PeerJ

An empirical evaluation of four variants of a universal species–area relationship

Daniel J. McGlinn, Xiao Xiao and Ethan P. White

Department of Biology and the Ecology Center, Utah State University, Logan, UT, USA

ABSTRACT

The Maximum Entropy Theory of Ecology (METE) predicts a universal species–area relationship (SAR) that can be fully characterized using only the total abundance (N) and species richness (S) at a single spatial scale. This theory has shown promise for characterizing scale dependence in the SAR. However, there are currently four different approaches to applying METE to predict the SAR and it is unclear which approach should be used due to a lack of empirical comparison. Specifically, METE can be applied recursively or non-recursively and can use either a theoretical or observed species-abundance distribution (SAD). We compared the four different combinations of approaches using empirical data from 16 datasets containing over 1000 species and 300,000 individual trees and herbs. In general, METE accurately downscaled the SAR ($R^2 > 0.94$), but the recursive approach consistently under-predicted richness. METE's accuracy did not depend strongly on using the observed or predicted SAD. This suggests that the best approach to scaling diversity using METE is to use a combination of non-recursive scaling and the theoretical abundance distribution, which allows predictions to be made across a broad range of spatial scales with only knowledge of the species richness and total abundance at a single scale.

INTRODUCTION

The species–area relationship (SAR) is a fundamental ecological pattern that characterizes the change in species richness as a function of spatial scale. The SAR plays a central role in predicting the diversity of unsampled areas (*Palmer, 1990*), reserve design (*Whittaker et al., 2005*), and estimating extinction rates due to habitat loss (*Brooks, da Fonseca & Rodrigues, 2004*). Applications involving the SAR depend strongly on the form of the relationship (*Guilhaumon et al., 2008*) which is known to change with spatial scale (*Palmer & White, 1994*; *McGlinn & Hurlbert, 2012*). Despite the scale-dependence of the SAR, a simple non-scale dependent model (the power-law) is still the most commonly used model for the SAR (*Tjørve, 2003*).

The Maximum Entropy Theory of Ecology (METE) is a unified theory that shows promise for characterizing a variety of macroecological patterns including the species-abundance distribution, a suite of relationships between body-size and abundance, and a number of spatial patterns including the species–area relationship (*Harte et al., 2008*;

Harte, Smith & Storch, 2009; *Harte, 2011*). METE adopts the inferential machinery of Maximum Entropy (MaxEnt; *Jaynes, 2003*) to solve for the most likely state of an ecological community (*Haegeman & Loreau, 2008*; *Haegeman & Loreau, 2009*) using only information on the total number of species, the total number of individuals, the total metabolic rate of all the individuals, and the area of the community.

METE predicts that all SARs follow a universal relationship between the exponent of a power-law characterizing the SAR at a particular scale and the ratio of richness and community abundance. The exponent of the SAR is scale dependent, decreasing with increasing spatial scale. Empirical evaluation of the theory suggests that METE is a promising model for the SAR (*Harte et al., 2008*; *Harte, Smith & Storch, 2009*); however, there are currently four different approaches to applying METE to predict the SAR and it is unclear which approach should be used due to a lack of empirical comparison.

There are two distinct versions of METE, recursive (where richness at different scales is obtained by consecutively halving or doubling of area; *Harte, Smith & Storch, 2009*) and non-recursive (where richness at different scales is obtained directly; *Harte et al., 2008*). These two versions predict somewhat different SARs. It is not clear *a priori* which of these versions of METE should be more accurate, and it has been suggested that the best approach should be based on empirical comparisons (*Harte, 2011*). In addition, the METE-SAR is derived using the species-abundance distribution (SAD). The SAD can either be predicted from N and S or the empirical distribution can be used. The most general use of METE for predicting diversity across scales relies on the use of the theoretical abundance distribution, but there have been no comparisons of METE-SAR predictions using theoretical and empirical SADs.

To understand which approach to METE is best for characterizing diversity across scales we conducted a thorough empirical comparison of the four different variants of the METE-SAR prediction: (1) recursive with predicted SAD, (2) recursive with observed SAD, (3) non-recursive with predicted SAD, and (4) non-recursive with observed SAD. Using 16 spatially explicit plant datasets we compared the form and accuracy of the predicted SAR across the four variations of METE at a wide-range of spatial scales and across a diverse set of plant communities with over 1000 species and 300,000 individual trees and herbs.

METHODS

Downscaling richness

The METE approach to predicting the SAR is a two-step application of the maximum entropy formalism (MaxEnt): (1) MaxEnt is first used to predict the SAD which represents the probability that a species has abundance n_0 in a community of area A_0 with S_0 species and N_0 individuals, $\Phi(n_0|N_0, S_0, A_0)$, and (2) MaxEnt is then used to predict the intra-specific, spatial-abundance distribution which represents the probability that n out of n_0 individuals of a species are located in a random quadrat of area A drawn from a total area A_0, $\Pi(n|A, n_0, A_0)$. The Π distribution is spatially implicit and does not contain information on the spatial correlation between cells. If the observed species abundance
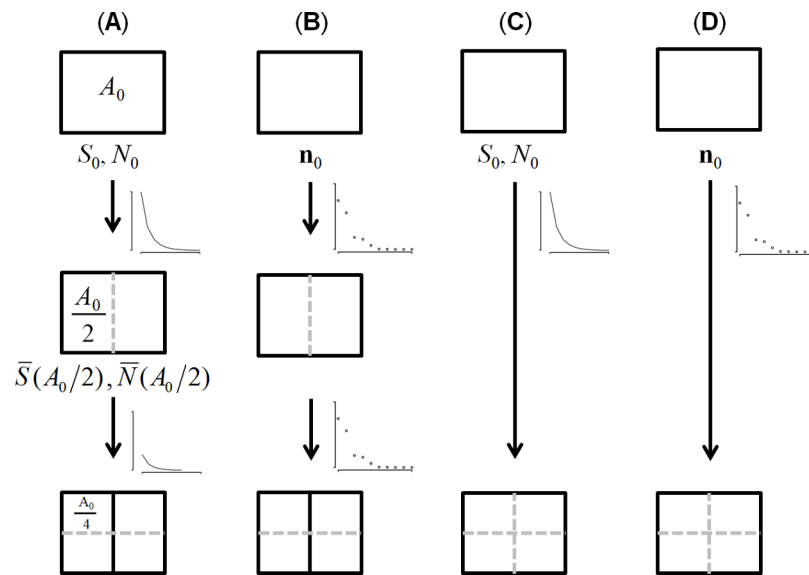
Figure 1 An illustration of the process for downscaling species richness from A_0 to $A_0/4$ across the four variants of METE. The recursive approach uses either (A) the theoretical SAD (inset curve) or (B) the observed SAD (inset points) to predict richness at $A_0/2$ and then the process is repeated to generate a prediction at $A_0/4$. In contrast, the non-recursive approach uses either (C) the theoretical SAD or (D) the observed SAD to predict richness at $A_0/4$ directly. S_0 is the total number of species, N_0 is the total number of individuals, and $\mathbf{n}_0$ is the vector of species abundances at the community scale (A_0) abundance.

distribution is used instead of the METE distribution, then only the Π distribution is solved for using MaxEnt.

There are no adjustable parameters in METE, and the solutions to Π and Φ only depend on the empirical constraints and possible system configurations (*Haegeman & Loreau, 2008*; *Haegeman & Loreau, 2009*; *Haegeman & Etienne, 2010*). If the observed SAD is not used then constraints on the average number of individuals per species (N_0/S_0) and on the upper bound of the number of individuals N_0 can be used to yield a truncated log-series abundance distribution (Fig. 1, *Harte et al., 2008*; *Harte, 2011*). To predict Π, METE places constraints on the average number of individuals per unit area ($n_0 A/A_0$) and on the upper bound of the total abundance of a species n_0. Although METE requires total metabolic rate to derive its predictions, this variable can be ignored when solving for the METE SAD or SAR (*Harte et al., 2008*; *Harte, Smith & Storch, 2009*; *Harte, 2011*).

There are two ways to downscale (and upscale) the Π distribution (Fig. 1). There is a recursive approach (Figs. 1A and 1B) in which the constraints at A_0 are used to solve for Π at $A_0/2$, which provides new constraints (i.e., predicted $S_{A_0/2}$ and $N_{A_0/2}$) for solving Π at $A_0/4$ and so on until richness is computed at every bisection of the total area $A_0/2^i$ where i is a positive integer (*Harte, Smith & Storch, 2009*; *Harte, 2011*, p 159). The recursive approach continually updates its prior information as it downscales richness. Alternatively we can use a non-recursive approach (Figs. 1C and 1D) in which we solve for Π at any area based only on the constraints at A_0 (*Harte et al., 2008*; *Harte, 2011*, p 243). The recursive approach may be more accurate because it continually upgrades its prior information or

less accurate due to error propagation thus only empirical comparisons can determine which approach is best used for prediction (*Harte, 2011*, p 160).

Harte (2011) provides the derivations for the Π and Φ distributions, so here we will only highlight the most relevant equations for differentiating the four METE variants. The MaxEnt solution to maximizing entropy for Π is:

$$\Pi(n|A, n_0, A_0) = \frac{1}{Z_\Pi} e^{-\lambda_\Pi n} \tag{1}$$

where λ_Π is the Lagrange multiplier and Z_Π is the partition function (*Harte, 2011*, Eq. 7.48). The partition function ensures normalization and it is defined as:

$$Z_\Pi = \sum_{n=0}^{n_0} e^{-\lambda_\Pi n} = \frac{1 - e^{-\lambda_\Pi(n_0+1)}}{1 - e^{-\lambda_\Pi}}. \tag{2}$$

The Lagrange multiplier can be solved for by defining Π in terms of its constraints which yields:

$$\bar{n} = \frac{n_0 A}{A_0} = \frac{\sum_{n=0}^{n_0} n x^n}{\sum_{n=0}^{n_0} x^n} \left(\frac{x}{1-x} - \frac{(n_0+1) \cdot x^{n_0+1}}{1 - x^{n_0+1}} \right) \tag{3}$$

where $x = e^{-\lambda_\Pi}$ to simplify notation. Although the METE prediction for Π can be solved numerically for any area, it is only known analytically for a special case in which the area A is half the total area A_0 (*Harte, 2011*, Eq. 7.51):

$$\Pi\left(n \middle| \frac{A_0}{2}, n_0, A_0\right) = \frac{1}{1+n_0}. \tag{4}$$

Equation (4) shows that METE predicts that all possible arrangements of n_0 individuals are equally likely across two equal area quadrats. The flat distribution characterized by Eq. (4) is identical to the prediction offered by the Hypothesis of Equal Allocation Probability (HEAP) model and therefore the recursive application of Eq. (4) to downscale Π generates the same set of Π distributions as the HEAP model (*Harte et al., 2005*; *Harte, 2007*; *Harte, 2011*):

$$\Pi\left(n \middle| \frac{A_0}{2^i}, n_0, A_0\right) = \sum_{q=n}^{n_0} \frac{\Pi\left(q \middle| \frac{A_0}{2^{i-1}}, n_0, A_0\right)}{(q+1)}, \quad i = 1, 2, 3, \ldots . \tag{5}$$

Thus for a given bisection of the total area (i.e., $A = A_0/2^i$) we can either use the recursive approach (Eq. (5)) or the non-recursive approach (Eq. (1)) to compute the Π distribution.

Expected richness is simply the sum of the individual probabilities of species occupancy. Table 1 gives the expressions for expected richness at $A_0/2$ given the four possible combinations of the choice of the downscaling approach and the choice of SAD to use. The equations in Table 1 will also hold for finer spatial scales except for the recursive, theoretical SAD approach which requires downscaling the SAD as well (*Harte, 2011*, Eq. 7.63).

Table 1 The four variants of METE formulated for expected species richness at $A_0/2$, given either the recursive or non-recursive method of downscaling and either the theoretical or observed SAD. These equations also hold for finer spatial scales except for the recursive, METE-SAD approach which requires downscaling the SAD as well (see *Harte, 2011*, Eq. 7.63). $\mathbf{n_0}$ is the vector of empirical abundances and $n_{0,j}$ is the abundance of the jth species at the community scale (A_0).

Species-abundance distribution	Method of downscaling	
	Recursive downscaling (Eq. (5))	**Non-recursive downscaling (Eq. (1))**
METE-SAD (truncated log-series distribution)	$S_0 \sum\limits_{n_0=1}^{N_0} \left[\left(1 - \sum\limits_{q=0}^{n_0} \dfrac{\Pi\left(q\mid\frac{A_0}{2},n_0,A_0\right)}{(q+1)}\right) \cdot \Phi(n_0\mid N_0,S_0) \right]$	$S_0 \sum\limits_{n_0=1}^{N_0} \left[\left(1 - \dfrac{e^{-\lambda_\Pi \cdot 0}}{Z_\Pi\left(n=0\mid\frac{A_0}{2},n_0,A_0\right)}\right) \cdot \Phi(n_0\mid N_0,S_0) \right]$
Observed-SAD ($\mathbf{n_0}$)	$\sum\limits_{j=1}^{S_0} \left(1 - \sum\limits_{q=0}^{n_{0,j}} \dfrac{\Pi\left(q\mid\frac{A_0}{2},n_{0,j},A_0\right)}{(q+1)}\right)$	$\sum\limits_{j=1}^{S_0} \left(1 - \dfrac{e^{-\lambda_\Pi \cdot 0}}{Z_\Pi\left(n=0\mid\frac{A_0}{2},n_{0,j},A_0\right)}\right)$

Empirical comparison

Testing METE's predictions requires spatially explicit, contiguous data from a single trophic level. We carried out an extensive search for data that met these requirements. This search resulted in a database of 16 communities (Table 2; see Table S1 for additional details). All of the datasets are terrestrial, woody plant communities with the exception of the serpentine grassland which is herbaceous. In the woody plant surveys, the minimum diameter at breast height (i.e., 1.4 m from the ground) that a tree must be to be included in the census was 10 mm with the exception of the Cross Timbers and Oosting sites where the minimum diameter was 25 and 20 mm respectively. Where datasets contained time-series information we selected a single census year from each dataset to analyze. *Harte (2011)* suggested that MaxEnt models will perform best when a single process such as the presence of a past disturbance is not dominating the system and rather a multitude of different interacting processes are operating. With this in mind, we attempted to choose the survey years that were the longest amount of time from known stand-scale disturbances (e.g., hurricane events).

For each dataset we constructed fully-nested, spatially-explicit SARs (Type IIA, *Scheiner, 2003*). Recursive METE only makes predictions for bisected areas so we restricted our datasets to areas that were square or rectangular with the dimensional ratio of 2:1. Due to the irregular shape of the Sherman and Cocoli sites we defined two separate 200 × 100 m subplots within each site (Fig. S1). We then calculated the results for each of the two subplots and reported the average.

To assess the accuracy of METE's predictions for the SAD and the four downscaling algorithms of the SAR, we computed the coefficient of determination about the one-to-one line: $R^2 = 1 - \sum_i (obs_i - pred_i)^2 / \sum_i (obs_i - \overline{obs_i})^2$ where obs_i and $pred_i$ are the ith log-transformed observed and METE-predicted values (abundance for the SAD, richness for the SAR) respectively. Log transformed richness was used to minimize the influence of the few very large richness values and because relative deviations are of greater interest in evaluating SARs than absolute differences. We used the python package METE (*White et al., 2013*), as well as a suite of project specific R and python scripts for our analysis. All R and Python code used to generate these analyses is archived in the supplemental materials and also available on GitHub (http://github.com/weecology/mete-spatial).

RESULTS

The four versions of METE all produced reasonable estimates of downscaled richness (Fig. 1). The R^2 values ranged from 0.944 for the recursive, observed-SAD model up to 0.997 for the non-recursive, observed-SAD (Fig. 2). Despite the high coefficient of determination, the recursive approach deviated systematically from the empirical data by underpredicting richness (Figs. 2 and 3). This deviation became larger at finer scales (Fig. 2). In contrast, the non-recursive approach showed no systemic deviations. The SAD was well characterized by the METE predictions ($R^2 = 0.95$); however, METE did on average predict slightly more uneven communities (i.e., predicted abundance was too low for rare species and too high for abundant species, Fig. S2). Overall, the

Table 2 Summary of the habitat type and state variables of the vegetation datasets analyzed. The state variables are total area (A_0), total abundance (N_0) and total number of species (S_0).

Site names	Habitat type	Ref.	A_0 (ha)	N_0	S_0
BCI	Tropical forest	a,b,c	50	205096	301
Sherman	Tropical forest	d	2	7622.5	174.5
Cocoli	Tropical forest	d	2	4326	138.5
Luquillo	Tropical forest	e	12.5	32320	124
Bryan	Oak-hickory forest	f,g,h	1.7113	3394	48
Big Oak	Oak-hickory forest	f,g,h	2	5469	40
Oosting	Oak-hickory forest	i	6.5536	8892	39
Rocky	Oak-hickory forest	f,g,h	1.44	3383	37
Bormann	Oak-hickory forest	f,g,h	1.96	3879	30
Wood Bridge	Oak-hickory forest	f,g,h	0.5041	758	19
Bald Mtn.	Oak-hickory forest	f,g,h	0.5	669	17
Landsend	Old field, pine forest	f,g,h	0.845	2139	41
Graveyard	Old field, pine forest	f,g,h	1	2584	36
UCSC	Mixed evergreen forest	j	4.5	5885	31
Serpentine	Serpentine grassland	k	0.0064	37182	24
Cross Timbers	Oak woodland	l	4	7625	7
Ranges			0.0064–50	669–205096	7–301

Notes.

[a] *Condit (1998).*
[b] *Hubbell et al. (1999).*
[c] *Hubbell, Condit & Foster (2005).*
[d] *Condit et al. (2004).*
[e] *Zimmerman et al. (1994).*
[f] *Peet & Christensen (1987).*
[g] *McDonald, Peet & Urban (2002).*
[h] *Xi et al. (2008).*
[i] *Palmer et al. (2007).*
[j] *Gilbert et al. (2010).*
[k] *Green, Harte & Ostling (2003).*
[l] *Arévalo (2013).*

inclusion of the observed SAD did not strongly improve the prediction of the SAR. For the non-recursive approach including the observed SAD improved the overall R^2 from 0.984 to 0.997 (Figs. 3C and 3D), but the accuracy of the recursive model actually decreased with the inclusion of the observed SAD (R^2 from 0.976 to 0.944, Figs. 3A and 3B).

Results were broadly consistent across datasets, with the exception of the serpentine grassland and Cross Timbers oak woodland. The serpentine community displayed a steeper non-saturating SAR in contrast to the other datasets, and was the only dataset where the recursive downscaling approach was more accurate (Fig. 2O). The oak community displayed a sigmoidal SAR, and in contrast to the other study sites the inclusion of the observed SAD for the oak community resulted in a large improvement in the predicted SAR (Fig. 2P).
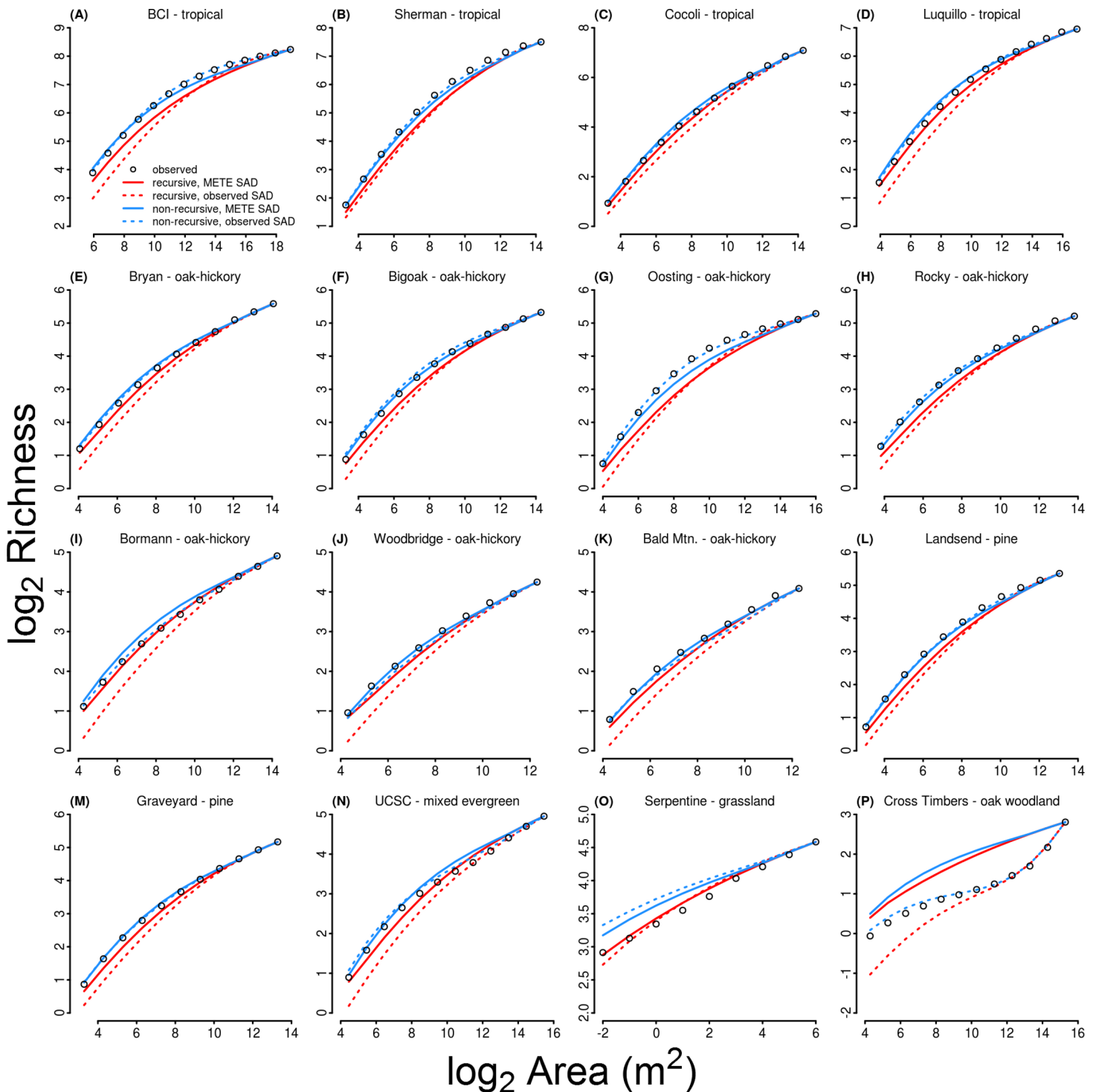
Figure 2 Empirical species–area relationships and the four versions of the METE model across the 16 sites. The habitat type of each site is given above each panel. The empirical averages are the open circles, the recursive approach is the red lines, the non-recursive approach is the blue lines, the curves using the observed SAD are dashed and those using the METE-SAD are solid.
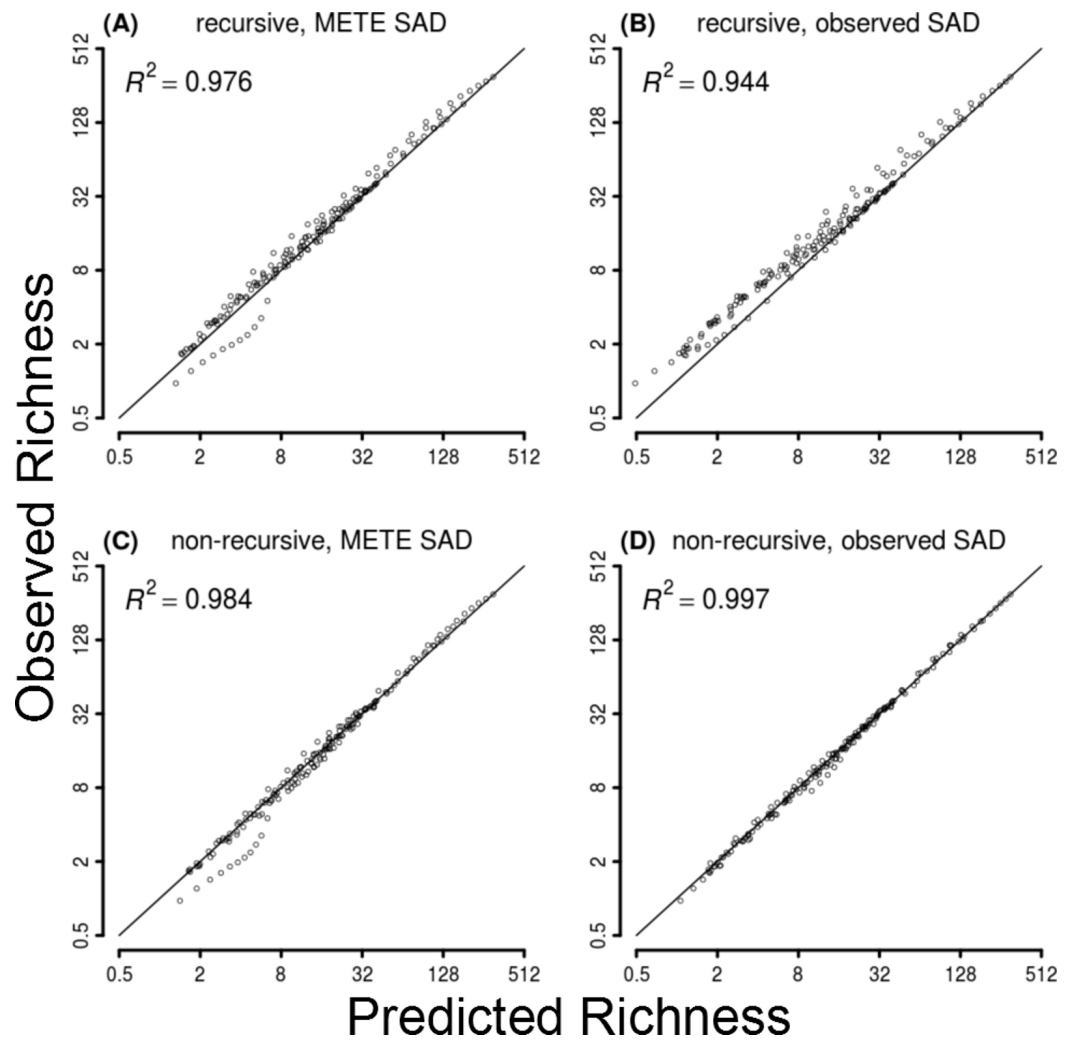
Figure 3 Observed vs predicted richness across datasets and spatial scales for the four METE SAR models. The R^2-value is computed with respect to the one-to-one line (diagonal).

DISCUSSION

All four variations of METE performed well at predicting species richness across scales (all $R^2 > 0.94$); however, some versions performed consistently better than others. The non-recursive approach outperformed the recursive version of METE in all but one dataset (the serpentine grassland). The recursive approach also showed small, but consistent, under-predictions for species richness. This means that the recursive approach predicted stronger intra-specific spatial aggregation than observed in the data. This finding is consistent with Harte's (*2011*) comparisons of the species-level spatial abundance distribution in which the recursive approach predicted greater aggregation than the non-recursive approach. Given that the recursive approach provides a poorer fit to empirical data and can only be applied at particular scales (i.e., $A_0/2, A_0/4, \ldots$), we recommend the use of the non-recursive approach for downscaling the SAR. However,

the recursive approach is currently the only means of providing a METE-based prediction for the distance decay relationship via the hypothesis of equal allocation probabilities approach in *Harte (2007)*, and the universal relationship between S/N and the slope of the SAR is currently only known for the recursive approach (*Harte, Smith & Storch, 2009*).

The SAR predictions were generally robust to using the predicted rather than observed SAD. Including the observed SAD increases the amount of information used to constrain the predictions, but it did not substantially increase the overall accuracy of the SAR predictions. This was primarily because the empirical SAD was well characterized by the METE-SAD, consistent with several other studies (*Harte et al., 2008*; *Harte, 2011*; *White, Thibault & Xiao, 2012*). Models in general, and MaxEnt models in particular, typically match empirical data better as increasing numbers of parameters or constraints are included in the analysis (*Haegeman & Loreau, 2008*; *Roxburgh & Mokany, 2010*; *Harte, 2011*). Therefore the naïve expectation for using the observed SAD is that the accuracy of the prediction should increase. However, this was generally only true for the non-recursive approach. This occurred because rarity and intraspecific aggregation interact in subtle ways to determine the shape of the SAR (*He & Legendre, 2002*; *McGlinn & Palmer, 2009*), and simply fixing one of these pieces of information does not guarantee improved predictive power. While using the observed SAD does improve the R^2 for the non-recursive form of METE, it only does so by ∼1%. Therefore N and S are generally sufficient to accurately downscale richness using METE across a wide range of habitat types. This is important because it should be possible to model geographic patterns of richness and abundance at a single scale to predict the SAD (*White, Thibault & Xiao, 2012*) and then use those modeled values to predict richness across scales.

Although METE yields accurate predictions for the SAR, its current form has limitations with respect to its extent of applicability and its ability to tie in more broadly with species–time and species–time–area relationships (*Rosenzweig, 1995*; *White et al., 2010*). Specifically, METE predictions are thought to be most relevant for single trophic level datasets that are spatially contiguous and relatively environmentally homogenous (*Harte, 2011*), thus constraining the applicability of METE. At the large spatial scales that are often of interest in conservation planning it is likely that a standard application of METE will fail once species ranges do not occupy all of A_0. These are also the scales at which the third phase of the triphasic SAR is expected to occur (*Allen & White, 2003*; *Storch, Keil & Jetz, 2012*), and METE does not predict this accelerating phase. However, *McGill (2010)* suggested that METE's local predictions could be connected with a broad-scale theory to predict a triphasic SAR. Additionally, METE does not currently make predictions through time; however, *Harte (2011)* suggests using Maximum Entropy Production (*Dewar, 2005*). It should be possible to extend METE to predict the species–time–area relationship (*White et al., 2010*) because this pattern, like the SAR, can be modeled in terms of the number of unique individuals sampled per unit area and time (*McGlinn & Palmer, 2009*).

Recently there have been two critiques of the METE spatial predictions. The universality of the relationship between the slope of the recursive METE-SAR and the ratio of N/S was questioned on the basis that the predicted METE-SAR for subsets of a community cannot

be added to yield the community based prediction (*Šizling et al., 2011*; *Šizling, Kunin & Storch, 2013*). However, *Harte et al. (2013)* argue that it is not a flaw of METE or a strong argument against universality.

Additionally, *Haegeman & Etienne (2010)* argued that a multivariate, spatially implicit analog of the univariate Π distribution that is derived using the non-recursive METE approach makes different predictions at different spatial scales (i.e., it is not scale consistent); however, they recognize that a spatially-explicit, scale-consistent version of this distribution may still exist. This critique does not apply to the recursive approach (DJ McGlinn, X Xiao, J Kitzes & EP White, unpublished data), but it may apply in other contexts such as the scaling of the SAD. The lack of scale-consistency in some of METE's predictions suggests that the choice of the anchor scale (A_0) may influence the theory's predictions; however, our results, which spanned a range of anchor scales (0.0064 to 50 ha), did not appear to change systematically with scale. Furthermore, *White, Thibault & Xiao (2012)* demonstrated that the METE-SAD accurately characterized empirical SADs across studies with a wide range of anchor scales. Although METE may not provide a universally applicable model of spatial structure in ecological systems and some of its predictions will depend on the anchor scale, our results as well as others suggest that METE can be used as a practical tool for inferring patterns of diversity and abundance from relatively little information.

We examined the down-scaling of richness; however, many conservation applications are interested in up-scaling richness or predicting diversity at a coarse unsampled scale using information at a fine scale. *Harte, Smith & Storch (2009)* demonstrated that recursive-METE accurately up-scaled tropical tree richness. Currently a formal examination of upscaling using the non-recursive approach is lacking. Thus, future investigations should examine the ability of different variants of METE to upscale richness across a range of spatial scales and ecological systems.

METE represents a useful practical tool for accurately predicting species richness across spatial scales. Among METE's four different approaches to predict the SAR, our analysis demonstrates that the non-recursive approach outperforms the recursive approach, and that using the observed rather than predicted SAD does not substantially improve accuracy. Therefore the METE prediction derived using the non-recursive approach and the predicted SAD will likely be the most useful for future applications involving the SAR.

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests

EPW is an Academic Editor for PeerJ.

Author Contributions

- Daniel J. McGlinn conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper.
- Xiao Xiao and Ethan P. White conceived and designed the experiments, contributed reagents/materials/analysis tools, wrote the paper.

Data Deposition

The following information was supplied regarding the deposition of related data:
https://github.com/weecology/mete-spatial.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.212.

REFERENCES

Allen AP, White EP. 2003. Effects of range size on species–area relationships. *Evolutionary Ecology Research* **5**:493–499.

Arévalo JR. 2013. Spatial analysis and structure of a cross-timber stand in the TallGrass Prairie Preserve (Pawhuska, Oklahoma). *Journal of Forestry Research* **24**:47–52 DOI 10.1007/s11676-013-0324-2.

Brooks T, da Fonseca GAB, Rodrigues ASL. 2004. Species, data, and conservation planning. *Conservation Biology* **18**:1682–1688 DOI 10.1111/j.1523-1739.2004.00457.x.

Condit R. 1998. *Tropical forest census plots.* Berlin, Germany and Georgetown, Texas: Springer-Verlag and R. G. Landes Company.

Condit R, Aguilar S, Hernandez A, Perez R, Lao S, Angehr G, Hubbell SP, Foster RB. 2004. Tropical forest dynamics across a rainfall gradient and the impact of an El Niño dry season. *Journal of Tropical Ecology* **20**:51–72 DOI 10.1017/S0266467403001081.

Dewar RC. 2005. Maximum entropy production and the fluctuation theorem. *Journal of Physics A: Mathematical and General* **38**:L371–L381 DOI 10.1088/0305-4470/38/21/L01.

Gilbert GS, Howard E, Ayala-Orozco B, Bonilla-Moheno M, Cummings J, Langridge S, Parker IM, Pasari J, Schweizer D, Swope S. 2010. Beyond the tropics: forest structure in a temperate forest mapped plot. *Journal of Vegetation Science* **21**:388–405 DOI 10.1111/j.1654-1103.2009.01151.x.

Green JL, Harte J, Ostling A. 2003. Species richness, endemism and abundance patterns: tests of two fractal models in a serpentine grassland. *Ecology Letters* **6**:919–928 DOI 10.1046/j.1461-0248.2003.00519.x.

Guilhaumon F, Gimenez O, Gaston KJ, Mouillot D. 2008. Taxonomic and regional uncertainty in species–area relationships and the identification of richness hotspots. *Proceedings of the National Academy of Sciences of the United States of America* **105**:15458–15463 DOI 10.1073/pnas.0803610105.

Haegeman B, Etienne RS. 2010. Entropy maximization and the spatial distribution of species. *The American Naturalist* **175**:E74–E90 DOI 10.1086/650718.

Haegeman B, Loreau M. 2008. Limitations of entropy maximization in ecology. *Oikos* **117**: 1700–1710 DOI 10.1111/j.1600-0706.2008.16539.x.

Haegeman B, Loreau M. 2009. Trivial and non-trivial applications of entropy maximization in ecology: a reply to Shipley. *Oikos* **118**:1270–1278 DOI 10.1111/j.1600-0706.2009.17560.x.

Harte J. 2007. Toward a mechanistic basis for a unified theory of spatial structure in ecological communities at multiple spatial scales. In: Storch D, Marquet PA, Brown JH, eds. *Scaling biodiversity*. Cambridge: Cambridge University Press, 101–126.

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

Harte J, Conlisk E, Ostling A, Green JL, Smith AB. 2005. A theory of spatial structure in ecological communities at multiple spatial scales. *Ecological Monographs* **75**:179–197 DOI 10.1890/04-1388.

Harte J, Kitzes J, Newman, Rominger AJ. 2013. Taxon categories and the universal species-area relationship: a comment on Šizling et al., "Between geometry and biology: the problem of universality of the species-area relationship". *The American Naturalist* **181**:282–287 DOI 10.1086/668821.

Harte J, Smith AB, Storch D. 2009. Biodiversity scales from plots to biomes with a universal species–area curve. *Ecology Letters* **12**:789–797 DOI 10.1111/j.1461-0248.2009.01328.x.

Harte J, Zillio T, Conlisk E, Smith AB. 2008. Maximum entropy and the state-variable approach to macroecology. *Ecology* **89**:2700–2711 DOI 10.1890/07-1369.1.

He FL, Legendre P. 2002. Species diversity patterns derived from species–area models. *Ecology* **83**:1185–1198.

Hubbell SP, Condit R, Foster RB. 2005. Barro Colorado forest census plot data. *Available at http://ctfs.arnarb.harvard.edu/webatlas/datasets/bci.*

Hubbell SP, Foster RB, O'Brien ST, Harms KE, Condit R, Wechsler B, Wright SJ, de Lao SL. 1999. Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science* **283**:554–557 DOI 10.1126/science.283.5401.554.

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

McDonald RI, Peet RK, Urban DL. 2002. Environmental correlates of oak decline and red maple increase in the North Carolina piedmont. *Castanea* **67**:84–95.

McGill BJ. 2010. Towards a unification of unified theories of biodiversity. *Ecology Letters* **13**:627–642 DOI 10.1111/j.1461-0248.2010.01449.x.

McGlinn DJ, Hurlbert AH. 2012. Scale dependence in species turnover reflects variance in species occupancy. *Ecology* **93**:294–302 DOI 10.1890/11-0229.1.

McGlinn DJ, Palmer MW. 2009. Modeling the sampling effect in the species–time–area relationship. *Ecology* **90**:836–846 DOI 10.1890/08-0377.1.

Palmer MW. 1990. The estimation of species richness by extrapolation. *Ecology* **71**:1195–1198 DOI 10.2307/1937387.

Palmer MW, Peet RK, Reed RA, Xi W, White PS. 2007. A multiscale study of vascular plants in a North Carolina piedmont forest. *Ecology* **88**:2674 DOI 10.1890/07-0796.1.

Palmer MW, White PS. 1994. Scale dependence and the species–area relationship. *American Naturalist* **144**:717–740 DOI 10.1086/285704.

Peet RK, Christensen NL. 1987. Competition and tree death. *BioScience* **37**:586–595 DOI 10.2307/1310669.

Rosenzweig ML. 1995. *Species diversity in space and time.* Cambridge: Cambridge University Press.

Roxburgh SH, Mokany K. 2010. On testing predictions of species relative abundance from maximum entropy optimisation. *Oikos* **119**:583–590 DOI 10.1111/j.1600-0706.2009.17772.x.

Scheiner SM. 2003. Six types of species–area curves. *Global Ecology & Biogeography* **12**:441–447 DOI 10.1046/j.1466-822X.2003.00061.x.

Šizling AL, Kunin WE, Šizlingová E, Reif J, Storch D. 2011. Between geometry and biology: the problem of universality of the species–area relationship. *The American Naturalist* **178**:602–611 DOI 10.1086/662176.

Šizling AL, Kunin, Storch D. 2013. Taxon-and-area invariances, maximum entropy, and the species–area relationship: a reply to Harte et al. *The American Naturalist* **181**:288–290 DOI 10.1086/668822.

Storch D, Keil P, Jetz W. 2012. Universal species–area and endemics-area relationships at continental scales. *Nature* **488**:78–81 DOI 10.1038/nature11226.

Tjørve E. 2003. Shapes and functions of species–area curves: a review of possible models. *Journal of Biogeography* **30**:827–835 DOI 10.1046/j.1365-2699.2003.00877.x.

White EP, Ernest SKM, Adler PB, Hurlbert AH, Lyons SK. 2010. Integrating spatial and temporal approaches to understanding species richness. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**:3633–3643 DOI 10.1098/rstb.2010.0280.

White EP, McGlinn DJ, Xiao X, Supp S, Thibault K. 2013. METE - software for analyzing Harte et al.'s maximum entropy theory of ecology. *Available at Figshare.* DOI 10.6084/m9.figshare.815905.

White EP, Thibault KM, Xiao X. 2012. Characterizing species abundance distributions across taxa and ecosystems using a simple maximum entropy model. *Ecology* **93**:1772–1778 DOI 10.1890/11-2177.1.

Whittaker RJ, Araujo MB, Paul J, Ladle RJ, Watson JEM, Willis KJ. 2005. Conservation biogeography: assessment and prospect. *Diversity and Distributions* **11**:3–23 DOI 10.1111/j.1366-9516.2005.00143.x.

Xi W, Peet RK, Decoster JK, Urban DL. 2008. Tree damage risk factors associated with large, infrequent wind disturbances of Carolina forests. *Forestry* **81**:317–334 DOI 10.1093/forestry/cpn020.

Zimmerman JK, Everham EM III, Waide RB, Lodge DJ, Taylor CM, Brokaw NVL. 1994. Responses of tree species to hurricane winds in subtropical wet forest in Puerto Rico: implications for tropical tree life histories. *Journal of Ecology* **82**:911–922 DOI 10.2307/2261454.