Chapter 3- Evaluating abundance distribution based signals of neutrality in terrestrial systems

Introduction

One of the fundamental goals of ecology is understanding what processes are important in structuring ecological communities. One of the major areas of debate surrounding this goal is whether simple neutral models that ignore differences between species can explain many of the empirical patterns observed in ecological systems (McGill et al. 2006, Rosindell et al. 2012, Matthews and Whittaker 2014). While there are multiple formulations of neutral theory, all models are based on the assumption that species and individuals are ecologically and demographically equivalent to one another, meaning that stochastic variation in birth, death, immigration, and speciation drives differences in a broad array of ecological patterns including the species abundance distribution, the species-area relationship, and the distance decay of similarity (Rosindell et al. 2011).

Early evaluations of neutral theory were based, in part, on comparing the fit of empirical species abundance distributions to the neutral prediction (e.g., Hubbell 2001, McGill 2003, Volkov et al. 2003). However, further evaluations of neutral theory suggested that comparisons based on the species abundance distribution were not sufficient for rigorous tests of neutrality (Volkov et al. 2005, 2006, McGill et al. 2006). This idea is further supported by work suggesting that species abundance distributions may contain little information about the detailed processes operating in ecological system more generally (Pielou 1975, White et al. 2012, Locey and White 2013). In contrast, recent work by Connolly et al. (2014) suggests that comparisons of species abundance distributions may be sufficient for evaluating whether or not neutral processes are dominant or whether other processes are important in structuring communities.

Building on work by Pueyo (Pueyo 2006), Connolly et al. (2014) were able to demonstrate that simulated neutral communities were typically better fit by negative-binomial distributions (referred

to as Poisson gamma distributions by Connolly et al. 2014) than by Poisson lognormal distributions. They then performed the same analysis on over 1000 marine communities, and showed that the empirical communities were better fit by the lognormal (Connolly et al. 2014). This suggests that, at least in marine environments, the shape of the species abundance distribution can be used to exclude neutral processes as the sole determinant of community structure. By focusing on the detailed fits of alternative models, this approach takes advantage of 'second-order effects', which have been proposed to provide an avenue for inferring ecological process based on patterns of community structure (Blonder et al. 2014).

While this approach has been well tested within marine communities, it has not yet been used in terrestrial systems. Here, we use Connolly et al.'s (Connolly et al. 2014) method to assess potential patterns of neutrality across a broad range of ecosystems and taxonomic groups. We tested this approach for vertebrate, invertebrate and plant communities in primarily terrestrial ecosystems. In total, we used abundance data from 16,218 communities from across to globe to determine whether we observe patterns that are more consistent with neutrality or non-neutrality.

Methods

Data

We compiled data from 9 distinct taxonomic groups and include birds, mammals, reptiles, amphibians, beetles, spiders, butterflies, trees, and bony fish from 16,218 distinct communities over all major biogeographic regions (Table 1, Figure 1). This dataset is a combination of the data compiled by White et al. 2012 (White et al. 2012) and the data described in Chapter 2 of this dissertation. The majority of the data are publicly available and were accessed through the EcoData Retriever (Morris and White 2013). These data included the US Geological Survey's North American Breeding Bird Survey (BBS; Pardieck et al. 2014), Mammal Community Database (MCDB; Thibault et al. 2011), US Forest Service Forest Inventory and Analysis (FIA; USDA Forest Service 2010), and Gentry's

Forest Transect Data Set (Gentry; Phillips and Miller 2002), and the data from Chapter 1. The North American Butterfly Association count data (NABC; North American Butterfly Association 2009) and the Audubon Society Christmas Bird Count (CBC; Society 2002) are not publicly available and were obtained through Memorandums of Understanding with their respective organizations.

Table 1: Description of total number of sites per taxa and dataset. Taxonomic groups are ordered by the total number of sites in the compiled dataset.

Taxa	Dataset(s)	Total sites
Trees	FIA, Gentry	10575
Birds	BBS, CBC	4768
Butterflies	NABC	400
Reptiles	Reptilia	138
Bony fish	Actinopterygii	161
Mammals	MCDB	103
Amphibians	Amphibia	43
Spiders	Arachnida	25
Beetles	Coleoptera	5

The locations of all of the georeferenced data are presented in Figure 1. Note that the data for reptiles, amphibians, bony fish, beetles, spiders, and butterflies are not represented, due to a lack of detailed location data.
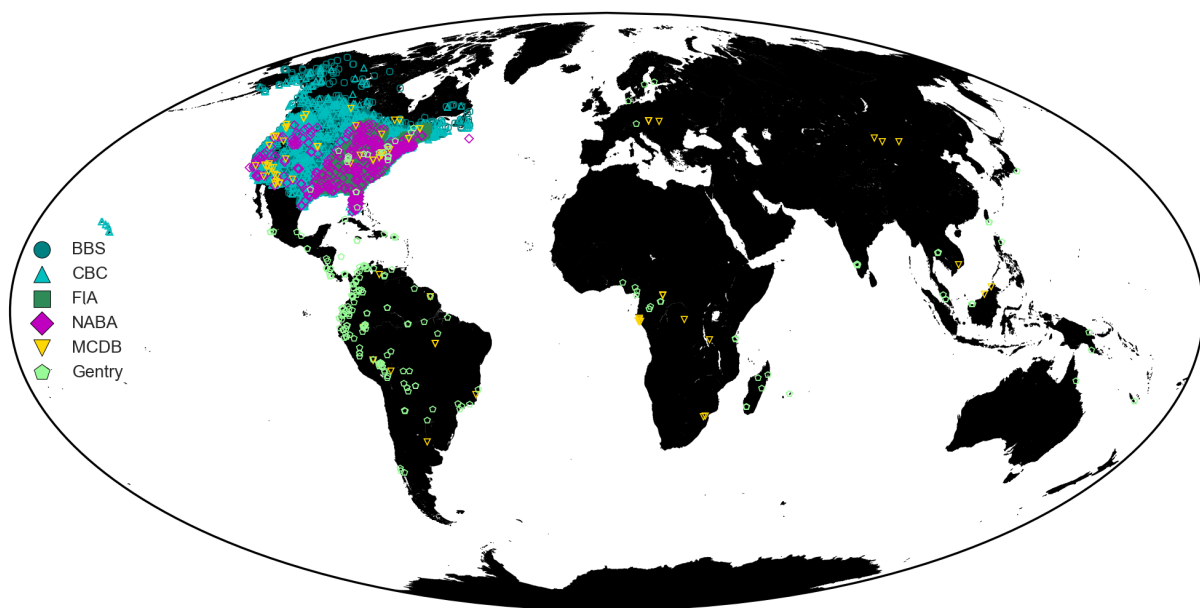
Figure 1: Map of the georeferenced portion of the data. Note that the data for reptiles, amphibians, bony fish, beetles, spiders, and butterflies are not represented. Redrawn from White et al. 2012

Analysis

Following Connolly et al., we used maximum likelihood methods for fitting and evaluating species abundance distributions models to data (the currently accepted best practice) (White et al. 2008, Connolly et al. 2014, Matthews and Whittaker 2014). This yielded fits of each distribution to each of the 16,000 communities in the dataset (Figure 2).
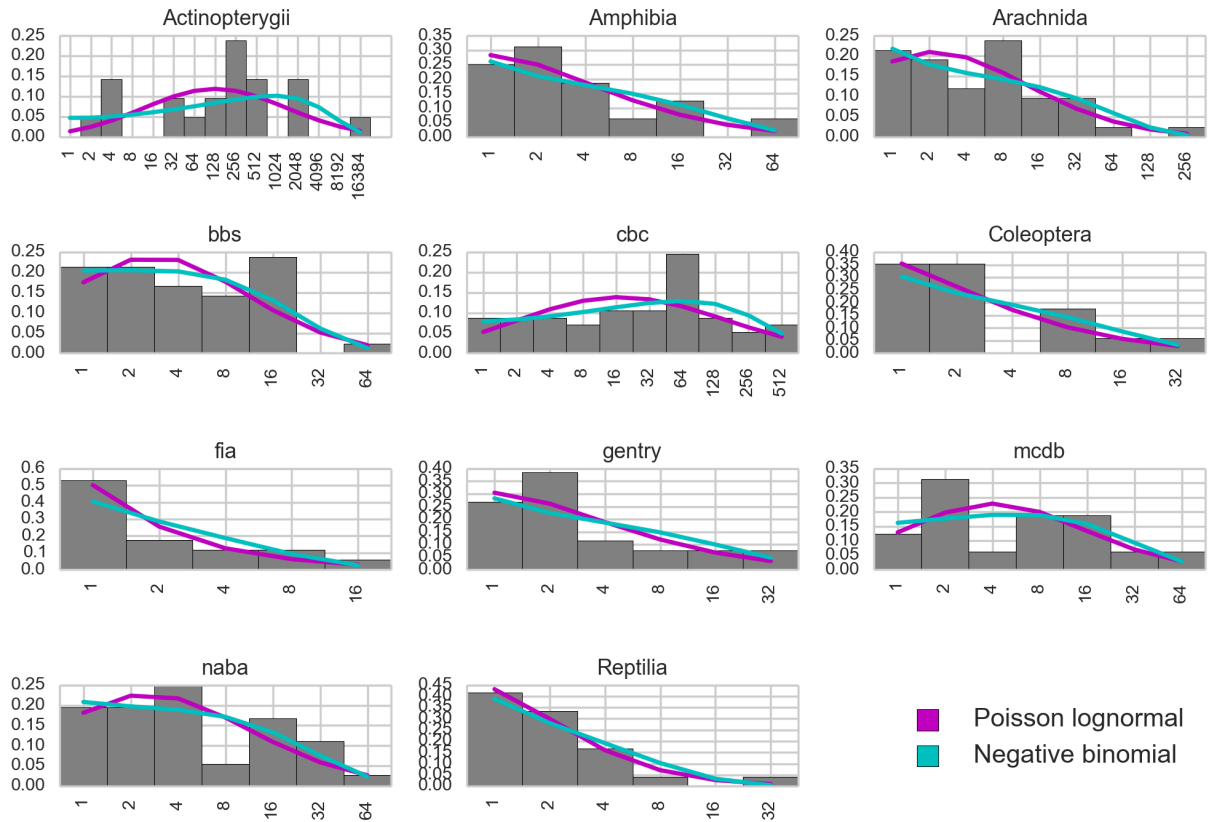

Figure 2: Preston plot of empirical data for single sites from each dataset with lines representing the Poisson lognormal and the negative binomial.

Connolly et al. used Akaike Information Criterion (AIC) weights to compare the fits of the negative-binomial and Poisson lognormal distributions to the empirical data. We modified this approach slightly by using weights calculated from the corrected Akaike Information Criterion (AICc) values, because AICc is more robust to small sample sizes (Burnham and Anderson 2002), which was a consideration for some communities. Model weights were calculated relative to the Poisson

lognormal, meaning that weights near zero support the negative-binomial as the better fitting model while weights near one support the Poisson lognormal as the better fitting model.

Following the approach of Connolly et al. (2014) we looked at the relationship between the AICc weight and the number of distinct abundance values in the dataset. Connolly et al. (2014) have argued that sites with more distinct abundance values provide greater power for differentiating between the fit of different models of the abundance distribution. Therefore, if the Poisson lognormal is superior to the negative-binomial distribution, the prediction is that AICc weights should approach one as the number of distinct abundance values increase. In Connolly et al.'s analysis, distinct abundance values greater than ~15 yielded AICc weights consistently above 0.8.

Data were first analyzed at the level of the individual site. We also evaluated the patterns of the average AICc and number of distinct abundance values for each dataset as a whole. These approaches differ somewhat from those of Connolly et al., in that: 1) do not present individual site level results; and 2) the structure of our data is different from Connolly et al.'s in that there are not natural spatial groupings, and therefore grouping at different spatial scales is less natural. Therefore, we have only analyzed the patterns at the site and whole dataset levels.

Results

The site level results show a large amount of scatter in the values of AICc both within and among datasets (Figure 3). Values in all datasets with reasonably large numbers of data points range from near zero to near 1. There is no consistent directional trend in AICc weight as a function of the number of distinct abundance values. In datasets where there is some directional trend in AICc with the number of distinct abundance values (Reptilia, BBS, FIA), the trend tends to be towards zero (i.e., a better fit for the negative binomial distribution) as opposed to the trend toward one (i.e., a better fit for the Poisson lognormal) observed by Connolly et al. in marine systems (Figure 3).

Averaging the AICc weights and distinct abundance values across all sites in a dataset yielded
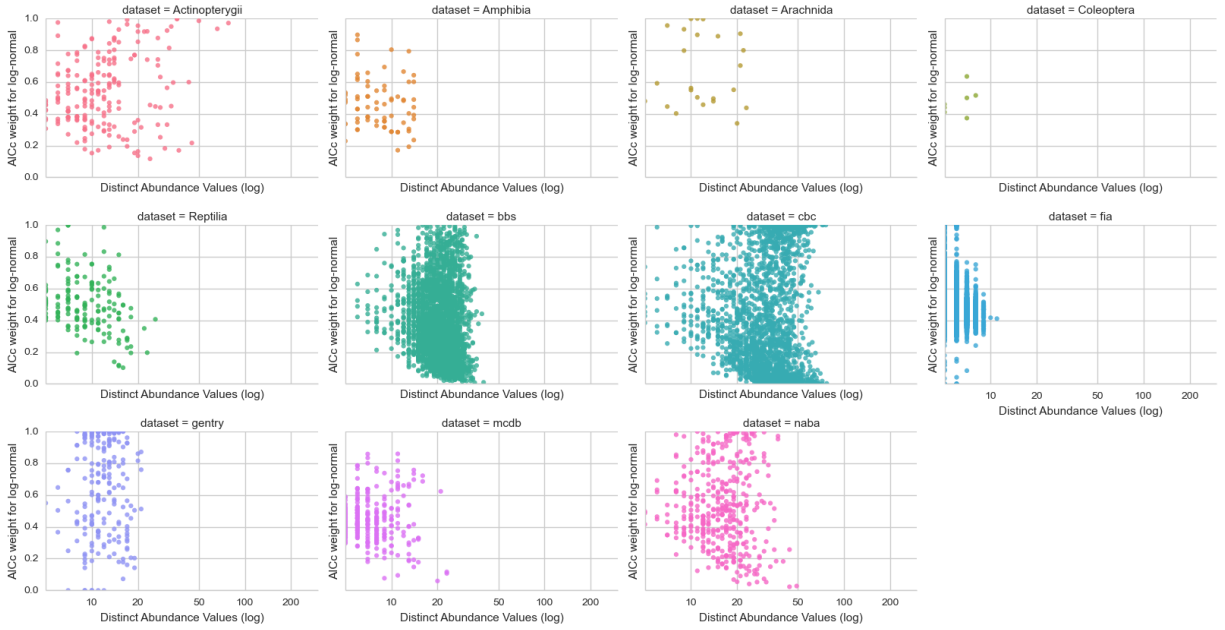
6

Figure 3: Log of distinct abundance values versus AIC weight of the lognormal distribution for each dataset.

similar results, with all average AICc values being between 0.35 and 0.7, and no notable trend in average AICc as a function of the number of distinct abundance values (Figure 4).

Discussion

The use of the species abundance distribution as a tool for identifying the processes operating in ecological systems has been widely questioned (Volkov et al. 2005, 2006, **???**, McGill et al. 2006), which makes recent results showing that it is possible to use the SAD to evaluate whether neutral processes are the dominant structuring process in ecological systems exciting (Connolly et al. 2014). In contrast to Connolly et al.'s results, which suggest that marine systems are demonstrably non-neutral in the structure of their SADs, our analysis suggests that terrestrial systems overall cannot be clearly defined as either neutral or non-neutral based on this type of analysis. Our results were consistent with our broad comparison of five different species abundance distribution models, which showed that it is difficult to identify a clear winning model (see details in Chapter 3 of this
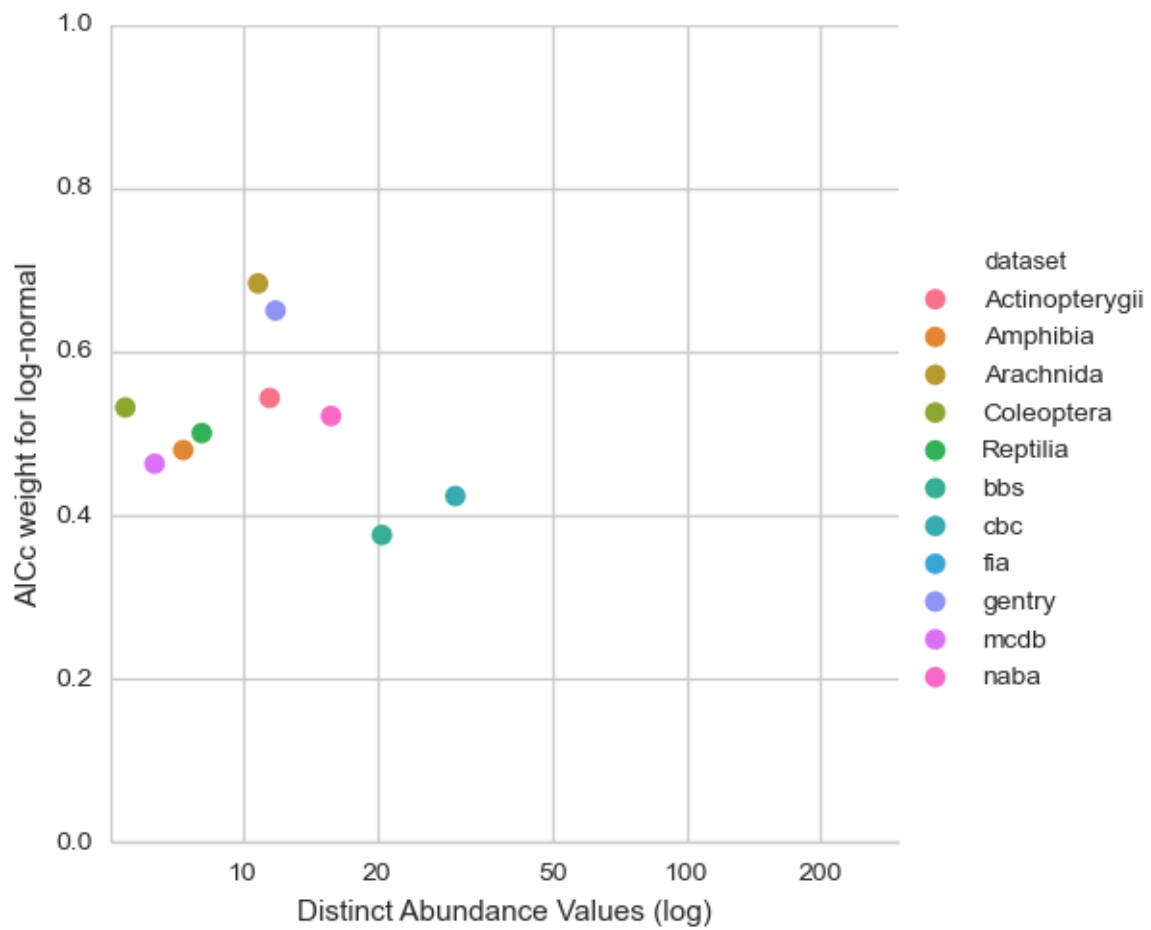
Figure 4: Average AIC weight of the lognormal distribution for each dataset.

dissertation). They are also consistent with a number of studies that have suggested that it should be difficult to identify underlying processes from the form of the abundance distribution alone (Pielou 1975, Volkov et al. 2005, 2006, McGill et al. 2006, White et al. 2012, Locey and White 2013, **???**). However, most of these studies focused on either terrestrial data or models originally based on terrestrial ecosystems. This suggests that there may be important differences between marine and terrestrial systems with regards to the processes operating in these systems and/or the ability to make inferences about these processes based on patterns like the SAD.

In combination, our results and those of Connolly et al. (2014) suggest that while marine systems are generally approximated by non-neutral dynamics, terrestrial systems show more variability between neutral and non-neutral dynamics. Several studies have noted that both patterns and processes may vary between marine and terrestrial systems (Webb 2012, Horne et al. 2015). While macroecological patterns have not been studied as extensively in marine systems as in terrestrial systems, marine and terrestrial systems do tend to exhibit many of the same general macroecological patterns (Webb 2012). However, Webb (2012) points out that while the same general patterns may occur, the processes generating those patterns may be different (Webb 2012). This could lead to subtle differences in the details of the patterns ('second order effects'), which Blonder et al. (2014) noted as the most promising avenue for identifying process using macroecological patterns.

<–I think this is better, and rewritten to be more precise in meaning.–> One key difference between terrestrial and marine systems is the way in which these systems have been differently impacted by anthropogenic activities. There has been a difference in the historical intensity and patterns of resource extraction in marine systems (Goudie 2013). Humans, being terrestrial, have been able to exploit terrestrial and coastal systems for a long period of time (Grayson 2001, **???**). Only recently in human history has technology advanced to intensively exploit non-coastal marine systems (resulting in such effects as fisheries collapses) (crain2008interactive; Jackson et al. 2001). One of the major differences is that in marine systems, anthropogenic efforts to extract resources have focused primarily on wild populations of consumers, while in terrestrial systems these efforts have focused more on land use for domesticated producers and consumers (Goudie 2013). These

differences lead to direct influences on marine species, but indirect effects in terrestrial systems (e.g., through land use changes)[marine, Jackson et al. (2001); Tittensor et al. (2009), terrestrial; haberl2007quantifying]. This legacy of marine exploitation and over-exploitation is a distinctly non-neutral influence on the structure of marine species abundance distributions that has the potential to produce a strong non-neutral signal in the SAD. In fact, there is an area of research using abundance distributions in marine systems to identify disturbed systems (Gray et al. 1979, Patil and Taillie 1982, Warwick 1986, Magurran 2013). There are several additional potential explanations for the difference in results between our study and the Connolly 2014 paper, some non-biological (spatial structuring, sampling intensity), others related to biological/ecological differences in the data.

Other significant differences also exist between terrestrial and marine systems that could result in the differences we observed. For example, marine and aquatic ecosystems can exhibit an inverted biomass pyramid when compared to terrestrial systems, dependant on the temporal scale of analysis (Trebilco et al. 2013). If species-abundance distributions are actually being structured based on a currency other than the number of individuals, e.g., biomass or resource use (Thibault et al. 2004, Connolly et al. 2005, McGill et al. 2007, Morlon et al. 2009), then this difference in the relationship between biomass and abundance could create a difference between marine and terrestrial systems. Another potential explanation comes from the core-occasional/core-transient species concept, in which core species, which are both common and temporally persistent, demonstrate a different shape of the species abundance distribution than transient species, which are rare and temporally variable (Magurran and Henderson 2003, Ulrich and Zalewski 2006, Magurran 2007). Differences in proporations of core and transient species occurring in terrestrial and marine systems could drive a difference in the general form of the abundance distribution, and greater variation in the proportions of core vs. transient species in terrestrial systems could drive the higher variation in the results. It is know that significant variation in the proportion of core vs. transient species exists in bird communities (Coyle et al. 2013), but there is little information on how variable these proportions are in marine systems.

While the vast majority of the data that we tested was terrestrial (approximately 99%), our data did

include 161 fish communities. Of these, the majority were freshwater, rather than marine. However we observed the same general pattern of results for for both freshwater and marine fish communities as we did for the other taxa. This leaves open the possibility that some non-biological difference between the two data compilations is driving the differences in the results.

One potential non-biological explanation for the difference between our results and Connolly et al.'s (2014) results is a difference in the spatial structure of the data: the data from Connolly 2014 is structured in natural spatial groupings, whereas the data that we used in this study is not. In this study, many of the sites are widely dispersed, or are not regularly dispersed over the landscape. These differences in spatial grouping may lead to results that are more consistent, due to spatial similarity than our widely dispersed sites.

Another potential non-biological explanation is related to sampling intensity. It is possible that the way in which marine communities are sampled is different from sampling of terrestrial communities, resulting in differing intensity of sampling that produce different patterns. However, the diversity of data we used covers a broad range of sampling intensities, from complete censuses (completely sampled trees above some minimum size cutoff; Forest Inventory and Analysis and Gentry), to well sampled but incomplete surveys (e.g., Christmas Bird Count), to incompletely sampled and incomplete taxonomic resolution (e.g. spiders and beetles).

In general, because of the diversity of data sources and types in our data compilation, it seems unlikely that the differences are due to the non-biological sampling differences rather than biological differences. Our compilation includes data collected at scales from a few square meters (e.g., invertebrate surveys) to 10s of square meters (Forest Inventory and Analysis) to 10s of hectares (Christmas Bird Count). As noted above they also include samples ranging from nearly complete surveys (trees) to communities where sampling of the local community is expected to be fairly sparse (invertebrates). Sampling also spans a broad array of general approaches including the use of traps, visual observations, identification by sound, and other approaches. The consistency of these results across diverse datasets makes it unlikely that any particular sampling approach/issue could

11

generate the observed results.

Intermediate to the sampling and biological explanations for the observed difference between marine and terrestrial systems is differences in the core constraints on the observed abundance distribution. A variety of approaches for modeling species abundance distributions suggest that species richness (S) and the total number of individuals (N) are important inputs that constrain the shape of the empirical pattern (Harte et al. 2008, Frank 2011, Harte 2011, White et al. 2012, Locey and White 2013). Consistent differences in the ratio of S/N for terrestrial vs. marine communities could provide another potential explanation. Further research needs to be done to determine if there is a difference in S/N ratios between the terrestrial data used in this study and the marine data used in Connolly 2014.

An increasing number of studies, including this one, suggest that there may be meaningful differences between marine and terrestrial systems in macroecological patterns [Webb (2012); horne2015]. While the 'first order' shape of these patterns may appear consistent between marine and terrestrial systems, there may be notable 'second order' differences (Blonder et al. 2014) related to differences in the processes driving the pattern (Webb 2012). This highlights the need for greater integration between the traditionally isolated analyses of marine and terrestrial systems in order to help understand differences in the processes driving these systems and the patterns that result (Beck et al. 2012, **???**) .

References

Beck, J., L. Ballesteros-Mejia, C. M. Buchmann, J. Dengler, S. A. Fritz, B. Gruber, C. Hof, F. Jansen, S. Knapp, H. Kreft, and others. 2012. What's on the horizon for macroecology? Ecography 35:673–683.

Blonder, B., L. Sloat, B. J. Enquist, and B. McGill. 2014. Separating macroecological pattern and process: comparing ecological, economic, and geological systems. PlOS one 9:e112850.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer.

Connolly, S. R., T. P. Hughes, D. R. Bellwood, and R. H. Karlson. 2005. Community structure of corals and reef fishes at multiple scales. Science 309:1363–1365.

Connolly, S. R., M. A. MacNeil, M. J. Caley, N. Knowlton, E. Cripps, M. Hisano, L. M. Thibaut, B. D. Bhattacharya, L. Benedetti-Cecchi, R. E. Brainard, and others. 2014. Commonness and rarity in the marine biosphere. Proceedings of the National Academy of Sciences:201406664.

Coyle, J. R., A. H. Hurlbert, and E. P. White. 2013. Opposing mechanisms drive richness patterns of core and transient bird species. The American Naturalist 181:E83–E90.

Frank, S. A. 2011. Measurement scale in maximum entropy models of species abundance. Journal of evolutionary biology 24:485–496.

Goudie, A. S. 2013. The human impact on the natural environment: past, present, and future. John Wiley & Sons.

Gray, J., M. Waldichuk, A. Newton, R. Berry, A. Holden, and T. Pearson. 1979. Pollution-induced changes in populations [and discussion]. Philosophical Transactions of the Royal Society B: Biological Sciences 286:545–561.

Grayson, D. K. 2001. The archaeological record of human impacts on animal populations. Journal of World Prehistory 15:1–68.

Harte, J. 2011. Maximum entropy and ecology: a theory of abundance, distribution, and energetics. Oxford University Press.

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

Horne, C. R., A. Hirst, D. Atkinson, and others. 2015. Temperature-size responses match latitudinal-size clines in arthropods, revealing critical differences between aquatic and terrestrial species. Ecology letters 18:327–335.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography (mPB-32). Princeton University Press.

Jackson, J. B., M. X. Kirby, W. H. Berger, K. A. Bjorndal, L. W. Botsford, B. J. Bourque, R. H. Bradbury, R. Cooke, J. Erlandson, J. A. Estes, and others. 2001. Historical overfishing and the recent collapse of coastal ecosystems. science 293:629–637.

Locey, K. J., and E. P. White. 2013. How species richness and total abundance constrain the distribution of abundance. Ecology letters 16:1177–1185.

Magurran, A. E. 2007. Species abundance distributions over time. Ecology Letters 10:347–354.

Magurran, A. E. 2013. Measuring biological diversity. John Wiley & Sons.

Magurran, A. E., and P. A. Henderson. 2003. Explaining the excess of rare species in natural species abundance distributions. Nature 422:714–716.

Matthews, T. J., and R. J. Whittaker. 2014. Fitting and comparing competing models of the species abundance distribution: assessment and prospect. Frontiers of Biogeography 6.

McGill, B. J. 2003. A test of the unified neutral theory of biodiversity. Nature 422:881–885.

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

McGill, B. J., B. A. Maurer, and M. D. Weiser. 2006. Empirical evaluation of neutral theory. Ecology 87:1411–1423.

Morlon, H., E. P. White, R. S. Etienne, J. L. Green, A. Ostling, D. Alonso, B. J. Enquist, F. He, A. Hurlbert, A. E. Magurran, and others. 2009. Taking species abundance distributions beyond individuals. Ecology Letters 12:488–501.

Morris, B. D., and E. P. White. 2013. The ecoData retriever: improving access to existing ecological data. PloS one 8:e65848.

North American Butterfly Association. 2009. NABA butterfly counts: 2009 report. NABA, Morristown, New Jersey, USA.

Pardieck, K. L., D. J. Ziolkowski Jr, and M.-A. Hudson. 2014. North american breeding bird survey dataset 1966 - 2013, version 2013.0. U.S. Geological Survey, Patuxent Wildlife Research Center.

Patil, G., and C. Taillie. 1982. Diversity as a concept and its measurement. Journal of the American Statistical Association 77:548–561.

Phillips, O., and J. S. Miller. 2002. Global patterns of plant diversity: alwyn h. gentry's forest transect data set. Missouri Botanical Garden Press St., Louis, Missouri.

Pielou, E. 1975. Ecological diversity. Wiley, New York.

Pueyo, S. 2006. Diversity: between neutrality and structure. Oikos 112:392–405.

Rosindell, J., S. P. Hubbell, and R. S. Etienne. 2011. The unified neutral theory of biodiversity and biogeography at age ten. Trends in ecology & evolution 26:340–348.

Rosindell, J., S. P. Hubbell, F. He, L. J. Harmon, and R. S. Etienne. 2012. The case for ecological neutral theory. Trends in ecology & evolution 27:203–208.

Society, N. A. 2002. The christmas bird count historical results. National Audobon Society, New York, New York, USA.

Thibault, K. M., S. R. Supp, M. Giffin, E. P. White, and S. M. Ernest. 2011. Species composition and abundance of mammalian communities: ecological archives e092-201. Ecology 92:2316–2316.

Thibault, K. M., E. P. White, and S. M. Ernest. 2004. Temporal dynamics in the structure and composition of a desert rodent community. Ecology 85:2649–2655.

Tittensor, D. P., B. Worm, and R. A. Myers. 2009. CHAPTER tWELVE mACROECOLOGICAL cHANGES iN eXPLOITED mARINE sYSTEMS. Page 310 Marine macroecology. University of Chicago Press.

Trebilco, R., J. K. Baum, A. K. Salomon, and N. K. Dulvy. 2013. Ecosystem ecology: size-based constraints on the pyramids of life. Trends in ecology & evolution 28:423–431.

Ulrich, W., and M. Zalewski. 2006. Abundance and co-occurrence patterns of core and satellite species of ground beetles on small lake islands. Oikos 114:338–348.

USDA Forest Service. 2010. Forest inventory and analysis national core field guide (phase 2 and 3). version 4.0. USDA Forest Service, Forest Inventory; Analysis.

Volkov, I., J. R. Banavar, F. He, S. P. Hubbell, and A. Maritan. 2005. Density dependence explains tree species abundance and diversity in tropical forests. Nature 438:658–661.

Volkov, I., J. R. Banavar, F. He, S. P. Hubbell, and A. Maritan. 2006. Theoretical biology: comparing models of species abundance (reply). Nature 441:E1–E2.

Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan. 2003. Neutral theory and relative species abundance in ecology. Nature 424:1035–1037.

Warwick, R. 1986. A new method for detecting pollution effects on marine macrobenthic communities. Marine biology 92:557–562.

Webb, T. J. 2012. Marine and terrestrial ecology: unifying concepts, revealing differences. Trends in ecology & evolution 27:535–541.

White, E. P., B. J. Enquist, and J. L. Green. 2008. On estimating the exponent of power-law frequency distributions. Ecology 89:905–912.

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.