

Overview

Communities are complex systems, in the technical sense: they have multitudes of components with context-dependent interactions, feedbacks, and stochasticity. Community ecology has a long tradition of trying to understand systems by focusing in detail on their constituent parts, but rules or even general principles have proven elusive (Lawton 1999, Simberloff 2004). A potentially complementary approach is to embrace and even harness ecological complexity. We can begin by focusing on the patterns that emerge when we blur the specifics and instead look at community properties in aggregate (Lawton 1999). Phenomena that emerge consistently in many systems may result from general processes that operate across systems, or may reflect constraints that restrict the possible outcomes of any random or systematic processes (Brown 1995, McGill 2019). Documenting these phenomena, and comparing them to our expectations based on potential generative processes or governing constraints – ranging from niche structure, to thermodynamic constraints on abundance and metabolic rate, to pure randomness – may help us rise above the fascinating but bewildering “mess” of community-level interactions and understand communities in simpler, more predictable, and more general terms (Lawton 1999, Brown 1995, McGill 2019).

To take this approach to ecological *communities*, we can summarize community properties with a combination of “state variables” and “community abundance distributions” (CADs). State variables describe an assemblage in terms of total individual abundance N , species richness S , and resource use as total energy use E (following, among others, Harte et al 2008). CADs relate the state variables to each other via community structure. I focus particularly on the species abundance distribution (the distribution of individual abundances among taxonomic categories in a community) and the size spectrum (the distribution of individuals, regardless of species, across body sizes within a community). These are measurable, sufficiently general to make meaningful comparisons across communities, and intimately linked to the processes and constraints we are interested in understanding.

So far, work on the species abundance distribution, the size spectrum, and the dynamics of abundance and resource use has yielded both “fuzzy generalities” (to borrow Lawton’s term) and major insights into the constraints that shape community dynamics in at least some systems (McGill 2019, Ernest et al 2008). However, progress on all of these fronts has stalled in the past decade due to a combination of barriers. The field struggles with a scarcity of assemblage-level individual size and metabolic data, a lack of quantitative methods for describing some CADs and relating them to qualitative expectations, and the puzzle of disentangling biological signal from patterns that emerge from null models or statistical constraints (White et al 2007, McGill et al 2007, Thibault et al 2011 *GEB*).

I will combine newly developed analytical tools for describing distributions and distinguishing between random and potentially process-driven patterns, data derived from new continent-scale monitoring programs and recently-established scaling relationships, and modern computational infrastructure to meet these obstacles and open a new chapter for macroecological approaches to community ecology. I will follow three intertwined strands to explore how abundance and resource use are distributed within terrestrial animal communities, how these distributions shift or remain stable over time, and what these distributions and dynamics can and cannot tell us about what drives community structure and function. Specifically, I will:

- I. Determine what, if any, *non-statistical* signal we can detect in the species abundance distribution, and test whether this signal corresponds to either broad community attributes or theoretical expectations for the response of macroecological patterns to perturbations
- II. Identify the common form(s) of the size spectrum in terrestrial animal communities, and compare observed forms to two major theories of resource availability and niche differentiation

- III. Estimate individual size and community level energy use data for North American bird communities over a forty-year timeseries, and test whether the temporal dynamics of size spectra and community-level energy use appear to be stable, shifting systematically, or indistinguishable from randomness

I. The species abundance distribution

Background

The species abundance distribution (SAD) has attracted considerable attention because it exhibits a strikingly consistent “hollow curve” form, with few common and many rare species, across nearly all communities (Lawton 1999, McGill et al 2007, Baldrige et al 2016). Somewhat paradoxically, this remarkable consistency has confounded efforts to evaluate theories that attempt to predict SADs. Most theories make similar predictions that correspond well to the empirical hollow curve, making it challenging to distinguish among competing theories (McGill et al 2007).

Recently, multiple lines of reasoning (Frank 2009 and 2019, Locey and White 2013, Harte et al 2011) have found that the SAD appears subject to a strong statistical constraint, which may interfere with our perception of the most informative aspects of the distribution. For a given number of individuals N divided among S species, there exists a finite array of *possible* or *probable* species abundance distributions – most of which turn out to be hollow curves. It is no surprise that most SADs, empirical and theoretical, are also hollow curves; in fact, it would be highly improbable to happen on an SAD that was *not* a hollow curve (Locey and White 2013). We stand to learn much more if we focus on how observed SADs compare to the distribution of possible, or probable, forms for the distribution. If we are unable to distinguish between observed SADs and what we would expect to emerge at random, there may not be much we can infer about biological process from the SAD. In contrast, if and when observed SADs deviate detectably from their statistically-most-likely forms, those deviations may betray biological processes driving systems away from randomness and offer new leverage for developing and testing theories (Locey and White 2013, Harte and Newman 2014).

I will build on the combinatorics-based approach developed by Locey and White (2013) to establish if and when we detect consistent deviations between observed SADs and their random baselines. Given that there exists a finite set of possible forms for an SAD with N individuals and S species (the “feasible set”), uniform samples from this set reflect the landscape of probable forms for that SAD (Locey and White 2013). By intensively sampling the feasible set for a given S and N , we can characterize both the nature and the strength of the statistical constraint on the SAD, and determine whether an observed SAD is surprising or unremarkable compared to its statistical baseline. I will apply this approach to a vast collection of empirical SADs (also analyzed in White et al 2012 and Baldrige et al 2016) to test whether we find consistent deviations from randomness, and the extent to which our capacity to detect these deviations depends on the statistical properties of the feasible set.

Although conceptual frameworks for *interpreting* any deviations we might find have mostly yet to be developed, we can begin with intuition borrowed from statistical mechanics (Harte and Newman 2014, Newman et al 2020). In systems at “steady state”, numerous processes may effectively cancel each other out and leave no detectable signal, but strong perturbations may temporarily overwhelm other processes and introduce large deviations from random expectations. I will explore this possibility using a database of community abundance data from experimentally-manipulated systems (Supp 2014, Ernest et al 2018).

Objectives and approach

1. Measure the magnitude and nature of deviations between empirical SADs and their statistical constraint

I will compare an extensive compilation of SADs (encompassing communities of trees, birds, mammals, and miscellaneous less-traditional taxa; White et al 2012 and Baldrige 2015) to their feasible sets. For each SAD, I will draw samples from the feasible set and compare the empirical distribution's shape, as measured by skewness and Simpson's evenness, to the distribution of shapes we would expect at random. I will test whether or not empirical distributions are unusually skewed or even relative to their feasible sets.

Because both the nature and specificity of likely forms for the SAD varies over large gradients of S and N, I will also test whether our capacity to *detect* deviations appears to be limited by the statistical properties of the feasible set. The more the elements of the feasible set converge on a particular form, they indicate a strong statistical constraint: it would be highly surprising to happen on a *different* form at random (Locey and White 2013). However, if there is a relatively even representation of different shapes within the feasible set, we may not be able to confidently say whether an observation is tracking or deviating from randomness. This phenomenon may be especially relevant for small communities (Jaynes 1957, Harte et al 2011, McGill 2019). I will test whether our ability to detect deviations suffers for small or less narrowly-constrained feasible sets. If this is the case, we may stand to learn the most from the SAD if we focus on large communities with highly specific statistical constraints.

2. Test whether manipulation increases deviations between observations and constraint

I will use community abundance data from the Manipulated Animal Community Database (Supp 2014), and the plant communities in experimentally manipulated plots at our field site near Portal, AZ (Ernest et al 2018), to compare the deviations associated with manipulated and unmanipulated animal and plant communities. If, in keeping with Newman and colleagues' suggestion (2020), disturbance drives communities away from their statistical baselines, manipulated communities should generally deviate more pronouncedly from their feasible set than controls.

II. The size spectrum

Background

The size spectrum is a potentially powerful representation of community structure, but remains surprisingly under-studied for terrestrial animal communities (Ernest 2005, 2013, White et al 2007, Thibault et al 2011 *GEB*). It relates numerical abundance to resource use and, to the extent body size is a proxy for functional traits, niche structure more broadly. Macroecological approaches to terrestrial animal size spectra have grappled with three major challenges: a lack of data at the appropriate scales, a lack of tools for comparing observation to theory, and the underappreciated possibility that fundamental constraints may generate "patterns" even without process to speak of (White et al 2007, Thibault et al 2011 *GEB*). In hopes of revitalizing this line of inquiry, I will compile the largest to-date database of animal size spectra, compare these observations to predictions from the leading relevant theories, and test whether it is possible to reliably distinguish between observation, process-based theoretical predictions, and simple null models.

Objectives and approach

1. Document size spectra for terrestrial mammals and birds

Thibault et al (2011, *GEB*) established a method for simulating community-level size data for determinate growers based on species' mean body sizes and a scaling relationship between body size and intraspecific size variation. I will use this method to re-create the database of bird community size data from Thibault et al (2011, *GEB*), and simulate an additional database of mammal community size data based on the Mammal Community Database (Thibault et al 2011, *Ecology*). Finally, I will use the recently-available NEON small mammal trapping data as gold-standard individual size data for rodent communities.

2. Compare empirical size spectra to theoretical predictions and null models

We have two major conceptual frameworks for interpreting size spectra, but have not evaluated them with high-quality data and appropriate null models. Holling's Textural Discontinuity Hypothesis (Holling 1992) suggests that resources are available to different body sizes in discontinuous pockets, and that the abundance of species and along the body size axis should correspond to these pockets. Scheffer and Van Nes (2006) simulated the evolutionary dynamics of species in competition along a trait axis, and uncovered a possibility they call "self-organized similarity": species can coexist by being very similar or very different, but intermediately-similar species cannot coexist. Both theories predict some degree of "clumpiness" in the size spectrum, with multiple modes separated by distinct gaps – in contrast to a uniform, hump-shaped, or power-law distribution.

Testing for clumpiness in empirical data is a non-trivial analytical challenge, because clumpiness is not exactly a rigorously defined statistical property. I will quantify clumpiness using a combination of an overlap metric developed to measure species similarity in rodent communities (Read et al 2018) and modularity metrics borrowed from network theory. We can represent a community as a network in which species are nodes and edges are weighted by the pairwise overlap in species' body size distributions. A highly clumped community is analogous to a highly modular network: some groups of species are strongly connected to each other, but weakly to others, and no species are moderately strongly connected.

Determinate growth, especially combined with a realistic abundance distribution, can produce a clumpy size spectrum even without other processes. I will adjust for this possibility by simulating communities in which species' sizes are randomly distributed relative to each other, but there is limited intraspecific variation in size. I will test whether observed communities exhibit greater clumpiness, consistent with both textural discontinuity and self-organized similarity, than would emerge simply due to determinate growth.

III. Temporal dimensions of abundance, energy use, and size structure

Background

While snapshots of community state variables and distributions are informative, we open entirely new horizons if we can incorporate *temporal* dimensions of community properties (Fisher et al 2010). Especially in an era of widespread changes, and apparent declines, in biodiversity, it is imperative to establish whether and in what ways *communities* are buffered against declines in overall function or shifts in composition or structure (Fisher et al 2010, Dirzo et al 2014). So far, macroecological studies on the temporal dimensions of size structure and community energy use have been limited by a lack of community timeseries with measurements of individual body size, let alone energy use. I will extend the method demonstrated in Thibault et al (2011, *GEB*) to estimate individual-level size and metabolic rate data over the *timeseries* of the Breeding Bird Survey (Pardieck et al 2019). I will use this database to expand our investigations of size structure into the temporal domain and characterize the dominant trends in community abundance, richness, and energy use for these communities over the past 40 years.

Objectives and approach

1. Simulate individual size data for the Breeding Bird Survey timeseries

Thibault et al (2011, *GEB*) estimated individual size data for the Breeding Bird Survey from 2003 to 2007. It is straightforward to extend the method to the full timeseries, in what will be the largest compilation of community timeseries with realistic estimates of individual size to date. The applications of this database extend well beyond the projects I outline here, and I will make it immediately available for other researchers to use.

2. Test whether size spectra are stable, shifting, or indistinguishable from random fluctuations

Size spectra may reflect the distribution of resources, or other dimensions of niche opportunity, over the body size axis. This is the intuition behind both the textural discontinuity hypothesis and self-organized similarity (Holling 1992, Scheffer and Van Nes 2006). More broadly, to the extent that size is a proxy for functional traits, the expectation that the distribution of traits in a community should be organized according to the available niches pervades community ecology (e.g. McGill et al 2006). If this is the case, the temporal dynamics of the size spectrum signal how a community is – or is not – changing substantively over time. A size spectrum that remains consistent over time, even as particular species come and go, could reflect a stable underlying niche structure (Holling 1992, Thibault et al 2011 *GEB*). If it shifts systematically, it could reflect changes in the types of opportunities available (as we have seen in the rodent community near Portal, AZ; Ernest et al 2008, Thibault et al 2010). However, either random fluctuations or systematic processes unrelated to body size could *also* produce stability or drift in size spectra. Before we can interpret apparently fixed or shifting size spectra in terms of ecological processes, we must establish whether we can detect signals of stability or systematic change that we are confident are not random.

To establish what we can learn from size spectra as they change over time, I will test if, and in what ways, the observed dynamics of size spectra are detectably different from random fluctuations. Rather than track entire size spectra over time, I will focus on dynamics of *the rate of change* of the size spectrum, as measured by the overlap between size spectra from different time steps. As a preliminary approach, I expect that non size-structured fluctuations will generate change in size spectra between time steps, but that the rate of change will not vary systematically over time. I reason that process-driven shifts in size spectra would operate over longer timescales and be more consistent than random fluctuations, and so communities undergoing systematic shifts should have periods of relatively rapid change bracketed with periods of relative stasis. I will use simulated communities with random, stable, and systematically shifting dynamics to validate these expectations and test the extent to which this approach can detect random or nonrandom dynamics. I will then use simulations based on the initial conditions for the communities in the Breeding Bird Survey as baselines to test whether these observed size spectra have trajectories that are more consistent with stability, systematic change, or random fluctuations.

3. Document trends in community-level energy use for North American birds, and relate them to trends in population and total abundance

We live in an age of unprecedented change to biodiversity, and the dominant narrative is one of declines in species richness, abundances of populations, and ecological “function” (Dirzo et al 2014). Somewhat surprisingly, the largest survey to date of trends in *community* timeseries found, on average, *no* trend in local species richness, roughly equal numbers of declining and increasing populations, and evidence for regulated trajectories of total abundance (Dornelas et al 2014, 2019, Gotelli et al 2017). However, we should not be too quick to take this to mean indicate that assemblages are wholly buffered against change, because richness and individual abundance do not tell us enough to distinguish between qualitatively different phenomena. Knowing that the number of individuals in a community is stable does not tell us whether the *community* is unchanging, or if, for example, large species are being replaced by small species – which would likely correspond with a decline in total energy use and biomass and a shift in the traits represented in the community.

I will test whether there are consistent trends in community energy use for the BBS communities, and whether these correspond to trends (or a lack thereof) in total and population-level abundance. This will provide a crucial dimension to our understanding of how biodiversity is shifting in the Anthropocene, and improve our context for interpreting the findings of abundance and richness-based studies.

Additional objectives

Reproducible and open science

I will continue to contribute to R packages in development in our lab and build packages to support my projects. I will release the databases I build on a public archive, with minimal restriction on re-use. I will release self-contained, tested research compendia for each of my papers, to facilitate review, replication, and re-use of methods. As by-products of several of my projects, I will retrace earlier authors' efforts: Ernest (2005), Thibault et al (2011 *GEB*), and potentially Gotelli et al (2018). I will document the results of these replications and release them as replication reports.

Teaching and course development

I will TA twice, beginning in Fall 2020. I will seek out opportunities to develop teaching materials for courses on data science and/or reproducible research. I am very interested in the potential for expanding the availability and quality of these trainings, especially using pedagogy that actively dismantles entrenched biases in who is best-served by learning opportunities. I will also complete The Carpentries' instructor training and volunteer to lead Data Carpentry workshops.

Undergraduate mentorship

Our lab expects to hire an REU student beginning in fall 2020, and I will assist at all stages of hiring and working with this student. I will help them identify and tackle a research project appropriate for their interests and training goals. As needed, I will provide training in field and computational methods and advise them on professional opportunities. I will continue to bring interested undergraduates to the field as volunteer research assistants. I will also continue to provide undergraduates guidance and advice on research opportunities and career trajectories in field and computational ecology via workshops with the UF Natural Resources Diversity Initiative and Wildlife Graduate Student Association.

References

- Baldrige, E. (2015). MiscAbundanceDB_main. <https://doi.org/10.6084/m9.figshare.95843.v4>
- Baldrige, E., Harris, D. J., Xiao, X., & White, E. P. (2016). An extensive comparison of species-abundance distribution models. *PeerJ*, 4, e2823. <https://doi.org/10.7717/peerj.2823>
- Brown, J. H. (1995). *Macroecology*. University of Chicago Press.
- Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., & Collen, B. (2014). Defaunation in the Anthropocene. *Science*, 345(6195), 401–406. <https://doi.org/10.1126/science.1251817>
- Dornelas, M., Gotelli, N. J., Shimadzu, H., Moyes, F., Magurran, A. E., & McGill, B. J. (2019). A balance of winners and losers in the Anthropocene. *Ecology Letters*, 22(5), 847–854. <https://doi.org/10.1111/ele.13242>
- Ernest, S. K. M. (2005). Body Size, Energy Use, and Community Structure of Small Mammals. *Ecology*, 86(6), 1407–1413. <https://doi.org/10.1890/03-3179>
- Ernest, S. K. M. (2013). Using Size Distributions to Understand the Role of Body Size in Mammalian Community Assembly. In F. A. Smith & S. K. Lyons, *Animal Body Size: Linking Pattern and Process across Space, Time, and Taxonomic Group*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226012285.001.0001>
- Ernest, S. K. M., Brown, J. H., Thibault, K. M., White, E. P., & Goheen, J. R. (2008). Zero Sum, the Niche, and Metacommunities: Long-Term Dynamics of Community Assembly. *The American Naturalist*, 172(6), E257–E269. <https://doi.org/10.1086/592402>
- Ernest, S. K. M., Yenni, G. M., Allington, G., Bledsoe, E. K., Christensen, E. M., Diaz, R. M., Geluso, K., Goheen, J. R., Guo, Q., Heske, E., Kelt, D., Meiners, J. M., Munger, J., Restrepo, C., Samson, D. A., Schutzenhofer, M. R., Skupski, M., Supp, S. R., Thibault, K., ... Valone, T. J. (2018). The Portal Project: A long-term study of a Chihuahuan desert ecosystem [Preprint]. *bioRxiv*. <https://doi.org/10.1101/332783>
- Fisher, J. A. D., Frank, K. T., & Leggett, W. C. (2010). Dynamic macroecology on ecological time-scales. *Global Ecology and Biogeography*, 19(1), 1–15. <https://doi.org/10.1111/j.1466-8238.2009.00482.x>
- Frank, S. A. (2009). The common patterns of nature. *Journal of Evolutionary Biology*, 22(8), 1563–1585. <https://doi.org/10.1111/j.1420-9101.2009.01775.x>
- Frank, Steven A. (2019). The common patterns of abundance: The log series and Zipf's law. *F1000Research*, 8, 334. <https://doi.org/10.12688/f1000research.18681.1>
- Harte, J., Zillio, T., Conlisk, E., & Smith, A. B. (2008). Maximum Entropy and the State-Variable Approach to Macroecology. *Ecology*, 89(10), 2700–2711. <https://doi.org/10.1890/07-1369.1>
- Harte, John. (2011). *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199593415.001.0001>

- Harte, John, & Newman, E. A. (2014). Maximum information entropy: A foundation for ecological theory. *Trends in Ecology & Evolution*, 29(7), 384–389. <https://doi.org/10.1016/j.tree.2014.04.009>
- Holling, C. S. (1992). Cross-Scale Morphology, Geometry, and Dynamics of Ecosystems. *Ecological Monographs*, 62(4), 447–502. <https://doi.org/10.2307/2937313>
- Jaynes, E. T. (1957). Information Theory and Statistical Mechanics. *Physical Review*, 106(4), 620–630. <https://doi.org/10.1103/PhysRev.106.620>
- Lawton, J. H. (1999). Are There General Laws in Ecology? *Oikos*, 84(2), 177. <https://doi.org/10.2307/3546712>
- Locey, K. J., & White, E. P. (2013). How species richness and total abundance constrain the distribution of abundance. *Ecology Letters*, 16(9), 1177–1185. <https://doi.org/10.1111/ele.12154>
- McGill, B. J. (2019). The what, how and why of doing macroecology. *Global Ecology and Biogeography*, 28(1), 6–17. <https://doi.org/10.1111/geb.12855>
- McGill, B. J., Enquist, B. J., Weiher, E., & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, 21(4), 178–185. <https://doi.org/10.1016/j.tree.2006.02.002>
- McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., Dornelas, M., Enquist, B. J., Green, J. L., He, F., Hurlbert, A. H., Magurran, A. E., Marquet, P. A., Maurer, B. A., Ostling, A., Soykan, C. U., Ugland, K. I., & White, E. P. (2007). Species abundance distributions: Moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters*, 10(10), 995–1015. <https://doi.org/10.1111/j.1461-0248.2007.01094.x>
- Newman, E. A., Wilber, M. Q., Kopper, K. E., Moritz, M. A., Falk, D. A., McKenzie, D., & Harte, J. (2020). Disturbance macroecology: A comparative study of community structure metrics in a high-severity disturbance regime. *Ecosphere*, 11(1), e03022. <https://doi.org/10.1002/ecs2.3022>
- Pardieck, K. L., Ziolkowski, D. J., Lutmerding, M., Aponte, V., & Hudson, M.-A. (2019). *North American Breeding Bird Survey Dataset 1966–2018, version 2018.0* [Csv]. U.S. Geological Survey. <https://doi.org/10.5066/P9HE8XYJ>
- Read, Q. D., Grady, J. M., Zarnetske, P. L., Record, S., Baiser, B., Belmaker, J., Tuanmu, M.-N., Strecker, A., Beaudrot, L., & Thibault, K. M. (2018). Among-species overlap in rodent body size distributions predicts species richness along a temperature gradient. *Ecography*, 41(10), 1718–1727. <https://doi.org/10.1111/ecog.03641>
- Scheffer, M., & van Nes, E. H. (2006). Self-organized similarity, the evolutionary emergence of groups of similar species. *Proceedings of the National Academy of Sciences*, 103(16), 6230–6235. <https://doi.org/10.1073/pnas.0508024103>
- Simberloff, D. (2004). Community Ecology: Is It Time to Move On?: (An American Society of Naturalists Presidential Address). *The American Naturalist*, 163(6), 787–799. <https://doi.org/10.1086/420777>

- Supp, S. (2014). Manipulated Animal Community Database. Figshare.
<https://doi.org/10.6084/m9.figshare.969831.v1>
- Thibault, K. M., Ernest, S. K. M., & Brown, J. H. (2010). Redundant or complementary? Impact of a colonizing species on community structure and function. *Oikos*, 119(11), 1719–1726. JSTOR.
- Thibault, K. M., Supp, S. R., Giffin, M., White, E. P., & Ernest, S. K. M. (2011). Species composition and abundance of mammalian communities. *Ecology*, 92(12), 2316–2316.
<https://doi.org/10.1890/11-0262.1>
- Thibault, K. M., White, E. P., Hurlbert, A. H., & Ernest, S. K. M. (2011). Multimodality in the individual size distributions of bird communities: Avian individual size distributions. *Global Ecology and Biogeography*, 20(1), 145–153. <https://doi.org/10.1111/j.1466-8238.2010.00576.x>
- White, E. P., Ernest, S. K. M., Kerkhoff, A. J., & Enquist, B. J. (2007). Relationships between body size and abundance in ecology. *Trends in Ecology & Evolution*, 22(6), 323–330.
<https://doi.org/10.1016/j.tree.2007.03.007>
- 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.
<https://doi.org/10.1890/11-2177.1>