Title
Disturbance Macroecology: An Information Entropy Approach for Cross-System Comparisons of Ecosystems in Transition

Author
Newman, Erica A.

Publication Date
2016

Peer reviewed|Thesis/dissertation

Disturbance Macroecology: An Information Entropy Approach for Cross-System Comparisons of Ecosystems in Transition

by

Erica Anna Newman

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in the

Energy and Resources Group

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John Harte, Co-Chair
Professor Max Alan Moritz, Co-Chair
Professor Steven R. Beissinger
Professor Scott L. Stephens

Spring 2016

Abstract

Disturbance Macroecology: An Information Entropy Approach for Cross-System Comparisons of Ecosystems in Transition

by

Erica Anna Newman

Doctor of Philosophy in Energy and Resources

University of California, Berkeley

Professor John Harte, Co-Chair

Professor Max Alan Moritz, Co-Chair

Little is known about how metrics of biodiversity and abundance scale in ecologically disturbed and disrupted systems. Natural disturbances have a fundamental role in structuring ecological communities, and the study of these processes and extension to novel ecological disruptions is of increasing importance due to global change and mounting human impacts. Numerous studies have demonstrated the importance of natural disturbance in determining basic ecological properties of an ecosystem, including species diversity, membership, and relative abundances of those species, as well as overall productivity. Although estimating ecological metrics at both the species and community level is of critical importance to conservation goals, predicting the impacts of disturbance and disruption, including anthropogenic changes, on ecosystems is a major problem for ecological theory for several reasons. Disturbances are diverse in type, create patches that are internally heterogeneous, interact with site-specific disturbance legacies, and have different effects over multiple spatial and temporal scales. In contrast, empirical studies providing the basis for development of models tend to focus on short time-scales and relatively homogeneous systems with steady-state dynamics. Sites that experience single disturbances or are part of disturbance regimes also pose a challenge to ecological theory because they represent open, non-equilibrium systems that are not tractable with equilibrium mathematics. Additionally, the spatial scale at which a disturbance is studied will affect the conclusions that are drawn about communities or their component species. Nevertheless, the ubiquity and importance of disturbance to ecosystems continues to motivate a search for generality in disturbance and landscape ecology. In this dissertation, I apply an information entropy based theory of macroecology to ecosystems in transition, or have otherwise experienced ecological disruption. This leads to comparable results between systems, and forms a basis for cross-system comparisons of ecosystems in transition.

The maximum information entropy inference procedure (MaxEnt) has been proven to produce the least-biased estimates of a probability distribution, given prior knowledge of a system. Empirical values make up the prior knowledge of the system, and constrain the mean, variance, or higher moments of a given distribution. An extension of the MaxEnt procedure, the Maximum

Entropy Theory of Ecology (METE) takes a macroecological approach to estimating plot- to landscape- to biome-scale species diversity, abundance, and energetics metrics, using only the relationships between four non-adjustable state variables S_0 (total species), A_0 (area under consideration), N_0 (total abundance), and E_0 (total metabolic energy), and no adjustable parameters to characterize the scaling of diversity and abundances of species in a system. Until this work, METE has mainly been tested in steady-state and minimally disturbed systems.

In Chapter 1, "Disturbance macroecology: a comparative study of plant species' abundances and distributions in different-age post-fire stands of Bishop pine (*Pinus muricata*)," I investigate how metrics of biodiversity and abundance are scale in a plant community that is largely structured by a dominant, disturbance-dependent species. We target two different-aged stands in a region of high wildfire activity, one a characteristically mature stand with a diverse understory, and one more recently disturbed by a stand-replacing fire 17 years previously. We compare the stands using various macroecological metrics of species richness, abundance and spatial distributions that are predicted by METE, which does not rely on steady-state or equilibrium assumptions, and is therefore well-suited to be a null model for ecosystems in transitional states. Ecological patterns in the mature stand more closely match METE predictions than do data from the more recently disturbed stand. This suggests METE's predictions are more robust in late successional, slowly changing, or steady-state systems than those in rapid flux with respect to species composition, abundances, and body sizes. These findings highlight the need for a macroecological theory that incorporates natural disturbance and other ecological perturbations into its predictive capabilities, because most natural systems are not in a steady state.

In Chapter 2, "Macroecology for management: Testing an information-entropy-based theory of macroecology against anthropogenic disruption of high-Sierra meadows, I investigate the extent to which anthropogenic changes to an ecosystem, in the form of grazing by large, introduced herbivores, are detectable using METE, and small (<1 ha) replicate census plots. Anthropogenically-induced ecological disruptions (anthropogenic disruptions) have been overlooked by macroecological theory because they represent ecosystems in various states of transition that result from non-natural selection on the community. While critically important to understand for conservation reasons, anthropogenic disruptions are, in general, not comparable to each other, nor to other ecological disturbances that are natural in origin. Here, we use METE to examine the effects of an anthropogenically-induced novel disturbance regime of grazing by horses in high Sierra Nevada meadows on the species-abundance distributions (SAD), number of singleton species, and the species-level spatial abundance distributions (SSADs) (a measure of spatial aggregation) for all species in three pairs of grazed and ungrazed meadows, each meadow containing a system of plots set up across a moisture gradient. We find that number of singleton species may be a better indicator of ecological disruption than the shape of the SAD in systems where the differences in community structure are subtle. We also find that the METE SSAD performs better than all other models tested for both grazed and ungrazed plots. We suggest ways of augmenting tests of the METE SSAD to refine theory for management relevance.

In Chapter 3, "Empirical tests of within- and across-species energetics in a diverse plant community," I (with my coauthors) test the metabolic predictions of METE for herbaceous plants in a subalpine meadow. METE is an extremely general macroecological theory that predicts spatial, abundance and metabolic rate distributions of species, and the interrelationships of these metrics for any system defined by a set of basic community state variables. It therefore

also predicts body-size distributions if a metabolic scaling relationship between metabolism and body size is assumed. Many fundamental properties of ecological systems and interactions are tied to body size, and a related metric, the metabolic rate distribution, both within and across species. Extensive tests of METE's macroecological predictions in multiple ecosystems and with multiple taxa generally support its species diversity and abundance predictions, but two related predictions had not been evaluated against full community census data until this study: the distribution of metabolic rates of individuals within species as a function of the abundance of the species, and the distribution of average individual metabolic rates across species. We show that while METE realistically predicts the distribution of individual metabolic rates across the entire community, the within and across species predictions generally fail. We also test the energy-equivalence type prediction that arises as a consequence of the prediction for the distribution of average individual metabolic rates across species. We suggest several possible explanations for the empirical deviations from theory, and distinguish between the expected deviations caused by ecological disturbance and those deviations that might be corrected within the theory.

Taken together, these results indicate that it is possible to extend macroecological theory to ecosystems that experience natural disturbances and other ecological disruptions. Because we find that there are regular empirical deviations from theory in ecosystems that have experienced some sort of disturbance, we can conclude that the values and ratios of the four state variables (A_0, S_0, N_0 and E_0) used by METE are not sufficient to describe the dynamics of real ecosystems. These regular deviations, however, are interesting in their own right, because they suggest where ecological processes may influence the shape of empirical macroecological distributions. This will provide a framework for comparing and eventually predicting the various effects of disturbance on biodiversity, in the contexts of disturbance regimes, anthropogenic change, and mixtures of both.

MJK: One of the main things you'll notice when you start paying attention to your environment… is all the synchronicities and overlapping patterns. It's kind of like nature throwing you a bone.

I: Is that comforting for you? To know that there's kind of a logic there?

MJK: It actually makes everything more confusing.

Maynard James Keenan, *paraphrased from an interview with Loud*

Dedication

For my grandmother, Eleonora Luce Szwed
who had wanted to be a doctor

to Dudley H. Towne
for being himself

and to my husband, David Hembry
because I love him very much

Acknowledgments

I have many people to thank for their support, encouragement and teachings during the better part of a decade that I have been writing this dissertation. Some of these people are thanked in the individual chapters where they made their contributions, but there are many people outside of those lists who have contributed to my becoming a better scientist and almost certainly a better person in the past years. My dissertation could have been many different things centered on the themes of disturbance ecology and anthropogenic changes to ecosystems. Not all of the studies I conducted during my time at UC Berkeley are represented here, but all of the people who contributed to the work I conducted during my time in graduate school deserve recognition.

First I would like to thank my committee for their diverse perspectives and the variety of contributions they have made to my work. My advisors John Harte and Max Moritz, and my committee members Steve Beissinger and Scott Stephens have offered me encouragement, mentorship, funding, field sites, places in their lab groups, opportunities to speak at public engagements, comments, edits, language and writing direction, and moments of philosophy; all the necessary pieces of a successful dissertation. I can't thank them enough. I especially need to thank John Harte for meeting with me as a college graduate, and encouraging me to join his lab group for nearly 10 years before I finally did. Truly I also have to thank all of the Energy and Resources Group professors, because they have all been mentors at various times in various ways, and have all shaped my education and growth as an academic. Duncan Callaway, Isha Ray, Richard Norgaard, David Anthoff, Dan Kammen, and of course, John Harte were a few individuals who had outsized effects on my time in graduate school. Kay Burns and Sandra Dovali were staff in ERG who made the whole experience possible, and I am deeply grateful to them, as well as Doty Valrey from Environmental Science, Policy and Management.

Through the University of California, Berkeley, I have overlapped with people who have changed my life. Marc-André Parisien, Eric Waller, and Luke Macaulay helped me find my feet when I arrived. In a rough transition from fieldwork back to classes and seemingly endless amounts of time in front of the computer, these people in particular helped me discover what I needed to get through. Elissa Sato became a good friend through shared UC Berkeley experiences, and for other "life" reasons.

While working and studying abroad, there have been many people who have welcomed me and my husband David into their homes and their lives. Jaqueline and Johnny Faraire, Minarii, Rocky, and Delano (and now Grace and Bruce!) made our time in French Polynesia into something much more meaningful than field station life. They were and continue to be a second home and family to us. Ravahere Tapatuarai, Jean-Yves Meyer, Élie and Zaza Poroi were incredible friends and field companions. Hinano Teavai-Murphy was an invaluable resource and a generous host at Berkeley's Gump Station.

Sachiko and Takayoshi Nishida provided the friendship and sanity we needed while living in Japan. Having been through the Japan-to-US-to-Japan-again culture shock themselves, they were the perfect foil to some of the challenges we faced there. They were truly delightful naturalists to spend time with on our series of excursions through the wilder parts of Japan. Also, they have a lot of cats, and I really liked those cats.

A large number of people helped me choose the academic paths I went down before I came to graduate school, and inspired me to think and work broadly. First, I want to thank my parents, Stuart and Jura Newman, for their support during these years. Janet Kuhn Francis demonstrated to me that it was possible to be elegant, compassionate, and uncompromisingly smart. Lincoln Sadler allowed me to burn with his prescribed fire crew, and this was one of the most eye-opening experiences I am likely to ever have. John Gerwin gave me a safe place to land while I pursued the study of birds, and he gave me connections and opportunities not usually afforded to field technicians. Jeff Beane, Jeff Humphries, and Kendrick Weeks taught me how to be a naturalist. Dudley Towne was also a huge influence on my life, and continues to be an inspiration to me almost 15 years after his death. He was truly his own person. Although I have always regretted not speaking at his memorial service to say what an incredible teacher and physicist he was, I now understand that his prodigious talent in those areas was just a sliver of who he was as a person. I am so grateful to have spent time with him not just as his student, but as his friend.

I have had a number of mentors and friends who were my support network while dealing with some of the more difficult aspects of life while still a student. It was a great blow to lose my grandmother, Eleonora Luce Szwed during these years while I always was so busy. She was one of my best friends. Hope and Bernie Gulker have been so much to me, as they have known me as a child, a teenager, and an adult, and have always been there to lend an ear or provide advice, structure, or friendship. Now that Bernie is no longer with us, I endeavor to be there for Hope the way she has always been there for me. I would also like to thank Dan Kammen, Ellen Simms and David Ackerly in particular for their mentorship and guidance through the trickier aspects of graduate school. I also want to thank Marcia Black and Robert Vargas for the time they spent with me. I am also deeply grateful to Dr. Richard Burg for helping me resolve chronic pain issues that posed a serious challenge to continuing any sort of computer-based work.

My former students are some of the people I am most happy to know from this whole experience. Mark Wilber is an extraordinary person, and has taught me how to be a better leader, and clearer-thinking scientist. His scientific ability and flexibility is striking, and I was very fortunate to be able to work with him. Natalie Lowell just gets it, by virtue of being the coolest. Henry Houskeeper reflected back to me that it doesn't matter if you're a student and you don't have your own lab yet, you can always be helpful and influential in someone else's life. Natalie, Henry, and Mark were my field technicians who I am now so, so happy to have as colleagues, and I am really proud of their successes. Ori Chafe, though not an academic, is inspiring to me because of her raw dedication to conservation, and I am truly lucky to have had her in my roster of talented field technicians.

In the later years of this process, Laura Mehrmanesh helped me manage my time and prioritize important relationships. Diane Hembry stepped in to care for me when I was sick, and provided so much care and support to me and David that I really can't thank her enough in these acknowledgements. I want to thank the entire Hembry family for being family.

Kevin Simmons stepped back into my life to continue sharing his puzzling imagination. Jim Crawford listened to me dream. Nels Anderson asked me about birds. Amy Jones, Robert Gordon and Jenni Polodna were the friends and people with artistic vision I ultimately needed in my life. Barbara Haya was there when I needed a Berkeley home. All of these people made my life better and more interesting just by including me in theirs. I also want to thank Ryan

Letourneau for being a constant in my life for the last 7 years, which was especially important to me while living overseas.

Finally, I am deeply thankful to – and for—my husband, David Hembry, who has really been everything to me. I can't list all the things he's helped me with and taught me… everything from California plants and biogeography, to Japanese history and language, to writing manuscripts and figuring out how to see all the National Parks. He met me with daily encouragement and optimism that made it possible for me to continue taking on the challenges I do. They say marriage is an adventure, but they don't usually say that marriage is so many adventures you will never be able to explain them all to anyone else, so just share them and be happy. David has helped me with so much of the practical stuff, and so much of the stuff that's not practical at all, like when he filled out stacks of paperwork in French to allow me to take the field station cat out of French Polynesia home to California. So I additionally thank David for Mo'o You Sing, who is the best cat.

Dudley H. Towne (on right) at the 1993 gay rights protest in Washington state

Table of Contents

CHAPTER 1

Disturbance macroecology: a comparative study of plant species' abundances and distributions in different-age post-fire stands of Bishop Pine (*Pinus muricata*)

Chapter 1: Disturbance macroecology: a comparative study of plant species' abundances and distributions in different-age post-fire stands of Bishop Pine (*Pinus muricata*)

Abstract

Previous macroecological studies have largely restricted their scope to relatively steady-state systems. How metrics of biodiversity and abundance are expected to scale in disturbance-dependent ecosystems is unknown. Evidence gathered from systems recovering from recent, major ecological disruptions is sparse and likely does not apply to disturbance-dependent communities. Furthermore, "disturbance" has often been used as a catch-all phrase to refer to any ecological disruption, and ecosystems that have evolved with a natural disturbance regime are conflated with those having undergone anthropogenic changes. In this study, we examine macroecological patterns in a fire-dependent community of Bishop pine (*Pinus muricata*). We target two different-aged stands in a region of high wildfire activity, one a characteristically mature stand with a diverse understory, and one more recently disturbed by a stand-replacing fire 17 years previously. We compare the stands using various macroecological metrics of species richness, abundance and spatial distributions that are predicted by the Maximum Entropy Theory of Ecology (METE). METE is an information-entropy based theory that has proven highly successful in predicting a wide variety of community- and species-level macroecological metrics across a wide variety of systems and taxa, that does not rely on steady-state or equilibrium assumptions, and is therefore well-suited to be a null model for ecosystems in transitional states. Ecological patterns in the mature stand more closely match METE predictions than do data from the more recently disturbed stand. This suggests METE's predictions are more robust in late-successional, slowly changing, or steady-state systems than those in rapid flux with respect to species composition, abundances, and body sizes. Our findings highlight the need for a macroecological theory that incorporates natural disturbance and other ecological perturbations into its predictive capabilities, because most natural systems are not in a steady state.

Introduction

Disturbance is pervasive in ecosystems, and it influences patterns of species diversity, abundance, and community membership over space and through time (Turner 1989). However, macroecology, the discipline concerned with large-scale patterns of diversity, has primarily focused on ecosystems that are perceived to be relatively stable (Fisher et al. 2010), in that they exhibit low variance in community structure through time (Turner et al. 1993). Ecosystems (and patches within ecosystems) that have recently undergone, or are continuing to undergo, natural disturbances (defined as those that are part of a repetitive disturbance regime *sensu* Turner 2010), anthropogenic changes, and other ecological disruptions are likely to be in flux with respect to species composition and richness, species-area relationships, distribution of abundances, and body sizes, and intraspecific spatial distributions of individuals. However, the dynamics of disturbed sites and entire disturbance regimes are not captured by standard macroecological study systems, which are often chosen because they are in or near steady states (e.g. most of the Center for Tropical Forest Science plots represent late-successional, primary forest) (Condit 1998 Ch.1).

Here, we restrict the use of the term "disturbance" to refer to "natural disturbances," which satisfy the following four characteristics: a) they cause mortality of individual organisms in a community; b) however, they do not cause mortality of all individuals in the community and therefore do not result exclusively in primary succession; c) they are part of a historic and repetitive "disturbance regime" (Turner 2010) with well-defined characteristics (Pickett and White 1985, Turner et al. 1998, Turner 2010); and d) the disturbance is "absolute" rather than "relative" (Pickett and White 1985, White and Jentsch 2001) in that each disturbance event is "a relatively discrete event in time that disrupts the ecosystem, community or population structure and changes the resources, substrate availability or physical environment" (White and Jentsch 2001). We differentiate (natural) disturbances from ecological "perturbations" and "disruptions," which will refer to any other processes that restructure an ecological community, including events that are natural in origin but not part of repetitive disturbance regimes (e.g., landslides, extremely rare weather events), and those that are novel and may be anthropogenic in origin (e.g., human impacts). A lava flow or landslide that kills or physically removes all plant individuals in the community and results in primary succession would therefore not be a disturbance, but rather a perturbation or disruption under our definition (criterion b). This strict operational definition of "disturbance" as synonymous with "natural disturbance" is consistent with its usage in several influential reviews of disturbance ecology (Pickett and White 1985, White and Jentsch 2001, Turner 2010).

Although natural disturbances have both large- and small-scale structuring effects in all ecosystems (Turner 1989), no macroecological study, to our knowledge, has addressed how metrics of biodiversity and abundance scale in disturbance-dependent ecosystems. Various studies on succession have led macroecologists to invoke "disturbance" broadly (including human activities, environmental variability, invasive species and so on) as a factor responsible for deviations from theoretical predictions or expected patterns, although it remains unclear whether macroecological patterns reported across ecosystems are properties of undisturbed, steady-state communities or are properties of all ecological systems. This failure to incorporate disturbance into macroecology poses a major challenge to the utility of this field in understanding ecological dynamics as well as global change. Synthesizing a "macroecology of disturbance" that incorporates quantitative macroecological metrics could have considerable consequences for conservation efforts, given that many ecosystems with active disturbance regimes (and the species that have evolved in them) rank among the most globally endangered (Turner 2010; and see Noss et al. 1995, Schlossberg and King 2015, Batllori et al. 2013). Distinguishing the effects of natural disturbances from those of anthropogenic changes is also important for predicting future states of ecosystems.

Past macroecological work that incorporates ecological disturbances of any type has predominantly focused on their effects on the shape of the species-abundance distribution (SAD). Although the SAD is well-studied (reviewed in McGill et al. 2007; White et al. 2012; Baldridge et al. 2015), the underlying shape of a "natural" SAD is debated (see for example Hill et al. 1995, Nummelin 1998, Hill and Hamer 1998; Ulrich et al. 2010), and various distributions have been proposed. Empirical support for each of these distributions is mixed. For the rank-abundance form of the SAD, a lognormal distribution is reported from many steady-state systems (Whittaker 1965; May 1975; Gray 1981; Ulrich et al. 2010), while other studies (Dennis and Patil 1979, Kempton and Taylor 1974) and "big data" methods showing that the log series distribution may be the most common across systems and taxa (White et al. 2012; Baldridge et

al. 2015). One study suggests the prevalence of the "double geometric" distribution (Alroy 2015). Disturbance or ecological perturbation is often invoked as responsible for a lognormal SAD (Bazzaz 1975; Hill and Hamer 1998; Kempton and Taylor 1974; Death 1996; Newman et al. 2014). Kempton and Taylor (1974) show in a comparative study that moth communities in undisturbed plots sites in the Rothamsted Insect Survey in England exhibit log series SADs, and the plots recovering from agricultural activity exhibit lognormal SAD. Certain ecological factors, sampling schema (Ulrich et al. 2010), detection issues (Tokeshi 1993), and mathematical processes (such as the central limit theorem) may also produce the lognormal (Tokeshi 1993). Work focusing on succession suggests a transition in the shape of the SAD from geometric in early successional stages to lognormal and subsequently log series in later stages (Gray and Mirza 1979; Whittaker 1975; Bazzaz 1975). Other macroecological metrics are much less well studied in the context of ecological disruption, although the species-area relationship (SAR) has been examined through experimental work with removal of seed predators (Supp et al. 2012), and the effects of ecological disruptions and perturbations are beginning to be investigated more broadly (Supp et al. 2014, Mayor et al. 2015).

The Maximum Information Entropy Theory of Ecology (METE) is a macroecological theory (Harte et al., 2008; 2009; Harte 2011; Harte and Newman 2014) that provides a statistical framework for linking the SAR, the SAD, and species-level spatial abundance distributions (SSADs), a metric quantifying the spatial distribution of individuals in a species over a given area. The METE framework is based on the principle of information entropy maximization, which allows the derivation of least-biased probability distributions that are constrained by prior knowledge. METE incorporates "prior knowledge" of the system being studied in the form of empirical values for state variables corresponding to species richness (S_0), total abundance (N_0), area under consideration (A_0), and total rate of metabolism of all organisms (E_0) as constraints.

Empirical tests of METE generally support its predicted forms for macroecological metrics, including the species-area relationships, species abundance distributions (Harte et al., 2008; 2009; White et al. 2012), species-level spatial abundance distributions, and certain metabolic predictions (Newman et al. 2014, Xiao et al. 2014) (but other spatial distribution and metabolic predictions are not strongly upheld; see McGlinn et al. 2015, Newman et al. 2014, Xiao et al. 2014). METE has accurately predicted these metrics for a range of natural communities, including herbaceous plants, trees, vertebrates and invertebrates, and in temperate, tropical, and montane environments, as well as isolated island communities (Harte et al. 2008; Harte et al. 2009; Harte 2011; Rominger et al. 2015). This study represents the first assessment of these common macroecological metrics for a plant community in a high-intensity natural disturbance regime.

In this study, we ask how well various macroecological metrics that describe community structure, specifically the SAR, SAD, and SSADs perform at the stand level, for a forest stand that has undergone a recent (17 yr previous) disturbance in Bishop pine forests and for a nearby, mature stand in the same disturbance regime (Brown et al. 1999) at Point Reyes National Seashore (PRNS) in California, USA. We hypothesize that the METE will more accurately predict these community structure metrics in the more mature plot (Mount Vision) because it has had a longer time since disturbance to reach steady-state dynamics, and METE's predictions will be less accurate for the more recently disturbed (Bayview) plot.

If an information-entropy based theory of macroecology (METE) performs equally well

for both the mature and disturbed plots, we would have supporting evidence that the information contained in the four state variables that constrain the predicted distributions is sufficient to describe ecosystems, regardless of what their disturbance status is. This would suggest that METE's successes are independent of the disturbance history of an ecosystem. Alternately, METE might not work for one or both of the different-aged plots, which means that the theory's four state variables do not contain adequate information to constrain the predicted distribution to the empirical distributions. Because METE is constrained to predict the maximum information entropy distributions only, the functional forms are fixed after the state variables are specified. Failures of METE to accurately predict ecological metrics in rapidly changing ecosystems would indicate the need to characterize deviations of real ecosystems from METE's predictions, or modify the theory to allow prediction of macroecolgical metrics in ecosystems with active disturbance regimes, undergoing succession, or experiencing perturbations generally.

Materials and Methods

Bishop pines: a plant community that experiences natural disturbance

This study focuses on Bishop pine (*Pinus muricata*) stands and their associated plant communities, which exhibit an unusual natural history. Bishop pine is endemic to the California Floristic Province in North America and has a patchy distribution along the coast of California, USA and Baja California, Mexico, including the California Channel Islands (Millar 1983, Millar 1986, Little 1971, Stephens and Libby 2006) (Figure 1). Mature stands (~40-120 years old) may have individuals that are widely spaced, and a moderately diverse understory of forbs and shrubs. Stand-replacing fires cause regeneration of the Bishop pines into a uniform age and size-class, "dog-hair" stand of trees that is nearly a monoculture with almost no understory. Dense stands have been shown to undergo a process of self-thinning (Harvey et al. 2011). Alternately, some have described additional thinning fires during the lifecycle of the trees that can restore the more open canopy and allow some trees to mature into large individuals (Brown et al. 1999, S. Stephens, *pers. comm*.). Although we found no evidence for such a process at our field sites in the wildfire records maintained by Point Reyes National Seashore since the establishment of the park in 1962, Brown et al. 1999 document frequent wildfires (every 8-9 years, on average) from the 1700's through 1945 for Olema Valley and general Point Reyes area, including 4 large fires in the early 20th century (1904, 1906, 1923, and 1945). Of these, we believe the 1923 fire is the most likely to have affected our study sites. The fires are thought to be human-caused, and surface rather than stand-replacing. More recently, patterns of fire severity leading to landscape heterogeneity are described in Forrestel et al. (2011), which focused on vegetation succession, especially with respect to Bishop pine communities following the October 1995 Vision Fire, which burned 12,354 acres (5000 hectares, or 50 km^2) within the National Park unit (NPS 2005). Forrestel et al. found that Bishop pines increased in extent by 85% and have an altered spatial distribution following this high-severity fire.

Species compositions between plots are not directly comparable as plant communities because the sites are exposed to different local climates. However, we are able to use our data to test hypotheses about the effects of intense natural disturbance on plant community structure from a macroecological perspective identifying overall species-level and community-level patterns.

Site descriptions

Field sites were chosen within the boundaries of Point Reyes National Seashore (PRNS), on the Pacific coast of California, USA, ~50 km northwest of San Francisco. According to data from 1964-2012, PRNS experiences a Mediterranean-type climate, with mild winters (monthly lows of 2-4 ^{o}C and highs of 15-17 ^{o}C) and cool summers (monthly lows of 6-9 ^{o}C to highs of 18-24 ^{o}C) with most of the ~100 cm annual rainfall occurring in winter, and a substantial amount of moisture received from fog drip in the summers (Dawson 1998; Forrestel et al. 2015).

We placed study plots in two Bishop pine (*Pinus muricata*) stands, each plot measuring 256 m^2 (16m x 16m), and censused each for all aboveground, live vascular plants ≥1 cm in height in April, 2012. At each site, live plants were censused (all individuals counted) as completely as possible with double-observers, and each plant's spatial location in the sampling grid was recorded with a cell number, representing a 1 m^2 subdivision of the larger plot. Plants were identified to species in the field when possible using the Jepson Manual (Baldwin et al. 2012) and other field guides for the local region (Howell et al. 2007, Keator and Heady 1981). In the cases where plants could not be identified to species, "morphospecies" (plants with a large number of shared characteristics) were given a unique species identifier for analysis, and reference notes and photographs were taken in the field. METE's predictions are robust to sampling only within a given taxonomic category or guild, and the lumping and splitting of taxa, provided that such decisions are made consistently (Harte et al. 2013).

The higher elevation "Bayview" plot at 252 m (825 ft) was placed in an area of PRNS that burned in the 1995 Vision fire. We also surveyed, but discarded, a pilot plot ("Hillview") in the 1995 Vision fire burn area because it had less than 10 vascular plant species and was therefore unsuitable for analysis by METE (Harte 2011; requirement that S_0>>1). Both the Bayview and Hillside plots can be characterized as "dog-hair" type stands of thin, closely-growing trees, in which the ages of the Bishop Pines are uniform, and the understory is sparse or absent. Six trees were cored at this site to create a record of variability of widths among the center rings of growing trees following an intense fire and a period of rapid growth. The slightly lower-elevation "Mount Vision" plot was located at 213 m (698 ft) in a mature Bishop pine stand with a more diverse and lush understory. Fourteen trees were cored at this site to determine the ages of all trees in the plot, and results were corroborated with aerial photographs of this area in the PRNS archive (see below). The two plots, which are 6.1 km (~3.8 mi) apart are shown in Figure 2, and locations and characteristics are summarized in Table 1.

Plant communities within Bishop pine forests at PRNS are highly patchy and exhibit high beta-diversity in the understory due to slope, elevation, and various local factors affecting climate variation, including exposure to ocean fog (Forrestel et al. 2011). As a result, species compositions between plots cannot be considered direct successional stages.

Establishing disturbance histories

We examined land-use history records (including aerial photographs, contemporary accounts, historical ranch maps, and post-wildfire incident records) in the archives at Point Reyes National Seashore, in consultation with National Park Service (NPS) archival staff. Other fire records examined include CALFIRE's Department of Forestry and Fire Protection FRAP Fire Perimeters (available online at http://frap.cdf.ca.gov/data, accessed in July 2015).

At each plot, trees were cored using increment borers (Haglöf Sweden™). Cores were stored in labeled paper straws until they could be glued into wooden mounts. The number of trees cored was severely limited by agreement in the National Parks permit, and a better estimate of growth height of initial growth year rings could not be obtained. See Appendix A for more information on tree ages, sampling, and curatorial information.

The maximum information entropy approach

Macroecology as a discipline has generally avoided studies of disturbed systems for at least two reasons: first, that disturbed systems are perceived as being in "transition" and unlikely to produce replicable, generalizable results; and second, assumptions of steady-state, equilibrium and stabilizing mechanisms in macroecological theory are common and often required in order to solve equations (see Hubbell 2001). In contrast, METE relies on the maximum information entropy inference procedure (MaxEnt) to predict least-biased probability distributions, given empirical constraints (Jaynes 1982), but invokes no explicit physical or ecological mechanisms (Harte 2011, Harte and Newman 2014). An application of the MaxEnt procedure, the "ASNE" version of METE (Harte and Newman 2014) uses only the relationships between four non-adjustable state variables that take on values from the system being measured: S_0 (total species), A_0 (total area under consideration), N_0 (total abundance), and E_0 (total metabolic energy). The state variables are static, not dynamic in this formulation, and there are no adjustable parameters characterizing the scaling of species diversity, abundances, and energetics in a system. Mathematical forms of empirical constraints arise from ratios of the state variables. More complete mathematical constructions of distributions are available in Harte (2011). Census data from multiple plots within the PRNS Bishop pine community are used here to test METE predictions for the species-area relationship (SAR), the species-abundance distribution (SAD), and the species-level spatial abundance distributions (SSAD).

Species-Area Relationship (SAR) and scale collapse

The Species-Area Relationship (SAR) describes how species diversity increases with increasing area. It is represented by $\overline{S}(A \mid N_0, S_0, A_0)$, where A is a sampled area within the total A_0 under consideration, and is calculable from the state variables as:

$$\overline{S}(A \mid N_0, S_0, A_0) = \sum_{species} \left[1 - \Pi(0 \mid A, n_0, A_0\right] \qquad (1)$$

Here, $\Pi(0 \mid A, n_0, A_0)$is the probability that a cell (or smaller area A within A_0) will be unoccupied by a given species, and $\left[1 - \Pi(0 \mid A, n_0, A_0)\right]$ is therefore the probability of occupancy by that species. Scale collapse is a property that emerges from the METE SAR when the local slope of the SAR at each spatial scale is graphed against the ratio of N/S measured at that scale (Harte et al. 2009), and we test that property for both plots compared to the METE predicted curve (Harte 2011, Harte et al. 2013, Wilber et al. 2015) with "z-D" scale collapse plots. The SAR calculated here is the recursive SAR, which was shown to make more accurate predictions than the non-recursive version of the same metric (McGlinn et al. 2013).

Species-Abundance Distribution (SAD)

The SAD, Φ, models the distribution of the total abundance of individuals N_0, across all species, S_0.

$$\Phi(n \mid S_0, N_0) = \frac{1}{\ln\left(\frac{1}{1-e^{-\beta}}\right)} \cdot \frac{e^{-\beta n}}{n} \tag{2}$$

Here, n represents the individuals within a given species, and β is related to the Lagrange multipliers λ_1 and λ_2, such that $\lambda_1 = \beta - \lambda_2$ and $\lambda_2 = S_0/(E_0 - N_0)$; and β satisfies the approximate relationship:

$$\beta \ln\left(\frac{1}{1-e^{-\beta}}\right) \approx \frac{S_0}{N_0} \tag{3}$$

following Harte et al. 2008, Harte et al. 2009, but using exact normalization. An exact expression for β, its derivation, and discussion of simplifying assumptions is available in Harte (2011), Chapter 7.5. The METE SAD is here compared to the continuous lognormal distribution common to many of the studies mentioned previously, although the Poisson lognormal has also been suggested as an appropriate comparison (McGill et al. 2007, White et al. 2012).

Species-level Spatial Abundance Distributions (SSADs)

The spatial distribution of individuals of a given species and their level of aggregation is predicted in METE as the SSAD, for which there is some support in the literature (McGlinn et al. 2015). With $\Pi(n \mid A, n_0, A_0)$ (or simply $\Pi(n)$) defined as the species-level spatial abundance distribution, the normalization constraint on the probability distribution for a given species:

$$\sum_{n=0}^{n_0} \Pi(n \mid A, n_0, A_0) = 1 \tag{4}$$

the additional constraint on the mean value of the number of individuals per cell:

$$\sum_{n=0}^{n_0} n \cdot \Pi(n \mid A, n_0, A_0) = \frac{n_0 A}{A_0} \tag{5}$$

and λ_Π representing the Lagrange multiplier associated with that constraint, Z_Π is defined to be the partition function that normalizes the solutions. We can write down the form of the solution that maximizes information entropy (Jaynes 1982):

$$\Pi(n) = \frac{1}{Z_\Pi} e^{-\lambda_\Pi n} \tag{6}$$

The partition function can be obtained by solving for it and the real-valued Lagrange multiplier simultaneously:

$$Z_{\Pi} = \sum_{n=0}^{n_0} e^{-\lambda_{\Pi} n} = \frac{1 - e^{-\lambda_{\Pi}(n_0+1)}}{1 - e^{-\lambda_{\Pi}}} \quad (7)$$

Additional steps in the solution are available in Harte (2011) Chapter 7.4.

The METE prediction for the SSAD of a given species is an "upper-truncated" (ut) geometric series, describing the frequency distribution of cells of the smallest sampled area (A) within the larger area A_0, which contain the maximum number of individuals n. "Upper truncation" or "right truncation" refers to the domain of the predicted distribution being limited to its physical maximum, rather than having an unbounded upper tail of probability (see White et al. 2012, Appendix A). The METE ut-geometric prediction is here modeled with the number of parameters (k) of the distribution equal to 1, which is 1 parameter fewer than a regular truncated geometric. This is because the corresponding parameters in the METE distribution, i.e. the Lagrange multiplier and the normalization by the partition function, are uniquely determined by the same state variable values, and there are no parameters that may be adjusted in this arrangement.

Applying METE to ecosystems in transition

As applied here, METE might accurately capture "snapshots" of rapidly changing ecosystems at an instant in time. The predictions of this ASNE version of METE are static in time. A dynamic version, while desirable (Fisher et al. 2010), is not yet available. Here, separate plots are placed to capture the macroecological patterns that characterize live, aboveground plant communities within separate patches in a disturbed landscape. As is the case for many macroecological studies, we here choose to study living, aboveground plants above a certain size threshold, which are a high-detectability system, with relatively large, sessile organisms. "Propagules" or organisms that persist through major disturbances in some form (as in a seed bank) are essential to a complete understanding of the disturbance ecology of a particular system, and may make up the larger part of plant abundances in an area. These propagules are not included in our measurements, but METE predictions hold for the remaining community of interest, and metrics scale with the measured community, rather than the full community. METE makes the least biased predictions of the community of interest as characterized by the measured state variables (Harte 2011) as a result of its underlying MaxEnt formalism (Jaynes 1982). Additionally, limiting METE to a focal taxon or taxa does not prevent accurate predictions (Harte et al. 2013), and issues of misidentification of species affect predictions in a minor way (Harte et al. 2013). We can therefore state with some certainty that tests of METE in this study are (1) robust to biases that arise in low-detectability systems, and (2) able to model the live plant community without accounting for other forms of biodiversity present in the system. These assumptions are useful to draw attention to, as they differ in important ways between the fields of macroecology and disturbance ecology. Moving towards a synthesis of the two fields will require careful study design that is informed by knowledge of both fields.

Analyses

Analyses for the macroecological metrics considered in this paper were carried out with Python (van Rossum, 2001) in the open-source project "Macroeco" (Kitzes et al. 2014, Kitzes and Wilber 2016). SAR, SAD and SSAD scripts were available from the beta version of this software (accessed June 2015). Other analyses were carried out in "R" versions 3.0.1 and 3.1.1 (R Core Team 2013, 2015). Models for SADs and SSADs were compared using Akaike's Information Criterion (AIC) value corrected for small sample sizes (AIC*c*). For SAR and z-D scale collapse model comparisons, we compare models with R^2 values derived from one-to-one predicted versus observed graphs (White et al. 2012, Appendix A), because no method is available to generate likelihood functions required for AIC comparisons.

Results

Summary statistics and calculated parameters

A total of 2330 individual plants in 32 species were censused in the two plots analyzed for this study. Species and their presence in each plot are shown in Table 2.

The Bayview plot, which burned in 1995, was censused at 17 years after the Vision Fire, and contained 16 species and a total of 486 individuals (148 of which were Bishop pines). Six tree cores showed that trees ranged in raw ring counts (a rough estimate of age) from 7 to 16 years (mean = 12.3). Density of Bishop pines in this plot was measured to be 0.58 trees/m^2 (or 5800 stems/ha), with a total basal area occupied by trees of 45.15 m^2/256 m^2 plot. The Mount Vision plot contained 27 species and 1844 individuals total (14 of which were Bishop pines). Bishop pine density in this plot measured 0.06 trees/m^2 (or 600 stems/ha), including the very few seedling trees in the plot. The total basal area occupied by trees measured to be 17.74 m^2/256 m^2 plot. Tree cores varied in age from an estimated 24 to 43 years old (mean = 34.1; mode = 33). Live tree density per hectare estimates are consistent with previous estimates (Harvey et al. 2011, 2014).

In both plots, Bishop pines were the only tree in the overstory, and were the largest plants in each plot by estimated biomass. See Figure 3 for histograms of dbh measurements for each site. For comparison to METE-predicted metrics, we calculate the value of the parameter β = 6.515 x 10^{-3} for the Bayview plot for the measured values N_0 = 486, S_0 = 16, and β = 2.429 x 10^{-3} for the Mount Vision plot, with measured values N_0 = 1844, S_0 = 27. The values for the Lagrange multipliers λ_1 and λ_2 are not independently calculable because state variable E_0 was not measured for either plot.

Species-Area Relationship and scale collapse

Generally, the METE prediction for the SAR appears to be a good fit for both data sets, whereas the z-D (scale collapse) predictions show more deviation from the METE prediction for the recently disturbed plot. See Figures 4 and 5.

To determine the best model fits for the SAR and z-D, comparisons of R^2 values on a one-to-one line for predicted versus observed distributions (White et al. 2012) were carried out for both the Bayview and Mount Vision plots. Best-fit power laws were calculated from the SAR data for each site (Figure 4) and applied to the z-D graphs (Figure 5). R^2 values for the mature

Mount Vision plot support METE's predicted SAR over the best-fit power-law predictions (R^2_{METE}= 0.991; $R^2_{PowerLaw}$= 0.989), whereas the power law fit is a better fit for the recently disturbed Bayview plot (R^2_{METE}= 0.977; $R^2_{PowerLaw}$= 0.998) on a ln-ln graph. The z-D plots show the local slope of the SAR plotted against ln(N/S) at every scale; visually each plot confirms the better fit of the METE-predicted distribution over the best-fit power-law for the mature plot.

Species-Abundance Distribution (SAD)

Model selection comparing AIC*c* values for both the Bayview and Mount Vision plots support the METE-predicted log series distribution over the lognormal distribution (Table 3) that is sometimes characteristic of disrupted systems. On visual inspection of the SAD graphs (Figure 6), there is the characteristic pattern of suppression of mid-abundance species in the Bayview plot that is characteristic of recently disrupted plots. The METE SAD does not capture this deviation, but because it fits the number of singleton species and the abundance of the most abundant species in this distribution, it wins out over the lognormal distribution by AIC*c* comparisons.

Species-level Spatial Abundance Distributions (SSADs)

SSADs were calculated for all species. Results are presented for the higher abundance species (with $n \geq 20$), comprising 14 species in the older Mt. Vision plot, and 4 species in the disturbed Bayview plot (Figures 7-9). In Figure 7, two alternate ways of presenting the same data are shown using the species (TRIBOR) as an example; first, a rank abundance plot (with rank corresponding to how many cells are occupied by a given level of abundance), and second, a cumulative density function (CDF). Note that the Poisson and binomial predictions appear to give the same results for each of the SSADs; this is because both models correspond to a null hypothesis of random placement, although the binomial has finite support and the Poisson is calculated with infinite support. Tables 4 and 5 summarize AIC*c* comparisons between the candidate distributions for the SSAD: binomial, Poisson, and METE ut-geometric predictions, for the Bayview disturbed and Mount Vision mature plots, respectively (negative binomial fits are excluded here because they are "best-fit" and are not characterized by a single shape parameter). Figures 8 and 9 show all CDF plots for the high-abundance species in the Bayview and Mount Vision plots, respectively. Figure 10 shows the distribution of AIC*c* weights for model fits for all species, for the binomial, Poisson, and METE ut-geometric predictions, with higher AIC*c* weights corresponding to better model fits.

We find that for the recently disturbed Bayview plot, SSADs for all 16 species have AIC*c* values supporting a Poisson distribution in 11 of 16 cases, with the 5 remaining cases supporting METE ut-geometric distribution. No species are best described as having a binomial SSAD. Bishop pine distribution is best described by the METE ut-geometric distribution, with next-best supported model having ΔAIC*c*= 3.6334. For the 16 most abundant species in the Mount Vision mature plot, AIC*c* values support a Poisson distribution in 7 cases, and 9 cases support METE ut-geometric distribution. For all species in the plot, the Poisson distribution is the best fit for 10 species, while 17 species' SSADs are described by the METE ut-geometric distribution (Fig. 10). Again, the distribution of Bishop pines is best described by the METE ut-geometric distribution, with next-best supported model (Poisson) having ΔAIC*c*=1.3961.

Discussion

Predicted and empirical distributions in different-aged stands

This study demonstrates how the field of macroecology may benefit from incorporating natural disturbance regimes. Macroecological predictions of METE perform well in the mature stand in a disturbance-dependent community for the SAR, SAD, and the SSAD of both the dominant species (Bishop pine) and general plant community, compared to other candidate distributions. These results conform to our expectations, because the mature (Mount Vision) stand exhibits similar constancy and demographic stability as the extremely stable Smithsonian plots where METE has proven successful previously (Harte et al. 2008, Harte 2011, Xiao et al. 2015). METE predictions have variable and lower success in the recently disturbed (Bayview) plot.

SSAD tests in both plots would benefit from higher sample size, both in terms of number of species and individuals within species, and replicate plots. Sampling issues confounded certain results: although some two species of vines (RUBURS, TOXDIV) are among the higher abundance species at all sites, their physical description was limited to presence or absence in cells, rather than true abundance counts. As a result, the physical distribution of these species is an artifact of sampling; a high number of cells containing a single individual. This suggests that (1) for species where only occupancy can be measured, candidate models for the SSAD make degenerate predictions that can not be differentiated, therefore the SSAD is not usefully applied; (2) meaningful sampling can only be carried out at a scale where there are multiple individuals of the same species of interest in multiple cells; and (3) sampling design for tests against METE should exclude species that have this problem (resulting in changes to state variable values that will reflect only the remainder, focal community).

Deviations from METE's predicted distributions

METE is an effective approach for "snapshot ecology" type studies where detection rates for the taxa studied are high. However, like many forms of macroecology (Fisher et al. 2010), METE is not a dynamic theory in the ASNE formulation. Our results from two sites are consistent with the idea that ecological perturbation results in lognormal SADs, and may even be consistent with the idea that the SAD transforms during successive successional stages from a geometric shape through a lognormal to a log series (Gray and Mirza 1979; Whittaker 1975; Bazzaz 1975). It is also apparent that time since disturbance affects the shape of the SAD and various other metrics in this study, including the shift of SSADs from the Poisson towards the METE ut-geometric.

We believe deviations from METE's predicted SAD and the more Poisson-type SSAD distributions for the general plant community in the younger stand of Bishop pines is likely explained by a lack of steady-state dynamics. As an information-entropy based statistical framework that employs state variables to describe the "macrostate" of an ecosystem or plot within that ecosystem, the static, ASNE version of METE and the MaxEnt mathematics underlying it automatically solves for the set of distributions that maximizes information entropy. This predicted state always corresponds to a steady-state solution. Although METE does not invoke explicit mechanisms of stability, the fixed distributions it predicts for a given set of state variables likely closely correspond to mature biological communities experiencing very little demographic fluctuations or other large shifts in community composition over time. This in

turn may explain why METE works better for the mature stand than for the younger, more recently disturbed stand in this study. However, it still leaves open the questions of how macroecology can account for disturbance in ecosystems, and what implications this has for predicting their ecological effects.

Other examples of notable deviations from METE's predictions have been observed in the Barro Colorado Island (BCI) forest plot, in the drought-affected Rocky Mountain (RM) meadow studied in Newman et al. 2014, and in some Hawaiian arthropod communities (Rominger et al. *in prep.*). For BCI, tree and seed-disperser extirpation on the island following its isolation from the mainland was a consequence of the construction of the Panama Canal. Time since isolation has been associated with an increasingly lognormal SAD. The lognormal SAD is also observed in the RM meadow during a period of unusual drought and high temperatures leading to a novel community of wildflowers that exhibited irregular phenology (Newman et al. 2014). In Hawaii case, the SAD shows higher-than-predicted numbers of singleton species. Deviation from the METE in this case may be caused by dispersal limitation and the relatively young age of the community (Rominger et al. *in prep.*). In each case, ecological context suggests that these systems are far from steady-state dynamics and provides insight into the ecological patterns observed.

Unifying macroecology with disturbance ecology

Until this study, no macroecological studies have focused on patterns in species diversity, spatial and abundance distributions in natural disturbance regimes. This study examines two very different successional states in a disturbance dependent ecosystem in an attempt to maximize the differences between METE's predictive abilities in communities with different disturbance histories.

We show that at the plot-scale, METE predictions are generally better supported for the more mature, less rapidly changing plot. Although METE has been demonstrated to work at the largest scale of ecosystems (Harte et al. 2009, Harte 2011, White et al. 2012, Harte and Kitzes 2015), it is unclear how well METE would predict various metrics at an intermediate "landscape scale" (~50,000-100,000 ha) that landscape contains multiple patches with different disturbance histories. Census information for this scale is generally lacking, and understanding biodiversity patterns at these large scales is part of the motivation to study macroecology (Brown 1995).

This study is a first step towards integrating macroecology into the study of landscapes undergoing natural disturbances; however, examining single successional states will not capture integral aspects of disturbance-dependent ecosystems. For example, the dynamics of the "shifting mosaic" of different successional states and patches that themselves may exhibit dynamic steady-states with respect to regional climate (Bormann and Likens 1979, Wu and Loucks 1995). As a form of "snapshot ecology" that incorporates no dynamics, it is possible that the ASNE version of METE may best be applied to patches within a disturbed landscape to characterize zones of different ages and disturbance histories. Alternately, using average or median values for the state variables S_0 and N_0 from multiple patches at different successional stages (or from plots with a mixture of successional stages) may adequately predict certain metrics (such as the SAR and SAD), while failing to predict other metrics (such as the SSADs and some or all metabolic metrics). The scaling of macroecological metrics may even provide insight into the scales at which disturbance-dependent ecosystems deviate most dramatically

from neutral, as patterns in disturbance-dependent landscapes are known to be scale-dependent (Wu 2004). These patterns may indeed change between landscapes characterized by large, infrequent disturbances (as in this study) and small, frequent disturbances (Romme et al. 1998, Turner and Dale 1998). Another opportunity to unify disturbance ecology with macroecology would be to use METE to test aspects of the Intermediate Disturbance Hypothesis, a long-standing idea in disturbance ecology with mixed support (see summaries in Fox 2013, Sheil and Burslem 2013). More studies in disturbance-dependent ecosystems would be necessary to evaluate these hypotheses.

Implications for conservation of Bishop pine forests

Although these Bishop pine forests provide a window onto the macroecology of natural disturbances, anthropogenic disruptions have become dominant in this ecosystem. Ongoing infestation by a non-native pathogen (Pitch Canker, *Fusarium circinatum*) is rapidly causing Bishop pine mortality over large, continuous areas within the Vision Fire burned area. Mortality in some places is close to 100% (Ben Becker, *pers. comm.*). This pathogen is likely to permanently affect stand structure and viability of Bishop pines, and may endanger them as a species. Understanding stand structure and heterogeneity in unaffected stands while they are still present may be critical to conservation efforts.

Acknowledgements

I thank Point Reyes National Seashore for providing permits, field sites, logistical support and facilities. This study was carried out under PRNS Park-assigned permit PORE-2012-SCI-0014, Activity #PORE-00572. I thank B. Becker for permitting assistance, C. Derooy, P.P. Creaseman and G. Dove for curatorial assistance, and M. Wilber, D. Hembry, A. Forrestel, M. Moritz, S. Beissinger, and S. Stephens for useful discussions and comments on earlier versions of this manuscript. I also thank K. Krasnow and S. Stephens for materials and training, and M. Wilber and K. Wilkin for field assistance. This research is funded in part by the Gordon and Betty Moore Foundation, and by the NSF through the Graduate Research Fellowship and grant NSF-EF-1137685.

Chapter 1 Tables and Figures

Tables.

Table 1. Locations and other descriptive metrics for research plots used in this study

Site name	**Bayview**	**Mount Vision**
GPS coordinates	38.0593628°, -122.8507065° (+/- 3.6m)	38.1028328°, -122.8933785° (+/-1.8m)
Elevation	251.46 m (825 ft)	212.75 m (698 ft)
Time since major disturbance	17 years	No recorded disturbance history, although tree age is maximum estimated 43 years
Slope, aspect	0^o, S	32^o, NE
Tree density/m^2	0.578	0.055
Trees cored; (cores available)	4; (6)	13; (14)
PRNS catalog numbers	PORE 18080 through PORE 18083; PRNS Accession number: PORE-00866	PORE 18084 through PORE 18096; PRNS Accession number: PORE-00866
Total species (S_0)	16	27
Total abundance (N_0)	486	1844
Total area (A_0)	256 m^2	256 m^2

Table 2. Presence of species by plot in two *Pinus muricata* stands of different ages. Species marked with a single asterix (*) are non-native.

Higher taxon	Common name	Species name	Analysis code	Presence in Bayview	Presence in Mt. Vision
Ferns and fern allies	Sword fern	*Polystichum munitum*	POLMUN	x	x
	Bracken fern	*Pteridium aquilinum*	PTEAQU	x	x
Gymnosperms	Bishop pine	*Pinus muricata*	PINMUR	x	x
Eudicots	California lilac	*Ceanothus thyrsiflorus*	CEATHY	x	–
	Cape ivy*	*Delairea odorata**	DELODO	–	x
	Wood strawberry	*Fragaria vesca*	FRAVES		x
	California coffeeberry	*Frangula californica* (syn. *Rhamnus californica*)	FRACAL	–	x
	Bedstraw	*Galium sp.* (possibly *porrigens*)	GALIUM	x	x
	Cow parsnip, pushki	*Heracleum maximum*	HERMAX	x	x
	Cat's ear*	*Hypochaeris radicata**	HYPRAD	–	x
	Red henbit*	*Lamium purpureum**	LAMPUR	–	x
	California honeysuckle	*Lonicera hispidula*	LONHIS	x	x
	False Solomon's seal, Slim Solomon	*Maianthemum stellatum*	MAISTE	–	x
	California man-root, wild cucumber	*Marah fabacea*	MARFAB	x	–
	Coast man-root	*Marah oregana* (syn. *Marah oreganus*)	MARORE	–	x
	Wax myrtle	*Morella californica*	MORCAL	–	x
	Tanoak	*Notholithocarpus densiflorus* (syn. *Lithocarpus densiflorus*)	NOTDEN	x	–
	Coast live oak	*Quercus agrifolia*	QUEAGR	x	–
	Canyon live oak	*Quercus chrysolepis*	QUECHR	–	x

	Flowering currant	*Ribes sanguineum*	RIBSAN	x	x
	California blackberry	*Rubus ursinus*	RUBURS	x	x
	Sheep's sorrel, sour weed	*Rumex acetosella**	RUMACE		x
	Red elderberry	*Sambucus racemosa*	SAMRAC		x
	Starflower	*Trientalis borealis*	TRIBOR		x
	Poison oak	*Toxicodendron diversilobum*	TOXDIV	x	x
	California huckleberry	*Vaccinum ovatum*	VACOVA	x	x
	UNKSP3	UNKSP3	UNKSP3	x	
	UNKSP5	UNKSP5	UNKSP5		x
Monocots	Unknown grass	UNKPOA1	UNKPOA1	x	x
	Unknown grass	UNKPOA2	UNKPOA2		x
	Unknown grass	UNKPOA2/3	UNKPOA3		x
	Unknown sedge	*Carex sp.*	CAREX		x

Table 3. Model comparisons of candidate Species Abundance Distributions (SADs) for the Bayview (disturbed) and Mount Vision (mature) plots. Here and following, k = number of parameters in model; AICc = Akaike's Information Criterion value corrected for small sample sizes; w_i = AICc weight (a measure of strength of evidence for each model); ΔAICc= difference of AICc value compared to the next best-supported model.

Plot	**Model**	***k***	**AIC*c***	**ΔAIC*c***	***w*$_i$**
Bayview	Lognormal	2	125.3246	5.1485	0.0708
	METE log series	**1**	**120.1761**	**0**	**0.9292**
Mount Vision	Lognormal	2	271.9128	3.7173	0.1349
	METE log series	**1**	**268.1956**	**0**	**0.8651**

Table 4. Model comparisons of Species-Specific Abundance Distributions (SSADs) for all species in the Bayview (17 years since disturbance) plot.

Species (rank in plot; total abundance)	**Model**	***k***	**AIC*c***	**ΔAIC**	***w*$_i$**
RUBURS	binm	2	503.8415	1.2021	0.3541
(r=1; n=204)	**pois**	**1**	**502.6394**	**0**	**0.6459**
	tgeo	1	633.8046	131.1651	0.0000
PINMUR	binm	2	538.7206	5.8831	0.0434
(r=2; n=148)	pois	1	536.4710	3.6334	0.1338
	tgeo	**1**	**532.8375**	**0**	**0.8228**
TOXDIV	binm	2	336.8804	1.8586	0.2826
(r=3; n=72)	**pois**	**1**	**335.0217**	**0**	**0.7158**
	tgeo	1	347.2462	12.2244	0.0016
RIBSAN	binm	2	145.8996	1.9218	0.2068
(r=4; n=20)	**pois**	**1**	**143.9778**	**0**	**0.5407**
	tgeo	1	145.5011	1.5233	0.2524
PTEAQU	binm	2	83.8498	3.2739	0.1094
(r=5; n=9)	pois	1	81.6494	1.0735	0.3286
	tgeo	**1**	**80.5759**	**0**	**0.5620**
MARFAB	binm	2	53.3378	1.9804	0.1598
(r=6; n=5)	**pois**	**1**	**51.3574**	**0**	**0.4303**
	tgeo	1	51.4544	0.0970	0.4099
CEATHY	binm	2	53.3378	1.9804	0.1598
(r=7; n=5)	**pois**	**1**	**51.3574**	**0**	**0.4303**
	tgeo	1	51.4544	0.0970	0.4099

Table 4. (*continued*)

Species (rank in plot; total abundance)	**Model**	***k***	**AIC*c***	**ΔAIC**	***w*$_i$**
VACOVA	binm	2	45.2554	1.9844	0.1584
(r=8; n=4)	**pois**	**1**	**43.2711**	**0**	**0.4273**
	tgeo	1	43.3332	0.0622	0.4142
QUEAGR	binm	2	45.2554	1.9844	0.1584
(r=9; n=4)	**pois**	**1**	**43.2711**	**0**	**0.4273**
	tgeo	1	43.3332	0.0622	0.4142
NOTDEN	binm	2	45.2554	1.9844	0.1584
(r=10; n=4)	**pois**	**1**	**43.2711**	**0**	**0.4273**
	tgeo	1	43.3332	0.0622	0.4142
UNKPOA1	binm	2	36.6677	1.9883	0.1573
(r=11; n=3)	**pois**	**1**	**34.6794**	**0**	**0.4250**
	tgeo	1	34.7144	0.0350	0.4177
LONHIS	binm	2	36.6677	1.9883	0.1573
(r=12; n=3)	**pois**	**1**	**34.6794**	**0**	**0.4250**
	tgeo	1	34.7144	0.0350	0.4177
GALIUM	binm	2	27.4003	1.9922	0.1564
(r=13; n=2)	**pois**	**1**	**25.4081**	**0**	**0.4234**
	tgeo	1	25.4235	0.0153	0.4202
UNKSP3	binm	2	17.0864	2.0000	0.1555
(r=14; n=1)	pois	1	15.0904	0.0039	0.4218
	tgeo	**1**	**15.0864**	**0**	**0.4227**
POLMUN	binm	2	17.0864	2.0000	0.1555
(r=15; n=1)	pois	1	15.0904	0.0039	0.4218
	tgeo	**1**	**15.0864**	**0**	**0.4227**
HERMAX	binm	2	17.0864	2.0000	0.1555
(r=16; n=1)	pois	1	15.0904	0.0039	0.4218
	tgeo	**1**	**15.0864**	**0**	**0.4227**

Table 5. Model comparisons of Species-Specific Abundance Distributions (SSADs) for the 16 most abundant species in the Mount Vision mature plot.

Species (rank in plot; total abundance)	**Model**	***k***	**AIC*c***	**ΔAIC**	***w*$_i$**
TRIBOR	binm	2	1973.8237	888.4117	0.0000
(r=1; n=657)	pois	1	1964.8166	879.4045	0.0000
	tgeo	**1**	**1085.4121**	**0**	**1.0000**
RUBURS	binm	2	514.9987	0.9987	0.3777
(r=2; n=256)	**pois**	**1**	**514.0000**	**0**	**0.6223**
	tgeo	1	711.7827	197.7827	0.0000
LAMPUR	binm	2	624.1456	109.7404	0.0000
(r=3; n=139)	pois	1	620.0188	105.6136	0.0000
	tgeo	**1**	**514.4051**	**0**	**1.0000**
TOXDIV	binm	2	448.9876	1.6406	0.3057
(r=4; n=128)	**pois**	**1**	**447.3470**	**0**	**0.6943**
	tgeo	1	490.8429	43.4959	0.0000
PTEAQU	binm	2	408.4793	1.6378	0.3060
(r=5; n=107)	**pois**	**1**	**406.8415**	**0**	**0.6940**
	tgeo	1	442.2202	35.3787	0.0000
LONHIS	binm	2	369.6808	1.7291	0.2964
(r=6; n=87)	**pois**	**1**	**367.9517**	**0**	**0.7036**
	tgeo	1	390.4842	22.5325	0.0000
UNKPOA2	binm	2	372.1632	1.9420	0.2726
(r=7; n=83)	**pois**	**1**	**370.2212**	**0**	**0.7199**
	tgeo	1	379.3697	9.1485	0.0074

Table 5. (*continued*)

Species (rank in plot; total abundance)	**Model**	***k***	**AIC*c***	**ΔAIC**	w_i
UNKPOA3	binm	2	368.9714	6.8254	0.0291
(r=8; n=77)	pois	1	366.7755	4.6295	0.0873
	tgeo	**1**	**362.1460**	**0**	**0.8836**
FRAVES	binm	2	324.0194	38.7963	0.0000
(r=9; n=53)	pois	1	320.6878	35.4647	0.0000
	tgeo	**1**	**285.2231**	**0**	**1.0000**
RUMACE	binm	2	338.0478	78.3158	0.0000
(r=10; n=46)	pois	1	331.2333	71.5012	0.0000
	tgeo	**1**	**259.7321**	**0**	**1.0000**
FRACAL	binm	2	221.2431	5.3742	0.0534
(r=11; n=35)	pois	1	219.0305	3.1615	0.1616
	tgeo	**1**	**215.8689**	**0**	**0.7850**
VACOVA	binm	2	186.7350	2.0360	0.1589
(r=12; n=28)	**pois**	**1**	**184.6991**	**0**	**0.4399**
	tgeo	1	184.8831	0.1840	0.4012
UNKPOA1	binm	2	175.5172	4.8365	0.0654
(r=13; n=25)	pois	1	173.2849	2.6042	0.1998
	tgeo	**1**	**170.6807**	**0**	**0.7347**
HYPRAD	binm	2	168.6391	12.8185	0.0016
(r=14; n=22)	pois	1	165.8832	10.0626	0.0065
	tgeo	**1**	**155.8206**	**0**	**0.9919**
POLMUN	binm	2	140.7537	1.9257	0.2026
(r=15; n =14)	**pois**	**1**	**138.8281**	**0**	**0.5307**
	tgeo	1	140.2046	1.3765	0.2667
PINMUR	binm	2	143.7426	3.5380	0.1022
(r=16; n=14)	pois	1	141.6007	1.3961	0.2983
	tgeo	**1**	**140.2046**	**0**	**0.5995**

Figures.

Figure 1. Map of Point Reyes National Seashore region in central coastal California, USA, showing study plot location and local distribution of Bishop Pines.

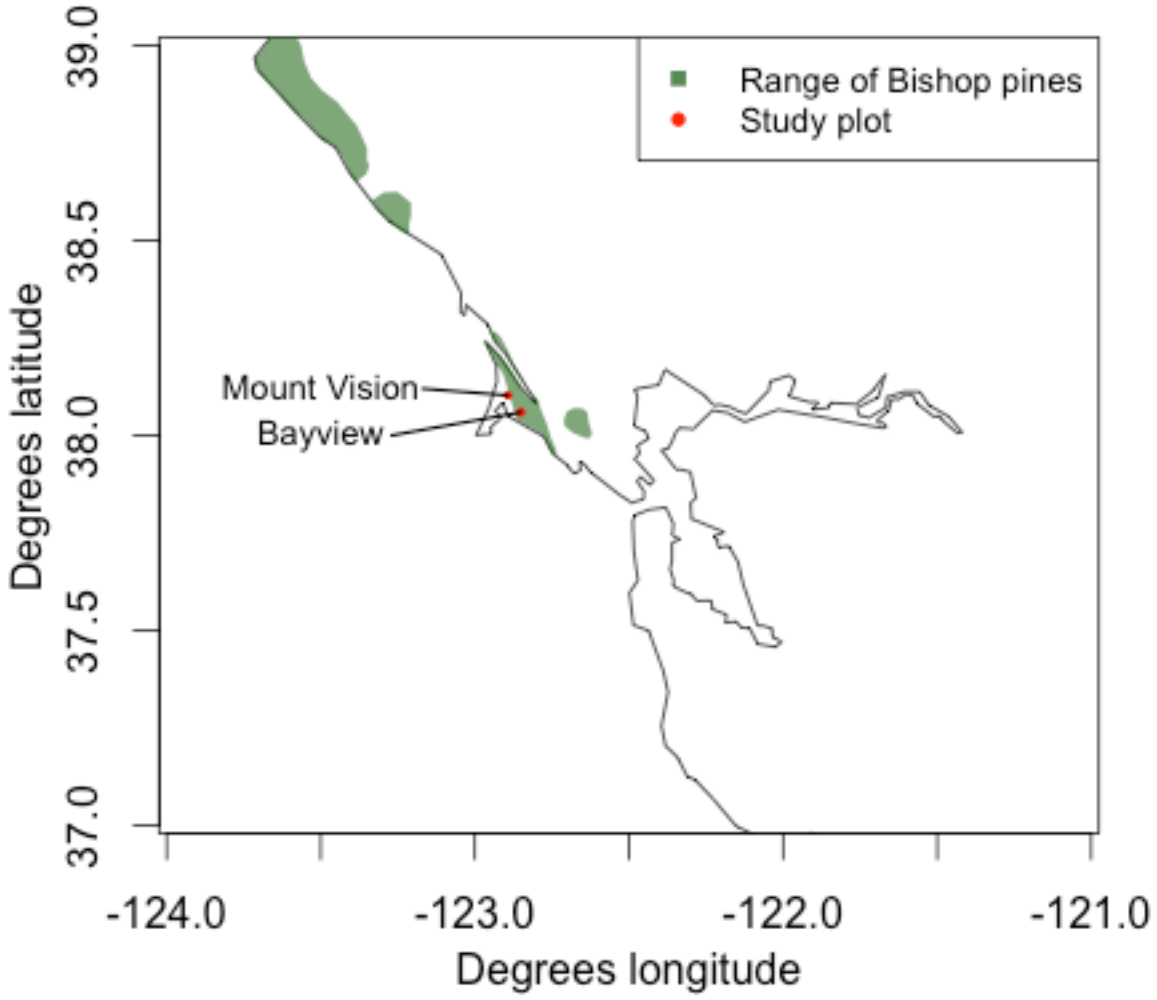


Figure 2. Photographs of (A) interior of mature Bishop pine stand at Mt. Vision site, and (B) side view of mature stand; (C) interior view of Bishop pine stand that burned in the 1995 Vision Fire at Bayview plot, and (D) exterior view of stand structure 17 years after the Vision Fire.

Figure 3. Distribution of dbh measurements in cm for *Pinus muricata* in the Bayview disturbed plot (17 years since fire) and the Mount Vision mature plot (~43 years max tree age).

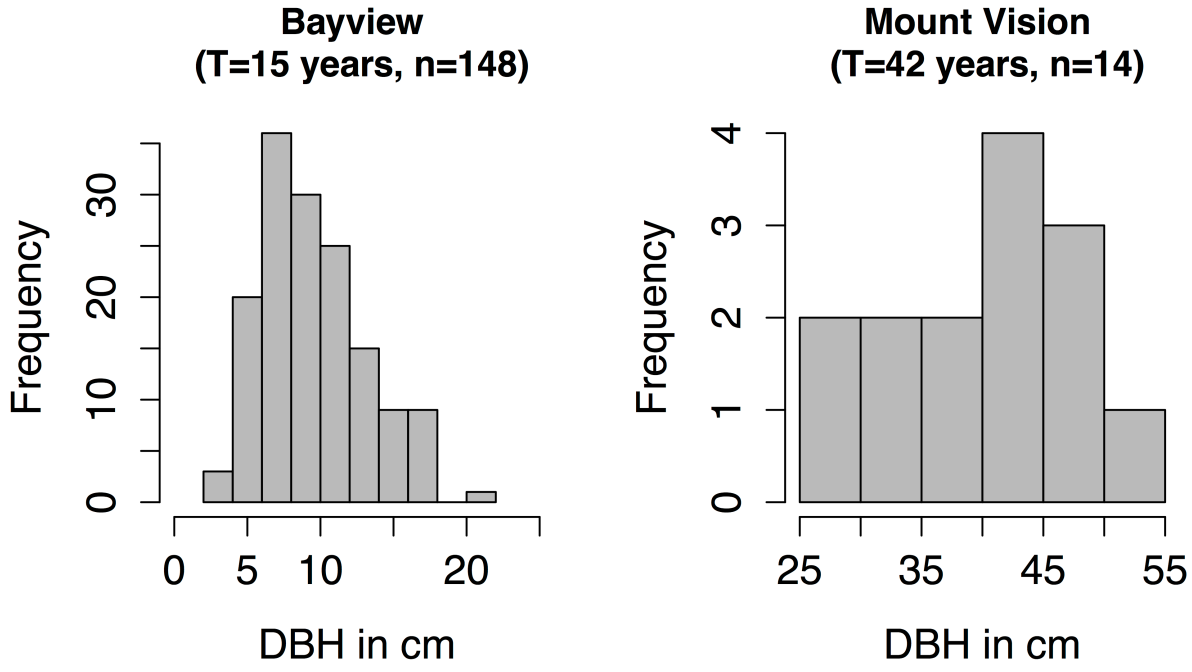


Figure 4. Species-Area Relationships for (a) the recently disturbed Bayview plot, and for (b) the mature Mount Vision plot. Empirical data are shown against the METE upper-truncated (ut) geometric prediction, and the ut binomial distribution for comparison.

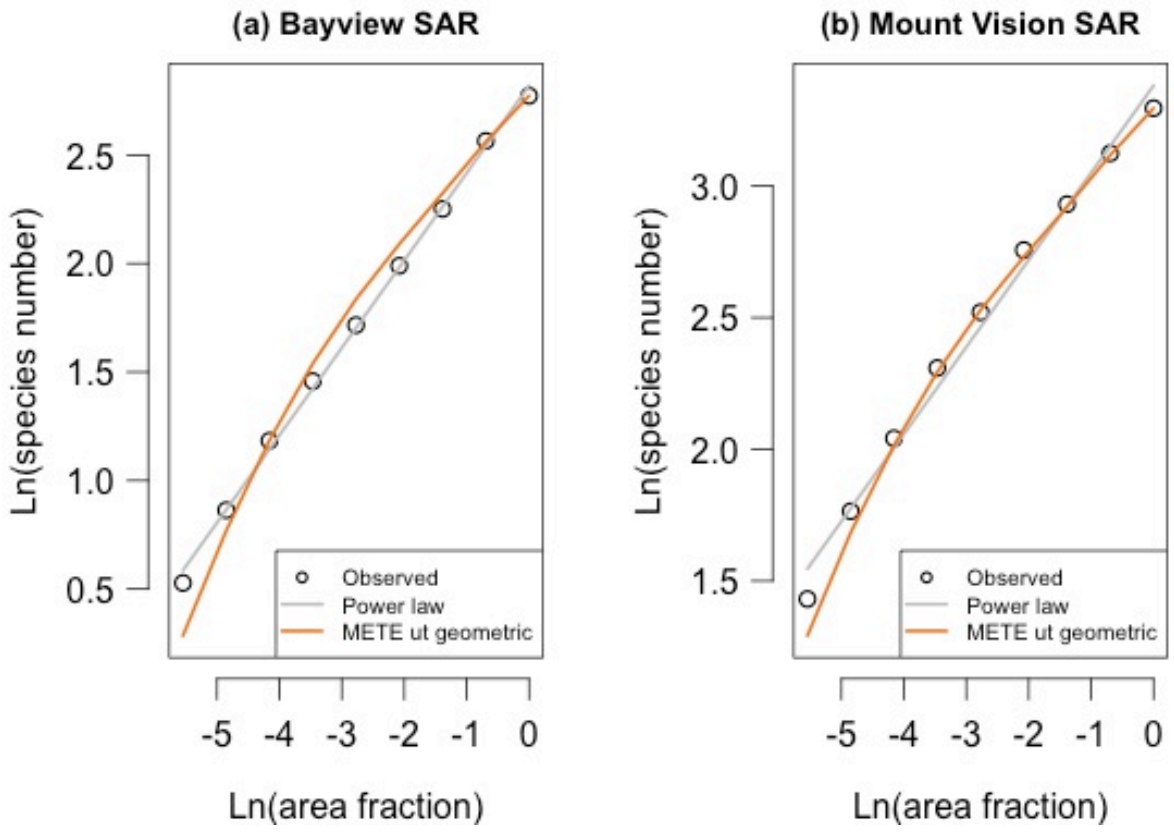


Figure 5. Universal scale collapse graphs, with METE-predicted and observed values illustrated at scales of *N/S* (total abundance/total species) for each plot each plot. The best-fit power law is shown for comparison.

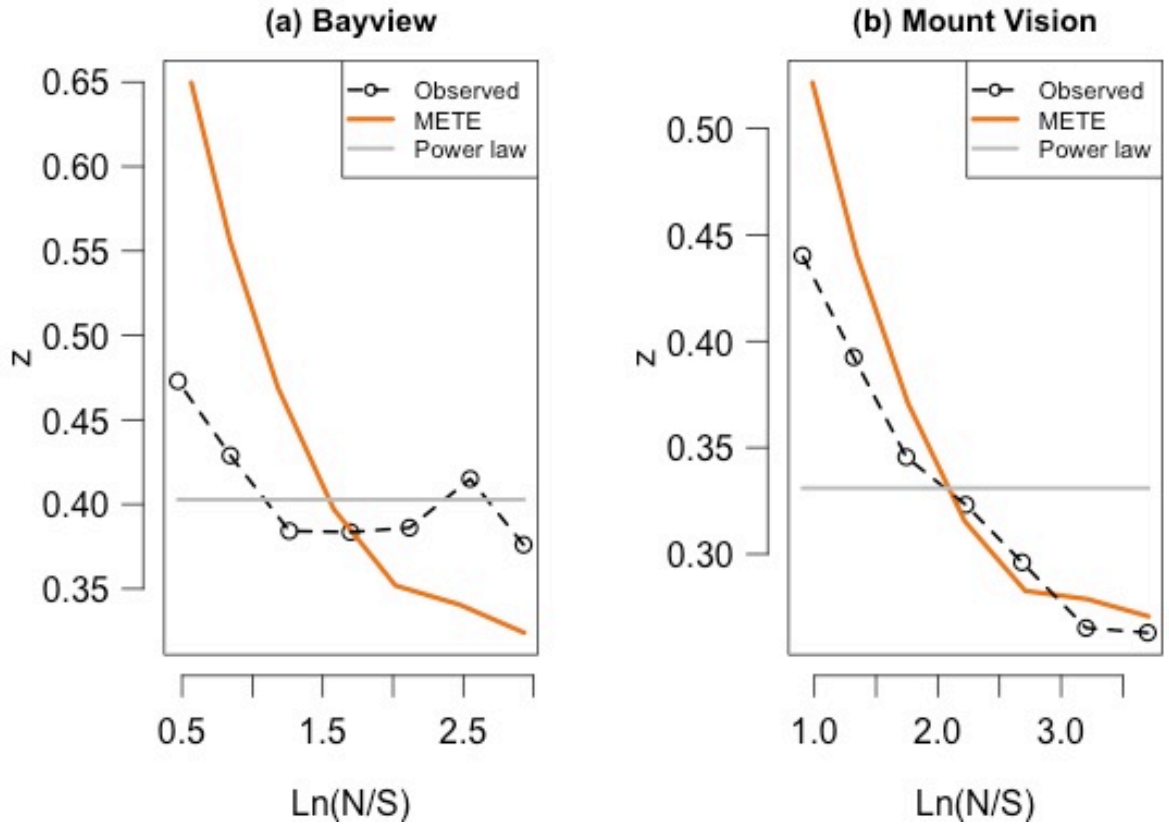


Figure 6. Empirical and METE-predicted ranked Species-Abundance Distributions and cumulative density functions for fire-adapted *Pinus muricata* stand in two plots: Bayview (a-b), the more recently disturbed, even-aged "dog-hair" stand that experienced a stand-replacing fire in the 1995 Vision fire, and (c-d) Mount Vision, the less recently disturbed, open stand mature trees with a more diverse understory.

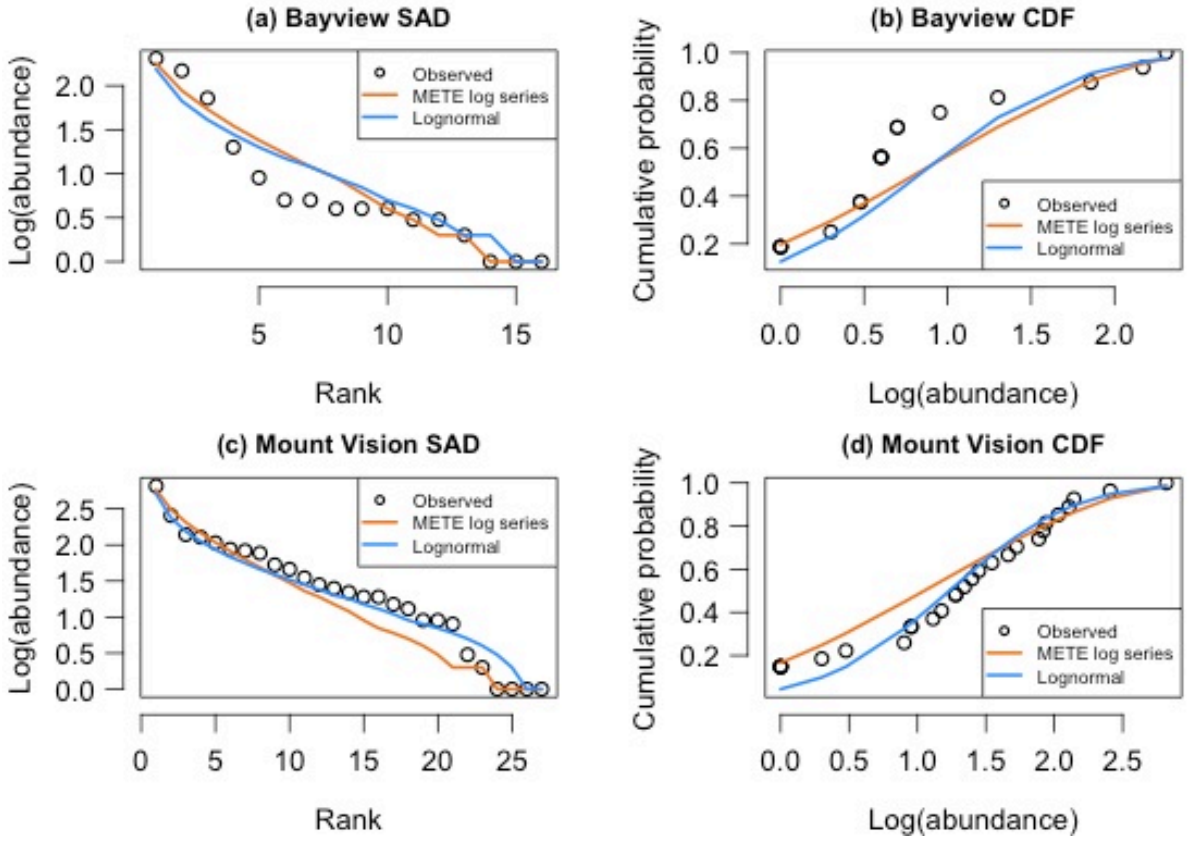


Figure 7. SSAD Example. Species shown is *Trientalis borealis* (TRIBOR) from the Mount Vision plot. In this example, binomial and Poisson distributions give the same predictions for the SSAD.

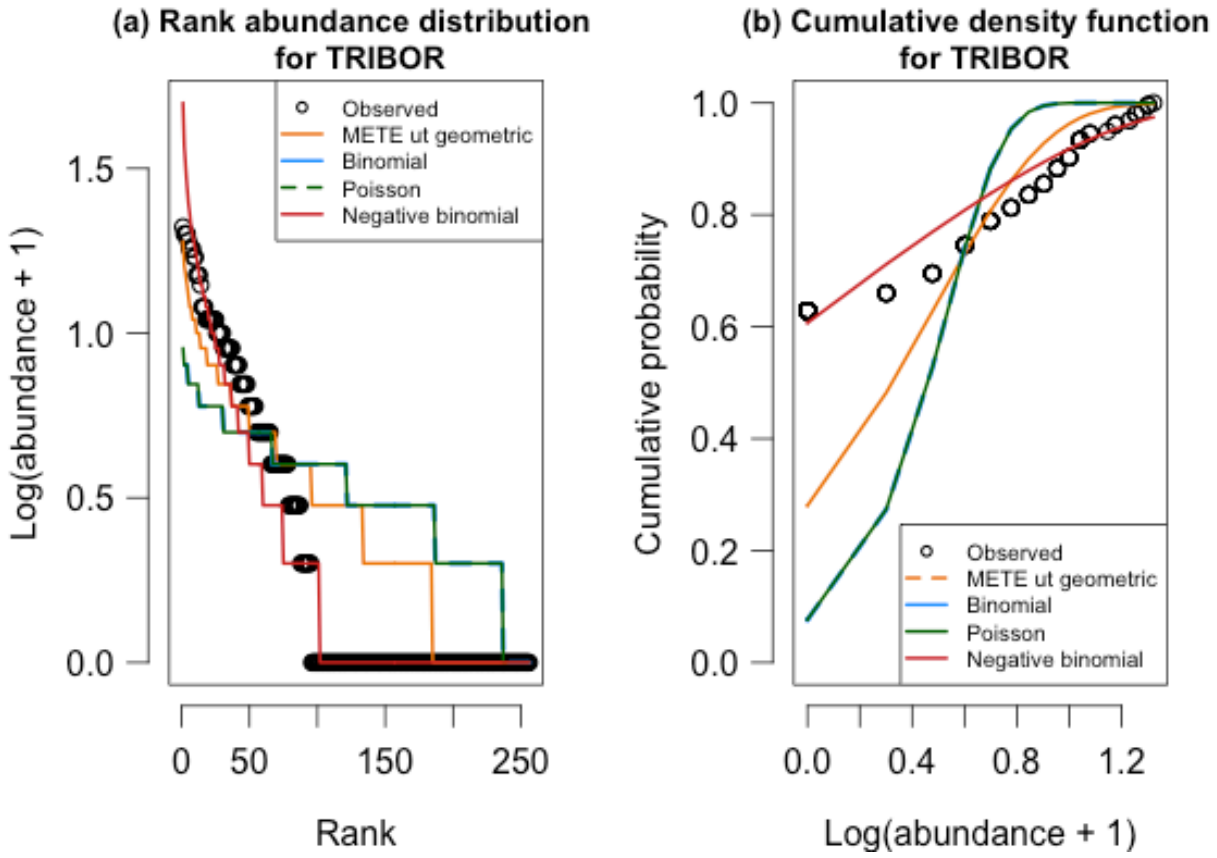


Figure 8. SSADs (shown as cumulative density functions) of the four most abundant species in the Bayview plot, which burned in the Vision Fire of 1995. Observed data are compared with a METE-predicted truncated geometric distribution and a Poisson distribution.

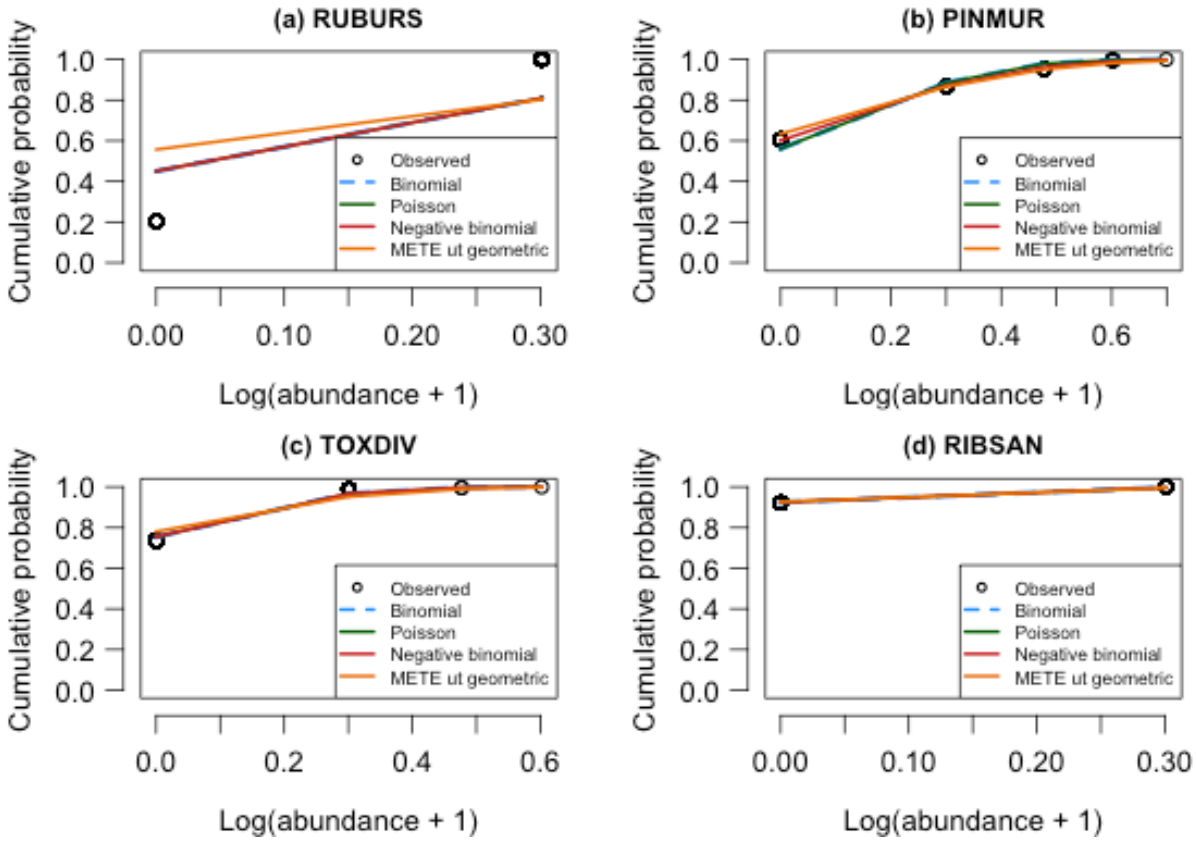
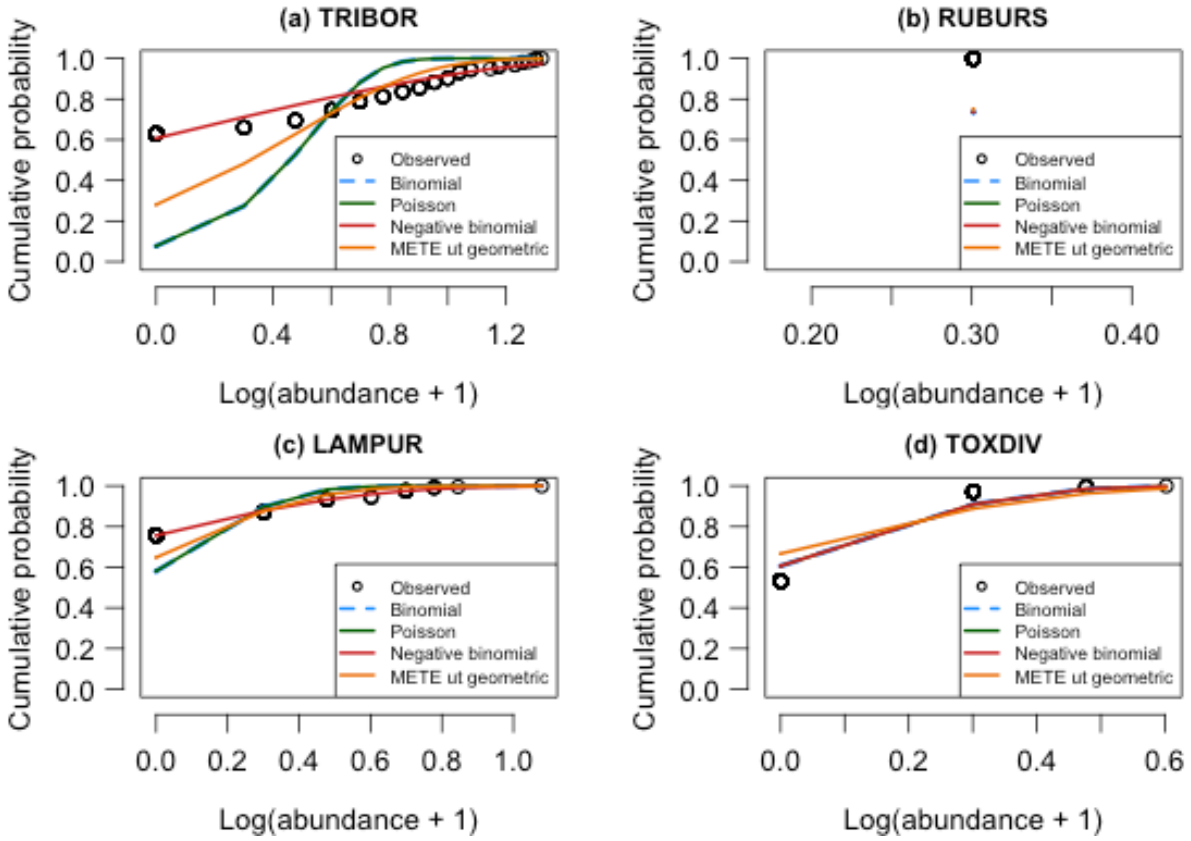
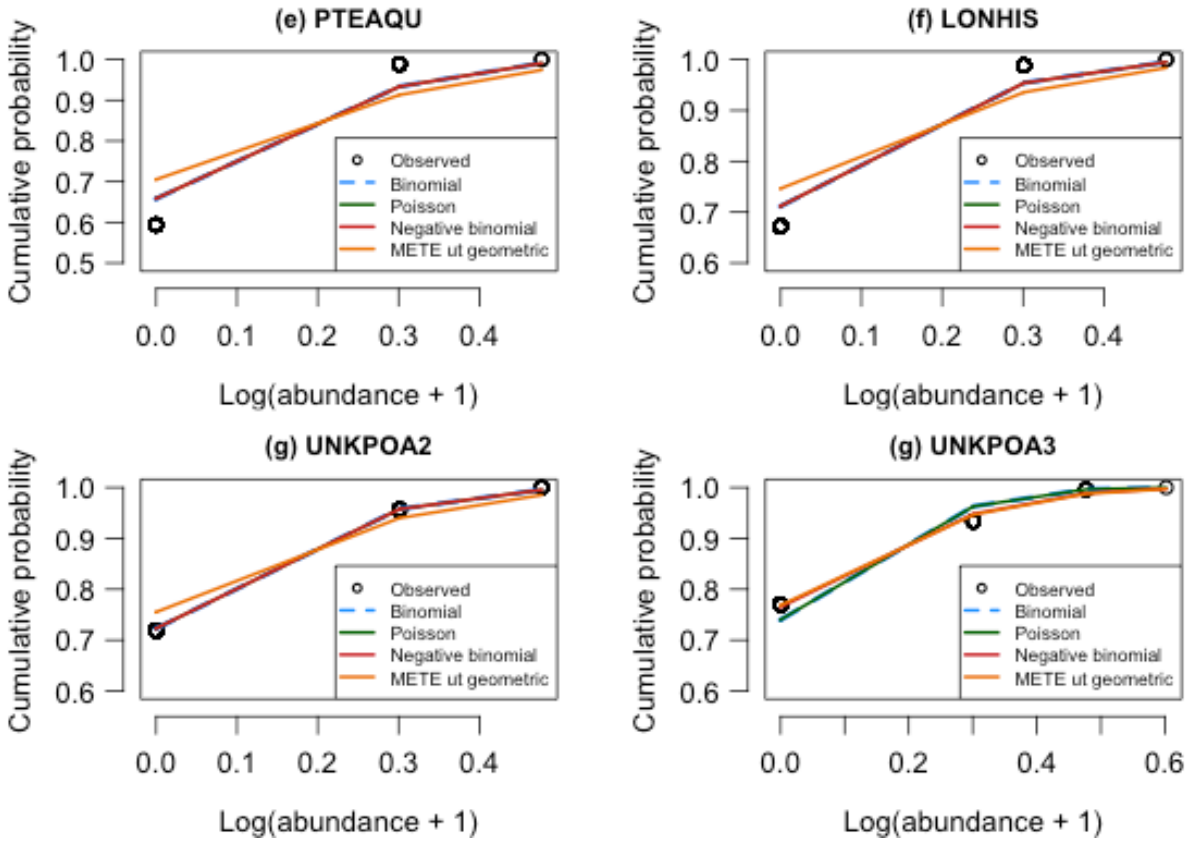
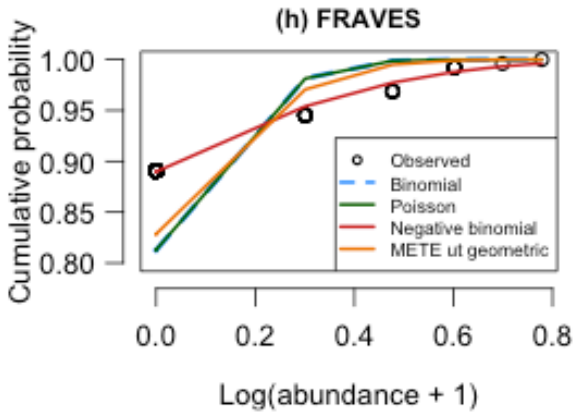
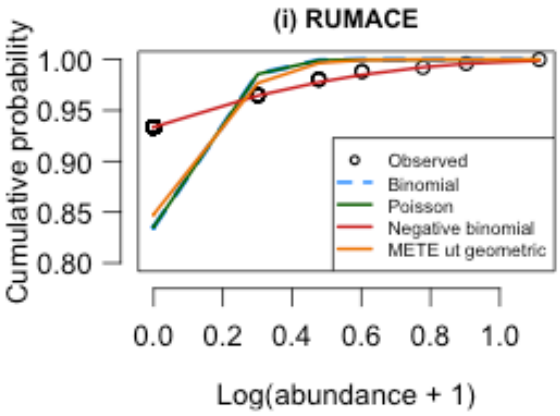


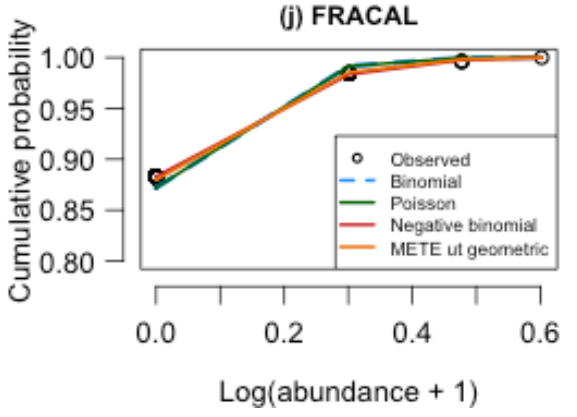
Figure 9. SSADs (shown as cumulative density functions on $\log_{10}$-transformed axes) of the most abundant species in the Mount Vision plot, a mature stand. Observed data are compared with a METE-predicted truncated geometric distribution and a Poisson distribution.

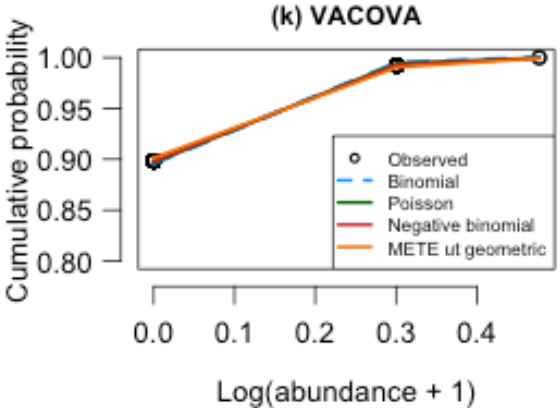


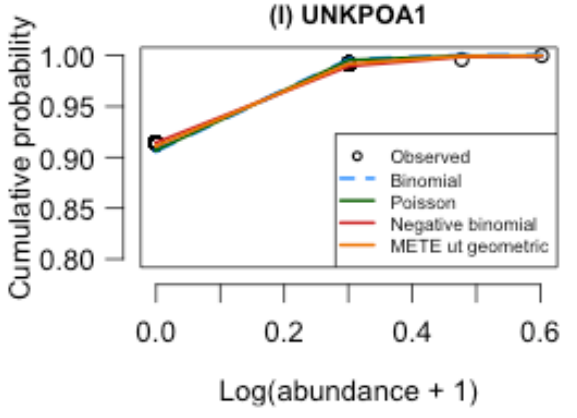

(e) PTEAQU
(f) LONHIS
(g) UNKPOA2
(g) UNKPOA3
Cumulative probability
Log(abundance + 1)
Observed
Binomial
Poisson
Negative binomial
METE ut geometric


(h) FRAVES
Cumulative probability
1.00
0.95
0.90
0.85
0.80
0.0
0.2
0.4
0.6
0.8
Log(abundance + 1)
Observed
Binomial
Poisson
Negative binomial
METE ut geometric


(i) RUMACE
Cumulative probability
1.00
0.95
0.90
0.85
0.80
0.0
0.2
0.4
0.6
0.8
1.0
Log(abundance + 1)
Observed
Binomial
Poisson
Negative binomial
METE ut geometric


(j) FRACAL
Cumulative probability
1.00
0.95
0.90
0.85
0.80
0.0
0.2
0.4
0.6
Log(abundance + 1)
Observed
Binomial
Poisson
Negative binomial
METE ut geometric


(k) VACOVA
Cumulative probability
1.00
0.95
0.90
0.85
0.80
0.0
0.1
0.2
0.3
0.4
Log(abundance + 1)
Observed
Binomial
Poisson
Negative binomial
METE ut geometric


(l) UNKPOA1
Cumulative probability
1.00
0.95
0.90
0.85
0.80
0.0
0.2
0.4
0.6
Log(abundance + 1)
Observed
Binomial
Poisson
Negative binomial
METE ut geometric

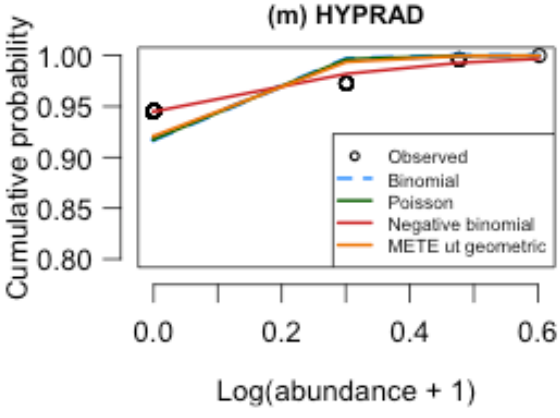
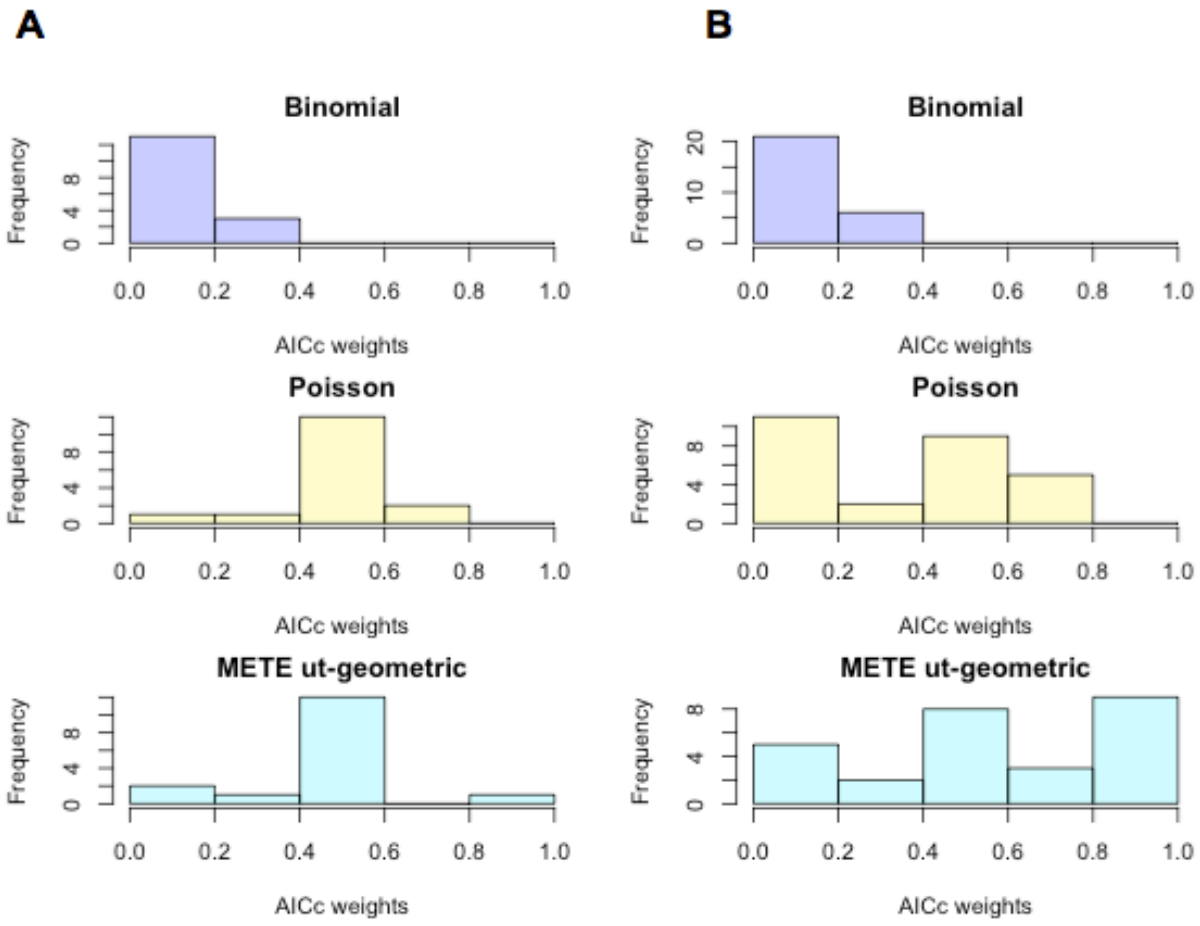

(m) HYPRAD
Cumulative probability
1.00
0.95
0.90
0.85
0.80
0.0
0.2
0.4
0.6
Log(abundance + 1)
Observed
Binomial
Poisson
Negative binomial
METE ut geometric

Figure 10. Histograms comparing SSAD models for all species by their AIC*c* weights. By column: (A) Bayview plot (site of Vision Fire) and (B) Mount Vision mature plot. Higher AIC*c* weights indicate better model fits.



CHAPTER 2

Macroecology for management: Testing an information-entropy-based theory of macroecology against anthropogenic disruption of high-Sierra meadows

Chapter 2: Macroecology for management: Testing an information-entropy-based theory of macroecology against anthropogenic disruption of high-Sierra meadows

Abstract

Anthropogenically-induced ecological disruptions (anthropogenic disruptions) have been overlooked by macroecological theory because they represent ecosystems in various states of transition that result from non-natural selection on the community. While critically important to understand for conservation reasons, anthropogenic disruptions are, in general, not comparable to each other, nor to other ecological disturbances that are natural in origin. The Maximum Entropy Theory of Ecology (METE) has demonstrated general applicability across a variety of systems, and including plant, arthropod and vertebrate taxa. Here, we use METE to examine the effects of an anthropogenically-induced novel disturbance regime of grazing by horses in high Sierra Nevada meadows on the species-abundance distributions (SAD), number of singleton species, and the species-level spatial abundance distributions (SSADs) (a measure of spatial aggregation) for all species in three pairs of grazed and ungrazed meadows, each meadow containing a system of plots set up across a moisture gradient. We find that number of singleton species may be a better indicator of ecological disruption than the shape of the SAD in systems where the differences in community structure are subtle. We also find that the METE SSAD performs better than all other models tested for both grazed and ungrazed plots. We suggest ways of augmenting tests of the METE SSAD to refine theory for management relevance.

Key words

Subalpine meadows, Sierra ecosystems, ecological perturbation, anthropogenic disruption, horse grazing, pack stock, species-abundance distributions, spatial distributions

Introduction

Macroecology, as a discipline, has historically focused on questions of species diversity, abundances, and how plot-level data can be scaled up to make predictions about diversity patterns at the scale of ecosystems or entire biomes (Brown 1995, Gaston and Blackburn 2008, Harte 2011). In contrast, many management-oriented studies have begun to emphasize ecosystem attributes such as dynamic landscapes and patchiness (Noss 1983, Pickett and White 1985, Pickett and Rogers 1997), which have been largely overlooked by macroecology (Fisher et al. 2010) despite occurring at spatial scales that have been the focus of much macroecological research. Adapting macroecology for biodiversity conservation and management at plot-to-landscape scales may be desirable because macroecological predictions scale well over space by design, and many macroecological metrics of interest can be measured with a simple census grid. In principle, macroecological methods would be able to reveal that ecosystems are lacking rare species that they should have, or differ in structure and composition from more desirable states (such as the differences between primary and secondary forests, or vegetation communities differing in their degree of invasion by non-native species). Ultimately, macroecological metrics, in aggregate, may yield a "signature of disturbance" or "signature of ecological disruption" in ecosystems that may be similar from ecosystem to ecosystem, and provide insight into how macroecological patterns change together under a variety of stressors.

A recently developed macroecological theory, the Maximum Entropy Theory of Ecology (METE), shows great potential as a framework for revealing patterns of biodiversity in a form that would be useful in support of land management decision-making. METE has been successful in making accurate predictions of many macroecological metrics, in a variety of ecosystems and for multiple taxa. These metrics include those we will examine in this chapter, all of which may be relevant to management issues: the species-abundance distribution (SAD; Harte et al. 2008, Harte 2011, White et al. 2012, Newman et al. 2014), the numbers of singleton species in study plots (or "singletons"; referring to those species represented by a single individual; Harte 2011, White et al. 2012, Harte and Kitzes 2015), and the species-level spatial abundance distribution (SSAD) of each species in each study area (Harte et al. 2008, Harte 2011, Newman et al. *manuscript* [Chapter 1 of this dissertation]). The SAD (and resulting measures of rarity) and SAR are metrics that measure biodiversity, while the SSAD is a measure of spatial clustering of individuals of a species in a given area. We choose these metrics from a large number of possible ones that could be calculated with the same methods because they are independent of area, and instead only rely on two of METE's "state variables": N, the total abundance of individuals, and S, the total number of species in the study plot (see Harte et al. 2008, Harte et al. 2009, Harte 2011, Harte and Newman 2014).

A small number of studies have shown that deviations from METE occur. Some of these deviations suggest that METE may be used to distinguish between steady-state and systems in transition. Others do not appear to be related to processes that may be of interest to biodiversity management, as not all deviations from METE might be considered a "signatures of disturbance" (some deviations may occur due to incomplete sampling, for example). Previous studies have demonstrated that recent natural ecological disturbance in a fire-adapted forest (Newman et al. manuscript [Chapter 1]), drought in Rocky Mountain meadows (Newman et al. 2014 [Chapter 3]), and isolation of a rainforest community at Barro Colorado Island in Panama resulting from construction all have SADs that show marked deviations from the METE-predicted Fisher log-series distribution (Harte 2011). The shape of the SAD is important in ecosystem and species-level management decisions because it has direct relevance to how many rare species can be maintained in an area. Similar studies focused on dominance and evenness in local communities describe these mediate inter- and intraspecific interactions with effects on coexistence and maintenance of rare species (Hillebrand et al. 2008).

For small scaling plots (<1 ha) to be useful as a macroecological tool, their results must distinguish between different ecological states, and either individual metrics or metrics in combination with one another must reveal a signature of disturbance or other ecological perturbation, that changes regularly between impacted and unimpacted plots. Here, we use an extensive novel dataset to examine fits to and deviations from METE in Sierra subalpine meadows. We focus in particular on the shape of the species-abundance distributions, predicted numbers singleton species, and the functional form of SSADs as determinants of how well METE performs in and distinguishes between unimpacted and heavily perturbed patches or communities within ecosystems. We compare community structure at the plot-level in subalpine Sierra Nevada meadows that have a recent history of grazing by horses (grazed plots) to those with no recorded history of use by horses or other "pack stock" animals non-native to the ecosystem (ungrazed plots). Effects of pack stock use have been indicated as drivers of ecological change in this ecosystem (DeBenedetti & Parsons 1979, McClaran and Cole 1993, Ostoja et al. 2014).

Because the METE SAD and SSAD work extremely well for natural systems in large-scale contexts (≥50 ha) and many of its metrics are explicitly designed to scale with area, we expect that METE will also make accurate predictions for the majority of unimpacted (ungrazed) small-scale plots we measure. The metrics we choose to focus on here are independent of area, and small-scale plots have the benefit of being replicable and easy to construct. However, small-scale plots have the potential downside of not capturing beta diversity and landscape heterogeneity well, which might limit the usefulness of METE's predictions. Based on previous cases of deviations from METE's predictions, we expect that METE will be more accurate in its predictions of metrics in ungrazed plots than grazed plots, which would mean that results between plots exposed to different environmental stressors would be distinguishable. If this is the case, there is potential for small-scale plots to be a useful management tool. Specifically, we predict (1) grazed plots' SADs will be more lognormal than their ungrazed comparison or control plots, regardless of moisture level (because grazed plots represent systems in transition, often associated with the lognormal SAD; Kempton and Taylor 1974, Bazzaz 1975, Gray and Mirza 1979), and therefore (2) grazed plots will contain fewer singleton species than their ungrazed comparison plots. METE should therefore (3) accurately predict the number singleton species within the ungrazed plots, but overestimate singletons in the more lognormal grazed plots. Finally, (4) METE's SSAD predictions should be the best fit in ungrazed plots when compared to other models, but may not be the best model in aggregate for all species when tested in grazed plots.

Methods

Site descriptions

Sierra Nevada subalpine meadows are found between the latitudes of 35-40^{o} N in eastern California and western Nevada, USA. These meadows generally occur in an elevation range of 2,450–3,600 m (Fites-Kaufman et al. 2007), and are characterized by their diverse, herbaceous flora, often comprising multiple plant communities, as well as two abiotic conditions: a shallow water table (generally <1m) and fine-textured soils (Weixelman et al. 2011). Occupying only ~3% of the land area (Keeler-Wolf et al. 2012, Viers et al. 2013), meadows are host to the majority of biodiversity in the Sierra Nevada (Ostoja et al. 2014). Water is known to be a limiting factor in Sierran meadows (Benedict 1983, Ratliff 1985, Fites-Kaufman et al. 2007, Berlow et al. 2003), and is a key determinant of plant community composition (Halpern 1985, Allen-Diaz 1991) and diversity (Benedict 1983). Sierra Nevada meadows are subject to multiple stressors (Lee 2013) affecting soil moisture levels, at least two of which are major conservation concerns, the first being a long-term drying trend (Darrouzet-Nardi et al. 2006) with exceptionally dry recent years (Dai 2013), and the second being animal grazing, including pack stock use. Agricultural grazing by sheep, cows, and horses has occurred since 1862, with use for sheep discontinued by 1900 (DeBenedetti & Parsons, 1979, Odion et al. 1988). Shortly after its inception, agricultural grazing by cows was described as having strikingly obvious effects on Sierran meadows, specifically denuding, degrading, and eroding them (Odion et al. 1988). The effects of pack stock animals (horses, mules and burros) for travel and recreation are not as clear. Recently, many studies have highlighted a complex impact of pack stock on meadows, such as an increase in bare ground (Moore et al. 2000, Cole et al. 2004, Lee 2013), a decrease in productivity (Cole et al. 2004) and in soil moisture (Shryock 2010, with similar findings for

cattle grazing in Berlow et al. 2002), and an increase in the incursion of woody species (D'Antonio et al. 2004, Darrouzet-Nardi et al. 2006). Meadow responses to grazing and pack stock use also vary with within-meadow water gradients (Lee *in prep.*), and drier local vegetation communities have lower resilience to these perturbations (Génin *in prep*.). Pack stock use has also been shown to have indirect effects including changes in soil chemistry; these and many other effects are reviewed in Ostoja et al. (2014). However, some studies show little or no effects of pack stock use in the longer term (on plant assemblages: Hopkinson et al. 2013, and arthropod communities: Holmquist et al. 2010). This may be a result of the inherent variability of meadows (Allen-Diaz 1991), and the dominant control of large-scale climatic and hydrologic factors (Wood 1975), as well more local processes. Due to this complex response to ecological processes and perturbations, Sierra Nevada meadows are an ideal system for using novel methods to assess the effect of perturbations on ecosystems.

Our study sites were located in subalpine meadows in Sequoia National Park in Tulare County, California (36°33'53"N, 118°46'24"W). Figure 1 shows study meadow locations within the Sequoia & Kings Canyon National Parks management unit (SEKI). Our study meadows were surrounded by conifer forests typical to this region (Fites-Kaufman et al. 2007), dominated by Lodgepole pine (*Pinus contorta*), also containing a mixture of Whitebark pine (*Pinus albicaulis*), Foxtail pine (*Pinus balfouriana*), and Sierra juniper (*Juniperus occidentalis* var. *australis*).

Pack stock use and plant community comparisons

Study meadows were chosen from a subset of a comprehensive National Parks Inventory of Sierra Nevada meadows (Berlow et al. 2013, Pyrooz 2015) based on aerial imagery, with pack stock use records that are compiled and maintained by the National Park Service (Holmquist et al. 2010, Matchett et al. 2015), and through expert opinion of the National Plant Ecologist, also in charge of pack stock monitoring in SEKI (Sylvia Haultain). In particular, our study meadows were pulled from a larger set of paired meadows identified in a concurrent study to have matching characteristics (through a method of optimizing Mahalanobis generalized distances), according to 27 geospatial, hydro-climatic, and vegetation covariates (Lee 2013), including pack stock use history, elevation, and proximity. After this initial selection of paired meadows, plots were placed within a meadow by matching the three species with highest percent cover between grazed and ungrazed comparisons at the same estimated soil moisture level. Relative moisture levels of plots were qualitatively estimated by soil composition and texture in three categories: mesic, mesic transition and xeric. Soil water content could have been measured in the field, but because soil moisture fluctuates over the course of a day, between precipitation events, and between years in these meadows, any measurements made over the course of one season in a single meadow would not adequately characterize average soil moisture levels. We therefore used a comparative approach between our data and an existing, longer-term data set that includes measurements of soil volumetric water content (VWC; Lee 2013, Lee *in prep*.) over two years and on a larger number of meadows (47). In the Lee (2013) study, VWC was measured at 12cm depth with a soil probe in an array of plots across individual meadows, and plot means were converted to a standardized measure to minimize potential effects of intra- and inter-annual variation in soil moisture between meadows.

For each plot in this study, we estimated the underlying moisture level that would have produced the plant communities we measured by comparing our data set with data available from

Lee (2013) for many of the same meadows. This was accomplished by merging the two datasets and using a non-metric multidimensional scaling (NMDS) to ordinate the combined data on a two-axis plane. NMDS is an ordination procedure that projects sites (species composition in one cell) onto a Euclidian plane while conserving to an extent the rank of dissimilarities between these sites. It extends classic ordination methods (e.g. CCA, PCA) by improving robustness to outliers and zeros, and allows the use of other dissimilarities beyond classic Euclidian distance that are known to pick up better the variations of vegetation communities along environmental gradients. NMDS was performed in R version 3.2.1 (R Core Team 2015), using functions provided by the R package *vegan* (Oksanen *et al.* 2015), using the square-root-transformed covers of the 10 most abundant species and a Bray-Curtis distance for the dissimilarity matrix.

Based on the ordination results, a 2D spline surface with VWC as the response variable was fit on the ordination results (effectively equivalent to fitting a Generalized Additive Model with VWC as response variable and the two axes ordination scores as explanatory variables). This allowed us estimate soil water content based on the species content for all sites of the ordination including sites where no VWC was measured (this work's dataset).

Data collection

Three sets of meadows were chosen as grazed and ungrazed pairs (those with known history of pack stocking, and those without), for a total of six study meadows spanning 22 km. Grazed and ungrazed meadows were selected to be within ≤3.5 km of one another, have similar elevations and size, and similar plant communities, differing mainly in intensity of use by pack stock animals (horses and mules), which are not native to the continent. Within each meadow, a series of three plots were set up at increasing distance from a stream (but within 100 meters of the stream edge), in order to capture a moisture gradient, for a total of 18 plots in six study meadows. Originally, the moisture gradient was conceived of as three moisture levels, from moist to dry: mesic, mesic transition, and xeric. These were later reclassified into 12 "intermediate" moisture-level plots (comprising mesic and mesic transition), and 6 "dry" plots (comprising the original xeric plots) based on plant community analyses; see Results. Plant communities in plots within each meadow were chosen in the field to be comparable between grazed and ungrazed meadows by matching the three most dominant species at each moisture level.

Data collection was carried out in the spring and summer of 2012. The three plots in each study meadow measured 4m x 4m (16 m^2) and were subdivided into 64 cells, each measuring 0.5 m x 0.5 m (0.25m^2). Two observers counted all live, aboveground plants above 1 cm. Observers estimated percent cover of all live plants by species, and bare ground, by cell in each plot. Bare was measured in an absolute range between 0-100%, but cover measurements of live plants was allowed to exceed 100% if plants overlapped in a top-down perspective (relative cover). Every species present in each cell as well as all species' abundances were recorded. Clumps and clusters of some grasses and other forb species were considered "individuals" in these surveys; this is permitted within the assumptions of METE, as any "individual" will scale up to more of the same type of count over space, i.e. counts of single stems will scale up to estimates of stems; counts of grass clumps will scale up to estimates of grass clumps. Plants that could not be identified to species level were photographed and described, and given a unique "unknown

species" designation for analysis. METE's predictions are tolerant to some unintentional lumping and splitting of taxa included in input data (Harte et al. 2013).

Macroecological metrics

We examine three predictions of METE in the context of grazed and ungrazed Sierra meadows: the SAD, and a discrete portion of the SAD: the number singleton species; and the SSAD for each species in each plot. Mathematical forms of these distributions are available from multiple sources (Chapters 1 and 3 of this dissertation, Harte et al. 2008, Harte 2011, Newman et al. 2013, Harte and Kitzes 2015 for SAD and abundance-related metrics; Chapter 1, Harte 2011 for the SSAD). As mentioned in the Introduction, we expect these metrics to change in predictable ways with ecosystem disruption, and potentially give differing patterns between grazed and ungrazed plots.

For fully-nested census data (as measured with gridded plots used in this study), the species abundance distribution (SAD), METE predicts the Fisher log-series distribution, which is a distribution with one parameter. We compare this to predictions of the lognormal distribution, with 2 parameters. We refer to the predictions of SSADs by METE as an upper-truncated (ut) geometric distribution, which models the ranked distribution of abundances of a species across cells in a plot. The ut-geometric has one parameter that constrains the shape of the distribution (and in the case of METE, this same parameter determines the point of upper truncation), and is identical to a negative binomial distribution with the shape parameter = 1. This shape parameter, called k, is not the same k as represents the number of parameters of a distribution, required for model comparisons. For clarity, we will refer to the shape parameter as k(nbd) in later discussion.

Analyses for METE SAD and SSAD predictions and comparison distributions were performed in Python (van Rossum, 2001) in the open-source project "Macroeco" (Kitzes et al. 2014, Kitzes and Wilber 2016). SAD and SSAD scripts from the beta version of this software (accessed June 2015) were used in this manuscript. Models for SADs and SSADs were compared using Akaike's Information Criterion value corrected for small sample sizes (AICc). Singleton species predictions were calculated in Mathcad (version 7.0), using the exact expression for calculating METE's Lagrange multipliers and term β as shown in equation (7.27) in Harte (2011).

Results

Pack stock use and plant community comparisons

Pack stock use in our study sites was available from 2004-2009. Use ranged from 527 stock-nights (number of animals*number of nights in a meadow) in a 5.2-ha meadow (meadow 4313), 635 stock-nights in a 4.7 ha meadow (meadow 3597), to 881 stock-nights in a 1.2 ha meadow (meadow 9054). Eighteen gridded plots were surveyed in this study, comprising 3 plots at different target moisture levels (mesic, mesic transition, and xeric) in 3 pairs of ungrazed and grazed meadows (at Rock Creek, Tyndall Creek, and Crabtree). Fifty-nine species present in all plots are listed in Table 1. In total, 100,491 individual plants were counted. Details of individual

plots, including geographic locations, number of species and number of individuals counted, and absolute cover values (compared to bare ground) are available in Table 2. Photographs of example cells from plots, and plots situated in their larger environment are shown in Fig. 2 and Fig. 3, respectively. Adjusting plant cover measured in each plot to "relative cover" equaling 100%, the dominant species with respect to relative cover are visualized in Figure 4, where we show these values for the 10 most dominant species by cover for each plot, arranged by moisture category. Vertical columns within graphs denote 64 cells within the plot, which are arranged by similarity to one another. We note the dominance of *Calamagrostis breweri* in most of the mesic and mesic-transition plots, and some of the xeric plots as well. Xeric plots contain a large amount of cover by *Carex filifolia*, which is much less dominant or completely absent from the moister plots.

We performed NMDS analyses on the species present and their abundances in each plot, and compared these results to a larger dataset (Lee 2013, Lee et al. *in prep*.) in order to estimate standardized volumetric water content for our plots (Fig. 5). From these analyses, we were able to group our plots into two categories: "intermediate" and "dry" moisture levels, corresponding to a previous description of the Sierra meadow plant community (Lee 2013). These categories correspond exactly to the original categories of mesic and mesic transition (both in intermediate) and xeric (in the new category of dry). All further analyses use this categorization.

Macroecological comparisons

The SAD was evaluated for each plot. Two candidate models were considered: the METE log series, and the log normal. We assessed model fits with AIC*c* values AIC*c* weights (Table 3). When collectively considering all plots, the METE distribution is a better fit than the log normal in 11 of 18 cases. "Wins," or best model fits assessed by AIC*c* weights, were considered grouping plots by grazing history, by moisture level, and then by both factors. Pearson's Chi-squared tests show that a lack of statistical significance in the proportion of wins by either METE or the lognormal distribution for any grouping of results (Fig. 6).

Considering only the right tail of the SAD, we expected METE to make highly accurate predictions of the number singleton species in ungrazed plots, and make less accurate predictions in grazed plots. We find this to be the case. Regression lines imposed on ln-ln graphs of predicted versus observed singleton species are shown in Fig. 7 for ungrazed and grazed plots. A regression line for ungrazed plots has a slope of 0.693 and an R^2 value of 0.598. For grazed plots, the slope of the regression line is 0.263 and its corresponding R^2 value is 0.317. A slope and R^2 value equal to 1 would indicate a perfect fit to data. Here, the proportion of variance explained in the ungrazed plots is about twice that explained in grazed plots.

SSADs were generated for every species in all 18 study plots. To compare model fits, AIC*c* values and AIC*c* weights were calculated for each of 3 models tested against data: the binomial, the Poisson, and the METE ut-geometric distributions. Following Baldridge et al. (2015), we consider the model with the highest AIC*c* weight to be the best-fit model, and compare the distribution of AIC*c* weights by model. We repeat this for intermediate ungrazed and grazed plots (Fig. 8) and xeric ungrazed and grazed plots (Fig. 9). We find that the METE ut-geometric distribution is the best model, based on multiple trials, to describe the SSADs of species for both moisture levels and both grazed and ungrazed plots. For ungrazed plots, AICc

models support METE's prediction for 151 out of 174 cases; for grazed plots, METE has 149 wins of 170 cases (Poisson distribution has 23/174 and 21/170 wins for ungrazed and grazed plots, respectively, while the binomial distribution has 0 wins in both cases). The distributions of AIC*c* weights of all models are remarkably similar for all categories of plots.

Discussion

Because of their high ecological variability (Allen-Diaz 1991) and their complex responses to ecological processes including disruptions and perturbations (Wood 1975), Sierra Nevada meadows are a prime system for testing innovative methods assessing the effects of disturbance and disruption on ecological systems, such as methods presented in this work. The full effects of pack stock use in Sierra Nevada meadows are difficult to measure because soils, plant and animal communities are differentially affected (DeBenedetti & Parsons 1979, McClaran and Cole 1993, Holmquist et al. 2010, Hopkinson et al. 2013, Ostoja et al. 2014), and effects may vary from meadow to meadow. Effects of pack stock use on rare species are statistically complicated to determine (Matchett et al. 2015) and need to be assessed on a species-by-species basis for conservation needs to be addressed. Here, we find that small (<1 ha) census plots can be used to measure community-level metrics, and that these plots can produce replicable results related to the presence of singleton species in a given area. Specifically we find that grazing by pack stock animals reduces the number of singleton species compared to ungrazed areas. It appears that pack stock grazing reduces the overall diversity and persistence of rare species in a local community.

Returning to our original hypotheses, we find mixed support for them. We did not find that grazed plots had more lognormal SADs than their ungrazed counterparts, but instead found highly variable results. METE performed at least as well as the lognormal, or better for all categories and all groupings of data, by moisture level and pack stock use. Elsewhere in macroecology, the lognormal distribution has previously been associated with systems in transition (Kempton and Taylor 1974, Whittaker 1975, Bazzaz 1975, Gray and Mirza 1979, Chapter 1 of this dissertation) and we therefore thought it might be a good model for anthropogenically-disrupted (i.e. grazed) plots. However, if the lognormal works as a descriptor of ecosystems in transition, then these results might imply that the new, grazed communities are more stable that we thought, and may be in a semi-permanent state of lower productivity (based on absolute cover measurements of bare ground). Alternately, the lognormal may be a good descriptor of temperate systems with many factors determining community structure, for reasons not related to the system being in a state of transition.

As hypothesized, we found that grazed plots contain fewer singleton species compared to ungrazed comparison plots, and that METE more accurately predicts the number of singleton species within the ungrazed plots than the grazed ones. However, METE generally overestimates singletons in both ungrazed and grazed plots. Given the extreme accuracy of predictions of singleton species among trees and arthropod communities (Harte and Kitzes 2015) and bird communities (White et al. 2012) from larger datasets, this result suggests that the number of singleton species shows promise as a discriminant of level of ecological disturbance or disruption in plots. This metric would be more powerful as a macroecological indicator if it varied with another indicator (like the METE or lognormal SAD) in predictable ways. By itself,

the number of singletons in this study (being overestimated by METE), may indicate that both sets of plots (grazed and ungrazed) are subject to larger-scale perturbations impacting overall community structure, like the ongoing drought in California during the year we conducted this survey, that swamps out the local differences between plot-types. This kind of hierarchical response to drivers of soil moisture has been mentioned previously in the literature (Allen-Diaz 1991, Wood 1975).

We also hypothesized that METE's SSAD should be the best fit model for species in ungrazed plots, but perform as well in grazed plots; we found that METE's SSAD was the best model, by number of wins, for all situations: grazed, ungrazed, dry and intermediate moisture levels. In contrast to the difference in patterns detected in different-aged stands in a disturbance-adapted pine forest [Chapter 1], the METEs predictions of the SSADs of the majority of plots in this study either indicates that METE is not sensitive to anthropogenic changes to systems (possibly because the grazed plots are not in transition, but rather, in a new stable state), or indicates that pack stock effects are too subtle to be picked up in the SSAD.

Analyses not presented in this manuscript (Newman, *unpublished*) indicate that best-fit negative binomial models, each with a different *k*(nbd) value, can provide model fits on a case-by-case basis. In this study, *k*(nbd) values were generally less than 1, indicating a more clustered SSAD than is predicted by METE. In the future, METE's SSAD could theoretically be extended to include density dependence, which would lead to a variable *k*(nbd) value, but the version of METE used in this manuscript (Harte 2011, Harte and Newman 2014) has a fixed *k*(nbd) value of 1.

Conclusions

Although macroecological methods are not typically used for management purposes, the technique of using macroecological scaling-plots can be adapted to small scales. Through use of relatively small scaling plots (on the order of 10's to 100's of m^2), macroecological estimations of community structure, including SADs, singleton species present in an area, and SSADs, can be compared between sites with differing land-use history, or in future applications, to judge the efficacy of restoration efforts. Traditionally, macroecological plots are large (~50 ha, such as the Smithsonian plots) and single plots are used to estimate ecological metrics of entire ecosystems, but the use of small plots can equally well be used to estimate metrics of smaller parts of an ecosystem, as we applied this method to a particular moisture zone in meadows. This method can provide support for conservation and land management decisions: small-scale plots with replicates, that can be deployed and measured over short timescales, on limited budgets, and without specialized equipment.

We assessed the applicability of METE's scaling-plot technique for small-scale plots in Sierra meadows, and found that the number of singleton species is a reliable indicator of ecological differences at the plot level. By itself, the SAD is not a good indicator of the ecological perturbations we studied in the Sierran meadow system, which may indicate that the impacted meadows are not in a state of transition or recovery, but rather, in a new stable state with lower productivity. Alternately, pack stock use may have low impacts on the system that do not affect the shape of the SAD or the spatial distribution of species generally (as assessed by

SSADs), but an effect is still detected in the number of singleton species measured compared to METE's predictions. The SSAD, a measure of spatial clustering of individuals in a species over space, is too-well predicted by METE in the majority of plots compared to the other models we tested to discriminate between different states (ungrazed or grazed) in the system we examined. This metric may show promise as an indicator of disturbance or disruption if tested more thoroughly against other candidate models, in particular, negative binomial models with fixed *k*(nbd) values. Like other plot-level studies, we found variance among results of plots, and suggest that if this technique is to be applied in a management context, replicate plots are required for lowering uncertainty in measurement results.

Acknowledgements

We thank Sequoia National Park, the National Parks Service, and the United States Geological Survey (USGS) for providing permitting and logistical support. Contributions of S. Haultain (NPS), J.R. Matchett (USGS), S. Ostoja (USFS) and Lucas Joppa (Microsoft Research) were critical for multivariate matching of meadow pairs for shared attributes. I thank Ori Chafe and Henry Houskeeper for their work in the field, and Eric Berlow for contributions in the early phase of this project and comments on this manuscript. Mark Wilber provided modeling support and development during various stages of this project, as well as manuscript comments. Alex Génin contributed to plant identifications and design of analyses. Steve Lee provided comparison data and many helpful comments. I thank Alex Génin and David Hembry, who also provided useful comments on this manuscript. I thank Marco Paliza-Carre for his many months of careful data entry. This research was funded by the Gordon and Betty Moore Foundation and a National Science Foundation Graduate Research Fellowship Program grant to EAN, and grant NSF-EF-1137685. This study was carried out under SEKI Park-assigned permit SEKI-2012-SCI-0436, Activity #SEKI-00348.

Chapter 2 Tables and Figures

Tables.

Table 1. Plant species recorded in this study

Category	6-Letter code	Species name	Common name
Fern			
	BOTSIM	*Botrychium simplex*	Little grapefern/ Yosemite moonwort
Gymnosperms			
	CON1UN	Unknown conifer seedling	
Monocots			
(grasses)	CALBRE	*Calamagrostis breweri*	Shorthair reedgrass
	ELYELY	*Elymus elymoides*	Squirrel tail grass/ bottlebrush squirreltail
	MUHFIL	*Muhlenbergia filiformis*	Pullup muhly/ slender muhly
	PHLAPL	*Phleum alpinum*	Alpine timothy
	TRISPI	*Trisetum spicatum*	Spike trisetum
	GRA1UN	Unknown bunchgrass sp.1	
	GRA2UN	Unknown bunchgrass sp.2	
	GRA3UN	Unknown bunchgrass sp.3	
	GRA5UN	Unknown grass sp.5	
	GRA7UN	Unknown bunchgrass sp.7	
	GRA8UN	Unknown bunchgrass sp.8	
(sedges)	CARFIL	*Carex filifolia*	Threadleaf sedge
	CAR1UN	*Carex* sp.	
	SED1UN	Unknown sedge sp.1	
	SED2UN	Unknown sedge sp.2	
(unknowns)	MON1UN	Unknown monocot sp.1	
	SPP5UN	Unknown gramminoid sp.5	
Eudicots			
	ACHMIL	*Achillea millefolium*	Yarrow
	ANTROS	*Antennaria rosea*	Rosy pussytoes

BOCLYA	*Boechera lyallii* (syn. *Arabis lyallii* var. *lyallii*)	Lyall's rockcress
CYMTER	*Cymopterus terebinthinus*	Turpentine cymopterus
DODALP	*Dodecatheon alpinum*	Alpine shooting star
ERIINC	*Eriogonum incanum*	Frosted buckwheat
ERI1UN	*Eriogonum* unknown sp.1	Buckwheat
ERI2UN	*Eriogonum* unknown sp.2	Buckwheat
ERILON	*Erigeron lonchophyllus*	Shortray fleabane
FRA1UN	*Fragaria virginiana*	Scarlet strawberry/ Virginia strawberry/ mountain strawberry
GAYDIF	*Gayophytum diffusum*	Spreading Groundsmoke
GENHOL	*Gentianopsis holopetala*	Sierra fringed gentian
GENNEW	*Gentiana newberryi* var. *tiogana*	Tioga gentian
IVECAM	*Ivesia campestris*	Field ivesia/ field mousetail
KALPOL	*Kalmia polifolia*	Mountain laurel/ bog laurel
LUPLEP	*Lupinus lepidus*	Dwarf lupine
MIMPRI	*Mimulus primuloides*	Primrose monkeyflower
OREALP	*Oreostemma alpigenum* syn. *Aster alpigenus*	Tundra aster
OREPEI	*Oreostemma peirsonii* (syn. *Aster peirsonii*)	Peirson's aster
PEDATT	*Pedicularis attollens*	Little elephant's head
PENHET	*Penstemon heterodoxus*	Sierra beardtongue
PERBOL	*Perideridia bolanderi* ssp. *bolanderi*	Bolander's yampah
PLA1UN	*Plantaginaceae* sp.	Plantain
POLDOU	*Polygonum douglasii*	Douglas' knotweed
POTGRA	*Potentilla gracilis* var. *fastigiata*	Slender cinquefoil
PYRLAN	*Pyrrocoma lanceolata*	Lanceleaf goldenweed/ intermountain pyrrocoma
RANALI	*Ranunculus alismifolius*	Plantainleaf buttercup

	RUM1UN	*Rumex* sp.	Dock
	SENINT	*Senecio integerrimus*	Tall western groundsel/ lambstongue ragwort
	SIBPRO	*Sibbaldia procumbens*	Creeping sibbaldia
	SOLMUL	*Solidago multiradiata*	Alpine goldenrod
	STELON	*Stellaria longipes*	Goldie's starwort/ longstalk starwort
	TRIMON	*Trifolium monanthum*	Mountain carpet clover
	VIOADU	*Viola adunca*	Western dog violet/ sand violet
(unknowns)	AST2UN	Unknown Asteraceae 2	
	SPP2UN	Unknown forb 2	
	SPP3UN	Unknown forb 3	
	SPP6UN	Unknown forb 6	
	SPP7UN	Unknown forb 7	
	SPP8UN	Unknown forb 8	

Table 2. Attributes of 18 plots distributed over 6 sites in this study

Site name	**Meadow number**	**GIS lat/long [and elevation centroid]**	**Treatment**	**Original moisture zone category**	**Total number of species (*S*) in plot**	**Total abundance (*N*) in plot**	**Absolute percent cover**
Pair A: Rock Creek	4313	36°29'31.64"	Grazed	mesic	17	8583	72.66
		118°18'36.92"		mesic transition	23	7983	66.88
		[3056 m]		xeric	17	1011	16.48
	505	36°28'38.90"	Control	mesic	23	9803	95.39
		118°18'48.78"		mesic transition	16	8044	62.5
		[3216 m]		xeric	12	1719	28.52
Pair B: Tyndall Creek	3597	36°37'39.52"	Grazed	mesic	27	8356	70.7
		118°23'33.19"		mesic transition	26	5439	59.77
		[3221 m]		xeric	13	1811	18.67
	3440	36°38'56.55"	Control	mesic	21	10261	93.75
		118°25'21.70"		mesic transition	21	9149	77.19
		[3355 m]		xeric	11	1567	26.09
Pair C: Crabtree	9054	36°33'17.65"	Grazed	mesic	18	7706	70.78
		118°20'53.95"		mesic transition	18	5103	53.44
		[3195 m]		xeric	11	1486	20.31
	5053	36°33'35.16"	Control	mesic	27	7647	86.09
		118°20'38.68"		mesic transition	22	4049	53.28
		[3233 m]		xeric	21	774	18.63

Table 3. Model comparisons for the species abundance distributions of each plot measured in this study

Treatment	Site	Meadow number	Original moisture zone	Model	k	AICc	w_i	ΔAIC
Grazed	Rock Creek	4313	mesic	Lognormal	2	238.1365	0.4639	0.2895
				METE log series	**1**	**237.8469**	**0.5361**	**0.0000**
	Tyndall Creek	3597	mesic	Lognormal	2	338.2366	0.3122	1.5799
				METE log series	**1**	**336.6567**	**0.6878**	**0.0000**
	Crabtree	9054	mesic	**Lognormal**	**2**	**224.1658**	**0.7143**	**0.0000**
				METE log series	1	225.9985	0.2857	1.8327
	Rock Creek	4313	mesic transition	Lognormal	2	306.5469	0.3253	1.4591
				METE log series	**1**	**305.0879**	**0.6747**	**0.0000**
	Tyndall Creek	3597	mesic transition	Lognormal	2	281.5527	0.0031	11.5464
				METE log series	**1**	**270.0063**	**0.9969**	**0.0000**
	Crabtree	9054	mesic transition	**Lognormal**	**2**	**235.3046**	**0.7310**	**0.0000**
				METE log series	1	237.3037	0.2690	1.9991
	Rock Creek	4313	xeric	Lognormal	2	178.6891	0.3264	1.4491
				METE log series	**1**	**177.2400**	**0.6736**	**0.0000**
	Tyndall Creek	3597	xeric	**Lognormal**	**2**	**153.1703**	**0.5325**	**0.0000**
				METE log series	1	153.4305	0.4675	0.2602
	Crabtree	9054	xeric	Lognormal	2	127.5642	0.0934	4.5450
				METE log series	**1**	**123.0192**	**0.9066**	**0.0000**

Ungrazed	Rock Creek	505	mesic	**Lognormal**	**2**	**322.6324**	**0.9128**	**0.0000**
				METE log series	1	327.3290	0.0872	4.6966
	Tyndall Creek	3440	mesic	Lognormal	2	286.9196	0.0537	5.7389
				METE log series	**1**	**281.1806**	**0.9463**	**0.0000**
	Crabtree	5053	mesic	Lognormal	2	358.1710	0.3480	1.2560
				METE log series	**1**	**356.9150**	**0.6520**	**0.0000**
	Rock Creek	505	mesic transition	**Lognormal**	**2**	**226.6055**	**0.7093**	**0.0000**
				METE log series	1	228.3891	0.2907	1.7836
	Tyndall Creek	3440	mesic transition	Lognormal	2	280.6535	0.3375	1.3492
				METE log series	**1**	**279.3044**	**0.6625**	**0.0000**
	Crabtree	5053	mesic transition	Lognormal	2	269.4951	0.0037	11.1865
				METE log series	**1**	**258.3086**	**0.9963**	**0.0000**
	Rock Creek	505	xeric	Lognormal	2	138.2299	0.4216	0.6327
				METE log series	**1**	**137.5972**	**0.5784**	**0.0000**
	Tyndall Creek	3440	xeric	**Lognormal**	**2**	**135.8331**	**0.9831**	**0.0000**
				METE log series	1	143.9638	0.0169	8.1308
	Crabtree	5053	xeric	**Lognormal**	**2**	**200.6889**	**0.6818**	**0.0000**
				METE log series	1	202.2127	0.3182	1.5238

Figures.

Figure 1. Locations of high-Sierra meadow study sites in Sequoia National Park, California

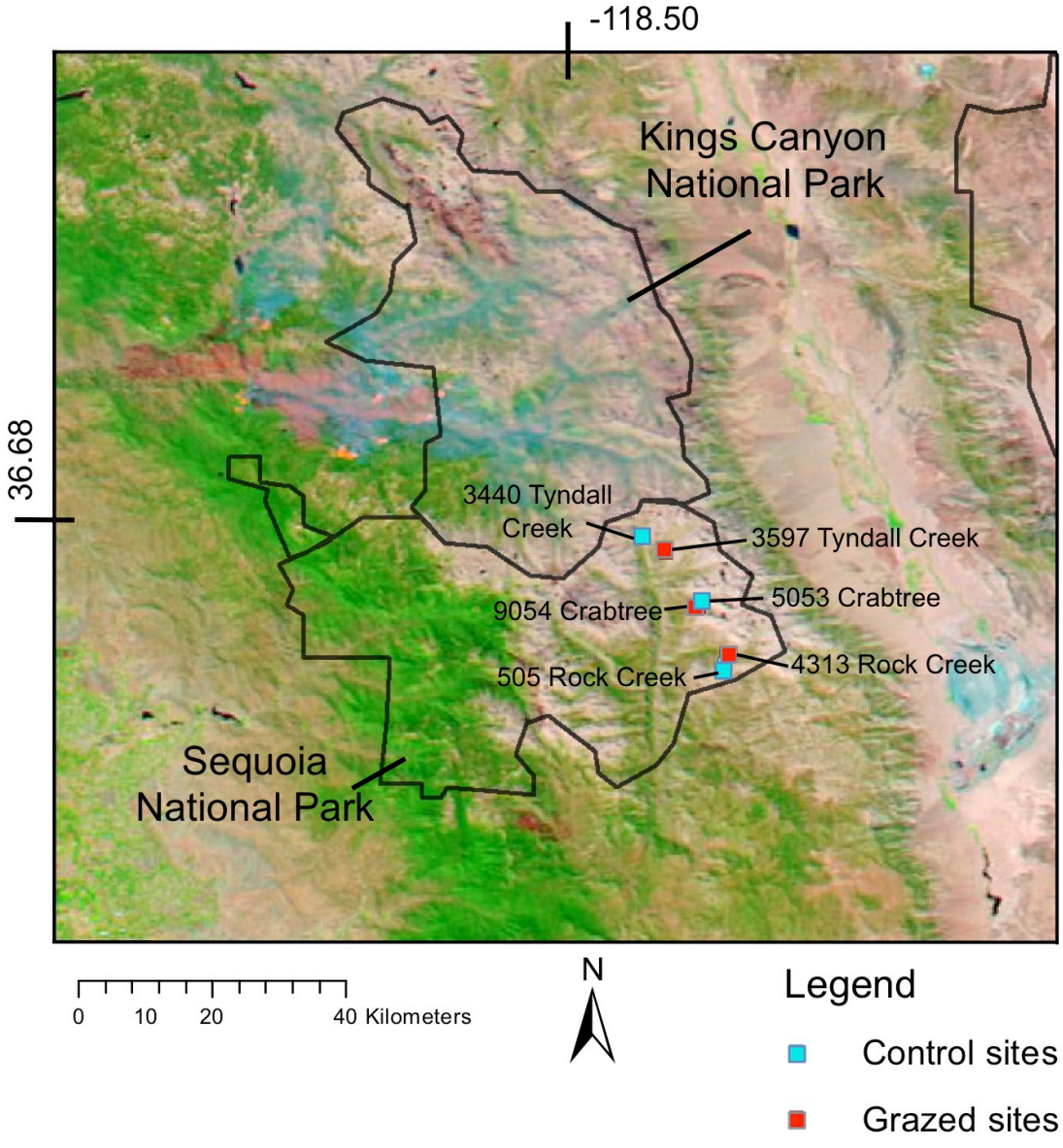


Figure 2. Examples of cells from sites in this study. (A) Intermediate ungrazed (meadow 505, plot closest to stream); (B) intermediate ungrazed (meadow 505, plot mid-distance from stream); (C) intermediate grazed (meadow 4313, plot mid-distance from stream); (D) xeric grazed (meadow 4313, plot furthest from stream).

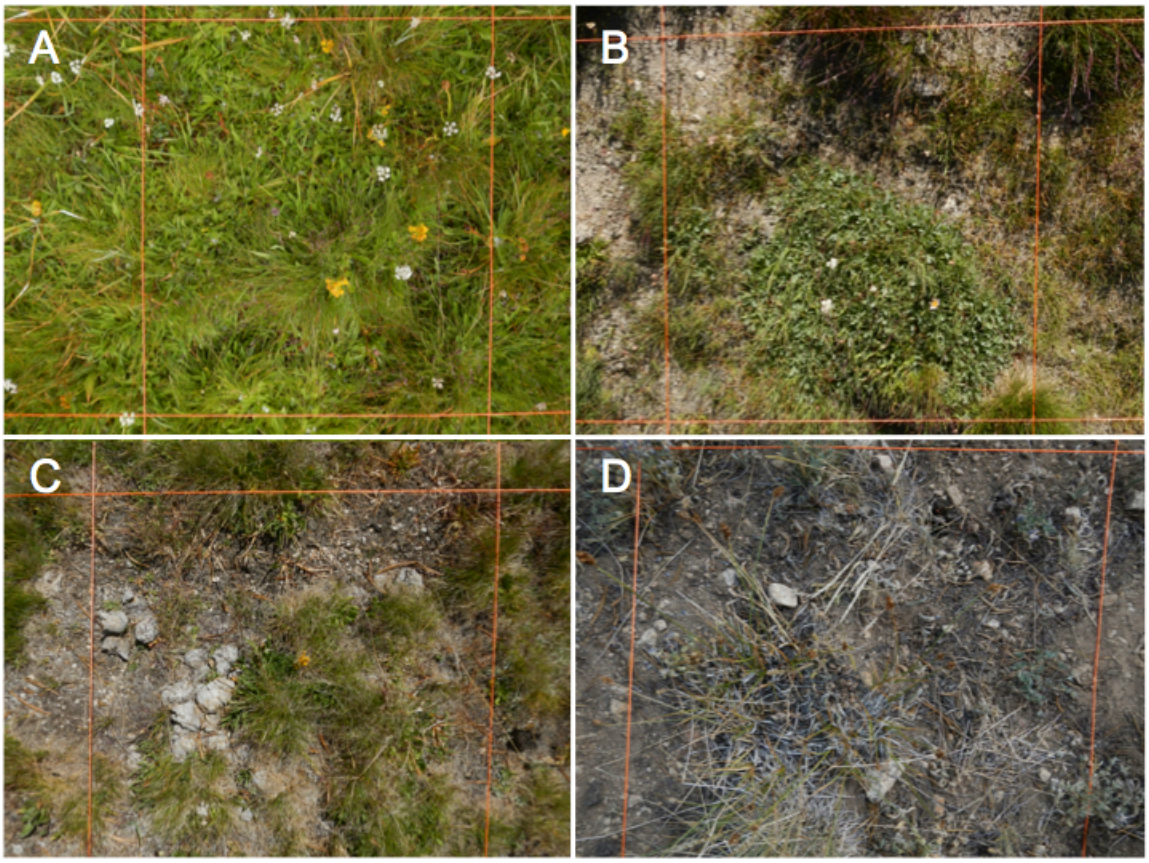


Figure 3. Photographs of two intermediate moisture plots and their immediate surroundings for: (A) meadow 505, ungrazed; and (B) meadow 3597 grazed (note horses in background).

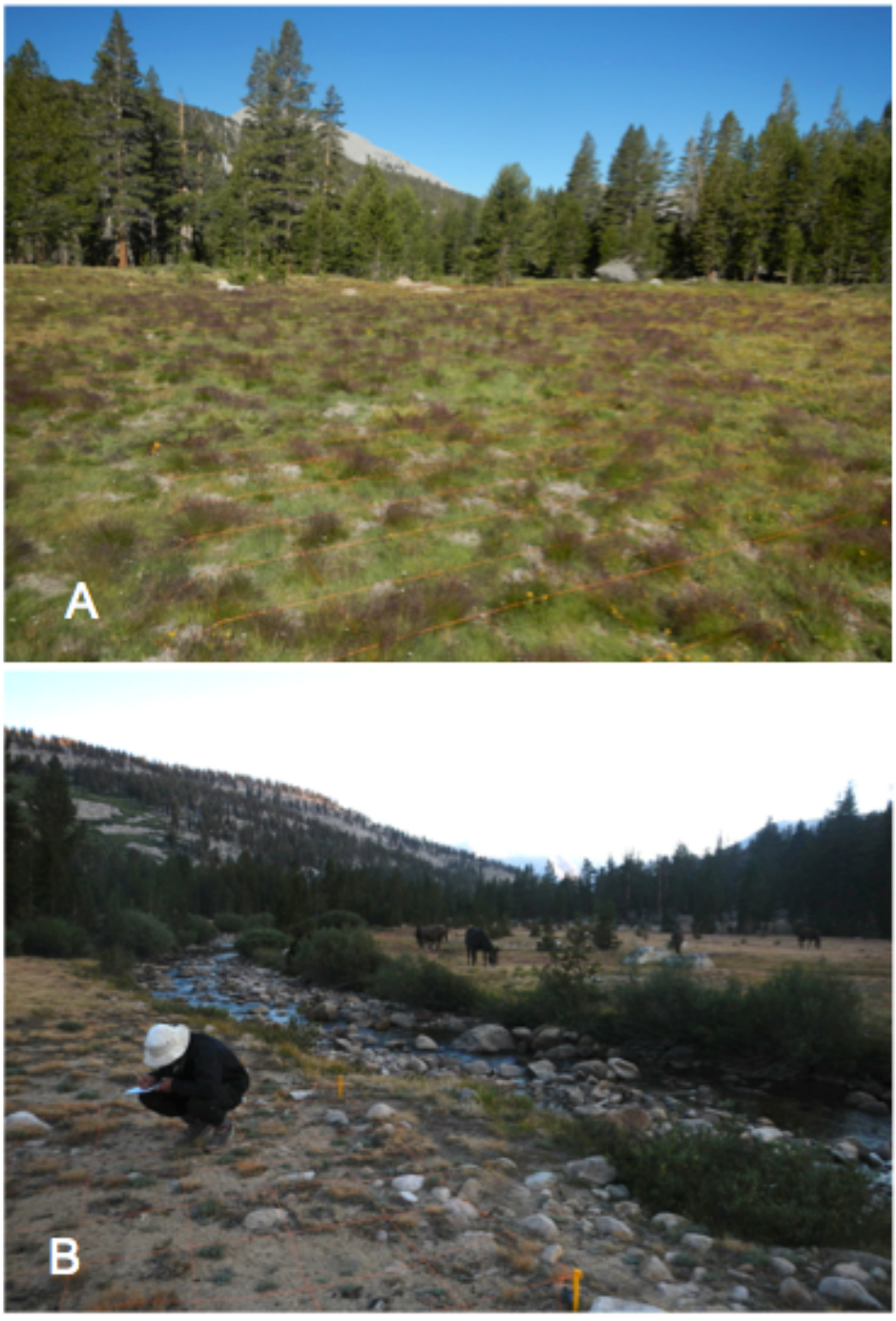


Figure 4. Relative percent covers of the 10 most dominant species, and all other species in each plot, by cell

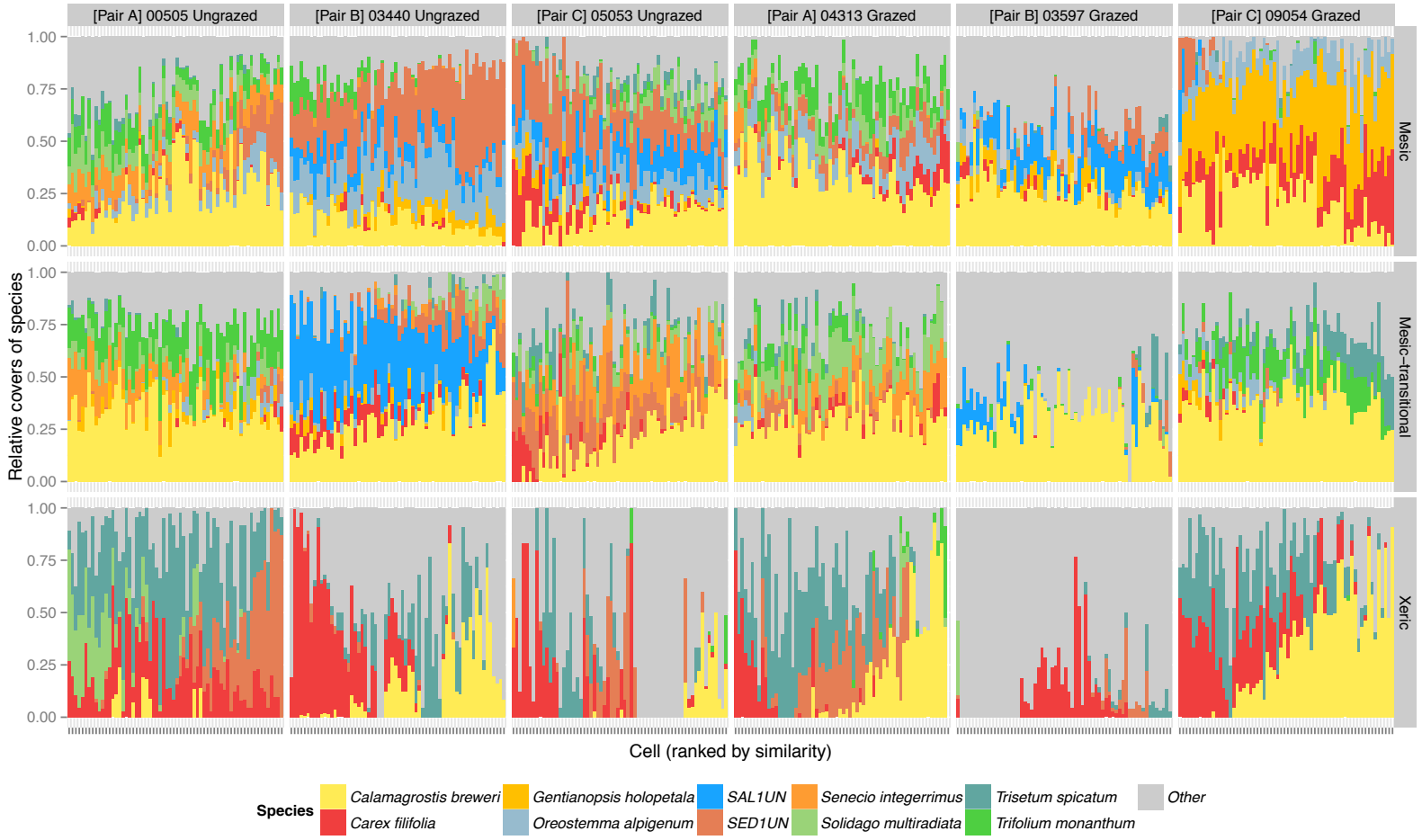


Figure 5. Comparative vegetation community structure and moisture levels of sites from this study (dots) and a larger data set incorporating hydrology measurements (background graph, Lee, 2013 and Lee, *in prep.*). Only the 12 most common species are shown here. Vegetation communities were compared between studies using a non-metric multidimensional scaling analysis.

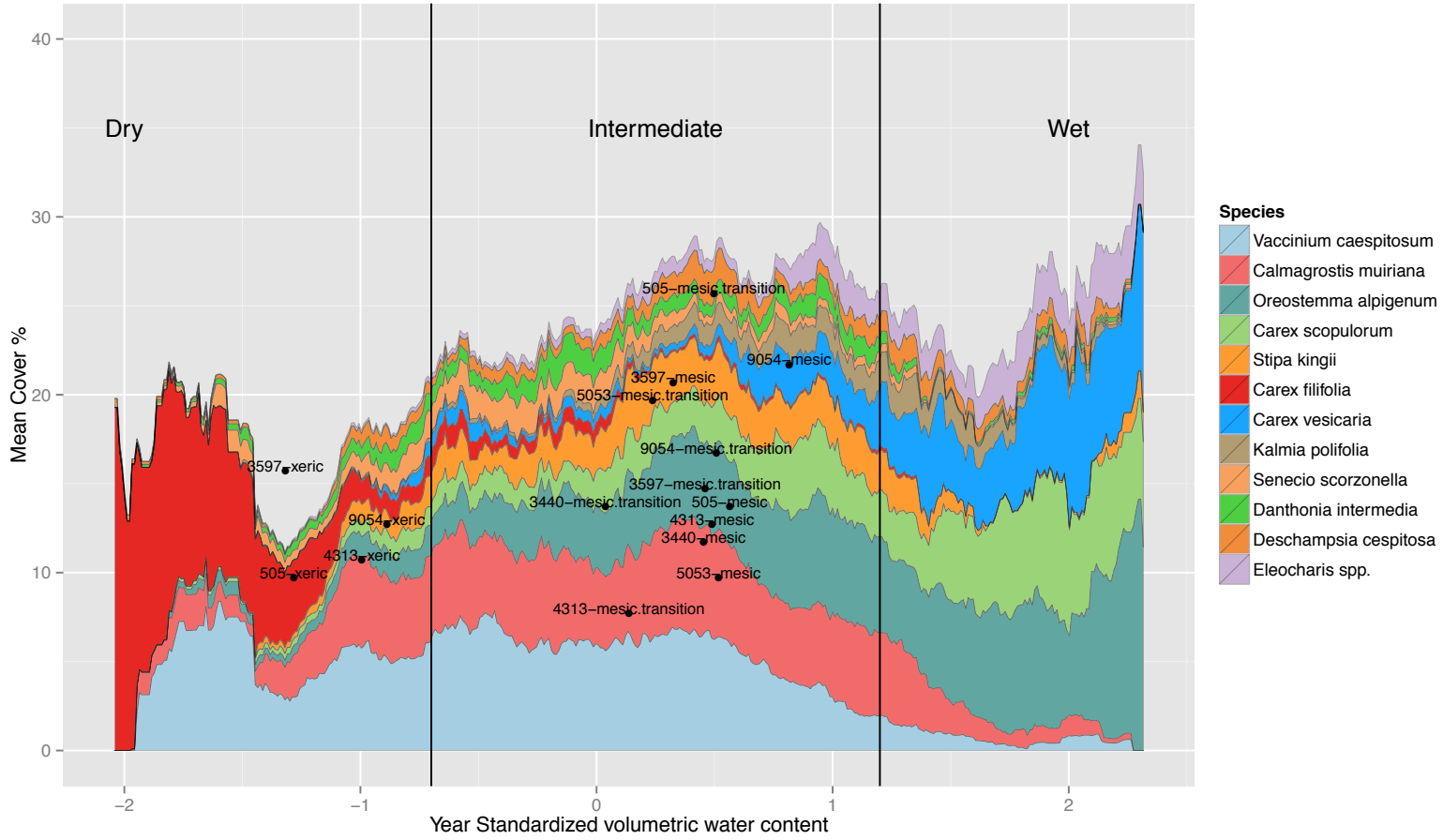


Figure 6. Number of model wins for species-abundance distribution candidate models, applied to 18 plots in this study. Results for all plots are first grouped by grazing history (GR=grazed, UNG=ungrazed), then by moisture level (INT=intermediate, DRY=dry), and finally, both by grazing history and moisture level. Pearson's Chi-squared tests (2-tailed tests with df=1) reveal no statistically significant differences between model performance in any category.

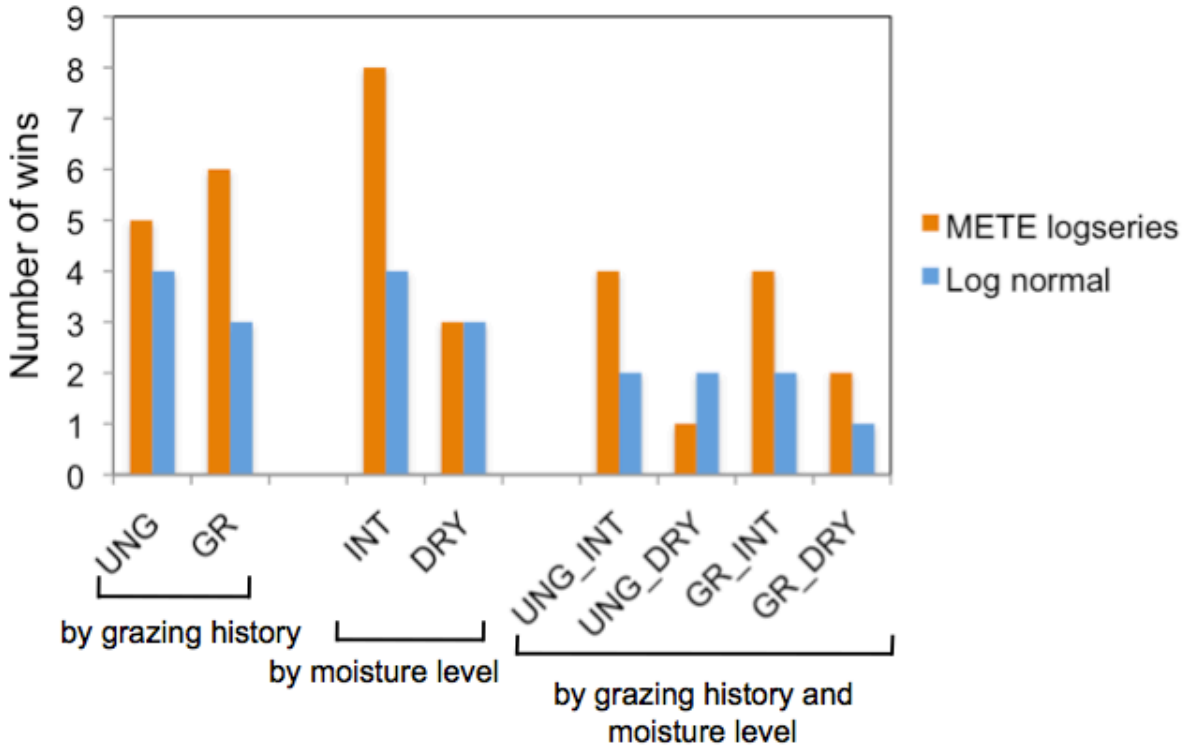


Figure 7. Observed versus predicted singleton (*n*=1 individual) species for all plots in this study, separated by (A) ungrazed, and (B) grazed plots. Predicted numbers arise from the METE species-abundance distribution. Data points on graphs each represent one plot (not jittered in these graphs). Open circles represent plots of intermediate moisture levels, and open squares represent dry plots. Equation and R^2 values are calculated for regression line.

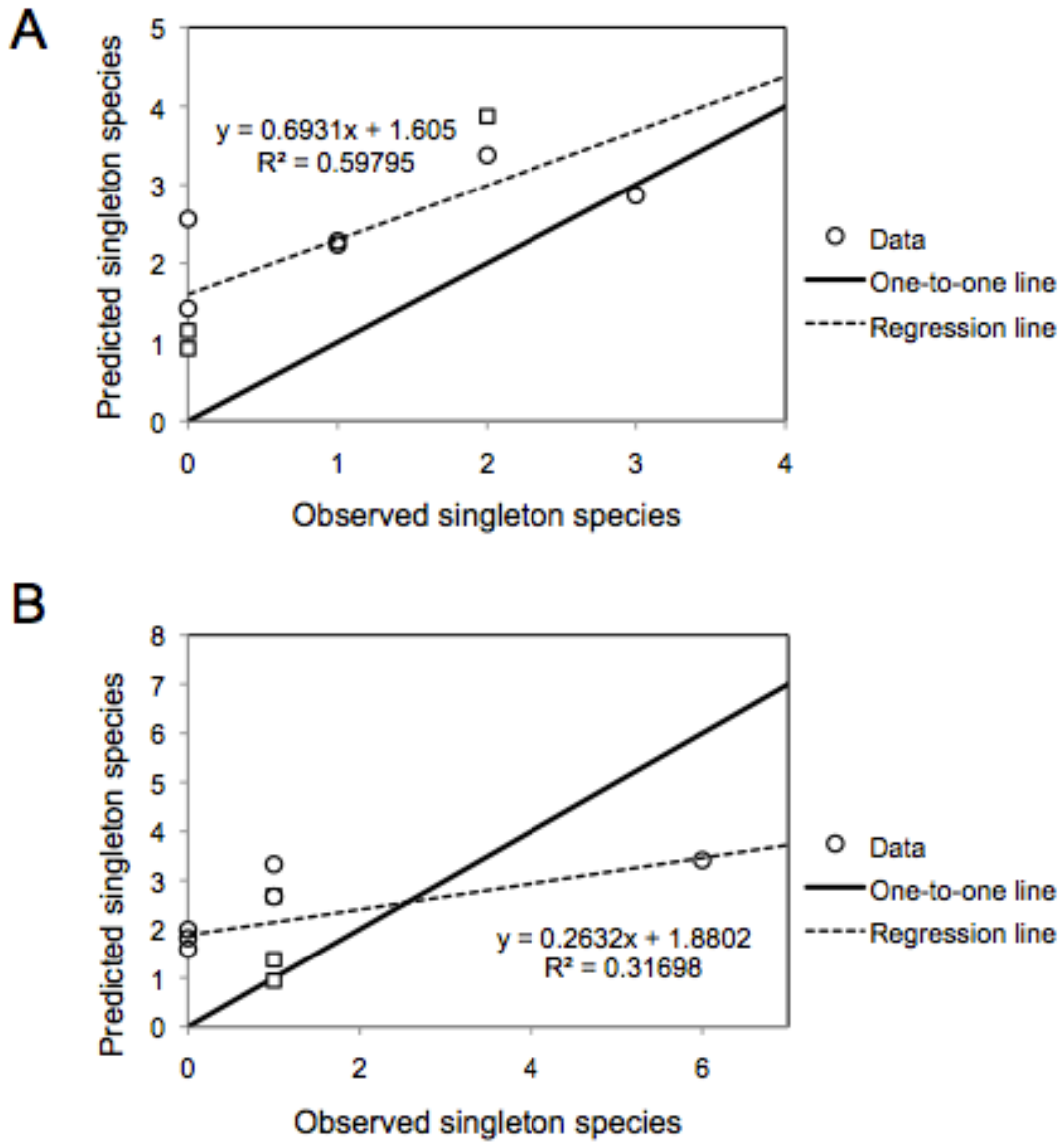


Figure 8. Histograms comparing SSAD models for all species by their AIC*c* weights. By column: (A) intermediate ungrazed and (B) intermediate grazed plots in this study.

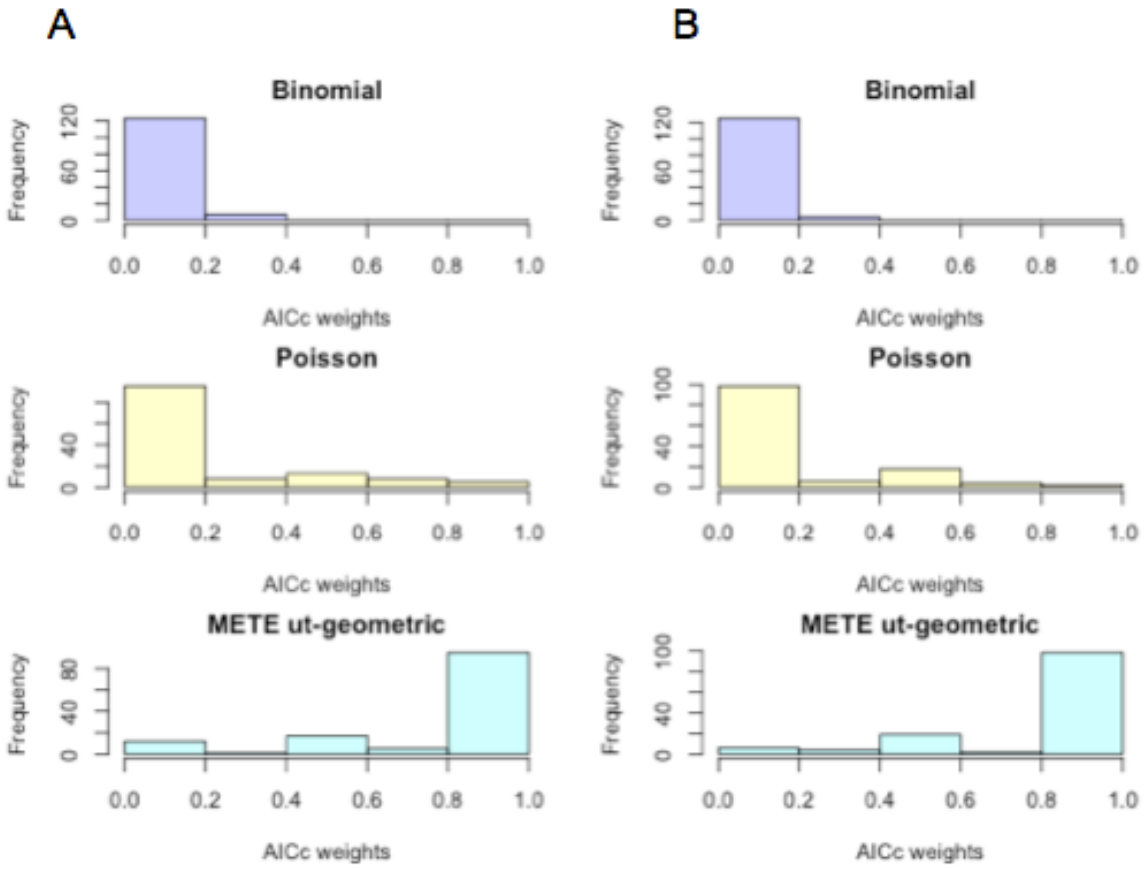
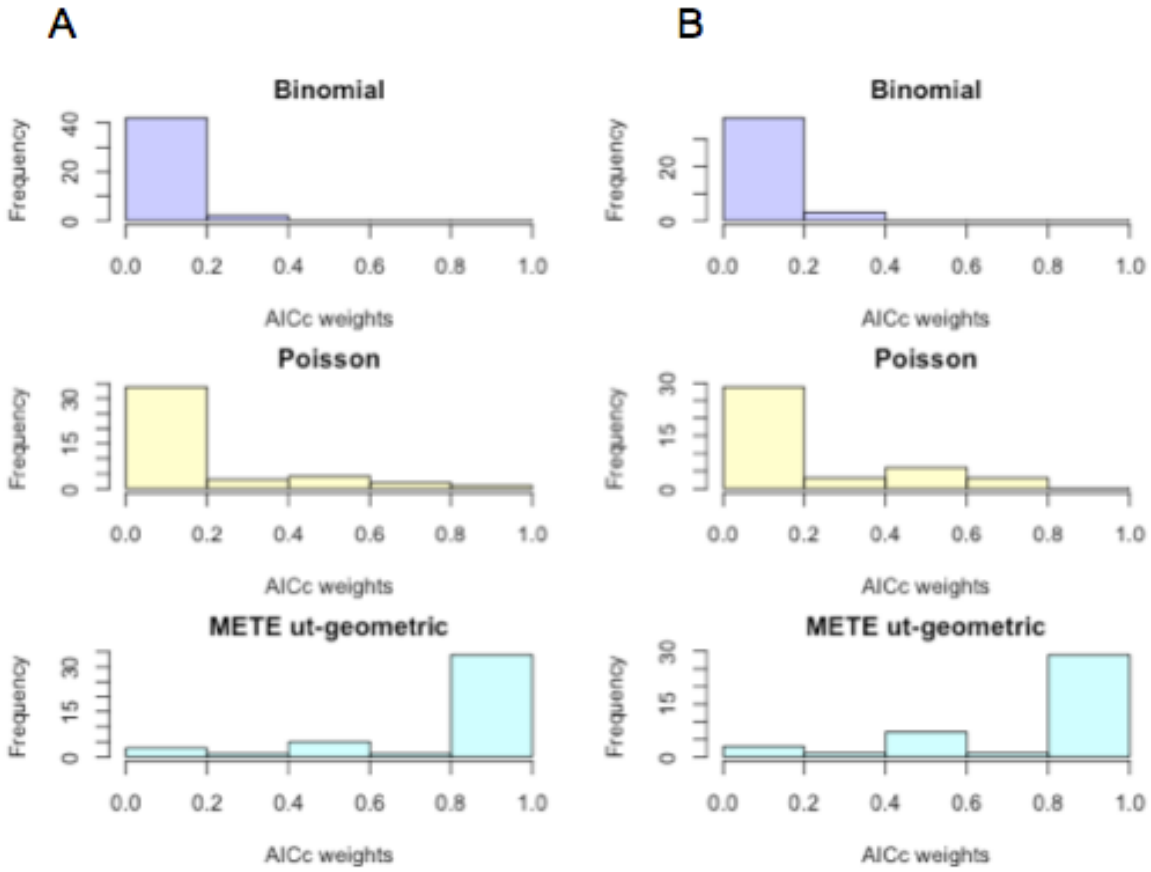


Figure 9. Histograms comparing SSAD models for all species by their AIC*c* weights. By column: (A) xeric ungrazed and (B) xeric grazed plots in this study.



CHAPTER 3

Empirical tests of within- and across-species energetics in a diverse plant community

Chapter 3: Empirical tests of within- and across-species energetics in a diverse plant community

This chapter has previously been published, and is reproduced here with permission from the publisher, the Ecological Society of America.

Newman, E. A., Harte, M. E., Lowell, N., Wilber, M., & Harte, J. (2014). Empirical tests of within-and across-species energetics in a diverse plant community. *Ecology* 95(10): 2815–2825.

Chapter 3: Empirical tests of within- and across-species energetics in a diverse plant community

Abstract

Many fundamental properties of ecological systems and interactions are tied to body size, and a related metric, the metabolic rate distribution, both within and across species. A previously proposed Maximum Entropy Theory of Ecology (METE) predicts numerous interrelated macroecological patterns, including spatial distributions of individuals within species, abundance distributions across species, species area relationships, and distributions of metabolic rates of all individuals within a community. Extensive tests of METE's macroecological predictions generally support the theory, but two related predictions have not been evaluated against full community census data: the distribution of metabolic rates of individuals within species as a function of the abundance of the species, and the distribution of average individual metabolic rates across species. We test the metabolic predictions of METE for herbaceous plants in a subalpine meadow and show that while this theory realistically predicts the distribution of individual metabolic rates across the entire community, the within and across species predictions generally fail. We also test the energy-equivalence type prediction that arises as a consequence of the prediction for the distribution of average individual metabolic rates across species. We suggest several possible explanations for the empirical deviations from theory, and distinguish between the expected deviations caused by ecological disturbance and those deviations that might be corrected within the theory.

Key words

metabolism, macroecology, scaling laws, Maximum Entropy Theory of Ecology, METE, information entropy, energy equivalence

Introduction

Many fundamental properties of ecosystems are tied to body sizes of their component members, both at the individual and species level. Within a species, body size often correlates to fitness and metabolism, whereas considering all species in a community, body size will correlate to metabolism, abundance, trophic activity, ecological and life history traits, and community stability. The distribution of body sizes within a community is potentially important for understanding community properties and function, as well as predicting conservation implications of invasive species or changes that will result following species extirpation in a community. Accordingly, body size distributions have attracted much theoretical and empirical attention (Damuth 1981, Damuth 1987, Niklas and Enquist 2001, Brown et al. 2004, Muller-Landau et al. 2006, White et al. 2007, Clauset et al. 2009, Mori et al. 2010, Harte 2011). Macroecologists have attempted to connect body size to abundances of species using power law models and experimental observations, but despite these efforts, there had not been an overarching theoretical framework for connecting metrics that describe abundance and metabolics until the development of the Maximum Entropy Theory of Ecology (METE) (Harte et al. 2008, 2009; Harte 2011).

METE is an extremely general macroecological theory that predicts spatial, abundance and metabolic rate distributions of species, and the interrelationships of these metrics for any system defined by a set of basic community state variables. It therefore also predicts body-size distributions if a metabolic scaling relationship between metabolism and body size is assumed. Unlike other unified theories of ecology and diversity (Hubbell 2001, de Aguiar et al. 2009, McGill 2010), METE uses no adjustable parameters, and is instead formulated using the procedure of maximizing information entropy, a technique that has been proven to yield least-biased probability distributions that are consistent with prior knowledge in the form of constraints on those distributions (Jaynes 1979; 1982; 2003). Constraints may take the form of the empirical ranges, means, or higher moments of values of one or more independent variables of a system, or outcomes of system measurements (Harte et al. 2008, Harte and Newman 2014).

As inputs, METE uses four measured state variables: S_0, the total number of species; N_0, the total number of individuals; E_0, the total metabolic rate of all individuals, and A_0, an arbitrary but specified spatial scale. Mathematical forms of macroecological metrics are derived from constraints arising from ratios of the four state variables. Various empirical tests of species-area relationships (SARs), endemics-area relationships (EARs), and species abundance distributions (SADs) indicate support for theoretical predictions of those metrics (Harte et al., 2008; 2009; White et al. 2012), but the metabolic predictions of METE require more empirical validation, including tests of the predicted within-species distribution of metabolic rates, and the predicted distribution of species' average metabolic rates across all species (see Xiao et al. 2013). In this study, we examine the diverse plant community of a subalpine meadow in the Rocky Mountains to investigate the extent to which METE makes realistic predictions about the metabolic rate distributions, both within and across species.

Methods

Site description

Sampling was carried out in the summer of 2012 in Gunnison National Forest, Gunnison County Colorado at a site located 6.8 km (4.2 mi) north of the Gothic townsite (plot location 39° 0' 20.29"N, 107° 1' 54.67"W), at an elevation of 3189 m. The study plot was sited in an area that minimized the extent of disturbance in the forms of Northern pocket gopher (*Thamomys talpoides*) mounds, and anthropogenic disturbance including cattle grazing, human presence, and invasive grasses and forbs. We chose a site that received constant moisture from a nearby creek to offset effects of the drought occurring in 2012, but overall disturbance level was still moderate. The dominant trees of this area are primarily Engelmann spruce (*Picea engelmannii*) and Subalpine fir (*Abies bifolia*), in stands surrounding but distinct from the meadow communities. Open meadows contain shrubby *Potentilla fruticosa* and large patches of the herbaceous *Veratrum californicum*, while moister areas of the meadow have large patches of mixed-species shrubby willows (dominated by *Salix planifolia*).

Sampling

A 4m x 4m plot was gridded into 1 m cells. Plants were censused and recorded with both a cell number and an x and y spatial coordinate reference within the cell for ease of relocating individual plants. After locating all individuals, we assigned each a unique identification and

counted all living leaves. We clipped a representative fraction of the living aboveground biomass at ground level from each plant for analysis (Cornelissen et al. 2003). We collected 20-50% of the live biomass of large plants, and collected small plants whole (see Appendix A for more detail). We placed samples into plastic bags and transported them to a lab in a cooler with ice to prevent leaf area shrinkage (following the methods of Blonder 2012), where they were rehydrated and placed on a flatbed scanner. We analyzed and recorded leaf and stem area separately for each sample using the software ImageJ 1.46 (Abramoff et al. 2004). We calculated total photosynthetic area of an individual plant by applying mean area per leaf sampled from that plant to its total leaf count, and adding its total stem area. This was done on an individual basis in order to take into account all plants' life stages and leaf size variation. Because only living material was included in this study, we excluded plants with all leaves senesced at the time of sampling, which excluded some individual plants and may have excluded one or more early blooming species altogether. Similarly, we excluded immature plants that measured less than 1 cm in height and breadth at the time of sampling.

Metabolic calculations

Allometric work by Niklas and Enquist (2001) has established that leaf area (A_L) scales as body mass (M) to the ¾ power ($A_L \propto M^{3/4}$) across species. Subsequent studies confirm this relationship between simple metrics of leaf area (or photosynthetic area, including green stems), leaf mass, body mass, and metabolic rates (Reich et al. 2006, Deng et al. 2008, Makarieva et al. 2008). Metabolic scaling theory states that $B \propto M^{3/4}$, where B is metabolic rate (Niklas and Enquist 2001). By this logic, it follows that $A_L \propto B$; that is, the relationship between leaf area and metabolic rate is approximately linear. Although the metabolic rate should in fact represent the instantaneous metabolism of each plant, simultaneous measurements of whole-plant photosynthetic processes were infeasible. A measure of total photosynthetic area, which is the result of the energy an individual has accrued over time (time integrated energy), is therefore suitable as a proxy of the time-averaged metabolic rate. Similar justification in other macroecological studies allows the estimation of metabolism from diameter at breast height (dbh) with scaling assumptions (see for example Muller-Landau et al. 2006, Rüger and Condit 2012, Xiao et al. 2013), and by organisms' mass (Glazier 2006).

Using the proportional relationship between total leaf area and total metabolic energy of a plant $A_L \propto B$, we can then estimate the METE state variable Energy (E_0), and the metabolic rate of an individual, ε, by units of normalized photosynthetic area for the herbaceous plants considered here. E_0 is related to ε by the relationship $E_0 = \sum_{(i = 1 \text{ to } No)} \varepsilon_i$, with the continuous variables ε and E_0 normalized through a linear transformation such that the individual with the lowest metabolic rate has $\varepsilon = 1$ (that is, $\varepsilon_i = a_i/\min(a_i)$, where a_i is photosynthetic area for individual i, and $\min(a_i)$ is the total photosynthetic area of the smallest individual in the community). We note that the quantity historically referred to in the metabolic literature as either "energy," "metabolic energy," or "respiratory metabolic rate" actually have physical units of power. We choose to use the term "metabolic rate distribution," and note that all "energy" labels and metrics of energy are actually in units of power.

We test three metrics predicted by METE for the subalpine plant community: the distribution of metabolic rates of all individuals in a community, the distribution of average individual metabolic rates across species (both of these being community-level metrics), and the

distribution of metabolic rates of individuals within species as a function of the abundance of the species (a species-level metric). See Fig. 1.

(1) The Individual-level Energy Distribution (IED) describes the distribution of metabolic rates over all of the individuals in the community, and is represented by $\Psi(\varepsilon)$:

$$\Psi(\varepsilon \mid S_0, N_0, E_0) = \lambda_2 \cdot \beta \frac{e^{-\gamma}}{(1-e^{-\gamma})^2} \tag{1}$$

Where ε is the metabolic rate of an individual, $\gamma = \lambda_1 + \lambda_2\varepsilon$; λ_1 and λ_2 are the Lagrange multipliers given by $\lambda_1 = \beta - \lambda_2$ and $\lambda_2 = S_0/(E_0 - N_0)$ respectively; and β satisfies the approximate relationship $\beta\ln(1/\beta) \approx S_0/N_0$ (Harte et al. 2008, Harte et al. 2009). A derivation of this expression and discussion of simplifying assumptions is available in Harte 2011, where a more exact expression for β is given. $\Psi(\varepsilon)$ is a continuous distribution on the domain of 1 to E_0. The integral of $\Psi(\varepsilon)$ from 1 to E_0 is 1, and so $\Psi(\varepsilon)$ is the normalized probability density function for metabolic rates of all individuals in the community.

(2) The Average Species Energy Distribution (ASED) is the distribution, across all the species in the community, of metabolic rates averaged over individuals within species, and is represented by $\nu(\bar{\varepsilon})$. The predicted form of the ASED of a community is:

$$\nu(\bar{\varepsilon} \mid S_0, N_0, E_0) = \frac{1}{\ln\left(\frac{S_0}{\beta}\right)} \cdot \frac{e^{\frac{-\beta}{\lambda_2(\bar{\varepsilon}-1)}}}{(\bar{\varepsilon}-1)} \tag{2}$$

where $(\bar{\varepsilon})$ is the average metabolic rate of the individuals for the species under consideration, and is defined on the continuous domain from 0 to N_0 (Harte et al. 2008, Harte et al. 2009). This exact expression replaces the approximate expression in Harte, 2011 (eq. 7.46) because the simplification used in Harte 2011, $\ln(1/\beta) >> \ln(S_0)$, or equivalently $\beta \cdot S << 1$, does not hold for our data. Here, the term $\ln(S_0/\beta) = \ln(1/\beta) + \ln(S_0)$ replaces $\ln(1/\beta)$.

A key relationship for understanding the ASED is the inverse relationship between the expected energy, $\bar{\varepsilon}$ for a species, and its abundance n (Harte et al. 2008, Harte 2011 Ch. 7):

$$\bar{\varepsilon} \approx 1 + \frac{1}{n\lambda_2} \tag{3}$$

Although $\bar{\varepsilon}$ is not exactly inversely proportional to n, it is the case for those species for which $n << 1/\cdot\lambda_2$, and therefore $1/(n \cdot \lambda_2) >> 1$. For our study system, the inequality holds for all species, and so the added 1 can be neglected.

Using this relationship between $\bar{\varepsilon}$ and n, we derive a probability distribution for $\bar{\varepsilon}$ from the species-abundance distribution. This probability distribution is the ASED itself (see Harte 2011, p.155). As a result of the functional form of Eq. (3), the smallest values of average energy per individual in a species are expected for the most abundant species, and conversely, the

largest values of average energy per individual within a species are expected for species with the lowest abundances.

Because the SAD is a continuous function of a discrete variable (n), the constructed ASED distribution reflects this discrete component. The predicted ASED is only defined for values of $\bar{\varepsilon} = (1+(1/(n\cdot\lambda_2))$, where n (the number of individuals of a species) is in the range [1, N] (and where $N = n_{max}$); however, empirical values of $\bar{\varepsilon}$ can take on any continuously defined value between 1 and E_0 . This leads to artificial effects described in Appendix F.

(3) The Species-level Energy Distribution (SED) is the distribution of metabolic rates across the individuals in a single species of abundance n. The SED is represented by $\Theta(\varepsilon|n)$, and is derived from the SAD, represented by $\Phi(n)$ (from Harte 2011):

$$\Theta(n \mid \varepsilon) = \frac{R(n,\varepsilon)}{\Phi(n)} \tag{4}$$

where $R(n,\varepsilon)$ is the "ecosystem structure function," a joint, conditional probability that a chosen species will have the abundance n, and that a given individual picked from that species will have a metabolic rate equal to that of the smallest individual (Harte 2011). The form of the ecosystem structure function is uniquely predicted by MaxEnt when the numerical values of the state variables are known. The approximation below is the form of the function used to predict SEDs by species (Harte 2011, Ch. 7):

$$\Theta(\varepsilon \mid n,S_0,N_0,E_0) \approx \lambda_2 n e^{-\lambda_2 n(\varepsilon -1)} \tag{5}$$

Summing $\Theta(\varepsilon|n)\cdot\varepsilon$ over all possible values of ε recovers Eq. (3). Predictions of SEDs by species incorporate the overall constraint that the smallest, lowest-energy individual in the entire study system must equal the actual value normalized to $\varepsilon = 1$. The SED metric is related to the ASED, in that the predicted distribution of the means of the SEDs is the ASED.

Letting r denote rank, the rank-metabolism relationships are derived from the SED and IED, assigning the rank of $r = 1$ for the individual with the largest metabolic rate. A maximum rank of $r = n$ (by species) for the SEDs, and $r = N_0$ for the IED is assigned to the individual with the lowest metabolic rate. Ranked metabolic distributions for the SED and IED are given by the following equations, respectively (Harte 2011, Ch.7):

$$\varepsilon_\theta(r,n) = 1 + (\lambda_2 n) \times \ln\left(\frac{1}{\frac{r-\frac{1}{2}}{n} + e^{-\lambda_2 n \times \varepsilon_{max}}}\right) \tag{6}$$

$$\varepsilon_\Psi(r) = \frac{1}{\lambda_2}\cdot \ln\left(\frac{\beta N_0 + r - \frac{1}{2}}{r-\frac{1}{2}}\right) - \frac{\lambda_1}{\lambda_2} \tag{7}$$

where ε_{max} is the individual in that species with the maximum metabolic rate (largest photosynthetic area) individual. To calculate ε_{max}, Eq. 7 is calculated with $r = 1$, yielding:

$$\varepsilon_{\Psi,\max} = \varepsilon_{\Psi}(r = 1) \approx \frac{1}{\lambda_2}\log(2\beta N_0) \quad (8)$$

(Harte 2011). This equation refers to the entire individuals pool for the IED, but can be easily modified for the SEDs, so that for the SEDs, $\varepsilon_{max} = \log((2n)/n)$. The rank-metabolism predictions for the SED is modified from Harte 2011 not to contain simplifying assumptions as described for Eq. 2, above.

Analysis

Many analyses and calculation of differential equations for the metrics considered in this paper were carried out with Python (van Rossum and Drake 2001) in the open-source project "Macroeco" (Kitzes et al. 2014). Analyses were carried out with ASED, IED and SED "metabolic analysis" scripts in the Beta version of this software. Predictions from Eqs. (2), (6) and (7) are compared to the measured metabolic rate and rank-metabolism data using slope and R^2 values on log(observed) vs. log(predicted value) graphs, and Kolmogorov–Smirnov tests are performed to determine the quality of METE predictions in the subalpine plot.

Results

A total of 877 plants comprising 31 species are included in this study. A complete species list for the plot and species' abundances are shown in Table 1. Calculated values for the Lagrange multipliers are $\lambda_1 = 7.0463 \times 10^{-3}$ and $\lambda_2 = 3.3806 \times 10^{-5}$ for $N_0 = 877$, $S_0 = 31$ and $E_0 = 917871.975$, and the parameter $\beta = 7.0801 \times 10^{-3}$. The range of total photosynthetic areas for individuals (a surrogate for body sizes) represented by our study system is greater than 38,200:1, that is, a 4-orders of magnitude range (comparable to the ranges represented by the trees of the Smithsonian datasets, which are often used to test macroecological theory). Minimum photosynthetic area for an individual plant is $a_{min} = 0.2793$ cm^2. Results for the IED, ASED, and SEDs are discussed in detail below. Graphs of the species- and community-level energy metrics are shown in Fig. 2, IED; Fig. 3, ASED; and Fig. 4, SEDs of the 12 most abundant species. SED results for all species with $n \geq 5$ are available in Appendix B with a summary of derived values in Appendix C, and results of the 15 most abundant species in order by average metabolic rate of an individual within a species are shown in Appendix D. Goodness of fit tests, including Kolmogorov–Smirnov (K-S) tests, are presented for all metrics in Appendix G.

Individual-level energy distribution

The METE prediction of the IED, the metabolic rate distribution for all individuals in a community, shows fairly good agreement with empirical data as a smooth, monotonically decreasing function of ranked metabolic rate, with increasingly negative slope (Fig. 2). Because the normalization of the IED requires the total predicted metabolic rate to equal the total metabolic rate of the empirical data, the curves necessarily intersect on this graph.

To test similarity between empirical and predicted curves of the IED, we construct a graph of log(observed) vs. log(predicted value) and note that a slope of nearly 1 and a high R^2 value (slope = 0.8362; R^2 = 0.9742) indicates that the IED performs well in predicting the metabolic rate distribution for all individuals in the community. However, a K-S test performed on 1000 bootstrapped samples does not support the null hypothesis that the observed and theoretical distributions are the same (D = 0.1653, P < 0.0001). But, whereas the K-S test may be more appropriate because it is less sensitive to normality of a distribution than some other goodness-of-fit metrics (like Shapiro-Wilk and Anderson-Darling), it tends to reject the null hypothesis at a high rate based on deviant values and has been criticized for producing misleading results (Babu and Feigelson 2006).

Average Species Energy Distribution

The distribution of average metabolic rates across species, the ASED, is predicted to have increasingly negative slope as a function of ranked energy if $\ln(\bar{\varepsilon})$ is plotted against rank, *r*, from largest average metabolic rate per species to smallest. Empirical data deviates from this pattern strongly, showing an inflected shape (Fig.3). A K-S test does not support the null hypothesis of no difference between observed and theoretical distributions (D = 0.5484, P < 0.0001, 1000 bootstrapped samples). Appendix H contains the predicted probability distribution for the ASED (*Nu*) and additional tests of goodness-of-fit. A notable feature of the ASED is that the empirical and predicted distributions do not have the same mean.

Species-level Energy Distributions

The METE-predicted distributions of metabolic rates across all individuals in a single species of abundance *n* are the SED distributions shown in Fig. 4 (for the 12 most abundant species in the plot). Graphs of SEDs are shown in order from smallest average photosynthetic area (ANDSEP, *Androsace septentrionalis*) to largest (HELQUI, *Helianthella quinquenervis*). Visual inspection of graphs of SEDs for these 12 most abundant species reveals a particular trend: METE overpredicts metabolic rates for species with low average size (size measured in normalized photosynthetic area), does well in approximating the shape and variance of mid-sized species, and underpredicts metabolic rates for species with large average size. This trend holds for all of the 12 most abundant species when ranked by average size of an individual by species, but is not generally upheld when all species are included (see Appendix D). Here we note a strong effect of abundance, in that this pattern is not upheld for species with 15-20 or fewer individuals (a limit which may be related to the log-normality of the empirical SAD, see Discussion).

This trend in the data is evident when the slopes of the log(observed) vs. log(predicted) values for each species are calculated, and then graphed in order from largest average metabolic rate to smallest average metabolic rate. See Fig. 5 and Appendix E, Figs. E.1-2. Species with individuals that are large on average are more likely to have slopes close to 1, and species that are smaller on average have much larger log(observed) vs. log(predicted) slopes. When log(slopes) are graphed against species ranked by average metabolic rate, a straight line regression is obtained with an R^2 value of 0.71. This regular deviation in slopes of predicted versus observed data on a log-log scale may reveal a bias in the theoretical prediction of the SEDs that is correlated to the average metabolic rate (or alternately, rank) of each species.

Discussion

The distributions IED*,* ASED*, and* SED

From goodness-of-fit tests, there is mixed evidence that the IED, a community-level metric that predicts the distribution of metabolic rates of all individuals, is successful. Although the IED fails the stringent K-S test, the shape of the predicted distribution is correctly monotonically decreasing with increasingly negative slope. The predicted IED is also constrained to have the same mean and total energy as the empirical data. The IED therefore shows promise as a candidate null model for an Individual Size Distribution (ISD) -type prediction, as discussed in White et al. 2007.

In its current mathematical formulation, the SED metric is not flexible enough to capture the range of body-sizes in the system we test here. Because the empirical data can be approximated with a decreasing linear fit on log(rank) vs. metabolic rate graphs (see Fig. 4), the exponential form of the SED predictions appears to be an adequate model, however, there is a clear pattern of bias in the slopes of the predicted SEDs related to average body size of a species as described above. This suggests that the exponential form of the SED may indeed be an adequate null model*,* but because the slopes of the predictions are incorrect, that the exponent itself could be adjusted to include information about the rank of the average metabolic rate of the species being considered (see Appendix E for further discussion). A modification of the exponent of the SED would give predictions closer to empirical values; however, tests of the SED with more than one dataset might better justify a change of the theoretical form.

Problems with the SED are expected to translate to problems with the ASED, because the ASED is the predicted distribution of the means of the SEDs. This, by itself, is a large enough effect to cause the ASED to fail, however, we also find that the means of the predicted and observed ASED do not match for unrelated reasons, and is the METE prediction therefore does not provide a function that is a useful description of this metric. This problem led us to look at a related metric, which we call the Total Species Energy Distribution (TSED), which has more tractable qualities as a community-level metabolic metric and reproduces results similar to energy equivalence as predicted by Damuth (1987). See Appendix K for a discussion of the failure of the ASED, and the properties and predictions of the TSED metric.

Distinguishing the role of disturbance from the failure of theory

Data from the subalpine meadow system show a log-normal trend in the empirical SAD rather than the predicted Fisher log-series curve (AIC evidence ratio = 20.09 in favor of the log-normal), with fewer than predicted singleton species (1 rather than 5); see Appendix I. A log-normal distribution of cover by species (Bazzaz 1975) or abundances is often observed in systems that have experienced recent ecological perturbations or natural disturbances (Hill and Hamer 1998, Kempton and Taylor 1974, Newman et al., *in preparation*, but see Ulrich et al. 2010). During our study year, unusually high early-season temperatures and lower than average snowpack in the previous winter likely had a perturbing effect on the plant community. In surrounding areas, novel plant communities arose in the year of this study as some plants failed to develop after early freezes, others developed according to photoperiod triggers (David Inouye, *pers. comm*.), and some plants showed advanced phenology due to higher than average growing-

season temperatures, and earlier than average snowmelt, a major influence on plant phenology in the region (Dunne et al. 2003).

From this empirical deviation from theory, we could conclude that the values and ratios of the four state variables (A_0, S_0, N_0 and E_0) are not sufficient to describe the dynamics of a real system, such as natural disturbance and post-disturbance. It has been suggested elsewhere that evolutionary processes including diversification and extinction may influence the shape of empirical macroecological distributions (Hubbell 2001, Harte 2011, Harte et al. 2013). These explanations, while plausible, would not allow us to distinguish between the role of disturbance in causing theory to fail, versus an inherent flaw with the mathematical construction of theory, so we investigate incorporating the empirical SAD into the METE predictions here.

Because of the aforementioned ecological perturbation and the community restructuring effect it may have had, we incorporate the empirical SAD in the elements of the SEDs and ASED that are explicitly dependent on the SAD (see Eq. 4). Predictions of the SEDs and ASED that result from using the empirical SAD and calculating the metrics numerically are presented in Appendix J, with the ASED and example SED shown in Figs. 6a-b. These analyses, while debatably not a direct test of METE's predictions, have allowed us determine that the log-normality of the empirical SAD does not greatly change the shape of the predicted SEDs and ASED, and merely shift the predictions by some normalization constant. The empirical SAD is therefore not the cause of the mismatch between data and theory. We conclude that the problem is likely located within the prediction of the SED because its exponential form appears to be correct except for the predicted slopes and intercepts of the SEDs (see Figs. 4 and 5). We therefore suspect that the predicted exponent of the SED equation is incorrect. This would further explain why the SEDs and ASED jointly fail. In spite of this, the IED remains a largely successful metric because it is free of this particular problem. These results are consistent with an independent study of tree communities (Xiao et al. 2013), suggesting that the observed relationships between the theoretical predictions and empirical data may be general, at least within plants.

Conclusions

The Maximum Entropy Theory of Ecology provides a theoretical basis for estimating various metrics of energetics in biotic communities. Empirical data from a subalpine meadow provide the first explorations of the realism of two of these predictions in a fully-censused community. Although we chose an herbaceous plant community because of its tractability and similarity of growth forms within the community, predictions of the community- and species-level metrics considered in this paper should hold generally for any taxon or group of taxa. Future work may reexamine the predicted forms of the metabolic metrics presented here. The IED, the metabolic rate distribution for all individuals in a community, appears successful as a theoretically-predicted measure of the individual energy distribution by linear comparisons of predicted versus observed data.

Given the poor performance of the METE SAD for this system, we incorporated the empirical, lognormal SAD and found that functional forms of the predicted SEDs are not highly affected by the empirical SAD. By association, METE's metabolic predictions should hold in ecosystems where the shape of the SAD is altered due to natural disturbance or ecological

perturbation, succession, or a variety of other ecological factors (Dornelas et al. 2009). Because it appears that the SED is exponential in form, the exponent itself ($\lambda_2 n$) is a likely cause of the highly regular pattern in mismatch between theory and data. The information included in the exponent therefore becomes the target for more theoretical inquiry and empirical testing.

Finally, considering the ASED, we show that the incorporation of the empirical SAD may provide predictions that are closer to empirical data; however, it METE fails to predict the correct form of this community-level metric, likely due to the ASED relationship to the SEDs. The predicted ASED also does not necessarily have the same mean as the empirical distribution. A simple conversion of the ASED recovers an energetic-equivalence prediction, the TSED, that is well-normalized and similar to the Damuth Rule. The TSED might be an effective prediction of global-scale metabolic rate distributions, but does not perform well for plot-level metabolic rate distributions.

Acknowledgements

We thank the Rocky Mountain Biological Laboratory for providing permitting, logistical support and facilities, and Gunnison National Forest for providing field sites. I thank J. Reithel, A.F. Messerman, D. Hembry, J. Kitzes, C. Levy, and A. Rominger for their assistance, and D. Hembry, E. P. White and X. Xiao for manuscript comments. B. Blonder provided analysis code and much appreciated guidance on methods. Zack Johnson contributed art appearing in this article. Funding was provided by the NSF in the form of the Graduate Research Fellowship Program, Research Experience for Undergraduates program, and grant NSF-EF-1137685.

Supplementary Material to Chapter 3

Appendix A. Measurements and calculations of photosynthetic area

Appendix B. SED predictions and observed values for all species with 5 or more individuals, including comparisons of predicted vs. observed values

Appendix C. Table. Species included in this study, ranked by average metabolic rate per species (a normalized measure of photosynthetic area), from largest to smallest

Appendix D. SED predictions and observed values for the 15 most abundant species

Appendix E. Additional analysis of SEDs

Appendix F. The functional form of the ASED (*Nu*) as derived from a discrete cumulative distribution function

Appendix G. Kolmogorov-Smirnov goodness-of-fit tests for ASED, IED, and all SEDs

Appendix H. Additional information on the Average Species Energy Distribution, including tests of goodness-of-fit

Appendix I. The Species-Abundance Distribution of a subalpine plant community

Appendix J. Incorporating the empirical SAD into ASED and SED predictions

Appendix K. Discussion of the normalization of the ASED, and exploration of a related metric, the Total Species Energy Distribution

Chapter 3 Tables and Figures

Tables.

Table 1. Species names, abundances, rank, and leaf-area measurements of 31 species of plants in a sub-alpine plot. All metabolic rates are in normalized units of leaf area.

6-Letter code	Species	Common name	Rank in plot by *n*	*n*	Minimum metabolic rate	Maximum metabolic rate	Average metabolic rate
BOEDRU	*Boechera stricta* (syn. *Boechera drumondii*)	Tower mustard	1	103	4.694	191.013	33.97
HELQUI	*Helianthella quinquenervis*	Five nerve heliantella	2	94	77.013	38251.85	4680.185
VIONUT	*Crocion nuttallii* (syn. *Viola nuttallii*)	Yellow montaine violet	3	93	1	1224.081	227.33
LUPARG	*Lupinus argenteus*	Silvery lupine	4	85	6.734	28164.32	736.251
ANDSEP	*Androsace septentrionalis*	Northern rock jasmine	5	62	1.134	35.431	8.193
HYMHOO	*Hymenoxys hoopesyii*	Sneezeweed	6	55	32.363	5719.091	956.284
FRAVES	*Fragaria vesca*	Alpine strawberry	7	54	15.119	1223.867	207.54
POTGRA	*Potentilla gracilis* var. *nuttallii* (syn. *Potentilla gracilis*)	Slender cinquefoil	8	47	8.443	4171.999	728.91
LIGPOR	*Ligusticum porteri* (syn. *Ligusticum porteri* var. *brevilobum*)	Osha	9	38	3.424	21632.616	3171.588
VICAME	*Vicia americana*	American vetch	10	33	6.248	582.964	128.913
IPOAGG	*Ipomopsis aggregata*	Scarlet gilia	11	23	7.075	467.952	72.744

Table 1 (*continued*)

6-Letter code	Species	Common name	Rank in plot by *n*	*n*	Minimum metabolic rate	Maximum metabolic rate	Average metabolic rate
THLMON	*Thlaspi montanum* (syn. *Noccaea montana*)	Candytuft, Alpine pennycress	12	20	5.821	1494.01	118.013
SENCRA	*Senecio crassulus*	Thick-leaved groundsel	13	19	8.018	1115.985	262.59
DELBAR	*Delphinium barbeyi*	Tall larkspur	14	16	73.235	17051.759	2243.05
ERIFOR	*Erigeron formosissimus*	Beautiful fleabane	15	14	14.089	14552.655	1578.719
LATLEU	*Lathyrus leucanthus* (syn. *Lathyrus lanzwertii* var. *leucanthus*)	Peavine, Rocky Mountain sweetpea	16	13	63.034	1857.99	564.596
SOLMUL	*Solidago multiradiata*	Goldenrod	17	12	7.449	408.856	109.939
POLDOU	*Polygonum douglasii*	Douglas's knotweed	18	12	5.862	157.921	35.843
ERIELA	*Erigeron elatior*	Pink-headed fleabane	19	11	15.921	1443.983	477.113
AGOGLA	*Agoseris glauca*	Mountain dandelion	20	11	20.513	714.54	192.236
DRASPE	*Draba spectabilis*	Showy draba	21	10	27.369	159.939	73.597
CASSUL	*Castilleja sulphurea*	Sulphur paintbrush	22	10	12.864	1805.139	537.313
EUCENG	*Eucephalus engelmannii* (syn. *Aster engelmannii*)	Englemann's aster	23	8	70.823	1589.375	607.619

Table 1 (*continued*)

6-Letter code	Species	Common name	Rank in plot by *n*	*n*	Minimum metabolic rate	Maximum metabolic rate	Average metabolic rate
ERISPE	*Erigeron speciosus*	Showy fleabane	24	8	65.799	1415.119	573.524
SENAMP	*Senecio amplectens* (syn. *Ligularia holmii*)	Showy alpine ragwort	25	7	39.841	1108.729	282.468
VIGMUL	*Heliomeris multiflora* (syn. *Viguiera multiflora*)	Showy goldeneye	28	2	294.612	458.466	376.539
OSMOCC	*Washingtonia occidentalis* (syn. *Osmorhiza occidentalis*)	Sweet cicely	29	2	4972.977	32451.402	18712.189
AQUCOE	*Aquilegia flavescens* (syn. *Aquilegia coerulea*)	Colorado columbine	30	2	24.894	23956.678	11990.786
TAROFF	*Taraxacum officinale*‡	Common dandelion	31	1	9.537	9.537	9.537

‡ indicates an exotic species

Figures.

Figure 1. Visual representation of what is modeled by (a) the Individual-level Energy Distribution (IED), (b) the distribution of average metabolic rates across species or Average Species Energy Distribution (ASED), and (c) one example of the distribution of metabolic rates across the individuals in a single species of abundance *n*, also known as the Species-level Energy Distribution (SED).

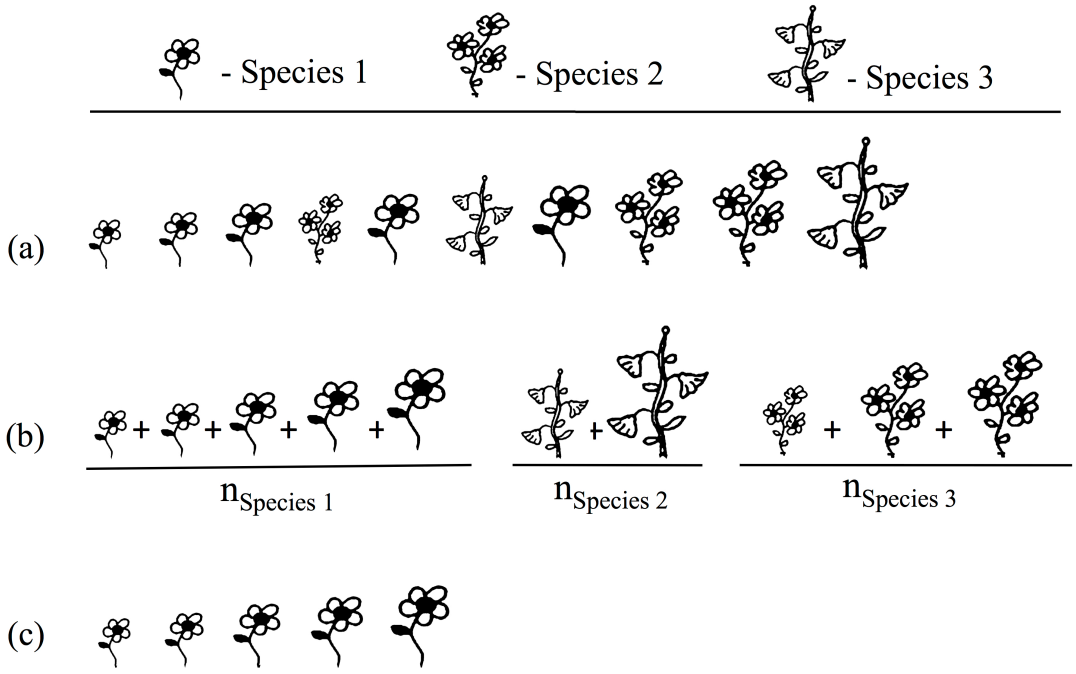


Figure 2. Comparison of empirical data to METE Individual-level Energy Distribution predictions for a subalpine plant community. This graph shows empirical and predicted IED values on a log(rank) versus log(metabolic rate) graph for all individuals in the plot.

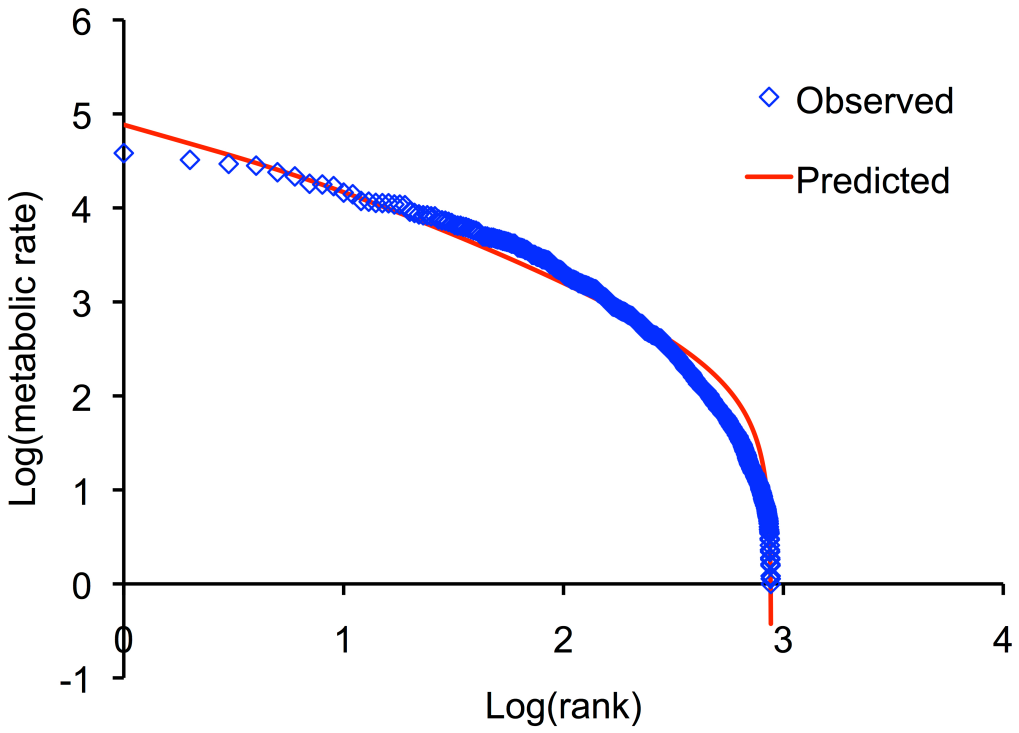


Figure 3. ASED predictions and empirical data for a subalpine plant community where species are ranked by average metabolic rate on a rank-log(metabolic rate) plot.

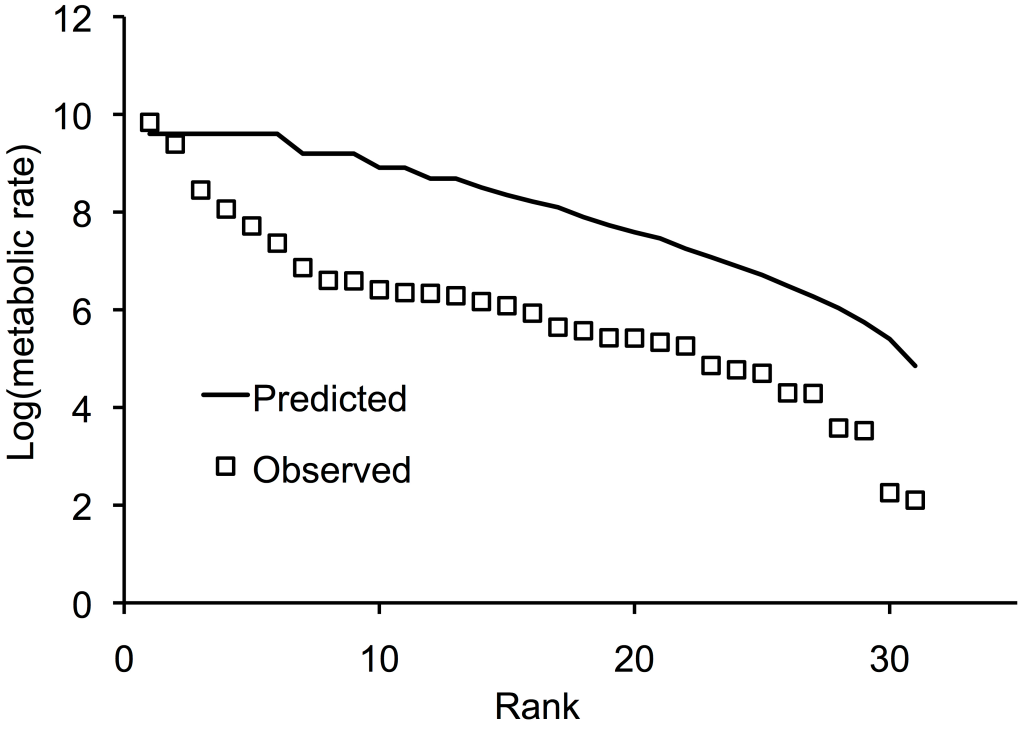


Figure 4. METE predicted and observed Species-level Energy Distributions (SED) for all individuals of the 12 most abundant species in the subalpine plot. Individuals are ranked from largest to smallest metabolic rate within the species, and graphs are presented in order from smallest to largest average metabolic rate per species.

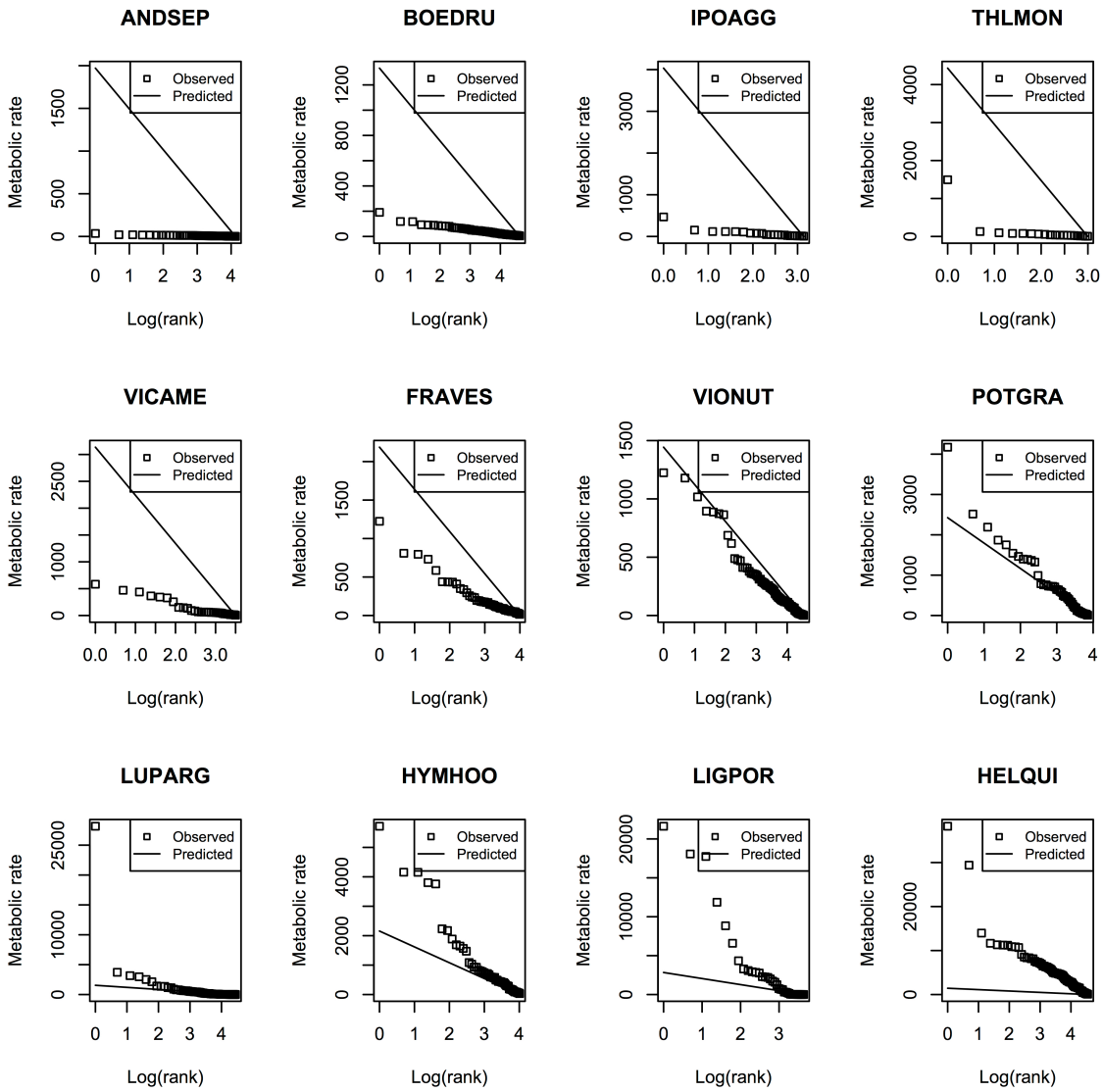


Figure 5. Slope of predicted versus observed data for all SEDs with n ≥ 5 on a log-rank graph. Each point represents one species. A good fit would be characterized by a trendline with slope = 0.

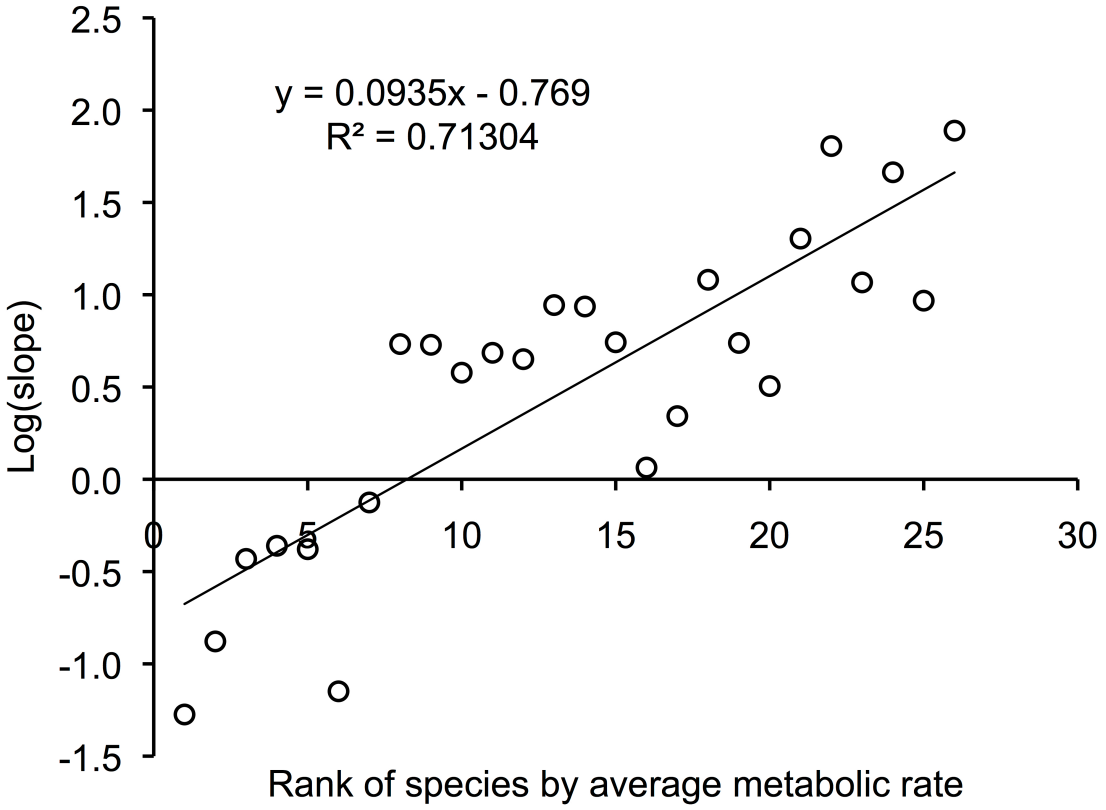
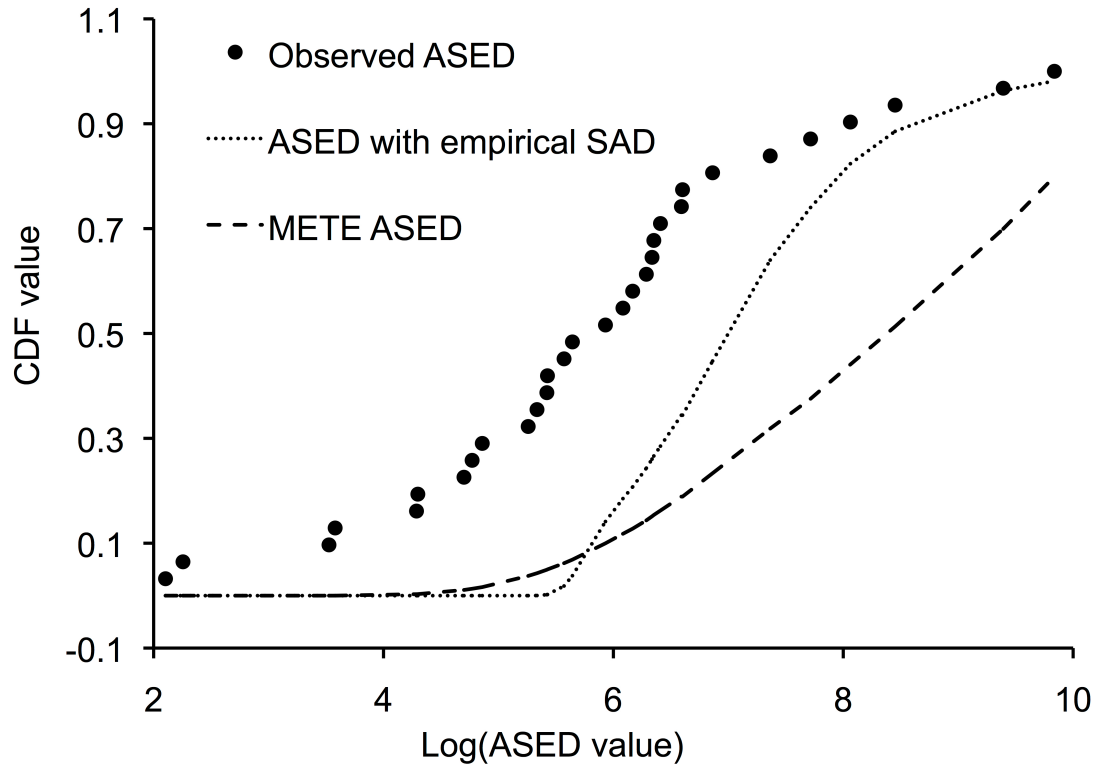
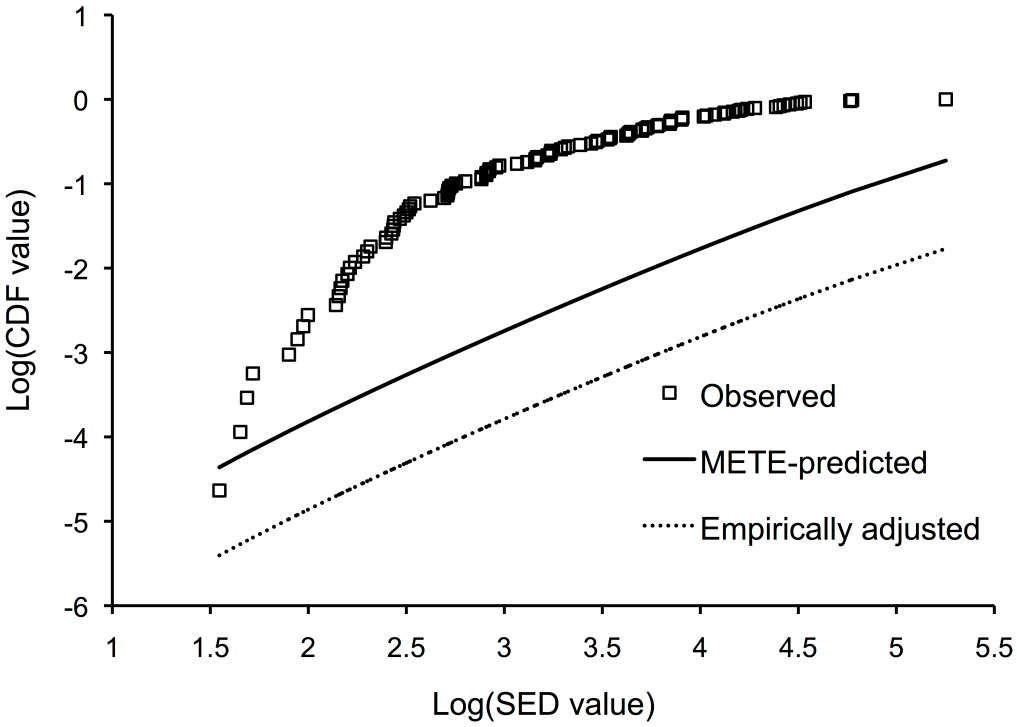


Figure 6(a-b). ASED and example SED showing a comparison of observed values, the METE prediction including the empirical adjustment, and the METE prediction. (a) Cumulative Density Function (CDF) of the ASED. (b) CDF of the SED for *Boechera stricta*.

(a)



(b)


1
0
-1
-2
-3
-4
-5
-6
Log(CDF value)
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
Log(SED value)
Observed
METE-predicted
Empirically adjusted

References

References for Chapter 1

Disturbance macroecology: a comparative study of plant species' abundances and distributions in different-age post-fire stands of Bishop Pine (*Pinus muricata*)

Alroy, J. (2015). The shape of terrestrial abundance distributions. *Science Advances*, 1(8), e1500082.

Baldridge, E., Xiao, X., & White, E.P. (2015). An extensive comparison of species-abundance distribution models. *bioRxiv*, 024802.

Baldwin, B.G., Goldman, D.H., Keil D.J., Patterson, R., Rosatti ,T.J., & Wilken, D.H. (2012). *The Jepson manual: vascular plants of California*, 2nd edn. University of California Press, Berkeley.

Batllori, E., Parisien, M.A., Krawchuk, M.A., & Moritz, M.A. (2013). Climate change‐induced shifts in fire for Mediterranean ecosystems. *Global Ecology and Biogeography*, 22(10), 1118-1129.

Bazzaz. F.A. (1975). Plant species diversity in old-field successional ecosystems in southern Illinois. *Ecology*, 56, 485-488.

Bormann, F. H., and G. E. Likens. 1979. Pattern and process in a forested ecosystem. Springer-Verlag, Berlin, Germany.

Brown, J. H. (1995). *Macroecology*. University of Chicago Press, Chicago.

Brown, P. M., Kaye, M. W., & Buckley, D. (1999). Fire history in Douglas-fir and coast redwood forests at Point Reyes National Seashore, California. Northwest Science, 73(3), 205-216.

Condit, R. (1998). Tropical forest census plots: methods and results from Barro Colorado Island, Panama and a comparison with other plots. Springer Science & Business Media.

Dawson, T. E. (1998). Fog in the California redwood forest: ecosystem inputs and use by plants. *Oecologia*, 117(4), 476-485.

Death, R. G. (1996). The effect of habitat stability on benthic invertebrate communities: the utility of species abundance distributions. *Hydrobiologia*, 317(2), 97-107.

Dennis, B. & Patil, G.P. (1979). Species abundance, diversity, and environmental predictability. In: Ecological Diversity in Theory and Practice (Ed. by J.F. Grassle, G. P. Patil, W. Smith and C. Taillie), pp. 93-114. International Cooperative Publishing House, Fairland, MD.

Fisher, J.A.D., Frank, K.T. & Leggett, W.C. (2010). Dynamic macroecology on ecological time-scales. *Global Ecology and Biogeography,* 19, 1-15.

Forrestel, A.B., Moritz, M.A. & Stephens, S.L. (2011). Landscape-scale vegetation change following fire in Point Reyes, California, USA. *Fire Ecology*, 7(2), 114-128.

Forrestel, A.B., Ramage, B.S., Moody, T., Moritz, M.A., & Stephens, S.L. (2015). Disease, fuels and potential fire behavior: Impacts of Sudden Oak Death in two coastal California forest types. *Forest Ecology and Management*, 348: 23-30.

Fox, J. W. (2013). The intermediate disturbance hypothesis should be abandoned. *Trends in Ecology & Evolution*, 28(2), 86-92.

Gray, J.S. & Mirza, F.B. (1979). A possible method for the detection of pollution-induced disturbance on marine benthic communities. *Marine Pollution Bulletin*, 10(5), 142-146.

Gray, J.S. (1981). Detecting pollution-induced changes in communities using the log-normal distribution of individuals among species. *Marine Pollution Bulletin*. 12, 173-176.

Hamer, K.C., and Hill, J.K. (2000). Scale‐dependent effects of habitat disturbance on species richness in tropical forests. *Conservation Biology*, 14(5): 1435-1440.

Harte, J. (2011). Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics. Oxford Univ. Press, Oxford UK.

Harte, J., & Kitzes, J. (2015). Inferring regional-scale species diversity from small-plot censuses. *PLOS one*, 10(2), e0117527.

Harte, J., Kitzes, J., Newman, E.A., & Rominger, A.J. (2013). Taxon categories and the universal species-area relationship. *The American Naturalist*, 181(2), 282-287.

Harte, J., & Newman, E.A. (2014). Maximum information entropy: a foundation for ecological theory. *Trends in Ecology & Evolution*. 29(7): 384-389

Harte, J., Smith, A., & Storch, D. (2009). Biodiversity scales from plots to biomes with a universal species-area curve. *Ecology Letters*, 12: 789-797.

Harte, J., Zillio, T., Conlisk, E. & Smith, A. (2008). Maximum entropy and the state variable approach to macroecology. *Ecology*, 89: 2700-2711.

Harvey, B.J., Holzman, B.A. & Davis, J.D. (2011). Spatial variability in stand structure and density-dependent mortality in newly established post-fire stands of a California closed-cone pine forest. *Forest Ecology and Management*, 262(11): 2042–2051.

Harvey, B.J., Holzman, B.A., & Forrestel, A.B. (2014). Forest resilience following severe wildfire in a semi-urban National Park. *Fremontia*, 42(3): 14-18.

Howell, J.T., Follette, W., Best, C. & Almeda, F. (2007). Marin Flora: An illustrated manual of flowering plants, ferns, and conifers of Marin County, California. San Francisco: California Academy of Sciences and California Native Plant Society.

Hill, J.K., Hamer, K.C., Lace, L.A., & Banham, W.M.T. (1995). Effects of selective logging on tropical forest butterflies on Buru, Indonesia. *Journal of Applied Ecology*, 754-760.

Hill, J. K., & K.C. Hamer (1998). Using species abundance models as indicators of habitat disturbance in tropical forests. *Journal of Applied Ecology*, 35: 458-460.

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

Jaynes, E.T. (1982). On the rationale of maximum entropy methods. *Proc. Instit. Elec. Electron. Eng.*, 70: 939-952.

Keator, G., & Heady, R. M. (1981). *Pacific Coast Fern Finder*. Nature Study Guild.

Kempton, R.A., & L.R. Taylor (1974). Log-series and log-normal parameters as diversity discriminants for the Lepidoptera. *Journal of Animal Ecology*, 43(2): 381-399

Kitzes, J., Lewis, C. & Wilber, M. (2014). macroeco v0.2. http://github.com/jkitzes/macroeco. Accessed online April 2013.

Kitzes, J. & Wilber, M. (2016). macroeco: Reproducible ecological pattern analysis in Python. *Ecography*.

Little, E.L., Jr. (1971). Atlas of United States trees, volume 1, conifers and important hardwoods: U.S. Department of Agriculture Miscellaneous Publication 1146, 9 p., 200 maps.

May, R. Patterns of species abundance and diversity. (1975). In: *Ecology and evolution of communities* (eds. Cody, M. L., & Diamond, J. M.). Harvard University Press, Cambridge, MA.

Mayor, S.J., Cahill Jr, J.F., He, F., & Boutin, S. (2015). Scaling Disturbance Instead of Richness to Better Understand Anthropogenic Impacts on Biodiversity. *PLOS one*, 10(5).

McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K., Dornelas, M., Enquist, B.J., Green, J.L., He, F. & Hurlbert, A.H. (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology letters*, 10(10), 995-1015.

McGlinn, D.J., Xiao, X., & White, E.P. (2013). An empirical evaluation of four variants of a universal species–area relationship. *PeerJ*, 1, e212.

McGlinn, D.J., Xiao, X., Kitzes, J., & White, E.P. (2015). Exploring the spatially explicit predictions of the Maximum Entropy Theory of Ecology. *Global Ecology and Biogeography*, 24(6), 675-684.

Millar, C.I. (1983). A steep cline in *Pinus muricata*. *Evolution*, 311-319.

Millar, C.I. (1986). The Californian closed cone pines (subsection Oocarpae Little and Critchfield): a taxonomic history and review. *Taxon*, 657-670.

National Park Service. "Vision Fire 10 Year Anniversary Fact Sheet #3—Incident Statistics." Retrieved online (October 26, 2015) at: http://www.nps.gov/pore/learn/management/upload/firemanagement_visionfire_factsheet3.pdf

Newman, E.A., Harte, M.E., Lowell, N., Wilber, M. & Harte, J. (2014). Empirical tests of within-and across-species energetics in a diverse plant community. *Ecology*, 95(10), 2815-2825.

Noss, R.F., LaRoe, E.T., Scott, J.M. (1995) Endangered ecosystems of the United States: A preliminary assessment of loss and degradation. U.S. Department of the Interior, National Biological Service.

Nummelin, M. (1998). Log-normal distribution of species abundances is not a universal indicator of rain forest disturbance. *Journal of Applied Ecology*, 35(3), 454-457.

Pickett S.T.A., & White P.S., eds. (1985). *The ecology of natural disturbance and patch dynamics*. Academic Press, Orlando, FL.

R Core Team (2013-2015). (Versions 3.0.1 and 3.1.3) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at http://www.Rproject.org/

Romme, W. H., Everham, E. H., Frelich, L. E., Moritz, M. A., & Sparks, R. E. (1998). Are large, infrequent disturbances qualitatively different from small, frequent disturbances?. *Ecosystems*, 1(6), 524-534.

Rominger, A.J., Goodman, K.R., Lim, J.Y., Armstrong, E.E., Becking, L.E., Bennett, G.M., ... & Gillespie, R.G. (2016). Community assembly on isolated islands: macroecology meets evolution. *Global Ecology and Biogeography*.

Rominger, A.J., Harte, J., Gruner, D. S, and Gillespie, R. G. *in prep*. The statistical mechanics of biodiversity in evolving island communities.

Schlossberg, S., & King, D. I. (2015). Measuring the effectiveness of conservation programs for shrubland birds. *Global Ecology and Conservation*, 4, 658-665.

Sheil, D., & Burslem, D. F. R. P. (2013). Defining and defending Connell's intermediate disturbance hypothesis: a response to Fox. *Trends in Ecology & Evolution*, 28(10), 571-572.

Stephens, S.L., & Libby, W.J. (2006). Anthropogenic fire and bark thickness in coastal and island pine populations from Alta and Baja California. *Journal of biogeography*, 33(4), 648-652.

Supp, S.R., Xiao, X., Ernest, S.K.M., & White, E.P. (2012). An experimental test of the response of macroecological patterns to altered species interactions. *Ecology*, 93(12), 2505-2511.

Supp, S.R., & Ernest, S.M. (2014). Species-level and community-level responses to disturbance: a cross-community analysis. *Ecology*, 95(7), 1717-1723.

Tokeshi, M. (1993). Species abundance patterns and community structure. *Advances in ecological research*, 24, 111-186.

Turner, M.G. (1989). Landscape ecology: the effect of pattern on process. *Annual review of ecology and systematics*, 171-197.

Turner, M.G. (2010). Disturbance and landscape dynamics in a changing world 1. *Ecology*, 91(10), 2833-2849.

Turner, M.G., Baker, W.L., Peterson, C.J. & Peet, R.K. (1998). Factors influencing succession: lessons from large, infrequent natural disturbances. *Ecosystems*, 1(6), 511-523.

Turner, M. G., & Dale, V. H. (1998). Comparing large, infrequent disturbances: what have we learned?. *Ecosystems*, *1*(6), 493-496.

Turner, M.G., Romme, W.H., Gardner, R.H., O'Neill, R.V., & Kratz, T.K. (1993). A revised concept of landscape equilibrium: disturbance and stability on scaled landscapes. *Landscape Ecology*, 8(3), 213-227.

Ulrich, W., Ollik, M. & Ugland, K.I. (2010). A meta-analysis of species–abundance distributions. *Oikos* 119, 1149-1155.

van Rossum, G. & Drake, F.L. (eds.) (2001). Python Reference Manual, PythonLabs, Virginia, USA. Available at http://www.python.org

White, E.P., Thibault, K.M. & Xiao, X. (2012). Characterizing species abundance distributions across taxa and ecosystems using a simple maximum entropy model. *Ecology* 93, 1772-1778.

White, P.S. & Jentsch, A. (2001). The search for generality in studies of disturbance and ecosystem dynamics. In: *Progress in Botany*. Springer, Berlin Heidelberg, pp. 399-450.

Whittaker, R.H. (1965). Dominance and Diversity in Land Plant Communities: Numerical relations of species express the importance of competition in community function and evolution. *Science*, 147(3655), 250-260.

Whittaker, R.H. (1975). Communities and Ecosystems, 2nd edn. Macmillan, New York.

Wilber, M. Q., Kitzes, J., & Harte, J. (2015). Scale collapse and the emergence of the power law species–area relationship. *Global Ecology and Biogeography*, 24(8), 883-895.

Wu, J. (2004). Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecology*, *19*(2), 125-138.

Wu, J., & Loucks, O. L. (1995). From balance of nature to hierarchical patch dynamics: a paradigm shift in ecology. *Quarterly review of biology*, 439-466.

Xiao, X., McGlinn, D.J., & White, E.P. (2015). A strong test of the maximum entropy theory of ecology. *The American Naturalist*, 185(3), E70-E80.

References for Chapter 2

Macroecology for management: Testing an information-entropy-based theory of macroecology against anthropogenic disruption of high-Sierra meadows

Allen-Diaz, B.H. (1991). Water tables and plant species relationships in Sierra Nevada meadows. *American Midland Naturalist* 126: 30–43

Baldridge, E., Xiao, X., & White, E. P. (2015). An extensive comparison of species-abundance distribution models. *bioRxiv*, 024802.

Bazzaz, F.A. (1975). Plant species diversity in old-field successional ecosystems in southern Illinois. *Ecology*, 56, 485-488.

Benedict, N. B. (1983). Plant associations of subalpine meadows, Sequoia National Park, California. *Arctic and Alpine Research*, 383-396.

Berlow, E. L., D'Antonio, C. M., & Reynolds, S. A. (2002). Shrub expansion in montane meadows: the interaction of local-scale disturbance and site aridity. *Ecological Applications*, *12*(4), 1103-1118.

Berlow, E. L., D'Antonio, C. M., & Swartz, H. (2003). Response of herbs to shrub removal across natural and experimental variation in soil moisture. *Ecological Applications*, *13*(5), 1375-1387.

Berlow, E. L., Knapp, R. A., Ostoja, S. M., Williams, R. J., McKenny, H., Matchett, J. R., Guo, Q., Fellers, G. M., Kleeman, P., Brooks, M. L. & Joppa, L. (2013). A network extension of species occupancy models in a patchy environment applied to the Yosemite toad (*Anaxyrus canorus*). *PloS one*, *8*(8), p.e72200.

Brown, J. H. (1995). *Macroecology*. University of Chicago Press, Chicago.

Cole, D.N., Van Wagtendonk, J.W., McClaran, M.P., Moore, P.E. & McDougald, N.K. 2004. Response of mountain meadows to grazing by recreational pack stock. *Journal of Range Management* 57: 153–160.

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature Climate Change*, *3*(1), 52-58.

D'Antonio, C. M., Berlow, E. L., & Haubensak, K. L. (2004). Invasive exotic plant species in Sierra Nevada ecosystems.

Darrouzet-Nardi, A., D'Antonio, C. M., & Dawson, T. E. (2006). Depth of water acquisition by invading shrubs and resident herbs in a Sierra Nevada meadow. *Plant and Soil*, *285*(1-2), 31-43.

DeBenedetti, S. H., & Parsons, D. J. (1979). Mountain meadow management and research in Sequoia and Kings Canyon National Parks: a review and update. In *Proceedings of the 1st Conference on Scientific Research in the National Parks* (pp. 1305-1311).

Fites-Kaufman, J.A., Rundel, P., Stephenson, N. & Weixelman, D.A. 2007. Montane and subalpine vegetation of the Sierra Nevada and Cascade ranges. In: Barbour, M.G., Keeler-Wolf, T., Schoenherr, A. (eds) *Terrestrial vegetation of California.* 3rd edition. pp. 456–501. University of California Press, Berkeley, CA, US.

Fisher, J.A.D., Frank, K.T. & Leggett, W.C. (2010). Dynamic macroecology on ecological time-scales. *Global Ecology and Biogeography,* 19, 1-15.

Gaston, K., & Blackburn, T. (2008). *Pattern and process in macroecology*. John Wiley & Sons.

Génin, A. (*in prep*). Identifying early warning signs of catastrophic shifts in alpine meadow ecosystems.

Gray, J.S. & Mirza, F.B. (1979). A possible method for the detection of pollution-induced disturbance on marine benthic communities. *Marine Pollution Bulletin*, 10(5), 142-146.

Halpern, C.B. (1985). Cooperative National Park Resources Studies Unit, Technical Report No. 20. Hydric Montane Meadows of Sequoia National Park, California: A Literature Review and Classification.

Harte, J. (2011). Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics. Oxford Univ. Press, Oxford UK.

Harte, J., & Kitzes, J. (2015). Inferring regional-scale species diversity from small-plot censuses. *PloS one*, *10*(2), e0117527.

Harte, J., Kitzes, J., Newman, E.A., & Rominger, A.J. (2013). Taxon categories and the universal species-area relationship. *The American Naturalist*, 181(2), 282-287.

Harte, J., & Newman, E.A. (2014). Maximum information entropy: a foundation for ecological theory. *Trends in Ecology & Evolution*. 29(7): 384-389

Harte, J., Smith, A. B., & Storch, D. (2009). Biodiversity scales from plots to biomes with a universal species–area curve. *Ecology letters*, *12*(8), 789-797.

Harte, J., Zillio, T., Conlisk, E., & Smith, A. B. (2008). Maximum entropy and the state-variable approach to macroecology. *Ecology*, *89*(10), 2700-2711.

Hillebrand, H., Bennett, D. M., & Cadotte, M. W. (2008). Consequences of dominance: a review of evenness effects on local and regional ecosystem processes. *Ecology*, *89*(6), 1510-1520.

Holmquist, J. G., Schmidt-Gengenbach, J., & Haultain, S. A. (2010). Does long-term grazing by pack stock in subalpine wet meadows result in lasting effects on arthropod assemblages?. *Wetlands*, *30*(2), 252-262.

Hopkinson, P., Hammond, M. & Bartolome, J. 2013. *SEKI meadows: paired grazed/ungrazed meadows species composition*. In: Sequoia and Kings Canyon National Parks Natural Resource Condition Assessment. National Park Service, Fort Collins, Colorado.

Keeler-Wolf, T., Moore, P.E., Reyes, E.T., Menke, J.M., Johnson D.N. & Karavidas D.L. (2012). Yosemite National Park vegetation classification and mapping project report. Natural Resource Technical Report NPS/YOSE/NRTR—2012/598. National Park Service, Fort Collins, Colorado.

Kempton, R.A., & L.R. Taylor (1974). Log-series and log-normal parameters as diversity discriminants for the Lepidoptera. *Journal of Animal Ecology*, 43(2): 381-399

Kitzes, J., Lewis, C. & Wilber, M. (2014). macroeco v0.2. http://github.com/jkitzes/macroeco. Accessed online April 2013.

Kitzes, J. & Wilber, M. (2016). macroeco: Reproducible ecological pattern analysis in Python. *Ecography*.

Lee, S.R. (2013). Contemporary pack stock effects on subalpine meadow plant communities in Sequoia and Yosemite National Parks. (2013). *Masters thesis.* University of California, Merced.

Matchett, J. R., Stark, P. B., Ostoja, S. M., Knapp, R. A., McKenny, H. C., Brooks, M. L., ... & Berlow, E. L. (2015). Detecting the influence of rare stressors on rare species in Yosemite National Park using a novel stratified permutation test. *Scientific reports*, *5*.

McClaran, M. P., & Cole, D. N. (1993). *Packstock in wilderness: use, impacts, monitoring, and management* (No. 04; USDA, FOLLETO 151.). US Department of Agriculture, Forest Service, Intermountain Research Station.

Moore, P. E., Cole, D. N., Van Wagtendonk, J. W., & McClaran, M. P. (2000). Meadow response to pack stock grazing in the Yosemite wilderness: integrating research and management.

Newman, E.A., Harte, M.E., Lowell, N., Wilber, M. & Harte, J. (2014). Empirical tests of within-and across-species energetics in a diverse plant community. *Ecology*, 95(10), 2815-2825.

Noss, R. F. (1983). A regional landscape approach to maintain diversity. *BioScience*, *33*(11), 700-706.

Odion, D. C., Dudley, T. L., & D'Antonio, C. M. (1988). Cattle grazing in southeastern Sierran meadows: ecosystem change and prospects for recovery. *Plant biology of eastern California. White Mountain Res. Station, Univ. Calif. Los Angeles*, 277-292.

Oksanen J, Guillaume Blanchet F, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2015) vegan: Community Ecology Package. R package version 2.0-10. Available online: http://CRAN.R-project.org/package=vegan.

Ostoja, S. M., Brooks, M. L., Moore, P. E., Berlow, E. L., Blank, R., Roche, J., ... & Haultain, S. (2014). Potential environmental effects of pack stock on meadow ecosystems of the Sierra Nevada, USA. *The Rangeland Journal*, *36*(5), 411-427.

Pickett, S. T., & Rogers, K. H. (1997). Patch dynamics: the transformation of landscape structure and function. In *Wildlife and landscape ecology* (pp. 101-127). Springer New York.

Pickett S.T.A., & White P.S., eds. (1985). *The ecology of natural disturbance and patch dynamics*. Academic Press, Orlando, FL.

Pyrooz (2015). Wet Meadow and Fen Mapping of Sequoia and Kings Canyon National Parks. National Park Service, U.S. Department of the Interior, Natural Resource Report NPS/SIEN/NRR-2015/968.

R Core Team (2015). (Version 3.2.1) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at http://www.Rproject.org/

Ratliff, R. D. 1985. Meadows in the Sierra Nevada of California: state of knowledge. Gen. Tech. Rep. PSW-GTR-84. Berkeley, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Forest and Range Experiment Station; 52 p

Shyrock, D.F. 2010. Influence of hydrology and recreational pack stock grazing on subalpine meadows of the John Muir and Ansel Adams Wilderness Areas, California. Master's thesis, Humboldt State University, Arcata, CA.

Viers, J. H., Purdy, S. E., and Peek, R. A., Fryioff-Hung, A., Santos, N. R., Katz, J. V. E., Emmons, J. D., Dolan, D. V., and Yarnell, S. M. (2013). Montane meadows in the Sierra Nevada: changing hydroclimatic conditions and concepts for vulnerability assessment. Centre for Watershed Sciences Technical Report (CWS-2013-01). University of California, Davis, CA.

Weixelman, D.A., Hill, B., Cooper, D.J., Berlow, E.L., Viers, J.H., Purdy, S.E., Merrill, A.G., & Gross, S.E. 2011. *A field key to meadow hydrogeomorphic types for the Sierra Nevada and Southern Cascade Ranges in California*. General Technical Report. R5-TP-034. Vallejo, CA. U.S. Department of Agriculture, Forest Service, Pacific Southwest Region, 34 pp.

White, E. P., Thibault, K. M., & Xiao, X. (2012). Characterizing species abundance distributions across taxa and ecosystems using a simple maximum entropy model. *Ecology*, *93*(8), 1772-1778.

Whittaker, R.H. (1975). Communities and Ecosystems, 2nd edn. Macmillan, New York.

Wood, S.H. 1975. *Holocene stratigraphy and chronology of mountain meadows, Sierra Nevada, California*. Ph.D. Dissertation, California Institute of Technology, Pasadena, California.

References for Chapter 3

Empirical tests of within- and across-species energetics in a diverse plant community

Abramoff, M. D., P. J. Magelhaes, and S. J. Ram. 2004. Image processing with ImageJ. Biophotonics International 11:36-42.

Babu, G. J., and E. D. Feigelson. 2006. Astrostatistics: Goodness-of-fit and all that!, in Astronomical Data Analysis Software and Systems XV (eds. C. Gabriel et al.), ASP Conf. #351, 127.

Bazzaz. F. A. 1975. Plant Species Diversity in Old-Field Successional Ecosystems in Southern Illinois. Ecology 56: 485-488.

Blonder B., V. Buzzard, I. Simova, L. Sloat, B. Boyle, R. Lipson et al. 2012. The leaf-area shrinkage effect can bias paleoclimate and ecology research. American Journal of Botany 99: 1756-1763.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. Ecology 85: 1771-1789.

Clauset, A., D. J. Schwab, and S. Redner. 2009. How many species have mass M?. The American Naturalist 173: 256-263.

Cornelissen, J. H. C., S. Lavorel, E. Garnier, S. Diaz, N. Buchmann, D. E. Gurvich, et al. 2003. A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Australian Journal of Botany 5: 335-380.

Damuth, J. Population density and body size in mammals. 1981. Nature 290: 699-700.

Damuth, J. 1987. Interspecific allometry of population density in mammals and other animals: the independence of body mass and population energy-use. Biological Journal of the Linnean Society 31:193–246.

de Aguiar, M. A. M., M. Baranger, E. M. Baptestini, L. Kaufman, and Y. Bar-Yam. 2009. Global patterns of species and diversity. Nature 460: 384–387.

Deng J.-M., T. Li, G.-X. Wang, J. Liu, Z.-L. Yu, et al. 2008. Trade-Offs between the Metabolic Rate and Population Density of Plants. PLoS one 3: e1799.

Dornelas, M., A. C. Moonen, A. E. Magurran, and P. Bàrberi. 2009. Species abundance distributions reveal environmental heterogeneity in modified landscapes. Journal of Applied Ecology 46: 666-672.

Dunne, J. A., J. Harte, and K. J. Taylor. 2003. Subalpine meadow flowering phenology responses to climate change: integrating experimental and gradient methods. Ecological Monographs 73: 69-86.

Glazier, D. S. 2006. The 3/4-power law is not universal: evolution of isometric, ontogenetic metabolic scaling in pelagic animals. BioScience 56: 325-332.

Harte, J., T. Zillio, E. Conlisk, and A. Smith. 2008. Maximum entropy and the state variable approach to macroecology. Ecology 89: 2700–2711.

Harte, J., A. Smith, and D. Storch. 2009. Biodiversity scales from plots to biomes with a universal species-area curve. Ecology Letters 12: 789–797.

Harte, J. 2011. Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics. Oxford Univ. Press, Oxford UK.

Harte, J., J. Kitzes, E. A. Newman, and A. J. Rominger. 2013. Taxon Categories and the Universal Species-Area Relationship (A comment on Šizling et al., "Beyond Geometry and Biology: The Problem of Universality and the Species-Area Relationship). The American Naturalist 181: 282-287.

Harte, J., and E. A. Newman. 2014. Maximum information entropy: a foundation for ecological theory. Trends in ecology & evolution. 29 (7): 384–389.

Hill, J. K., and K. C. Hamer. 1998. Using species abundance models as indicators of habitat disturbance in tropical forests. Journal of Applied Ecology 35: 458-460.

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

Jaynes, E. T. 1979. Where do we stand on maximum entropy? Pages 15–118 in R. D. Levine and M. Tribus, eds. The maximum entropy formalism. MIT Press, Cambridge, MA.

Jaynes, E. T. 1982. On the rationale of maximum entropy methods. Proceedings of the IEEE 70: 939–952.

Jaynes, E. T. 2003. Probability: the logic of science. Cambridge University Press, Cambridge.

Kempton, R. A. and L. R. Taylor. 1974. Log-series and log-normal parameters as diversity discriminants for the Lepidoptera. The Journal of Animal Ecology 381-399.

Kitzes, J., C. Lewis, and M. Wilber. 2014. macroeco v0.2. http://github.com/jkitzes/macroeco. Accessed online April 2013.

Makarieva, A. M., V. G. Gorshkov, B.-L. Li, S. L. Chown, P. B. Reich, and V. M. Gavrilov. 2008. Mean mass-specific metabolic rates are strikingly similar across life's major domains:

evidence for life's metabolic optimum. Proceedings of the National Academy of Sciences of the USA 105: 16994–16999.

McGill, B. J. 2010. Towards a unification of unified theories of biodiversity. Ecology Letters 13:627–642.

Mori, S., K. Yamaji, A. Ishida, S. G. Prokushkin, O. V. Masyagina, A. Hagihara, et al. 2010. Mixed-power scaling of whole-plant respiration from seedlings to giant trees. Proceedings of the National Academy of Sciences 107: 1447-1451.

Muller-Landau**,** H. C., R. S. Condit, J. Chave, S. C. Thomas, S. A. Bohlman, S. Bunyavejchewin, et al. 2006. Testing metabolic ecology theory for allometric scaling of tree size, growth, and mortality in tropical forests. Ecology Letters 9: 575-588.

Niklas, K. J., and B. J. Enquist. 2001. Invariant scaling relationships for interspecific plant biomass production rates and body size. Proceedings of the National Academy of Sciences USA 98: 2922-2927.

Reich, P. B., M. G. Tjoelker, J. L. Machado, and J. Oleksyn. 2006. Universal scaling of respiratory metabolism, size and nitrogen in plants. Nature 439: 457-461.

Rüger N., and R. Condit. 2012. Testing metabolic theory with models of tree growth that include light competition. Functional Ecology 26: 759–765.

Ulrich, W., M. Ollik, and K. I. Ugland. 2010. A meta-analysis of species–abundance distributions. Oikos 119:1149–1155.

van Rossum, G. and F. L. Drake (eds), Python Reference Manual, PythonLabs, Virginia, USA, 2001. Available at http://www.python.org

White, E. P., S. K. Ernest, A. J. Kerkhoff, and B. J. Enquist. 2007. Relationships between body size and abundance in Ecology. Trends in Ecology and Evolution 22: 323-30.

White, E. P., K. M. Thibault, and X. Xiao. 2012. Characterizing species abundance distributions across taxa and ecosystems using a simple maximum entropy model. Ecology 93: 1772-1778.

Xiao, X., D. J. McGlinn, and E. P. White. 2013. A strong test of the Maximum Entropy Theory of Ecology. arXiv preprint arXiv:1308.0731

Appendix 1

Supplementary Material to Chapter 1

Appendix A: Tree coring data and metadata

Twenty cores were sampled from 17 individual trees, under National Park Service permitting.

Park-assigned Study or Activity #: PORE-00572 and

Park-assigned Permit #: PORE-2012-SCI-0014

The accession number for all of the specimens is PORE-00866, with catalog number range PORE 18080-PORE 18096.

Tree core specimens are property of the National Park Service, and are housed under a long-term loan agreement with the Laboratory of Tree-Ring Research (LTRR) at the University of Arizona.

Terms

tree_coring_code – Unique code given to each tree cored at Point Reyes

row – The row of the plot in which the tree was found

column – The column of the plot in which the tree was found

dbh_cm – The diameter at breast height (dbh) in cm of the tree

date – The date which the tree was cored in (MM/DD/YYYY) format

height_to_core_cm – the height from the ground at which the core was taken (in cm)

core_type – either a half-core or a through-core. Half core is sampled to halfway through the diameter of the tree at breast height; through-core is a full core sampled all the way through the tree's diameter at breast height

corer – initials of the person who did the coring

notes – notes describing particular aspects of the tree or coring procedure

rings – number of tree rings counted in tree-core sample

raw_date_sprout – raw year of germination of the tree based on number of rings counted

estimated_date_sprout – estimated year of germination of the tree based on (number of rings counted + estimated missing rings from years of growth not apparent at height of coring). This measure accounts for first growth ring only growing to ~35 cm height.

A1. Site Information: Bayview Trailhead

Location: Point Reyes National Seashore, California
Dates: Site sampled from April 15th to April 16th 2012
Samplers: LS, BC, JH, EAN, MQW

Notes: This was the second site that we sampled at the Bayview Trailhead. This site represents the more recently disturbed site used to test METE predictions. At this site, we made a 16m X 16m grid and gridded every 1m for a total of 256 cells. Within each cell we counted the abundance of every plant. For plants such as *Rubus ursinus* (California blackberry, field code RUBURS) that were vines and occurred everywhere, we recorded their presence only (rather than abundance) within the cell. We measured the dbh (cm) of all *Pinus muricata* (Bishop pine, field code PINMUR) within the grid. PINMUR trees were cored near (but not within) the plot.

Table A1. Records for Bayview site tree core data.

TREE_CORING_CODE	DBH_CM	DATE	HEIGHT_TO_CORE_CM	CORE_TYPE	CORER	NOTES	RINGS	RAW_DATE_SPROUT	ESTIMATED_DATE_SPROUT
PTRY001 (PORE 18080)	N/A	4/15/12	32	through-core	MQW	none	11	2001	(2001)
PTRY002 (PORE 18081, "A")	N/A	4/15/12	28	half-core	BC	Sample came from the same tree as PTRY003	12	2000	(1998)
PTRY003 (PORE 18081, "B")	N/A	4/15/12	131	half-core	EAN	Sample came from the same tree as PTRY002	13	1999	(1998)
PTRY004 (PORE 18082)	4.4	4/15/12	134	through-core	EAN	Thin tree. Tree was cored at breast height. Core broke at far end.	7	2005	(2003)
PTRY005 (PORE 18083, "A")	20	4/15/12	120	half-core	LS	Approximately 150-200m from BayviewTH site. Same tree as PTRY006	15	1997	(1996)
PTRY006 (PORE 18083, "B")	22	4/15/12	32	half-core	LS	Sample came from the same tree as PTRY005	16	1996	(1996)

A2. Site Information: Mt. Vision

Location: Mount Vision Road, Point Reyes National Seashore, California

GPS: 10 S 0509348, 4217230 ±1.8m for center of the site. 27 degrees from center point of plot

Elevation: 698 ft

Clinometer reading: -32 degrees from the horizontal

Dates: Site sampled from April 19th to April 20th 2012

Samplers: EAN, MQW

Notes: This was the third site that we censused during the macroecology project at Point Reyes. This site was located on Mt. Vision Road and was an undisturbed stand of mature Bishop Pine (PINMUR). At this site, we made a 16m x 16m grid and gridded every 1m for a total of 256 cells. Consistent with the methods applied to the first field plot, within each cell we counted the abundance of every plant. For plants such as *Rubus ursinus* (California blackberry, field code RUBURS) that were vines and occurred everywhere, we recorded their presence only (rather than abundance) within the cell. We measured the dbh (cm) of all *Pinus muricata* (Bishop pine, field code PINMUR) within the grid. PINMUR trees were cored near (but not within) the plot.

There are 14 mature *Pinus muricata* (PINMUR) in the site. This file only contains 13 unique cores (PTRY020 is a core from the same tree as PTRY009). The PINMUR located in cell I3 (dbh: 28cm) was not cored.

Table A2. Records for Mount Vision site tree core data.

TREE_CORING_CODE	ROW	COLUMN	DBH_CM	DATE (M/D/Y)	HEIGHT_TO_CORE_CM	CORE_TYPE	CORER
PTRY007 (PORE 18084)	3	C	50.1	4/20/12	62.5	half-core	MQW
PTRY008 (PORE 18085)	10	C	28.5	4/20/12	55	half-core	MQW
PTRY009 (PORE 18086, "A")	11	C	44.5	4/20/12	69.5	half-core	MQW
PTRY010 (PORE 18087)	6	H	27.7	4/21/12	73	half-core	MQW
PTRY011 (PORE 18088)	13	G	34.6	4/20/12	66	half-core	MQW
PTRY012 (PORE 18089)	2	L	41.3	4/20/12	73	half-core	MQW
PTRY013 (PORE 18090)	5	L	48.7	4/20/12	62.5	half-core	MQW
PTRY014 (PORE 18091)	13	J	36.5	4/20/12	51	half-core	MQW
PTRY015 (PORE 18092)	2	P	43.3	4/20/12	80	half-core	EAN
PTRY016 (PORE 18093)	2	P	45.5	4/21/12	82	half-core	MQW
PTRY017 (PORE 18094)	9	O	34.1	4/21/12	49.5	half-core	MQW
PTRY018 (PORE 18095)	14	N	45.5	4/20/12	65	half-core	MQW
PTRY019 (PORE 18096)	16	M	40	4/20/12	61	half-core	MQW
PTRY020 (PORE 18086, "B")	11	C	44.5	4/20/12	46.5	half-core	MQW

Table A2. *continued*. Records for Mount Vision site tree core data (continued).

TREE_CORING_CODE	NOTES	RINGS	RAW_DATE_SPROUT	ESTIMATED_DATE_SPROUT
PTRY007 (PORE 18084)	Tree slanting downhill. Cored on downhill side	35	1977	(1976)
PTRY008 (PORE 18085)	Tree slanting uphill	23	1989	(1988)
PTRY009 (PORE 18086, "A")	Cored from downhill slope	42	1970	(1969)
PTRY010 (PORE 18087)	Tree branches to a Y at ~35-40cm from the ground. One branch of the tree is broken and the other is a full-grown PINMUR. Cored the full-grown PINMUR	30	1982	(1981)
PTRY011 (PORE 18088)	Cored on uphill side	41	1971	(1970)
PTRY012 (PORE 18089)	Cored on downhill side. Tree is sloping slightly left	37	1975	(1974)
PTRY013 (PORE 18090)	Tree asymmetrical. Cored on downhill side	36	1976	(1975)
PTRY014 (PORE 18091)	Tree has lots of bends in its trunk	33	1979	(1978)
PTRY015 (PORE 18092)	Merged with another tree. Cored on opposite side of neighboring tree.	33	1979	(1978)
PTRY016 (PORE 18093)	Tree splits into two distinct trunks at ~40cm from the ground. Cored at 82 cm from the ground to get a trunk that is distinct from PTRY015	38	1974	(1973)
PTRY017 (PORE 18094)	Tree has lots of branches near the base that affected where it could be cored. Pith was hit exactly.	26	1986	(1985)
PTRY018 (PORE 18095)	Tree is misshapen 3m up the trunk	36	1976	(1975)
PTRY019 (PORE 18096)	Tree trunk is pretty straight. Cored on the uphill side. Core has some rot or insect damage that makes rings unclear.	~29	~1983	(1982)
PTRY020 (PORE 18086, "B")	Same tree and PTRY009. Tree is possibly diseased to two samples were taken. This sample was taken 90 degrees from the previous sample towards the nearest edge of the plot. A few rings close to the bark may have been lost. Inner rings intact.	33	1979	(1978)

Appendix 2

Supplementary Material to Chapter 3

Appendix A. Measurements and calculations of photosynthetic area

Identification of individuals and calculation of photosynthetic area of each individual plant was done on a species-by-species basis. For species that grow by rhizomes or from other underground attachments, such as *Vicia americana*, *Fragaria vesca*, and *Boechera drumondii*, we uncovered roots to see which aboveground plant structures were part of the same individual, or assumed that leaf structures emerging from the ground in a line were attached underground and therefore part of the same individual. For other plants, we distinguished annuals from perennials. The annuals *Androcase septentrionalis* and *Taraxacum officianale* were counted as individuals by rosette unless there were shared root structures. For perennials, aboveground structures were assumed to be part of the same plant if they were growing in patches or clumps. Distances between structures considered part of the same "clump" varied by species, such that rosettes of the small *Ipomopsis aggregata* were considered separate individuals if their centers were more than 3 cm apart and there was no sign of attachment between roots, whereas stems of the much larger *Lupinus argentus* were counted as individuals if they were 20 or more cm apart. Individuals were assessed during counting, and validated during collection of aboveground biomass.

In all cases of plants with both leaves and stems, the leaves were counted and scanned separately from the stems. Both types of structures were scanned separately and included in the final accounting for photosynthetic area (because of significant contribution of stems; Pfanz and Aschan 2001, Valladares et al. 2003, Flexas et al. 2012), defined as leaves + green stems (see, for example Boutin and Keddy 1993), or simply all aboveground biomass for these herbaceous plants. Leaves collected from rosettes were counted, labeled, and scanned separately from leaves growing on stems. For example, we counted rosettes and stems for *Thlaspi montanum,* which grows in large mats containing thousands of individual leaves. Rosettes were counted individually for a mat considered to be an individual plant, and stems were counted separately. At least 7-10 rosettes were plucked and each individual leaf was counted, resulting in an estimation of leaves/rosette. Similarly, leaves/stem was calculated from at least 7-10 stems. Calculated averages of individual leaf area obtained from scans of collected leaves were then applied to the total number of rosettes x leaves/rosette. Total photosynthetic area was then calculated as (area per leaf) x (rosettes) x (leaves/rosette) + (area per stem leaf) x (total stems) x (leaves/stem) + (area per stem) x (total stems). We applied similar methods to other plant species based on their growth form using similar reasoning.

A note on taxonomy: lupine individuals were grouped into the category "*Lupinus sp*." for analysis, as the taxonomy of that group is in flux (although *Lupinus argentus* is likely the dominant species).

Literature cited

Boutin, C., and P. A. Keddy. 1993. A functional classification of wetland plants. Journal of Vegetation Science 4:591-600.

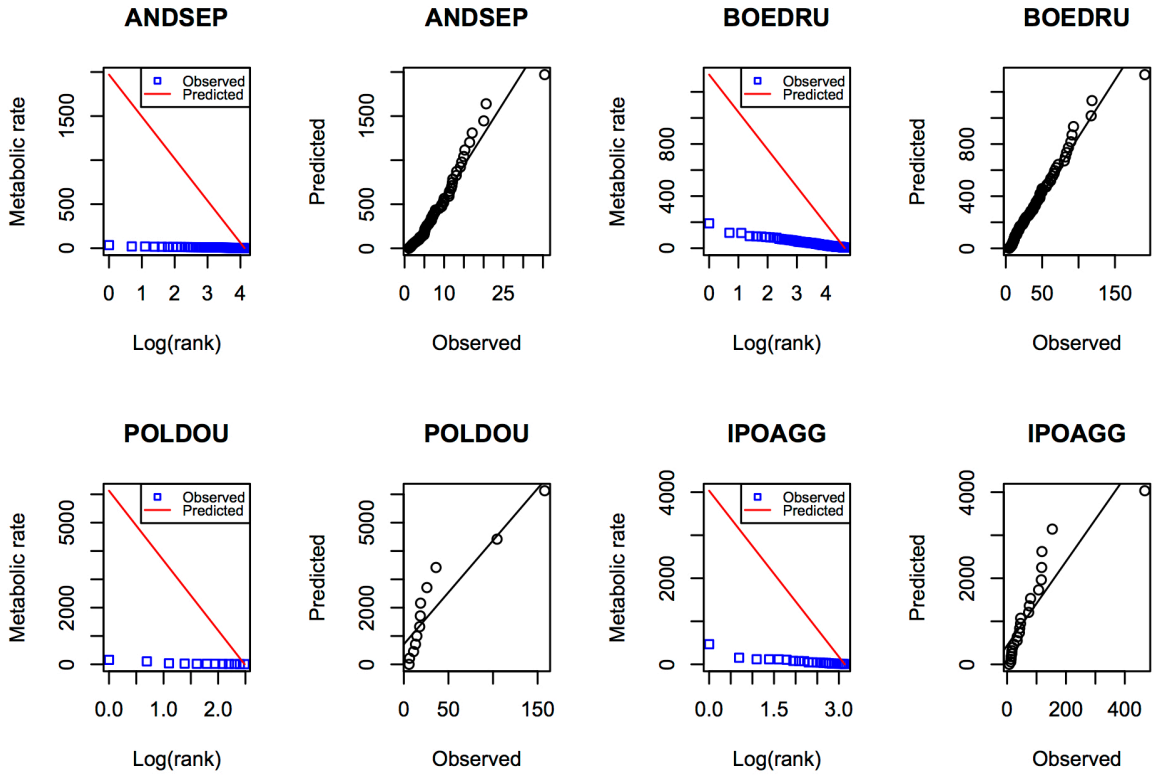
Flexas, J., F. Loreto, and H. Medrano, eds. 2012. Terrestrial photosynthesis in a changing environment: a molecular, physiological, and ecological approach. Cambridge University Press.

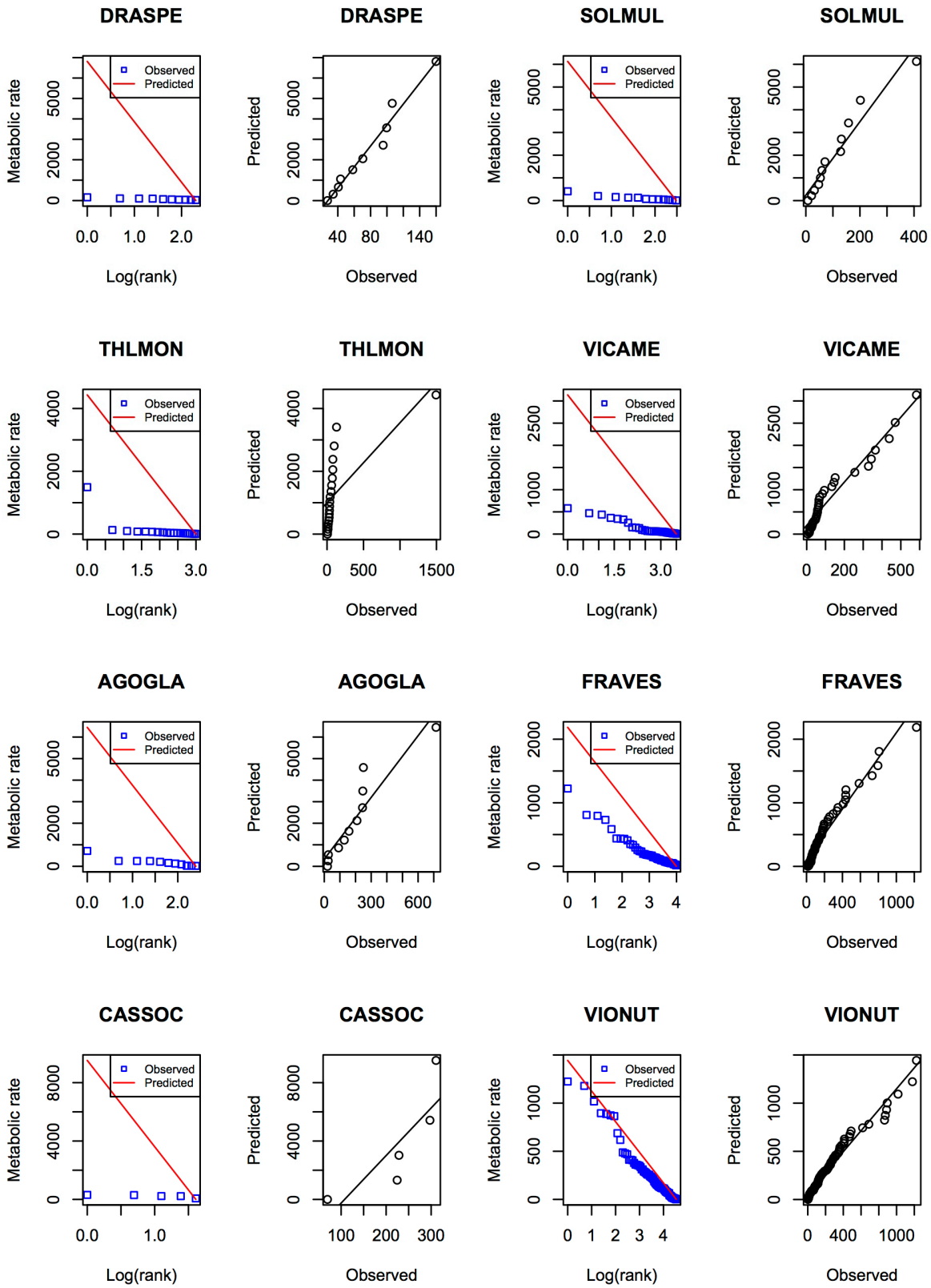
Pfanz, H., and G. Aschan. 2001. The existence of bark and stem photosynthesis in woody plants and its significance for the overall carbon gain. An eco-physiological and ecological approach. Progress in Botany 62: 477–510.

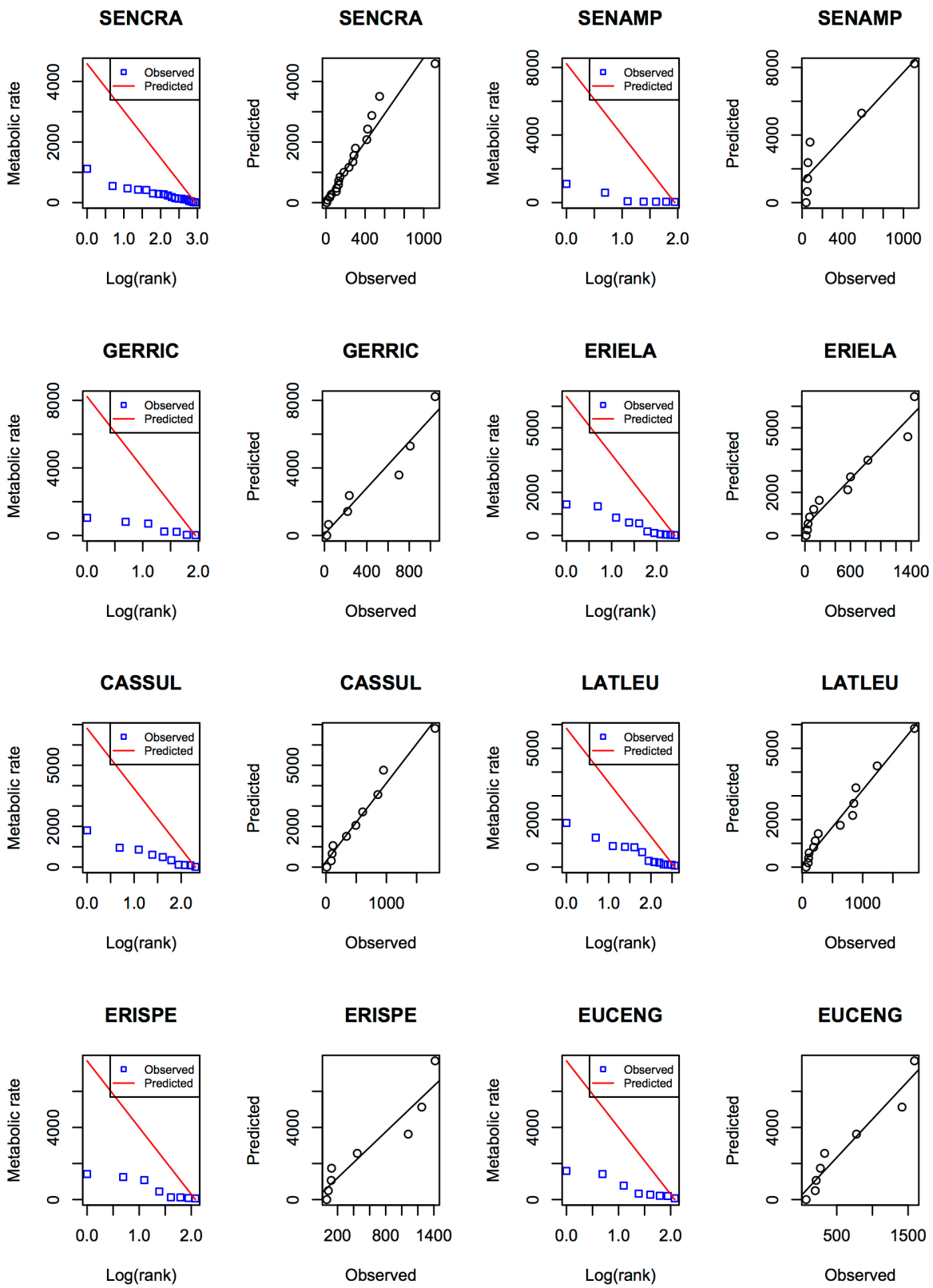
Valladares, F., L. G. Hernández, I. Dobarro, C. García-Pérez, C., R. Sanz, and F.I. Pugnaire. 2003. The ratio of leaf to total photosynthetic area influences shade survival and plastic response to light of green-stemmed leguminous shrub seedlings. Annals of Botany, 91: 577-584.

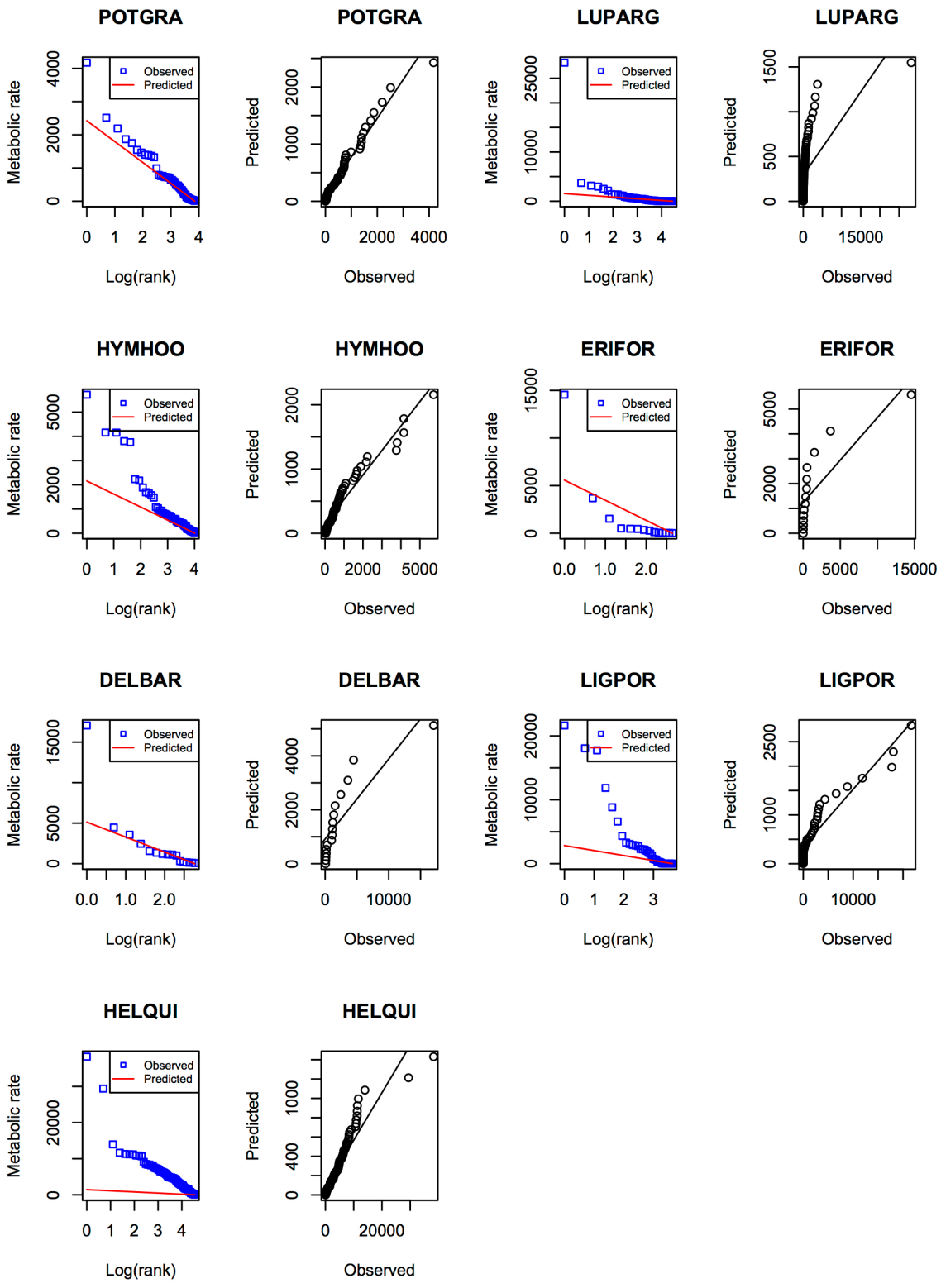
Appendix B. SED predictions and observed values for all species with 5 or more individuals, including comparisons of predicted vs. observed values

Figure B.1 SED predictions and observed values for all species with 5 or more individuals are arranged from smallest to largest by average metabolic rate of an individual of a species. Within each graph, individuals are ranked from largest to smallest individual on the x-axis. Predicted vs. observed values on a linear scale are included for each species (open circles), with a linear regression line (solid black line).




DRASPE
Metabolic rate
Log(rank)
Observed
Predicted
DRASPE
Predicted
Observed
SOLMUL
Metabolic rate
Log(rank)
Observed
Predicted
SOLMUL
Predicted
Observed
THLMON
Metabolic rate
Log(rank)
Observed
Predicted
THLMON
Predicted
Observed
VICAME
Metabolic rate
Log(rank)
Observed
Predicted
VICAME
Predicted
Observed
AGOGLA
Metabolic rate
Log(rank)
Observed
Predicted
AGOGLA
Predicted
Observed
FRAVES
Metabolic rate
Log(rank)
Observed
Predicted
FRAVES
Predicted
Observed
CASSOC
Metabolic rate
Log(rank)
Observed
Predicted
CASSOC
Predicted
Observed
VIONUT
Metabolic rate
Log(rank)
Observed
Predicted
VIONUT
Predicted
Observed


SENCRA
Metabolic rate
Log(rank)
Observed
Predicted
SENCRA
Predicted
Observed
SENAMP
Metabolic rate
Log(rank)
SENAMP
Predicted
Observed
GERRIC
Metabolic rate
Log(rank)
GERRIC
Predicted
Observed
ERIELA
Metabolic rate
Log(rank)
ERIELA
Predicted
Observed
CASSUL
Metabolic rate
Log(rank)
CASSUL
Predicted
Observed
LATLEU
Metabolic rate
Log(rank)
LATLEU
Predicted
Observed
ERISPE
Metabolic rate
Log(rank)
ERISPE
Predicted
Observed
EUCENG
Metabolic rate
Log(rank)
EUCENG
Predicted
Observed


POTGRA
Observed
Predicted
Metabolic rate
Log(rank)
POTGRA
Predicted
Observed
LUPARG
Observed
Predicted
Metabolic rate
Log(rank)
LUPARG
Predicted
Observed
HYMHOO
Observed
Predicted
Metabolic rate
Log(rank)
HYMHOO
Predicted
Observed
ERIFOR
Observed
Predicted
Metabolic rate
Log(rank)
ERIFOR
Predicted
Observed
DELBAR
Observed
Predicted
Metabolic rate
Log(rank)
DELBAR
Predicted
Observed
LIGPOR
Observed
Predicted
Metabolic rate
Log(rank)
LIGPOR
Predicted
Observed
HELQUI
Observed
Predicted
Metabolic rate
Log(rank)
HELQUI
Predicted
Observed

Appendix C. Table. Species included in this study, ranked by average metabolic rate per species (a normalized measure of photosynthetic area), from largest to smallest

Slopes, intercepts, and R^2 values of a linear regression of predicted SED values versus observed data are shown. Metrics are not calculated (NC) for species with fewer than 5 individuals.

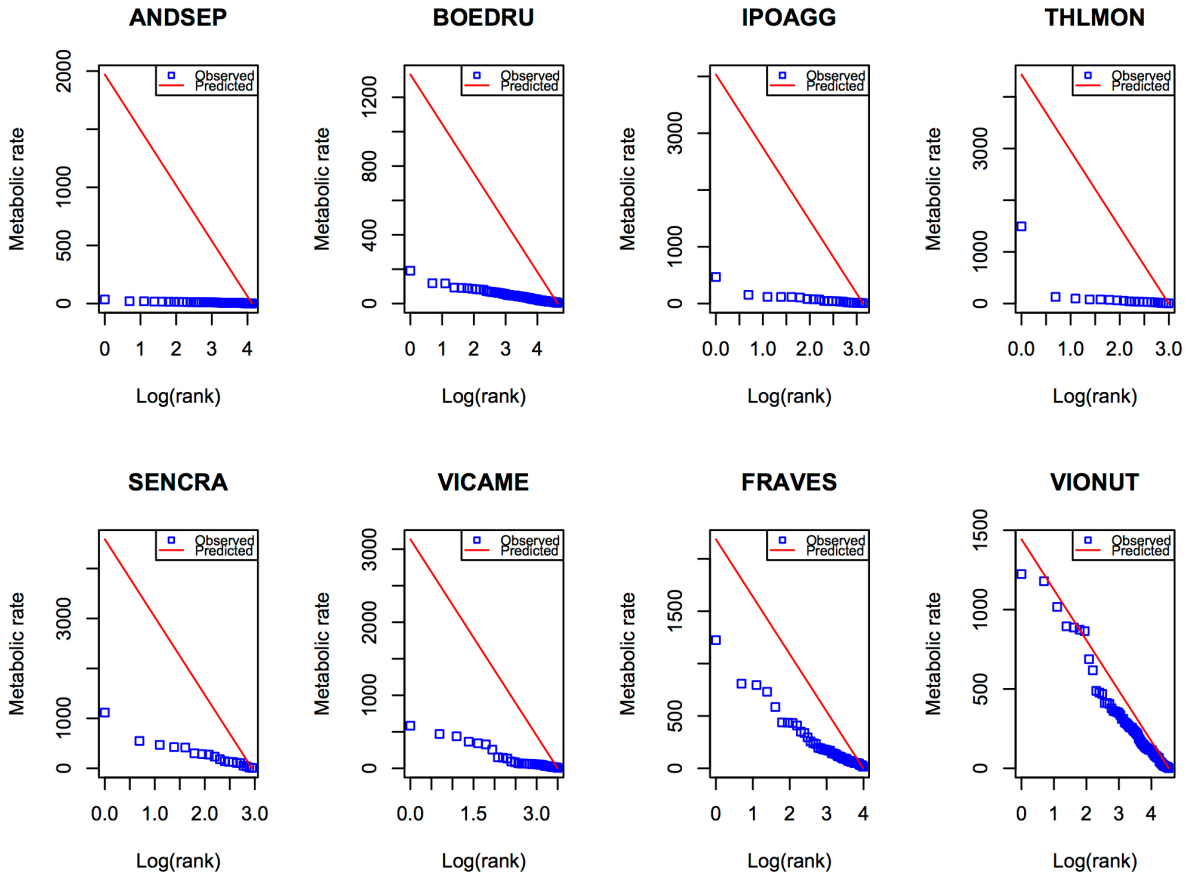
Table C.1

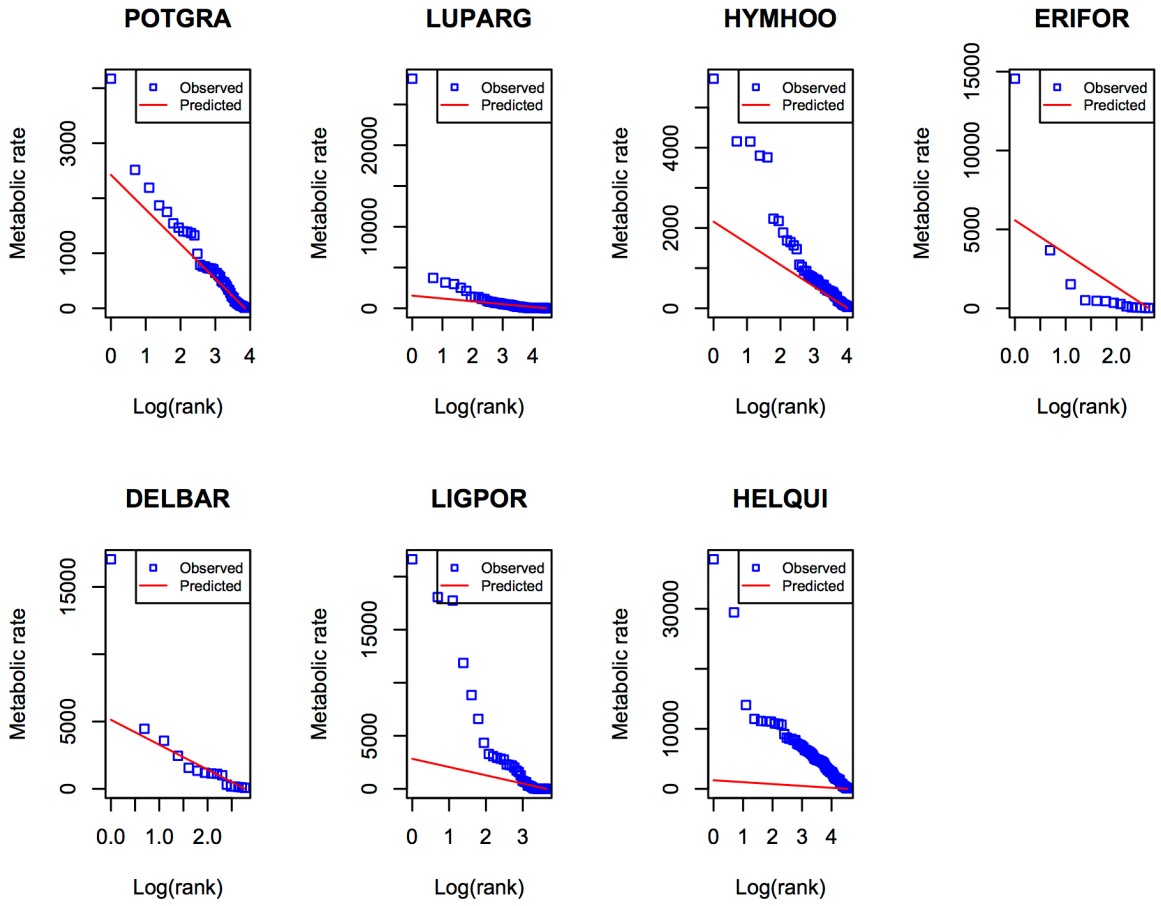
6-Letter code	Rank by ave. metabolic rate of species	Intercept	Slope	R^2	Rank by abundance	Abundance	Ave. metabolic rate
ANDSEP	31	-159.043	77.446	0.9687	5	62	8.193
(TAROFF)	30	NC	NC	NC	31	1	9.537
BOEDRU	29	-28.243	9.287	0.9865	1	103	33.97
POLDOU	28	744.391	46.067	0.8768	18	12	35.843
IPOAGG	27	419.74	11.659	0.8266	11	23	72.744
DRASPE	26	-1844.204	63.888	0.9582	21	10	73.597
SOLMUL	25	181.127	20.142	0.9628	17	12	109.939
THLMON	24	1076.99	3.199	0.5263	12	20	118.013
VICAME	23	181.74	5.479	0.9402	10	33	128.913
AGOGLA	22	289.64	12.05	0.8932	20	11	192.236
FRAVES	21	88.762	2.2	0.9653	7	54	207.54
CASSOC	20	-4142.9	42.72	0.5108	27	5	226.134
VIONUT	19	54.631	1.158	0.9702	3	93	227.33
SENCRA	18	78.885	5.525	0.9582	13	19	262.59
SENAMP	17	1582.521	8.633	0.8775	25	7	282.468
(VIGMUL)	16	NC	NC	NC	28	2	376.539
GERRIC	15	168.053	8.79	0.8787	26	7	438.361
ERIELA	14	470.67	4.476	0.9189	19	11	477.113
CASSUL	13	254.251	4.845	0.9819	22	10	537.313
LATLEU	12	81.52	3.781	0.9639	16	13	564.596
ERISPE	11	470.842	5.353	0.8347	24	8	573.524
EUCENG	10	257.171	5.404	0.9012	23	8	607.619
POTGRA	9	79.31	0.749	0.9788	8	47	728.91
LUPARG	8	295.3	0.071	0.4064	4	85	736.251

HYMHOO	7	135.787	0.418	0.9456	6	55	956.284
ERIFOR	6	1373	0.436	0.7053	15	14	1578.719
DELBAR	5	979.284	0.37	0.7676	14	16	2243.05
LIGPOR	4	352.8	0.132	0.8976	9	38	3171.588
HELQUI	3	65.659	0.053	0.8906	2	94	4680.185
(AQUCOE)	2	NC	NC	NC	30	2	11990.786
(OSMOCC)	1	NC	NC	NC	29	2	18712.189

Appendix D. SED predictions and observed values for the 15 most abundant species

Figure D.1 SED predictions and observed values for 15 most abundant species are arranged from smallest to largest by average metabolic rate. Individuals are ranked from largest to smallest individual by metabolic rate on the x-axis.




POTGRA
LUPARG
HYMHOO
ERIFOR
DELBAR
LIGPOR
HELQUI
Observed
Predicted
Metabolic rate
Log(rank)

Appendix E. Additional analysis of SEDs

Figure E.1. Slopes of predicted versus observed data for all SEDs of species with $n \geq 5$, organized by average photosynthetic area (equivalent to average metabolic rate) of an individual within a species. Each point represents one species, and species are in rank order from largest to smallest average photosynthetic area (equivalent to ranking by metabolic rate) of the individuals of that species. An exponential trendline is shown on the graph. A good fit of the predicted versus observed data points would be a slope of 1, however, we see variable, increasing slope here.

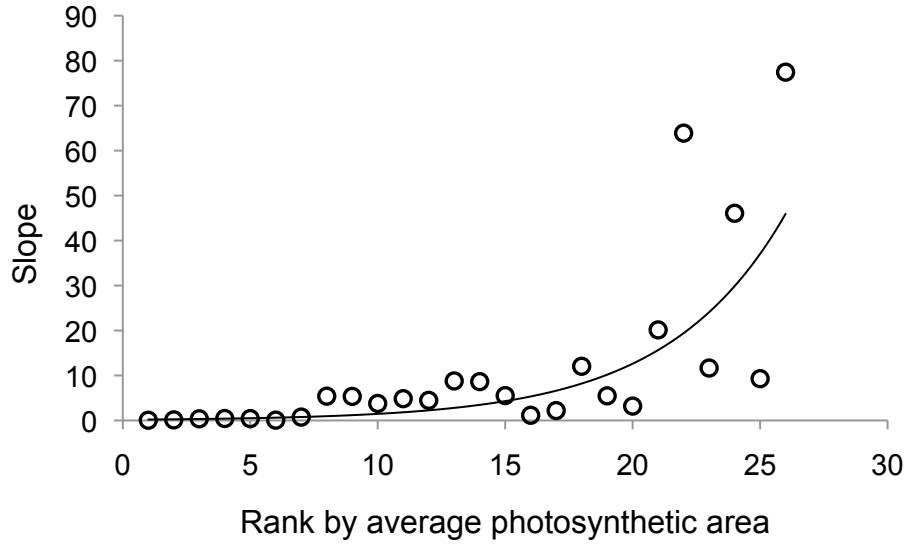
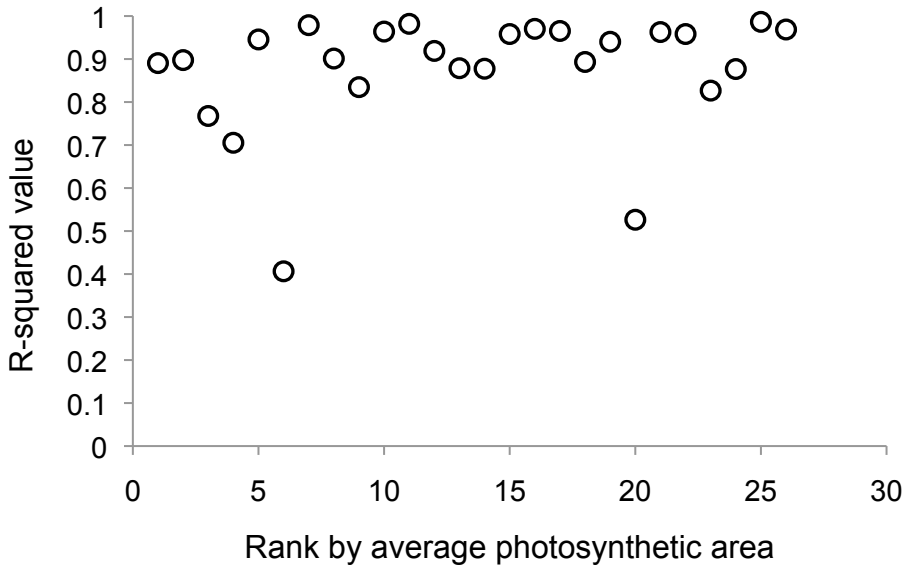


Figure E.2. R^2 values for predicted SED distributions versus observed data. Species are in rank order from largest to smallest average photosynthetic area (metabolic rate). Outlier points with low R^2 values are *Lupinus argentus* (rank 6 by average size) and *Thlaspi montanum*, (rank 20 by average metabolic rate), both of which have one very large outlying individual. High R^2 values naturally result from ranked data that falls along a straight line on a log-log graph of an exponential distribution.



One way to interpret these results is to say that the exponential model is a good model, however, the slopes are incorrect. We note that the intercepts are also incorrect for this model, which may indicate that one or more important constraints are missing from the model.

Appendix F. The functional form of the ASED (*Nu*) as derived from a discrete cumulative distribution function

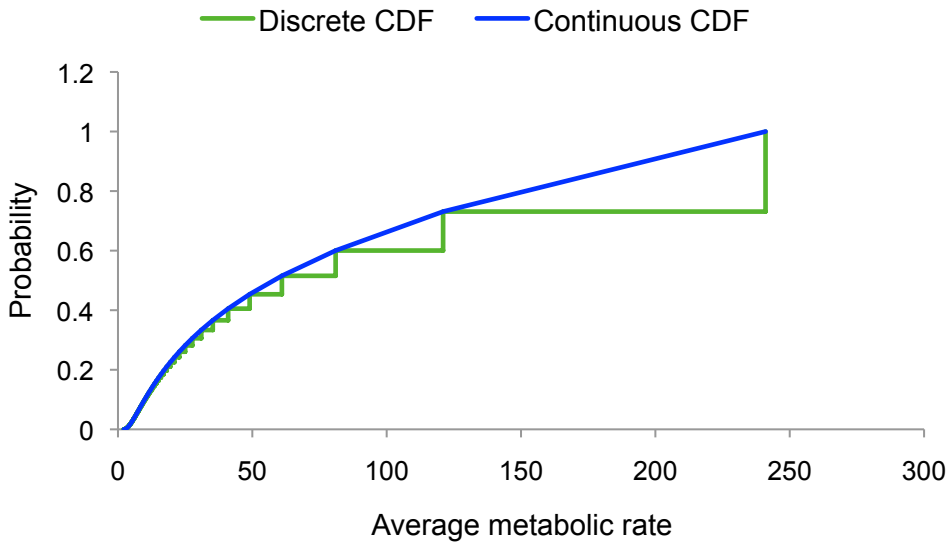
Because the SAD is a discrete distribution, the theoretical ASED distribution will reflect that discrete nature. However, as a function of a continuous variable, $\bar{\varepsilon}$, the empirical ASED may take on any value in the range of 1 to ε_{max}. The METE-predicted ASED is only defined for $\bar{\varepsilon}$ values given by $\bar{\varepsilon} = (1+(1/(n \cdot \lambda_2))$, where n (the number of individuals of a species) is in the range $[1,N]$ (and where $N = n_{max}$). Although we would ultimately like to model metabolic rate as a continuous distribution (conforming to our understanding of metabolic "energy" as a continuous variable in nature), there is no closed form rank-metabolic rate distribution for the ASED. The method suggested in Harte, 2011 (eq. 7.44, p. 155), allows us to calculate the ASED numerically and create a model that is close to a continuous distribution. We use this numerically-estimated continuous formulation of the ASED for the all analyses in this paper, and discuss the problems of doing so below.

Various artifacts arise from trying to predict a continuous variable, metabolic "energy" (or metabolic rates, defined in units of power), from a rank-metabolic rate distribution constructed from a discrete probability distribution. For example, when using a discrete ASED graphed on rank-metabolic rate and rank-log(metabolic rate) graphs, predicted values for multiple species with large average metabolic rates have a maximum predicted value, and values in the mid-range may be repeated. A continuous probability distribution would instead predict the ASED to be a strictly decreasing function (rather than a stepped function) of ranked average metabolic rates.

Graphically, the values $\bar{\varepsilon}_1$, $\bar{\varepsilon}_2$, ... $\bar{\varepsilon}_N$ (where $\bar{\varepsilon}_1$ is the smallest value and $\bar{\varepsilon}_N$ is the largest value) get progressively more spread out approaching $\bar{\varepsilon}_N$. With this in mind, we build a rank-metabolic rate distribution using the cumulative distribution function (CDF) through the following steps: (1) calculate an observed CDF for the observed energy values, and (2) use the cumulative probability values in the observed CDF and find the $\bar{\varepsilon}$ in the METE-predicted CDF with the same cumulative probability value. This is the predicted metabolic rate value. Therefore, constructing a CDF from this type of probability distribution causes large steps on the x-axis between each gain in probability at large $\bar{\varepsilon}$.

For example, as in Figure F.1, for three observed probability values of 0.81, 0.88, and 0.92, the continuous CDF (blue line) estimates a different $\bar{\varepsilon}$ for each probability value, while a discrete CDF (green line) has $\bar{\varepsilon}$ falling in the gap between two defined metabolic rate values and therefore must be assigned to either the lower or higher energy value. Therefore, different CDF values may end up with the same predicted $\bar{\varepsilon}$, leading to a step-like, artifactual pattern that may not resemble patterns of natural systems. This explains why it is a general feature of the predicted, discrete ASED that we see a flat line of predicted values for the first few ranks, *i.e.* the highest $\bar{\varepsilon}$ values, on rank abundance graphs, and may see repeated values elsewhere in the distribution.

Figure F.1. Discrete predictions of *Nu* distribution compared to idealized continuous curve. These example distributions were generated with $N = 200$, $S = 20$, and total energy $E = 5000$.



Literature cited

Harte, J. 2011. Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics. Oxford Univ. Press, Oxford UK.

Appendix G. Kolmogorov-Smirnov goodness-of-fit tests for ASED, IED, and all SEDs

Table G.1. The following goodness-of-fit tests represent how well theoretical distributions fit empirical data. Kolmogorov-Smirnov (K-S) average statistics for 1000 bootstrapped samples for the ASED, IED, and all SEDs with $n \geq 5$ individuals are presented here. Metabolic distributions that produced K-S bootstrapped p-values > 0.05 are marked with an asterisk. P-values less than 0.05 are considered not to support the null hypothesis that the predicted and observed distributions are the same.

Metric	Species	K-S statistic (bootstrapped average)	K-S bootstrapped p-value	Number observed (*n*)
ASED	all	0.5806	0.0001	31
IED	all	0.1653	0	877
SED	AGOGLA	0.8182	0.0007	11
	ANDSEP	0.9355	0	62
	BOEDRU	0.6893	0	103
	CASSOC*	0.8	0.0794	5
	CASSUL*	0.6	0.0524	10
	DELBAR*	0.1875	0.9523	16
	DRASPE	0.9	0.0002	10
	ERIELA*	0.5455	0.0747	11
	ERIFOR	0.5714	0.0188	14
	ERISPE*	0.625	0.087	8
	EUCENG	0.625	0.087	8
	FRAVES	0.4074	0.0002	54
	GERRIC*	0.7143	0.053	7
	HELQUI	0.766	0	94
	HYMHOO*	0.1818	0.3255	55
	IPOAGG	0.8261	0	23
	LATLEU*	0.4615	0.1265	13
	LIGPOR	0.3684	0.0109	38
	LUPARG	0.3059	0.0006	85
	NOCMON	0.85	0	20
	POLDOU	0.9167	0	12

POTGRA*	0.1277	0.8439	47
SENAMP*	0.7143	0.053	7
SENCRA	0.6316	0.0007	19
SOLMUL	0.8333	0.0002	12
VICAME	0.6061	0	33
VIONUT*	0.1828	0.0894	93

Appendix H. Additional information on the Average Species Energy Distribution, including tests of goodness-of-fit

Figure H.1. Predicted continuous probability distribution for the ASED (*Nu*) based on state variable values. Here, $\bar{\varepsilon}$ ("epsilon_bar") represents the average metabolic rate of an individual in a species, estimated by normalized photosynthetic area. The continuous distribution was estimated numerically, and normalized on the range $[\bar{\varepsilon}_{\min} = 1 + \frac{1}{N \cdot \lambda_2}, \bar{\varepsilon}_{\max} = 1 + \frac{1}{\lambda_2}]$.

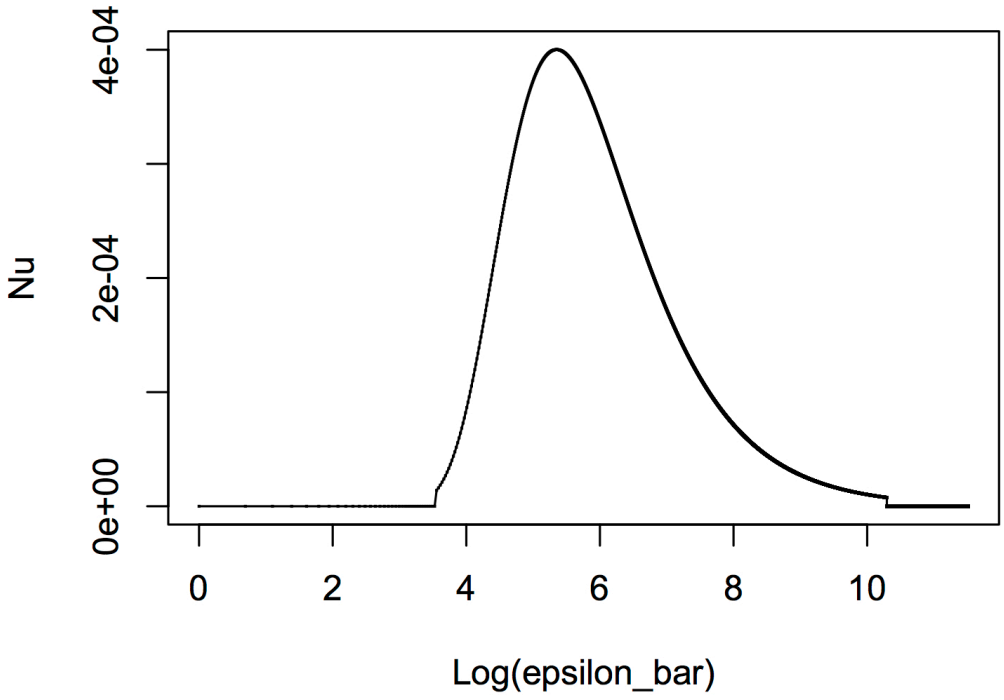
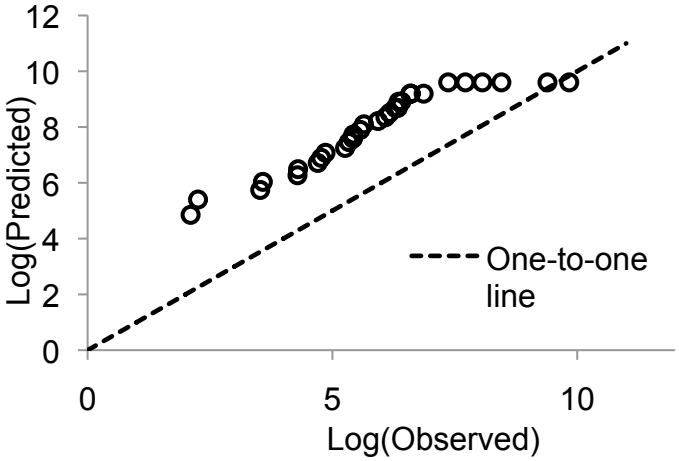


Figure H.2 ASED predictions and empirical data for a subalpine plant community where species are ranked by average energy: (a) discrete predicted versus observed values on a log-log graph; and (b) continuous predicted versus observed values on a log-log graph. Although we use the discrete values for all analyses in this paper, we note that use of the continuous ASED does not change the fit to the predicted versus observed values in a major way.

(a)



(b)

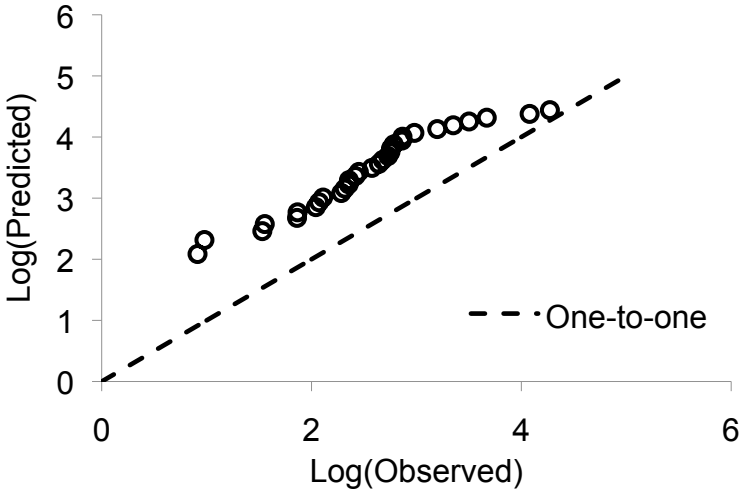
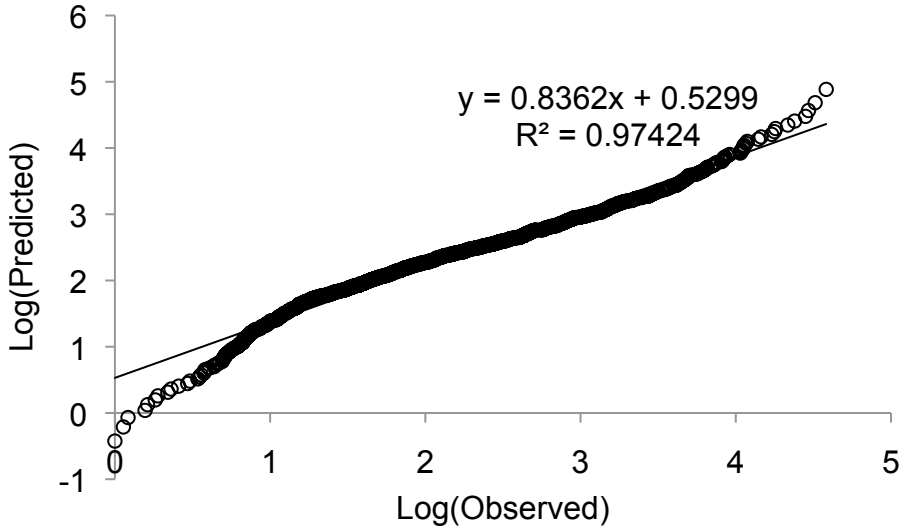


Figure H.3 Comparison of empirical data to METE Individual-level Energy Distribution predictions for a subalpine plant community. The figure shows predicted (discrete) versus observed values on a log-log graph. A high R^2 value is expected as a fit to rank-order data.



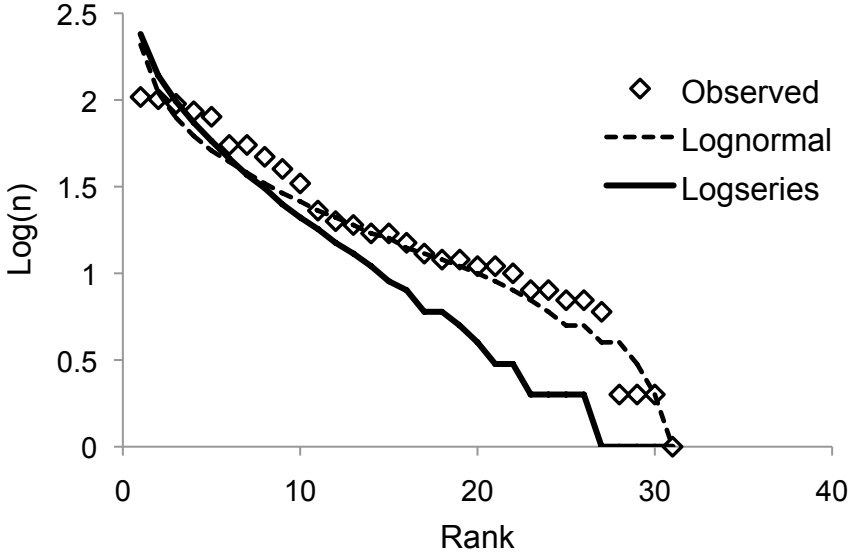
Appendix I. The Species-Abundance Distribution (SAD) of a subalpine plant community

METE predicts a Fisher log-series distribution for the SAD, but many ecosystems show a more log-normal pattern. Figure I.1 shows the comparison of empirical ranked abundances against the METE prediction and the lognormal distribution for the subalpine system. AIC weights for comparison to the distributions are:

Distribution	**Number of Parameters**	**AIC(weight)**
Lognormal	2	0.9526
Log-series	1	0.0474

This yields an evidence ratio of 20.09 in favor of the log-normal distribution.

Figure I.1. The Species-Abundance Distribution (SAD) of a subalpine plant community. Observed values on a rank-log(*n*) plot are closer to a log-normal distribution than they are to the METE-predicted log-series distribution.



Appendix J. Incorporating the empirical SAD into ASED and SED predictions

Both the SED and the ASED METE predictions explicitly incorporate the METE-predicted SAD in their derivations. While METE predicts a log-series SAD, the empirical SAD from a community does not necessarily follow this prediction. If the shape of the empirical SAD does not match the predicted SAD, one may expect that the ASED and SED METE predictions would not match empirical observations. To account for this possible source of error, we computationally derive the ASED and SED predictions using the empirical SAD rather than the METE-predicted SAD.

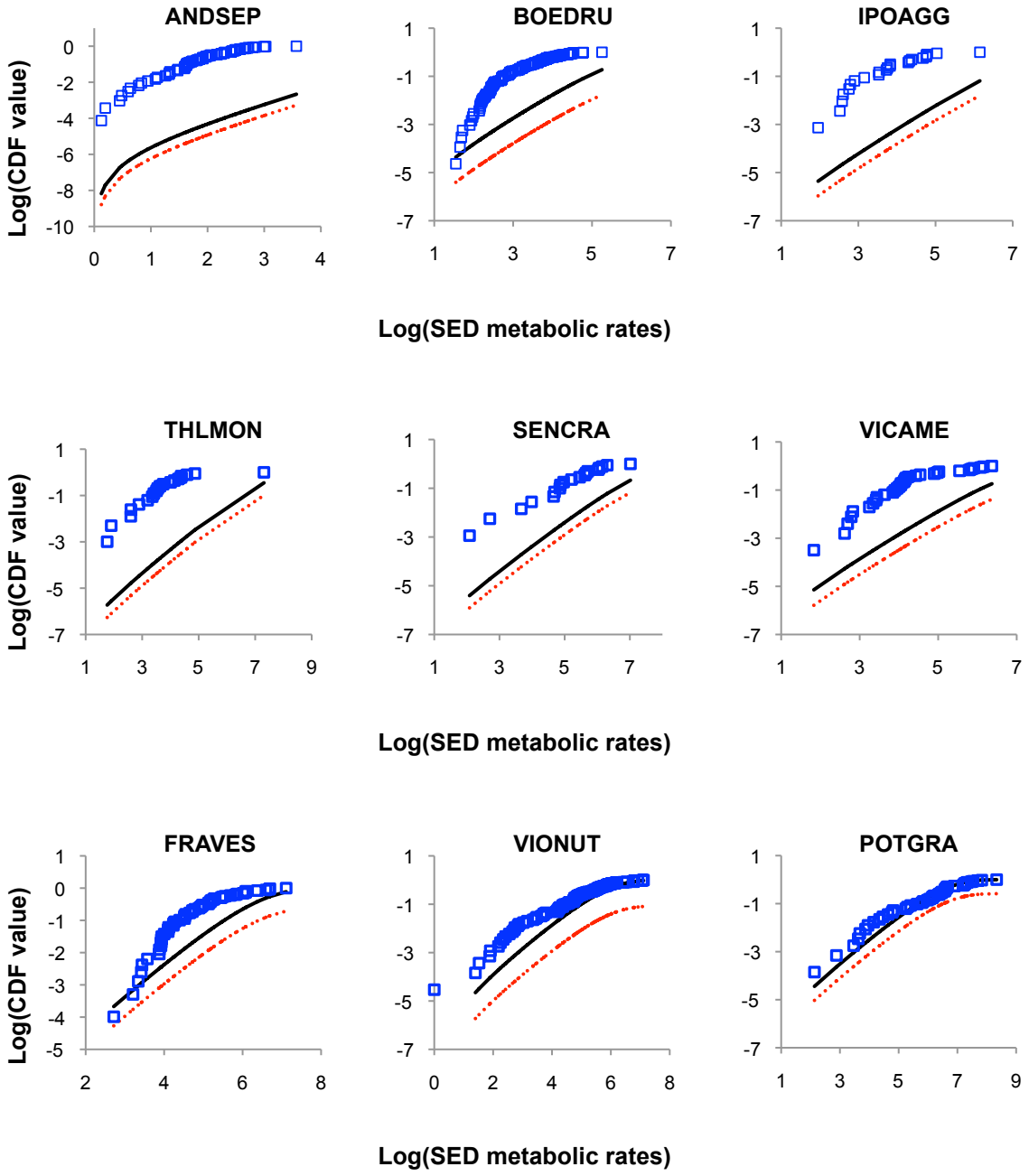
To obtain the empirically-based SED, we first computed the empirical SAD using Gaussian kernel density estimation (Jones *et al.* 2001), discretized, and renormalized the density estimate such that it has support from 1 to *N*. We then applied equation 7.7 from Harte (2011), restated here:

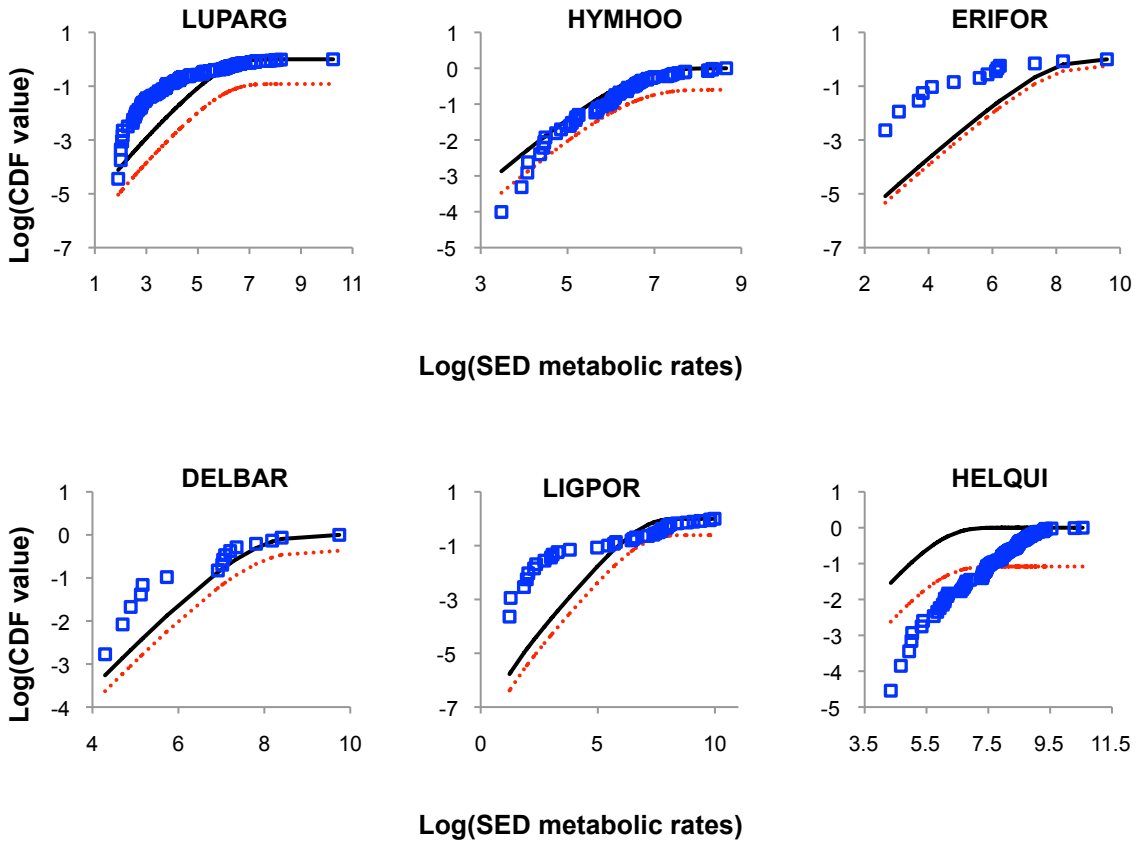
$$\Theta(\varepsilon \mid n) = \frac{R(n,\varepsilon)}{\Phi(n)} \tag{J.1}$$

and used our renormalized density kernel in place of *Phi*(*n*). The resulting SED distribution was not normalized, and normalizing the empirically-derived SED gives exactly the same distribution as the METE SED. This is obvious from the above equation because *Phi* (the SAD) influences the predicted SED of a species through a single value, *Phi*(*n*), where *n* is the abundance of the species. This value is the normalization factor of the SED and, because $R(n,\varepsilon)$ in Eq. J.1, above, is the same for the METE and empirically-derived SED, it has a single, unique value if the SED is properly normalized. Therefore, regardless of the shape of the SAD, all properly normalized SEDs will be identical. If the SED is not normalized, the resulting psuedo-distribution will have the same shape as the METE SED, but will be shifted by some factor (see Figure J.1). To change the shape of the SED (*i.e.* make it not an exponential distribution) one would have to alter the bivariate distribution $R(n,\varepsilon)$. Considering that $R(n,\varepsilon)$ is the foundation of METE and our goal in this paper is to test the energy predictions of METE, altering $R(n,\varepsilon)$ to incorporate the information from the empirical SAD would no longer be a test of METE.

Note: some figures in this section are printed in the main manuscript, and are reprinted here for comparison.

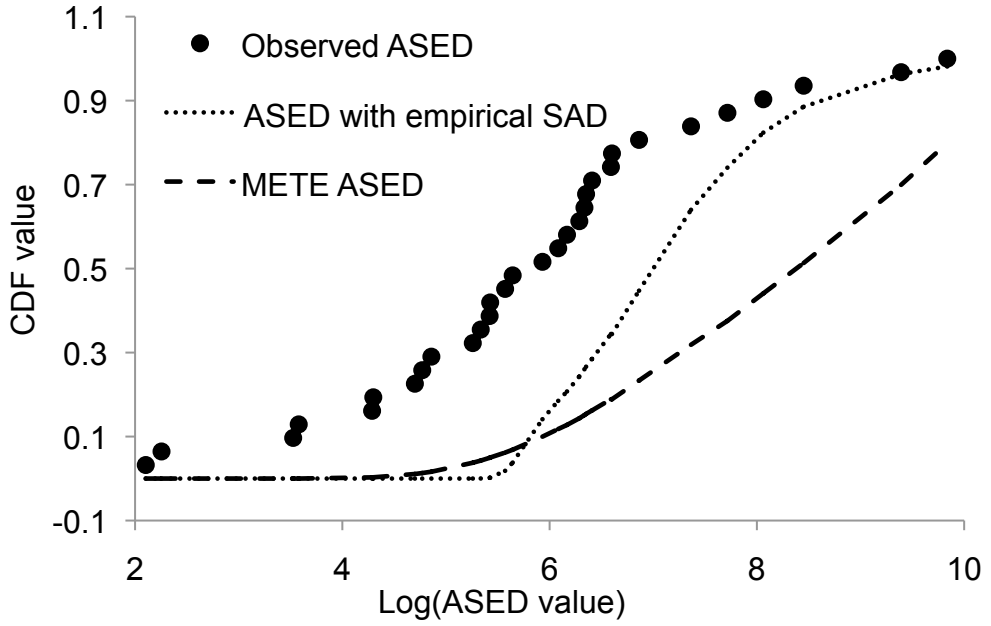
Figure J.1 SEDs by species, METE predictions, and empirically-adjusted METE predictions. Data are shown in blue squares, METE predictions are shown as a solid, black line, and METE predictions that incorporate the empirical SAD are shown as red, dotted lines.





To obtain the empirically-based ASED, we computed the empirical SAD using the method described above and then applied equation 7.44 from Harte (2011). The resulting empirically-derived ASED distribution had the same general shape as the METE-predicted ASED and is a better fit to to the observed ASED, especially in that it shows an inflection point, which is consistent with a unimodal *Nu* distribution (Figure J.2).

Figure J.2 Cumulative Density Function of the discrete ASED, showing a comparison of the observed ASED, the ASED including the empirical SAD, and the METE-predicted ASED.



These results strongly demonstrate that the non-METE empirical SAD we observed in this study is not the reason that the METE-predicted SED and ASED failed to match empirical data. Both the ASED and the SED predictions appear to be robust to the shape of the SAD, in that the functional form of the distributions do not change, but are only shifted away from empirical data by some (unknown) factor related to improper normalization. Although it may be too broad a claim to state that the SED is a metric that is robust to the effects of ecological perturbation on a community, it is valid to state that the SED is robust to factors that change the shape of the SAD.

We conclude that either the functional form of the exponent for the SED is incorrect, and/or there may be some fundamental information that the METE ASED and SED are missing. One solution that has been proposed is that higher taxonomic levels pose additional constraints on the energy distributions, a topic currently being explored by Harte and Rominger, et al. (*in preparation*). Both solutions point to problems with theory, which leaves open certain theoretical questions for further research.

Literature cited

Harte, J. 2011. Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics. Oxford Univ. Press, Oxford UK.

Jones E, Oliphant T, Peterson P. SciPy: Open Source Scientific Tools for Python. 2001. URL http://www.scipy.org/

Appendix K. Discussion of the normalization of the ASED, and exploration of a related metric, the Total Species Energy Distribution

Although it was expected that the ASED should be normalized such that the predicted and observed distributions should represent the same total metabolic rates, and the means of those distributions should be E_0/S_0, neither of these expectation were upheld for reasons that were not obvious immediately, but are discussed below.

The non-normalized ASED lead us to investigate a community-level metric that is normalized and should, by design, have a mean equal to E_0/S_0. The total species energy distribution (TSED) is similar to the ASED but equation 7.42 in Harte (2011) is multiplied by n to obtain the total energy per species (see Eq. 3 in the main text):

$$\varepsilon_{Tot} = n\bar{\varepsilon} \approx n + \frac{1}{\lambda_2} \tag{K.1}$$

The distribution for the total energy per species, the TSED, is then obtained using the change of variable formula for a discrete probability distribution (as the ASED was calculated without the Jacobian and was considered discrete in this analysis). The TSED, represented here by ω, is given by the equation:

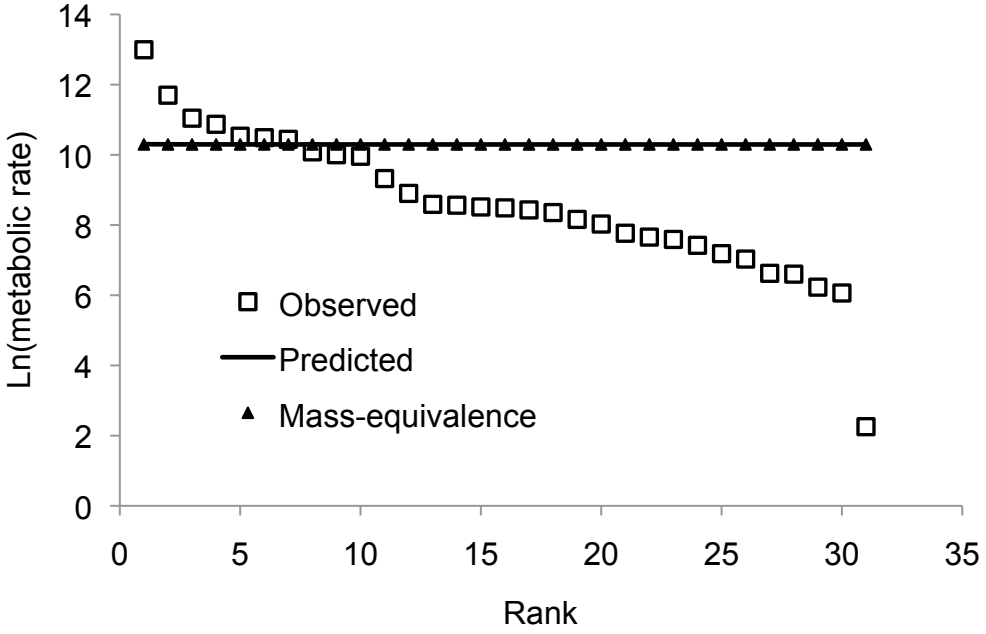
$$\omega = \Phi(n(\varepsilon_{Tot})) \tag{K.2}$$

where:

$$n(\varepsilon_{Tot}) = \varepsilon_{Tot} - \frac{1}{\lambda_2} \tag{K.3}$$

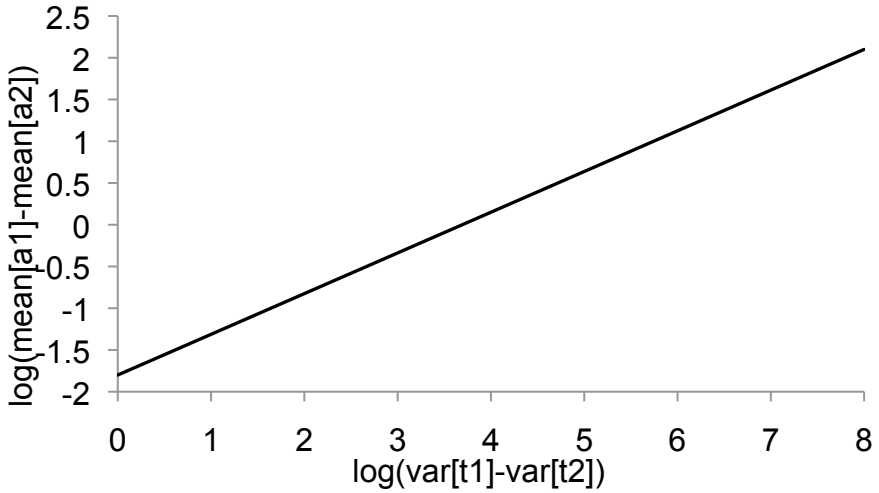
We confirmed that the mean of the observed and predicted TSED are equal to E_0/S_0. We compare empirical data to the METE predictions and an "energy equivalence" prediction (Damuth 1987), which is a model that specifies that each species in a given area will have the same total energy, or total metabolic rate. See Figure K.1.

Figure K.1 METE TSED prediction compared to empirical data and to an energy-equivalence prediction. The METE TSED is almost indistinguishable from the energy-equivalence prediction for this dataset.



There is a simple justification for why the *Nu* predicted ASED does not have to have the same mean as the observed ASED, even though observed and predicted TSEDs must have the same means. To illustrate, consider two vectors, t_1 and t_2, of finite length m, which will represent the observed and predicted TSED, respectively. The vector t_1 can be thought of as the TSED obtained by sampling a community of m species: each element in the vector is a species, and the value of the element is the total energy observed across all individuals in that species. The vector t_2 is the TSED predicted by some theory (in our case METE), which can be directly compared to t_1. Using the METE prediction for t_2, t_1 and t_2 must have the same mean. Let's then take another vector of length m called s_1. In our case, s_1 will be the empirical SAD where m is the number of species in the community. If we divide (using element-wise division) each of the original vectors (t_1 and t_2) by the new vector (s_1), we produce something that has the property of the empirical and observed ASEDs: $a_1 = t_1/s_1$ and $a_2 = t_2/s_1$. Although it might appear that the means of a_1 and a_2 have to be equal because the means of t_1 and t_2 are equal, straightforward algebra shows that this is only the case in a few unique situations. In general, the difference in the means of a_1 and a_2 varies as a function of the difference in the variance between t_1 and t_2. As the difference in the variance between t_1 and t_2 increases, the difference in the mean between a_1 and a_2 increases following a power-law (see Figure K.2). For our results, the difference in the variance between the observed and predicted TSEDs is very large, and the variance in the SAD is also large. Although not shown on the graph, the difference between the means of the two vectors a_1 and a_2 is also influenced by the variance of s_1, or in our case, the empirical SAD. This demonstrates why the means of the observed and predicted ASED do not have to match in practice.

Figure K.2 The difference of the means of two simple functions $a_1 = t_1/s_1$ and $a_2 = t_2/s_1$, resulting from the variance of their component vectors t_1 and t_2, as plotted on a log-log graph.



Finally, comparing the predictive value of METE versus energy equivalence, we expect METE makes predictions that are consistent with energy equivalence where variation in body size (or energy of individuals) is small, and expect METE to make more accurate predictions when the range of body sizes (or energy) in a community or within a species is large. Energy equivalence will only hold under conditions where $n << 1/\lambda_2$, alternately written as $n << (E_0-N_0)/S_0$ (see discussion in Harte 2011). The predicted TSED will often be very close to a horizontal line (energy equivalence) because n is typically small compared to $1/\lambda_2$, which is what we observe in this dataset.

Literature cited

Damuth, J. 1987. Interspecific allometry of population density in mammals and other animals: the independence of body mass and population energy-use. Biological Journal of the Linnean Society 31:193–246.

Harte, J. 2011. Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics. Oxford Univ. Press, Oxford UK.