Mechanisms in macroecology: AWOL or purloined letter? Towards a pragmatic view of mechanism

Brian J. McGill and Jeffrey C. Nekola

B. J. McGill (mail@brianmcgill.org), School of Natural Resources, Univ. of Arizona, Tucson, AZ 85719, USA. – J. C. Nekola, Dept of Biology, Univ. of New Mexico, Albuquerque, NM 87131, USA.

Ecologists often believe the discovery of mechanism to be the central goal of scientific research. While many macroecologists have inherited this view, to date they have been much more efficient at producing patterns than identifying their underlying processes. We discuss several possible attitudes for macroecologists to adopt in this context while also arguing that in fact macroecology already has many mechanisms that are ignored. We briefly describe six of these: central limit theorem, fractals, random sampling and placement, neutral theory (and descendents), concordance of forces, and maximum entropy. We explore why these mechanisms are overlooked and discuss whether they should be. We conclude that macroecology needs to take a more pragmatic, less ideological approach to mechanism. We apply this viewpoint to the recent controversy over maximum entropy and suggest that maximum entropy needs to be viewed more pragmatically and less ideologically.

How vital is the identification of mechanism to the scientific progress of ecology, and specifically, macroecology? While some iconoclasts (Peters 1991) suggest mechanism is not vital, many important figures in the field (Schoener 1986, Tilman 1987, Rosenzweig 1991, Brown 1999b) suggest that mechanism documentation is central to the field. This is made clear through the prevalent meme that ecological and macroecological investigation should proceed in two steps: 1) find patterns and 2) find the processes underlying the patterns (Rosenzweig 1991, Brown 1999b, Gaston and Blackburn 1999, Lawton 1999). Although anecdotal, our personal experiences with reviewers and questioners after seminars suggests an almost obsession with mechanism, with phrases like 'all you've done is explain a pattern with another pattern' or 'it's only science if you can explain it in terms of population processes' being commonplace.

There can be little argument that mechanism is useful and desirable in science. Some key benefits of understanding mechanism include:

1. Curiosity – knowing the 'why' of things is certainly rewarding and arguably a major motivation.
2. Extrapolation/generalizability – one is never sure how far one can extend a phenomenological pattern into new taxa, new scales or new conditions: patterns are useful for interpolation but not necessarily for extrapolation (Peters 1991, Helmuth et al. 2005, Peterson et al. 2007). Even strong proponents of a phenomenological approach (Peters 1991) acknowledge this limitation. Understanding mechanism can give insight into how far a prediction can be extended into new domains (Dunham and Beaupre 1998), something becoming increasingly important in this world of global change.
3. Connecting disciplines – mechanism almost by its nature takes a reductionist approach and necessarily spans across levels of organization. Thus communities can be described by population processes, populations can be described by individual physiology and behaviour, which can in turn be described by cellular biology and the laws of physics and chemistry. This bridging of fields is not only intellectually satisfying, but it is productive since new findings at one level can be leveraged into novel predictions at another.

However, we macroecologists must also honestly and collectively acknowledge that after a century and a half of practice (von Humboldt and Ross 1852, Darwin 1859) and two decades as a named field (Brown and Maurer 1989) we have had little success at identifying mechanism. In a series of review articles several of the leading figures in macroecology identified this shortcoming (Brown 1999b, Gaston and Blackburn 1999). While some feel that the tide is turning (Smith et al. 2008) due to a few notable successes (West et al. 1997), on the whole our personal assessment is that macroecology is still more adept at producing patterns – and lists of potential mechanisms – than it is on practicing the strong inference that produces decisive tests which identify underlying mechanism. Species abundance distributions are a particularly egregious example (McGill et al. 2007) where probably a half a dozen or more mechanisms are proposed

per year, with no real consensus having yet been achieved on which are in fact important. In a recent editorial in the International Biogeography Society newsletter, Carsten Rahbek says "macroecology appears to have hit a wall" and notes that "We … need conceptual advancements. The spatial ecological literature is surprisingly poor in formulated hypotheses". The most recent book on macroecology (Gaston and Blackburn 2000) lists half a dozen to a dozen possible mechanisms for each major pattern described, but in almost no case are these authors able to decisively conclude which are actually in operation. In short it is easy to make the case, as suggested in the title of this paper, that mechanisms in macroecology are AWOL (a military acronym in most English speaking countries for Absent-WithOut-Leave, which has strong negative connotations).

It should first be noted that macroecologists do not deserve scorn for this failing. Discovering mechanism in macroecology is hard: by definition the scales involved are far too large to conduct the manipulative experiments (Brown and Maurer 1989) that are largely regarded as the best (and by some the only) route to identifying mechanism (McArdle 1996, Dunham and Beaupre 1998, Resetarits Jr and Bernardo 1998). Even though we all recognize that correlation cannot prove causation, macroecologists are left with no other option than to muddle along with observational data and correlation/regression.

If deducing mechanism is believed central to scientific endeavours, what can macroecologists do? There are at least four possible strategies:

1. *Wait longer* – fields do not mature overnight, and finding mechanism can only occur after the patterns are first found and then refined. Wiegert (1988) suggests there are three stages: What? (patterns), How? (details of patterns – how general are they, what are the patterns in numerical parameters, etc) and only lastly Why? (i.e. mechanism). A good case can be made that macroecology is only in the 2nd (How?) stage and it is premature to expect mechanisms (3rd or Why?) stage (Gaston and Blackburn 1999).
2. *Shape up and work smarter* – others think the lack of mechanism is due to the powerful pull created by the ease with which patterns can be explored: a high school student can do it with nothing more than an internet connection and a spreadsheet program. Many feel that macroecologists are in general not using the available analytical tools as best they could, and that as a field we must pay more attention to natural experiments, microcosm (bottle) experiments, null models, statistical issues and procedures such as spatial and taxonomic autocorrelation (Blackburn and Gaston 1998, Gaston and Blackburn 1999, Blackburn 2004) and path analysis (Shipley 2003), while also incorporating longer time series (McGill et al. 2005, Agrawal et al. 2007). As Robert MacArthur observed a long time ago "Astronomy was a respected science long before ecology, and Copernicus and Galileo never moved a star" (Brown 1999a). Based on this train of thought, as we become cleverer, mechanism will follow.
3. *Mechanism is not required* – Peters, who was a macroecologist before the field had been so named, argued

that the myopic quest for mechanism was useless and probably damaging (Peters 1991). Other scientific fields, such as physics, have advanced far without mechanism. For instance, Newton merely observed an inverse square law worked well in predicting gravitational attraction, yet knew nothing about the underlying cause. In the subsequent four hundred years since Newton we have not significantly elucidated this, although Einstein pushed the problem back one level by equating gravity to curved space-time. Einstein himself was so disenchanted with the lack of mechanism in quantum mechanics that he famously uttered his plea for more mechanism, because "God does not play dice with the universe". Yet, seven decades later these mechanisms have yet to be identified, if in fact they exist at all. In some complex systems such as climatology, where the mechanisms are well known (there are about seven physical laws that govern the atmosphere and oceans, and which can be easily incorporated into a computer model; Peixóto and Oort 1984) the resulting dynamics are so complex that disciplinary researchers are entirely comfortable pursuing correlational patterns (Wallace and Gutzler 1981, Barnston and Livezey 1987).

4. *Look again, macroecology already has mechanisms* – A final alternative is that macroecology has already found mechanisms, but that they are overlooked either because they are not like those we were expecting or because they are so unpalatable that we reject them. This is the "purloined letter" option of the title – a reference to a story by Edgar Allen Poe in which the villain successfully hid a critical letter by unexpectedly leaving it in plain site on the fire place mantle, contrary to the expectations of the searchers.

A good case can be made for all four of the above arguments – likely all of them are true and relevant. Elsewhere one of us (McGill 2003a) has argued for some potential paths along the lines of no. 2 above and we find ourselves in full agreement with the suggestions of others (Blackburn and Gaston 1998, Gaston and Blackburn 1999). We also feel that no. 3 deserves more attention than commonly given, not because we think the extreme position stated by Peters (1991) is true but because the ecological obsession with mechanism is not characteristic of all other successful scientific fields. But in this paper we wish to explore in some detail the fourth argument, which has not often been made. Specifically, we suggest that macroecology has produced a number of mechanisms and practiced strong inference in decisively rejecting some of them, but they are not recognized due to their unexpected nature and failure to fit into our preconceived notions of what ecological mechanisms should look like.

Within ecology broadly there has been a great deal of discussion and dispute about what are the appropriate sources of mechanisms. MacArthur (1968) and May (1976) found mechanism in population dynamic models that described species interactions and could be aggregated to explain community structure. Schoener (1986), however, dismisses such approaches as "phenomenological" and argued that the true mechanisms for community ecology are to be found at the individual level through behaviour and physiology. Still others

argue that mechanism needs to trace back even farther to the "first principles of geometry, biology (e.g. natural selection), physics and chemistry" (West and Brown 2005). Finally, some adopt a pragmatic view that mechanism is simply those things that elucidate and advance science (Rosenzweig 1991).

In spite of this diversity of opinions it is clear that macroecology, with a few exceptions (West et al. 1997), is not looking for mechanism in any of these areas. Perhaps this is not surprising – with the change of scales and reframing of questions implicit in macroecology the nature of mechanism will almost certainly change as well. Where, then, do macroecologists seek mechanism? Brown clearly expected to find them at least in part in the same niche and feeding processes studied in community ecology (Brown and Maurer 1989, Brown 1995). Later, Brown (1999b) and Lawton (1999) iterated towards a broadly similar vision where mechanism was found through reductionism via physics and physiology mixed with the unique biological mechanism of evolution. Maurer (1999, 2005) however argued that an important consequence of macro-scales is that statistical-mechanics-like approaches may also be important, with individuals and species playing the role of a multitude of complexly different particles which need to be averaged across. This view has been reiterated by the fact that macroecological patterns tend to have exact correlates across multiple non-ecological systems, suggesting that their ultimate mechanisms may also reside outside of ecology (McGill 2003a, Nekola and Brown 2007).

The remainder of this paper is divided into three sections. In the middle third of the paper we explore whether these are in fact mechanisms and what the implications of their success are for mechanistic search in macroecology. Specifically, we will argue for adopting a more pragmatic view that mechanisms are things that are useful in advancing our abilities to understand and predict – no more and no less. Finally we will evaluate these six mechanisms against this standard. But first, we review six different general mechanisms (often adopted from physics and/or statistical mechanics) that may be important in a macroecological perspective to provide a common ground for the discussion.

Central limit theorem

The central limit theorem (CLT) states that a sum of a set of random variables will approach a normal distribution as the number of variables approaches infinity. In practice, this pattern is achieved at far less than an infinite number of variables, with normality often being approached in no more than four to five. When random variables are multiplied together, causing their log transforms to be added, a normal distribution on a log-transformed scale (the lognormal distribution) is generated (MacArthur 1960, May 1975, Limpert et al. 2001, McGill 2003a).

The simplest to prove (weak) version of the CLT requires that the random variables be independent and identically distributed (IID). This proof requires only a few lines combining the use of characteristic functions in probability with the Taylor series. However, probability theorists have spent much of the 20th century progressively weakening these conditions. While the resultant proofs have become vastly more complicated, they demonstrate that sums of non-identically distributed and correlated random variables also converge to normality often only a little bit slower than in the IID case (Grimmett and Stirzaker 1992). The normal distribution also emerges as the output of MaxEnt when only the first two moments (mean and variance) are known. More profoundly, MaxEnt can be used to prove both the weak and strong versions of the CLT with reference to sums of random variables (Johnson 2004).

There has been enormous debate on the applicability of the CLT to ecology. Although lognormality is most often discussed in terms of species abundances (SADs), many other macroecological distributions are at least close to lognormal in shape, including the distribution of body sizes across species (i.e. average body size per species), the distribution of body sizes across individuals (i.e. ignoring species), and the distribution of species range sizes (Brown 1995, Gaston and Blackburn 2000). MacArthur (1960) and later more mathematically-precisely May (1975) suggested that SAD lognormality is generated by the multiplication of species growth rates over time. Yet Pielou (1975) and later authors (Williamson and Gaston 2005, Alonso et al. 2008) state that this is an incorrect interpretation since the growth rates of different species are not sampled from the same distribution. How can some of the top intellects in mathematical ecology fail to agree on such a basic mathematical point? The reality is that probability theory, more than any other field of math, is at points only a thin veneer over philosophy and differences in interpretation over what counts as a 'random variable' or 'instantiation of a distribution' are subtle. In the end, this debate is moot in this particular case as both May (1975) and McGill (2003a) provide examples of how the CLT can lead to SAD lognormality without invoking assumptions regarding species growth rates.

One reason for the success of the lognormal is that it can accommodate a wide variety of shapes. Although always forming a perfectly symmetrical Gaussian on a log-scale, on an arithmetic scale the lognormal changes shape drastically depending on the coefficient of variation (CV=σ/μ). When the CV is close to zero the lognormal is nearly normal. It is little appreciated that although human height is often thought a canonical example of the normal distribution, it is equally well fit by the lognormal and it can be argued that height as an outcome of ontogenetic growth is better modelled as a product of random variables (Limpert et al. 2001). When the CV is large, the lognormal approaches the power distribution with an exponent of -1 (i.e. 1/x) (Montroll and Shlesinger 1982). It is also increasingly well understood (Magurran and Henderson 2003, McGill 2003b, Green and Plotkin 2007) that sampling processes, especially in a biologically realistic context that allows for both spatial and temporal autocorrelation, can distort the shape of a lognormal curve giving it skew along a logscale. As a result, it is hard to reject the lognormal as a possible descriptor of most species abundance datasets (data that remains with the mode at 1 even on a log scale being one of the few exceptions, Southwood 1996).

Fractal geometry and chaotic systems

While fractal dimension simply represents a descriptive statistic, it is important to remember that fractal and fractal-like patterns are readily generated via a wide array of non-linear dynamical (e.g. chaotic) systems (Mandelbrot 1982). As a result, system

fracticality can be used as a rubric to capture the impact of non-linear dynamics within complex systems. Classical fractal geometry assumes that the pattern of interest is self-similar or scale-free – that is to say that they look the same at any spatial or temporal scale (equivalent under any degree of magnification). However, even Mandelbrot (1982) noted that such absolute self-similarity was not required for a system to operate in a fractal fashion. In practical applications, an additional qualification is added: that the patterns are stochastically self-similar. That is to say that they are not perfectly regular like Koch's snowflake, but that there is random variation at all scales constrained such that any statistical treatments of measures of curvature, fractal dimension, general shape, etc are constant across scales.

The impact of fractal geometry on ecological process has been considered for decades (Halley et al. 2004). Some have explored the evolution of 1/f noise from the fractal properties of time series (Halley 1996). Others (Palmer 1988, Ritchie and Olff 1999, Haskell et al. 2002) have generated biodiversity models predicated on the fractal distribution of resources. Harte et al. have demonstrated how spatial fracticality may give rise to the power law form of the species-area relationship (Harte 2008), and that a fractal distribution of individuals can lead to realistic species abundance distributions (Harte et al. 1999). These theories have subsequently been extensively elaborated and tested (Green et al. 2003, Green and Ostling 2003). Kunin and colleagues (Kunin et al. 2000, Hartley et al. 2004) sought to use fractality to extrapolate from easily measured smaller scales up to important but hard to measure larger scales. Others have explored fractally branching trees (Mouillot et al. 2000).

However, the ecological world does not seem very fractal: it is very strongly scale differentiated. Tests of fractal structure typically find that fractal dimension rarely remains constant across more than 1–2 orders of observational magnitude (Palmer 1988, Kunin 1998, Kunin et al. 2000, Green et al. 2003, Hartley et al. 2004). Most recently authors have begun exploring multifractals which assume that the aggregate property of fractal dimension is not constant but changes in some predictable fashion across scales (Scheuring and Riedi 1994, Borda-de-Agua et al. 2002). However, theoretical use of multifractals requires an a priori specification of the change of fractal dimension. As this relationship is more often emprically derived, multifractals border on losing their status as a general mechanism, and may perhaps be best thought of as representing sophisticated tool for measuring and describing data.

Random sampling and placement

Many authors have wondered how many patterns in macro-ecology can be explained by random sampling from a larger regional pool of individuals or species. This idea goes back to the (at the time) radical theory of island biogeography (MacArthur and Wilson 1967) and the contentious debate over null models (Connor and Simberloff 1979, Gotelli and Graves 1996).

Increasingly, sampling theory is evolving in ecology from being the null model used to test the existence of pattern to becoming an outright explainer and predictor of that pattern. Arrhenius (1921) and later Coleman (1981) showed that power-law like species area relationships can be gener-

ated from simple random sampling of a species abundance distribution in combination with an assumption that number of individuals is proportionate to area. MacArthur and Wilson's theory of island biogeography (MacArthur and Wilson 1967) made island biota essentially a random subset of the mainland. Colwell (Colwell and Hurtt 1994, Colwell and Lees 2000) attempted to explain the latitudinal richness gradient by placing species ranges down at random within a finite domain (the MDE or mid-domain effect).

In the end, many of these theories have been shown to be at best approximate. In particular, many authors (Condit et al. 2000, Conlisk et al. 2009) have shown that individuals within a species are clumped and not randomly (Poisson) distributed in space, with more accurate models being produced when this effect is taken into account (Leitner and Rosenzweig 1997, He and Gaston 2003). Similarly, several authors (Hawkins and Diniz-Filho 2002, Connolly et al. 2003, Kerr et al. 2006) have shown that random placement alone is not enough to fully explain variations in diversity across landscapes. However, random placement of species with intra-specific clumping does a good job of parsimoniously producing many key macroecological patterns (McGill and Collins 2003, Harte et al. 2005, 2008).

Neutrality, regional replacement and dispersal limitation

The unified neutral theory of biogeography (Bell 2000, Hubbell 2001) claims to explain many patterns in macro-ecology by assuming neutral population dynamics. These authors touted neutrality as the key innovation. Here neutral means that the per-capita demographic rates of a species are on average equal and do not depend in a species-specific fashion on the environmental background of other species and the abiotic environment. The response was both intensely positive (Alonso et al. 2006) and negative (Clark 2008). It has become increasingly clear, however, that several key aspects of sensu strictu neutralilty are easily and repeatedly empirically falsifiable, namely in terms of white noise demographic variability over many generations (Clark and McLachlan 2003, McGill et al. 2005), the relative unimportance of species-specific responses to the environment (Gilbert and Lechowicz 2004) and the unimportance of interactions with other species (Wootton 2005).

Chave et al. (2002) showed that the same patterns could be produced under a variety of assumptions, both neutral and non-neutral. Although tempting to turn this into a black/white debate and conclude neutral theory is wrong, a more productive route is to try to identify what features of neutral theory overlap with those of Chave et al. and see if a more general, empirically supportable principle emerges. In particular neutral theory has three key ingredients: a) neutral drift of populations in local communities, b) dispersal limitation (in the spatially explicit version), and c) a replenishment of individuals (and hence species) in the local community from the regional community. All of Chave's models include b and c but only half of the models include a, yet they all produce the same results.

This suggests that perhaps ingredients b and c might be the most relevant in explaining observed patterns. It would

be a mistake to simply equate dispersal limitation with neutral theory, however. The fact that spatially explicit dispersal models produce different (and more realistic results) than models that are well-mixed has been known for a long time (Janzen 1970, Levin 1974, Durett and Levin 1998). Moreover non-neutral dispersal limitation has effects similar to neutral dispersal limitation, suggesting that dispersal limitation, and not neutrality, is key (Holyoak and Loreau 2006). The more novel contribution has been the second factor which one might call the "regional replacement" or "species pool theory" (Zobel 1997). Empirical comparisons to date all suggest that this factor may be the essential process underlying much macroecological pattern (Magurran and Henderson 2003, McGill 2003b, Dolman and Blackburn 2004, also see Maurer and McGill 2004). This process requires the following features:

1. a local pool of individuals (i.e. a community) that is tracked;
2. an exogenous source of individuals and species identity (usually a regional pool of individuals but occasionally very high rates of local speciation);
3. periodic death followed by replacement of individuals from the local pool by the exogenous source;
4. an intermediate replacement rate of replacement – if the entire community turns over each iteration (replacement rate of 100%) or if the local community is populated just once and then never replaced (replacement rate of 0%), then the local community is just a sample of the regional pool and echoes most of the properties of the regional pool (Green and Plotkin 2007).

Under these conditions, regional replacement controls the distributions of abundances within the local community and in particular of the rare species. In neutral theory all of the conditions of regional replacement are met along with the addition of neutral dynamics. In Chave et al's work all of the regional replacement conditions are met with the addition of either neutral drift or niche-based population dynamics with or without density dependence. While these are suggestive of the pivotal role of the regional replacement, confirmation comes from three independent studies which directly model just regional replacement and produce identical results (McGill 2003b, Dolman and Blackburn 2004, Zillio and Condit 2007).

Thus of the three aspects of the unified neutral theory of biogeography, neutrality appears dispensable while dispersal limitation and regional replacement appear critical. Ironically these same two features arguably were first prominently featured in a model in the original theory of island biogeography (MacArthur and Wilson 1967). Further work clearly needs to be done to explore the potential of dispersal limitation and regional replacement as a general mechanism.

Concordance of causes

Another very general mechanism might be thought of as concordance of causes. Ecological systems are complex, and so, unlike Newtonian physics where it is often possible to reduce systems to a single factor (e.g. gravity for heavy, aero-

dynamic objects falling through the air), ecological patterns are invariably the outcome of multiple factors interacting with each other (Quinn and Dunham 1983). Under one view these multiple causal factors are all finely balanced, and slight changes in circumstances tip the balance towards one force or another, leading to extreme contingency (Simberloff 2004). But it is equally possible to adopt the view of Mandelbrot (1963) who suggested that a pattern will emerge when it is supported by many forces, while noise will be created when the forces in play lead to contradictory patterns. Mandelbrot specifically was seeking an explanation for the ubiquity of power distributions and pointed out that it is the only distribution which survives the transformations of summation, mixing, and taking the maximum. For example, what if the oft-observed correlation between abundance and occupancy (Gaston et al. 2000) is not produced by a single mechanism? At least half-a-dozen mechanisms have been hypothesized (e.g. passive sampling, abundance and occupancy both being positively correlated with niche breadth or position, abundance and occupancy both being a function of position within a geographic range, etc. Gaston et al. 2000). And it is possible that all of these act simultaneously and the pattern is robust and frequently observed because most likely mechanisms all lead to a positive correlation. This view is not too far from that found in Gaston and Blackburn's review of macroecology (Gaston and Blackburn 2000) where they list multiple causes and assess the relative strength of each. An idea similar to concordance is also found in Darveau et al.'s alternative (2002) to the fractal branching network mechanism for allometric scaling (West et al. 1997).

Maximum entropy

The last few years have seen a spate of papers importing the statistical mechanics approach of maximum entropy to biogeography (Phillips et al. 2006, Phillips and Dudik 2008) and ecology (McGill 2006, Shipley et al. 2006, Pueyo et al. 2007, Dewar and Porté 2008; see also Table 1). Although unappreciated, MacArthur (1955) also used MaxEnt in ecology 50 years ago.

Jaynes (1957) introduced MaxEnt as a logical extension of the framework of information theory (Shannon 2001, Cover et al. 2006) and demonstrated that it could produce most of the major results in statistical mechanics. Statistical mechanics is the study of systems with a multitude of particles that ignores the complex details of each particle (microstate) and instead makes statements either about macrostates (emergent properties) or their statistical distributions. A classic example is an ideal gas, where the position and velocity of each particle describe the microstate (6N measurements in a system with N particles due to x, y and z components of position and velocity). Statistical mechanics instead makes statements about temperature (a macrostate or emergent property) and about the distribution of some microstate property of the particles (e.g. energy). The potential relevance to macroecology is obvious and has been promoted by Lotka (1925) and Maurer (1999, 2005). While statistical mechanics existed as a field of physics investigation long before the development of MaxEnt (McQuarrie and Allan 2000), in ecology the use of statistical mechanics approaches seems to be limited to the application of MaxEnt.

Table 1. Which generic mechanisms explain which macroecological patterns? An "X" indicates that mechanism can be used to explain that pattern. Some patterns (e.g. last row) appear to have not been explained by any general mechanisms.

	MaxEnt	CLT	Random sample	Fractals	Regional replacement and dispersal limitation
Hollow curve SAD on an arithmetic scale	X	X		X	X
Lognormal range size		X			
Lognormal body size	X	X			
Species area relationship	X		X	X	X
Damuth's rule	X				
Nestedness			X		X
Abundance/occupancy/range size correlations			X		
Relationship of diversity to latitude			X		
Correlation of productivity and diversity					

MaxEnt is a general purpose tool. Its goal is to describe the probability distribution, p(n) giving the probability that a particle is in microstate n (i.e. has property n). The process follows three steps:

1. take as input a set $k=1..K$ of 'constraints' relating a microstate property, n, to an aggregate (i.e. macrostate or emergent) property of the form

$$c_k = \Sigma_n \, f_k(n)p(n).$$

Here a weighted average across the microstates of some function, f of n, is aggregated up to the macrostate level. A continuous formulation using integrals is equally possible. The constraint $f_0(n)=c_0=1$, giving $\Sigma_n \, p(n)=1$ (making a valid probability distribution) is always added. Probably the second most common constraint is $f_k(n)=n$ (giving $c_k = \Sigma_n \, p(n)=\bar{n}=<n>$) meaning that the average value of the ensemble is known. Two other common constraints in macroecology are $f_k(n)=n^2$ (constraint on variance $<n^2>$) and $f_k(n)=\log(n)$, and hence a constraint on the geometric average that is usually considered inappropriate in physics (Montroll and Shlesinger 1982) but possibly with justification in macroecology.

2. Typically the set of possible states {n} (of size N) is much larger than the number of constraints K. In this case, p is not fully defined. Thus, some rule is needed to choose among the set of feasible p-distributions – which has been reduced from the list of all possible distributions by the constraints but which can still be quite large – in a linear world it would have dimensionality of N-K. MaxEnt picks out a specific p by maximizing entropy, $H=\Sigma_n \, p(n)\log(p(n))$ subject to the constraints specified in step 1. It should be noted that while this entropy formula is the same as the Shannon diversity measure these two concepts should not be conflated (i.e. MaxEnt ≠ maximum diversity). Maximization of a function subject to constraints is a common problem in optimization solved by using Lagrange multipliers, λ_k and solving equations where 1st derivatives are set equal to zero (a generalization of the univariate, unconstrained approach to finding a maximum by setting the first derivative to zero). Details can be found in any advanced calculus or optimization text as well as several recent ecology papers (Shipley et al. 2006, Pueyo et al. 2007, Haegeman and Loreau 2008).

3. K+2 equations result (setting equal to zero the first derivatives with respect to n and each of the λ_k – including the sums–to–one constraint). This results in a general solution known as the Gibbs distribution: $p(n)=\exp[\Sigma_k \, \lambda_k \, f_k(n)]/ \, \Sigma_n \, \exp[\Sigma_k \, \lambda_k \, f_k(n)]$. The Gibbs distribution includes the normal, lognormal, exponential and logseries distributions as special cases. To calculate p_n one need only know the λ_k that come from solving the derivatives of the constraint equations set equal to zero. The Gibbs distribution is also found as the shape of the curve used in multivariate logistic regression, providing a nice tie to a common statistical method (He 2010)

As such, MaxEnt can be seen as a balance between deterministic factors (the constraints) and stochastic factors (the entropy maximization). MaxEnt can also be seen as a heuristic for indentifying the minimum necessary set of constraints (forces). Three minor variations on this process are possible and may be a source of confusion.

First, MaxEnt can be applied in either a theoretical or an empirical context. In the theoretical context, the c_k are parameters and the goal is to produce a functional form, p(n), (e.g. the logseries) parameterized by the c_k (i.e. the λ_k are functions of the c_k) giving the probability that a species has n individuals (Alexeyev and Levich 1997, Levich 2000, Pueyo et al. 2007, Dewar and Porté 2008, Harte et al. 2008). In the empirical case, the c_k are observed values, the λ_k can be solved to actual numbers, and it is assumed that the distribution is the multinomial distribution and so the goal is calculate the p_i, giving, for example, the probability that an individual is of species i, or in short the relative abundance of each species (Alexeyev and Levich 1997, Levich 2000, Shipley et al. 2006).

Secondly, instead of simply maximizing entropy one may switch to a Bayesian framework and incorporate a prior distribution (Pueyo et al. 2007). Then one maximizes relative entropy (the increase in entropy for the final p versus the Bayesian prior) known as MaxREnt. In this framework of relative entropy, the maximization of absolute entropy is equivalent to assuming a uniform distribution as a prior. Of course, as in any Bayesian context, this opens a question of what the appropriate prior is. Jaynes argued for a least information prior (Jaynes 1968, Pueyo et al. 2007) but Jeffrey's Priors or informative priors based on actual prior information or expert opinion could also be incorporated. Pueyo et al.

(2007) argue that a lack of prior is appropriate for efforts to predict the relative abundance of specific species (i.e. the empirical case in the previous paragraph such as Shipley et al. 2006), but that attempts to predict the shape of an SAD (i.e. theoretical applications) should use a least information prior. However, Harte et al. (2008) uses no prior for their theoretical application. There is of course no one right answer – it is a decision that must be made based on reasonability and results.

Finally, there are three different justifications given for why one would maximize entropy.

a) The second law of thermodynamics states that entropy will increase over time. The bringing into ecology of this justification requires great caution as it becomes reasoning by analogy. Simply because the universe maximizes the disorder (entropy) of atomic particles over time does not mean that Mother Nature will maximize the disorder of, for example, species in a community. Indeed, one definition of life is 'those systems that locally reverse entropy maximization'.

b) Maximum probability – it can be shown fairly simply using Stirling's approximation that maximizing entropy simultaneously chooses the macrostate that has the highest probability (i.e. has the most possible microstates leading to that macrostate; Pueyo et al. 2007, Haegeman and Loreau 2008). This was the original interpretation of entropy in statistical mechanics, and clearly could be applicable to ecology as a general probabilistic rule.

c) Minimum information – Jaynes (1957) interpretation of MaxEnt was derived from information theory (Shannon 2001, Cover et al. 2006). Increasing entropy means decreasing information. Thus, as Shipley (2009) notes, an information theoretic interpretation is a rule of logic – take the probability distribution that increases the information in p the least amount possible while being consistent with the constraints – in short only add information found in the constraints. Any other rule for selecting p such as taking the center of feasible solution or minimizing entropy or Simpson's concentration (Haegeman and Loreau 2008), while possibly having logic of their own outside of the information framework, are adding information in the technical sense on top of that provided by the constraints.

Very often those using justification c follow Jaynes and use priors and hence MaxREnt, while those using justifications a and b tend to use no prior (equivalently the equiprobable prior) and hence to use MaxEnt.

Given this general machinery for reproducing a distribution of microstates given only a few constraints (or known values) on macrostates, what can MaxEnt bring to ecology? Most obviously, it could reproduce previously known patterns relating to probability distributions, which it does with considerable success (Table 2). To date ecological MaxEnt practitioners have focused most heavily on the species abundance distribution (McGill et al. 2007).

Why do we ignore and reject these mechanisms?

We have provided a summary of six different mechanisms that seem to exist and are being successfully applied in macroecology today but are often overlooked or not recognized as mechanistic. It is interesting (and indicative of the issues this paper seeks to examine) just how strongly negatively some authors react to some of the mechanisms listed above. Examples include strongly negative reactions to lognormality (Williamson and Gaston 2005), neutrality (Clark 2008) and random sampling/placement (Diamond and Gilpin 1982). We think there are two main reasons the above mechanisms keep being rejected or ignored: they are 'too' random and they 'have nothing to do with ecology'.

With regards to the first critique, we ask why random processes that produce pattern should not be considered a type of mechanism? It is interesting to note that stochastic thinking has been taking an increasingly prominent role in ecology, first as noise on a deterministic skeleton (Ludwig 1975, Cushing et al. 1998), and now increasingly as the source of pattern itself (perhaps beginning with MacArthur and Wilson 1967, and reaching a peak with Hubbell 2001). But this has been resisted every step of the way (Clark 2008). We think this bias against stochastic models may be partly driven by adopting the sense of mechanism used by our cousins in biology, molecular and cell biologists, where mechanism is quite explicit, concrete and deterministic – this protein bumps into that protein causing a conformation change and exposing an enzymatic site. But physics has seen a progression from deterministic laws such as those of Newton and Maxwell to stochastic techniques and laws such as those of quantum mechanics and statistical mechanics. Now as molecular and cellular biologists have begun scaling up into systems biology, stochastic models are of increasing importance. Why should ecology be any different? Perhaps the increasing incorporation of stochasticity is a sign of disciplinary maturity. We thus see stochastic models as being just another class of models that should be judged on their ability to elucidate our understanding of the world and to make novel predictions.

Although it is verbalized less often, we think considerable resistance is also based on the "it has nothing to do with ecology" argument. Ecology is relatively unique as a field because many scientists choose their discipline not because they are enamoured with the process of science but because they are enthralled with the organisms they study and the landscapes those organisms live in. Few physicists felt a loss at reducing atoms to randomly colliding abstractions to develop statistical mechanics, but many ecologists feel a loss at reducing organisms to grist for an abstract, general principle like the central limit theorem or maximum entropy. But therein lies an irreconcilable problem: many of the patterns ecologists like to hold up as their own, such as the species abundance distribution, species area relationship or distance decay of similarity are in fact frequently repeated across entirely non-ecological systems such as geology, meteorology, economics, computer science, sociology, and the arts (Gaston et al. 1993, Limpert et al. 2001, McGill 2003a, Mitzenmacher 2004, Nekola and Brown 2007). Are we willing to throw these essential ecological descriptors away as not being truly 'ecological'? Or, should we be willing to accept mechanisms

Table 2. A summary of applications of MaxEnt (or MaxRelEnt) to ecology.

Reference	Prior*	Empirical (E) or theoretical (T)	Constraints	Results
Montroll and Shlesinger 1982, Pueyo et al. 2007	equiprobable	T	$\langle n \rangle$	• exponential distribution
Montroll and Shlesinger 1982, Pueyo et al. 2007	equiprobable	T	$\langle n \rangle$, $\langle n^2 \rangle$	• normal distribution
MacArthur 1955	equiprobable	T	none	• criteria which lead to greater routes of energy flow through a food web
MacArthur 1960	equiprobable	T	$\langle n \rangle$	• broken stick (i.e. exponential) SAD
Alexeyev and Levich 1997, Levich 2000	equiprobable	E, T	resource consumption $\leqslant$ available	• SAD curve, abundances of individual species
Shipley et al. 2006	equiprobable	E	abundance-weighted trait means	• Specific predictions of abundances of specific species
Pueyo et al. 2007	n^{-1}	T	$\langle n \rangle$	• logseries SAD $cn^{-1}exp(-\kappa n)$
Pueyo et al. 2007	n^{-1}	T	$\langle n \rangle$, $\langle \log n \rangle$	• generalized logseries: $cn^{-\beta}exp(-\kappa n)$
Pueyo et al. 2007	n^{-1}	T	$\langle n \rangle$, $\langle \log n \rangle$, $\langle \log^2 n \rangle$	• lognormal SAD
Harte 2008	equiprobable	T	$\langle n \rangle = N_0/S_0$, $\langle e \rangle = E_0/S_0$ (i.e. constraints on average abundance and energy across species)	• logseries SAD distribution of resource use Damuth's law
Harte 2008	equiprobable	T	$\langle n \rangle = n_0 A/A_0$ (i.e. constraints on average abundance across cells)	• distribution of local occupancies, species area curve (requires constraints/results of previous line also)
Dewar and Porté 2008	Π $1/(n+1)$	T	$\langle n \rangle = N$ $\langle nr \rangle = R$ (i.e. constraints on average abundance and total resource use across species subject to a special form of resource use distribution)	• particular SAD (complex, but similar in shape to lognormal) • relationship between abundance and resource use
Banavar and Maritan 2007	Π $1/(n+c)$	T	$\langle n \rangle$	• generalized logseries $c(n+\lambda)^{-1}exp(-\kappa n)$
Phillips et al. 2006, Phillips and Dudik 2008	equiprobable	E	$\langle f(e_i) \rangle = E_i$ (i.e. the average value of some function such as linear, quadratic or threshold of some environmental variable such as average precipitation matches the same average at sites where the species was observed	• species ranges (roughly the probability that a species is found at each site)
He (2010)	equiprobable	T	general	• If richness goes up towards tropics while no. of constraints is constant, evenness must go up in the tropics • the lagrange mutlipliers and maxent distribution are analogous to logistic regression

*an equiprobable prior is essentially equivalent to no prior and hence to MaxEnt while any other prior implies MaxREnt.

that are general enough to be common between all these fields?

Moving forward with mechanism in macroecology?

How should macroecology ecology address the issue of mechanism? We earlier identified four paths. As already noted, we believe all of these are viable and should be pursued. But more generally we think we need to change our mindset about mechanisms. Mechanism has often been used in a very dogmatic and even sanctimonious way: "my work is better than yours because it is more mechanistic", leading to ideological squabbles over what constitutes mechanism. Instead, we suggest a return to the view of Rosenzweig (1991) who espoused a more pragmatic view. Even MacArthur (in choosing entropy over alternatives as the property to be maximized in his analysis of how diversity affects stability) argued "It should be pointed out that choice among

various functions to define stability precisely rests only upon usefulness of the definition" (MacArthur 1960).

In short, mechanism is good when it is useful – no more no less. Specifically, mechanism is good when it improves our understanding or our predictive abilities (likely by enhancing our confidence in extrapolative predictions). Mechanisms are only likely able to improve understanding and prediction when they can be tested and falsified. And just because predictable patterns exist that cannot be mechanistically explained (like the law of gravity) does not make them unworthy of scientific consideration. While mechanisms from any source are good if they're useful, mechanisms developed from some ideologically pure source or direction may have little applicability to the real world. Box 1 summarizes eight viewpoints that we need to adopt in a pragmatic view of mechanism in macroecology.

How do the above six general mechanisms fit the pragmatic criterion? The jury is still out and there is more work to do, but it seems entirely possible that a pragmatic view will embrace at least some of these. It is first important to

---

Box 1 – Dos and don'ts of a pragmatic view of mechanism

1. Do *expect mechanisms to be useful* by: 1) increasing our understanding how the world works, 2) improving our ability to make predictions, especially to extrapolate into new circumstances

2. Do *continue to practice strong inference*. Test and reject proposed mechanisms. Ironically some of the seemingly general mechanisms listed above have been subject to Plattian strong inference (Platt 1964, McGill 2003a) and been tested and found wanting (e.g. fractals, random sampling and placement), which suggests we have made progress on mechanism in macroecology

3. Do seek to *use more sophisticated approaches* to find mechanism including natural experiments, microcosms, long time series (e.g. paleontological data), path analysis, and rigorous null hypotheses.

4. Don't *assume mechanism* is everything or dismiss things that are "only patterns", unless we choose as scientists to also ignore the law of gravity.

5. Don't claim or expect that there *is only one valid domain* of science to produce mechanisms. If a mechanism is useful it is good, regardless of its origin.

6. Don't create a *false dichotomy between ecological and non-ecological mechanisms*. For example the central limit theorem may operate when just four or five forces are in play. The CLT can usefully give a mathematical form while not requiring more detailed exploration into the four or five forces operating. The same with the concordance mechanism.

7. Don't expect patterns that are more general than ecology to have mechanisms that are specific to ecology (e.g. lognormal distributions). On the other hand do expect patterns that are unique to ecology (e.g. more species in the tropics) to be driven by mechanisms specific to ecology. Overall, probability distributions or patterns that relate S, N and A seem to fit general mechanisms well. Patterns that are correlations, especially with organism traits or environmental variables, and dynamics over time series have not been well explained by general mechanisms.

8. Don't exaggerate the success or importance of a reductionist, mechanistic approach in other fields of science (e.g. physics is surprisingly non-mechanistic yet very successful).

9. A pragmatic view of mechanism takes into account the difficulty of measuring the relevant data and parameters. Notably many of these general mechanisms have parameters that are easily measured.

---

note that all six of these mechanisms are testable. They have been measured against empirical data and in many cases found to fit quite well and in a few cases (most often fractals and random placement) found to fit poorly. Secondly, by the very nature of their generality, these mechanisms have the benefit of parsimony. Especially general mechanisms like MaxEnt or regional replacement/dispersal limitation have been shown to lead to unified theories tying together many patterns in a single framework. Probably the biggest test of these mechanisms is whether they are useful. Despite wishes that it not be true, all of these mechanisms explain how the universe works: why should constraints imposed by mathematical logic be excluded from consideration as valid explanations for pattern? And, many of these quite successfully make novel predictions or focus research. The regional replacement suggested the ecologically relevant distinction of transient versus permanent members of communities which has been empirically borne out (Magurran and Henderson 2003, Ulrich and Ollik 2004). Shipley and colleague's (2006) work is predictive down to the abundance of individual species, something rarely achieved in ecology. Harte's MaxEnt work (Harte et al. 2008) has pointed out that many patterns are driven by just a few key state variables (S, N, E for a given area A) and highlighted the importance of further work on understanding the drivers of those few factors. Fractals and modified random placement have been used to make extrapolative predictions to larger, harder to sample spatial scales (Kunin 1998, Kunin et al. 2000, He and Gaston 2003) or to move from presence/absence to abundance (He and Gaston 2000, Conlisk et al. 2009). In short, a pragmatic view of mechanism would embrace these general mechanisms (although not in exclusion to other, more explicitly biological mechanisms). Box 2 lists several important future research directions in general mechanisms.

A pragmatic view of mechanism applied to MaxEnt

We wish to briefly return to the topic of this overall forum and apply the above pragmatic view of mechanism to some of the controversies surrounding the recent uses of MaxEnt in ecology (Shipley et al. 2006, Pueyo et al. 2007, Harte et al. 2008). To date only the Shipley et al. paper (hereafter SVG2006) has been published long enough to evoke a published response, but SVG2006 has already received much commentary (Marks and Muller-Landau 2007, Roxburgh and Mokany 2007, Haegeman and Loreau 2008, Shipley 2009), suggesting the use of MaxEnt will be as controversial as the other general mechanisms examined above. Specifically we think a pragmatic view would note the following things about the MaxEnt controversy to date:

1. *MaxEnt has made some very strong, useful predictions.* SVG2006 predicted the relative abundance of individual species. This has been a surprisingly and disappointingly hard task with many attempts but few consistent patterns (Murray et al. 2002). Similarly Harte and colleagues (Harte et al. 2008) work has to stand as one of the most spectacularly parsimonious and unified theories seen in community ecology to date. From a

pragmatic point of view MaxEnt would appear to be very strong in the prediction front.

2. *MaxEnt is moderately circular but so are most other models*. A major criticism of SVG2006 has been that it is circular – the abundances predicted are derived from abundance weighted trait means. But it should be noted that MaxEnt is not alone in being inherently circular (the c_k are used to predict properties of the system in which they are measured): so are most models. Although desirable, it is very rare to find a model that is parameterized independent of the data it is fitting. Even linear regression fits the line using the data, something that is statistically well accounted for by the idea of degrees of freedom. This is a key issue in MaxEnt: are the degrees of freedom large? Haegaman and Loreau (2008) point out that the degrees of freedom in physics are typically very large (many particles, few constraints) while it is noticeably smaller in SVG2006 (namely nine constraints and 30 variables predicted). Similarly, Roxburgh and Mokany (2007) show that the accuracy of prediction in SVG2006 goes down with increasing degrees of freedom. But this is all to be expected. Ecology has fewer degrees of freedom than physics (less organisms than molecules versus more forces leading to more constraints). And the link between degrees of freedom and goodness of fit is expected. A pragmatic view will assess MaxEnt on whether the degrees of freedom are large enough, especially in relation to other accomplishments in ecology.

3. *Quality of prediction is moderate*. Marks and Muller-Landau (2007) and Haegeman and Loreau (2008) both suggest that the predictive accuracy of SVG2006 is less than the $r^2 = 0.96$ originally claimed when methods such as logarithmic or square root transforms of abundance, cross validation, and RMSE instead of r^2 are used. This is true, but they still find r^2 of 0.30 and up. This should be compared to other attempts to predict abundance in communities where the best r^2's are quite often in the 0.10–0.20 range (Murray et al. 2002, Russo et al. 2003, McGill 2008). A pragmatic view seeks good prediction in comparison to other ecological methods for the question.

4. *MaxEnt is not the best possible predictor* – Several authors show that center of feasible set methods (e.g. Haegeman and Loreau 2008) are better predictors than MaxEnt. This is interesting and possibly important. A pragmatic view would embrace these predictors as better if they prove equally general and easy to apply, although an eye to parsimony might still prefer MaxEnt as containing no hidden assumptions (Shipley 2009). In any case, in a pragmatic view, it is not an invalidation of the basic question and approach, just a refinement in optimization methods.

5. *MaxEnt is unfortunately highly tuneable and sensitive to the exact set of constraints* – probably the greatest caution with regard to the application of MaxEnt is that it is highly tuneable. If one knows the outcome desired, one can choose constraints to produce that outcome (Montroll and Shlesinger 1982). And the outcome is highly sensitive to perturbation – adding a constraint often drastically changes the result. For example, adding a variance constraint to the mean constraint changes the monotonic exponential distribution to the unimodal normal distribution (Table 2), but in ecology who is to say which of these constraints belongs or does not? For example in Harte et al.'s work (2008), removing the energy constraint or adding a stoichiometric constraint on nitrogen or adding variance constraints or adding a prior would all significantly change the resulting predictions. Some of the advantages and interpretations of this fact are highlighted in Pueyo et al. (2007). The empirical approach such as found in SVG2006 is less subject to this issue due to the lack of decision about what constraints to use. But this remains an area requiring careful scrutiny to be sure that strong, novel, general predictions are produced other than just the already known patterns, or else MaxEnt will be just a fancy form of curve fitting (McGill 2003a).

Summary

Macroecology has struggled to find mechanisms to match its success at finding patterns. Broadly, the field of ecology has perceived this as a major failure. We argue, however, that there are already a lot of mechanisms that often get overlooked due to their general nature. We thus recommend a more pragmatic view of mechanism: the goal of science is to explain and predict and anything that furthers that is good regardless of whether it is a pattern or a process or what level of biological organization it comes from. In this view, the overlooked mechanisms are viable and deserving of more attention. And in particular, MaxEnt shows promise while still needing to prove itself, but the measuring stick should be pragmatic and measured in shades of gray and not absolute and ideological.

References

Agrawal, A. A. et al. 2007. Filling key gaps in population and community ecology. – Front. Ecol. Environ. 5: 145–152.

Alexeyev, V. L. and Levich, A. P. 1997. A search for maximum species abundances in ecological communities under conditional diversity optimization. – Bull. Math. Biol. 59: 649–677.

Alonso, D. et al. 2006. The merits of neutral theory. – Trends Ecol. Evol. 21: 451–457.

Alonso, D. et al. 2008. The implicit assumption of symmetry and the species abundance distribution. – Ecol. Lett. 11: 93–105.

Arrhenius, O. 1921. Species and area. – J. Ecol. 9: 95–99.

Banavar, J. and Maritan, A. 2007. The maximum relative entropy principle. – Arxiv preprint cond-mat/0703622.

Barnston, A. G. and Livezey, R. E. 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. – Monthly Weather Rev. 115: 1083–1126.

Bell, G. 2000. The distribution of abundance in neutral communities. – Am. Nat. 155: 606–617.

Blackburn, T. M. 2004. Method in macroecology. – Bas. Appl. Ecol. 5: 401–412.

Blackburn, T. M. and Gaston, K. J. 1998. Some methodological issues in macroecology. – Am. Nat. 151: 68–83.

Borda-de-Agua, L. et al. 2002. Species-area curves, diversity indices and species abundance distributions: a multifractal analysis. – Am. Nat. 159: 138–155.

Brown, J. H. 1995. Macroecology. – Univ of Chicago Press.

Brown, J. H. 1999a. The legacy of Robert MacArthur: from geographical ecology to macroecology. – J. Mammal. 80: 333–344.

Brown, J. H. 1999b. Macroecology: progress and prospect. – Oikos 87: 5–14.

Brown, J. H. and Maurer, B. A. 1989. Macroecology: the division of food and space among species on continents. – Science 243: 1145–1150.

Chave, J. et al. 2002. Comparing classical community models: theoretical consequences for patterns of diversity. – Am. Nat. 159: 1–23.

Clark, J. S. 2008. Beyond neutral science. – Trends Ecol. Evol. 24: 8–15.

Clark, J. S. and McLachlan, J. S. 2003. Stability of forest biodiversity. – Nature 423: 635–638.

Coleman, B. 1981. Random placement and species area relations. – Math. Biosci. 54: 191–215.

Colwell, R. K. and Hurtt, G. C. 1994. Nonbiological gradients in species richness and a spurious Rapoport effect. – Am. Nat. 144: 570.

Colwell, R. K. and Lees, D. C. 2000. The mid-domain effect: geometric constraints on the geography of species richness. – Trends Ecol. Evol. 15: 70–76.

Condit, R. et al. 2000. Spatial patterns in the distribution of tropical tree species. – Science 288: 1414–1418.

Conlisk, E. et al. 2009. Improved abundance prediction from presence–absence data. – Global Ecol. Biogeogr. 18: 1–10.

Connolly, S. R. et al. 2003. Indo-Pacific biodiversity of coral reefs: deviations from a mid-domain model. – Ecology 84: 2178–2190.

Connor, E. F. and Simberloff, D. 1979. The assembly of communities: chance or competition. – Ecology 60: 1132–1140.

Cover, T. M. et al. 2006. Elements of information theory. – Wiley.

Cushing, J. M. et al. 1998. Nonlinear population dynamics: models, experiments and data. – J. Theor. Biol. 194: 1–9.

Darveau, C. et al. 2002. Allometric cascade as a unifying principle of body mass effects on metabolism. – Nature 417: 166–170.

Darwin, C. 1859. On the origin of species. – Clows and Sons.

Dewar, R. C. and Porté, A. 2008. Statistical mechanics unifies different ecological patterns. – J. Theor. Biol. 251: 389–403.

Diamond, J. M. and Gilpin, M. E. 1982. Examination of the "null" model of connor and simberloff for species co-occurrences on Islands. – Oecologia 52: 64–74.

Dolman, A. M. and Blackburn, T. M. 2004. A comparison of random draw and locally neutral models for the avifauna of an English woodland. – BMC Ecol. <http://www.biomedcentral.com/1472-6785/4/8>.

Dunham, A. E. and Beaupre, S. J. 1998. Ecological experiments: scale, phenomenology, mechanism, and the illusion of generality. Experimental ecology: issues and perspectives. – Oxford Univ. Press, pp. 27–49.

Durett, R. and Levin, S. A. 1998. The importance of being discrete (and spatial). – Theor. Popul. Biol. 46: 363–394.

Gaston, K. J. and Blackburn, T. M. 1999. A critique for macroecology. – Oikos 84: 353.

Gaston, K. J. and Blackburn, T. M. 2000. Pattern and process in macroecology. – Blackwell.

Gaston, K. J. et al. 1993. Comparing animals and automobiles – a vehicle for understanding body size and abundance relationships in species assemblages. – Oikos 66: 172–179.

Gaston, K. J. et al. 2000. Abundance–occupancy relationships. – J. Appl. Ecol. 37: S39–S59.

Gilbert, B. and Lechowicz, M. J. 2004. Neutrality, niches, and dispersal in a temperate forest understory. – Proc. Natl Acad. Sci USA 101: 7651–7656.

Gotelli, N. J. and Graves, G. R. 1996. Null models in ecology. – Smithsonian Inst. Press.

Green, J. L. and Ostling, A. 2003. Endemics–area relationships: the influence of species dominance and spatial aggregation. – Ecology 84: 3090–3097.

Green, J. L. and Plotkin, J. B. 2007. A statistical theory for sampling species abundances. – Ecol Lett. 10: 1037–1045.

Green, J. L. et al. 2003. Species richness, endemism and abundance patterns: tests of two fractal models in a serpentine grassland. – Ecol. Lett. 6: 919–928.

Grimmett, G. R. and Stirzaker, D. R. 1992. Probability and random processes. – Clarendon Press.

Haegeman, B. and Loreau, M. 2008. Limitations of entropy maximization in ecology. – Oikos 117: 1700–1710.

Halley, J. M. 1996. Ecology, evolution and 1/f-noise. – Trends Ecol. Evol. 11: 33–37.

Halley, J. M. et al. 2004. Uses and abuses of fractal methodology in ecology. – Ecol. Lett. 7: 254–271.

Hanski, I. and Gyllenberg, M. 1997. Uniting two general patterns in the distribution of species. – Science 275: 397–400.

Harte, J. 2008. From spatial pattern in the distribution and abundance of species to a unified theory of ecology: the role of maximum entropy methods. – Appl. Optimization 102: 243.

Harte, J. et al. 2005. A theory of spatial structure in ecological communities at multiple spatial scales. – Ecol. Monogr. 75: 179–197.

Harte, J. et al. 1999. Self-similarity in the distribution and abundance of species. – Science 284: 334–336.

Harte, J. et al. 2008. Maximum entropy and the state variable approach to macroecology. – Ecology 89: 2700–2711.

Hartley, S. et al. 2004. Coherence and discontinuity in the scaling of species' distribution patterns. – Proc. R. Soc. Lond. B 271: 81–88.

Haskell, J. P. et al. 2002. Fractal geometry predicts varying body size scaling relationships for mammal and bird home ranges. – Nature 418: 527–530.

Hawkins, B. A. and Diniz-Filho, J. A. F. 2002. The mid-domain effect cannot explain the diversity gradient of Nearctic birds. – Global Ecol. Biogeogr. 11: 419–426.

He, F. 2010. Maximum entropy, logistic regression, and species abundance. – Oikos 119: 578-582.

He, F. and Gaston, K. J. 2000. Occupancy-abundance relationships and sampling scales. – Ecography 23: 503–511.

He, F. and Gaston, K. J. 2003. Occupancy, spatial variance, and the abundance of species. – Am. Nat. 162: 366–375.

Helmuth, B. et al. 2005. Biophysics, physiological ecology, and climate change: does mechanism matter? – Annu. Rev. Physiol. 67: 177–201.

Holyoak, M. and Loreau, M. 2006. Reconciling empirical ecology with neutral community models. – Ecology 87: 1370–1377.

Hubbell, S. P. 2001. A unified theory of biodiversity and biogeography. – Princeton Univ. Press.

Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. – Am. Nat. 104: 501–529.

Jaynes, E. T. 1957. Information theory and statistical mechanics. – Phys. Rev. 106: 620–630.

Jaynes, E. T. 1968. Prior probabilities. – Syst. Sci. Cybernetics, IEEE Trans. 4: 227–241.

Johnson, O. 2004. Information theory and the central limit theorem. – Imperial College Press.

Kerr, J. T. et al. 2006. The missing Madagascan mid-domain effect. – Ecol. Lett. 9: 149–159.

Kunin, W. E. 1998. Extrapolating species abundance across spatial scales. – Science 281: 1513–1515.

Kunin, W. E. et al. 2000. Scaling down: on the challenge of estimating abundance from occurrence patterns. – Am. Nat. 156: 560–566.

Lawton, J. H. 1999. Are there general laws in ecology? – Oikos 84: 177–192.

Leitner, W. A. and Rosenzweig, M. L. 1997. Nested species-area curves and stochastic sampling: a new theory. – Oikos 79: 503–512.

Levich, A. P. 2000. Variational modelling theorems and algocoenoses functioning principles. – Ecol. Modell. 131: 207–227.

Levin, S. A. 1974. Disturbance, patch formation, and community structure. – Proc. Natl Acad. Sci. USA 71: 2744–2747.

Limpert, E. et al. 2001. Log-normal distributions across the sciences: keys and clues. – BioScience 51: 341–352.

Lotka, A. J. 1925. Elements of mathematical biology. – Williams and Wilkins.

Ludwig, D. 1975. Persistence of dynamical systems under random perturbations. – SIAM Rev. 17: 605–641.

MacArthur, R. 1955. Fluctuations of animal populations, and a measure of community stability. – Ecology 36: 533–536.

MacArthur, R. 1960. On the relative abundance of species. – Am. Nat. 94: 25–36.

MacArthur, R. H. 1968. The theory of the niche. – In: Lewontin, R. (ed.), Population biology and evolution. Syracuse Univ. Press, pp. 159–176.

MacArthur, R. H. and Wilson, E. O. 1967. The theory of island biogeography. – Princeton Univ. Press.

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

Mandelbrot, B. B. 1963. New methods in statistical economics. – J. Political Econ. 71: 421–440.

Mandelbrot, B. B. 1982. The fractal geometry of nature. – W. H. Freeman and Co.

Marks, C. O. and Muller-Landau, H. C. 2007. Comment on``From plant traits to plant communities: a statistical mechanistic approach to biodiversity". – Science 316: 1425.

Maurer, B. A. 1999. Untangling ecological complexity. – Univ. of Chicago Press.

Maurer, B. A. 2005. Statistical mechanics of complex ecological aggregates. – Ecol. Complexity 2: 71–85.

Maurer, B. A. and McGill, B. J. 2004. Neutral and non-neutral macroecology. – Bas. Appl. Ecol. 5: 413–422.

May, R. M. 1975. Patterns of species abundance and diversity. – In: Cody, M. L. and Diamond, J. M. (eds), Ecology and evolution of communities. Belknap Press of Harvard Univ. Press, pp. 81–120.

May, R. M. 1976. Theoretical ecology. Principles and applications. – WB Saunders Company.

McArdle, B. H. 1996. Levels of evidence in studies of competition, predation and disease. – N. Z. J. Ecol. 20: 7–15.

McGill, B. 2003a. Strong and weak tests of macroecological theory. – Oikos 102: 679–685.

McGill, B. J. 2003b. Does Mother Nature really prefer rare species or are log-left-skewed SADs a sampling artefact? – Ecol. Lett. 6: 766–773.

McGill, B. J. 2006. A renaissance in the study of abundance. – Science 314: 770–771.

McGill, B. J. 2008. Exploring predictions of abundance from body mass using hierarchical comparative approaches. – Am Nat. 172: 88–101.

McGill, B. and Collins, C. 2003. A unified theory for macroecology based on spatial patterns of abundance. – Evol. Ecol. Res. 5: 469–492.

McGill, B. J. et al. 2005. Community inertia of Quaternary small mammal assemblages in North America. – Proc. Natl Acad. Sci. USA 102: 16701–16706.

McGill, B. J. et al. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. – Ecol. Lett. 10: 995–1015.

McQuarrie, D. A. and Allan, D. 2000. Statistical mechanics. – Univ. Science Books Sausalito, California.

Mitzenmacher, M. 2004. A brief history of generative models for power law and lognormal distributions. – Internet Math. 1: 226–251.

Montroll, E. and Shlesinger, M. F. 1982. On 1/f noise and other distributions with long tails. – Proc. Natl Acad. Sci. USA 79: 3380–3383.

Mouillot, D. et al. 2000. The fractal model: a new model to describe the species accumulation process and relative abundance distribution (RAD). – Oikos 90: 333–342.

Murray, B. R. et al. 2002. How plant life-history and ecological traits relate to species rarity and commonness at varying spatial scales. – Austral Ecol. 27: 291–310.

Nekola, J. C. and Brown, J. H. 2007. The wealth of species: eco- logical communities, complex systems and the legacy of Frank Preston. – Ecol. Lett. 10: 188–196.

Palmer, M. W. 1988. Fractal geometry: a tool for describing spatial patterns of plant communities. – Plant Ecol. 75: 91–102.

Peixóto, J. P. and Oort, A. H. 1984. Physics of climate. – Rev. Mod. Phys. 56: 365–429.

Peters, R. H. 1991. A critique for ecology. – Cambridge Univ. Press.

Peterson, A. T. et al. 2007. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Max- ent. – Ecography 30: 550–560.

Phillips, S. J. and Dudik, M. 2008. Modeling of species distribu- tions with Maxent: new extensions and a comprehensive evalu- ation. – Ecography 31: 161–175.

Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – Ecol. Modell. 190: 231–259.

Pielou, E. C. 1975. Ecological diversity. – Wiley.

Platt, J. R. 1964. Strong inference. – Science 146: 347–353.

Pueyo, S. et al. 2007. The maximum entropy formalism and the idiosyncratic theory of biodiversity. – Ecol. Lett. 10: 1017–1028.

Quinn, J. F. and Dunham, A. E. 1983. On hypothesis testing in ecology and evolution. – Am. Nat. 122: 602–617.

Resetarits Jr, W. J. and Bernardo, J. (eds) 1998. Experimental ecology: issues and perspectives. – Oxford Univ. Press.

Ritchie, M. E. and Olff, H. 1999. Spatial scaling laws yield a synthetic theory of biodiversity. – Nature 400: 557–560.

Rosenzweig, M. L. 1991. Habitat selection and population interac- tions: the search for mechanism. – Am. Nat. 137: S5–S28.

Roxburgh, S. H. and Mokany, K. 2007. Comment on" From plant traits to plant communities: a statistical mechanistic approach to biodiversity". – Science 316: 1425b.

Russo, S. E. et al. 2003. Size–abundance relationships in an Ama- zonian bird community: implications for the energetic equiva- lence rule. – Am. Nat 161: 267–283.

Scheuring, I. and Riedi, R. H. 1994. Application of multifractals to the analysis of vegetation pattern. – J. Veg. Sci. 5: 489–489.

Schoener, T. W. 1986. Mechanistic approaches to community ecol- ogy: a new reductionism. – Am. Zool. 26: 81–106.

Shannon, C. E. 2001. A mathematical theory of communication. – ACM SIGMOBILE Mobile Comput. Comm. Rev. 5: 3–55.

Shipley, B. 2003. Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference. – Cambridge Univ. Press.

Shipley, B. 2009. Limitations of entropy maximization in ecology: a reply to Haegeman and Loreau. – Oikos 118: 152–159.

Shipley, B. et al. 2006. From plant traits to plant communities: a statistical mechanistic approach to biodiversity. – Science 314: 812–814.

Simberloff, D. 2004. Community ecology: is it time to move on? – Am. Nat. 163: 787–799.

Smith, F. A. et al. 2008. Macroecology: more than the division of food and space among species on continents. – Progr. Phys. Geogr. 32: 115–138.

Southwood, T. R. E. 1996. The Croonian lecture, 1995: natural communities: structure and dynamics. – Philos. Trans. R. Soc. Lond. 351: 1113–1129.

Tilman, D. 1987. The importance of the mechanisms of interspecific competition. – Am. Nat. 129: 769–774.

Ulrich, W. and Ollik, M. 2004. Frequent and occasional species and the shape of relative-abundance distributions. – Div. Distr. 10: 263–269.

von Humboldt, A. and Ross, T. 1852. Personal narrative of travels to the equinoctial regions of America, during the years 1799–1804. – HG Bohn.

Wallace, J. M. and Gutzler, D. S. 1981. Teleconnections in the geo- potential height field during the Northern Hemisphere winter. – Monthly Weather Rev. 109: 784–812.

West, G. B. and Brown, J. H. 2005. The origin of allometric scaling laws in biology from genomes to ecosystems: towards a quanti- tative unifying theory of biological structure and organization. – J. Exp. Biol. 208: 1575–1592.

West, G. B. et al. 1997. A general model for the origin of allometric scaling laws in biology. – Science 276: 122–126.

Wiegert, R. G. 1988 Holism and reductionism in ecology: hypoth- eses, scale and systems models – Oikos: 53: 267–269.

Williamson, M. and Gaston, K. J. 2005. The lognormal distribu- tion is not an appropriate null hypothesis for the species abun- dance distribution. – J. Anim. Ecol. 74: 1–14.

Wootton, J. T. 2005. Field parameterization and experimental test of the neutral theory of biodiversity. – Nature 433: 309–312.

Zillio, T. and Condit, R. 2007. The impact of neutrality, niche dif- ferentiation and species input on diversity and abundance dis- tributions. – Oikos 116: 931–940.

Zobel, M. 1997. The relative of species pools in determining plant species richness: an alternative explanation of species coexistence? – Trends Ecol. Evol. 12: 266–269.