, so that

$$\beta_{jk} \sim \text{Normal}(\mu_k^{g_j}, \tau_k^{g_j}) \quad (2.24)$$

$$(2.25)$$

with group-level mean and precision parameters $\mu_k^{g_j}, \tau_k^{g_j}$ for each trait k , but found that all species ended up assigned to the same group. Still, using Bayesian priors to constrain the model helped reduce the effective dimensionality of parameter space, by imposing an overall level of regularization to parameter distributions as well as individually for each trait.

Static food-web data may also take the form of continuous-valued link strengths, and the general approaches used here could naturally be adapted to continuous-valued probability distributions.

2.5 Markov-chain Monte Carlo sampling

Some Bayesian models are formulated in such a way that the full posterior distribution is analytically tractable given the data, but the models described here do not possess that feature, aside from the group model once the grouping already been determined. Markov-chain Monte Carlo (MCMC) methods enable stochastic sampling from the posterior distribution in a way that can be tuned to balance exploration of the parameter space with exploitation of contiguous areas of high density. MCMC was originally invented to explore the high-dimensional state space of interacting molecules using the fast computing machines of 1953¹ [28].

The general idea of MCMC is to set up a sequence of dependent samples $\theta_1, \theta_2, \dots$ that is guaranteed to converge to a target distribution, in this case the posterior distribution of the model. The basic Markov-chain Monte Carlo method is the Metropolis-Hastings algorithm [28, 19], which proceeds as follows. Starting from the current sample, a change is proposed, drawn from a *proposal distribution* over possible changes, $q(\theta \rightarrow \theta^*)$. This sample is either rejected, in which case the current sample is repeated, or the proposed sample is accepted as the new sample. The acceptance probability,

$$r(\theta \rightarrow \theta^*) = \min \left\{ 1, \frac{f(\theta^*) q(\theta^* \rightarrow \theta)}{f(\theta) q(\theta \rightarrow \theta^*)} \right\}, \quad (2.26)$$

guarantees that the sequence of samples will converge to the posterior distribution, $f(\theta|D) \propto f(\theta)f(D|\theta)$, the prior times the likelihood.

A useful special case of Metropolis-Hastings is known as Gibbs sampling [14], which can be used if the distribution of a single parameter or group of parameters, conditional on fixed values of all other parameters, is analytically tractable. By using the conditional distribution $f(\theta|\phi)$ as the

¹The calculations were performed on the Los Alamos MANIAC computer, which was of course built to simulate weapons [18]. The senior author on the paper is Edward Teller (the father of the hydrogen bomb), and this paper is by far his most-cited work. According to Google Scholar, it had 25,141 citations at the time of writing.

proposal distribution, where ϕ represents all the parameters except θ , we have the identities

$$q(\theta^* \rightarrow \theta) = f(\theta|\phi) \tag{2.27}$$

$$q(\theta \rightarrow \theta^*) = f(\theta^*|\phi) \tag{2.28}$$

$$f(\theta^*) = f(\phi)f(\theta^*|\phi) \tag{2.29}$$

$$f(\theta) = f(\phi)f(\theta|\phi) \tag{2.30}$$

and therefore

$$r(\theta \rightarrow \theta^*) = \min \left\{ 1, \frac{f(\phi)f(\theta^*|\phi)}{f(\phi)f(\theta|\phi)} \frac{f(\theta|\phi)}{f(\theta^*|\phi)} \right\} \tag{2.31}$$

$$= 1. \tag{2.32}$$

In other words, with Gibbs sampling, all proposals are accepted, meaning that repeated sampling from the conditional distributions of different individual parameters is guaranteed to converge to the posterior distribution.

2.5.1 Metropolis-coupled MCMC

Although the Metropolis-Hastings algorithm is guaranteed to converge to the target distribution given infinite time, local maxima in the likelihood surface can cause a chain to become stuck for long periods. One approach to avoiding this problem, known as “Metropolis coupling,” involves running multiple chains in parallel. One chain, the “cold chain,” explores the target distribution, while the other chains, “hot chains,” explore low-likelihood configurations more freely. Periodically, swaps are proposed between chains, allowing good configurations discovered on hot chains to propagate toward the cold chain.

Rather than exploring the target distribution $f(\theta|D) \propto f(\theta)f(D|\theta)$, heated chains explore

$$f_\tau(\theta|D) \propto f(\theta) [f(D|\theta)]^\tau \quad \tau \in [0, 1], \quad (2.33)$$

where τ is a heating parameter. We use linearly spaced values of τ , with the hottest chain exploring the prior ($\tau = 0$) and the coldest chain exploring the posterior ($\tau = 1$).

Swap moves are standard Metropolis-Hastings proposals, but rather than considering a change to a single chain, they consider a change to the joint distribution of two chains. The acceptance probability is thus the ratio of the joint distribution after and before the move:

$$r((\theta_i, \theta_j) \rightarrow (\theta_j, \theta_i)) = \frac{f(\theta_j) [f(D|\theta_j)]^{\tau_i} f(\theta_i) [f(D|\theta_i)]^{\tau_j}}{f(\theta_i) [f(D|\theta_i)]^{\tau_i} f(\theta_j) [f(D|\theta_j)]^{\tau_j}} \quad (2.34)$$

$$= \left[\frac{f(D|\theta_i)}{f(D|\theta_j)} \right]^{\tau_j - \tau_i}, \quad (2.35)$$

where θ_i, θ_j are the configurations that begin in chains i and j , and τ_i, τ_j are the heat parameters of the two chains.

The use of multiple heated chains also provides a means of estimating the marginal likelihood, as described in 2.6.

2.5.2 MCMC for the group model

For the group model, the samples θ consist of hyperparameters for the model variant as well as the group count K and assignment vector \mathbf{G} . The link probabilities \mathbf{P} governing links between groups are not included, because the likelihood function would not be compatible between partitions with different values of K . One possible solution to this problem would be to include \mathbf{P} in the sampling procedure, restrict K to a particular number for a particular run, and then appropriately weight runs with different values of K . Another approach is reversible-jump MCMC [17], which appropriately handles a mapping between two different parameter spaces as part of the Metropolis-

Hastings proposal ratio.

In fact, numerically sampling values of \mathbf{P} is unnecessary, because the *marginal likelihood* of a partition given model hyperparameters—that is, the posterior distribution, conditional on values of α , β , and \mathbf{G} , integrated over all possible values of \mathbf{P} —is analytically tractable, and can be used directly in the MCMC procedure.

For a beta prior over link probabilities, the likelihood of \mathbf{G} , α and β marginalized over all possible values of \mathbf{P} is

$$f(\mathbf{A}|\mathbf{G}, \alpha, \beta) = \int_{\mathbf{P}} f(\mathbf{P}|\alpha, \beta) f(\mathbf{A}|\mathbf{G}, \mathbf{P}) d\mathbf{P} \quad (2.36)$$

$$= \prod_{i=1}^K \prod_{j=1}^K \int_0^1 \frac{1}{B(\alpha, \beta)} p_{ij}^{\alpha-1} (1-p_{ij})^{\beta-1} p_{ij}^{Y_{ij}} (1-p_{ij})^{Z_{ij}} dp_{ij} \quad (2.37)$$

$$= \prod_{i=1}^K \prod_{j=1}^K \int_0^1 \frac{1}{B(\alpha, \beta)} p_{ij}^{Y_{ij}+\alpha-1} (1-p_{ij})^{Z_{ij}+\beta-1} dp_{ij} \quad (2.38)$$

$$= \prod_{i=1}^K \prod_{j=1}^K \frac{B(Y_{ij} + \alpha, Z_{ij} + \beta)}{B(\alpha, \beta)}, \quad (2.39)$$

where $B(\cdot)$ is the beta function. Similarly, for a uniform prior over link probabilities, the marginal likelihood of a particular partition is simply

$$f(\mathbf{A}|\mathbf{G}) = \prod_{i=1}^K \prod_{j=1}^K B(Y_{ij} + 1, Z_{ij} + 1). \quad (2.40)$$

The proposal distribution for partitions is straightforward. Because the conditional distribution for the assignment of a single species to a group is discrete, Gibbs sampling can be applied by simply enumerating the prior and likelihood for all possible changes to the group.

For continuous parameters, many proposal distributions are possible, including uniform or normal proposals centered on the current parameter value, or a multiplier proposal, which is appropriate for parameters constrained to be positive. The range of the proposal must be chosen to avoid

excessive acceptance, which implies too narrow a search, or excessive rejection, which keeps the search stuck due to too-bold proposals.

The proposal distributions just described are all implemented in the `mc3kit` software package (Section 2.7).

2.6 Model selection via marginal likelihood

The Bayesian framework provides a natural way to make probabilistic inferences based on a particular model. However, we also want to be able to choose between different models by quantifying their relative goodness of fit. One approach to Bayesian model selection can be framed directly in terms of Bayes' rule, mirroring the process for estimating the posterior distribution over parameters for a single model.

Consider two models, M_1 and M_2 , to which we assign prior weight $\Pr(M_1)$ and $\Pr(M_2)$. After the data has been observed, we can calculate the posterior probability of the models using Bayes' rule:

$$\Pr(M_1|D) = \frac{\Pr(M_1)\Pr(D|M_1)}{\Pr(D)}, \quad (2.41)$$

$$\Pr(M_2|D) = \frac{\Pr(M_2)\Pr(D|M_2)}{\Pr(D)}, \quad (2.42)$$

where the denominator is equal to the probability of observing the data unconditional of the particular model at play, $\Pr(D) = \Pr(M_1)P(D|M_1) + \Pr(M_2)P(D|M_2)$. The probabilities $\Pr(D|M_1) = \int_{\theta_1} f(\theta_1)f(D|\theta_1) d\theta_1$ and $\Pr(D|M_2) = \int_{\theta_2} f(\theta_2)f(D|\theta_2) d\theta_2$ are the marginal likelihoods of the two models, corresponding to the denominator in Equation 2.2. If we give the two models equal prior weight, then the relative posterior weight of the two models is simply given by the marginal likelihoods. This reasoning extends naturally to any number of models.

The ratio of the marginal likelihoods is often called the Bayes factor [23, 24, 26], and is equal to the posterior odds ratio of the two models, assuming equal prior weight:

$$B_{12} = \frac{\Pr(D|M_1)}{\Pr(D|M_2)} \quad (2.43)$$

The Bayes factor provides a convenient way to compare models: if $B_{12} = 10$, then we consider support for model M_1 to be ten times stronger than model M_2 . In AIC-based model selection, the Bayes factor is analogous to a ratio of Akaike weights [10].

The marginal likelihood of a model is the likelihood averaged over the prior distribution. That is, it is the likelihood one would expect by randomly sampling parameters from the prior distribution:

$$f(D|M) = \int_{\theta} f(\theta) f(D|\theta) d\theta. \quad (2.44)$$

This value serves as a useful measure of model fit because it directly incorporates the dependence of the likelihood on uncertainty in parameter values, implicitly penalizing extra degrees of freedom [7]. If an additional parameter improves the maximum likelihood but decreases the average likelihood, the model suffers from overfitting relative to the simpler model.

2.6.1 Marginal likelihood estimation

Calculating the marginal likelihood exactly requires integrating over all model parameters, including partitions. As it is impossible to enumerate all possible partitions for networks of any significant size, we would like to use MCMC to estimate the marginal likelihood for the sake of comparison among different models. Marginal likelihood estimates derived from a single chain, such as the harmonic mean estimator of Raftery [26], converge very slowly, because MCMC fails to sample sufficiently from low-likelihood areas. However, it is possible to use the information gathered about low-likelihood areas in heated chains using a technique called thermodynamic integration [27, 6], or path sampling [13].

Assuming a continuum of heated chains, the thermodynamic estimator for the log-marginal likelihood is

$$\log \hat{\mathcal{L}}(M) = \int_0^1 \frac{1}{m} \sum_{i=1}^m \pi(\theta_{i,\tau}) \log \mathcal{L}(\theta_{i,\tau}) d\tau \quad (2.45)$$

where m is the number of samples in the MCMC output, and $\theta_{i,\tau}$ is a single sample from the output in a chain with heat parameter τ [6]. With a finite number of chains, we use the trapezoid rule to numerically integrate this integral, using uneven spacing of heats to improve the estimate [11].

2.6.2 Empirical Bayes: maximum marginal likelihood and AIC

A main feature of the Bayesian framework is the use of prior distributions for parameters, but in the absence of actual prior information about parameter distributions, the choice of priors can be arbitrary, particular at the top level of a hierarchical model. One way to avoid prior dependence is to abandon the strict Bayesian framework at the top level of organization, and identify maximum-likelihood estimates of those parameters. The likelihood function being maximized, however, is the marginal likelihood of the Bayesian model, conditional on the top-level parameters. This approach is known as maximum marginal likelihood, or empirical Bayes [31].

Using Bayesian terminology, this procedure can be seen as assigning a uniform prior distribution to all top-level parameters, and then identifying the posterior mode of those parameters, marginalized over all other parameters. From the frequentist perspective, we have simply designed a model with a complicated shape; the parameters of the model are our top-level parameters, and we are seeking to maximize the likelihood function of those parameters.

Bayesian model selection breaks down under this approach, because when top-level parameters are given unbounded uniform priors—which are not, in fact, proper probability distributions—the marginal likelihood may be undefined. The Bayesian view does, however, give us a simple mechanism for finding estimates of top-level parameters using Markov-chain Monte Carlo: perform

an MCMC run using wide uniform priors for top-level parameters, and estimate the joint mode of those parameters from MCMC samples. Once top-level parameter modes are estimated, we can perform a second MCMC run with the parameters fixed at their estimated modes. The marginal likelihood, conditional on fixed top-level parameters, should now be well-defined and calculable as described above.

However, marginal likelihoods cannot be directly compared, because we have eliminated the penalizing effect of degrees of freedom in the top-level parameters. Instead, we can employ the Akaike information criterion (AIC), a model-selection criterion developed for maximum-likelihood inference [1], using the maximum marginal likelihood estimates as part of our calculation. Intuitively, AIC corrects the likelihood for overfitting, since adding parameters will tend to improve the maximum likelihood score. Technically, AIC is based on an estimate of the Kullback-Leibler divergence, a measure of how far the model is from representing the “truth” [10]. Although the large number of data points in typical food-web data sets makes this point moot, it is generally advisable to use the sample size-corrected AIC [22]:

$$\text{AIC}_c = 2k - 2 \log \mathcal{L} + \frac{2k(k+1)}{n-k-1}, \quad (2.46)$$

where \mathcal{L} is the marginal likelihood estimate given the top-level parameters, k is the number of top-level parameters, and n is the sample size, equal to, e.g., the number of possible links in a food web.

An quantity analogous to the Bayes factor [26] can be constructed based on AIC scores. The ratio of evidence of model A with AIC score $\text{AIC}_c^{(A)}$, compared against model B with AIC score $\text{AIC}_c^{(B)}$, can be calculated as

$$r_{AB} = \exp \left[\frac{\Delta \text{AIC}_c}{2} \right], \quad (2.47)$$

where

$$\Delta\text{AIC}_c = \text{AIC}_c^{(B)} - \text{AIC}_c^{(A)}, \quad (2.48)$$

so that if, e.g., $\Delta\text{AIC}_c = 2$, then model A has $e \approx 2.718$ times as much support as model B. When the evidence ratios of several models are normalized to sum to 1, these are sometimes referred to as AIC weights [10].

2.7 The mc3kit software library

A large side-effect of the research presented here was a general-purpose Java library for performing Markov-chain Monte Carlo inference, mc3kit. In mc3kit, models are defined as a directed acyclic graph of conditional dependencies among random variables and probability distributions, and the software automatically generates proposals in a way that takes advantage of the dependency structure to eliminate redundant calculations. The software includes built-in support for group inference, and generally allows inference of structures that modify the dependency structure of other random variables. In addition, the software architecture supports multi-chain algorithms such as Metropolis-coupled MCMC, and automatically parallelizes them when running on multi-core computers. mc3kit uses the SQLJet library [36] to save program state and posterior samples in the widely supported SQLite database format [35].

Currently, models must be implemented directly in Java, but support for describing models in simple structured text formats is planned. mc3kit still needs substantial work before it is ready for general use, but the source code is available on GitHub at <https://github.com/edbaskerville/mc3kit>.

2.7.1 Bayesian models as directed acyclic graphs

All the models in this dissertation can be described as directed acyclic graphs that specify the conditional dependencies of different random variables (data and model parameters). For example, in the group-niche model of Chapter 4, each possible link in the food web has a probability p_{ij} that depends on the parameter n_i of species i and parameters c_j, r_j of species j . Each of these parameters depends on a prior distribution for its group, which have parameters that in turn depend on prior distributions for trophic levels. When the MCMC algorithm proposes a new value for parameter n_i , for example, link probabilities need to be recalculated, but only the ones in row i of the matrix. If the group prior mean μ_n^g changes, then the probability density of n_i parameters in that group need to be recalculated, but nothing else does, including the link probabilities themselves.

Rather than requiring optimizations based on these structures to be coded by hand, mc3kit uses the dependency-graph structure to automatically infer which calculations need to be redone. Furthermore, mc3kit allows the dependency structure to be rewired in response to, e.g., changes in group structure, so that when a species is moved from group 1 to group 2, the software can simply rewires its prior distributions to those for group 2, and efficiently updates calculations accordingly.

2.7.2 Proposal distributions and automatic tuning

mc3kit includes proposal distributions for group assignment via Gibbs sampling as well as normal and multiplier Metropolis-Hastings proposals for continuous-valued parameters, which are automatically tuned to a target acceptance rate during a burn-in period. Development versions have included experimental support for Gibbs proposals of continuous parameters with conjugate priors, a main feature of software such as BUGS [9], but non-conjugate models and multi-chain inference have been the focus of development to date; better support for Gibbs sampling is planned.

The package also includes support for so-called differential evolution MCMC (DEMC) algorithms [8], using the past history of samples to propose changes to multiple variables simultaneously in a way that takes correlation structure into account.

2.7.3 Parallel multi-chain architecture

Multi-chain MCMC algorithms such as Metropolis-coupled MCMC (Section 2.5.1) may include both proposals that take place on individual chains as well as coupled proposals that involve two or more chains. In `mc3kit`, these algorithms are described as a series of steps, each of which consists of tasks that involve one or more chains, which are executed in maximally parallel fashion.

First, the dependency structure among tasks is inferred from the sequence of steps. For example, step 1 could include proposals to each parameter individually (one task per chain); step 2 could consist of DEMC proposals to blocks of variables (one task per chain); and step 3 could consist of swap proposals between pairs of adjacent chains (one task per pair of chains). If the third step involves a swap proposal between chains 1 and 2, then it cannot begin until the DEMC proposal step is finished for chain 1 and for chain 2, each of which depends independently on the single-parameter proposal step.

Then, tasks are executed in parallel based on this dependency structure. The tasks in step 1 are dispatched via thread pool. When a task on one chain finishes, it automatically dispatches its corresponding task in step 2 without waiting for the other chains. Since step 3 involves multiple chains, a completed task in step 2 dispatches its corresponding step-3 task only if all other dependent tasks have completed. Because the execution engine is based on the task dependency graph rather than global synchronization, it is more tolerant of variability in execution time between chains.

2.8 Lessons learned

Although Bayesian models and Markov-chain Monte Carlo approaches are very appealing conceptually, they present a number of practical challenges. The group-niche models in Chapter 4 are conceptually straightforward but present an extremely high-dimensional integral to be approximated, both in terms of discrete parameters (partitions) and in continuous parameters (niche space). These models in particular motivated the design of `mc3kit`, which, for better or for worse, takes a

shotgun approach to the primary challenge, convergence.

The reality, and the magic, of Markov-chain Monte Carlo methods is that they do not attempt to fully integrate over parameter space, but can only sample, and in the case of these high-dimensional models, can only sample rather sparsely. For large problems that require MCMC, it is impossible to know if MCMC is, in fact, effective. Convergence diagnostics can give some indication, but if important parts of parameter space are simply unreachable via the initial conditions and proposal mechanisms used, they will still indicate convergence. Unfortunately, there is no solution to this aside from trying new proposals and getting the same result—and if there were, we would not need MCMC. That said, the chains do settle down in practice, giving us marginal likelihood estimates and parameter distributions that can be compared within and between models. These are useful, but, to be mathematically precise, must include the caveat that they result from the range of parameters identified by the MCMC *so far*.

The mc3kit package takes a general-purpose, shotgun approach to convergence, for better or for worse. Rather than requiring the model-builder to tune proposals for every parameter and every model, mc3kit provides automatic tuning to rule-of-thumb acceptance rates based on idealized models that may not, unfortunately, be ideal for all models, as well as a general-purpose mechanism for multidimensional proposals. Additionally, mc3kit uses $(MC)^3$ (hence the name) in order to more freely explore parameter space. The idea is to provide several mechanisms for exploring the space in order to reduce the chance of getting stuck.

Still, a watched Markov chain never converges. More advanced Bayesian inference techniques such as variational Bayes [3] or Hamiltonian MCMC [29] would be worth exploring. Additionally, it would be interesting to try to reformulate all the models in this dissertation in a way that enabled more of the parameters to be analytically integrated out, reducing the degrees of freedom that the MCMC needs to cope with.

A conceptual point worth noting is that the Bayesian-ness of a model is in the eye of the beholder. The so-called marginal likelihood is simply a likelihood: it is the probability of observing

the data, given a model that happens to include a hierarchical structure of probability distributions and happens to have no free parameters. Bayesian discourse attaches the labels prior, hyperprior, etc. to those distributions, but they could also be called “random effects”. The arbitrary nature of the parameters controlling the top-level parameter distributions motivates the use of AIC and maximum marginal likelihood in Chapter 4, which fixes the top level of parameters at values estimated to maximize the marginal likelihood, conditional on those values.

In short, Bayesian models and MCMC approaches are useful and powerful, but are also not for the faint of heart. A well-known piece of advice not followed enough in the following chapters: start with simpler approaches; think hard about how to build models that are computationally tractable; and complexify from there.

References

- [1] H. Akaike. “A new look at the statistical model identification”. In: *IEEE Transactions on Automatic Control* 19.6 (1974), pp. 716–723. DOI: 10.1109/TAC.1974.1100705.
- [2] S. Allesina and M. Pascual. “Food web models: a plea for groups”. In: *Ecology Letters* 12.7 (2009), pp. 652–662. DOI: 10.1111/j.1461-0248.2009.01321.x.
- [3] H. Attias. “Inferring parameters and structure of latent variable models by variational Bayes”. In: *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc. 1999, pp. 21–30.
- [4] E. B. Baskerville, A. P. Dobson, T. Bedford, S. Allesina, T. M. Anderson, and M. Pascual. “Spatial guilds in the Serengeti food web revealed by a Bayesian group model”. In: *PLoS Computational Biology* 7.12 (2011), e1002321. DOI: 10.1371/journal.pcbi.1002321.
- [5] T. Bayes. “An Essay towards solving a Problem in the Doctrine of Chances”. In: *Philosophical Transactions of the Royal Society of London* (1763), pp. 370–418. URL: <http://www.jstor.org/stable/105741>.

- [6] P. Beerli and M. Palczewski. “Unified framework to evaluate panmixia and migration direction among multiple sampling locations”. In: *Genetics* 185.1 (May 2010), pp. 313–26. DOI: 10.1534/genetics.109.112532.
- [7] B. M. Bolker. *Ecological Models and Data in R*. Princeton University Press, 2008.
- [8] C. J. ter Braak and J. A. Vrugt. “Differential evolution Markov chain with snooker updater and fewer chains”. In: *Statistics and Computing* 18.4 (2008), pp. 435–446. DOI: 10.1007/s11222-008-9104-9.
- [9] BUGS. URL: <http://www.mrc-bsu.cam.ac.uk/bugs/>.
- [10] K. P. Burnham and D. Anderson. *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*. Springer, 2002.
- [11] B. Calderhead and M. Girolami. “Estimating Bayes factors via thermodynamic integration and population MCMC”. In: *Computational Statistics & Data Analysis* 53.12 (2009), pp. 4028–4045. DOI: 10.1016/j.csda.2009.07.025.
- [12] T. Ferguson. “A Bayesian analysis of some nonparametric problems”. In: *Annals of Statistics* 1.2 (1973), pp. 209–230. DOI: 10.2307/2958008.
- [13] A. Gelman and X. Meng. “Simulating normalizing constants: from importance sampling to bridge sampling to path sampling”. In: *Statistical Science* 13.2 (1998), pp. 163–185. DOI: 10.2307/2676756.
- [14] S. Geman and D. Geman. “Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 6 (1984), pp. 721–741. DOI: 10.1109/TPAMI.1984.4767596.
- [15] M. Girvan and M. Newman. “Community structure in social and biological networks”. In: *Proceedings of the National Academy of Sciences* 99.12 (2002), p. 7821. DOI: 10.1073/pnas.122653799.

- [16] P. J. Green and S. Richardson. “Modelling heterogeneity with and without the Dirichlet process”. In: *Scandinavian Journal of Statistics* 28.2 (2001), pp. 355–375. DOI: 10.1111/1467-9469.00242.
- [17] P. Green. “Reversible jump Markov chain Monte Carlo computation and Bayesian model determination”. In: *Biometrika* 82.4 (1995), p. 711. DOI: 10.1093/biomet/82.4.711.
- [18] F. H. Harlow and N. Metropolis. “Weapons Simulation Leads to the Computer Era”. In: *Los Alamos Science* (1983), p. 133.
- [19] W. K. Hastings. “Monte Carlo sampling methods using Markov chains and their applications”. In: *Biometrika* 57.1 (1970), pp. 97–109. DOI: 10.1093/biomet/57.1.97.
- [20] P. W. Holland, K. B. Laskey, and S. Leinhardt. “Stochastic blockmodels: first steps”. In: *Social Networks* 5.2 (1983), pp. 109–137. DOI: 10.1016/0378-8733(83)90021-7.
- [21] J. P. Huelsenbeck and M. A. Suchard. “A nonparametric method for accommodating and testing across-site rate variation”. In: *Syst. Biol.* 56.6 (Dec. 2007), pp. 975–87. DOI: 10.1080/10635150701670569.
- [22] C. M. Hurvich and C.-L. Tsai. “Regression and time series model selection in small samples”. In: *Biometrika* 76.2 (1989), pp. 297–307. DOI: 10.1093/biomet/76.2.297.
- [23] H. Jeffreys. “Some tests of significance, treated by the theory of probability”. In: *Proceedings of the Cambridge Philosophical Society* 31 (1935), pp. 203–222. DOI: 10.1017/S030500410001330X.
- [24] H. Jeffreys. *Theory of Probability*. 3rd. The International Series of Monographs on Physics. Oxford: Clarendon Press, 1961.
- [25] E. Jones, T. Oliphant, and P. Peterson. *SciPy: Open source scientific tools for Python*. 2001–. URL: <http://www.scipy.org/>.
- [26] R. Kass and A. Raftery. “Bayes factors”. In: *Journal of the American Statistical Association* 90.430 (1995), pp. 773–795. DOI: 10.1080/01621459.1995.10476572.

- [27] N. Lartillot and H. Philippe. “Computing Bayes factors using thermodynamic integration”. In: *Systematic Biology* 55.2 (2006), pp. 195–207. DOI: 10.1080/10635150500433722.
- [28] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. “Equation of state calculations by fast computing machines”. In: *Journal of Chemical Physics* 21.6 (1953), p. 1087. DOI: 10.1063/1.1699114.
- [29] R. Neal. “MCMC Using Hamiltonian Dynamics”. In: *Handbook of Markov Chain Monte Carlo*. 2011, pp. 113–162.
- [30] R Development Core Team. *R: A Language and Environment for Statistical Computing*. ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria, 2010. URL: <http://www.r-project.org>.
- [31] H. Robbins. *An empirical Bayes approach to statistics*. 1956.
- [32] R. P. Rohr, H. Scherer, P. Kehrl, C. Mazza, and L.-F. Bersier. “Modeling food webs: exploring unexplained structure using latent traits”. In: *The American Naturalist* 176.2 (2010), pp. 170–177. DOI: 10.1086/653667.
- [33] B. E. Sedio, E. B. Baskerville, M. D. Hunter, S. J. Wright, J. W. Westbrook, J. R. de Wet, and C. R. Dick. “Character displacement of anti-herbivore defenses facilitates the coexistence of closely-related rainforest trees”. In: (*in prep.*) (2013).
- [34] R. Sokal and C. Michener. “A statistical method for evaluating systematic relationships”. In: *University of Kansas Science Bulletin* 38 (1958), pp. 1409–1438.
- [35] *SQLite*. URL: <http://www.sqlite.org>.
- [36] *SQLJet*. URL: <http://sqljet.com>.
- [37] J. H. Ward Jr. “Hierarchical grouping to optimize an objective function”. In: *Journal of the American Statistical Association* 58.301 (1963), pp. 236–244. DOI: 10.1080/01621459.1963.10500845.

- [38] E. Xing, M. Jordan, and R. Sharan. “Bayesian haplotype inference via the Dirichlet process”.
In: *Journal of Computational Biology* 14.3 (2007), pp. 267–284. DOI: 10.1089/cmb.2006.0102.

Chapter 3

Spatial Guilds in the Serengeti Food Web

This chapter was previously published as:

Baskerville, Edward B., Andy P. Dobson, Trevor Bedford, Stefano Allesina, T. Michael Anderson, and Mercedes Pascual. (2011) Spatial Guilds in the Serengeti Food Web Revealed by a Bayesian Group Model. *PLoS Computational Biology* 7(12): e1002321. doi:10.1371/journal.pcbi.1002321

3.1 Introduction

Food webs, networks of feeding relationships in ecosystems, connect the biotic interactions among organisms with energy flows, thus linking together population dynamics, ecosystem function, and network topology. Ecologists have been using this powerful conceptual tool for more than a century [12, 8, 17]. One particularly relevant aspect of food webs is the subdivision of species into compartments or modules, a feature that has been proposed to contribute to food web stability by constraining the propagation of disturbances through a network [36]. In this definition, compartments are alternately referred to as modules, clusters, or “communities” [20], and are defined by high link density within groups and low link density between them. A large literature has considered the presence of compartments of food webs, with early work concluding that compartmentalization results primarily from habitat boundaries, not from dynamical effects [46], although continuing the-

oretical work has shown that compartmentalization can affect stability [61, 62]. One recent study shows that niche structure can result in compartmentalization [22], but the relationship between compartments and spatial habitat structure remains the strongest empirical pattern identified [29, 51].

Although compartmental structure may be significant at one scale of analysis, compartments alone do not account for much of the topological structure in food webs. Recent work with a probabilistic model considers a more flexible notion of groups, allowing link density to be high or low within any group or between any pair of groups [2]. Groups can thus represent compartments in the previous sense, but can also represent trophic guilds or roles [7, 33], sets of species that feed on, and are fed on, by similar sets of species. By fitting models of this type to data, the dominant topological pattern in the network can be found, which may include compartments, trophic guilds, or some combination of the two. The initial application of this model to empirical food webs from different ecosystems has revealed a predominance of trophic guilds rather than compartments [2].

Two major challenges limit the application of this model in resolving the group structure of food webs and interpreting its biological basis. First, most food webs have poor resolution of primary producers; plants in terrestrial systems and phytoplankton in aquatic ones are typically represented by a few nodes that are highly aggregated taxonomically. Some are aggregated at multiple trophic levels, e.g., the Coachella Valley web [48]; others aggregate only the primary producers, e.g., the El Verde rainforest [49], which identifies basal taxa as categories of plant parts. Another recently published Serengeti food web includes highly aggregated primary producers and varying levels of aggregation at other trophic levels [64]. Some webs that do include high resolution of plants include plant-herbivore bipartite networks, notably one lowland food-web from Papua New Guinea [44], and plant-insect-parasitoid “source webs” [34, 41]. Because primary producers form the base of the food web, high resolution in those groups can facilitate a much better understanding of how spatial organization and habitat type percolate up the web, and how higher trophic levels cut across the habitat structure at lower levels.

Second, some technical problems have hindered the use of probabilistic models in analyzing group structure. Early food web models served as null models for food web structure and were tested by generating model webs and comparing summary statistics against data from real webs [13, 67]. More recently, a more rigorous approach for measuring the goodness of fit of a model has been provided by maximum likelihood and model selection [1, 2]. Two problems still remain within this framework. One is technical: standard model-selection criteria are not applicable to “discrete parameters” such as group membership. The second problem is more fundamental: there are many almost equally good arrangements, and it is desirable to extract information not just from a single best arrangement, but also from the rest of the ensemble.

The Bayesian approach is gaining popularity in ecological modeling due to the philosophical and conceptual appeal of explicitly considering uncertainty in parameter estimation as well as its methodological flexibility [38]. This approach is especially well-suited for handling uncertainty in complex food web models, and allows us to overcome the limitations of the previous implementation of the group model. In network inference, there are only a few examples of complete Bayesian models [24, 45] and a few examples of MCMC for maximum-likelihood inference [11, 66], but Bayesian inference in phylogenetics has been long established [68, 35], and provides a clear methodological analogue.

In this paper, we address the group structure of a newly assembled food web for the large mammals and plants of the Serengeti grassland ecosystem of Tanzania by applying a computational approach to the identification of groups based on Bayesian inference. We specifically ask whether the structure that emerges reflects the underlying spatial dimension, as delineated by the different plant communities that characterize different sub-habitats within the ecosystem, or whether it is determined by trophic dimensions in the form of species guilds that share functional roles.

The Serengeti has been studied as an integrated ecosystem for almost five decades [56, 55, 57], and because of widespread popular familiarity with the consumer-resource dynamics of lions, hyenas, wildebeest, zebra and grasses, it provides a strong intuitive test for probabilistic food web

models. Furthermore, all the primary producers in this Serengeti web are identified to the genus or species level. The plant diversity encompasses a number of distinct grass, herb, and woody plant communities on different soils and across a rainfall gradient [40]. This well-documented structure allows us to examine the extent to which habitat structure defines network topology at multiple trophic levels. Although not yet a comprehensive community web, with the addition of more taxa, such as those in another recently published Serengeti web [64], this data set can become the most highly-resolved terrestrial web available.

3.2 Results

3.2.1 The Serengeti Food Web Data Set

We compiled the Serengeti food web from published accounts of feeding links in the literature [58, 10, 14, 23, 42, 60, 59, 54, 9, 30, 40, 39, 63, 31] along with some links known from personal observation. With a few exceptions, the taxa included are large mammalian carnivores and herbivores and the plant diets of the herbivores. In its current form it is not a comprehensive community web, nor does such a terrestrial web yet exist. Another recently published Serengeti food web is largely complementary, containing many bird, mammal, and invertebrate species not included here, but without high resolution of plants [64]. We have not included invertebrates (insects and parasitic helminths) or birds, but are adding data for these groups for future studies.

The compiled food web consists of 592 feeding links among 161 species (129 plants, 23 herbivores, and 9 carnivores). 507 of the links are herbivorous, and 85 are predatory. The fraction of all possible links (connectance, $C = L/S^2$), ignoring all biological constraints, is equal to 0.023. We attribute the low connectance, as compared to other existing food-web data sets, to the high taxonomic resolution of the plant community.

3.2.2 Performance of Model Variants

We compared marginal likelihood estimates of different model variants to determine which one best describes the Serengeti food web (see Methods). First, we find unequivocal support for the use of group-based models in describing the Serengeti food web, as compared with simple null models that ignore group structure, either by treating each species as its own group or by combining all species into a single group (Table 3.1). We also find that a flexible group model that allows for high or low connectance between and within groups vastly outperforms a compartmental model that restricts between-group connectance to be lower than within-group connectance, with a posterior odds ratio (Bayes factor) of 9.4×10^{306} against the compartmental model.

Table 3.1: Marginal likelihood estimates and Bayes factors relative to best model.

Group model	Partition prior	Link prior	Log MLE	Δ MLE	Bayes factor
One group	—	Uniform	-2828.60	-1556.82	1.5×10^{-676}
161 groups	—	Beta	-2828.60	-1556.82	1.5×10^{-676}
161 groups	—	Uniform	-17967.07	-16695.28	5.2×10^{-7251}
Compartmental groups	Dirichlet process	Beta	-1978.76	-706.97	1.1×10^{-307}
Flexible groups	Uniform	Uniform	-1710.83	-439.04	9.0×10^{-192}
Flexible groups	Uniform	Beta	-1404.32	-132.53	2.9×10^{-58}
Flexible groups	Dirichlet process	Uniform	-1455.32	-183.54	2.9×10^{-80}
Flexible groups	Dirichlet process	Beta	-1271.78	0	1

Additionally, the use of flexible priors vastly improves the fit of the basic model, for both link probability parameters and network partitions. The model variant with beta prior for link probabilities and Dirichlet process prior for partitions performed best. Next, in order, were (1) the model with beta link probability prior and uniform partition prior, (2) the model with uniform link probability prior and Dirichlet process partition prior, and (3) the model with both uniform priors. The strongest variant surpassed its closest competitor by 133 units of (natural) log-likelihood, corresponding to a posterior odds ratio of 3.4×10^{57} against the worse one, and surpassed the model with both uniform priors by 439 units of log-likelihood, a posterior odds ratio of 1.1×10^{191} . In all cases, 95% confidence intervals on the marginal likelihood estimates were less than one unit of

log-likelihood, far less than the differences between models. Given this unequivocal support, we consider results only from the best model variant.

3.2.3 Identification of Model Parameters

We used samples from the posterior distribution to summarize model hyperparameters controlling link probabilities and partitions. The posterior mean number of groups K is 14.1 (95% credible interval 12, 17), and the mean value of the Dirichlet process parameter χ is 3.0 (1.5, 5.1) (Figure 3.1). The prior expectation of χ was 1.0 and the prior expectation of K was 5.3. The finding of posterior values substantially greater than prior values strongly supports the presence of detailed group structure in the Serengeti food web.

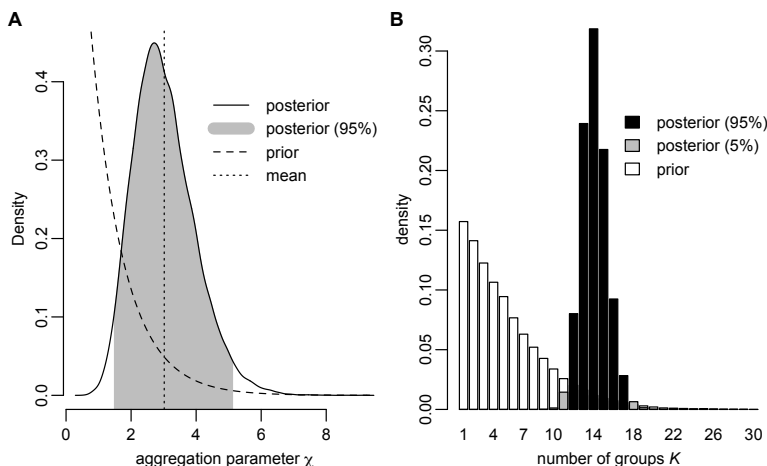


Figure 3.1: Posterior distributions and prior expectations of aggregation parameter χ and group count K . Mean values for beta distribution parameters are $\alpha = 0.044$ (0.027, 0.067) and $\beta = 0.80$ (0.43, 1.30) (Figure 3.2). The corresponding beta prior has support concentrated near 0, since most species do not feed on most other species (Figure 3.3).

3.2.4 Consensus Partition

The posterior output includes 30,000 partitions of the Serengeti food web into groups, nearly all distinct from each other. One partition appears 6 times; two partitions appear 3 times; 14 partitions appear 2 times, and the rest appear only once. For the sake of interpretation, we formed a consensus

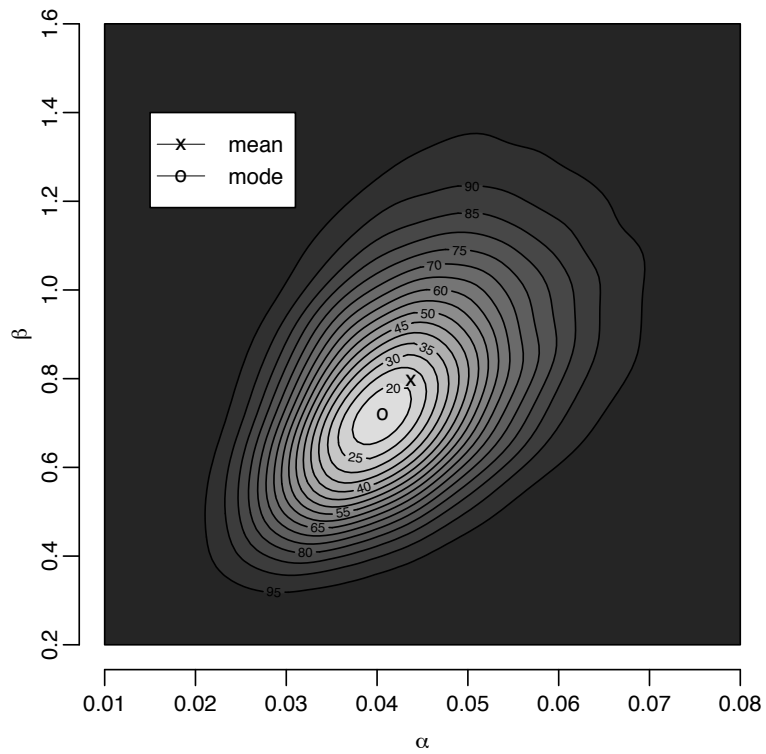


Figure 3.2: Posterior distributions of link density parameters α and β . Color brightness indicates posterior density, estimated using the `ks` multivariate kernel density estimation package for R [16]. Contours indicate cumulative density. The α parameter is significantly lower than 1, indicating departure from a uniform distribution.

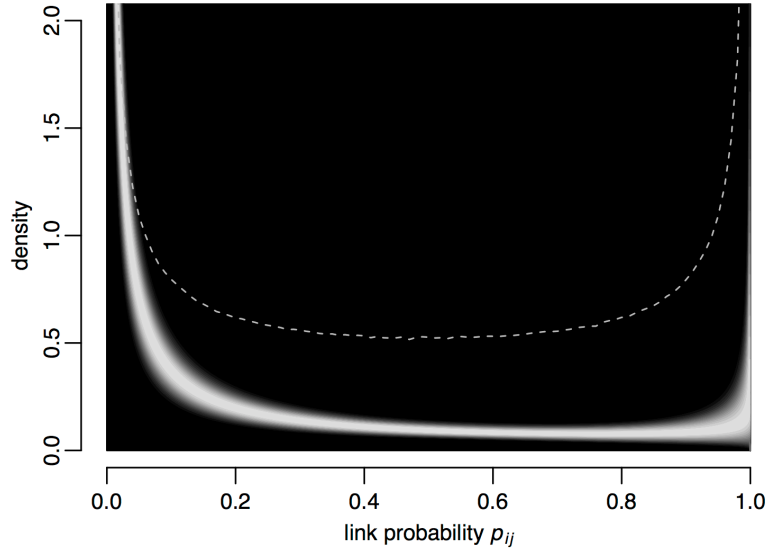


Figure 3.3: Distribution of link probability parameters. The prior distribution for link probability parameters, integrated over the priors for beta distribution parameters α and β , is indicated with a dotted line. The heat map shows beta distributions corresponding to the posterior distribution for α and β , with lightness indicating the posterior density of the parameter values.

partition (Table 3.2) of 14 groups from the affinity matrix (Figure 3.4), which represents the fraction of partitions in all posterior samples in which pairs of species appear in the same group. On average, the consensus partition differs from sampled partitions by 5.6%, calculated as the fraction of species pairs that are assigned to the same group in one partition but to different groups in the other. By comparison, on average, individual sampled partitions differ from other sampled partitions by 7.9%. In addition, every sampled partition differs on average from the others by more than the consensus partition does, indicating the value of the consensus approach.

3.2.5 Groups Identified in the Serengeti Food Web

The groups identified in the Serengeti food web in the consensus partition contain trophically similar species, with all groups restricted to a single trophic level (plants, herbivores, or carnivores). The consensus partition, with 14 groups, is shown in Table 3.2. The partition includes 2 groups of carnivores (groups 1–2), 4 groups of herbivores (groups 3–6), and 8 groups of plants (groups 7–14).

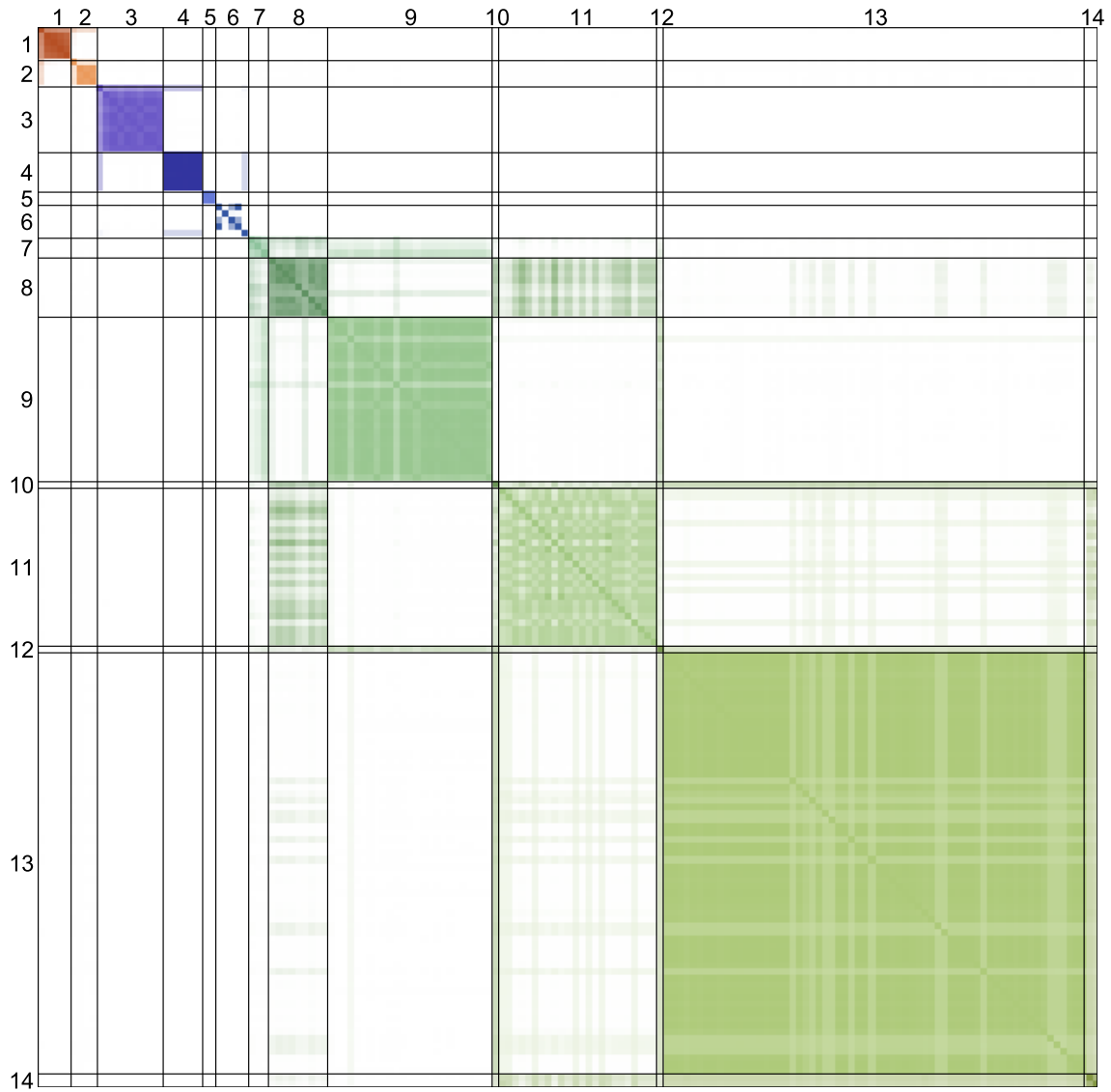


Figure 3.4: Affinity matrix. Species are identically ordered top to bottom and left to right according to the consensus partition as listed in Table 3.2. Hue indicates group identity; color saturation indicates the fraction of partitions in which species occupy the same group. Note that this image conveys information about group membership, not network connectivity.

Table 3.2: Groups identified in the Serengeti food web using a 14-group consensus partition.

Group 1	<i>Acinonyx jubatus</i> , <i>Crocuta crocuta</i> , <i>Lycaon pictus</i> , <i>Panthera leo</i> , <i>Panthera pardus</i>
Group 2	<i>Canis aureus</i> , <i>Canis mesomelas</i> , <i>Caracal caracal</i> , <i>Leptailurus serval</i>
Group 3	<i>Damaliscus korrigum</i> , <i>Hippopotamus amphibius</i> , <i>Kobus ellipsiprymnus</i> , <i>Ourebia ourebi</i> , <i>Pedetes capensis</i> , <i>Phacochoerus africanus</i> , <i>Redunca redunca</i> , <i>Rhabdomys pumilio</i> , <i>Taurotragus oryx</i> , <i>Tragelaphus scriptus</i>
Group 4	<i>Aepyceros melampus</i> , <i>Alcelaphus buselaphus</i> , <i>Connochaetes taurinus</i> , <i>Equus quagga</i> , <i>Nanger granti</i> , <i>Eudorcas thomsonii</i>
Group 5	<i>Heterohyrax brucei</i> , <i>Procavia capensis</i>
Group 6	<i>Giraffa camelopardalis</i> , <i>Loxodonta africana</i> , <i>Madoqua kirkii</i> , <i>Papio anubis</i> , <i>Syncerus caffer</i>
Group 7	<i>Digitaria scalarum</i> , <i>Dinebra retroflexa</i> , <i>Hyparrhenia rufa</i>
Group 8	<i>Chloris gayana</i> , <i>Combretum molle</i> , <i>Digitaria diagonalis</i> , <i>Duosperma kilimandscharica</i> , <i>Eragrostis cilianensis</i> , <i>Microchloa kunthii</i> , <i>Sporobolus festivus</i> , <i>Sporobolus fimbriatus</i> , <i>Sporobolus spicatus</i>
Group 9	<i>Acacia tortilis</i> , <i>Andropogon greenwayi</i> , <i>Aristida</i> spp., <i>Balanites aegyptiaca</i> , <i>Bothriochloa insculpta</i> , <i>Brachiaria semiundulata</i> , <i>Croton macrostachyus</i> , <i>Cynodon dactylon</i> , <i>Digitaria macroblephara</i> , <i>Eragrostis tenuifolia</i> , <i>Eustachys paspaloides</i> , <i>Grewia bicolor</i> , <i>Harpachne schimperi</i> , <i>Heteropogon contortus</i> , <i>Hibiscus</i> spp., <i>Hyparrhenia filipendula</i> , <i>Indigofera hochstetteri</i> , <i>Panicum coloratum</i> , <i>Panicum maximum</i> , <i>Pennisetum mezianum</i> , <i>Sida</i> spp., <i>Solanum incanum</i> , <i>Sporobolus ioclados</i> , <i>Sporobolus pyramidalis</i> , <i>Themeda triandra</i>
Group 10	<i>Pennisetum stramineum</i>
Group 11	<i>Acacia seyal</i> , <i>Acacia xanthophloea</i> , <i>Andropogon schirensis</i> , <i>Chloris pycnothrix</i> , <i>Chloris roxburghiana</i> , <i>Crotalaria spinosa</i> , <i>Cymbopogon excavatus</i> , <i>Digitaria milanjana</i> , <i>Digitaria ternata</i> , <i>Echinochloa haploclada</i> , <i>Eragrostis exasperata</i> , <i>Euphorbia candelabrum</i> , <i>Hyperthelia dissoluta</i> , <i>Kigelia africana</i> , <i>Lonchocarpus eriocalyx</i> , <i>Olea</i> spp., <i>Panicum deustum</i> , <i>Panicum repens</i> , <i>Phragmites mauritanus</i> , <i>Psilolemma jaegeri</i> , <i>Sarga versicolor</i> , <i>Setaria pallidefusca</i> , <i>Setaria sphacelata</i> , <i>Typha capensis</i>
Group 12	<i>Acacia senegal</i>
Group 13	<i>Abutilon</i> spp., <i>Acalypha fruticosa</i> , <i>Acacia robusta</i> , <i>Achyranthes aspera</i> , <i>Albizia harveyi</i> , <i>Albuca</i> spp., <i>Allophylus rubifolius</i> , <i>Aloe macrosiphon</i> , <i>Aloe secundiflora</i> , <i>Blepharis acanthodioides</i> , <i>Capparis tomentosa</i> , <i>Pennisetum ciliare</i> , <i>Cissus quadrangularis</i> , <i>Cissus rotundifolia</i> , <i>Commelina africana</i> , <i>Commiphora merkeri</i> , <i>Commiphora schimperi</i> , <i>Cordia ovalis</i> , <i>Croton dichogamus</i> , <i>Cyperus</i> spp., <i>Cyphostemma</i> spp., <i>Digitaria velutina</i> , <i>Diheteropogon amplexens</i> , <i>Emilia coccinea</i> , <i>Eragrostis aspera</i> , <i>Eriochloa nubica</i> , <i>Ficus glumosa</i> , <i>Ficus ingens</i> , <i>Ficus thonningii</i> , <i>Grewia fallax</i> , <i>Grewia trichocarpa</i> , <i>Heliotropium steudneri</i> , <i>Hibiscus lunariifolius</i> , <i>Hoslundia opposita</i> , <i>Hypoestes forskoolii</i> , <i>Iboza</i> spp., <i>Indigofera basiflora</i> , <i>Ipomoea obscura</i> , <i>Jasminum</i> spp., <i>Kalanchoe</i> spp., <i>Kedrostis foetidissima</i> , <i>Kyllinga nervosa</i> , <i>Lippia ukambensis</i> , <i>Maerua cafra</i> , <i>Ocimum</i> spp., <i>Pappea capensis</i> , <i>Pavetta assimilis</i> , <i>Pavonia patens</i> , <i>Pellaea calomelanos</i> , <i>Phyllanthus sepialis</i> , <i>Pupalia lappacea</i> , <i>Rhoicissus revouilii</i> , <i>Sclerocarya birrea</i> , <i>Senna didymobotrya</i> , <i>Sansevieria ehrenbergii</i> , <i>Sansevieria suffruticosa</i> , <i>Solanum dennekense</i> , <i>Solanum nigrum</i> , <i>Sporobolus pellucidus</i> , <i>Sporobolus stapfianus</i> , <i>Tricholaena teneriffae</i> , <i>Turraea fischeri</i> , <i>Ximenia cafra</i> , <i>Ziziphus</i> spp.
Group 14	<i>Boscia augustifolia</i> , <i>Commiphora trothae</i>

On average, plant groups contain more species than herbivore and carnivore groups (16.1, 5.8, and 4.5, respectively). As evident in the affinity matrix, the carnivore and herbivore groups are well-defined, including several individual species or pairs of species with distinct diets. Plant groups demonstrate mild overlap, indicating a partially hierarchical relationship between smaller groups and larger groups. Figures 3.5, 3.7, and 3.6 show three alternate views of the food web, organized by the 14-group consensus partition. Except for carnivore group 1, there are no connections within groups, and partitions are defined by targeted, directed connections between specific pairs of groups. For actual link densities between groups in the consensus partition, see Table 3.3.

3.2.6 Habitat Signature and Food-Web Structure

Overall, plants of the same habitat type are significantly more clustered in groups than random according to weighted Shannon entropy. (Lower values of weighted entropy indicate higher levels of clustering; see Methods.) Mean weighted entropy across all posterior partitions is 1.25, compared to a randomized mean value of 1.39 ($p < 0.0005$).

Furthermore, the four largest plant groups reflect significant overrepresentation of four different habitat types, and either significant underrepresentation or no significant signal for other habitat types. In group 13, kopje plants are significantly overrepresented, comprising 36.7% of the group, compared to a random expectation of 18.1% ($p \approx 0$). Group 9 contains 60.4% grassland plants compared to a random expectation of 41.5% ($p = 0.02$), and includes 40.4% of individual species records in the plot data. All of the identified riparian species occur in group 11, comprising 31.8% of the group, compared with a 6.3% random expectation ($p \approx 0$). Finally, woodland plants comprise 66.7% of group 8, compared with a random expectation of 25.6% ($p = 0.01$). This result holds across all individual sampled partitions in the posterior output; each one includes four different groups with significant overrepresentation of kopje, grassland, riparian, and woodland habitat.

Plant groups are coupled by groups of herbivores, which are in turn coupled by groups of carnivores. Large migratory grazers (group 4, wildebeest, zebra, and gazelles) feed plant groups that

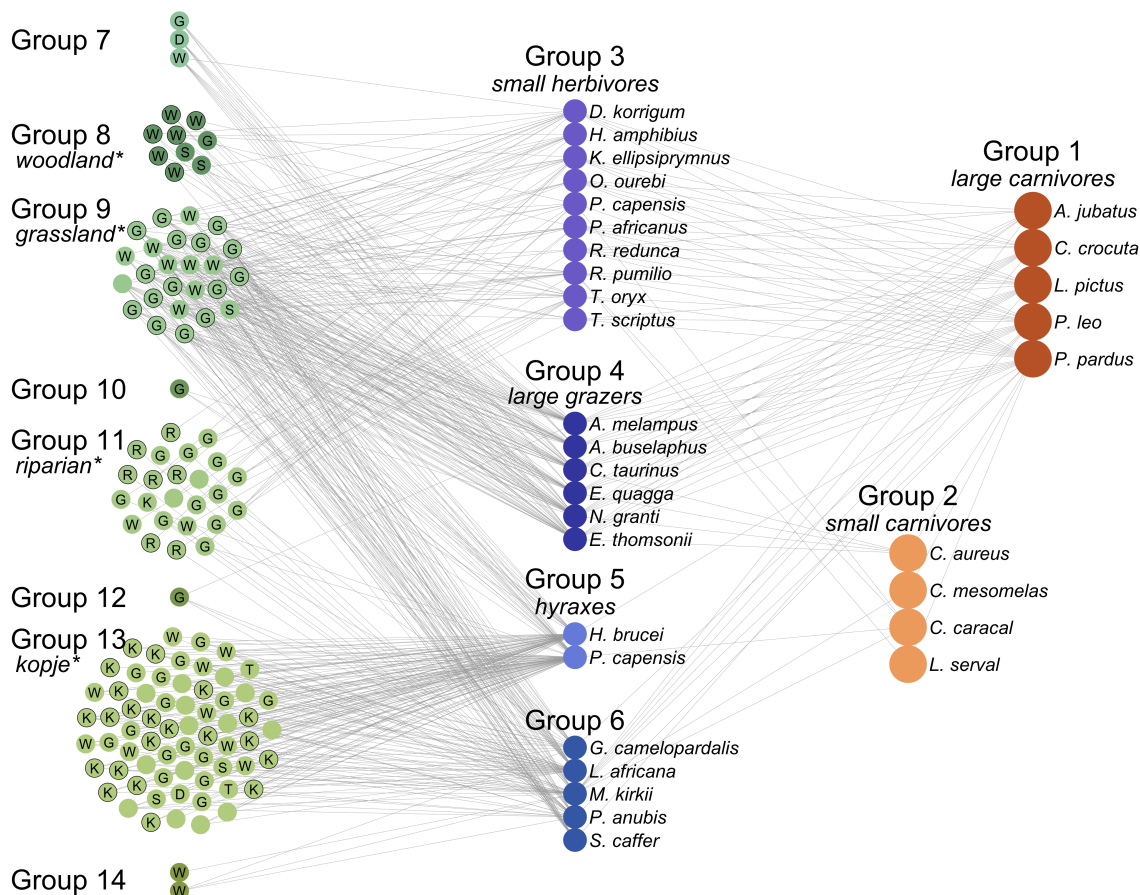


Figure 3.5: The Serengeti food web. The network is shown organized and colored by group according to the consensus partition listed in Table 3.2, and arranged by trophic level from left (plants) to right (carnivores). Plants are identified by the first letter of identified habitat type, if available: (G)assland, (W)oodland, (R)iparian, (K)opje, (S)hrubland, (T)hicket, and (D)isturbed. Plant groups are labeled by significantly overrepresented habitat types, and species assigned to the overrepresented type are labeled with black borders. An interactive version of this figure is available at <http://edbaskerville.com/research/serengeti-food-web/>.

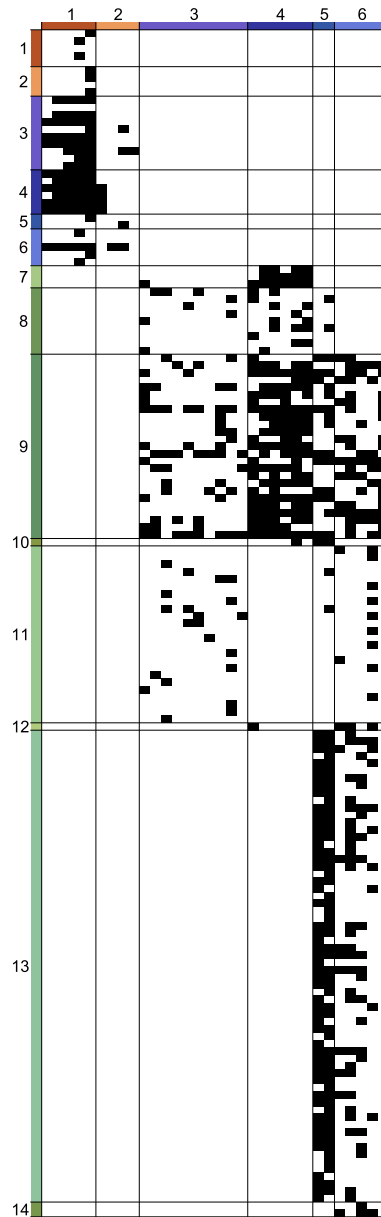


Figure 3.6: Adjacency matrix ordered by groups. Species are identically ordered top to bottom and left to right according to the consensus partition as listed in Table 3.2. White matrix entries indicate that the species in the column feeds on the species in the row. Columns that would indicate prey of plant groups are omitted. Note that in a modular network according to the standard definition, links would be concentrated on the diagonal of the adjacency matrix, since they occur within groups. By contrast, here links are concentrated in off-diagonal blocks.

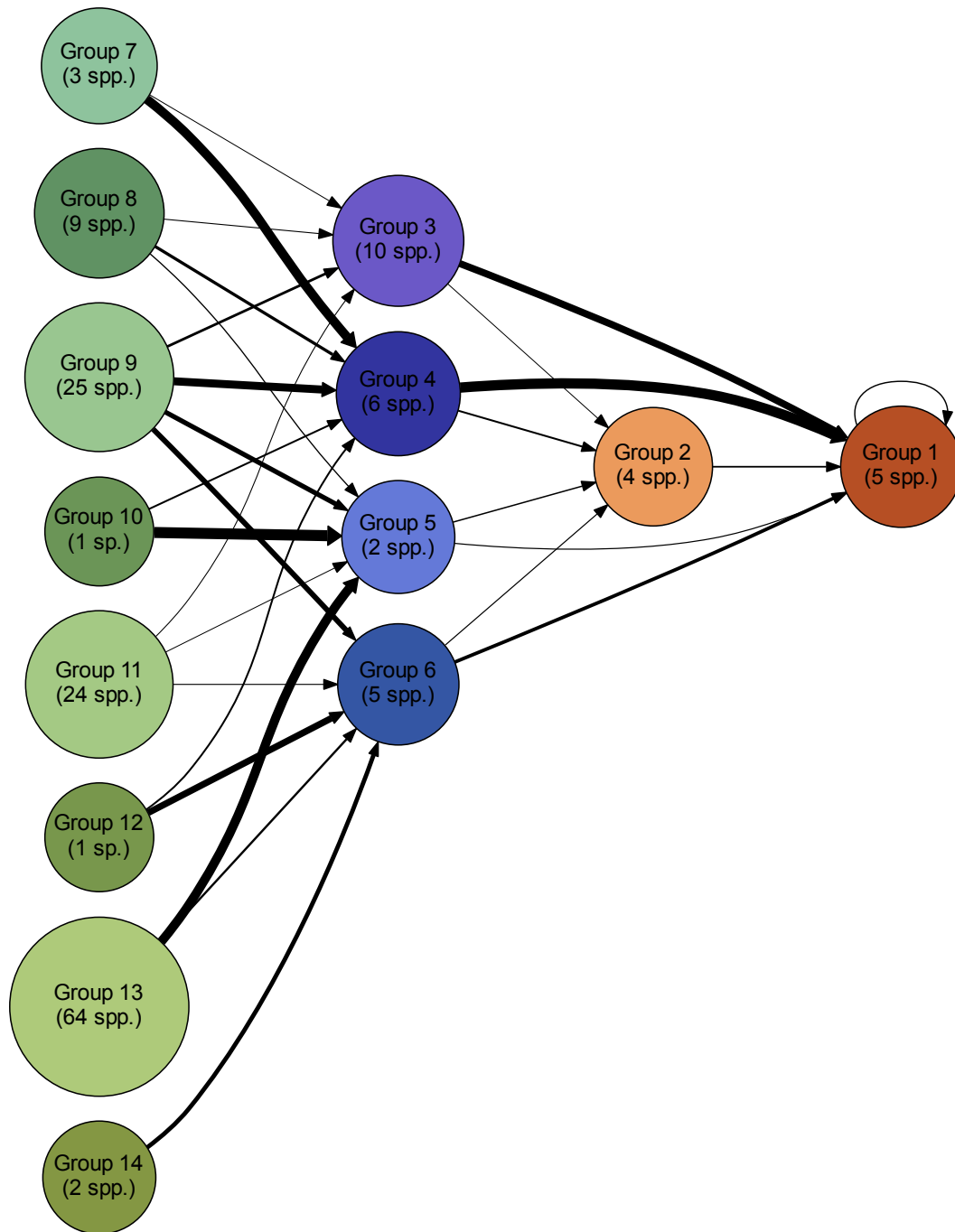


Figure 3.7: Network layout of aggregated groups. Nodes in the network are aggregated and colored by group according to the consensus partition listed in Table 3.2, and arranged by trophic level from left (plants) to right (carnivores). Line thickness indicates the link density between groups. Node area increases with the number of species in a group.

Pred. Group	Prey Group	Link Density
1	1	0.12
1	2	0.15
1	3	0.64
1	4	0.93
1	5	0.10
1	6	0.32
2	3	0.07
2	4	0.17
2	5	0.12
2	6	0.10
3	7	0.03
3	8	0.09
3	9	0.24
3	11	0.09
4	7	0.83
4	8	0.28
4	9	0.73
4	10	0.17
4	12	0.17
5	8	0.11
5	9	0.44
5	10	1.00
5	11	0.04
5	13	0.83
6	9	0.49
6	11	0.09
6	12	0.60
6	13	0.20
6	14	0.40

Table 3.3: Link densities between groups in the consensus partition.

include the dominant grass species in the ecosystem (group 9), predominantly riparian species (group 11), and a mixture of woodland species (*Combretum molle*, *Digitaria diagonalis*, *Duosperma kilimandscharica*, and others) and less common species (group 8). Group 7 represents a specific case where very high trophic similarity brings two spatially separate plants into the same group. *Hyparrhenia rufa* is found mainly in the north, and is a significant component of zebra and wildebeest diets during the dry season; in contrast, *Digitaria scalarum* dominates much of the plains and is eaten in large amounts by migrants during the rainy season when their nutritional needs are at a maximum due to calving and lactation. However, they are grouped together because of their mutual inclusion in the diets of the migratory species. Herbivores feeding in the longer grasslands, woodlands and in riparian habitats (group 3) couple groups 9 and 11. The hyraxes (group 5) and group 6 (giraffe, elephant, buffalo, and others) couple group 13, which bears a strong kopje signature, to groups biased toward other habitats. At the highest trophic level, the large carnivores (group 1) integrate across all the herbivore groups; smaller carnivores (group 2) show more specialized diets, reflecting the more distinct habitats in which they are usually found.

3.3 Discussion

3.3.1 Spatial Guilds in the Serengeti Food Web

In order to analyze the group structure of the Serengeti food web, we used a flexible Bayesian model of network structure that includes no biological information aside from a set of nodes representing species and links representing their interactions. The groups that emerge from an otherwise blind classification of species make remarkable biological sense, and moreover reveal detailed patterns between habitat structure and network topology that expert intuition alone cannot. Species are divided into trophic guilds that reveal a strong relationship between the habitat structure of plant, herbivore, and carnivore groups and the structure of the network. At the coarsest scale, the groups in the Serengeti food web correspond to carnivores, herbivores, and plants. The further subdivisions

that emerge within the trophic levels reveal connections between habitat types and feeding structure. This deeper analysis is made possible by high resolution at the plant level along with information about the habitat occupancy of different plants. Since different habitat types occupy distinct spatial locations in the Serengeti, the group structure thus reflects in part the flow of energy up the food web from different spatial locations, with herbivores integrating spatially separated groups of plants, and carnivores integrating spatially widespread herbivores. A priori, it was not clear precisely what kind of group structure would emerge in the Serengeti web from the use of the group model. In general, the more complex the web, the more useful these methods will be in helping to disentangle the complexity.

The food web presented here included only plants and mammals, but we hypothesize that the general conclusions will be largely robust to the addition of more species. Although the addition of birds, reptiles, invertebrates, and pathogens will likely add a significant number of new groups, they should not significantly modify the derived structure for mammals, since the insect-bird links reflect an almost parallel food web. To the extent that insect herbivores further differentiate plants, plant groupings may be affected, but we expect that the larger tendency for groups to reflect habitat structure will remain.

Recently, interesting theoretical and empirical work has highlighted the relationship between observed patterns of food-web structure and energy flow that seemingly relates to the trophic guild structure in the Serengeti. Rooney and colleagues [52] give evidence that real ecosystems may be dominated by nested sets of fast and slow “energy channels,” each of which represents a food chain of trophic guilds. They suggest that this pattern may have a strong stabilizing effect, based on theoretical work by McCann on spatially coupled food webs [37]. The group structure for the Serengeti web that emerges from our analysis supports a pattern of spatial coupling at multiple trophic levels: the grasslands have very high turnover rates compared to those of the kopjes and woodlands. This suggests a similar pattern of fast and slow energy channels to those described by Rooney and colleagues, with fast energy flow up through the highly seasonal but very productive grasses of the

short-grass plains. These are almost completely consumed by wildebeest and zebra during their peak calving season, which are then in turn consumed by large predators (lions and hyenas). Although the migratory species of wildebeest and zebra form a crucial and major component of the diet of the predatory species, their high abundance and presence in open habitat places them at a lower per capita risk of predation. In contrast, the resident herbivore species living on kopjes and in the woodlands reproduce at slower rates and are consumed at higher per capita rates by large carnivores during the time when the carnivores are unable to feed on migratory wildebeest and zebra.

These patterns emerge directly from the topology of the food web without being explicitly labeled as different habitats upfront as was done in previous empirical work [52], showing that topological analysis can reveal structures that may be very significant for food-web dynamics. They are subtly different, however, from the proposed pure fast and slow chains, in that they incorporate the migration of the keystone species in the ecosystem, so the fastest energy chain is seasonally ephemeral and may only operate for three to four months in any year. We suspect that even within the sub-habitats of kopjes and woodlands there are similarly nested faster and slower chains that involve species for which we are still collating data (e.g., birds, small mammals, and insects). More generally, we see the identification of important structures in empirical food webs via probabilistic models as important for grounding future investigations into the relationship between structure and dynamics in empirical pattern.

3.3.2 Bayesian Analysis of Food-Web Structure

In this paper, we used a probabilistic model to analyze the structure of a single food web, an approach we have seen in only one other study based on a probabilistic version of the niche model [66] (see supporting text S1 for more discussion of probabilistic modeling of food webs). This approach has proved fruitful in Bayesian phylogenetics, where the combinatorial challenges are similar. Moreover, we view the group model as only a starting point for richer modeling efforts to help identify relevant

processes that influence the structure of ecological communities.

In fact, the Bayesian approach described here provides a powerful general framework for encoding hypotheses about the structure of food webs and comparing models against each other, and we see it as a natural next step in the current trend of representing food-web models in a common way. Simple abstract models such as the niche model and the group model used here act as proxies for the high-dimensional trait space that determines feeding relationships in an ecosystem. The identification of actual traits that correspond to groups (or niche dimensions) is another valuable direction, so far followed primarily by finding correlations between compartments/groups [51] or niche values [67] and traits such as body size or phylogenetic relatedness. A more sophisticated, rigorous approach is to directly incorporate such traits into the probabilistic models themselves, either as predictors or as informed Bayesian model priors. Although the current work does not employ that approach, the results from the habitat analysis suggest that including additional information directly in the model would be a valuable approach.

The use of flexible, hierarchical priors for model parameters is another useful innovation of the Bayesian framework. The number of groups identified by the model increases dramatically with the use of a flexible beta prior distribution for link probability parameters. In that model variant, we effectively introduce two degrees of freedom to the model (the beta distribution parameters) but dramatically reduce the effective degrees of freedom of the link probability parameters. Note that we penalize parameters by using the marginal likelihood for model selection, so that the model selection represents a balance between goodness of fit and model complexity. Moreover, this structure makes intuitive sense: since most link probability parameters are simply zero, they should not be penalized. An alternate approach is to remove and add parameters to the model, but this hierarchical technique is much easier to implement in practice.

Advanced Markov-chain Monte Carlo methods make it possible to accurately estimate marginal likelihoods for probabilistic network models. Unlike information criteria such as AIC or BIC, an accurate estimate of the marginal likelihood provides a direct measurement of goodness of fit that

takes into account the degrees of freedom in a model without making any asymptotic assumptions about parameter distributions [6], and can handle discrete parameters such as partitioning into groups that are not properly handled by AIC and BIC.

Additionally, the Bayesian approach also serves as a means to avoid fundamental issues inherent in network models with a large parameter space. One recent study has shown that, even in relatively small networks, a large number of good solutions exist for the standard modularity criterion [43, 21]. A maximization algorithm is thus guaranteed to find a single local maximum of many—possibly even the best one, but certainly not one that captures the full range of good solutions. This problem arises whether the quantity to be maximized is a heuristic such as modularity or a likelihood value. The group model and other parameter-rich models presumably suffer from similar degeneracy problems. In the present case, we find that nearly every partition sampled from the posterior distribution is unique. Although MCMC sampling cannot reproduce the full posterior distribution, it is an important step in the right direction. Philosophical arguments aside, one of the main reasons for maximizing likelihood or modularity is simply that a single solution is far more tractable than a distribution. The consensus partitioning heuristic used here is an attempt to find a single partition that represents the posterior distribution reasonably well for the sake of interpretation and presentation (see Methods). More sophisticated approaches to collapsing partitions will be welcomed. However, since the Bayesian approach provides direct access to uncertainty in the form of the posterior distribution, quantitative analyses should be done across the whole distribution, and we follow that approach here.

3.3.3 Conclusion

The group model, based on the simple notion that groups of species may have similar feeding relationships to other groups, reveals that trophic guilds are the topologically dominant type of group in the Serengeti food web. The model also reveals an interesting relationship between habitat structure and network structure that corroborates recent ideas on spatial coupling in food webs. A theoret-

ical study with a dynamical model suggests that this type of structure may contribute to ‘stability’ in the sense of the persistence of species [37]. Now, by using group structures directly inferred from empirical webs, we can better guide investigations into the relationship between structure and various aspects of stability, for example robustness to secondary extinctions [15, 5]. Although the Bayesian modeling approach is not new to network analysis in general [24, 45], it remains relatively rare. The Bayesian group model, and, more importantly, the general framework for modeling and model selection, naturally extend to other kinds of biological networks, such as metabolic and regulatory networks [27] and networks describing other ecological interactions such as pollination [4]. We advocate this framework as a way to build stronger ties between hypothesis formulation, model building, and data analysis.

3.4 Methods

3.4.1 The Bayesian Group Model

In this work, we use Bayesian probabilistic models to analyze food webs; for a general introduction to the Bayesian modeling approach and details on the specific models used here, please see supporting text S1. We employ a generative model based on groups that treats the food-web network, represented as the presence or absence of each possible feeding link, as data. The group model [2], known as a stochastic blockmodel in the statistical literature [65], was previously treated in a maximum-likelihood framework. In a Bayesian framework, both data and model parameters are treated probabilistically, making the object of inference a posterior distribution over model parameters rather than a point estimate. For a general overview of Bayesian inference, see section 3 of supporting text S1.

The group model (supporting text S1, section 2) divides species into some number of groups K , thus determining a *partition*. All possible links between any pair of groups are assigned the same probability of existing, p_{ij} , for consumer group j and resource group i . If a between-group link

probability p_{ij} is close to one, then there are likely to be many links with a species from group j feeding on a species from group i . A highly compartmental network can be generated by having lower between-group link probabilities p_{ij} (for $i \neq j$) than within-group link probabilities p_{ii} .

3.4.2 Priors and Model Variants

In general, priors may incorporate informed knowledge about the system, but in this case we simply use them to encode different variants of the same basic model. We use two distributions for partitions and two distributions for link probabilities, which are combined to form four different model variants. We also consider several null models for comparison.

Partition Prior

For partitions, we employ two prior distributions: (1) a uniform distribution and (2) a distribution generated by the Dirichlet process, sometimes referred to as the “Chinese restaurant process” [18]. Alternative (2) is controlled by an aggregation parameter χ that is in turn drawn from an exponential distribution with mean 1. The uniform distribution assigns equal prior probability to each possible partition, irrespective of the number of groups. Because there are far more ways to partition the network at an intermediate, but relatively high, number of groups, the uniform prior implicitly biases the model toward that number. For example, for a network of 100 nodes, there is an *a priori* expectation of 41.9 groups. In contrast, the hierarchically structured Dirichlet process prior provides flexibility via the aggregation parameter χ . When χ is large, partitions tend to have many small groups; when χ is small, partitions tend to have fewer groups, with a skewed group-size distribution. See section 3.1 of supporting text S1 for mathematical details and a fuller discussion.

Link Probability Prior

The two alternative prior distributions used for link probabilities p_{ij} are (1) a uniform distribution between 0 and 1, and (2) a beta distribution with shape parameters α and β , which are in turn governed by exponential distributions with mean 1. With α and β fixed at their means, alternative

(2) reduces to a uniform distribution; at other values, the distribution may take a uniform, convex, concave, or skewed shape. The second alternative is thus structured hierarchically, with exponential *hyperpriors* for α and β governing the beta prior for link probabilities p_{ij} . For more details, see section 2.1 of supporting text S1.

Null Models

We also consider two simple models without groups as null comparisons: (1) a directed random graph model (i.e., one group) with a uniform prior on a single link probability parameter p , and (2) a fully parameterized model, with each species in its own group, and a 161×161 link probability parameter matrix \mathbf{P} , also with a uniform link probability prior.

Finally, in order to explicitly restrict the model to detecting compartmental structure, we also consider a modification that requires all between-group link probabilities p_{ij} to be less than corresponding within-group link probabilities p_{ii} and p_{jj} . This is accomplished by adding a parameter q_{ij} for each between-group probability, and setting p_{ij} equal to $q_{ij} \min(p_{ii}, p_{jj})$.

3.4.3 Markov-chain Monte Carlo Sampling

We sample from the posterior distribution of model parameters using a Markov-chain Monte Carlo technique known as Metropolis-coupled MCMC, or (MC)³ [19], which improves mixing between multiple modes of the posterior distribution, and also allows improved estimation of the marginal likelihood [32]. Software for performing MCMC sampling was implemented in Java, and is available from the corresponding author on request. A full treatment of MCMC is given in supporting text S1, section 4, including details on applying the method to the group model.

3.4.4 Bayesian Model Selection

In order to select a good model variant, we employ the marginal likelihood, the probability of data given a model integrated over all model parameters (partitions and link probability parameters). This approach extends the use of Bayes' rule to model selection as well as inference of parameter

values. The ratio of the marginal likelihoods for two models is often called the Bayes factor [25, 26, 28], and determines the posterior odds ratio of two models given equal prior odds. For details on marginal likelihood-based model selection, see text S1, section 5.

3.4.5 Consensus Partitions

The output of an MCMC simulation includes a large number of network partitions representing draws from the posterior distribution. As these partitions are potentially all distinct from each other, but represent similar tendencies of species to be grouped together, it is useful to try to summarize the information contained in all the samples in a more compact form. To do this, we construct an affinity matrix with entries equal to the posterior probability that two species are grouped together. We use the affinity matrix to then form a consensus partition, using an average-linkage clustering algorithm (see supporting text S1, section 6). The affinity matrix is akin to the co-classification matrix previously used to identify uncertainty in end-points in a simulated annealing algorithm [53].

3.4.6 Habitat Signature

In order to test the overall presence of habitat signature in plant groups, we assigned plants to habitat types via one of three methods based on data availability. For plants present in 133 plots sampled from around the Serengeti [3], we assigned them to the habitat type of the plot in which they were most abundant; plot habitat types were assigned via a separately compiled map of habitat boundaries [50]. Some plants were available from a study of kopje forbs [47]. Finally, some were assigned from personal knowledge of the system.

We used a randomization test to measure the overall clustering of habitat in groups across sampled partitions. The habitat signature of an individual group i was measured as the Shannon entropy—low entropy indicates an uneven distribution—of the assignment of species to habitats, $H_i = \sum_j \frac{n_{ij}}{n_i} \log(n_{ij}/n_i)$, where j is the habitat, n_i is the group size, and there are n_{ij} species as-

signed to habitat j within the group. The overall clustering signature for a partition was the average of the individual group entropies, weighted by the size of the groups, $\bar{H} = \sum_i \frac{n_i}{n} H_i$, for total species count n . The p-value for the statistic is the probability that a partition drawn from the posterior distribution has overall clustering greater than or equal to a randomized partition with groups of identical size.

To test clustering significance of a specific habitat type in a specific grouping of species, we calculated the p-value as the probability that a randomized group of the same size would have as many or more species assigned to the chosen habitat type.

References

- [1] S. Allesina, D. Alonso, and M. Pascual. “A general model for food web structure”. In: *Science* 320.5876 (2008), pp. 658–661. DOI: 10.1126/science.1156269.
- [2] S. Allesina and M. Pascual. “Food web models: a plea for groups”. In: *Ecology Letters* 12.7 (2009), pp. 652–662. DOI: 10.1111/j.1461-0248.2009.01321.x.
- [3] T. M. Anderson, J. Shaw, and H. Olf. “Ecology’s cruel dilemma, phylogenetic trait evolution and the assembly of Serengeti plant communities”. In: *Journal of Ecology* 99.3 (2011), pp. 797–806. DOI: 10.1111/j.1365-2745.2011.01795.x.
- [4] J. Bascompte, P. Jordano, C. Melián, and J. Olesen. “The nested assembly of plant-animal mutualistic networks”. In: *Proceedings of the National Academy of Sciences* 100.16 (2003), p. 9383. DOI: 10.1073/pnas.1633576100.
- [5] A. Bodini, M. Bellingeri, S. Allesina, and C. Bondavalli. “Using food web dominator trees to catch secondary extinctions in action”. In: *Philosophical Transactions of the Royal Society B* 364.1524 (2009), p. 1725. DOI: 10.1098/rstb.2008.0278.
- [6] B. M. Bolker. *Ecological Models and Data in R*. Princeton University Press, 2008.

- [7] T. Burns. “Lindeman’s contradiction and the trophic structure of ecosystems”. In: *Ecology* 70.5 (1989), pp. 1355–1362. DOI: 10.2307/1938195.
- [8] L. Camerano. “Dell’equilibrio dei viventi merce la reciproca distruzione”. In: *Accademia delle Scienze di Torino* 15 (1880), pp. 393–414.
- [9] T. Caro. *Cheetahs of the Serengeti Plains: Group Living in an Asocial Species*. University of Chicago Press, 1994.
- [10] R. Casebeer and G. Koss. “Food habits of wildebeest, zebra, hartebeest and cattle in Kenya Masailand”. In: *African Journal of Ecology* 8.1 (1970), pp. 25–36. ISSN: 1365-2028. DOI: 10.1111/j.1365-2028.1970.tb00827.x.
- [11] A. Clauset, C. Moore, and M. E. J. Newman. “Hierarchical structure and the prediction of missing links in networks”. In: *Nature* 453.7191 (2008), pp. 98–101. DOI: 10.1038/nature06830.
- [12] J. E. Cohen. “Lorenzo Camerano’s contribution to early food web theory”. In: *Frontiers in Mathematical Biology*. Vol. 100. Lecture Notes in Biomathematics. Berlin: Springer, 1994, pp. 351–359.
- [13] J. E. Cohen, F. Briand, and C. M. Newman. *Community Food Webs: Data and Theory*. Biomathematics. Springer, 1990. ISBN: 0387511296.
- [14] S. Cooper, K. Holekamp, and L. Smale. “A seasonal feast: long-term analysis of feeding behaviour in the spotted hyaena (*Crocuta crocuta*)”. In: *African Journal of Ecology* 37.2 (1999), pp. 149–160. ISSN: 1365-2028. DOI: 10.1046/j.1365-2028.1999.00161.x.
- [15] J. Dunne and R. Williams. “Cascading extinctions and community collapse in model food webs”. In: *Philosophical Transactions of the Royal Society B* 364.1524 (2009), p. 1711. DOI: 10.1098/rstb.2008.0219.

- [16] T. Duong. “ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R”. In: *Journal of Statistical Software* 21.7 (2007), pp. 1–16. URL: <http://www.jstatsoft.org/v21/i07/>.
- [17] C. S. Elton. *Animal Ecology*. Text-Books of Animal Biology. London: Sidgwick & Jackson, Ltd., 1927.
- [18] T. Ferguson. “A Bayesian analysis of some nonparametric problems”. In: *Annals of Statistics* 1.2 (1973), pp. 209–230. DOI: 10.2307/2958008.
- [19] C. Geyer. “Markov chain Monte Carlo maximum likelihood”. In: *Proceedings of the 23rd Symposium on the Interface*. Ed. by E. Keramidas. Computing Science and Statistics. 1991, pp. 156–163.
- [20] M. Girvan and M. Newman. “Community structure in social and biological networks”. In: *Proceedings of the National Academy of Sciences* 99.12 (2002), p. 7821. DOI: 10.1073/pnas.122653799.
- [21] B. Good, Y. De Montjoye, and A. Clauset. “Performance of modularity maximization in practical contexts”. In: *Physical Review E* 81.4 (2010), p. 46106. DOI: 10.1103/PhysRevE.81.046106.
- [22] R. Guimera, D. Stouffer, M. Sales-Pardo, E. Leicht, M. Newman, and L. Amaral. “Origin of compartmentalization in food webs”. In: *Ecology* 91.10 (2010), pp. 2941–2951. DOI: 10.1890/09-1175.1.
- [23] R. Hansen, M. Mugambi, and S. Bauni. “Diets and trophic ranking of ungulates of the northern Serengeti”. In: *Journal of Wildlife Management* 49.3 (1985), pp. 823–829. DOI: 10.2307/3801717.
- [24] P. Hoff, A. Raftery, and M. Handcock. “Latent space approaches to social network analysis”. In: *Journal of the American Statistical Association* 97.460 (2002), pp. 1090–1098. DOI: 10.1198/016214502388618906.

- [25] H. Jeffreys. “Some tests of significance, treated by the theory of probability”. In: *Proceedings of the Cambridge Philosophical Society* 31 (1935), pp. 203–222. DOI: 10.1017/S030500410001330X.
- [26] H. Jeffreys. *Theory of Probability*. 3rd. The International Series of Monographs on Physics. Oxford: Clarendon Press, 1961.
- [27] H. Jeong, B. Tombor, R. Albert, Z. Oltvai, and A. Barabási. “The large-scale organization of metabolic networks”. In: *Nature* 407.6804 (2000), pp. 651–654. DOI: doi : 10 . 1038 / 35036627.
- [28] R. Kass and A. Raftery. “Bayes factors”. In: *Journal of the American Statistical Association* 90.430 (1995), pp. 773–795. DOI: 10.1080/01621459.1995.10476572.
- [29] A. E. Krause, K. A. Frank, D. M. Mason, R. E. Ulanowicz, and W. W. Taylor. “Compartments revealed in food-web structure”. In: *Nature* 426.6964 (2003), pp. 282–285. DOI: 10.1038/nature02115.
- [30] H. Kruuk. *The Spotted Hyaena*. University of Chicago Press, 1972.
- [31] J. Lamprecht. “On diet, foraging behaviour and interspecific food competition of jackals in the Serengeti National Park, East Africa”. In: *Z. Säugetierkd.* 43 (1978), pp. 210–223.
- [32] N. Lartillot and H. Philippe. “Computing Bayes factors using thermodynamic integration”. In: *Systematic Biology* 55.2 (2006), pp. 195–207. DOI: 10.1080/10635150500433722.
- [33] J. J. Luczkovich, S. P. Borgatti, J. C. Johnson, and M. G. Everett. “Defining and Measuring Trophic Role Similarity in Food Webs Using Regular Equivalence”. In: *Journal of Theoretical Biology* 220.3 (2003), pp. 303–321. DOI: 10.1006/jtbi.2003.3147.
- [34] N. Martinez, B. Hawkins, H. Dawah, and B. Feifarek. “Effects of sampling effort on characterization of food-web structure”. In: *Ecology* 80.3 (1999), pp. 1044–1055. DOI: 10.1890/0012-9658(1999)080[1044:EOSEOC]2.0.CO;2.

- [35] B. Mau, M. Newton, and B. Larget. “Bayesian phylogenetic inference via Markov chain Monte Carlo methods”. In: *Biometrics* 55.1 (1999), pp. 1–12. DOI: 10 . 1111 / j . 0006 - 341X . 1999 . 00001 . x.
- [36] R. M. May. *Stability and Complexity in Model Ecosystems*. Princeton University Press, 1973.
- [37] K. McCann, J. Rasmussen, and J. Umbanhowar. “The dynamics of spatially coupled food webs”. In: *Ecology Letters* 8.5 (2005), pp. 513–523. DOI: 10 . 1111 / j . 1461 - 0248 . 2005 . 00742 . x.
- [38] M. A. McCarthy. *Bayesian Methods for Ecology*. Cambridge University Press, 2007. ISBN: 0521850576.
- [39] S. McNaughton. “Ecology of a grazing ecosystem: the Serengeti”. In: *Ecological Monographs* 55.3 (1985), pp. 259–294. ISSN: 0012-9615. DOI: 10 . 2307 / 1942578.
- [40] S. McNaughton. “Serengeti grassland ecology: the role of composite environmental factors and contingency in community organization”. In: *Ecological Monographs* 53.3 (1983), pp. 291–320. ISSN: 0012-9615. DOI: 10 . 2307 / 1942533.
- [41] J. Memmott, N. Martinez, and J. Cohen. “Predators, parasitoids and pathogens: species richness, trophic generality and body sizes in a natural food web”. In: *Journal of Animal Ecology* 69.1 (2000), pp. 1–15. DOI: 10 . 1046 / j . 1365 - 2656 . 2000 . 00367 . x.
- [42] M. Murray. “Comparative nutrition of wildebeest, hartebeest and topi in the Serengeti”. In: *African Journal of Ecology* 31.2 (1993), pp. 172–177. ISSN: 1365-2028. DOI: 10 . 1111 / j . 1365 - 2028 . 1993 . tb00530 . x.
- [43] M. Newman. “Modularity and community structure in networks”. In: *Proceedings of the National Academy of Sciences* 103.23 (2006), p. 8577. DOI: 10 . 1073 / pnas . 0601602103.

- [44] V. Novotny, S. E. Miller, L. Baje, S. Balagawi, Y. Basset, L. Cizek, K. J. Craft, F. Dem, R. A. I. Drew, J. Hulcr, J. Leps, O. T. Lewis, R. Pokon, A. J. A. Stewart, G. Allan Samuelson, and G. D. Weiblen. “Guild-specific patterns of species richness and host specialization in plant–herbivore food webs from a tropical forest”. In: *Journal of Animal Ecology* 79.6 (2010), pp. 1193–1203. DOI: 10.1111/j.1365-2656.2010.01728.x.
- [45] Y. Park, C. Moore, and J. S. Bader. “Dynamic networks from hierarchical Bayesian graph clustering”. In: *PLoS One* 5.1 (2010), e8118. DOI: 10.1371/journal.pone.0008118.g001.
- [46] S. L. Pimm and J. H. Lawton. “Are food webs divided into compartments?” In: *Journal of Animal Ecology* (1980), pp. 879–898. DOI: 10.2307/4233.
- [47] M. Poelchau and S. Mistry. “Forb diversity and community similarity of kopjes in the Serengeti National Park, Tanzania”. In: *African Journal of Ecology* 44.1 (2006), pp. 38–46. DOI: 10.1111/j.1365-2028.2006.00611.x.
- [48] G. Polis. “Complex trophic interactions in deserts: an empirical critique of food-web theory”. In: *American Naturalist* 138.1 (1991), pp. 123–155. DOI: 10.2307/2462536.
- [49] D. Reagan and R. Waide. *The food web of a tropical rain forest*. University of Chicago Press, 1996.
- [50] D. N. Reed, T. M. Anderson, J. Dempewolf, K. Metzger, and S. Serneels. “The spatial distribution of vegetation types in the Serengeti ecosystem: the influence of rainfall and topographic relief on vegetation patch characteristics”. In: *Journal of Biogeography* 36.4 (2009), pp. 770–782. DOI: 10.1111/j.1365-2699.2008.02017.x.
- [51] E. L. Rezende, E. M. Albert, M. A. Fortuna, and J. Bascompte. “Compartments in a marine food web associated with phylogeny, body mass, and habitat structure”. In: *Ecology Letters* 12.8 (2009), pp. 779–788. DOI: 10.1111/j.1461-0248.2009.01327.x.
- [52] N. Rooney, K. McCann, and J. Moore. “A landscape theory for food web architecture”. In: *Ecology Letters* 11.8 (2008), pp. 867–881. DOI: 10.1111/j.1461-0248.2008.01193.x.

- [53] M. Sales-Pardo, R. Guimera, A. Moreira, and L. Amaral. “Extracting the hierarchical organization of complex systems”. In: *Proceedings of the National Academy of Sciences* 104.39 (2007), p. 15224. DOI: 10.1073/pnas.0703740104.
- [54] G. Schaller. *The Serengeti Lion*. University of Chicago Press, 1972.
- [55] A. R. E. Sinclair and P. Arcese, eds. *Serengeti II: Dynamics, Management, and Conservation of an Ecosystem*. University Of Chicago Press, 1995. ISBN: 0226760316.
- [56] A. R. E. Sinclair and M. Norton-Griffiths, eds. *Serengeti, Dynamics of an Ecosystem*. University Of Chicago Press, 1979. ISBN: 0226760286.
- [57] A. R. E. Sinclair, C. Packer, S. A. Mduma, and J. M. Fryxell, eds. *Serengeti III: Human Impacts on Ecosystem Dynamics*. University Of Chicago Press, 2008.
- [58] A. Sinclair, S. Mduma, and J. Brashares. “Patterns of predation in a diverse predator-prey system”. In: *Nature* 425.6955 (2003), pp. 288–290. ISSN: 0028-0836. DOI: 10.1038/nature01934.
- [59] L. M. Talbot and M. H. Talbot. “The Wildebeest in Western Masailand, East Africa”. In: *Wildlife Monographs* 12 (1963), pp. 3–88. DOI: 10.2307/3830455.
- [60] L. Talbot and M. Talbot. “Food preferences of some East African wild ungulates”. In: *East African Agricultural and Forestry Journal* 27.3 (1962), pp. 131–138.
- [61] J. Teng and K. McCann. “Dynamics of compartmented and reticulate food webs in relation to energetic flows”. In: *American Naturalist* 164.1 (2004), pp. 85–100. DOI: 10.1086/421723.
- [62] E. Thébault and C. Fontaine. “Stability of ecological communities and the architecture of mutualistic and trophic networks”. In: *Science* 329.5993 (2010), p. 853. DOI: 10.1126/science.1188321.
- [63] D. Vesey-FitzGerald. “Grazing succession among East African game animals”. In: *Journal of Mammalogy* 41.2 (1960), pp. 161–172. ISSN: 0022-2372. DOI: 10.2307/1376351.

- [64] S. N. de Visser, B. P. Freymann, and H. Olf. “The Serengeti food web: empirical quantification and analysis of topological changes under increasing human impact”. In: *Journal of Animal Ecology* 80.2 (2011), pp. 484–494. DOI: 10.1111/j.1365-2656.2010.01787.x.
- [65] Y. Wang and G. Wong. “Stochastic blockmodels for directed graphs”. In: *Journal of the American Statistical Association* 82.397 (1987), pp. 8–19. DOI: 10.1080/01621459.1987.10478385.
- [66] R. J. Williams, A. Anandanadesan, and D. Purves. “The probabilistic niche model reveals the niche structure and role of body size in a complex food web”. In: *PLoS One* 5.8 (2010), e12092. DOI: 10.1371/journal.pone.0012092.g001.
- [67] R. J. Williams and N. D. Martinez. “Simple rules yield complex food webs”. In: *Nature* 404.6774 (2000), pp. 180–183. DOI: 10.1038/35004572.
- [68] Z. Yang and B. Rannala. “Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo method”. In: *Molecular Biology and Evolution* 14.7 (1997), p. 717.

Chapter 4

Groups and Niches in Food-Web Structure

4.1 Introduction

Food webs, networks of feeding links in an ecosystem, provide a valuable abstraction for understanding ecological complexity. Nodes in the network are species, possibly aggregated taxonomically or functionally; links represent feeding relationships between species. Food webs thus connect biodiversity directly to a fundamental ecosystem process in a form amenable to mathematical and computational analysis. As food-web structure has a strong effect on the stability and robustness of ecosystems, a better understanding of that structure should enable a better understanding of the response of ecosystems to natural fluctuations and man-made perturbations.

Two basic ideas have underpinned much of the discourse on food-web structure: (1) the organization of species into a low-dimensional niche space of feeding interactions [9]; and (2) the organization of the food web into compartments or groups [26, 3]. In recent years, the use of probabilistic generative models has made it increasingly possible to directly test these and other hypotheses against food web data sets.

The dominance of the feeding-niche idea emerged primarily from the observation of a simple pattern in data. Empirical food-web networks appear to exhibit the property of *intervality*: species can be ordered such that the diet of every species consists of most of the species within a largely contiguous interval in that ordering [8]. In other words, as realized in the topology of the food web, niche

space appears to be one-dimensional. The most straightforward hypothesis for low dimensionality is that a single biological trait, body size, dominates feeding interactions [36]. Statistical modeling has found that although a body-size effect is certainly present, it cannot alone explain empirical food-web structure [29, 37]. Additionally, the statistically detectable dimensionality of this niche space has been found to increase to two dimensions for webs with large numbers of species [39].

The focus on compartmentalization in food webs has been inspired in part by a decades-old observation that, at least under certain assumptions, modular structures are much more stable than non-modular structures [23]. Compartmentalization has been typically seen as indicative of habitat boundaries [26, 21]. More recently, network modularity metrics have been used to identify compartmentalization in the sense of groups of species that have significantly more links among them than would be expected from a random null model [21, 27]. However, these significance tests often do not even take into account the directed nature of trophic links, and even when they do, most detail in link structure is lost at this level of description [33]. A better approach is to describe the web as groups of species that have statistically similar patterns of linkage to other groups, i.e., trophic guilds [3]. The nature of the groups is not specified, allowing meaningful groups to be found whether they emerge from trait clustering, habitat structure, or other processes. Another important realization is that much of the detected compartmentalization in food webs may simply result from niche-space constraints, since collections of species with overlapping interval diets will be more compartmentalized than random [15].

Probabilistic generative models have emerged as a powerful tool for testing hypotheses about food-web structure and for analyzing the details of individual food webs. These models define probability distributions over food webs, providing an objective basis for comparisons among models when fit to data. The niche model, which encoded strict one-dimensional intervality into a generative model, displayed surprising correspondence with the properties of empirical food webs as compared with the previous effort, the cascade model [38, 9], but because of strict intervality it could not be fit to data. Generalizations of the niche model have relaxed the assumption of strict

intervality and have enabled objective, likelihood-based formulations of model fit [2, 34, 37]. The group model, under which all the links between a pair of groups has the same probability, reveals that trophic guild structure gives a good coarse-grained view of food-web structure [3].

A group-based model proved useful in analyzing a plant-mammal food web for the Serengeti ecosystem of Tanzania [10, 5]. In the Serengeti web, which is unusually detailed at the plant level, the group model revealed guild structure at multiple trophic levels as well as the patterns of linkages between the different guilds. Additionally, plant-level guilds reflected statistically detectable patterns of habitat structure. The group model thus provided a meaningful bird's eye view of the network structure. The study also introduced the use of the Bayesian statistical framework for food-web models, allowing uncertainty in group structure to be inferred in a natural way.

To date niches and groups have been considered separately in food-web models, and, depending on the data set, one or the other may give a better statistical description of link structure [3, 39]. However, they complement each other conceptually: at a broader scale of organization, groups of species may share similar feeding patterns. At a finer scale, these feeding patterns may be constrained by a low-dimensional niche space. A natural direction for model development is to combine these two approaches. Furthermore, the Bayesian framework provides a straightforward mechanism for hierarchical model structure, enabling a niche-based model to be embedded in a group structure.

Both niche and group structure are related to the biological traits of species, even if a single trait such as body size does not explain the low dimensionality of a niche space. This fact should be mirrored in part in the evolutionary relationships among the species: under the simplest hypothesis, closely related species should eat similar species and be eaten by similar species, because closely related species will tend to have similar phenotypes. The groups and niche-space parameters inferred from a model can be tested against these traits, revealing the strength of the linkage.

In fact, a recent evolutionary model has shown that low niche dimensionality can emerge from an underlying trait space of very low or very high dimensionality precisely due to this mecha-

nism [31]. Under this model, shared evolutionary history and adaptive evolution produce clusters of species with similar traits. As the number of trait dimensions increases, intervality decreases at first, but then increases again: although the trait space is high-dimensional, closely related species tend to have very similar trait values, and thus high overlap in their predators and prey.

In this study, we present a model that combines the coarse-grained structure of the group model with the constraints of a low-dimensional niche space using a hierarchical Bayesian formulation, and use this model to extend the analysis of the Serengeti food web. We compare the goodness of fit of one- and two-dimensional versions of this model, with and without inferred group structure, alongside two versions of the group model. We then use the inferred groups and niche-space model parameters to further analyze the Serengeti web, including tests for whether body size and evolutionary history are reflected in the niche space of feeding relationships.

4.2 Methods

We fitted six different models, using empirical Bayes [28], to a plant-mammal food web from the Serengeti ecosystem of Tanzania [10, 5]. The six models, detailed in Section 4.2.1, include two versions of the group model and four niche-based models with groups. We inferred model parameters using Markov-chain Monte Carlo (MCMC) methods (section 4.2.4), and estimated model fit using the Akaike information criterion (AIC) calculated from the maximum marginal likelihood estimate (Section 4.2.3).

In order to detect the signature of body size and evolutionary relationships, we tested inferred group structure and niche-model parameters using post-hoc randomization tests (Section 4.2.5).

4.2.1 Models

All models were formulated using a hierarchical structure, but rather than giving arbitrary Bayesian prior distributions to top-level parameters, they were fixed at maximum marginal likelihood (empirical Bayes) estimates 4.2.2.

Name	Description
Group	Group model. The probability of a link between two groups is determined by a single parameter assigned to that pair of groups. Groups are inferred within trophic levels.
Group-DC	Degree-corrected group model. Like the group model, but the between-group probability is multiplied by per-species parameters for prey and predator.
Niche-1D-TL	One-dimensional niche model with trophic-level groups. The probability of a link is determined by the distance of the prey from the center of the predator’s feeding niche, normalized by the size of the feeding niche. Species within trophic levels are given common priors for niche-space parameters.
Niche-1D-G	One-dimensional niche model with groups within trophic levels. Groups are inferred within trophic levels, and species within the same group are given common priors.
Niche-2D-TL	Two-dimensional niche model with trophic-level groups.
Niche-2D-G	Two-dimensional niche model with groups within trophic levels.

Table 4.1: Models

Two changes in model structure were made from the previous study in order to reduce model complexity and computation time: plants were assumed to have no prey a priori, and group structure was inferred within trophic levels (plants, herbivores, carnivores) rather than across all species.

In the group-based models, groupings (set partitions) were given a Dirichlet-categorical prior distribution [14] with a maximum number of groups of 32, 16, and 8 for carnivores, herbivores, and plants, respectively, with the Dirichlet concentration parameter δ_ℓ for each trophic level ℓ fixed at a maximum marginal likelihood estimate.

The individual models are detailed below, and summarized in Table 4.1.

Group model (Group)

The group model [3], also known as a stochastic blockmodel [17], organizes species into groups. Conditional on group structure, all links between each directed pair of groups have the same probability of occurring, p_{ij} for group j feeding on group i . Link probabilities p_{ij} are given a beta

prior:

$$p_{ij} \sim \text{Beta} [\alpha_p, \beta_p] \quad (4.1)$$

with α_p and β_p fixed. The beta prior is conjugate to the likelihood, so that, conditional on \mathbf{g} and α_p, β_p , the marginal likelihood of the model is analytically tractable, and the MCMC does not have to integrate over individual values of p_{ij} .

Degree-corrected group model (Group-DC)

One shortcoming of Model A is the tendency to group nodes of similar degree. In other real-world networks, such as social networks, stochastic blockmodels show substantially improved fit when degree-correction parameters are added to each node [19], so that the degree distribution can be fit independently from the group structure.

In the degree-corrected version of the group model, each link probability is adjusted according to species-specific parameters. Each possible prey i is assigned an in-degree correction parameter θ_i , and each possible predator j is assigned an out-degree correction parameter ϕ_j . The probability p_{ij} that species j feeds on species i is then

$$p_{ij} = \theta_i \phi_j q_{g_i g_j} \quad (4.2)$$

where g_i, g_j are the groups containing species i, j , respectively, and $q_{g_i g_j}$ is a parameter corresponding to that pair of groups. Group-pair parameters and degree-correction parameters are drawn

from beta priors:

$$\theta_i \sim \text{Beta} \left[\alpha_\theta^{(\ell_i)}, \beta_\theta^{(\ell_i)} \right] \quad (4.3)$$

$$\phi_i \sim \text{Beta} \left[\alpha_\phi^{(\ell_i)}, \beta_\phi^{(\ell_i)} \right] \quad (4.4)$$

$$q_{g,h} \sim \text{Beta} [\alpha_q, \beta_q] \quad (4.5)$$

for each species i and each pair of groups g, h , where $\alpha_\theta^{(\ell_i)}, \beta_\theta^{(\ell_i)}, \alpha_\phi^{(\ell_i)}, \beta_\phi^{(\ell_i)}$ are fixed top-level parameters corresponding specific to the trophic level ℓ_i of species i , and α_q, β_q are also fixed.

Niche models (Niche-1D-TL, Niche-1D-G, Niche-2D-TL, Niche-2D-G)

In a D -dimensional niche model, each prey species i is assigned a location \mathbf{n}_i in D -dimensional niche space, determining who eats it. Each predator species i is assigned a feeding center \mathbf{c}_i and feeding radius \mathbf{r}_i , determining whom they eat. The probability p_{ij} that species j feeds on species i is

$$q_{ij} = \exp \left[-x_{ij}^{s_{\ell_j}} \right], \quad (4.6)$$

where s_{ℓ_j} is a shape parameter for trophic level ℓ_j , and x_{ij} is the Euclidean distance between \mathbf{n}_i and \mathbf{c}_j , normalized by \mathbf{r}_j in each dimension:

$$x_{ij} = \sqrt{\sum_{d=1}^D \left[\frac{c_j^{(d)} - n_i^{(d)}}{r_j^{(d)}} \right]^2}. \quad (4.7)$$

This formulation is identical in one dimension but different in multiple dimensions from the previously described probabilistic niche model, which uses a multiplicative formulation across dimensions [39]; preliminary tests indicated that this distance-based formulation yielded a better fit.

The parameters for each species i are drawn from priors determined by its group membership

g_i :

$$n_i^{(d)} \sim \text{Normal} \left[\mu_{n,d}^{(g_i)}, \tau_{n,d}^{(g_i)} \right] \quad (4.8)$$

$$c_i^{(d)} \sim \text{Normal} \left[\mu_{c,d}^{(g_i)}, \tau_{c,d}^{(g_i)} \right] \quad (4.9)$$

$$r_i^{(d)} \sim \text{Exponential} \left[\lambda_{r,d}^{(g_i)} \right] \quad (4.10)$$

where d is the dimension, $\mu_{n,d}^{(g_i)}$ and $\mu_{c,d}^{(g_i)}$ are normal distribution means; $\tau_{n,d}^{(g_i)}$ and $\tau_{c,d}^{(g_i)}$ are normal distribution precisions (inverse variance); and $\lambda_{r,d}^{(g_i)}$ is an exponential distribution rate parameter. In the variants of the model without inferred groups, different priors are defined for each trophic level, with prior hyperparameters fixed. In the group-based variants, niche-space parameters are defined for each group, which are in turn drawn from a hyperprior for each trophic level, whose parameters are fixed:

$$\mu_{n,d}^{(g)} \sim \text{Normal} \left[\mu_{\mu_{n,d}}^{(\ell_g)}, \tau_{\mu_{n,d}}^{(\ell_g)} \right] \quad (4.11)$$

$$\tau_{n,d}^{(g)} \sim \text{Exponential} \left[\lambda_{\tau_{n,d}}^{(\ell_g)} \right] \quad (4.12)$$

$$\mu_{c,d}^{(g)} \sim \text{Normal} \left[\mu_{\mu_{c,d}}^{(\ell_g)}, \tau_{\mu_{c,d}}^{(\ell_g)} \right] \quad (4.13)$$

$$\tau_{c,d}^{(g)} \sim \text{Exponential} \left[\lambda_{\tau_{c,d}}^{(\ell_g)} \right] \quad (4.14)$$

$$\lambda_{r,d}^{(g)} \sim \text{Exponential} \left[\lambda_{\lambda_{r,d}}^{(\ell_g)} \right]. \quad (4.15)$$

4.2.2 Identification of top-level parameters via maximum marginal likelihood

Because of the arbitrary nature of assigning prior distributions to top-level parameters, an empirical Bayes framework was used, so that top-level parameters were fixed at values estimated to maximize the marginal likelihood. Initial Markov-chain Monte Carlo runs were performed with wide uniform priors on top-level parameters, and estimates of maximum marginal likelihood values were calculated using a simple histogram method from posterior samples. Final runs for analysis and

model selection were then performed using top-level parameters fixed at these estimates.

4.2.3 Model selection

Maximum marginal likelihoods were calculated from posterior samples. Because models included different numbers of top-level parameters, sample-size-corrected AIC [1, 18], calculated from maximum marginal likelihoods, served as a model-selection criterion:

$$\text{AIC}_c = 2k - 2 \log \mathcal{L} + \frac{2k(k+1)}{n-k-1} \quad (4.16)$$

where \mathcal{L} is the marginal likelihood estimate given the top-level parameters, k is the number of top-level parameters, and n is the sample size, equal to the number of possible links in the web.

The ratio of evidence of model A with AIC score $\text{AIC}_c^{(A)}$ to model B with AIC score $\text{AIC}_c^{(B)}$ can be calculated as

$$r_{AB} = \exp \left[\frac{\Delta \text{AIC}_c}{2} \right] \quad (4.17)$$

where

$$\Delta \text{AIC}_c = \text{AIC}_c^{(B)} - \text{AIC}_c^{(A)} \quad (4.18)$$

so that if, e.g., $\Delta \text{AIC}_c = 2$, then model A is weighted by a factor of $e \approx 2.718$ relative to model B. This is analogous to the Bayes factor [20], and for models with the same number of top-level parameters, it is exactly equal if the fixed values of top-level parameters are considered part of the model specification. When the relative weights of several models are normalized to sum to 1, these are sometimes referred to as AIC weights [7].

4.2.4 Markov-chain Monte Carlo methods

Model parameters were inferred using Metropolis-coupled Markov-chain Monte Carlo [13] using the mc3kit software package 2.7. Real-valued parameters were sampled using the Metropolis-Hastings algorithm [24, 16] and differential-evolution MCMC with snooker updating [6], with acceptance rates tuned during the burn-in period to a goal of 0.25. Group assignments were sampled using Gibbs sampling [12]. Marginal likelihoods were estimated from multi-chain MCMC output using thermodynamic integration [22].

4.2.5 Randomization tests for taxonomy and body size

We tested for a signal of evolutionary relationships and body size among species using randomization tests across the posterior distribution in both group identity and in niche-space traits. All tests were performed only within trophic levels.

For all distance-based tests, we first constructed a distance matrix based on the trait being tested. Overall taxonomic distance was calculated as 1, 2, 3, or 4 for species in different genera, families, orders, and classes, respectively. Single-level (e.g., genus) distance was 0 for species in the same taxon, 1 for different taxa. Body-size distance was calculated simply as the absolute value of the difference in log-body size.

To test group identity, for each posterior sample we computed the mean distance from each species to the other species in its group for the inferred grouping and for a randomized grouping of the same size; the p-value is the fraction of samples whose mean distance was greater than or equal to than the randomization.

To test relationships with niche-space parameters, for each posterior sample we calculated the Euclidean distance between all niche-space parameters (\mathbf{n} , \mathbf{c} , \mathbf{r}) together and individually, for each pair of species, and calculated the correlation between parameter distances and empirical trait distance. We repeated the calculation with species identities randomized; the p-value is the fraction of samples whose mean correlation was less than or equal to the randomization.

The taxonomy was normalized using the Catalogue of Life database, April 2013, accessed via the Encyclopedia of Life application programming interface [30, 11]. Body size data was gathered from the Animal Diversity Web [25].

4.3 Results

We fitted all six models to the Serengeti food web, compared the models, and performed more detailed analysis on the best one. The Serengeti food web was previously compiled from the literature and contains 592 feeding links among 161 species (129 plants, 23 herbivores, and 9 carnivores), with 507 links in the diets of herbivores and 85 in the diets of herbivores [5].

4.3.1 Model selection

We assessed the fit of each model by calculating the Akaike information criterion based on the marginal likelihood estimate with top-level parameters fixed. Results are shown in Table 4.2. The best-performing model was the two-dimensional niche model with groups (Niche-2D-G). The second- and third-best models, with overlapping confidence intervals, were the one-dimensional niche model with trophic-level groups (Niche-1D-TL) and the degree-corrected group model (Group-DC); the weight of evidence for these models relative to the best model was on the order of 1×10^{-6} . The group model (Group), two-dimensional niche model with trophic-level groups (Niche-2D-TL), and one-dimensional niche model with groups (Niche-1D-G) followed in that order, with substantially worse support.

4.3.2 Group organization

We analyzed the membership of the groups across the posterior distribution by looking at both the number of groups found and the species identity within those groups.

The posterior distributions of group size at each trophic level is shown in Figure 4.1. At the carnivore level, 95% of posterior samples included only a single group. For herbivores, 61% of

	Name	Log MMLE	k	AIC_c	AIC_c 95% CI	ΔAIC_c	Evidence Ratio
1	Niche-2D-G	-1113.0	17	2259.9	(2260.2, 2259.6)	0	1
2	Niche-1D-TL	-1129.3	14	2286.7	(2287.4, 2285.9)	26.7	1.6×10^{-6}
3	Group-DC	-1128.6	15	2287.1	(2287.8, 2286.4)	27.2	1.3×10^{-6}
4	Group	-1148.7	5	2307.4	(2307.7, 2307.2)	47.5	4.9×10^{-11}
5	Niche-2D-TL	-1131.5	24	2310.9	(2311.9, 2310.0)	51.0	8.4×10^{-12}
6	Niche-1D-G	-1174.9	12	2373.8	(2374.1, 2373.6)	113.9	1.9×10^{-25}

Table 4.2: Model-selection results. Log MMLE is the natural logarithm of the maximum marginal likelihood estimate. k is the number of fixed top-level parameters. AIC_c is provided along with 95% bootstrap confidence intervals. ΔAIC_c and evidence ratios are given relative to the best model.

samples contained two groups, and 34% had three groups. Plant groupings had a broader spectrum, with a posterior mode of 7 groups (22% of samples).

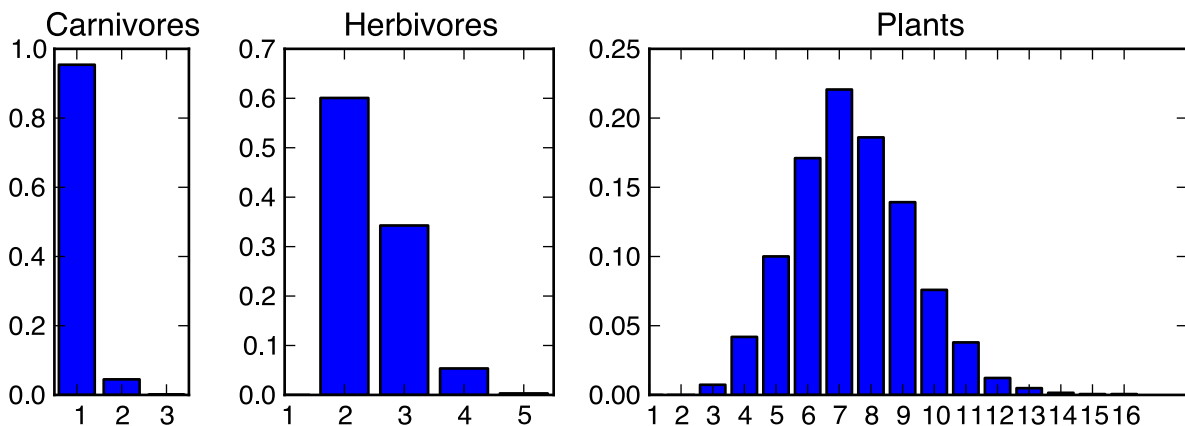


Figure 4.1: Group counts

Herbivore and plant groupings were highly variable across the posterior distribution, so in order to facilitate interpretation we constructed an affinity matrix, representing the fraction of samples in which each pair of species appears in the same group, as the basis for a hierarchical clustering.

The hierarchical clustering for the herbivores is shown in Figure 4.3. The error-minimizing flat cluster contains two groups labeled in the figure. The two large groups are almost supersets of the groups identified in the group-model analysis [5], with the exception of the displacement of *Madoqua*

kirkii and *Taurotragus oryx* to group H1, and *Hippopotamus amphibius*, and *Syncerus caffer* to group H2; these species were also the ones most weakly associated with the other members of their group. The identity of the top-level groups was identical for all clustering algorithms and both error metrics; the clustering from Ward's algorithm is shown [35].

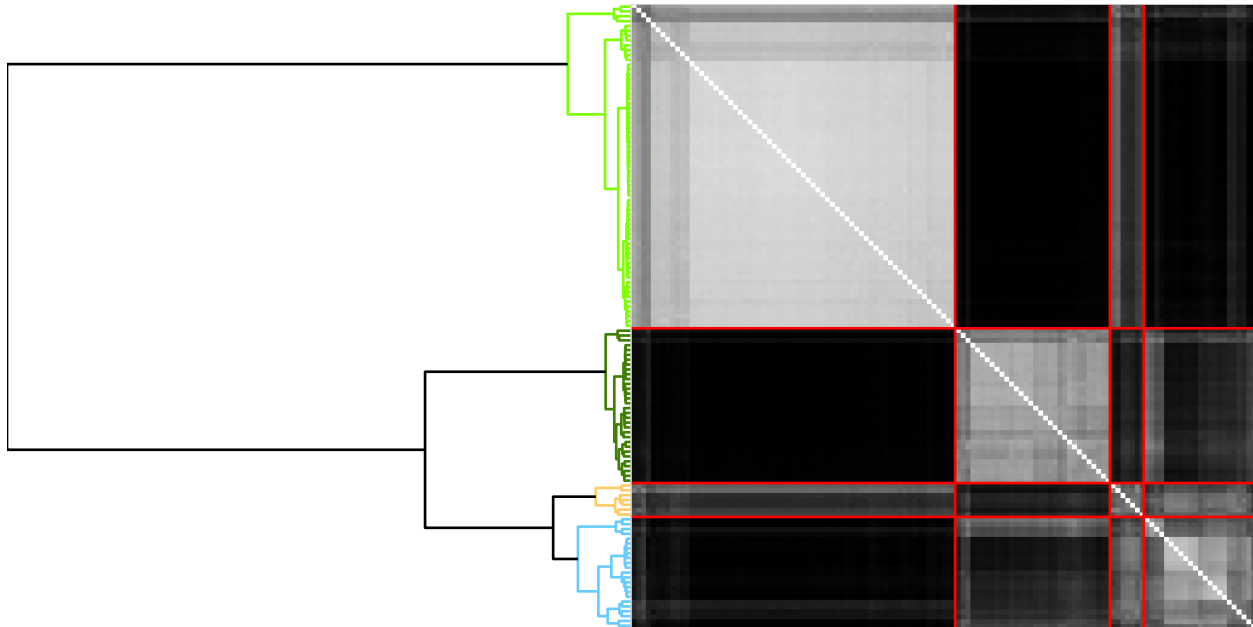


Figure 4.2: Plant groups.

The plant clustering is shown in Figure 4.2. Three large groups of plants were consistently identified across clustering methods. The composition of additional smaller groups varied with the clustering method used; Ward's algorithm was chosen for the figure because it generated a parsimonious four-group clustering and an error similar to the other methods.

The full adjacency matrix of the food web is shown, ordered by groups, in Figure 4.4.

4.3.3 Niche space

Each sample from the posterior distribution represents a possible arrangement of species in niche space. Figure 4.5 is a visualization of the niche space, taken from the sample with the highest poste-

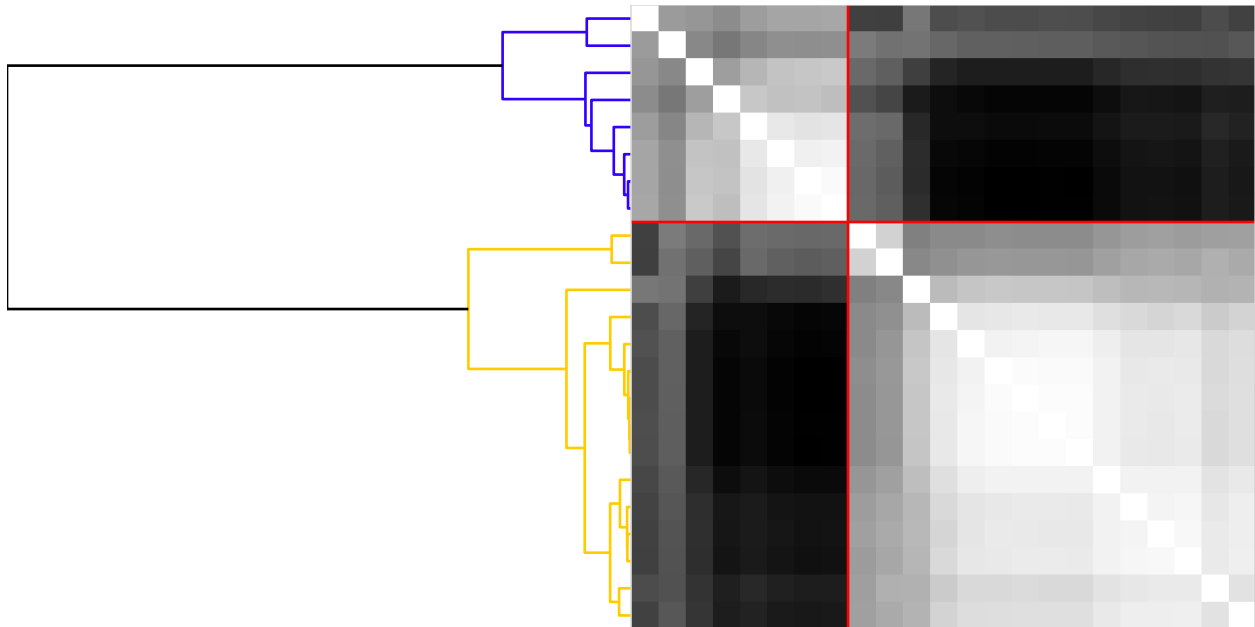


Figure 4.3: Herbivore groups.

rior density, showing prey as points and predator niches as ellipses, with the background indicating the expected number of predators for a prey located at that point in niche space. Species are visibly clustered in niche space by group, in terms of both their \mathbf{n} parameter, controlling their predators and their \mathbf{c} and \mathbf{r} parameters, controlling their prey. This observation verifies the statistical support for the usefulness of including group structure as part of the model. The grouping shown results from consensus across the posterior; as substantial uncertainty was present in group identity, some mixing of groups is not surprising to see in this picture.

4.3.4 Phylogenetic signal in groups and niche space

In order to test for phylogenetic signal in the food web, we tested whether inferred groups and niche-space parameters reflected taxonomic information about species (Section 4.2.5).

The relationship between taxonomic distance and group membership was strongly detectable for plants, and weakly so for herbivores. For plants, the mean within-group taxonomic distance was

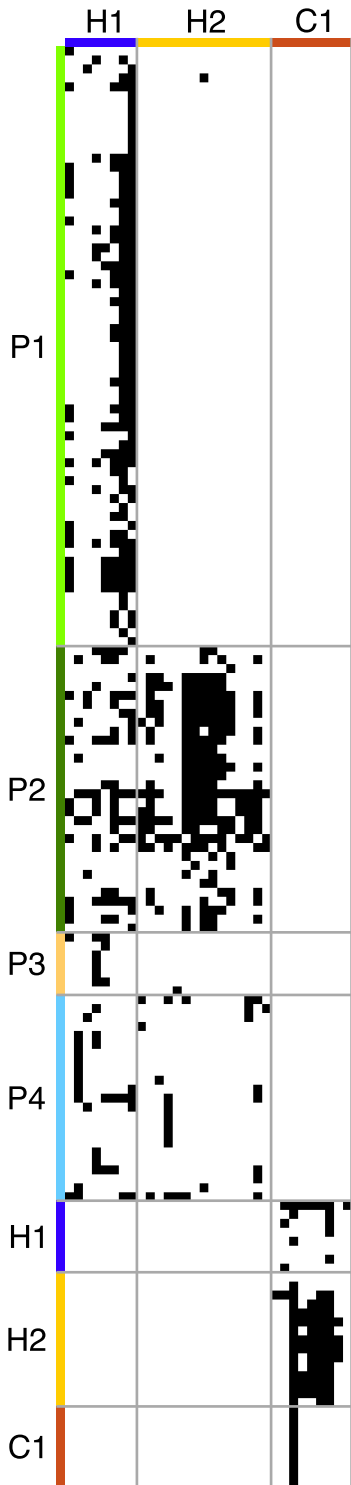


Figure 4.4: Adjacency matrix by groups.

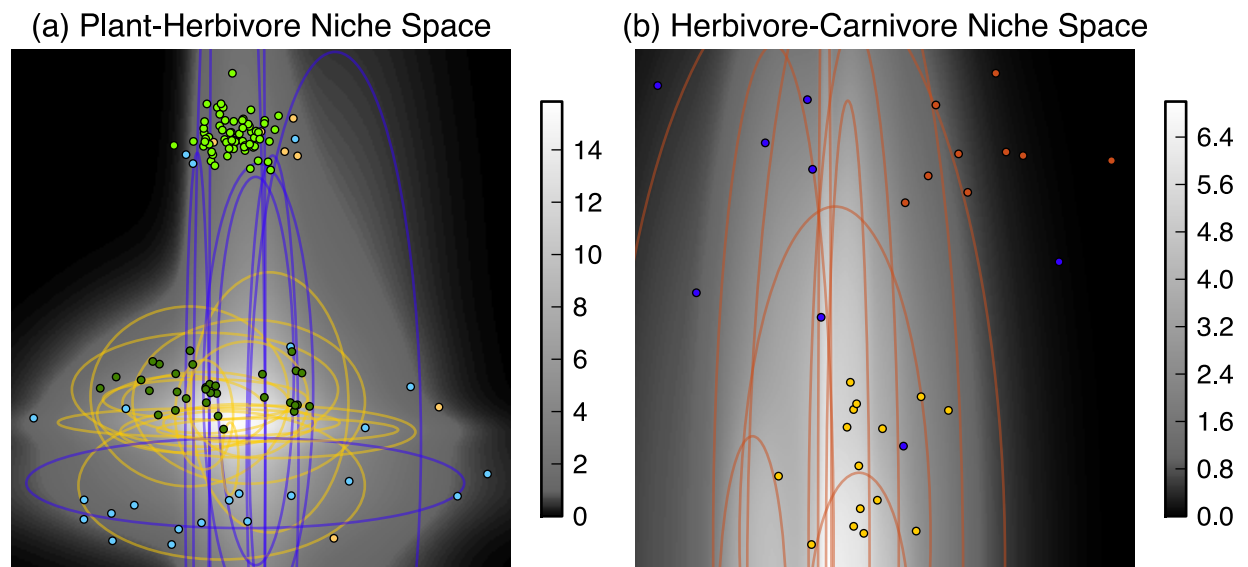


Figure 4.5: Plant-herbivore and herbivore-carnivore regions of niche space. Background gray levels indicate the expected number of predators for prey situated at different points in niche space. Dots indicate n parameters of prey species, and ellipses indicate the 50% probability boundary of feeding by predator species. Dots and ellipses are colored according to the groups shown in Figure 4.4. Plant-herbivore and herbivore-carnivore portions of niche space were perfectly partitioned, so they are shown in separate panels.

2.82, compared with a randomized value of 3.09 ($p \approx 0$), and was detectable at the class ($p \approx 0.0$), order ($p \approx 0.0003$), and family ($p \approx 0.0$) levels (genus, $p \approx 0.18$). For herbivores, the mean within-group taxonomic distance was 2.04, compared to a randomized mean of 2.17 ($p \approx 0.04$). This effect was detectable only at the family level ($p \approx 0.03$).

Niche-space parameters were unequivocally related to phylogeny at all three trophic levels. For plants, the correlation between taxonomic distance and niche-parameter distance was 0.08 (randomized 1×10^{-4} , $p \approx 0$). This relationship was detectable at the class and genus levels in both dimensions of niche space, but restricted to the second dimension at the order and family levels. For herbivores, the correlation was 0.51 (randomized 0.002, $p \approx 0$), detectable at the order, family, and genus levels. For carnivores, the correlation was 0.50 (randomized 0.01, $p \approx 0$), detectable at the family and genus levels.

4.3.5 Signal of body size

We also tested whether species were clustered in niche-parameter space by body size, and whether directional correlations were present between niche-space parameters and body size.

Herbivore body size was clustered in parameter space, with an overall correlation coefficient of 0.14 between parameter-space distance and body-size distance ($p \approx 0.01$). This correlation existed primarily in the **n** parameters, with a correlation coefficients of 0.16 ($p \approx 0.01$) and 0.11 ($p \approx 0.05$) in niche dimensions 1 and 2, respectively. Carnivore body-size clustering was only weakly detectable, with a correlation coefficient of 0.2 but with $p \approx 0.06$.

However, no directional correlations between body size and niche-space parameters were detectable for carnivores or herbivores.

4.4 Discussion

In this study, we combined two different approaches to modeling food-web structure: (1) the niche model, in which species feed in contiguous regions of an abstract niche space that acts as a proxy

for the traits involved in ecological interactions; and (2) the group model, based on the notion that food webs have identifiable groups with similar linkage patterns to other groups. The Bayesian hierarchical modeling framework enabled an intuitive relationship between groups and niche-space parameters, so that, conditional on group structure, species in the same group had the same prior parameter distribution. The model thus allowed the simultaneous inference of both group structure and niche relationships from network data.

The constraints of niche space alone cause a particular kind of group structure: since feeding ranges in the niche space will overlap, groups of species with overlapping ranges will have similar trophic relationships. This connection has been quantitatively explored in a previous study, using the notion of compartmentalization from the point of view of prey and predators [15]. More generally, niche structure is an abstraction of relationship between biological traits that determine feeding interactions, and group structure is a different, coarser abstraction that may emerge from clusters of traits or habitat structure. The results here confirm the validity of the niche-space representation of food webs to describe diet structure, reiterating results from other work [37, 39, 33]. The results also confirm that group structure also is a useful descriptor of food-web structure at a higher level of organization.

Ultimately, group and niche structure both derive from the biological traits of species. The strong relationship between taxonomic distance, group structure, and niche-space parameters can be seen simply as the effect of species traits on feeding structure: a default assumption is that closely related species will have similar traits, and thus will engage in similar feeding relationships. Based solely on taxonomic information and body size, this basic pattern is detectable in the Serengeti food web. However, due to competition-driven evolutionary processes, one might expect that closely related plant species may differ substantially from each other in defense traits, and that these differences could affect the spatial structure of plant communities, as detected in *Psychotria* communities in relation to insect herbivores [32]. In the Serengeti, a study of plant competitive traits demonstrated interesting, but subtle, patterns in the relationships among phylogeny, trait evolution, spatial

structure, and environmental gradients, and it would be valuable to extend this analysis to defense traits and food-web structure [4]. A natural approach for such an extension would be to use discriminative, rather than generative models, that predict feeding links directly from empirical trait data, as in the *Psychotria* study, or from a combination of empirical data and inferred traits, akin to a study that combined body size with inferred traits [29].

The low effective dimensionality of feeding niche space has long been assumed to be the result of a low-dimensional trait space controlling feeding relationships, but recent theoretical work using an evolutionary trait model has demonstrated that low-dimensional orderings can arise from a low-dimensional or a very high-dimensional trait space [31]. The mechanism that produces low-dimensional orderings in the model is, essentially, shared evolutionary history. Under the model, niche-space traits evolve according to a random process, and extinctions occur due to absence of prey and die to competitive exclusion. Because exclusion is weak enough for very similar species, clusters of closely related species with similar trait values can emerge. Although trait dimensionality may be high, clusters of closely related species with similar traits have a high degree of diet overlap, resulting in high food-web intervality; adaptive evolution of predators to prey traits increases similarity, and thus intervality, even further. In light of this result, the weak signal of body size on food-web structure demonstrated in several studies [39, 29] need not contradict the high degree of intervality present in food webs.

Beyond the success of the two-dimensional niche model with groups, there were several surprising patterns in the comparison among models. The degree-corrected group model gave a substantially better fit than all but two of the niche models tested, and only weakly worse than one of those two. This result suggests that the group model has been hampered by the necessity of using group structure to fit the degree distribution of networks, as has been demonstrated in other kinds of networks [19]. Furthermore, much of the success of the probabilistic niche model may have been due primarily to its ability to better reproduce the degree distribution of the individual species. We thus hypothesize that a degree-corrected group model may outperform the probabilistic niche model in

some larger webs.

Furthermore, adding group structure within trophic levels improved the fit of the niche model in two dimensions, but not in one. This is likely a reflection of the lack of one-dimensional intervality in the web: the necessary presence of gaps in the niche structure may necessitate more idiosyncratic, species-specific ranges, whereas in two dimensions the similarities between species are smoother, and thus amenable to clustering in the same distribution.

The two-dimensional niche model suffers from the curse of too many parameters, but the Bayesian grouping technique allows the dimensionality of parameter space to be effectively reduced, turning the second-worst model into the best model. This in and of itself is a lesson in model complexity, albeit an old one: a model may be abstractly very good, but if dimensionality is not constrained it may provide a poor fit to data. Similarly, with the one-dimensional model, an attempt at constraining dimensionality will fail if the inherent dimensionality is too high.

Two primary ideas, niche space and group structure, have both provided important insights into food-web structure. By combining the two into a single model, we have shown that, in the Serengeti food web, the two approaches are not merely reflections of the same information, but are in fact both important components of how food webs are organized. The group structure also improves detection of niche space, since group-based constraints on niche-space parameters allow a higher-dimensional model to be recovered. Furthermore, the group structure and niche-space parameters, although purely inferred from the link data, reflect the biological traits of species as seen through evolutionary relationships.

References

- [1] H. Akaike. “A new look at the statistical model identification”. In: *IEEE Transactions on Automatic Control* 19.6 (1974), pp. 716–723. DOI: 10.1109/TAC.1974.1100705.

- [2] S. Allesina, D. Alonso, and M. Pascual. “A general model for food web structure”. In: *Science* 320.5876 (2008), pp. 658–661. DOI: 10.1126/science.1156269.
- [3] S. Allesina and M. Pascual. “Food web models: a plea for groups”. In: *Ecology Letters* 12.7 (2009), pp. 652–662. DOI: 10.1111/j.1461-0248.2009.01321.x.
- [4] T. M. Anderson, J. Shaw, and H. Olf. “Ecology’s cruel dilemma, phylogenetic trait evolution and the assembly of Serengeti plant communities”. In: *Journal of Ecology* 99.3 (2011), pp. 797–806. DOI: 10.1111/j.1365-2745.2011.01795.x.
- [5] E. B. Baskerville, A. P. Dobson, T. Bedford, S. Allesina, T. M. Anderson, and M. Pascual. “Spatial guilds in the Serengeti food web revealed by a Bayesian group model”. In: *PLoS Computational Biology* 7.12 (2011), e1002321. DOI: 10.1371/journal.pcbi.1002321.
- [6] C. J. ter Braak and J. A. Vrugt. “Differential evolution Markov chain with snooker updater and fewer chains”. In: *Statistics and Computing* 18.4 (2008), pp. 435–446. DOI: 10.1007/s11222-008-9104-9.
- [7] K. P. Burnham and D. Anderson. *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*. Springer, 2002.
- [8] J. E. Cohen. *Food webs and niche space*. Vol. 11. Princeton University Press, 1978.
- [9] J. E. Cohen, F. Briand, and C. M. Newman. *Community Food Webs: Data and Theory*. Biomathematics. Springer, 1990. ISBN: 0387511296.
- [10] A. Dobson. “Food-web structure and ecosystem services: insights from the Serengeti”. In: *Philosophical Transactions of the Royal Society B* 364.1524 (2009), pp. 1665–1682. DOI: 10.1098/rstb.2008.0287.
- [11] *Encyclopedia of Life API*. URL: <http://www.eol.org/api>.

- [12] S. Geman and D. Geman. “Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 6 (1984), pp. 721–741. DOI: 10.1109/TPAMI.1984.4767596.
- [13] C. Geyer. “Markov chain Monte Carlo maximum likelihood”. In: *Proceedings of the 23rd Symposium on the Interface*. Ed. by E. Keramidas. Computing Science and Statistics. 1991, pp. 156–163.
- [14] P. J. Green and S. Richardson. “Modelling heterogeneity with and without the Dirichlet process”. In: *Scandinavian Journal of Statistics* 28.2 (2001), pp. 355–375. DOI: 10.1111/1467-9469.00242.
- [15] R. Guimera, D. Stouffer, M. Sales-Pardo, E. Leicht, M. Newman, and L. Amaral. “Origin of compartmentalization in food webs”. In: *Ecology* 91.10 (2010), pp. 2941–2951. DOI: 10.1890/09-1175.1.
- [16] W. K. Hastings. “Monte Carlo sampling methods using Markov chains and their applications”. In: *Biometrika* 57.1 (1970), pp. 97–109. DOI: 10.1093/biomet/57.1.97.
- [17] P. W. Holland, K. B. Laskey, and S. Leinhardt. “Stochastic blockmodels: first steps”. In: *Social Networks* 5.2 (1983), pp. 109–137. DOI: 10.1016/0378-8733(83)90021-7.
- [18] C. M. Hurvich and C.-L. Tsai. “Regression and time series model selection in small samples”. In: *Biometrika* 76.2 (1989), pp. 297–307. DOI: 10.1093/biomet/76.2.297.
- [19] B. Karrer and M. E. Newman. “Stochastic blockmodels and community structure in networks”. In: *Physical Review E* 83.1 (2011), p. 016107. DOI: 10.1103/PhysRevE.83.016107.
- [20] R. Kass and A. Raftery. “Bayes factors”. In: *Journal of the American Statistical Association* 90.430 (1995), pp. 773–795. DOI: 10.1080/01621459.1995.10476572.

- [21] A. E. Krause, K. A. Frank, D. M. Mason, R. E. Ulanowicz, and W. W. Taylor. “Compartments revealed in food-web structure”. In: *Nature* 426.6964 (2003), pp. 282–285. DOI: 10.1038/nature02115.
- [22] N. Lartillot and H. Philippe. “Computing Bayes factors using thermodynamic integration”. In: *Systematic Biology* 55.2 (2006), pp. 195–207. DOI: 10.1080/10635150500433722.
- [23] R. M. May. “Will a large complex system be stable?” In: *Nature* 238 (1972), pp. 413–414. DOI: 10.1038/238413a0.
- [24] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. “Equation of state calculations by fast computing machines”. In: *Journal of Chemical Physics* 21.6 (1953), p. 1087. DOI: 10.1063/1.1699114.
- [25] P. Myers, R. Espinosa, C. S. Parr, T. Jones, G. S. Hammond, and T. A. Dewey. *The Animal Diversity Web*. 2013. URL: <http://animaldiversity.org>.
- [26] S. L. Pimm and J. H. Lawton. “Are food webs divided into compartments?” In: *Journal of Animal Ecology* (1980), pp. 879–898. DOI: 10.2307/4233.
- [27] E. L. Rezende, E. M. Albert, M. A. Fortuna, and J. Bascompte. “Compartments in a marine food web associated with phylogeny, body mass, and habitat structure”. In: *Ecology Letters* 12.8 (2009), pp. 779–788. DOI: 10.1111/j.1461-0248.2009.01327.x.
- [28] H. Robbins. *An empirical Bayes approach to statistics*. 1956.
- [29] R. P. Rohr, H. Scherer, P. Kehrl, C. Mazza, and L.-F. Bersier. “Modeling food webs: exploring unexplained structure using latent traits”. In: *The American Naturalist* 176.2 (2010), pp. 170–177. DOI: 10.1086/653667.
- [30] Y. Roskov, T. Kunze, L. Paglinawan, T. Orrell, D. Nicolson, A. Culham, N. Bailly, P. Kirk, T. Bourgoign, G. Baillargeon, F. Hernandez, and A. De Wever. *Species 2000 & ITIS Cata-*

- logue of Life, 2013 Annual Checklist*. URL: <http://www.catalogueoflife.org/annual-checklist/2013/info/cite>.
- [31] A. G. Rossberg, Å. Brännström, and U. Dieckmann. “Food-web structure in low-and high-dimensional trophic niche spaces”. In: *Journal of The Royal Society Interface* 7.53 (2010), pp. 1735–1743. DOI: 10.1098/rsif.2010.0111.
- [32] B. E. Sedio, E. B. Baskerville, M. D. Hunter, S. J. Wright, J. W. Westbrook, J. R. de Wet, and C. R. Dick. “Character displacement of anti-herbivore defenses facilitates the coexistence of closely-related rainforest trees”. In: (*in prep.*) (2013).
- [33] P. P. Staniczenko, M. J. Smith, and S. Allesina. “Selecting food web models: the good, the bad, and the ugly”. In: (*in prep.*) (2013).
- [34] D. Stouffer, J. Camacho, and L. Amaral. “A robust measure of food web intervality”. In: *Proceedings of the National Academy of Sciences* 103.50 (2006), p. 19015. DOI: 10.1073/pnas.0603844103.
- [35] J. H. Ward Jr. “Hierarchical grouping to optimize an objective function”. In: *Journal of the American Statistical Association* 58.301 (1963), pp. 236–244. DOI: 10.1080/01621459.1963.10500845.
- [36] P. Warren and J. Lawton. “Invertebrate predator-prey body size relationships: an explanation for upper triangular food webs and patterns in food web structure?” In: *Oecologia* 74.2 (1987), pp. 231–235. DOI: 10.1007/BF00379364.
- [37] R. J. Williams, A. Anandanadesan, and D. Purves. “The probabilistic niche model reveals the niche structure and role of body size in a complex food web”. In: *PLoS One* 5.8 (2010), e12092. DOI: 10.1371/journal.pone.0012092.g001.
- [38] R. J. Williams and N. D. Martinez. “Simple rules yield complex food webs”. In: *Nature* 404.6774 (2000), pp. 180–183. DOI: 10.1038/35004572.

- [39] R. J. Williams and D. W. Purves. “The probabilistic niche model reveals substantial variation in the niche structure of empirical food webs”. In: *Ecology* 92.9 (2011), pp. 1849–1857. DOI: 10.1890/11-0200.1.

Chapter 5

Food-Web Robustness and the Dual Nature of Compartments

5.1 Introduction

A food-web network provides a partial description of the interdependencies in an ecosystem. The network circumscribes the pathways for the flow of matter and energy via feeding relationships, and, from the perspective of a single species in the web, delineates all the food chains that support it. Nature contains complex food webs that persist over time with many species and many feeding relationships, and the desire to understand the persistence of these webs has inspired decades of speculation and theory. However, their persistence is not guaranteed, and the disproportionate ability for human activity to cause extinctions as compared with the rest of the natural world makes it especially important for us to understand the properties of these systems. Toward this end, ecologists have come up with a number of approaches to try to investigate how complex food webs persist.

Food webs have been studied with respect to many different measures of how systems responds to perturbations. Stability and resilience are often used as general terms, but also refer to specific definitions in terms of local asymptotic analysis of the equilibrium of a dynamical system [14, 17]. Persistence typically refers to the continued long-term existence of species in a system, and does not depend upon a particular mathematical formulation of dynamics. Robustness has also been used in various ways, and the term overlaps in part with stability, resilience, and persistence, but in studies of ecological networks, robustness usually refers to the ability of a network to withstand the

secondary extinction of species in response to an initial loss of species [8]. That definition is used here.

An important line of inquiry has focused on how the topology of food-web networks affects these various measures, a dominant theme of which has been the relationship between compartmentalization in food webs and various measures of stability. The presence of compartments, groups of species with higher internal connectivity within groups than between groups, has long been noted in food webs, particularly as resulting from habitat boundaries [22]. More recently, undirected network modularity has been used to identify the presence of significant compartmentalization in food webs as compared random-graph null models [13, 23].

The notion that compartmentalization plays an important role in stabilizing food webs has emerged from a number of studies. Early work by Robert May, in which randomly generated matrices were interpreted as linearizations of dynamical systems, showed that matrices with higher connectance within diagonal blocks were more likely to be linearly stable, which has been construed as evidence for the importance of compartmentalization [15], although it has been recognized that the original formulation poorly relates to predator-prey networks [4]. Thébault and Fontaine studied randomly generated bipartite antagonistic networks, finding that a measure of dynamical resilience was higher in more modular networks, but that in fact species persistence decreased slightly with modularity [28]. Most recently, Stouffer and colleagues identified a positive statistical relationship between undirected network modularity in empirical food-web networks and the persistence of species in dynamical simulations based on the networks when controlling for the number of links in the food web but nothing else [27].

However, definitions of compartmentalization vary somewhat across studies, making comparisons difficult, and most standard compartmentalization measures are inherently limited in their ability to capture food-web structure. The most widely used measure of compartmentalization, network modularity [9], is calculated as the deviation in the number of within-module links as compared to a null model, typically a random graph, normalized by the maximum possible deviation.

This definition, which ignores the direction of feeding links, was used in the work by Stouffer and colleagues [27]. A modification that considers links only from the prey perspective or only from the predator perspective has also been used [10], but not in a direct study of dynamical effects. Network modularity can also be straightforwardly adapted to bipartite networks, as in the work by Thébault and Fontaine [28], so that species are *a priori* assigned to the appropriate trophic level, and only links between these levels are considered. Both these approaches, however, are limited by maximizing the number of links within groups, and groups will necessarily include species at two or more trophic levels. By instead maximizing the number of links either within *or* between groups, as in the group model of Allesina and Pascual [2], differing numbers of groups may be found at different trophic levels. This approach demonstrates that more natural groupings of species are not strictly compartmentalized across trophic levels, but may instead, for example, include multiple groups of species at lower trophic levels coupled by smaller numbers of groups at higher trophic levels, as in the Serengeti web studied in Chapters 3 and 4.

The results from Chapter 4 show that both flexible group structure and contiguous diets in a low-dimensional niche space are important for describing food-web structure. Also, general, niche-based models are much better descriptors of food-web structure than compartmentalized structures according to standard modularity metrics [25]. Furthermore, the niche model of Williams and Martinez [29] generates food webs with high network modularity according to the standard definition, as well as from prey-focused and predator-focused views [10], indicating that much of the “compartmentalization” detected in food webs may be largely due to the underlying niche space. Since the standard definition of compartmentalization describes a small part of empirical food-web structure when defined as significant modularity relative to a random graph model, it likely is also a relatively poor indicator of the stability and robustness properties of empirical food webs.

These results invite two directions for further investigations. First, we should verify that better-fitting models of structure in fact do a better job at capturing the robustness properties of empirical webs. Second, we should use the structures contained in good models as a basis for further

exploration.

Group-based models provide an aggregated approximation to the overall structure of a food web in terms of functional trophic groups, and perhaps therefore serve as a more useful, flexible way of thinking about the effects of compartmentalization. Therefore, one useful line of inquiry is to systematically investigate the robustness properties of various trophic group structures. In the Serengeti web, for example, four main plant groups were linked differentially to two main herbivore groups, which were in turn linked to the carnivores, forming a familiar pyramid structure among statistical aggregations of species. This structure mirrors a coupled-chain pattern that seems to be more stable than separated chains [16], but at the scale of many species.

The use of dynamical models when investigating the consequences of different structures on robustness is appealing but challenging. Studies based on fully specified dynamics have the clear advantage of considering dynamical cascades that include both bottom-up and top-down effects, the importance of which is abundantly clear from real-world examples [20]. However, particularly when divorced from testable data from a particular system, this approach requires many assumptions about the details of the dynamics. A simpler structure-only approach, in which only bottom-up extinctions are considered, cannot predict top-down cascades, but also cannot predict top-down cascades incorrectly. Structure-only approaches thus serve as a useful starting point by revealing a reasonably confident subset of the fragilities of a food web.

The purpose of this study is to pursue two simple investigations into food-web structure and robustness: first, to test whether better-fitting food-web models better capture the robustness of empirical food webs, and second, to examine how simple trophic group structures differ with respect to robustness, as measured by the ability of a food web to withstand bottom-up extinctions caused by the targeted loss of prey species.

5.2 Methods

We performed two experiments, one testing the ability of food-web models to capture robustness, and the other comparing the effect of different trophic structures on robustness. For both experiments, we measured robustness as the maximum extinction area, a quantity that measures the effect of the secondary extinction pattern caused by an extinction sequence (Section 5.2.1).

In the first experiment, we used fits from several different food-web models, each of which encodes a different hypothesis about food-web structure, to generate random food webs. We compared how the goodness of fit of models related to how close the extinction area of generated webs was to the extinction area of the empirical web. The details of the models and food webs are described in Section 5.2.2.

In the second experiment, we generated webs from a variety of trophic group structures (Section 5.2.3), with the number of species and links held constant but with links differentially allocated, and compared the generated distribution of extinction areas. We also calculated a measure of the degree of group organization, *tritrophic groupiness*, for each structure (Section 5.2.4). Finally, we repeated the experiment with niche-like intervality introduced within the group structure, to see how robustness changed.

5.2.1 Maximum extinction area

When a species is removed from a web, the removal may generate secondary extinctions. Considering only bottom-up extinctions, which unlike top-down extinctions can be directly inferred from food-web structure, any non-basal species that lose all prey species as a result of the extinction will also go extinct. This process can be visualized using an extinction curve, which shows the cumulative total number of extinctions as a function of the cumulative number of species removed. The area under the curve is a measure of effectiveness of the extinction sequence in damaging the web, which has been termed the extinction area [3]. If no secondary extinctions occur, the curve follows

a 1:1 line, and the extinction area is 0.5; if most of the species go extinct as a result of a few primary extinctions, the extinction area is close to 1. The maximum extinction area out of all extinction sequences is a measure of the robustness of the web.

In order to efficiently identify a damaging sequence of extinctions, an eigenvector-based algorithm can be used to identify the relative importance of species, under the recursive definition that important species serve as prey for other important species [3]. This algorithm is analogous to Google PageRank [21], which assumes that important web pages are linked to by important pages. This algorithm does a better job at identifying near-worst extinctions sequences than other approximation algorithms for many empirical food webs, and is computationally much cheaper than, e.g., a genetic algorithm-based search for the worst sequence.

For group-based structures, the fact that species within the same group are statistically identical can additionally be used to identify the sequence of groups expected to cause the maximum extinction area. With a small enough number of groups, all possible permutations of group removal sequences can be explored. When a group is removed, if the expected number of links for a species in a remaining group is less than 1, that group is also removed as a secondary extinction, and counts toward the extinction area calculation. A “partial extinction area” can also be calculated, where a group that loses prey species is considered to be partially extinct, with the extinction fraction equal to the fraction of links missing out of the original total.

5.2.2 Models and data

This analysis is based on model fits provided by the authors of work of normalized maximum likelihood (NML) for food-web model selection [25]. NML is defined as the maximum likelihood for a data set divided by the maximum likelihood integrated over all possible data sets of the same size, and emerges as the appropriate quantity for model selection when model quality is defined according to the minimum description length (MDL) principle. Under MDL a model is good if it provides an efficient encoding for transmitting the data, thus effectively capturing regularities. This

approach avoids the assumption of an underlying truth as in AIC-based model selection, and the specification of prior distributions for parameters as in the Bayesian framework.

The authors provided fits for six different food webs (Table 5.1) under eight different models (Table 5.2). The best fits for each model type was used for comparison.

Source	S	L	C
Kongsfjorden [12]	252	1124	0.018
Lough Hyne [24]	326	4262	0.040
Reef [19]	210	2065	0.047
St. Marks [5]	116	1128	0.084
Weddell Sea [11]	381	10182	0.070
Ythan Estuary [7]	77	307	0.052

Table 5.1: Food webs used for robustness comparison. S is the number of species in the web; L is the number of links; and $C = L/S^2$ is the connectance.

In order to compare the effectiveness of models at capturing the structural robustness of a web, we used each fitted model to generate 500 replicate webs. We measured the robustness of the original web and of the generated webs as the estimate of the maximum extinction area 5.2.1 produced by the

5.2.3 Robustness of different group structures to bottom-up extinctions

In order to examine differences in robustness between different large-scale trophic group structures, we generated tritrophic webs with the same number of species and links, but with species organized into different group structures and links allocated differentially among pairs of groups. We held the number of species held constant at 32 top basal, 64 middle, and 128 basal species, and the number of links constant at 256 between the top and middle levels and 256 between the middle and basal levels, so that $\frac{1}{8} = 0.125$ of the possible directed links between each pair of levels are realized.

We considered 10 different divisions of species into groups and various patterns of link allocation for each division into groups, for a total of 31 different link allocations (Figure 5.1). These included all possible divisions of each trophic level into one or two groups ($2^3 = 8$ different combinations) as

Name	Description
Random	Random model. All links have equal probability p .
Cascade	Cascade model [6]. Species are ordered, and species consume prey with lower order with probability p , and consume prey with higher order with probability q , so that the matrix is essentially an upper-triangular half of a random graph.
Modular	Modular or anti-modular model [25]. Species are organized into groups, and the probability of consuming a species within one's group is p , and the probability of consuming a species outside the group is q .
Hybrid	Hybrid modular/cascade model [25]. Species are organized into groups and are ordered within groups, so that cascade structure is enforced within groups with probabilities p for lower-order species and q for higher-order species, and between groups the probability of a link is r .
Group	Group model [1]. Species are organized into groups, and probabilities p_{ij} are assigned between pairs of groups, so that species in group j have probability p_{ij} of feeding on species in group i .
NicheMP	Minimum potential niche model [2]. Species are ordered, and feed within a contiguous interval with probability p_i and outside with probability $q_i = 0$, so that all links must be contained in the diet.
Niche2	Niche2 model [25]. Species are ordered, and like NicheMP feed within a contiguous interval with probability p_i , but can also feed outside the interval with probability $q_i \neq 0$.
Niche3	Niche3 model [25] Species are ordered and assigned three probabilities: the probability of feeding within an interval, p_i ; the probability of feeding above the interval, q_i ; and the probability of feeding below the interval, r_i .

Table 5.2: Food-web models used for robustness comparison.

well as structures with 1, 2, 4 and 4, 4, 4 groups at the top, middle, and basal levels, respectively. In order for the group structure to be meaningful, link structure was necessarily asymmetrical, so that either some groups were allocated more links than others, some pairs of groups were more highly connected than other pairs, or both.

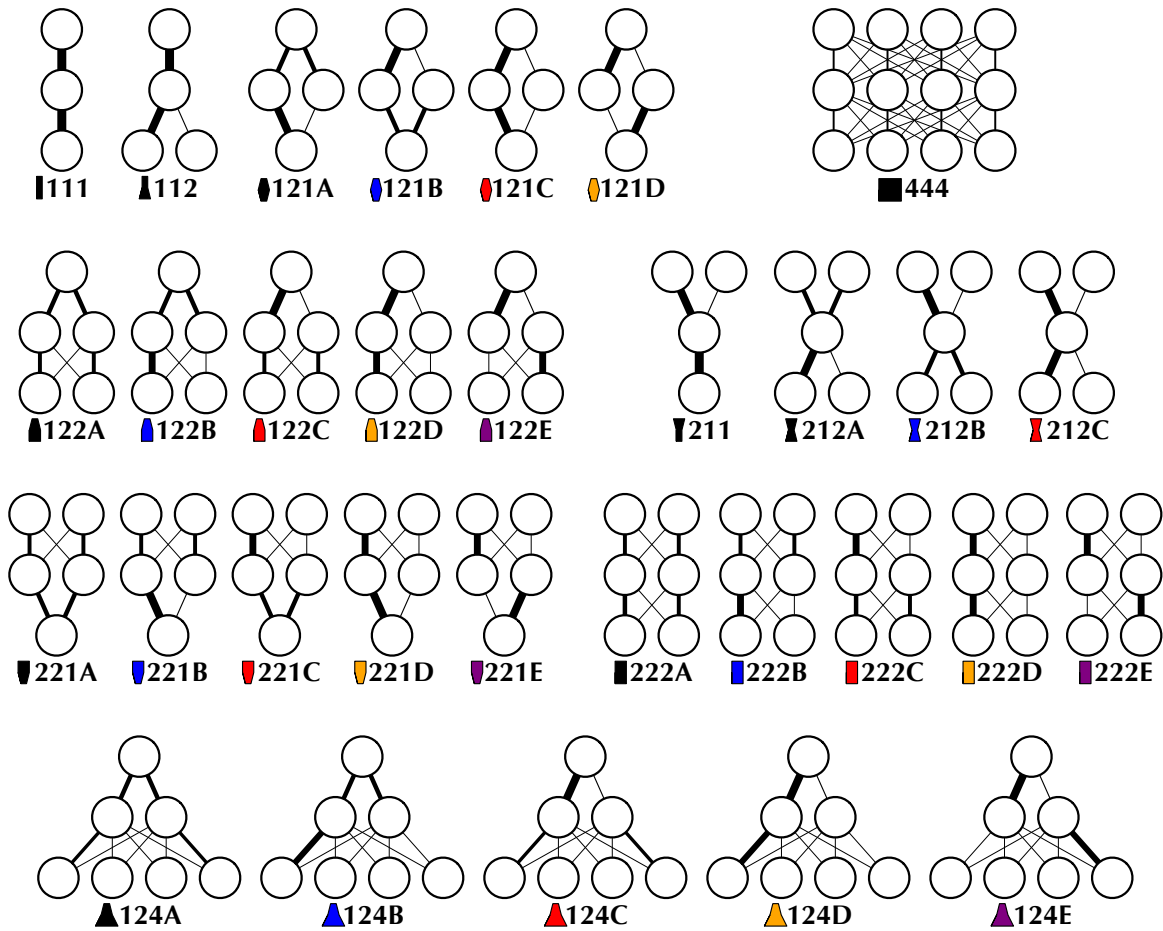


Figure 5.1: Group structures for robustness experiment. Circles represents groups of species, and the relative thickness of edges represents the fraction of links between two trophic levels assigned to the connected pair of groups. All link allocation asymmetries are 7:1, as described in the text.

Since some patterns of link asymmetry are mirror images, and thus topologically equivalent, or isomorphic, for the patterns with one or two groups at each level we considered all non-isomorphic

combinations of asymmetry patterns, with all links being allocated in either a 1:1 or 7:1 ratio between groups or pairs of groups.

For the structure 124, link allocation between the two middle and one top group is either 1:1 or 7:1. Under the structure with left-right symmetry between the basal and middle levels, link allocation is as follows: 392 links to the strongest connection; 56 links for the weak connection on the same side; and 32 links to the connections on opposite sides. With asymmetry between the basal and middle levels, one of the middle groups is assigned 7/8 of the links: 686 allocated to the strong connection; 98 to the weak connection on the same side; and 56 to each group on the other side. The other middle group is given 98, 14, 8, and 8 links, respectively.

For the structure 444, within compartments 224 links are allocated between the basal and middle levels; 56 links are allocated between the middle and top levels; and remaining links are distributed evenly to groups outside the compartment.

Each group structure thus specifies the division of each trophic level into 1, 2, or 4 groups as well as the number of links L_{gh} allocated to each pair of groups g, h between adjacent trophic levels. In order to compare the robustness of different group structures, we generated webs satisfying these constraints as well as ensuring that all basal species have predators and all middle and top species have prey by assigning links uniformly randomly according to this algorithm:

1. For each pair of groups g, h at adjacent trophic levels, uniformly randomly select a prey species from group g and a predator species from group h for each of L_{gh} links assigned to this pair of groups.
2. For each basal species with no predators, choose a link uniformly randomly from the links connected to species in the same group that have more than one predator. Reassign the chosen link to this species.
3. For each middle and top species with no prey, choose a link uniformly randomly from among species in the same group that have more than one prey. Reassign the chosen link to this

species.

For each generated web, we calculated the extinction area generated by the eigenvector ranking algorithm and by the group-based algorithm (Section 5.2.1), and used the maximum of the two as an estimate of the maximal extinction area.

Finally, we repeated the entire experiment, but with niche model-like intervality introduced into each generated web. After step 1 in the generation algorithm, the diets of each species is made to occupy an interval region for each of its prey groups, but to minimize the modification and preserve the overall group structure, intervality is not enforced across groups.

5.2.4 Tritrophic groupiness

A measure of the group structure of a network called groupiness has recently been proposed [26]. Groupiness is analogous to modularity [18], but is based on the number of links between pairs of groups instead of links within groups. As with modularity, groupiness is measured relative to a null model. Groupiness is defined as the sum across all directed pairs of groups of the absolute difference between the actual and expected number of links, divided by the maximum possible value of this quantity, so that it varies from 0 to 1.

For the tritrophic structures considered here, the appropriate null model is two random bipartite networks: one between the basal and middle levels, and one between the middle and top. The formula for tritrophic groupiness is

$$G_{TT} = \frac{1}{2 \sum_{t=1}^2 L_{t,t+1}} \sum_{t=1}^2 \sum_{i=1}^{n_t} \sum_{j=1}^{n_{t+1}} \left| L_{ij} - \frac{n_i}{n_t} \frac{n_j}{n_{t+1}} L_{t,t+1} \right|, \quad (5.1)$$

where t is the lower trophic level in the pair being considered, $L_{t,t+1}$ is the total number of links between the two trophic levels, n_i is the number of species in group i , n_t is the number of species at trophic level t , and L_{ij} is the number of links between groups i and j .

5.3 Results

5.3.1 Model fits and robustness

In general, the extinction area of generated webs was much closer to the extinction area of the original web for better-fitting webs (Figure 5.2). In all cases, the best-fitting web was one of the niche-based models, and the closest match to the extinction area was also a niche-based model, but these were not necessarily the same. Cascade, hybrid cascade-group, modular, and random models gave very poor fits; the group model was somewhere in between.

The random, modular, and group-based models predicted substantially higher robustness than the original web (Figure 5.3). The niche-based models also overpredicted robustness, but only slightly; cascade and hybrid models were inconsistent in this regard.

5.3.2 Group structure and robustness

A plot of the distribution of extinction area for all structures is shown in Figure 5.4. In general, robustness to bottom-up extinctions is lower when species at higher trophic levels are dependent on a smaller number of species at lower trophic levels. The asymmetric allocation of links to particular groups produces this effect, as does compartmentalization, even when compartments have the same number of links per species.

The most robust structures, clustered to the left of Figure 5.4, have a single group at the basal level. The one other structure in that cluster with two basal groups (212B) in fact has effectively only one, since there is only one middle group and link allocation is symmetric between the basal and middle levels. (In fact, it should have been excluded from the experiment, since the group division is meaningless.)

The least robust structure, is 124E, characterized by a top group that is strongly linked to a middle group with few supporting basal links. Among the next four least robust structures include 124B and 124D, which similarly have a strong dependence of the top group upon a small number

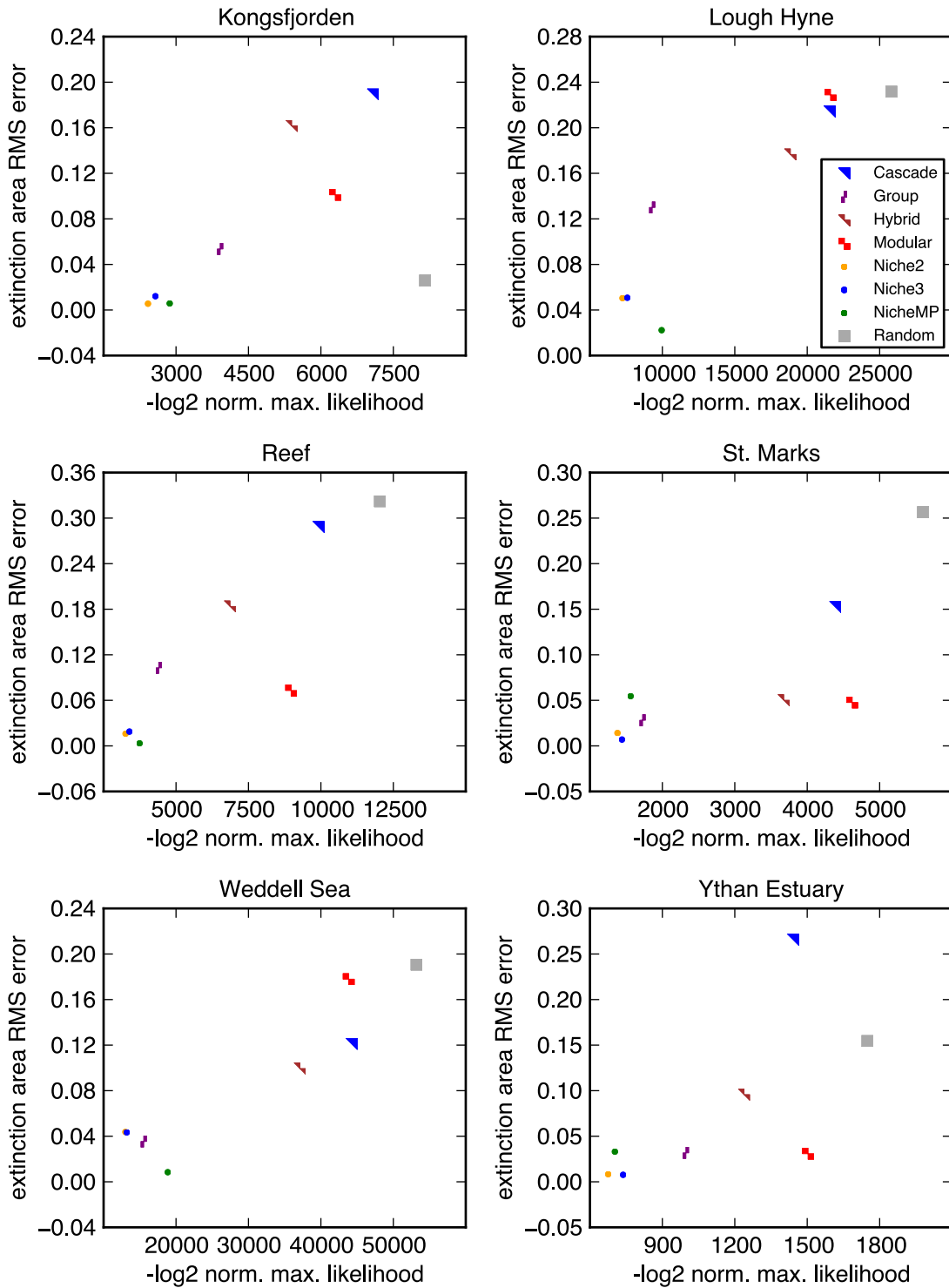


Figure 5.2: Root-mean-square error of predicted extinction area against model fit for each food web. In all cases, both the best-fitting model and the model with lowest error is one of the niche models, but not necessarily the same one.

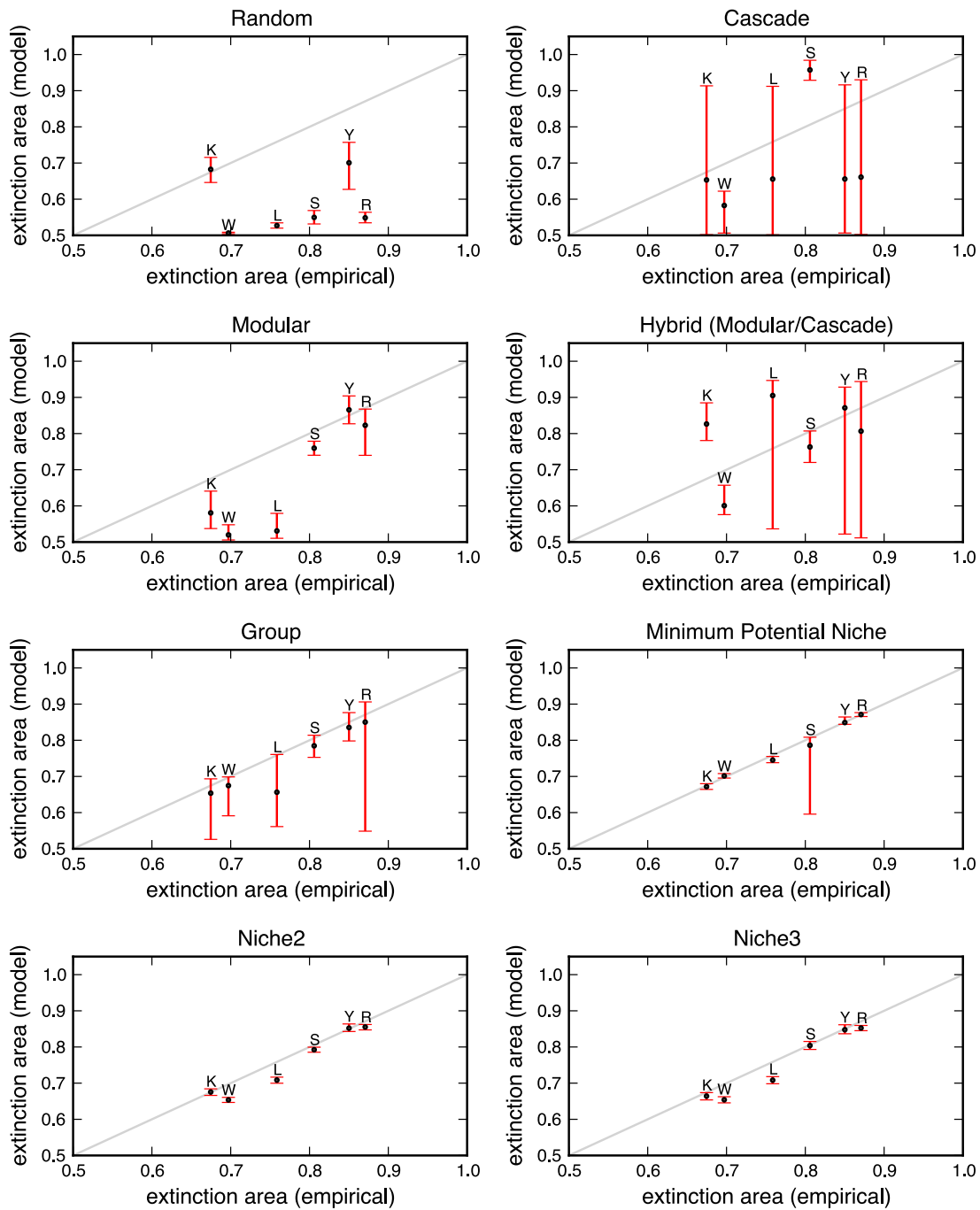


Figure 5.3: Predicted vs. empirical extinction area for each model. Means and 90% intervals of generated distributions are shown. Food webs are labeled by the first letter of their name.

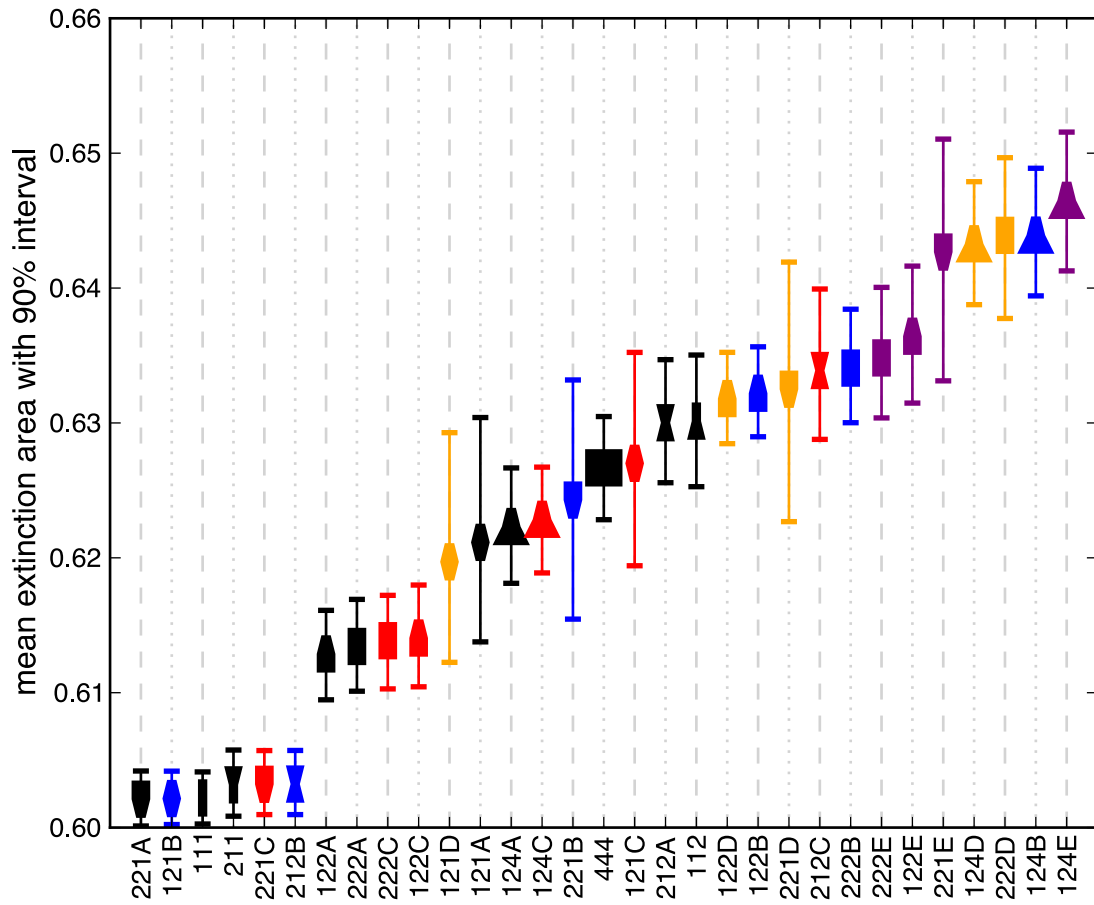


Figure 5.4: Extinction area by group structure. Means and 90% intervals are shown for each group structure, in sorted order by mean. Labels correspond to the structures shown in Figure 5.1.

of basal species, since all the asymmetric 124 structures have a higher degree of basal asymmetry than the 2-group structures. Also in the top 5 are 221E, which has a similar strong-weak linkage as 124E, and 222D, which has one robust chain and one highly fragile one.

The structures 111, 222A, and 444 form a simple test of the effect of compartmentalization. In all three, compartments are equivalent to each other; no group is allocated more links than any other group. However, 222A is less robust than 111, since its top species depend more on one group of basal species than another; and 444 is less robust than 222A.

The structure 124A is similarly more robust than 444, since species at higher trophic levels depend on a wider range of species at lower trophic levels.

The structures with the lowest tritrophic groupiness were the most robust, but many structures with differing extinction area had identical groupiness (Figure 5.5), so this metric alone does not adequately capture the robustness of the web.

5.3.3 Intervalsity and robustness

The addition of interval niche structure within pairs of groups decreased robustness in all cases, as shown in Figure 5.6. Structures 111 and 221C, most notably, went from being two of the most robust to being two of the least robust; conversely, structure 124E hardly changed.

5.4 Discussion

This chapter includes two intentionally simple investigations into the relationship between structure and robustness. In the first, we find that, generally speaking, better-fitting models of food web structure, unsurprisingly, do a better job at capturing the structural robustness of empirical food webs. Among these webs, niche-based models best captured the structure, and group-based models did a reasonable job, although as shown in Chapter 4, improved group models and combined group-niche models may perform better for some webs. In the second investigation, we looked at the robustness of simple tritrophic group structures by varying only the asymmetric allocation of links

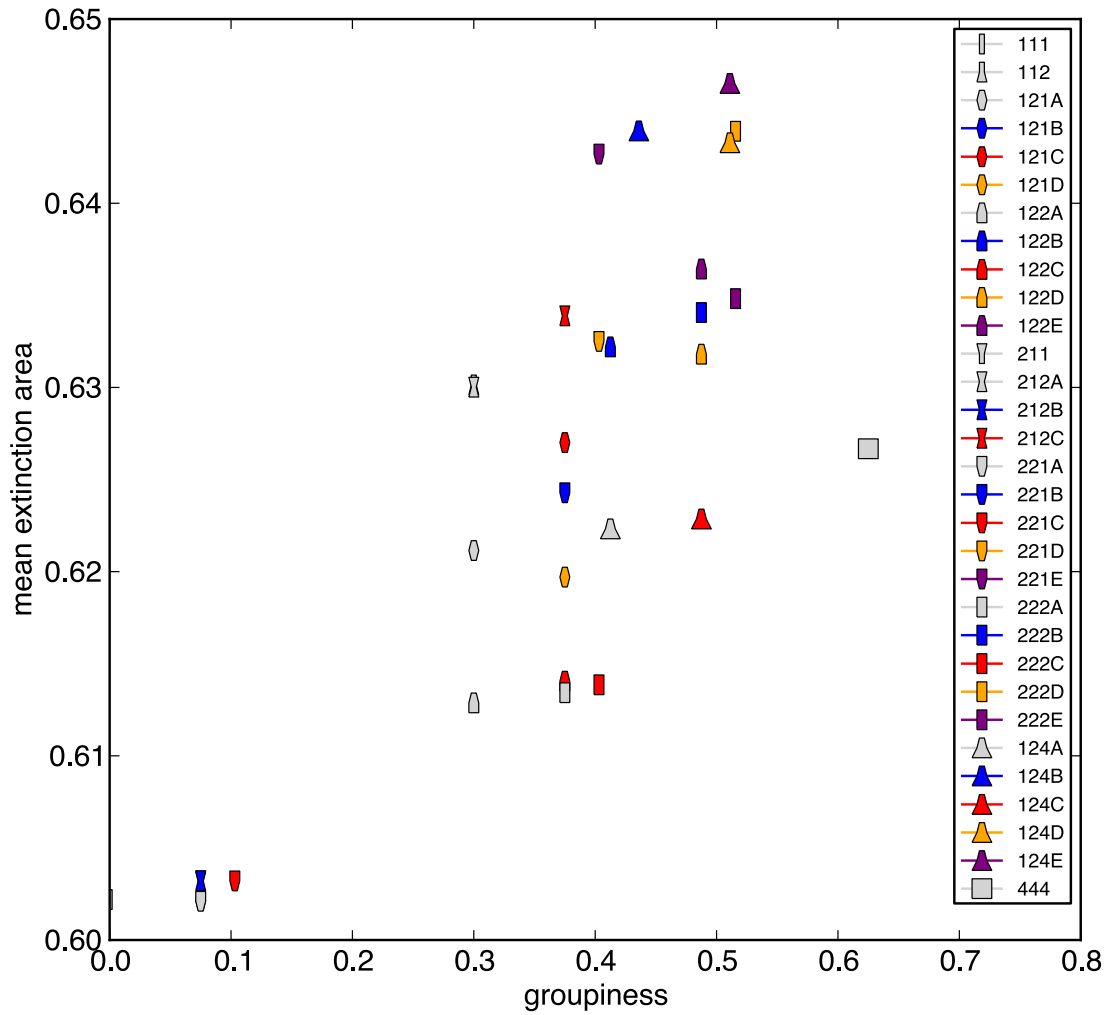


Figure 5.5: Mean extinction area vs. groupiness of different model structures. The least “groupy” structures are the most robust, and the least robust are among the most groupy, but several groups of structures with identical groupiness vary in robustness.

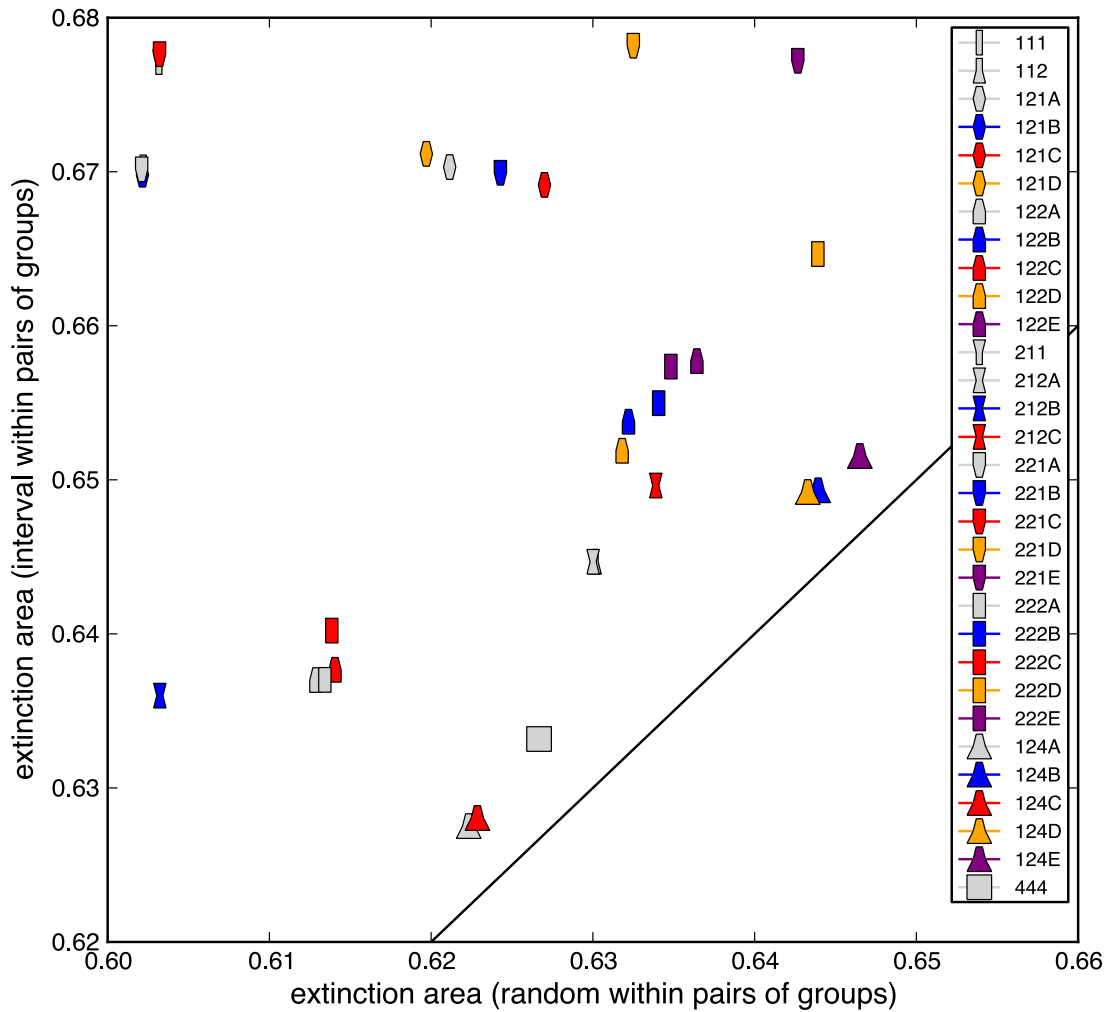


Figure 5.6: Effect of intervality on robustness. Mean extinction area with interval structure introduced within pairs of groups is plotted against mean extinction area without interval structure. A diagonal 1:1 line is shown, indicating that extinction area increased for all group structures when intervality was added.

to different pairs of groups at adjacent trophic levels, and found that compartmentalized structures are less robust than non-compartmentalized structures. The key point is that, where bottom-up extinctions are concerned, what matters is the presence of species at higher trophic levels that are dependent on a small number of species at lower trophic levels.

Past discussions of compartmentalization in food webs have been hindered both by the inconsistent use of terminology and by the use of mathematical tools that are fundamentally mismatched to food webs. Both these problems have led to the emergence of a conventional wisdom that holds, firstly, that food webs are highly compartmentalized, secondly, that compartments stabilize food webs, and thirdly, that the compartmentalization detected in food webs stabilizes them. All three of these statements may be true to a certain extent. Formulated vaguely, the notion that there are collections of species in food webs that are more highly connected than other collections seems true. Group-based models reveal trophic guilds of species, and pairs or tritrophic chains of these guilds may be highly connected. Niche-based models also produce these kinds of compartmentalized dependency chains [10]. It may further be true that, the results presented here notwithstanding, such structures aid dynamical stability, in which case that fact may be important for their stability in nature. However, the use of standard measures of network modularity tested against inappropriate null models reveals relatively little about food-web structure, as evidenced by the poor performance of the modular models, and should therefore be avoided when investigating these questions [25]. The investigation into the robustness of trophic group structures presented here is one attempt at a more careful approach.

An interesting tension arises from thinking about robustness to dynamical perturbations as compared with bottom-up extinctions alone. An intuitive basis for the notion that compartmentalization stabilizes food webs is that perturbations are more likely to be contained within a compartment, and some support for this notion exists in dynamical studies [15, 27]. This intuition translates to a caricatured trophic structure consisting of tritrophic chains of highly connected groups of species with relatively few links between groups in different chains: perturbations within one vertical block of

species should be less likely to spread to adjacent blocks than if the chain boundaries were erased and species were equally likely to link to any species at the adjacent trophic level.

However, for bottom-up extinctions, the argument goes the other way: if you select a random top species and follow it down the web to the set of basal species it depends upon, the more compartmentalized the links, the fewer species, and thus the more susceptible the top species is to bottom-up extinction. That basic idea explains why compartmentalized structures are less robust to bottom-up extinction cascades. This does not contradict dynamical results on the stabilizing nature of compartments, but it means that we need to investigate this question with an eye toward disentangling opposing bottom-up and top-down effects on robustness.

Interestingly, group-based and especially modular models tend to predict higher food-web robustness than the empirical web being fit. A possible explanation for this pattern relates to the robustness argument just presented. In niche-based models, which fit the data very well, the extremely high trophic similarity of species feeding in the same region of niche space results in compartmentalization of links [10]. Group-based models can successfully fit clusters of species with high trophic similarity, but the constraints of a niche-based model seem to induce even higher similarity among adjacent species, and the links are thus even more compartmentalized in this sense than group-based models. As a result, it makes sense that niche-structured webs would be more susceptible to bottom-up extinction cascades. Adding intervality to simple group structures reduced robustness, confirming this notion: in fact, the least compartmentalized group structure became one of the least robust structures once intervality was added, since interval overlaps were less constrained by group boundaries.

An understanding of the relationship between topological food-web patterns and robustness remains challenging, but we now have a much better grasp of how best to characterize and describe those patterns. Group structures are present at a broad scale, but may in large part result from clustering in a constrained niche space. Here, we have shown that better food-web models better capture robustness properties. We have also demonstrated how asymmetric link allocation and

compartmentalization in trophically structured webs can reduce robustness to bottom-up extinctions, leaving open the important question of how opposing bottom-up and dynamical top-down cascades interact in real sequences of extinctions. We leave this for future work, along with a more difficult and fundamental question: how does robustness emerge from, and feed back upon, patterns of structure and dynamics?

References

- [1] S. Allesina, D. Alonso, and M. Pascual. “A general model for food web structure”. In: *Science* 320.5876 (2008), pp. 658–661. DOI: 10.1126/science.1156269.
- [2] S. Allesina and M. Pascual. “Food web models: a plea for groups”. In: *Ecology Letters* 12.7 (2009), pp. 652–662. DOI: 10.1111/j.1461-0248.2009.01321.x.
- [3] S. Allesina and M. Pascual. “Googling food webs: can an eigenvector measure species’ importance for coextinctions?” In: *PLoS Computational Biology* 5.9 (2009), e1000494. DOI: 10.1371/journal.pcbi.1000494.
- [4] S. Allesina and S. Tang. “Stability criteria for complex ecosystems”. In: *Nature* 483.7388 (2012), pp. 205–208. DOI: 10.1038/nature10832.
- [5] R. R. Christian and J. J. Luczkovich. “Organizing and understanding a winter’s seagrass foodweb network through effective trophic levels”. In: *Ecological Modelling* 117.1 (1999), pp. 99–124. DOI: 10.1016/S0304-3800(99)00022-8.
- [6] J. E. Cohen, F. Briand, and C. M. Newman. *Community Food Webs: Data and Theory*. Biomathematics. Springer, 1990. ISBN: 0387511296.
- [7] J. E. Cohen, D. N. Schittler, D. G. Raffaelli, and D. C. Reuman. “Food webs are more than the sum of their tritrophic parts”. In: *Proceedings of the National Academy of Sciences* 106.52 (2009), pp. 22335–22340. DOI: 10.1073/pnas.0910582106.

- [8] J. A. Dunne, R. J. Williams, and N. D. Martinez. “Network structure and biodiversity loss in food webs: robustness increases with connectance”. In: *Ecology Letters* 5.4 (2002), pp. 558–567. DOI: 10.1046/j.1461-0248.2002.00354.x.
- [9] M. Girvan and M. Newman. “Community structure in social and biological networks”. In: *Proceedings of the National Academy of Sciences* 99.12 (2002), p. 7821. DOI: 10.1073/pnas.122653799.
- [10] R. Guimera, D. Stouffer, M. Sales-Pardo, E. Leicht, M. Newman, and L. Amaral. “Origin of compartmentalization in food webs”. In: *Ecology* 91.10 (2010), pp. 2941–2951. DOI: 10.1890/09-1175.1.
- [11] U. Jacob. “Trophic dynamics of Antarctic shelf ecosystems: food webs and energy flow budgets”. PhD thesis. Bremen, Univ., Diss., 2005, 2005.
- [12] U. Jacob, A. Thierry, U. Brose, W. E. Arntz, S. Berg, T. Brey, I. Fetzer, T. Jonsson, K. Mintenbeck, and C. Mollmann. “The Role of Body Size in Complex Food Webs: A Cold Case”. In: *The Role of Body Size in Multispecies Systems*. Vol. 45. Advances In Ecological Research. Amsterdam: Elsevier, 2011, pp. 181–223.
- [13] A. E. Krause, K. A. Frank, D. M. Mason, R. E. Ulanowicz, and W. W. Taylor. “Compartments revealed in food-web structure”. In: *Nature* 426.6964 (2003), pp. 282–285. DOI: 10.1038/nature02115.
- [14] R. M. May. *Stability and Complexity in Model Ecosystems*. Princeton University Press, 1973.
- [15] R. M. May. “Will a large complex system be stable?” In: *Nature* 238 (1972), pp. 413–414. DOI: 10.1038/238413a0.
- [16] K. McCann, J. Rasmussen, and J. Umbanhowar. “The dynamics of spatially coupled food webs”. In: *Ecology Letters* 8.5 (2005), pp. 513–523. DOI: 10.1111/j.1461-0248.2005.00742.x.

- [17] M. G. Neubert and H. Caswell. “Alternatives to resilience for measuring the responses of ecological systems to perturbations”. In: *Ecology* 78.3 (1997), pp. 653–665. DOI: 10.1890/0012-9658(1997)078[0653:ATRFMT]2.0.CO;2.
- [18] M. Newman. “Modularity and community structure in networks”. In: *Proceedings of the National Academy of Sciences* 103.23 (2006), p. 8577. DOI: 10.1073/pnas.0601602103.
- [19] S. C. Opitz. *Trophic interactions in Caribbean coral reefs*. 1085. The WorldFish Center, 1996.
- [20] M. L. Pace, J. J. Cole, S. R. Carpenter, and J. F. Kitchell. “Trophic cascades revealed in diverse ecosystems”. In: *Trends in Ecology & Evolution* 14.12 (1999), pp. 483–488. DOI: 10.1016/S0169-5347(99)01723-1.
- [21] L. Page, S. Brin, R. Motwani, and T. Winograd. *The PageRank Citation Ranking: Bringing Order to the Web*. Technical Report 1999-66. Previous number = SIDL-WP-1999-0120. Stanford InfoLab, Nov. 1999. URL: <http://ilpubs.stanford.edu:8090/422/>.
- [22] S. L. Pimm and J. H. Lawton. “Are food webs divided into compartments?” In: *Journal of Animal Ecology* (1980), pp. 879–898. DOI: 10.2307/4233.
- [23] E. L. Rezende, E. M. Albert, M. A. Fortuna, and J. Bascompte. “Compartments in a marine food web associated with phylogeny, body mass, and habitat structure”. In: *Ecology Letters* 12.8 (2009), pp. 779–788. DOI: 10.1111/j.1461-0248.2009.01327.x.
- [24] J. O. Riede, B. C. Rall, C. Banasek-Richter, S. A. Navarrete, E. A. Wieters, M. C. Emmerson, U. Jacob, and U. Brose. “Scaling of Food-Web Properties with Diversity and Complexity Across Ecosystems”. In: *Ecological Networks*. Vol. 42. Advances In Ecological Research. Amsterdam: Elsevier, 2010, pp. 139–170.
- [25] P. P. Staniczenko, M. J. Smith, and S. Allesina. “Selecting food web models: the good, the bad, and the ugly”. In: (*in prep.*) (2013).

- [26] P. P. Staniczenko, M. J. Smith, and S. Allesina. “Significance of modularity in food webs”. In: *(in prep.)* (2013).
- [27] D. B. Stouffer and J. Bascompte. “Compartmentalization increases food-web persistence”. In: *Proceedings of the National Academy of Sciences* 108.9 (2011), pp. 3648–3652. DOI: 10.1073/pnas.1014353108.
- [28] E. Thébault and C. Fontaine. “Stability of ecological communities and the architecture of mutualistic and trophic networks”. In: *Science* 329.5993 (2010), p. 853. DOI: 10.1126/science.1188321.
- [29] R. J. Williams and N. D. Martinez. “Simple rules yield complex food webs”. In: *Nature* 404.6774 (2000), pp. 180–183. DOI: 10.1038/35004572.

Chapter 6

Conclusion

This dissertation examined the importance of groups and niches to food-web structure, developed methodology for identifying these patterns, linked these patterns to ecological and evolutionary relationships among species, and finally explored how these patterns relate to one aspect of network robustness. The Bayesian and empirical Bayes methods described in Chapter 2 enabled the modeling work of Chapters 3 and 4, which identified both large-scale group structure and a low-dimensional niche space as important components of the structure of the Serengeti food web. These patterns were linked to habitat structure and evolutionary history, a proxy for biological traits. Chapter 5 demonstrated how, when only bottom-up extinctions are considered, compartmentalized structures tend to reduce network robustness, and, consequently, that structures like those seen in the Serengeti, with groups higher trophic levels coupling groups at lower trophic levels, are more robust.

This chapter discusses directions for food-web structural models and methodology (Section 6.1); how the coupled-group pattern seen in the Serengeti relates to robustness and work on spatial coupling (Section 6.2); and, finally, broader directions for food-web studies (Section 6.3).

6.1 Models of food-web structure: where to go from here

Over the last several years, improvements in methodology have revised, clarified, and strengthened our understanding of the structure of food webs. The encoding of hypotheses about food-web structure into probabilistic models that can be fit to data and compared on a more objective basis

tremendously increased our ability to separate conventional wisdom from useful pattern. Although we are left with fewer overarching laws as a result, we are much more confident in the patterns that remain, giving us a sounder basis for future explorations into how the topology food-web networks emerges from and feeds back on the processes that build, destroy, and reshape ecosystems.

6.1.1 What patterns have food-web models revealed?

The one overwhelming pattern that has emerged from careful application of food-web models to data is the validity of the concept of the low-dimensional feeding niche [8]. Several different probabilistic formulations support this conclusion, as compared with other model structures using model selection techniques [28, 30, 22]. The work in this dissertation on the Serengeti food web also supports that conclusion, even though the Serengeti data set is limited to three specific components of a much larger ecosystem. The other seeming laws of food-web structure have given way to more nuanced descriptions of specific webs.

A more difficult topic to disentangle is whether food webs are organized into compartments, groups, modules, etc. Although the misuse of modularity metrics has led to the overinterpretation of compartmentalization in food webs [23], the group model has shown that it is certainly possible to decompose food webs into meaningful collections of functional trophic groups, between which link structure is compartmentalized. Surveys across webs show that niche-based models do better than group-based models, indicating that the group structure emerges in part from the constraints of niche space, a result that seems consistent with work demonstrating compartmentalization in niche-model food webs [13]. However, the work presented in Chapter 4 demonstrates that, at least in the Serengeti data set analyzed there, clustering among species can be detected above and beyond the constraints of the niche space.

6.1.2 Methodological developments

Since the introduction of the cascade model, we have seen the adoption of a series of improved statistical methodologies for food-web models. The use of maximum-likelihood model fitting, and subsequently model selection via AIC, placed food web models on a stronger, more objective foundation for comparison [1, 2].

Due to some problems inherent with AIC, recent work is emerging based on the minimum description length (MDL) principle, which views statistical models essentially as compression algorithms for data [22]. Under MDL, good models are those that minimize the description length of the data, that is, those that minimize the number of bits needed to transmit both model parameters plus the number of bits needed to specify the additional information content present in the data itself. Although an old idea among information theorists, it has only recently been unified with maximum-likelihood modeling via a quantity known as normalized maximum likelihood, which involves integrating a model over all possible data sets in order to determine the amount of information content that separates observed data from a model description. Staniczenko and Allesina applied this approach to food-web models that can be represented blocks of Erdős-Renyi random graphs, which include the cascade model [8], the group model [2], as well as various versions of the niche model [29, 25, 1].

The work in this dissertation introduced the use of the full Bayesian framework for food-web models, which is a very flexible approach for building food-web models but also computationally challenging (see Chapters 2, 3, and 4). The first main advantage of these models is that model parameters, as well as the data, are modeled probabilistically, so that the model-inference process provides samples from a distribution of model parameters, and thus estimates of uncertainty, rather than a single point estimate that must be further explored via, e.g., profile likelihoods. The second main advantage is that, with parameters formulated probabilistically, it becomes straightforward to also model them hierarchically, with parameters for, say, individual species being drawn from

a distribution for their group, and the parameters controlling the group distribution being drawn from a top-level prior. This flexibility made the group-niche model of Chapter 4 possible.

The main disadvantage of the full Bayesian method is computational, since the goal of the inference process is to generate independent samples from the full posterior distribution over model parameters. Markov-chain Monte Carlo methods provide a practical and conceptually straightforward way of sampling from the posterior, but, as I learned the hard way, implementation details can be challenging to deal with, and convergence is never guaranteed, especially for models with as many parameters as the group-niche model. These challenges motivated the development of the mc3kit software package (Section 2.7, which like existing packages such as BUGS or JAGS allows sampling from arbitrary hierarchical structures, and additionally implements parallel multi-chain inference as well manipulation of the dependency graph as part of the inference process.

6.2 Compartments, stability, robustness, and spatial coupling

The main finding Chapter 5 is that, when considering bottom-up extinctions alone, compartmentalized link patterns tend to reduce robustness. This result is relatively straightforward from first principles, but the discourse has been so strongly skewed toward the notion that compartmentalization stabilizes food webs [24] that, from the view of a vague interpretation of that idea, it seems surprising. This is not to say, however, that compartmentalized structures are not stabilizing, merely that, when top-down extinction cascades are not considered, they decrease the robustness of food webs in response to targeted bottom-up removal.

One of the general link patterns observed in the Serengeti, and caricatured in the robustness experiment, is the coupling of groups of species at lower trophic levels by groups of species at higher trophic levels. This pattern is more compartmentalized than random link assignment between trophic levels, and is thus less robust. However, relative to patterns with strict compartments spanning all trophic levels, this pattern is less compartmentalized: species share prey distributed across multiple groups at the level below them. From a purely bottom-up perspective, this pattern provides

link redundancy, compared to strict compartmentalization, to species at higher trophic levels: if a species loses one of the groups below, it still has links in the remaining group.

Interestingly, this pattern of link redundancy seems to be coherent with results on dynamic stability by McCann [17], in which individual species were arranged in this kind of pattern, with species at higher trophic levels coupling food chains at lower trophic levels. The dynamic argument, which has some support from simulations, is similar to the structural one: if one of the prey species is reduced in abundance, the predator species can pick up the slack from the other food chain. Patterns of link redundancy thus may correspond in this case to patterns of dynamic stabilization.

The pattern of trophic coupling is linked to the idea that larger organisms having a wider spatial range couple smaller organisms with a smaller range. Work by the same group shows empirical support for this pattern [19], and certainly in the Serengeti food web herbivores have a wider range than plants—whose habitat structure is reflected in inferred trophic groups—and carnivores have a wider range than herbivores. Links will thus be compartmentalized at least in part based on the differing ranges at different trophic levels. At larger scales, there should be major habitat boundaries that induce strict compartmentalization. At smaller scales, though, it seems plausible that this pattern of coupling, which also seems to confer both structural robustness and dynamical stability, would be more dominant.

6.3 Future directions for the food web

Food web networks have have proven to be a useful representation of the complexity of ecosystems, and there is much more to learn from them and about them. Today, however, the greatest opportunities for advancing our understanding lie in extending them past the simple binary-network representations studied in this dissertation. The integration of techniques and ideas from probabilistic modeling of food-web structure with other kinds of data presents an opportunity to build much more useful and predictive models for comprehending and managing ecosystems. Furthermore, although feeding relationships are fundamental to ecosystems, they are mediated by a wide

range of other processes that today are understood for the most part only as isolated components, not as part of a complex network of interactions.

The representation of food webs as simply the presence or absence of feeding links between pairs of species alone gives a surprisingly challenging view of ecosystems, because the link topology alone is quite complex. But species are not equally coupled to each other, and better data on the strength of the links in food webs should improve our understanding of food-web structure [10, 5]. The probabilistic modeling approach generalizes naturally to weighted networks, and some work along these lines is taking place (Jarrett Byrnes, personal communication).

In statistical parlance, the food-web models used for the Serengeti analyses are generative: that is, they define a probability distribution over food webs conditional on nothing but the model. Discriminative models, which are used to predict data from other data—and thus define a conditional probability distribution—should be very useful in systems where species trait data is available. Such models have had relatively limited use to date for multitrophic food webs, with the exception of a body-size model, partly because of data availability, and partly because food-web modeling emerged from a mindset that emphasized general laws, not statistical prediction. Chapter 2 covers one such model, which was used to relate the traits of *Psychotria* plant species, including mechanical and chemical defense traits, with DNA observations of the species identified in the guts of herbivores; that model is also an example of the use of raw observations of interaction frequency rather than a qualitative feeding network [21].

An important reason to understand food-web structure is so that we can better predict ecosystem dynamics, particularly in response to human-mediated change. With an eye toward building predictive dynamical models, food web topology, weighted or unweighted, could serve as a useful constraint on model structure. When network structure is unknown or only partially known, it may instead be an object to be inferred from dynamical data to then serve as information for future prediction [9]. The general niche-based models of food-web structure provide information about the general shape we should expect from inferred networks, and may thus provide a useful constraint

even in the absence of topological information from a particular system. Species traits should also prove useful in the context of such models. Modern experience with statistical prediction, via machine learning, has shown that a wide variety of data can be usefully combined into models and will typically outperform simple models built to test Popperian hypotheses [7, 6]. The intuitive reason is this: in a highly connected system, the answer to the question “does x affect y ” is yes, so by developing techniques that can integrate all available data—including sparse, noisy data that alone would be termed non-significant—we can build much better predictive models. These models may be much more opaque to human understanding, but may also give better predictive results, and should enable better management of complex ecosystems.

Ecosystems include a wide variety of interactions aside from feeding relationships, and an important research direction for ecology is a network-based view that integrates these disparate interactions. Many network studies have already focused on mutualisms, such as pollination networks, in isolation [15, 26]. However, integrated views are rare, though they have been called for [14]. As we build up more and more specific data about interactions in different ecosystems, we may see a trend akin to what has happened in systems biology [18], where generic views of cells as networks have been replaced by very system-specific models built of a multitude of interaction pathways among genes, transcription factors, and metabolites, but with an eye toward mathematical modeling and the understanding of large-scale, emergent behaviors.

Still, feeding relationships capture the essence of energy and nutrient flow through a food web, and are thus fundamental in a way that other interactions are not. One promising way to expand our view is to view non-trophic effects as interaction modification of trophic links [31, 11], or, to use another terminology, as representing trait-mediated indirect interactions [27, 20]. To date, most studies of interaction modification and trait-mediated interactions have been in small empirical or model systems, with a few exceptions [4, 12]. An important question is whether this modified view of the food web will require us to fundamentally rethink notions of stability, robustness, and even the basic conceptualization of an ecosystem in food-web terms; or if, essentially, these complex

higher-order interactions “average out” at the scale of the food web.

6.4 Theory is dead, long live theory!

The bulk of this dissertation focused on the application of statistical methods to detecting patterns in the structure of food webs based on relatively simple ideas, but which ultimately fit fairly specific structures in particular data sets. Furthermore, I speculate that, to the extent that we seek better predictive models for particular systems, we will increasingly employ general-purpose machine learning methods that can adapt to regularities in data. In some cases, those regularities will yield insight and understanding; in many others, they may just be a way to build a better prediction machine for complex ecological systems. In the age of Google, the strange notion that large quantities of data will obviate the need for scientific theory of any kind is gaining traction [16, 3]. When the goal is to predict the behavior of systems for which we have data that can characterize the range of behaviors, the theory of theory-free science seems plausible.

But if we want to expand understanding past the boundaries of predicting things we have already observed, theory is vital: it is our mechanism for guiding exploration into the unknown. Even as we improve statistical methods and predictive models, we should reflect on how little we still know about how tangled webs of living and nonliving matter and energy assemble, collapse, reform, and evolve, and be inspired to think about them in new ways unconstrained by the distractions of data: on paper, in the computer, and alone in the forest.

References

- [1] S. Allesina, D. Alonso, and M. Pascual. “A general model for food web structure”. In: *Science* 320.5876 (2008), pp. 658–661. DOI: 10.1126/science.1156269.
- [2] S. Allesina and M. Pascual. “Food web models: a plea for groups”. In: *Ecology Letters* 12.7 (2009), pp. 652–662. DOI: 10.1111/j.1461-0248.2009.01321.x.

- [3] C. Anderson. “The end of theory: the data deluge makes the scientific method obsolete”. In: *Wired Magazine* 16.7 (2008).
- [4] R. Arditi, J. Michalski, and A. H. Hirzel. “Rheagogies: modelling non-trophic effects in food webs”. In: *Ecological Complexity* 2.3 (2005), pp. 249–258. DOI: 10.1016/j.ecocom.2005.04.003.
- [5] J. Bascompte, C. J. Melián, and E. Sala. “Interaction strength combinations and the over-fishing of a marine food web”. In: *Proceedings of the National Academy of Sciences of the United States of America* 102.15 (2005), pp. 5443–5447. DOI: 10.1073/pnas.0501562102.
- [6] C. M. Bishop and N. M. Nasrabadi. *Pattern recognition and machine learning*. Vol. 1. Springer, 2006.
- [7] L. Breiman. “Statistical modeling: The two cultures (with comments and a rejoinder by the author)”. In: *Statistical Science* 16.3 (2001), pp. 199–231. DOI: 10.1214/ss/1009213726.
- [8] J. E. Cohen, F. Briand, and C. M. Newman. *Community Food Webs: Data and Theory*. Biomathematics. Springer, 1990. ISBN: 0387511296.
- [9] J. M. Drake. “Computational methods for identifying structure in ecological networks”. 2011.
- [10] M. C. Emmerson and D. Raffaelli. “Predator-prey body size, interaction strength and the stability of a real food web”. In: *Journal of Animal Ecology* 73.3 (2004), pp. 399–409. DOI: 10.1111/j.0021-8790.2004.00818.x.
- [11] A. J. Golubski and P. A. Abrams. “Modifying modifiers: what happens when interspecific interactions interact?” In: *Journal of Animal Ecology* 80.5 (2011), pp. 1097–1108. DOI: 10.1111/j.1365-2656.2011.01852.x.
- [12] A. Goudard and M. Loreau. “Nontrophic interactions, biodiversity, and ecosystem functioning: an interaction web model”. In: *American Naturalist* 171 (2008), pp. 91–106. DOI: 10.1086/523945.

- [13] R. Guimera, D. Stouffer, M. Sales-Pardo, E. Leicht, M. Newman, and L. Amaral. “Origin of compartmentalization in food webs”. In: *Ecology* 91.10 (2010), pp. 2941–2951. DOI: 10.1890/09-1175.1.
- [14] T. C. Ings, J. M. Montoya, J. Bascompte, N. Blüthgen, L. Brown, C. F. Dormann, F. Edwards, D. Figueroa, U. Jacob, and J. I. Jones. “Review: Ecological networks—beyond food webs”. In: *Journal of Animal Ecology* 78.1 (2009), pp. 253–269. DOI: 10.1111/j.1365-2656.2008.01460.x.
- [15] P. Jordano, J. Bascompte, and J. M. Olesen. “The ecological consequences of complex topology and nested structure in pollination webs”. In: *Plant-Pollinator Interactions: From Specialization to Generalization*. Ed. by N. M. Waser and J. Ollerton. University of Chicago Press Chicago, Illinois, 2006, pp. 173–199.
- [16] K. Kelly. *The Google Way of Science*. 2008. URL: http://www.kk.org/thetechnium/archives/2008/06/the_google_way.php.
- [17] K. McCann, J. Rasmussen, and J. Umbanhowar. “The dynamics of spatially coupled food webs”. In: *Ecology Letters* 8.5 (2005), pp. 513–523. DOI: 10.1111/j.1461-0248.2005.00742.x.
- [18] J. K. Nicholson and J. C. Lindon. “Systems biology: metabonomics”. In: *Nature* 455.7216 (2008), pp. 1054–1056. DOI: 10.1038/4551054a.
- [19] N. Rooney, K. McCann, G. Gellner, and J. C. Moore. “Structural asymmetry and the stability of diverse food webs”. In: *Nature* 442.7100 (2006), pp. 265–269. DOI: 10.1038/nature04887.
- [20] O. J. Schmitz, V. Krivan, and O. Ovadia. “Trophic cascades: the primacy of trait-mediated indirect interactions”. In: *Ecology Letters* 7.2 (2004), pp. 153–163. DOI: 10.1111/j.1461-0248.2003.00560.x.

- [21] B. E. Sedio, E. B. Baskerville, M. D. Hunter, S. J. Wright, J. W. Westbrook, J. R. de Wet, and C. R. Dick. “Character displacement of anti-herbivore defenses facilitates the coexistence of closely-related rainforest trees”. In: (*in prep.*) (2013).
- [22] P. P. Staniczenko, M. J. Smith, and S. Allesina. “Selecting food web models: the good, the bad, and the ugly”. In: (*in prep.*) (2013).
- [23] P. P. Staniczenko, M. J. Smith, and S. Allesina. “Significance of modularity in food webs”. In: (*in prep.*) (2013).
- [24] D. B. Stouffer and J. Bascompte. “Compartmentalization increases food-web persistence”. In: *Proceedings of the National Academy of Sciences* 108.9 (2011), pp. 3648–3652. DOI: 10.1073/pnas.1014353108.
- [25] D. Stouffer, J. Camacho, and L. Amaral. “A robust measure of food web intervality”. In: *Proceedings of the National Academy of Sciences* 103.50 (2006), p. 19015. DOI: 10.1073/pnas.0603844103.
- [26] E. Thébault and C. Fontaine. “Stability of ecological communities and the architecture of mutualistic and trophic networks”. In: *Science* 329.5993 (2010), p. 853. DOI: 10.1126/science.1188321.
- [27] E. E. Werner and S. D. Peacor. “A review of trait-mediated indirect interactions in ecological communities”. In: *Ecology* 84.5 (2003), pp. 1083–1100. DOI: 10.1890/0012-9658(2003)084[1083:AROTII]2.0.CO;2.
- [28] R. J. Williams, A. Anandanadesan, and D. Purves. “The probabilistic niche model reveals the niche structure and role of body size in a complex food web”. In: *PLoS One* 5.8 (2010), e12092. DOI: 10.1371/journal.pone.0012092.g001.
- [29] R. J. Williams and N. D. Martinez. “Simple rules yield complex food webs”. In: *Nature* 404.6774 (2000), pp. 180–183. DOI: 10.1038/35004572.

- [30] R. J. Williams and D. W. Purves. “The probabilistic niche model reveals substantial variation in the niche structure of empirical food webs”. In: *Ecology* 92.9 (2011), pp. 1849–1857. DOI: 10.1890/11-0200.1.
- [31] J. T. Wootton. “The nature and consequences of indirect effects in ecological communities”. In: *Annual Review of Ecology and Systematics* (1994), pp. 443–466. DOI: 10.2307/2097320.