ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/268019236

Maximum entropy and ecology. A theory of abundance, distribution, and energetics

Article · January 2011

DOI: 10.1093/acprof:oso/9780199593415.001.0001

CITATIONS
110

READS
128

1 author:



John Harte
University of California, Berkeley
238 PUBLICATIONS **11,953** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



DynaMETE View project

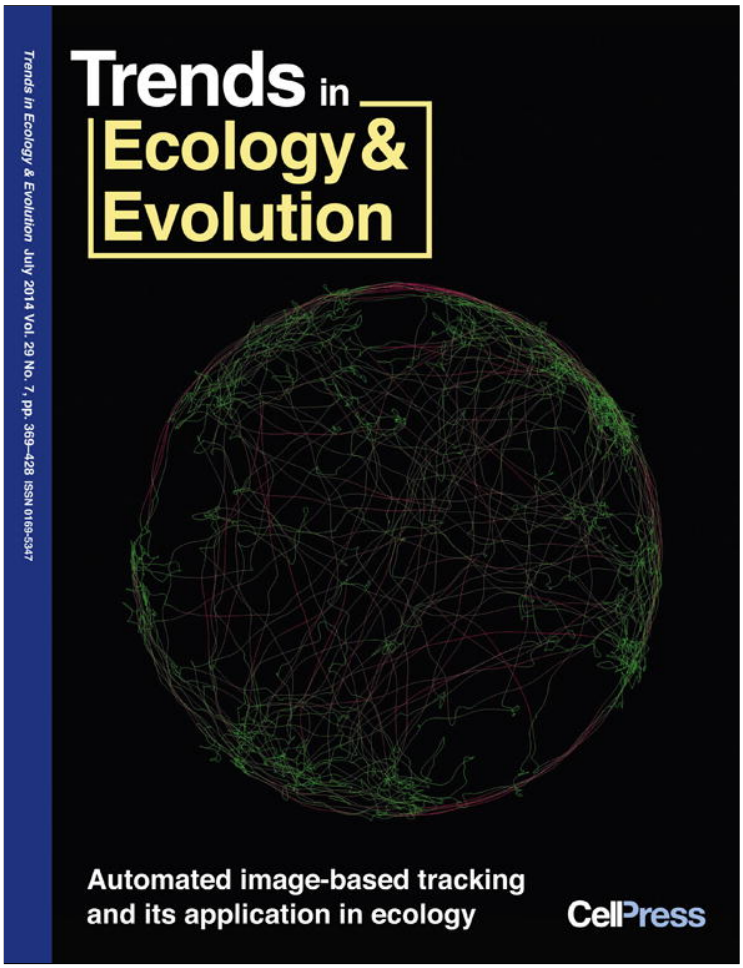


Alpine-Treeline Warming Experiment View project

All content following this page was uploaded by John Harte on 12 October 2016.

The user has requested enhancement of the downloaded file.

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/authorsrights

Author's personal copy

Opinion

CellPress

Maximum information entropy: a foundation for ecological theory

John Harte[1,2] and Erica A. Newman[1,2]

[1] Energy and Resources Group, University of California at Berkeley, 310 Barrows Hall, Berkeley, CA 94720, USA
[2] Department of Environmental Science, Policy, and Management, University of California at Berkeley, 130 Mulford Hall, Berkeley, CA 94720, USA

The maximum information entropy (MaxEnt) principle is a successful method of statistical inference that has recently been applied to ecology. Here, we show how MaxEnt can accurately predict patterns such as species–area relationships (SARs) and abundance distributions in macroecology and be a foundation for ecological theory. We discuss the conceptual foundation of the principle, why it often produces accurate predictions of probability distributions in science despite not incorporating explicit mechanisms, and how mismatches between predictions and data can shed light on driving mechanisms in ecology. We also review possible future extensions of the maximum entropy theory of ecology (METE), a potentially important foundation for future developments in ecological theory.

The MaxEnt principle in ecology

MaxEnt is a widely accepted statistical inference procedure [1,2] that has advanced predictive capacity in topics as diverse as thermodynamics [1,2], economics [3], forensics [4], imaging technologies [5–7], and recently ecology [8–20]. MaxEnt has been proven to produce the least-biased predictions of the shapes of probability distributions consistent with prior knowledge constraining those distributions [1,2] (Box 1). Its introduction to ecology has led to two major advances in landscape-level inference. First is the development of an ecological niche modeling software named 'MaxEnt' [8], which has facilitated mapping species distributions, conservation planning [16,17], and predicting wildfire activity [18]. The second advance, which is the focus here, is to theory building in ecology [11–14,19,20], and is exemplified by the METE [12–14].

MaxEnt is a powerful method of predicting probability distributions, but it is not immediately obvious that it can provide a foundation for building an ecological theory of biodiversity. If it can do so, such a theory would differ from traditional ecological theories and models built around explicit choices of dominant driving mechanisms. In fact, it is the complexity of mechanisms in ecology that motivates this statistical approach to theory building. Given that a vast number of mechanisms influence organisms and their interactions, and so many traits distinguish organisms, it is difficult to select the most influential of these and build theory upon them. The MaxEnt method avoids having to make that selection. As we show, MaxEnt can provide accurate predictions of patterns in macroecology, and also help identify the mechanisms that matter most.

Below, we describe the structure, successes, and failures of METE. We present reasons why theory lacking explicit mechanisms can nevertheless successfully predict patterns in ecology, and explain how failures in such theory can help identify dominant mechanisms. We also review how and why the MaxEnt formalism works. Finally, we discuss prospects for MaxEnt becoming the foundation for a macroecological theory of biodiversity.

What is METE and what can it predict?

The METE is a spatially explicit theory of biodiversity, abundance, and resource allocation, based on the principle of maximization of information entropy (MaxEnt; Box 1). METE is applicable across habitat types, spatial scales, and choices of taxonomic groups.

By analogy with the state variables such as pressure and volume that describe the macroscopic state of a thermodynamic system, METE describes an ecological community with ecological state variables, such as the total number of individuals and species, and total metabolic rate, within some specified taxonomic group and within some specified geographic area. In thermodynamics, the distribution of molecular velocities and other detailed properties of the system can be inferred from the constraints imposed by its state variables. Similarly, from the constraints imposed by the ecological state variables, METE can predict, across spatial scales and without adjustable parameters, the shapes of distributions, such as that of individual across species, species across space, and body sizes across individuals within ecological communities.

A variety of specific models, characterized by different choices of state variables, can be used to create different realizations of the theory. The original model, ASNE, predicts how metabolic energy is distributed over individuals, and how individuals are distributed over species and area (Box 2). ASNE specifies as state variables the area under consideration, A_0; the number of species within some selected taxonomic group, S_0; the total number of individuals in those species, N_0; and the total metabolic rate of all those individuals, E_0. From the constraints imposed by these state variables, application of the MaxEnt inference

Corresponding author: Harte, J. (jharte@berkeley.edu).
Keywords: ecological theory; information entropy; macroecology; MaxEnt; maximum entropy theory of ecology.

0169-5347/
© 2014 Elsevier Ltd. All rights reserved. http://dx.doi.org/10.1016/j.tree.2014.04.009

CrossMark

Author's personal copy

Box 1. What is MaxEnt?

The MaxEnt principle is a rigorously proven inference procedure that yields least-biased predictions consistent with prior knowledge [1,2].

What does 'entropy' mean in this context?
The word 'entropy' refers here to information entropy, rather than thermodynamic entropy. Information entropy is a quantitative measure of uncertainty about an outcome of a draw from a probability distribution.

What is 'rigorously proven' about information entropy?
It has been shown [1,2] that the more uniform (flat and smooth) a probability distribution is, the larger is its information entropy. A uniform probability distribution assigns equal probability to many outcomes, which means that the uncertainty about outcomes of future similar probabilistic events (and, therefore, information entropy) is high. A sharply peaked probability distribution is more informative about outcomes, in that outcomes close to the peak value are more likely than outcomes that are more distant from the peak. Maximizing information entropy is equivalent to maximizing the uncertainty remaining after the shape of the relevant probability distribution is known.

What does 'least-biased' mean?
Bias arises when we make assertions that embody explicit or implicit information that is not compelled by prior knowledge. MaxEnt makes no assumptions about underlying distributions of outcomes of probabilistic events. Instead, it satisfies the constraints imposed by prior knowledge, but otherwise maximizes the uncertainty remaining by making the probability distribution as uniform as possible.

What is 'prior knowledge'?
Prior knowledge of a system is empirical data that can be expressed as constraints on probability distributions describing system outcomes. It can comprise some combination of measured values of moments of a distribution; for example, the range, mean, and variance of a variable. In the case of a theory formulated in terms of state variables, ratios of these variables often constrain the first moments of the distributions being predicted.

Why is maximization of information entropy necessary; why not just let every probability distribution be uniform?
A uniform distribution will, in most instances, be incompatible with the constraints imposed by prior knowledge.

How is information entropy maximized?
Often, maximization of information entropy for a distribution can be obtained with algebraic methods, as shown with examples in [14]. The formal procedure uses the calculus of variations and the Lagrange multiplier technique [38]. Table I shows examples of probability distributions that maximize information entropy subject to specified constraints (see also [39]).

Table I. MaxEnt probability distributions consistent with specified constraints

Constraint[a]	Form of $p(x)$ predicted by MaxEnt[b,c]
[minimum value, maximum value]; (range)	Uniform distribution
$\langle x \rangle$; (mean)	Discrete geometric or continuous exponential distribution [d]: $e^{-\lambda x}$
$\langle x \rangle$ and $\langle x^2 \rangle$; (mean and variance)	Gaussian (normal) distribution
$\langle \log(x) \rangle$, $\langle \log^2(x) \rangle$	Lognormal distribution
$\langle \log(x) \rangle$	Power-law distribution: $x^{-\lambda}$

[a]The symbol $\langle a \rangle$ refers to the mean value of a for all cases in the first column. The variable x can be either continuous or discrete, depending on context.

[b]λ = Lagrange multiplier.

[c]Normalization constants for distributions are not shown (see [39] for further examples).

[d]A special case: if x = all integers between x_{min} and x_{max}, and $\langle x \rangle = (x_{max} + x_{min})/2$, then $p(x) = 1/(x_{max} - x_{min} + 1)$. This would apply to a six-sided fair die, with $p(x) = 1/6$.

procedure results in predictions for a large number of probability distributions describing patterns in macroecology. Figure 1 summarizes many of the predictions of the ASNE model and Box 2 describes in more detail the mathematical core of ASNE.

Many examples of successful predictions of the ASNE version of METE have been published [12–15,21,22]. Among its most surprising and accurate predictions is the predicted form of SAR. ASNE predicts that all SARs collapse onto a universal curve when the local slope of the SAR at any scale, A, is plotted against the ratio of N/S at that scale [13]. Given this feature of scale collapse, any ecosystem may be compared to any other ecosystem, regardless of spatial scale, the taxa that have been censused, or average size of an organism being considered. As shown in Figure 2, the ASNE prediction reasonably describes empirical SARs across various habitat types, taxa, and spatial scales, and does so more convincingly than the

Box 2. The structure of METE

The accuracy of many of the predictions of METE derives in part from the MaxEnt inference procedure (Box 1). The choice of appropriate state variables also influences the success of theory derived from MaxEnt, as does the choice of how the probability distributions to be obtained from MaxEnt are defined.

Two probability distributions form the core of ASNE. The 'ecosystem structure function' $[R(n, \varepsilon | S_0, N_0, E_0)]$ is a joint conditional distribution over abundance (n) and metabolic rate (ε). Here, $R \cdot d\varepsilon$ is the probability that if a species is picked from the species pool, then it has abundance n, and if an individual is picked at random from that species, then its metabolic energy requirement is in the interval $(\varepsilon, \varepsilon + d\varepsilon)$.

The second core probability distribution is the 'species-level spatial abundance distribution' or SSAD, represented by $\Pi(n|A, n_0, A_0)$, which describes the spatial aggregation of individuals of a species. If a species has n_0 individuals in an area A_0, then $\Pi(n)$ is the probability that it has n individuals in an area A within A_0. Values and ratios of the state variables A_0, S_0, N_0, and E_0 constrain R and Π; however, there are no adjustable parameters. The predictions of ASNE for the shape of the probability distributions describing patterns in macroecology derive from R and Π (Figure 1, main text).

Another model within the METE framework resembles ASNE but adds a state variable, W_0, corresponding to a resource constraint, such as water. Just as E_0 is the total rate of energy utilization by the system, W_0 would be the rate of water use by all the individuals. In this extension of the ASNE model (ASNEW), the ecosystem structure function becomes $R(n, \varepsilon, w | S_0, N_0, E_0, W_0)$. Here, the new ecosystem structure function R is defined in analogy with the ASNE R, above. All the metrics shown in Figure 1 in main text, along with additional ones, such as the distribution of water utilization rates across individuals, can be derived using the new form of R. Consequences for macroecological predictions of this extension of ASNE, as well as a discussion of including nested state variables corresponding to higher-order taxonomic categories such as genus or family, are discussed in section 'How can METE be extended beyond ASNE?' in the main text.

Author's personal copy

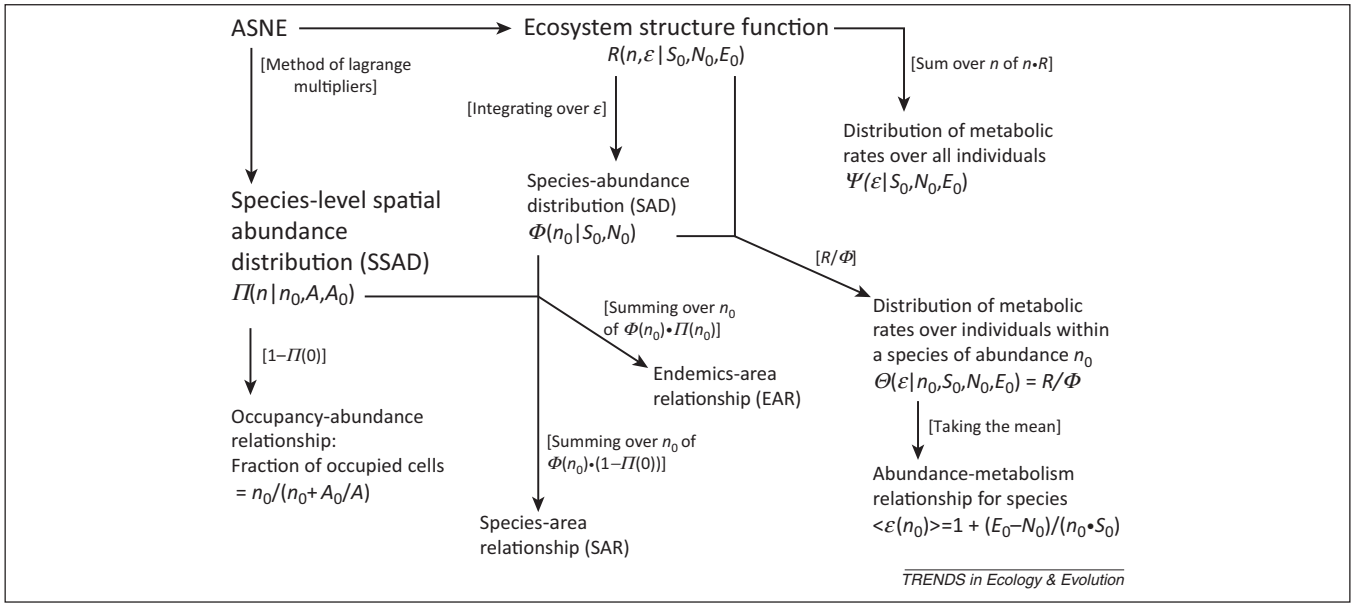


Figure 1. The structure of the ASNE version of maximum entropy theory of ecology (METE). This chart shows how metrics of ecology derive from knowledge of the two core distributions, R and Π, which in turn are predicted from the MaxEnt principle applied to the ecological state variables A, S, N, and E.

power-law SAR. The predicted form of the SAR has been used to estimate species richness at the biome scale from knowledge of state variables in small plots within the biome ([13], J. Harte and J. Kitzes, 2014 unpublished).

The species-abundance distribution (SAD; Φ in Figure 1) predicted by ANSE is the Fisher log series function, $\Phi \sim (1/n)\exp(-\beta n)$, where β is determined by the state variables. This distribution often, but not always, describes empirical SADs accurately [12,14,15]. Several studies [14,21,22] provide confirmation that the prediction by METE of the distribution of metabolic rates across all individuals in a given area (Ψ in Figure 1) is realistic. Related to this, ASNE predicts a relationship between abundance and average metabolic rate of individuals within a species (Figure 1), a prediction consistent with the Damuth rule [23] relating abundance and average body size of individuals in a species. Theoretical implications of the mixed success of this prediction are discussed below in 'How can METE be extended beyond ASNE?'.

Detailed derivations and extensive tests of the many predictions of METE are published elsewhere [12–15,21,22]. Although more testing is needed, METE is arguably the beginnings of a unified theory of macroecology.

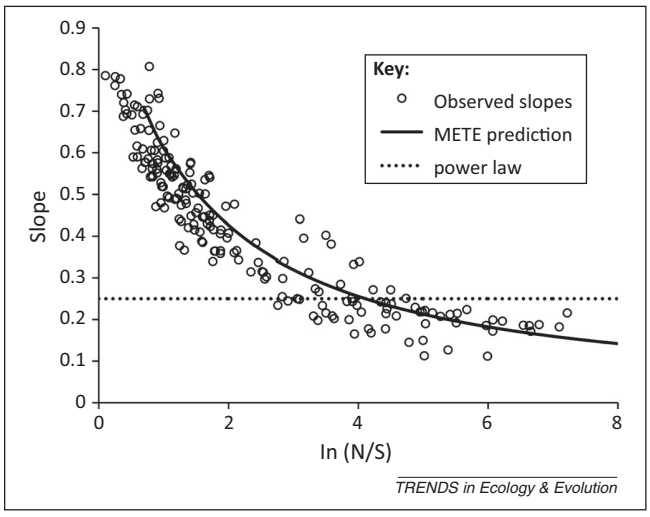


Figure 2. Universal scale collapse of species–area relationships (SARs) in the maximum entropy (MaxEnt) theory of ecology. Predicted and observed values of the slope of an SAR at every spatial scale are plotted as a function of the value of the ratio of abundance to species richness at each scale. SAR data are derived from a variety of ecosystems, taxa, and spatial scales from multiple sources; for derivation of the theoretical prediction from MaxEnt and for sources of data, see [13,14]. Note that the often-assumed power-law form of the SAR with slope ~1/4 corresponds to all data points lying on the horizontal broken line.

How can a theory with no explicit mechanisms yield accurate predictions?

A frequent criticism of MaxEnt-based theory, including METE, is that logic without mechanism cannot possibly be the basis of fundamental theory in science [24]. Here are four possible responses.

Relevant mechanisms are sufficiently embodied in values of state variables

This response asserts that mechanism is implicit, not explicit in METE. The prior knowledge used in METE is knowledge of the state variables; the values of these state variables (species richness, etc.) are in turn the outcome of numerous mechanisms. From this perspective, METE produces accurate results because meaningful information about the mechanisms driving the shapes of the metrics of macroecology (such as abundance and body size distributions) is captured by the empirical values and ratios of the state variables. The theory does not predict the values of the state variables, but once it incorporates their numerical values into constraints, accurate predictions result. The role of mechanism, then, is to determine the values of the state variables [14,15,21,25].

Much scientific theory and explanation is nonmechanistic

At the deepest level, physics and other fields of science have to varying degrees moved away from seeking

Author's personal copy

process-based theoretical foundations. For example, all the fundamental results of thermodynamics and statistical mechanics, such as the Maxwell–Boltzmann distribution of molecular kinetic energies, can be derived in the absence of a mechanistic framework, by applying MaxEnt under the constraints imposed by knowledge of the state variables (e.g., pressure, volume, or number of molecules) of a thermodynamic system [1,2]. Attempts to base quantum theory on some underlying stochastic mechanism always fail, and there is no mechanism from which the Schrödinger equation follows. In short, it has been demonstrated in many areas of science that asking where the mechanism is amounts to asking the wrong question.

A possible rebuttal to this is that ecology clearly is not physics. Major features of physical theories, such as the identical nature of particles, do not hold for species or their members, and the number of demonstrably relevant mechanisms that influence ecosystem structure is enormous [26–30]. This leads to the third response.

Given that there are so many mechanisms, processes, and trait-specific interactions at work in an ecosystem, they can effectively be ignored

Perhaps there are so many niches and traits, types of interaction, and evolutionary strategies for achieving sufficient fitness, that any particular mechanism is not strong enough to alter macroecological patterns over multiple spatial scales [31]. It has also been argued that each species in a quasi-steady state system has achieved an equivalent (or symmetric [32]) seat at the ecological table; Hubbell has used this argument to defend the assumptions behind neutral theory [33]. Either argument could be used to motivate theoretical approaches in ecology that ignore highly complex interactions.

MaxEnt is a way to identify relevant mechanisms

Consider a putatively parallel situation in thermodynamics, the ideal gas law, $PV = nRT$ (where, for a defined system, P is pressure, V is volume, n is the number of moles of a gas, R is the universal gas constant, and T is temperature). This relation can be derived using MaxEnt [1,2] and it is remarkably accurate in most situations. However, observations of its failure at sufficiently high pressure led to the discovery of the Van der Waals force that results from dipole–dipole interactions between molecules. Some of the predictions of METE, such as SAD, also appear to fail in stressed, highly disturbed, or rapidly changing ecosystems [14]. Should certain METE metrics fail in predictable ways under certain ecological conditions, these patterns could lead to identification and better understanding of the dominant mechanisms that govern dynamic changes in ecosystems. In this way, METE might reveal, rather than assume in advance, dominant mechanistic drivers of ecosystem structure.

Why might MaxEnt sometimes yield inaccurate predictions?

It is well established that the MaxEnt inference procedure often yields highly accurate predictions. This is because biases lead to incorrect predictions, and the MaxEnt procedure minimizes at least one major source of bias: basing predictions on hidden and unwarranted assumptions (Box 1). There is no guarantee that MaxEnt will yield accurate predictions, but it will minimize that source of bias and yield the most accurate prediction possible given the constraints imposed by prior knowledge.

When MaxEnt fails to make accurate predictions, it can mean, trivially, that the prior knowledge obtained for the system was incorrectly measured, contained bias, or that mistakes occurred in the analysis. It is important to note that the term 'information' in 'information theory' does not imply truth or usefulness of that information. If both data and application of the mathematical methods are correct, then inaccurate solutions might be obtained for three reasons: (i) the available prior knowledge was insufficient to infer an accurate answer (although the least-biased inference will be produced given that incomplete knowledge). This might be the case if, for example, the system is changing so rapidly that prior knowledge is continually obsolete. Just as the ideal gas law will not be useful in a defined volume of turbulent air, prior knowledge in ecosystems undergoing rapid change will likely not adequately inform MaxEnt models; (ii) incorrect assumptions were made about which variables are influential. Analogously, if the ideal gas law were formulated using a state variable that describes the surface area, rather than volume, of the physical container of a gas, it would produce incorrect predictions; and (iii) additional constraints in the application of MaxEnt can draw predictions away from least-biased estimations. For example, in modeling the geographic presence of a species in the species distribution model application of MaxEnt, such constraints can include incorrect assumptions about relations of relevant ecological variables. Other constraints that potentially introduce bias include application of smoothing factors after model estimations and use of highly collinear variables.

How can METE be extended beyond ASNE?

METE can potentially take various forms. The structure of ASNE, a specific realization of METE, is described in Box 2 and its predictions are summarized in Figure 1. Other realizations of METE arise from selecting different sets of state variables.

If, for example, one adds to the ASNE state variables a new resource variable (call it W_0 for water), the resulting ASNEW version of METE will produce a SAD that differs from the Fisher log series predicted by ASNE, with the $(1/n)$ term altered to $(1/n^2)$ [14]. In general, with a total of r resources (including E_0) to be allocated, the predicted SAD is a product of an exponential and a term $1/n^{r+1}$. This modification increases the predicted fraction of species that are rare. Intuitively, this makes sense; either a greater number of limiting resources might provide more specialized opportunities for rare species to survive, or induce rarity in species that would in other contexts be abundant, although these alternatives have not yet been tested. Thus, METE provides a framework in which the degree of rarity in a community can be related to the number of resources driving macroecological patterns. This example demonstrates how METE, a nonmechanistic theory, can provide insight into what mechanisms might be driving macroecological patterns.

Author's personal copy

Evidence for the need to extend METE beyond ASNE stems from studies [21,22] examining the distributions of metabolic rates across individuals and species. Although ASNE accurately predicts the distribution of metabolic energy across individuals [14,21,22], empirical distributions of the individuals-averaged metabolic rates across species deviate considerably from the ASNE prediction [21,22]. Moreover, empirically, there is mixed empirical support for the Damuth rule [23] or for the related energy equivalence principle [34], with many data sets showing considerable scatter when log(metabolic rate) or log(body size) is plotted against log(abundance).

To possibly address this, METE has recently been extended from the original ASNE version to one in which knowledge of the numbers of units at coarser taxonomic category provide additional state variables and, therefore, additional constraints (J. Harte *et al.*, 2014, unpublished). For example, if the number of families in the community, F_0, is also known, then the AFSNE version of METE predicts new metrics, such as the family–area relationship, and the distribution of species richness over the families. Most predictions of the original theory are left essentially unchanged. However, the predicted relationship between abundance and average metabolic rate of the individuals in a species dramatically changes when one or more additional taxonomic categories are added. On a log-log plot of species abundance versus metabolic rate, AFSNE predicts a series of parallel lines with slope −1, with each line corresponding to a family with a fixed number of species. AFSNE also makes additional predictions of universality, such as the families–area relationship, but does not alter the universal collapse prediction for the SAR. Further tests of these relationships are needed. Incorporating higher-level taxonomic information is one promising method of resolving the discrepancies in the metabolic rate predictions of ASNE [21].

What should one conclude if the more taxonomically resolved theory, ASFNE, is more accurate than the original ASNE model? Additional valid constraints on a probability distribution improve MaxEnt predictions, and knowing what types of additional constraint are most effective in that regard can provide clues about driving mechanisms. For example, improving the accuracy of METE predictions by including higher taxonomic categories as state variables can indicate that the macroecological patterns are shaped by evolutionary history as well as by extant ecological mechanisms.

More evidence for a need to modify ASNE stems from observations that ASNE predictions that are valid for systems in a relative steady state tend to be less valid in systems undergoing relatively rapid change as a result of natural diversification processes or anthropogenic disturbance [14]. In such systems, the fraction of rare species is often greater than (A. Rominger, 2014, personal communication) or less than ([35], E.A. Newman *et al.*, 2014, unpublished) that predicted by the Fisher log series distribution. Additional resource constraints (e.g., ASNEW) might account for more rarity than predicted by ASNE, whereas, more generally, a dynamic version of METE, perhaps based on a mechanistically based theory of the time-evolution of state variables, or based on Tsallis entropy [36] (Box 3), might resolve the failure of ASNE predictions in changing systems.

Finally, work by Williams [37] has shown that including the number of linkages in a food web as a state variable can provide insight into the distribution of trophic linkages across nodes in a trophic network. The inherent flexibility of state variable-based theory gives rise to a family of possible realizations that can be empirically compared to determine which set of state variables provide the most explanatory power.

Box 3. Measures of information entropy and a frequentist motivation for Shannon entropy

The Shannon form of the information entropy of a probability distribution $p(n)$ is defined as $I = -\sum_n p(n)*\log[p(n)]$ [40]. The form of Shannon information entropy shown here assumes that all prior information about $p(n)$ is contained in the constraints imposed upon $p(n)$. Additional prior knowledge can be imposed by replacing $\log[p(n)]$ in the formula for I with $\log[p(n)/q(n)]$, where q(n) is a prior distribution. METE uses only constraints derived from state variables when performing the entropy maximization calculation.

In addition to its application to inference of probability distributions, the Shannon formula is the basis of a diversity index familiar to ecologists [41,42]. This diversity measure incorporates information about total species richness, as well as the evenness of the distribution of individuals or biomass across species.

The Shannon formula is the most commonly applied measure of information entropy and it can be shown to be the unique measure that satisfies a set of plausible axioms about information [43]. However, other measures of information entropy exist. For example, relaxing the assumption that entropy must be additive for disjoint systems results in an entropy measure known as Tsallis entropy [36], which has multiple applications in physics. As discussed briefly in [14], the Shannon entropy formula appears most appropriate in ecology based on empirical evidence, but further study of Tsallis entropy in ecological contexts may be useful for modeling ecological systems that have internal elements exhibiting long-range interactions or memory effects, or systems that are far from equilibrium (C. Tsallis, 2014, personal communication).

In statistical physics, where systems may comprise large numbers of objects (e.g., a container of gas with $\sim 10^{24}$ molecules), a 'counting-of-states' argument also leads to the conclusion that Shannon entropy will be maximized. The actual distribution of molecules across the possible energy or spatial states available to them will be, by chance alone, that distribution which is compatible with the largest number of permutations of how energy is allocated across the molecules, or the molecules are distributed through a volume (the mathematics of this argument is described in [44]). This frequentist, or counting-of-states, argument requires both an assumption about the distinguishability or indistinguishablity of the objects, as well as there being a large number of them. The arguments set forth in [1,2] do not require those assumptions and, therefore, are more general.

Concluding remarks

The maximum entropy method can enrich understanding of biodiversity. Information from a small number of state variables, such as total number of taxa and individuals and the total rate of metabolic activity, in a prescribed area and for a prescribed taxonomic group, can be used to predict detailed patterns in the distribution, abundance, and energetics of individuals and species across wide ranges of spatial scales (Figure 1). Sizeable deviations from the patterns predicted by the ASNE version of METE are observed sometimes, such as in systems undergoing relatively rapid ecological or evolutionary change. Apparent

Author's personal copy

regularities in deviations from theory suggest the possibility of extending the theory with other state variables and the identification of mechanisms important in determining large-scale ecological patterns.

Various opportunities to extend METE exist. For example, the current ASNE version of METE is a static, not yet a dynamic, theory of macroecology. ASNE uses 'species' as the default taxonomic level; the theory is readily extended to higher taxonomic levels, and would benefit from including phylogenetic information. Finally, although ASNE assumes that the allocation of a single resource, energy, is sufficient for prediction, it can be readily extended to include additional resources.

In summary, the MaxEnt inference procedure provides a new approach to the study of patterns of biodiversity. Instead of building models by assuming the dominance of particular driving mechanisms, the MaxEnt approach uses empirical constraints (state variables in the case of METE) to predict the probability distributions characterizing patterns in macroecology. Ideally, dominant mechanisms can then be inferred from the patterns of failure of nonmechanistic theory. In tandem with mechanism-based models, these approaches can help achieve a predictive understanding of biodiversity.

Acknowledgments

We gratefully acknowledge discussions with, and manuscript edits from, David Hembry, Tom Jackson, Justin Kitzes, Max Moritz, Andrew Rominger, Mark Wilber, and Yu Zhang, as well as two anonymous reviewers. Funding was provided by the Gordon and Betty Moore Foundation, and by the US National Science Foundation (NSF) through the Graduate Research Fellowship Program and grant NSF-EF-1137685.

References

1 Jaynes, E.T. (1957) Information theory and statistical mechanics. *Phys. Rev.* 106, 620
2 Jaynes, E.T. (1982) On the rationale of maximum entropy methods. *Proc. Inst. Elec. Electron. Eng.* 70, 939–952
3 Judge, G. and Miller, D. (1996) *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, John Wiley
4 Roussev, V. (2010) Data fingerprinting with similarity digests. In *Advances in Digital Forensics VI* (Chow, K-P. and Shenoi, S., eds), pp. 207–226, Springer
5 Frieden, B.R. (1972) Restoring with maximum likelihood and maximum entropy. *J. Opt. Soc. Am.* 62, 511–518
6 Skilling, J. (1984) Theory of maximum entropy image reconstruction. In *Maximum Entropy and Bayesian Methods in Applied Statistics* (Justice, J.H., ed.), pp. 156–178, Cambridge University Press
7 Gull, S.F. and Newton, T.J. (1986) Maximum entropy tomography. *Appl. Opt.* 25, 156–160
8 Phillips, S.J. *et al.* (2006) Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190, 231–259
9 Phillips, S.J. (2008) Transferability, sample selection bias and background data in presence-only modelling: a response to Peterson et al. (2007). *Ecography* 31, 272–278
10 Elith, J. *et al.* (2011) A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* 17, 43–57
11 Dewar, R.C. and Porté, A. (2008) Statistical mechanics unifies different ecological patterns. *J. Theor. Biol.* 251, 389–403
12 Harte, J. *et al.* (2008) Maximum entropy and the state variable approach to macroecology. *Ecology* 89, 2700–2711
13 Harte, J. *et al.* (2009) Biodiversity scales from plots to biomes with a universal species area curve. *Ecol. Lett.* 12, 789–797
14 Harte, J. (2011) *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics*, Oxford University Press
15 White, E.P. *et al.* (2012) Characterizing species abundance distributions across taxa and ecosystems using a simple maximum entropy model. *Ecology* 93, 1772–1778
16 Elith, J. *et al.* (2006) Novel methods improve prediction of species distributions from occurrence data. *Ecography* 29, 129–151
17 Franklin, J. (2009) *Mapping Species Distributions: Spatial Inference and Prediction*, Cambridge University Press
18 Parisien, M-A. and Moritz, M.A. (2009) Environmental controls on the distribution of wildfire at multiple spatial scales. *Ecol. Monogr.* 79, 127–154
19 Pueyo, S. *et al.* (2007) The maximum entropy formalism and the idiosyncratic theory of biodiversity. *Ecol. Lett.* 10, 1017–1028
20 Shipley, B. *et al.* (2006) From plant traits to plant communities: a statistical mechanistic approach to biodiversity. *Science* 314, 812–814
21 Newman, E.A. *et al.* (2014) Empirical tests of within- and across-species energetics in a diverse plant community. *Ecology* http://dx.doi.org/10.1890/13-1955.1
22 Xiao, X. *et al.* (2014) A strong test of the Maximum Entropy Theory of Ecology. *arXiv* 1308.0731
23 Damuth, J. (1981) Population density and body size in mammals. *Nature* 290, 669–700
24 Clark, J.S. (2009) Beyond neutral science. *Trends Ecol. Evol.* 24, 8–15
25 McGill, B.J. and Nekola, J.C. (2010) Mechanisms in macroecology: AWOL or purloined letter? Towards a pragmatic view of mechanism. *Oikos* 119, 591–603
26 Pickett, S.T.A. and White, P.S., eds (1985) *The Ecology of Natural Disturbance and Patch Dynamics*, Academic Press
27 White, P.S. and Jentsch, A. (2001) The search for generality in studies of disturbance and ecosystem dynamics. *Prog. Bot.* 62, 399–450
28 Parrish, J.K. and Edelstein-Keshet, L. (1999) Complexity, pattern, and evolutionary trade-offs in animal aggregation. *Science* 284, 99–101
29 Stephens, P.A. *et al.* (1999) What is the Allee effect? *Oikos* 87, 185–190
30 Thompson, J.N. (2005) *The Geographic Mosaic of Coevolution*, University of Chicago Press
31 Frank, S.A. (2009) The common patterns of nature. *J. Evol. Biol.* 22, 1563–1585
32 Hubbell, S. (2006) Neutral theory and the evolution of ecological equivalence. *Ecology* 87, 1387–1398
33 Hubbell, S. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press
34 White, E. *et al.* (2007) Relationships between body size and abundance in ecology. *Trends Ecol. Evol.* 22, 323–330
35 Kempton, R.A. and Taylor, L.R. (1974) Log-series and log-normal parameters as diversity discriminants for the Lepidoptera. *J. Anim. Ecol.* 43, 381–399
36 Tsallis, C. (1988) Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* 52, 479–487
37 Williams, R.J. (2010) Simple MaxEnt models explain food web degree distributions. *Theor. Ecol.* 3, 45–52
38 Arfken, G.B. and Weber, H.J. (2005) *Mathematical Methods for Physicists*. (6th edn), Academic Press
39 Frank, S. and Smith, E. (2011) A simple derivation and classification of common probability distributions based on information symmetry and measurement scale. *J. Evol. Biol.* 24, 469–484
40 Shannon, C.E. (1948) A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423
41 Pielou, E.C. (1969) *An Introduction to Mathematical Ecology*, Wiley-Interscience
42 Spellerberg, I.F. and Fedor, P.J. (2003) A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the 'Shannon–Wiener' Index. *Global Ecol. Biogeogr.* 12, 177–179
43 Khinchin, A.I. (1957) *Mathematical Foundations of Information Theory*, Dover
44 Haegeman, B. and Loreau, M. (2009) Trivial and non-trivial applications of entropy maximization in ecology: a reply to Shipley. *Oikos* 118, 1270–1278

View publication stats