CONCEPTS & SYNTHESIS
EMPHASIZING NEW IDEAS TO STIMULATE RESEARCH IN ECOLOGY

Ecology, 89(10), 2008, pp. 2700–2711

MAXIMUM ENTROPY AND THE STATE-VARIABLE APPROACH TO MACROECOLOGY

J. Harte,[1,3] T. Zillio,[2] E. Conlisk,[1] and A. B. Smith[1]

[1]*Energy and Resources Group, University of California, 310 Barrows Hall, Berkeley, California 94720 USA*
[2]*Department of Renewable Resources, University of Alberta, 114th Street 89 Avenue, Edmonton, Alberta T6G 2H1 Canada*

Abstract. The biodiversity scaling metrics widely studied in macroecology include the species–area relationship (SAR), the scale-dependent species–abundance distribution (SAD), the distribution of masses or metabolic energies of individuals within and across species, the abundance–energy or abundance–mass relationship across species, and the species-level occupancy distributions across space. We propose a theoretical framework for predicting the scaling forms of these and other metrics based on the state-variable concept and an analytical method derived from information theory. In statistical physics, a method of inference based on information entropy results in a complete macro-scale description of classical thermodynamic systems in terms of the state variables volume, temperature, and number of molecules. In analogy, we take the state variables of an ecosystem to be its total area, the total number of species within any specified taxonomic group in that area, the total number of individuals across those species, and the summed metabolic energy rate for all those individuals. In terms solely of ratios of those state variables, and without invoking any specific ecological mechanisms, we show that realistic functional forms for the macroecological metrics listed above are inferred based on information entropy. The Fisher log series SAD emerges naturally from the theory. The SAR is predicted to have negative curvature on a log–log plot, but as the ratio of the number of species to the number of individuals decreases, the SAR becomes better and better approximated by a power law, with the predicted slope z in the range of 0.14–0.20. Using the 3/4 power mass–metabolism scaling relation to relate energy requirements and measured body sizes, the Damuth scaling rule relating mass and abundance is also predicted by the theory. We argue that the predicted forms of the macroecological metrics are in reasonable agreement with the patterns observed from plant census data across habitats and spatial scales. While this is encouraging, given the absence of adjustable fitting parameters in the theory, we further argue that even small discrepancies between data and predictions can help identify ecological mechanisms that influence macroecological patterns.

Key words: biodiversity; endemics–area relationship; energy distribution; macroecology; maximum entropy; metabolic theory; spatial pattern; spatial scaling; species–abundance distribution; species–area relationship; state variables.

Introduction

Macroecology, the study of patterns in the distribution of abundances and body sizes of organisms across species and across spatial and temporal scales, is a central focus of ecology. Knowledge of macroecological metrics, such as the species–area relationship (SAR) relating average species richness found in a census cell to the area of that cell, the endemics-area relationship (EAR) describing the average number of species unique to a census cell of specified area, the species–abundance distribution (SAD) describing the fraction of species with specified abundance, the abundance–energy relationship (AER) relating a species metabolic energy requirement to its abundance, and the scaling properties of the spatial distributions of the individuals within species, is critical to designing efficient censusing strategies, predicting extinction rates under habitat loss, estimating species diversity and abundances from incomplete census data, and deciphering the processes that most influence ecosystems (Preston 1948, May 1975, Brown 1995, Rosenzweig 1995, Gaston and Blackburn 2000, Kinzig and Harte 2000, Hubbell 2001).

Manuscript received 23 August 2007; revised 21 January 2008; accepted 21 February 2008. Corresponding Editor: B. E. Kendall.
[3] E-mail: jharte@berkeley.edu

Both statistical and mechanistic models have been used to try to understand and predict the form of macroecological metrics. Various statistical models have been suggested to describe subsets of the metrics listed above, with the focus generally on the clustering properties of occurrences of individuals at multiple spatial scales. The Coleman (1981) random-placement model (RPM) assumes individuals within species are placed randomly on a landscape, yielding a binomial distribution for occurrences at any scale. The spatial distributions of individuals within species generally exhibit more aggregation than predicted by the RPM (Kunin 1998, Gaston and Blackburn 2000, He and Gaston 2000, Plotkin et al. 2000, He and Legendre 2002, Green et al. 2003, Harte et al. 2005). Models of spatial distribution that incorporate aggregation have also been explored, including a negative binomial distribution (NBD) model (He and Gaston 2000), a fractal model (Kunin 1998), a Poisson cluster model (Plotkin et al. 2000), and the HEAP model (Harte et al. 2005). None of these models of the spatial structure of occurrences provide a comprehensive framework for macroecological synthesis because to predict community metrics such as the SAR, they need to be augmented with empirical knowledge of the SAD. Moreover, the ad hoc nature of the statistical assumptions that underlie these models is unsatisfying, thus motivating interest in other approaches.

Spatially explicit models that assume knowledge of some subset of processes, usually drawn from birth, death, dispersal, speciation, migration, extinction, and abiotic niche differentiation, have been advanced to attempt to predict and understand the origins of patterns in the abundance and distribution of species across space and time (see reviews by Hubbell 2001, Chave 2004, Leibold et al. 2004). Based on specific mechanistic assumptions, numerous parameterized macroecological models have been proposed for predicting the functional forms of subsets of the macroecological metrics (Sugihara 1980, Hubbell 2001). While considerable attention is given to whether one or another function that is derived from a mechanistic model best fits empirical abundance distributions or species–area curves, there has been less attention to the goal of developing a unifying framework that can adequately, and possibly without arbitrarily fitted parameters, describe the central tendencies observed for the entire range of macroecological metrics of interest, across taxa, habitats, and spatial scales.

The approach proposed here for achieving that goal builds on the concept of state variables. These are properties of a system that comprise the conditions whose specification is necessary to implement theory, but whose determination lies outside the theory. Thermodynamics provides an analogy. If the thermodynamic system is a container filled with η moles of a gas species, the state variables can be taken to be the volume of the container, the temperature of the gas, and the number of moles, and from them statistical properties, such as the Boltzmann energy distribution, and the thermodynamic properties of the system, such as the law $PV = \eta RT$ can be derived in terms of R, the fundamental gas constant. In the Gibbs-Jaynes formulation of thermodynamics, the machinery needed to make these derivations is the maximum entropy method (MaxEnt), which we also exploit here.

The basic idea of MaxEnt, due to Jaynes (1957, 1965, 1979, 1982) and building on the work of Shannon (1948), is that information entropy provides a logical basis for inferring the forms of probability distributions. To understand how the concept of entropy, a measure of disorder, applies to information, we consider the concept of "least biased" in the context of a probability distribution. Suppose we seek to infer the shape of a probability distribution $p(n)$, and that all of our knowledge about this distribution is that it satisfies a set of constraints that arise from our prior information about the system. Then the least-biased inference of the shape of the distribution is that which is as smooth and flat as possible subject to the known constraints. Any other distribution would be assuming information that is not captured by the prior knowledge embodied in the constraints; any other distribution would contain structures such as peaks and dips that implied additional constraints that we cannot claim to know.

This prior information might take the form of explicit knowledge of the mean, or the variance, or some combination of moments of $p(n)$. In ecology, $p(n)$ might, for example, be the probability that a species has n of its individuals in a 1 hectare census-plot randomly chosen within a 100-ha preserve. A constraint on this probability distribution might then be that the mean of the distribution is 0.01 times the known total abundance of the species in the preserve.

Two questions then arise: what is the rigorous mathematical definition of a function that, within its constraints, is "as smooth and flat as possible" and what is the procedure for finding that function? In an intuitive sense, those peaks and dips referred to above reduce the entropy of the distribution in that such mathematical structures represent a form of order. Thus Shannon and Jaynes were led to look to entropy for a more rigorous way of quantifying the notion of "least biased." It was formally proven (Jaynes 1957) that the shape of the least biased distribution can be inferred by maximizing information entropy, $S_{\mathrm{I}} = -\Sigma_n p(n)\log[p(n)]$, subject to the known constraints. Moreover, Jaynes proved that S_I was the unique function possessing the necessary properties of an entropy measure whose constrained maximum generated the least biased inference. The mathematical procedure used to find the function that yields the constrained maximum of information entropy is the method of Lagrange multipliers (see Appendix A). The MaxEnt method has been applied to many areas of science (Lorenz 2003, Dewar 2005, Elith et al. 2006, Phillips et al. 2006, Schneidman et al. 2006, Shipley et al. 2006, Laurent and Cai 2007).

We emphasize that making predictions using MaxEnt is an application of mathematical logic. It is a rigorously

proven mathematical procedure for inferring the most likely probability distribution, if our knowledge about that distribution can be incorporated as a set of constraints on the distribution. Physical or biological processes are incorporated in the MaxEnt procedure to the extent that the constraints arise from physical or biological principles or phenomena. Can MaxEnt fail to make accurate predictions? Yes, because we either may be assuming incorrect constraints or we may be ignoring other constraints. If the constraints are correctly, but possibly incompletely described, then the MaxEnt procedure provides a way of determining when additional information, for example about some dominant processes leading to additional constraints, needs to be imposed if we are to make accurate predictions. We give an example of that in *Discussion*.

Application of the MaxEnt concept to any complex system requires a decision as to the fundamental entities (molecules in the thermodynamic example) and specification of a set of state variables. In the theory presented here, we will derive probability distributions defined on two kinds of entities, "individuals" and "species," but we do not need to narrowly define either. Thus we take the "species" to mean any well-defined set of groups of "individuals." The groups could be the taxonomic species (as will be assumed here), but they might also be genera, or even trait groups. And "individuals" can be defined in the usual sense of individuals in community ecology (as will be assumed here), but other choices can be readily accommodated. The only criteria for choosing the entities in a MaxEnt application is that they are unambiguously defined in a manner that allows specification of the numerical values of the state variables and thus the constraints on the probability distributions. In our application of MaxEnt, we shall see that the constraints arise from knowledge of the ratios of four state variables (area, species richness, total number of individuals, and total metabolic rate).

The theory we present here is "null" in the sense that, like the RPM (Coleman 1981), it makes falsifiable predictions and assumes no explicit mechanisms. Compared to the RPM, our theory predicts a much wider range of phenomena, including not just the spatial distributions of the individuals within each species, but also the SAD, the SAR, the EAR, and the AER. As with any null model or theory, its value derives in part from the nature of its failures. Systematic discrepancies between data and prediction can point the way to identification of important mechanisms that should not be ignored. Indeed, we present this theory with the expectation that further tests, with a wider range of taxa, habitats and spatial scales, will assist in identifying mechanisms that are essential to the understanding of macroecology.

Formulation of the Theory

Here we define the state variables, the macroecological probability distributions, and the constraints on those distributions that are then used, in the MaxEnt framework, to derive predictions for the major metrics of macroecology.

Consider an ecosystem of area A_0 and within it a group of organisms such as the trees on a 50-ha plot. We assume prior specification of four state variables: total area, A_0, of the ecosystem; total number of individual trees, N_0, within A_0; summed metabolic energy rate of all the trees, E_0, within A_0; and total number of tree species, S_0, within A_0. We show that these four parameters (A_0, N_0, E_0, and S_0), coupled with the MaxEnt method, are sufficient to allow us to infer numerous macroecological metrics, including the species–abundance distribution $\Phi(n)$ describing the fraction of species with n individuals, the energy consumption probability density $\psi(\varepsilon)$ describing the fraction of the N_0 individuals with energy demand between ε and $\varepsilon + d\varepsilon$, the energy–abundance relationship describing the relationship between the abundance of a species and the average energy requirement of the individuals within that species, the species–area relationship $S(A)$ describing the average number of species in cells of area A within A_0, the endemics–area relationship $E(A)$ describing the average number of species unique to a cell of area A, and the species-level spatial abundance distributions $P_A^{(j)}(n)$ defined as the probability of finding n individuals of species j in a randomly selected habitat cell of area A within A_0.

To accomplish this, we apply MaxEnt to two probability distributions from which all the macroecological metrics listed above can be derived. The first is a joint probability density, $R(n, \varepsilon)$, over the S_0 species in A_0. $R(n, \varepsilon)d\varepsilon$ is the probability that if a species is picked at random from the species list in A_0, then it has abundance n and if an individual is picked at random from that species, then its metabolic requirement is in the interval ε, $\varepsilon + d\varepsilon$. R satisfies the following constraints:

$$\sum_{n=1}^{N_0} \int_0^{E_0} R(n, \varepsilon) d\varepsilon = 1 \tag{1a}$$

$$\sum_{n=1}^{N_0} \int_0^{E_0} nR(n, \varepsilon) d\varepsilon = \frac{N_0}{S_0} \tag{1b}$$

$$\sum_{n=1}^{N_0} \int_0^{E_0} n\varepsilon R(n, \varepsilon) d\varepsilon = \frac{E_0}{S_0}. \tag{1c}$$

The first equation is a normalization condition, the second expresses the constraint on the average number of individuals per species, and the third expresses the constraint on the average total energy requirement of all the individuals within a species. We note that R is a function of two variables, consistent with there being two independent ratios of the three assumed state variables N_0, S_0, E_0, and thus two independent mean values that constrain R. The upper limits of the integrals over ε and the sums over n are simply the maximum possible energy available to an individual and the

maximum number of individuals available to a species. Strictly speaking, the upper limit on n should be $N_0 - S_0 + 1$ because each species that is present has to have at least one individual, but in all our applications $N_0 \gg S_0$ so this correction is quite insignificant.

The second core probability distribution is the spatial abundance distribution $P_A^{(j)}(n)$ defined above. As we shall see, our theory predicts that for fixed scale, A/A_0, the j dependence of $P_A^{(j)}(n)$ depends only on the total abundance, n_0, of species j in A_0, and so henceforth we specify the function as $P_A^{(n_0)}(n)$. In addition to the normalization condition,

$$\sum_{n=0}^{n_0} P_A^{(n_0)}(n) = 1 \tag{2}$$

a constraint on P arises because we know the total number of individuals, n_0, at the largest scale A_0 and thus we have

$$\sum_{n=0}^{n_0} nP_A^{(n_0)}(n) = n_0 \frac{A}{A_0}. \tag{3}$$

The quantity n_0 in $P_A^{(n_0)}(n)$ is a species-level state variable, such that knowledge of it for any particular species provides the constraint (Eq. 3) that determines, under MaxEnt, the shape of that species' spatial distribution. But MaxEnt predicts the explicit function (Eq. 8a) describing the SAD, based on knowledge of the state variables A_0, S_0, N_0, E_0, and so we predict the distribution of n_0 values, if not their actual values, based only on knowledge of those four state variables. Because the community level scaling metrics such as the SAR and the EAR can be determined from the predicted form of the SAD and the $P_A^{(n_0)}(n)$'s by summing appropriate expressions (Eqs. 10 and 11), we do not need to know the measured values of the n_0's to determine these community metrics.. Thus the symbol n_0 plays a dual role in our theory. On the one hand, for purposes of testing the predicted spatial abundance distributions of individual species, its value is prior knowledge. On the other hand, for purposes of predicting community-level metrics it is a summation variable and no prior knowledge of the S_0 values that it takes on is needed because MaxEnt predicts the distribution of those values in the form of the SAD.

From knowledge of R, and by integrating over all possible values of the energy variable, we can derive the species–abundance distribution $\Phi(n)$ and the energy distribution $\Psi(\varepsilon)$:

$$\Phi(n) = \int_0^{E_0} R(n,\varepsilon)d\varepsilon \tag{4a}$$

$$\Psi(\varepsilon) = \frac{S_0}{N_0}\sum_{n=1}^{N_0} nR(n,\varepsilon) \tag{4b}$$

where $\Phi(n)$ is the fraction of species with abundance n, while $\Psi(\varepsilon)d\varepsilon$ is the probability that an individual picked from the entire individuals pool (N_0) has an energy requirement between ε and $\varepsilon + d\varepsilon$.

From the definition of $R(n, \varepsilon)$, another metric, the conditional probability,

$$\theta(\varepsilon|n) = R(n,\varepsilon)/\Phi(n) \tag{4c}$$

can be formed. $\theta(\varepsilon\,|\,n)d\varepsilon$ is the probability that if an individual is selected from a species with known abundance n, then its energy requirement is in the interval ε, $\varepsilon + d\varepsilon$.

Moreover, the SAR can be derived from knowledge of $P_A^{(n_0)}(n)$ and $\Phi(n)$. In particular, the expected number of species on a patch of area A is the total number of species in A_0 times the sum over abundances, n_0, of the product of (the probability a species, with abundance n_0 in A_0, is present on the patch) $\times$ (the fraction of species with abundance n_0). Noting that the probability of presence is 1 minus the probability of absence, we have

$$S(A) = S_0\sum_{n_0=1}^{N_0}[1 - P_A^{(n_0)}(0)]\Phi(n_0). \tag{5}$$

The testable metrics predicted by the theory thus include $\Phi(n)$, $\Psi(\varepsilon)$, $\theta(\varepsilon\,|\,n)$, $P_A^{(n_0)}(n)$, and $S(A)$. Moreover, although Φ, Ψ, and θ lack an explicit spatial scale index, they are predicted at any spatial scale finer than A_0 because the values of the state variables that constrain R, namely N_0 and S_0, are readily computed at finer scale, $N_A = (A/A_0)N_0$, and $S(A)$ is obtained from Eq. 5.

Predictions of the Theory

From the constraint equations (1a–1c), MaxEnt results in

$$R(n,\varepsilon) = \frac{e^{-\lambda_1 n}e^{-\lambda_2 n\varepsilon}}{\sum_{n'=1}^{N_0}\int_0^{E_0} e^{-\lambda_1 n'}e^{-\lambda_2 n'\varepsilon'}d\varepsilon'} \tag{6}$$

where the λ_i are called Lagrange multipliers and are completely specified by the values of N_0, E_0, and S_0. The mathematical details are given in Appendices A and B.

While no approximations have to be made to obtain numerical solutions, it is straightforward to show that if terms of magnitude e^{-S_0} are dropped compared to terms of order 1, then the Lagrange multipliers are determined to a very good approximation by (see Appendix B for this and other derivations):

$$\lambda_2 = \frac{S_0}{E_0} \tag{7a}$$

$$\lambda_1 \ln(\lambda_1^{-1}) = \frac{S_0}{N_0}. \tag{7b}$$

Using Eqs. 4a and 7, it follows from straightforward integration of Eq. 6 over the energy variable that the abundance distribution is given by

$$\Phi(n) = \frac{1}{\ln(\lambda_1^{-1})}\frac{e^{-\lambda_1 n}}{n}. \tag{8a}$$

Similarly, using Eqs. 4b and 7, from summation of Eq. 6 over abundance n, it follows that:

TABLE 1. Properties of the six sites used to evaluate the maximum entropy method (MaxEnt) predictions.

Property	Location					
	BCI	Cocoli	Sherman	Luquillo	San Emilio	Serpentine
A_0	50 ha	2 ha	1.96 ha	10.24 ha	9.68 ha	64 m^2
S_0	302	237	193	122	138	24
N_0	214 034	4381	6822	46 919	12 851	37 182
Variance in observed rank–log(abundance) explained†	0.98	0.97	0.99	0.99	0.99	0.99
Slope of observed vs. predicted log(abundance)	0.91	0.97	0.98	0.99	0.94	1.02
Variance in observed log–log SAR explained‡	0.99	n.a.	n.a.	n.a.	0.99	0.99
Slope of observed vs. predicted log(S) from SAR	0.99	n.a.	n.a.	n.a.	1.00	1.19

Notes: A_0, S_0, and N_0 are state variables. The last four rows list measures of the reasonableness of theoretical predictions for the species–area relationship (SAR) and the scale-dependent species–abundance distribution (SAD). For descriptions of the censuses, see Condit et al. (2004) on Barro Colorado Island (BCI), Cocoli, and Sherman; Zimmerman et al. (1994) and Thompson et al. (2002, 2004) on Luquillo; Enquist et al. (1999) on San Emilio; and Green et al. (2003) on Serpentine. The 782 species grids used for Fig. 3a, b are the "in-sample" grids from Conlisk et al. (2007). Only species with total abundance $n_0 > 2$ were used here. We did not have access to the data required to do the analyses needed to fill in the six entries marked "n.a."

† Fraction of variance in observed rank–log(abundance) distribution explained by theory.

‡ Fraction of variance in observed log–log species–area relationship (SAR) explained by theory.

$$\Psi(\varepsilon) = \lambda_1\lambda_2 \frac{e^{-(\lambda_1+\lambda_2\varepsilon)}}{[1-e^{-(\lambda_1+\lambda_2\varepsilon)}]^2}. \quad (8b)$$

This expression for $\Phi(n)$ is the Fisher log-series distribution (Fisher 1943), which we see emerges as a direct consequence of MaxEnt and in which the value of the parameter λ_1 in the exponent is determined by S_0 and N_0. Indeed, the Lagrange multiplier λ_1 is just Fisher's α divided by N_0. For the data sets we have examined, $\lambda_1 n \ll 1$ for all but the most abundant species, and thus for most species the species–abundance distribution (Eq. 8a) is well approximated by the geometric series distribution $\Phi(n) \propto 1/n$.

The conditional probability distribution $\theta(\varepsilon \mid n)$ is then derived from Eq. 4c:

$$\theta(\varepsilon|n) \cong \lambda_2 n e^{-\lambda_2 n\varepsilon} \quad (8c)$$

where the equivalence is increasingly exact as $\exp(-S_0) \to 0$.

The form of the species-level spatial abundance distributions, $P_A^{(n_0)}(n)$, derived from MaxEnt is

$$P_A^{(n_0)}(n) = \frac{e^{-\lambda_P n}}{\sum_{k=0}^{n_0} e^{-\lambda_P k}} \quad (9)$$

where the Lagrange multiplier, λ_P is determined by Eq. B.5 as a function of the abundance of the species, n_0, and spatial scale, A/A_0.

The species–area relationship (SAR) is uniquely specified from Eqs. 5, 8a, and 9. In particular,

$$S(A) = S_0 \sum_{n_0=1}^{N_0} [1 - P_A^{(n_0)}(0)]\Phi(n_0)$$

$$= S_0 \sum_{n_0=1}^{N_0} \left[1 - \frac{1}{\sum_{n=0}^{n_0} e^{-\lambda_P n}}\right] \frac{1}{\ln(\lambda_1^{-1})} \frac{e^{-\lambda_1 n_0}}{n_0}. \quad (10)$$

The summand of this equation is the product of the probability that a species has abundance n_0 times the probability that a species with abundance n_0 will be present in a cell of area A. Here the area-dependence enters the right-hand side entirely via the area-dependence of λ_P. We emphasize that this expression for $S(A)$ has no adjustable parameters; in our theory, the species–area relationship is completely determined by the measured values of S_0 and N_0.

Finally, the endemics–area relationship (EAR) is given by

$$E(A) = S_0 \sum_{n_0=1}^{N_0} P_A^{(n_0)}(n_0)\Phi(n_0) = S_0 \sum_{n_0=1}^{N_0} \frac{e^{-\lambda_P n_0}}{\sum_{n=1}^{n_0} e^{-\lambda_P n}} \Phi(n_0). \quad (11)$$

Here the summand is the product of the probability that a species has abundance n_0 times the probability that such a species has all of its n_0 individuals in a cell of area A. We note that for the SAR and EAR predictions to be valid, both the predicted SAD and predicted values of the $P_A^{(n_0)}(n)$ at $n = 0$ and n_0, at all tested spatial scales, would generally have to be valid.

ARE THE PREDICTIONS REASONABLE?

Our main purpose here is to compare the MaxEnt predictions for the various macroecological metrics with observed data from a variety of spatially explicit vegetation data sets. Table 1 summarizes the six data sets that we examined. It is not our primary intent here to focus on comparisons of the MaxEnt predictions with the predictions from other theories and models, although we do present likelihood comparisons for the SAR predictions. One reason we do not dwell on such measures is that objective methods for comparing theories that differ in scope (number of different metrics that they predict) are not known to us. Secondly, we do not claim that there is no other theory or model that

makes a better fitting prediction for any particular metric than does ours.

Species–abundance distribution

Consider, first, the predicted Fisher log series form for the SAD (Eqs. 7b and 8a). Empirical support for a Fisher log series species–abundance distribution has been demonstrated for many taxa/habitat combinations (Hubbell 2001, Pueyo 2006), and in some instances it has been shown (Green et al. 2003, Loehle 2005, Martin and Goldenfeld 2006) that the limiting case of the geometric series distribution provides an excellent fit to data. We note that this limit corresponds to the Lagrange multiplier $\lambda_1 \rightarrow 0$, or from decreasing S_0/N_0. A geometric distribution, $\Phi(n) = c/n$, results in a rank-abundance graph that is a straight line when plotted as log(abundance) vs. rank, with a negative slope if low rank corresponds to high abundance; the exponential term in the log series distribution bends that straight line upward at low rank. In Fig. 1a–c we examine the SAD, expressed as a rank-abundance relationship, for three vegetation communities with very different S_0/N_0 values of 0.000645 (serpentine grassland), 0.0026 (Luquillo 10.24-ha tropical forest), and 0.0541 (Cocoli 2-ha wet tropical forest). As shown in Fig. 1, the importance of the exponential term in the predicted SAD, reflected in the magnitude of the upturning deviation from a straight line, does indeed increase empirically as S_0/N_0 increases. Moreover, the comparisons of observation and prediction in the figure suggest that our predicted SADs are in reasonable agreement with data. Table 1 presents the critical parameters, and the R^2 values and slopes of graphs of predicted-vs.-observed abundances in the vegetation data sets that we have examined.

Some authors have argued that lognormal distributions or power-law abundance distributions with exponent different from 1 provide superior fits to some data sets, but we have not seen compelling evidence that the data are sufficient to discriminate amongst such functions given sampling uncertainties. This point has been made convincingly by Loehle (2006), who demonstrated that repeated sampling, with random error, from a log series or a geometric distribution can generate rank-abundance curves that resemble the lognormal as the noise term increases, strengthening the conclusion that distinguishing these distributions is difficult. Others have argued that data comparisons using rank-abundance distributions are not as reliable as is the use of binning procedures to directly test probability distributions. However, a substantial literature (Preston 1948, Bulmer 1974, Williamson and Gaston 2005, Gray et al. 2006, Loehle 2006) has shown that binning procedures introduce arbitrariness and that if data are generated from a Fisher log-series distribution and then the binned pdf is plotted, the result can appear either peaked or monotonic decreasing, thereby causing a "true" log series to be confused with a lognormal.

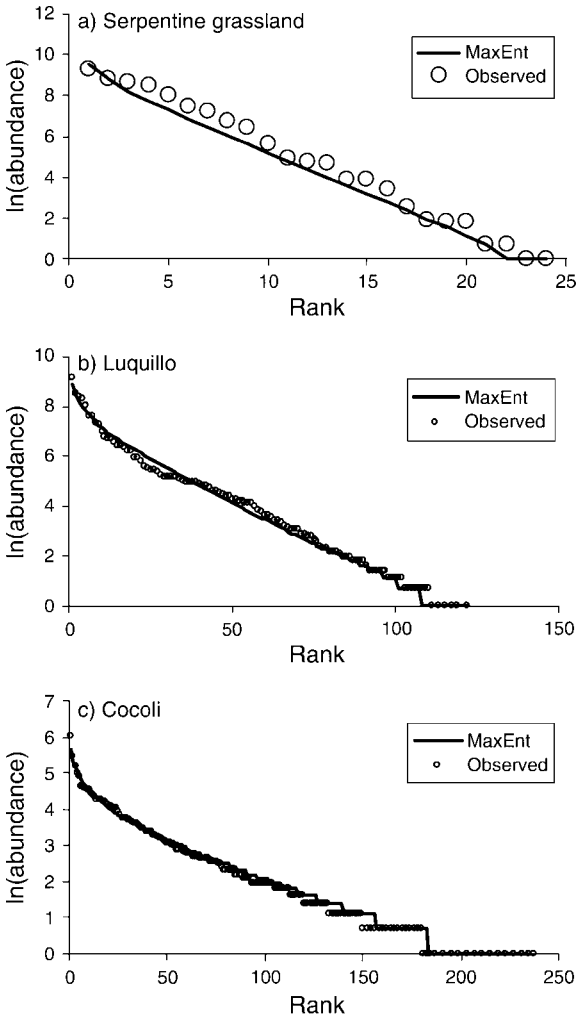

FIG. 1. Observed species–abundance distributions and maximum entropy method (MaxEnt) predictions (Eqs. 7b and 8a), plotted as rank vs. ln(abundance) distributions, for vegetation data at three census sites: (a) Serpentine, (b) Luquillo, and (c) Cocoli. See Table 1 for data sources and relevant state-variable values.

We also note that the Fisher log series is considered appropriate for the SAD of many communities (May 1975, Magurran 2005 and references therein), and is also the curve obtained from neutral dynamics (Hubbell 2001, Volkov et al. 2003), which is a robust result even in the presence of small contribution from non-neutral dynamic (Zillio and Condit 2007). Obviously, other curves have been proposed for the species abundance of communities, including the lognormal, the zero-sum multinomial (Hubbell 2001) and many others. We conclude that the SAD predicted by MaxEnt is a reasonable candidate for a null distribution.

Species-level spatial abundance distribution

The species-level spatial abundance distributions, $P_A^{(n_0)}(n)$, as determined by Eqs. 9 and B.5 are predicted

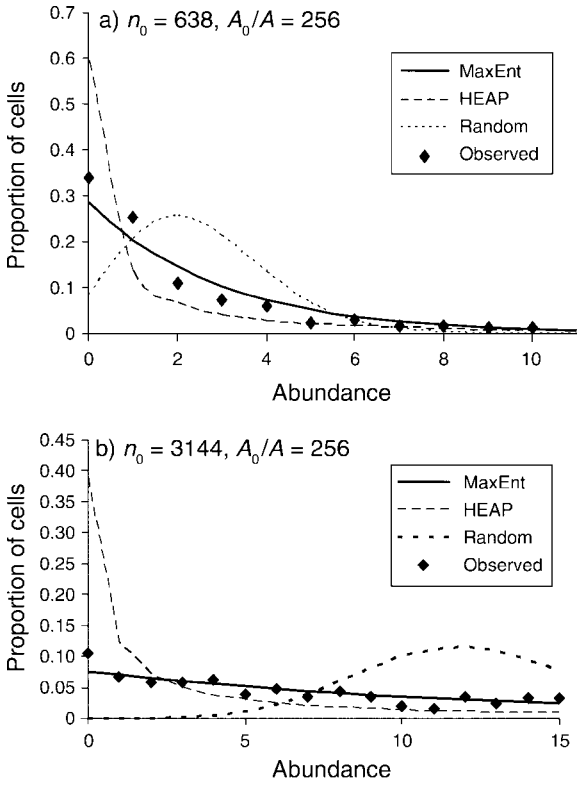

FIG. 2. Observed species-level spatial abundance distributions and predictions from MaxEnt (Eq. 9, Appendix B: Eq. B.5), HEAP (Harte et al. 2005), and the random placement model (Coleman 1981). Representative distributions are shown for two species in the Barro Colorado Island, Panama (BCI) plot at a spatial scale $A = A_0/256$ (where A_0 is a state variable [ecosystem area] and total abundance n_0 varies): (a) $n_0 = 638$, (b) $n_0 = 3144$. The x-axis is the number of individuals in a cell, and the y-axis is the fraction of cells with that abundance.

by our theory to be monotonically, and in particular exponentially, decreasing functions of n for all scales $A < A_0/2$, exponentially increasing functions of n for all scales $A > A_0/2$, and constant (independent of n) at scale $A = A_0/2$. The prediction of a uniform distribution at $A = A_0/2$ is identical to the hypothesis of equal allocation probabilities (HEAP) model prediction at that scale (Harte et al. 2005). At finer scales, and particularly for the higher abundance species, our MaxEnt prediction for $P_A^{(n_0)}(n)$ deviates from the HEAP prediction and from the random placement model prediction (RPM), with MaxEnt predicting fewer cells with "absences," or in other words a smaller value of $P_A^{(n_0)}(0)$, than in the HEAP model (Harte et al. 2005) and a larger value than in the RPM. We note that both HEAP and MaxEnt predict that $P_A^{(n_0)}(n)$ is a monotonically decreasing function of n for $A < A_0/2$, whereas the RPM predicts monotonicity only if $n_0 < A_0/A$.

Previously, we showed (Harte et al. 2005) that the HEAP model systematically over-estimated aggregation at smaller scales for nearly all the more-abundant plant species (typically those ranked in the top abundance quartile) in both tropical forest and temperate grassland plots. Related to this, for these same species it over-predicted the number of cells in which the species is absent. The MaxEnt prediction for $P_A^{(n_0)}(n)$, which is intermediate in its predicted aggregation between HEAP and the RPM, estimates aggregation reasonably well. To illustrate these trends, two representative examples are shown in Fig. 2, but to more systematically evaluate the theory, we note that as the scale ratio A_0/A increases, the MaxEnt prediction for $P_A^{(n_0)}(n)$ approaches the negative binomial distribution (NBD) model distribution with NBD parameter $k = 1$, where $k = \langle n \rangle^2/(\sigma^2 - \langle n \rangle)$ and $\langle n \rangle$ is the mean number of individuals per cell and σ^2 is the variance in the number of individuals per cell. Whereas our distribution is conditional on n_0, the total abundance in A_0, the NBD distribution that it approaches is unconditional on total abundance, and thus these two distributions have a different interpretation.

Consider a census area A_0 on which a uniform grid has been imposed, with $M \equiv A_0/A$ cells of size A. For a single species with n_0 individuals in A_0, let m denote the fraction of the M cells that are occupied by the species. The expected value $\langle m \rangle$ of m is $1 - P_A^{(n_0)}(0)$. For $M > 3$ and $n_0 > 2$, it follows from Eqs. 9 and B.4 that this expected value is very accurately approximated by $n_0/(M + n_0)$:

$$\langle m \rangle = n_0/(M + n_0). \quad (12)$$

The model predicts that occupancy values will cluster about this expected value. To test the prediction, data points (m, n_0) were gathered for 782 species with $n_0 > 2$ from six plant censuses (see Table 1). Grids were 16×16 ($M = 256$). Fig. 3a plots m against $n_0/(M + n_0)$. The points do cluster along the 1:1 line.

To further test the merits of Eq. 12 against alternative predictions, we can embed the MaxEnt prediction (Eq. 12) within an infinite family of functions, characterized by one adjustable clustering parameter k:

$$\langle m \rangle = 1 - [kM/(kM + n_0)]^k. \quad (13)$$

This equation has a negative binomial interpretation (but see two paragraphs above) with k the clustering parameter. Eq. 12 is a special case of Eq. 13 with $k = 1$ and for $M > 3$ and $n_0 > 2$. MaxEnt predicts $k = 1$. Using the same data for 782 species, the parameter k in Eq. 13 was estimated as the value of k which minimizes the sum of squared differences between m and $1 - [kM/(kM + n_0)]^k$, thus providing a best fit estimate of k. To two decimal places, the estimate was $k = 0.91$, indeed close to $k = 1$, with little difference in mean square error between $k = 0.91$ and $k = 1$. See Fig. 3b. For 8×8 grids instead of 16×16 grids, the estimate of k to two decimal places was $k = 1.01$, almost $k = 1$ exactly.

Species–area relationship

Comparisons of predicted (from Eq. 10) and observed SARs are shown in Fig. 4. The three data sets we

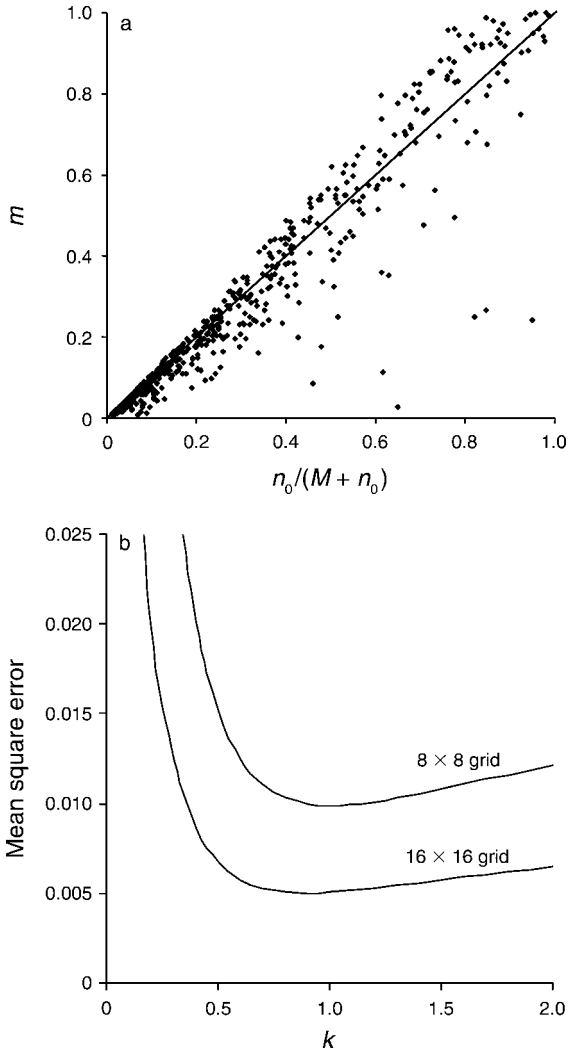

FIG. 3. Cell occupancies. (a) Observed cell occupancy fractions are plotted against MaxEnt predictions (Eq. 12) for 782 plant species (see Table 1); where m is the fraction of cells occupied by each species. The one-to-one line of perfect agreement is also shown. (b) Goodness of fit of the more flexible predictor of Eq. 13, which involves an adjustable parameter, k, with $k = 1$ corresponding to MaxEnt. The vertical axis variable "mean square error" is the sum of squared errors from Eq. 13, divided by 782, as a function of k. Panel (a) shows data only for 16 × 16 grids ($M = 256$ cells), whereas Panel (b) shows results for both 8 × 8 grids ($M = 64$) and 16 × 16 grids ($M = 256$).

examined include one (serpentine grassland) which is well described by a power-law function and two forest plots that show distinct curvature on a log-log plot at all scales. Based just on the S_0 and N_0 values at each of the sites, MaxEnt predicts nearly exact power-law behavior for the serpentine site (with slightly lower slope than is observed) and well-matching curvature for the other two sites. Considering that no parameters are available to adjust in the predicted SARs, which derive solely from knowledge of S_0 and N_0, we suggest that the predictions our theory makes for the SARs in Fig. 4 are quite reasonable.

We note that the serpentine site, with the smallest value of S_0/N_0 of our data sets, and thus with the SAD closest to a pure geometric distribution, is also the site with an SAR closest to a power law. In our MaxEnt theory, this is not a coincidence; we find, numerically, that as S_0/N_0 decreases, and the geometric series SAD becomes a better approximation, the predicted SAR becomes better and better approximated by a power-law. The predicted slope, z, lies in the range of 0.15–0.20, decreasing as S_0/N_0 decreases. To our knowledge, this connection has not been previously established.

We compared the accuracy of the MaxEnt prediction for the SAR with that of several other fitted functional forms that are sometimes advocated: a power law ($S = cA^z$), a Monod function of a power law ($S = cA^z/(1 + bA^z)$ [Lomolino 2001]), and the logarithmic function (S

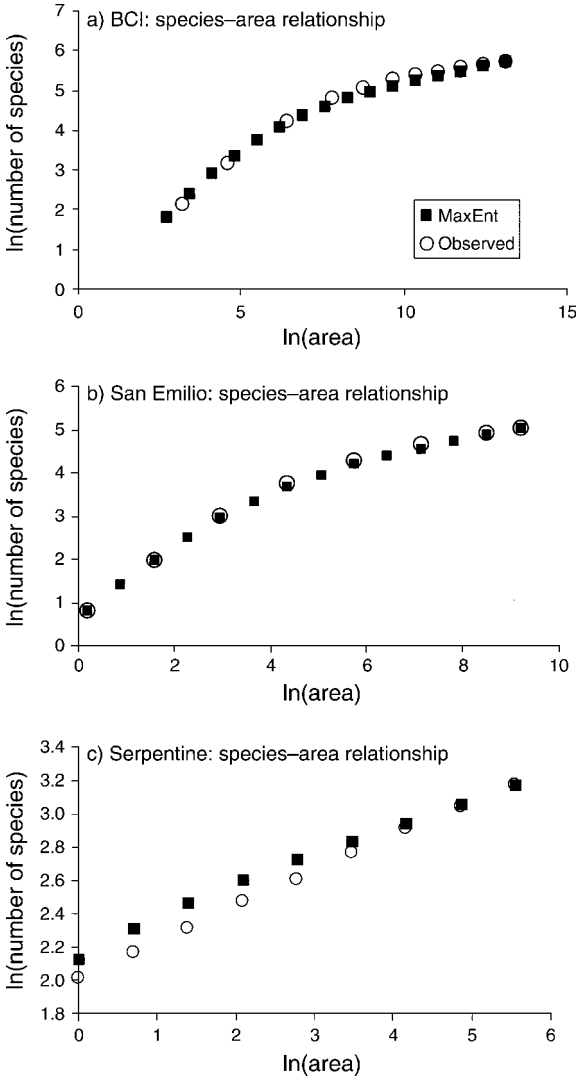

FIG. 4. Observed species–area relationship (SAR) and MaxEnt prediction (Eq. 10) for three vegetation data sets: (a) BCI, (b) San Emilio, and (c) Serpentine. See Table 1 for data sources.

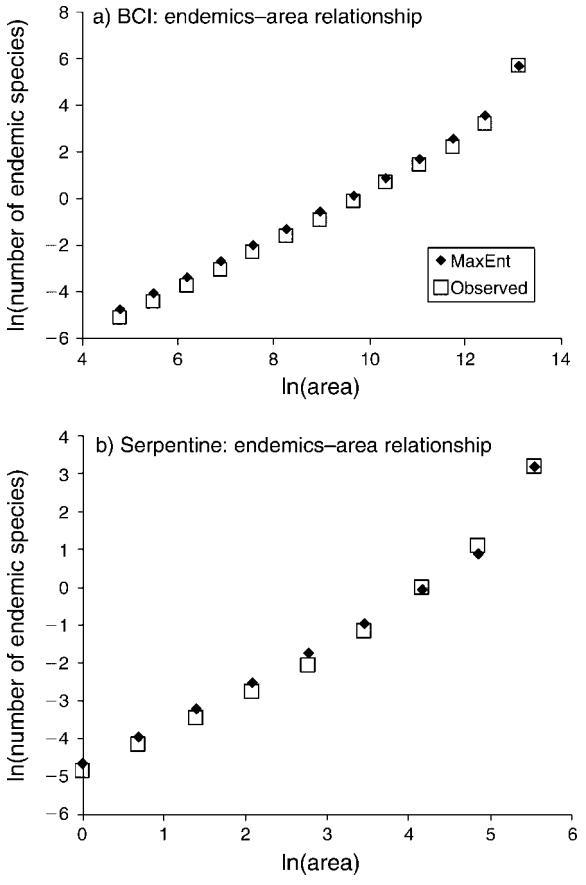

FIG. 5. Observed endemics–area relationship and MaxEnt prediction (Eq. 11) for two vegetation data sets: (a) BCI and (b) Serpentine. See Table 1 for data sources. Negative logarithms arise because the averages over cells of the number of endemic species per cell at the smaller spatial scales are less than 1.

$= c \log(A)$ [Gleason 1922]). To account for different numbers of fitting parameters, we carried out likelihood tests using the Akaike information criterion (AIC) and expressed results as AIC relative weights (Anderson et al. 2000). To make fair comparisons with MaxEnt, for which the state variable S_0 is pre-specified, we evaluated the fractional rather than absolute residual sums of squares, thus giving relatively more weight to the small-scale data. For the San Emilio and BCI data, the MaxEnt relative AIC weights are >0.99, meaning that the MaxEnt model is overwhelmingly preferred by the AIC criterion, but for the serpentine plot the best-fitted power law and Lomolino functions outperform MaxEnt because of the slope discrepancy noted above.

Endemics–area relationship

Comparisons of predicted (from Eq. 11) and observed EARs are shown in Fig. 5. The two data sets we examined are the BCI forest and the serpentine grassland. Based only on the two state variables, S_0 and N_0 as input, the predictions are reasonable at both sites. Considering that no parameters are available to adjust in the predicted SARs, which derive solely from knowledge of S_0 and N_0, we suggest that the predictions our theory makes for the EARs in Fig. 5 are quite reasonable. Interestingly, where as MaxEnt predicted nearly exact power-law behavior for the serpentine site SAR and distinct curvature on a log-log plot for the BCI SAR, if we ignore the data point at area $= A_0$, MaxEnt correctly predicts a more nearly power-law behavior for the BCI EAR and greater curvature for the serpentine site EAR.

Energy and energy–abundance distributions

For reasons explained in the following section, our theory is formulated with total metabolic energy requirement, rather than total biomass, as a state variable, leading to predictions for the distribution of energy requirements across all the individuals, $\Psi(\varepsilon)$, or across individuals within a species with known abundance, $\theta(\varepsilon \mid n)$. To evaluate these predictions in more detail, we need information about metabolic energies of individuals.

Typically forest data sets contain information about the diameters at breast height (dbh's) of individuals, which is two steps removed from actual metabolic energy data. First, we would need to convert available dbh values to individual masses using allometric relationships, and then we would need to convert mass to metabolism using a metabolic scaling rule. The central tendency of the mass–metabolism relationship is well described by the 3/4-power scaling rule of metabolic theory (West et al. 1997, Brown et al. 2004) for most taxa that have been examined, although there is ample scatter in vegetation data around that prediction (Enquist et al. 1999). Different allometric scaling rules have been proposed to relate tree mass to dbh. An 8/3 power scaling rule is often assumed in tests of metabolic theory (Enquist et al. 1999), so that metabolic rate $\sim$ mass$^{3/4}$ $\sim$ (dbh)$^{(8/3)(3/4)}$ $\sim$ basal area. On the other hand, in a critical review of allometric scaling for trees, Muller-Landau et al. (2006) found that the scaling exponents for smaller trees were somewhat smaller than 8/3. Thus direct testing of the energy distribution predictions (Eqs. 8b and 8c) is not possible starting with just data on dbh values.

Nevertheless, we can examine the plausibility of the predicted distribution, $\theta(\varepsilon \mid n)$, of the energy requirements of individuals within species by noting that from Eq. 8c it follows that the total energy requirement of a species with abundance n, which is given by $\int n\varepsilon\theta(\varepsilon \mid n)\, d\varepsilon$, is approximately independent of n. This is the "energy-equivalence" rule that has been tested previously and is claimed to capture central tendencies of data (Enquist et al. 1999), although considerable scatter around the central tendency does exist. Moreover, the energy requirement of an average individual in a species varies as the inverse of its abundance. Thus if the metabolic energy requirement scales as $m^{3/4}$ (West et al. 1997, Banavar et al. 2002, Brown et al. 2004) we obtain the Damuth rule (Damuth 1981) relating average mass and abundance across species: $n \sim m^{-3/4}$. Again, there is considerable scatter, but the general trend of mass–

abundance data is consistent with this rule. Our results are best interpreted as suggesting that more sophisticated testing of the energy predictions, probably with animal data to avoid the uncertainties of allometry, is needed.

Discussion and Conclusion

One could consider either additional or different state variables, and as a consequence derive different probability distributions. We have chosen area A_0 as the first state variable because it is the obvious measure of the physical scale of the system, in analogy with the state variable, volume, in thermodynamics. We have chosen the state variable S_0 because of the central role that species richness plays in ecology and in macroecological metrics, although we could apply the same theoretical methods to the total number of genera, G_0, and then derive abundance distributions over the genera as well as genera–area relationships. We have chosen total abundance and metabolic energy rate as the remaining two state variables because they scale additively, increasing linearly with area in complete nested designs (that is, when the data from all nonoverlapping cells of a specified area are averaged). Moreover, the individual organism and its energy requirements are of fundamental importance in biology and so those two state variables are intuitively reasonable ones to base theory upon. They also share a close analogy with the number of molecules and the total internal energy in thermodynamic systems.

In thermodynamics, temperature and pressure are termed "intensive" variables in the sense that their values in a combined system are each the weighted averages of their values in the component systems. In contrast, volume and energy are termed "extensive" variables because their values are additive when systems are combined. In our theory, A_0, N_0, and E_0 are extensive variables, but S_0 is intermediate between an intensive and an extensive variable; it neither adds linearly nor is it averaged when systems are adjoined and thus has no analogy in thermodynamics. Pursuing this further, in thermodynamics the intensive variable temperature emerges as the inverse of the Lagrange multiplier in the MaxEnt derivation of the Boltzmann energy distribution; in our theory, the inverses of the Lagrange multipliers are neither intensive nor extensive, and we do not know if they can be associated with a generalized ecological "temperature."

If an additional extensive state variable that was allometrically independent of N_0 and E_0 were used to provide further constraints, then that would alter the form of the predicted species abundance distribution. In particular, if an additional state variable, Q_0, is introduced, and q labels the value for an individual in a species with abundance n, leading to the constraint equation $\langle nq \rangle = Q_0/S_0$, then we would define a new joint distribution $R(n, \varepsilon, q)$ and the species abundance distribution would be given by an integral of R over ε and q. It is readily shown that this would yield $\Phi(n) \sim e^{-\lambda n}/n^2$. More generally, with K additive quantities (total E_0, N_0, plus $K - 2$ others), we would get $\Phi(n) \sim e^{-\lambda n}/n^{K-1}$. A plot of ln(abundance) vs. log(rank) would then exhibit a straight-line portion at large rank (small abundance), with slope $-(K - 1)$ if $K > 2$. The power of n in the denominator of the SAD, and thus the slope of its small-n behavior, is unaffected by the magnitude of the Lagrange multipliers; the slope only depends on the number of multipliers associated with additive quantities. Hence, if additional additively conserved quantities exist, they can be identified by such a test; we see no evidence for them in the SAD's for the six vegetation data sets we have examined. We note that if the energy constraint is ignored, then the resulting SAD is a Boltzmann (exponential) distribution, which is readily falsified because it hugely underpredicts the frequency of species with small abundances.

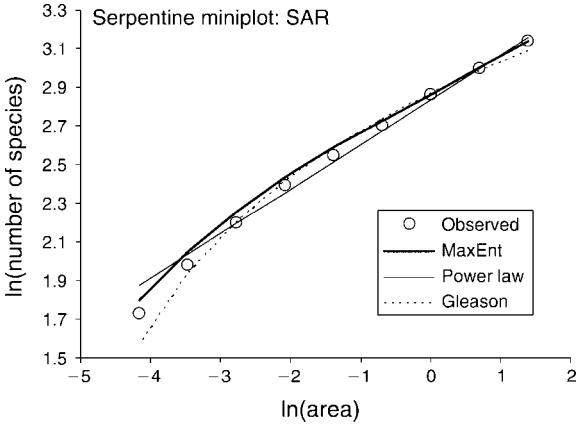

Fig. 6. Observed and predicted species–area relationships for a "miniplot" within the serpentine 64-m^2 plot and located deliberately to avoid the relatively barren quadrant. The functional form for the Gleason function is given in *Are the predictions reasonable? Species–area relationship*.

We re-emphasize that MaxEnt is a mathematically proven method for inferring the most likely probability distribution if our knowledge about that distribution can be described as a set of constraints on the distribution. If it should turn out that the distribution obtained from maximizing $-\Sigma_n\, p(n)\log(p(n))$ subject to the available constraints fails to adequately predict the data, then that indicates that some of the assumed constraints did not really hold and/or that additional constraints do hold but were not included in the constraint equations. The MaxEnt procedure is thus a framework for determining when additional assumptions, for example about dominant mechanisms or environmental heterogeneities leading to additional constraints, need to be imposed. In our application, knowledge of these mechanisms would supplement our knowledge of the constraints that follow by definition from the ratios of the values of the macrovariables, S_0, N_0, and E_0.

To pursue this possibility, we noted that one quadrant of the serpentine plot was largely rocky outcrop, with proportionally few individuals and species. The question

then arises: if one examines more homogeneous subplots within A_0, will the SAR slope discrepancy disappear? In fact it does, as shown in Fig. 6. The MaxEnt prediction for the SAR on a "miniplot" located within the larger serpentine plot but not including the barren quadrant matches the data quite well, performing better than the fitted power-law, the Lomolino function (not shown in figure), or the Gleason function. The relative likelihood weights, AIC corrected, are: MaxEnt (0.94), power law (0.02), Lomolino (0.005), and Gleason (0.035). Alternatively, and more in the spirit of constrained entropy maximization, one can constrain the geometry of A_0 to be the L-shaped plot that excludes the nearly bare quadrant, and again MaxEnt predicts the SAR accurately. We suggest, then, that large-scale heterogeneity can provide an additional constraint in MaxEnt applications in macroecology.

An analogy is found in the thermodynamics of ideal gases, where MaxEnt results in the equation of state for the most probable configuration compatible with the state variables P, V, and T. There, too, when a deviation is observed it means that explicit mechanisms need to be considered (e.g., van der Waals forces).

It is possible to imagine other ways of applying MaxEnt in ecology, for example starting with alternatives to our Eqs. 1–5. In fact during the review process, we became aware of two other manuscripts (Banavar and Maritan 2007, Dewar and Porte 2007). Dewar and Porte, in particular, use a constraint on resource consumption that is analogous to the one on energy that we use here, but their approach requires input knowledge of the full distribution of resource consumption rates of the species, while in our approach we predict the distribution. In both of those papers, and in another MaxEnt application, (Pueyo et al. 2007) the log series SAD is found. Pueyo et al. use prior probabilities to derive that result, whereas the other two papers make use of a "maximum relative entropy" framework. To our knowledge our theory, alone, links together energetics, diversity, abundance, and spatial scaling to predict relatively accurately and without adjustable parameters the SAD (Eq. 8a), the SAR (Eq. 10), the EAR (Eq. 11), and the spatial distributions of individuals within each of the species (Eq. 9). Moreover, our theory predicts the new, relatively unexplored energy distribution metrics given by Eqs. 8b and 8c.

Our state-variable approach to macroecology, combined with MaxEnt, provides a useful null theory of macroecology in that it yields reasonable parameter-free predictions for the central tendencies in empirical patterns in the distribution and abundance of species and individuals for the data sets that we have examined. The theory predicts an SAR that approaches power-law behavior, with slope $z \approx 0.15$ as the ratio of the total number of individuals to the total number of species in the ecosystem increases. The theory also provides a unified theoretical justification for the energy-equivalence rule, the Damuth rule, and the Fisher log-series species–abundance distribution.

Future studies of the applicability of the ideas presented here to animal, microbial, and additional vegetation data sets will illuminate the degree to which the maximum entropy principle provides a useful framework for a general unified description of the diverse metrics and observed patterns in macroecology. Its failures will spotlight the situations in which consideration of additional constraints, perhaps arising from explicit mechanisms ignored here, will be required to understand the central tendencies in macroecological patterns.

Acknowledgments

We thank the U.S. National Science Foundation and the Miller Foundation at UC–Berkeley for support. We are grateful to Richard Condit, Brian Enquist, Jill Thompson, Jessica Green, and Danielle Svelha for their willingness to share data; to Michael Bloxham and Fred Cummings for useful conversations; and to Susan Carey for assistance in data analysis. Two anonymous reviewers provided helpful suggestions for improving the manuscript. For access to data, we also thank the Luquillo Long Term Ecological Research Program; the Institute for Tropical Ecosystem Studies, University of Puerto Rico; the International Institute for Tropical Forestry (U.S. Department of Agriculture); the University of California McLaughlin Natural Reserve; the National Science Foundation; and the Andrew W. Mellon Foundation. This work benefited from participation in the workshop "Towards a Unified Theory of Biodiversity" conducted at the National Center for Ecological Analysis and Synthesis, a Center funded by NSF (Grant #DEB-0553768), the University of California, Santa Barbara, and the State of California.

Literature Cited

Anderson, D. R., K. P. Burnham, and W. L. Thompson. 2000. Null hypothesis testing: problems, prevalence, and an alternative. Journal of Wildlife Management 64:912–923.

Banavar, J. R., J. Damuth, A. Maritan, and A. Rinaldo. 2002. Supply-demand balance and metabolic scaling. Proceedings of the National Academy of Sciences (USA) 99:10506–10509.

Banavar, J., and A. Maritan. 2007. The maximum relative entropy principle. ⟨http://arxiv.org/abs/cond-mat/0703622⟩

Brown, J. H. 1995. Macroecology. University of Chicago Press, Chicago, Illinois, USA.

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.

Bulmer, M. 1974. On fitting the Poisson lognormal distribution to species–abundance data. Biometrics 30:101–110.

Chave, J. 2004. Neutral theory and community ecology. Ecology Letters 7:241–253.

Coleman, B. 1981. On random placement and species–area relations. Journal of Mathematical Biosciences 54:191–215.

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

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

Dewar, R. 2005. Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. Journal of Physics A: Mathematical and General 38:L371–L381.

Dewar, R., and A. Porte. 2007. Statistical mechanics explains macroecological patterns. ⟨http://arxiv.org/abs/q-bio/0703061⟩

Elith, J., et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29: 129–151.

Enquist, B. J., G. B. West, W. L. Charnov, and J. H. Brown. 1999. Allometric scaling of production and life history variation in vascular plants. Nature 401:907–911.
Fisher, R. A. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. Journal of Animal Ecology 12:42–58.
Gaston, K. J., and T. M. Blackburn. 2000. Pattern and process in macroecology. Blackwell Publishers, London, UK.
Gleason, H. A. 1922. On the relation between species and area. Ecology 3:158–162.
Gray, J., A. Bjorgesaeter, and K. Ugland. 2006. On plotting species abundance distribution. Journal of Animal Ecology 75:752–756.
Green, J. L., J. Harte, and A. Ostling. 2003. Species richness, endemism, and abundance patterns: tests of two fractal models in a serpentine grassland. Ecology Letters 6:919–928.
Harte, J., E. Conlisk, A. Ostling, J. Green, and A. Smith. 2005. A theory of spatial structure in ecological communities at multiple spatial scales. Ecological Monographs 75:179–197.
He, F., and K. J. Gaston. 2000. Estimating species abundance from occurrence. American Naturalist 156:553–559.
He, F., and P. Legendre. 2002. Species diversity patterns derived from species area models. Ecology 83:1185–1198.
Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey, USA.
Jaynes, E. 1957. Information theory and statistical mechanics. Physical Review 106:620–630.
Jaynes, E. 1965. Gibbs vs. Boltzmann entropies. American Journal of Physics 33:391–398.
Jaynes, E. 1979. Where do we stand on maximum entropy? Pages 15–118 *in* R. Levine and M. Tribus, editors. The maximum entropy formalism. MIT Press, Cambridge, Massachusetts, USA.
Jaynes, E. 1982. On the rationale of maximum entropy methods. Proceedings of the IEEE 70:939–952.
Kinzig, A., and J. Harte. 2000. Implications of endemics–area relationships for estimates species extinctions. Ecology, 81: 3305–3311.
Kunin, W. E. 1998. Extrapolating species abundances across spatial scales. Science 281:1513–1515.
Laurent, R., and X. Cai. 2007. A maximum entropy method for combining AOGCMs for regional intra-year climate change assessment. Climatic Change 82:411–435.
Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, R. D. Holt, J. B. Shurin, R. Law, D. Tilman, M. Loreau, and A. Gonzalez. 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7:601–613.
Loehle, C. 2006. Species abundance distributions result from body size-energetics relationships. Ecology 87:2221–2225.
Lomolino, M. V. 2001. The species–area relationship: new challenges for an old pattern. Progress in Physical Geography 25:1–21.
Lorenz, R. 2003. Full steam ahead—probably. Science 299:837–838.
Magurran, A. E. 2005. Species abundance distributions: pattern or process? Functional Ecology 19:177–181.
Martin, H., and N. Goldenfeld. 2006. On the origin and robustness of power-law species area relationships in ecology. Proceedings of the National Academy of Sciences (USA) 103: 10310–10315.
May, R. M. 1975. Patterns of species abundance and diversity. Pages 81–120 *in* M. L. Cody and J. M. Diamond, editors. Ecology and evolution of communities. Belknap Press, Cambridge, Massachusetts, USA.
Muller-Landau, H.C., et al. 2006. Testing metabolic ecology theory for allometric scaling of tree size, growth, and mortality in tropical forests. Ecology Letters 9:575–588.
Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231–251.
Plotkin, J. B., M. D. Potts, N. Leslie, N. Manokaran, J. LaFrankie, and P. S. Ashton. 2000. Species–area curves, spatial aggregation, and habitat specialization in tropical forests. Journal of Theoretical Biology 207:81–99.
Preston, F. 1948. The commonness and rarity of species. Ecology 29:254–283.
Pueyo, S. 2006. Diversity: between neutrality and structure. Oikos 112:392–405.
Pueyo, S., F. He, and T. Zillio. 2007. The maximum entropy formalism and the idiosyncratic theory of biodiversity. Ecology Letters 10(11):1017–1028. [doi: 10.1111/j.14610248.2007.010096.x]
Rosenzweig, M. L. 1995. Species diversity in time and space. Cambridge University Press, Cambridge, Massachusetts, USA.
Schneidman, E., M. J. Berry, R. Segev, and W. W. Bialek. 2006. Weak pairwise correlations imply strongly correlated network states in a neural population. Nature 440:1007–1012.
Shannon, C. 1948. A mathematical theory of communication. Bell System Technical Journal 27:379–423.
Shipley, B., D. Ville, and E. Garnier. 2006. From plant traits to plant communities: a statistical mechanistic approach to biodiversity. Science 314:812–814.
Sugihara, G. 1980. Minimum community structure: an explanation of species abundance patterns. American Naturalist 116:770–787.
Thompson, J., N. Brokaw, J. K. Zimmerman, R. B. Waide, E. M. Everham, D. J. Lodge, C. M. Taylor, D. Garcia-Montiel, and M. Fluet. 2002. Land use history, environment, and tree composition in a tropical forest. Ecological Applications 12:1344–1363.
Thompson, J., N. Brokaw, J. K. Zimmerman, R. B. Waide, E. M. Everham, and D. A. Schaefer. 2004. The Luquillo forest dynamics plot. Pages 540–550 *in* E. Losos and E. Leigh, editors. Tropical forest diversity and dynamism: findings from a large-scale plot network. University of Chicago Press, Chicago, Illinois, USA.
Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan. 2003. Neutral theory and relative species abundance in ecology. Nature 424:1035–1037.
West, G. B., J. H. Brown, and B. J. Enquist. 1997. A general model for the origin of allometric scaling laws in biology. Science 276:122–126.
Williamson, M., and K. Gaston. 2005. The lognormal distribution is not an appropriate null hypothesis for the species–abundance distribution. Journal of Animal Ecology 74:409–422.
Zillio, T., and R. Condit. 2007. The impact of neutrality, niche differentiation and species input on diversity and abundance distributions. Oikos 116(6):931–940.
Zimmerman, J. K., E. M. Everham, R. B. Waide, D. J. Lodge, C. M. Taylor, and V. L. Brokaw. 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.

APPENDIX A

The MaxEnt machinery (*Ecological Archives* E089-153-A1).

APPENDIX B

Mathematical details (*Ecological Archives* E089-153-A2).