CHAPTER 13

Ecological Gradient Theory: A Framework for Aligning Data and Models

Gordon A. Fox, Samuel M. Scheiner, and Michael R. Willig

Understanding the heterogeneous nature of species distributions is central to ecology as embodied in the first fundamental principle of its general theory (Table 1.3). As early as the 18th century, it was noted that species richness differed across the globe (von Humboldt 1808; Hawkins 2001). Today, it is widely recognized that species richness changes along a variety of gradients. Some gradients are spatial (e.g., latitude, depth, elevation; Willig et al. 2003), but may reflect underlying or correlated environmental variation (e.g., solar insolation with latitude). In this chapter, we do not discuss gradients from a purely spatial perspective, leaving such consideration to Colwell (Chapter 14). Here, we focus on gradients of species richness that pertain to environmental characteristics (e.g., disturbance, salinity, succession; Grace 1999). Gradients with respect to productivity are probably the most widely discussed of these ecological gradients (Waide et al. 1999). As productivity increases, species richness may increase, decrease, or assume a hump-shape or a U-shape; the pattern may change with geographical or ecological scale (Mittelbach et al. 2001).

Our goal is to further develop a constitutive theory of environmental gradients of species richness as first promulgated by Scheiner and Willig (2005). We expand on that effort by refining the propositions of that theory, revealing hitherto concealed assumptions, and providing a conceptual framework that further unifies seemingly disparate models. We also examine an oft-cited model in detail, show that it is interpreted incorrectly by many, and discuss

approaches for revising it. Besides improved understanding of the particular theory examined here, this exercise illustrates the process of theory development, emphasizing its dynamic nature.

Domain of the theory and its models

The domain of our constitutive theory is environmental gradients in species richness. Literally, *gradient* refers to the slope of a curve; in this case the curve is richness as a function of some environmental characteristic. Slopes range from negative infinity to positive infinity. Although most thinking about these gradients has concerned continuous variation in the environment (so that richness describes a smooth curve with continuous derivatives), there is no logical, biological, or mathematical reason why this must be so. Indeed, one can imagine a limiting case: a threshold in some environmental variable x, such that locations with $x < x_{crit}$ have dramatically lower richness than locations with $x > x_{crit}$. Rather than a smooth curve, the graph will be flat except at x_{crit}, where the change in richness will be immediate and represented by a vertical line. Mathematically, this is described by a step function, which has a slope of 0 everywhere but at the step itself, where the slope is infinite. While real ecological examples are probably less extreme than this, the step function is instructive in making clear that the theory of species richness gradients should account for very sudden changes in richness in ecological space, as well as gradual changes in richness for continuously varying environments. Models used to study richness under continuous environmental variation are likely to take a different form than those used for a small number of discretely different environments; we focus on the continuous case unless stated otherwise.

Ecological gradients occur in spatial contexts, but the theory itself is not necessarily spatial; in its broadest sense, the theory refers to species richness as a function solely of some environmental characteristic. Most intuitively think of ecological gradients as occurring over space, like the gradient from drier to wetter soils that occurs along a hillside. This intuition can be misleading, as the theory encompasses environmental variation occurring in any spatial or temporal pattern, on any spatial or temporal scale. The gradient need not be spatially contiguous or arranged so that the most similar environments are nearest to each other. For example, the theory may apply to a landscape consisting of randomly distributed patches in which environmental characteristics do not show spatial autocorrelation. The pattern of spatial or temporal contiguity and autocorrelation, or lack thereof, can determine which models are appropriate for consideration in any particular situation. The models that

we consider here are not spatially explicit, although spatially explicit versions are possible.

That said, particular models may be relevant only to particular spatial or temporal scales, in that it is likely that different mechanisms (e.g., competition, speciation) will dominate at particular scales. The appropriate scale of a model generally depends on the assumptions of the model itself and on the biology of the taxa under consideration, rather than on an a priori scale. A critical distinction is whether the set of sites or collections of species under consideration draws on organisms from a single pool of species (a metacommunity; Leibold Chapter 8), or from multiple pools. Again, the importance of various ecological processes will differ in these instances. Although the constitutive theory that we describe can apply to gradients at any scale, our focus in this chapter will be models with domains at the regional scale (10s to 100s of km^2) and with mechanisms operating in ecological time. These are the spatiotemporal scales for which most models of environmental gradients in species richness have been developed.

To clarify the relationship between the theory of ecological gradients and spatial issues, it is useful to consider the relationship between gradient theory and species-area theory. Both theories involve predictions of richness as a function of another variable: resource or stressor concentration in gradient theory, and area in species-area theory (Fig. 13.1). Richness is, of course, a function of both area and resources or stressors; however, we cannot yet draw a surface connecting the two-dimensional graphs in Fig. 13.1 unless we assume that there is no interaction between area and resources or stressors. Currently both of these theories only permit limited views of such a relationship: gradient theory predicts richness as a function of resources or stressors for a fixed area (along a single plane slicing the three-dimensional space perpendicular to the area axis). Changing assumptions about the landscape (how patches of different resource levels are arranged in space, relative to dispersal processes) lead to different models under gradient theory. By contrast, species-area theory predicts richness along a single plane slicing the three-dimensional space perpendicular to the resource axis, and also requires assumptions about landscape-level variation. This suggests that a complete theory of species richness may have landscape-related dimensions in addition to the area and resource or stressor dimensions. The identities and number of these axes represents a problem that has yet to be explored.

Although a number of models examine aspects of richness gradients, few have clearly defined the relevant characteristics of the species pool. Models often fail to indicate whether taxonomic or ecological attributes delimit the
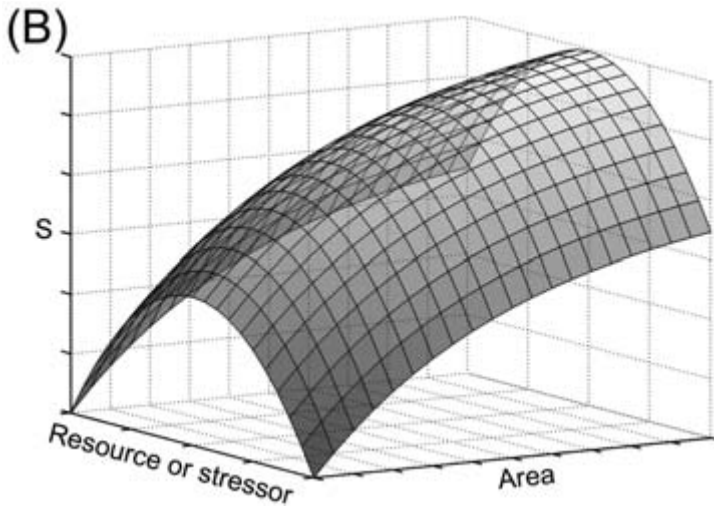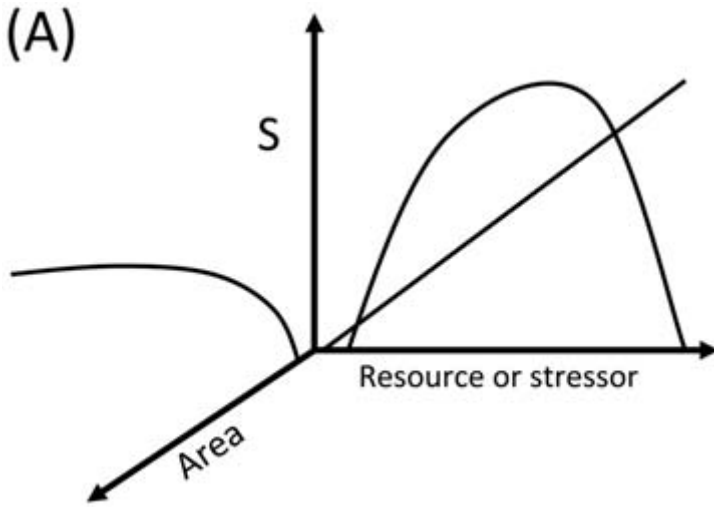
Figure 13.1 Gradient theory and species-area theory both predict species richness (S) by fixing the value of the other axis (area and resource/stressor, respectively) and making landscape-level assumptions. (A) For gradient theory the relationship can take a variety of forms; we depict three here. For species-area theory the relationship is generally assumed to be monotonically increasing or asymptotic. (B) We do not yet know how to draw the surface connecting these graphs in general, because we do not know whether area can interact with the resource or stressor axis, and in most cases would need additional axes for landscape-level variables. The surface shows what the resulting model would look like for a unimodal resource gradient in a uniform landscape without interactions between the axes.

species under consideration. For example, analyses may focus on all species within a clade and a particular level in the taxonomic hierarchy (e.g., a family), a functional guild (e.g., diurnal foliage-gleaning insectivores), an ensemble (e.g., frugivorous bats), or a trophic level (e.g., herbivores). There has been little consideration as to when a model should concern one or the other of these species pools. A species pool may also have a spatial component that is often ill-defined (e.g., biome specific or continental). A related issue is whether there is a single pool or a number of different pools, which again may be strictly a spatial phenomenon (e.g., a gradient that stretches over multiple continents) or be ecological (e.g., a gradient that involves clades that specialize on different conditions).

As we show later, extant models treat species as identical in resource requirements, dispersal ability, and extinction probability. Clearly this violates something that is probably better documented than any other fact in ecology—species differ from one another. In practice, the models make this assumption, but in the literature, the species are assumed only to be roughly equivalent. Most authors refer to these models—or to related species-abundance models—as applying to limited groups such as particular taxa or guilds. For example, herbaceous annual plants might be thought of as roughly equivalent, whereas herbaceous plants and trees are certainly not, because herbs and trees have very different mortality patterns. It is less clear whether herbaceous perennials are roughly equivalent to one another, or whether seed-eating birds are roughly equivalent to seed-eating rodents because they consume the same resources. The exact meaning of "roughly equivalent" requires exploration within particular models. Indeed, the extent to which the assumption of equivalent species can be violated remains a thorny problem for theoretical and empirical research.

Although models within the domain of this constitutive theory are often described as models of species diversity, they are more precisely models of species richness or species density (richness per unit area). Species richness is well defined. In contrast, there are many different definitions of diversity (Whittaker et al. 2001). All involve consideration of species richness, but also include the relative abundances or importances of the species. Some ecologists use the terms richness and diversity interchangeably. In almost all cases, discussions of "diversity gradients" are really discussions of "richness gradients." This is not a semantic argument, as gradients of different aspects of biodiversity (e.g., richness vs. evenness vs. diversity) can be quite different or even independent of each other (Stevens and Willig 2002; Wilsey et al. 2005; Chalcraft et al. 2009).

A theory of environmental gradients of species richness

Our theory rests on four propositions (Table 13.1), set within a conceptual framework (Fig. 13.2). All models of gradients in species richness use the first two propositions, whereas only some include one or both of the last two. These propositions are not universal statements about the world: we do not claim that all propositions hold under all circumstances. Rather, the propositions are statements about the structure of current models of ecological gradients.

Our four propositions are of different kinds. The first proposition is a

Table 13.1 The domain, background assumptions, and propositions that constitute the theory of species richness gradients. Propositions 1 and 2 are used by all models, whereas propositions 3 and 4 are used only by some.

Domain	Environmental gradients in species richness. The gradient can extend over very short spatial distances or be global, or it can extend over short or very long periods of time.
Assumptions	Systems are at equilibrium at some spatial or temporal scale. [most models]
	The species under consideration are roughly equivalent in their resource requirements, dispersal abilities, and extinction probabilities.
	Each species restricts itself more than it restricts other species.
	Local assemblages tend to be in persistent states. [local extinction models]
	The regional species pool contains only species that can coexist with one another. [random placement models]

Propositions

1. A gradient implies one or more limiting resources or conditions that differ in space or time.

2. In a uniform environment of fixed area, more individuals lead to more species.

3. Within an area of fixed size or a unit of time of fixed duration, the variance of an environmental factor increases with its mean.

4. All nonmonotonic relationships require a tradeoff in organismal, population, or species characteristics with respect to the environmental gradient.
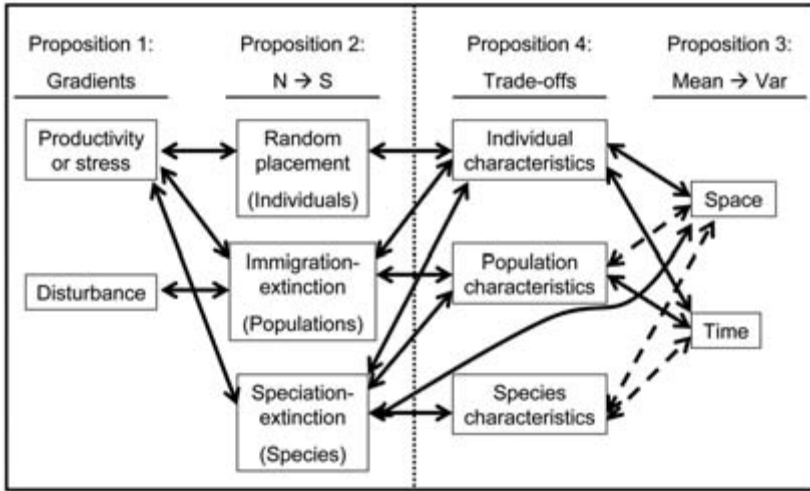
Figure 13.2 A diagram indicating how the four propositions (Table 13.1) can be assembled into different models. The vertical dotted line separates the two propositions (1 and 2) that must be included in any model from the two (3 and 4) that are optional. Solid arrows indicate propositions that have been linked in at least one model. Not all possible combinations of linkages appear in current models. Of the 43 possible models based on unique combinations of linkages, only 9 have been developed to date. Dashed arrows indicate linkages that have not been made. Additional linkages could be developed between propositions 2 and 3, but are not included in the diagram for clarity. The absent linkages between propositions 2 and 4 may not exist because of incommensurate timescales; however, we do not preclude the development of such linkages.

definition, establishing the essence of an environmental gradient in abundance. The second proposition encompasses several mechanisms that can be derived from first principles, each of which assures that the environmental gradient in abundance is also a gradient in richness. The third proposition is a description of a common empirical pattern, or is a general statement about ecological variation. The fourth proposition comprises a heterogeneous mix of mechanisms that derive from other domains and theories that influence the nature of environmental gradients in species richness.

The scale of the data or the model, including aspects of grain and extent, determine the particular mechanisms in effect for each of the propositions. Propositions 1 and 2 (Table 13.1) are functions of the extent, the range of environmental conditions encompassed by the data or being described by a model. Proposition 3 is a function of grain, the sizes of sampled patches or local communities. These scales are always determined by the biology of the

species under consideration. A failure to recognize the scales within which particular mechanisms operate has led to a misapplication of a much-cited model (Wright 1983) that has been used to explain global species richness gradients (see below). Recognizing such misapplication is an example of how the process of theory formalization (as illustrated throughout this book) can provide critical insights and guide future research.

Gradients

Our first proposition is that variation characterizes a limiting environmental factor X, which affects variation in the number of individuals that can persist in a sample location of a particular size, thus creating an environmental gradient in abundance. The abundance gradient exists in space and time, although the environmental factor need not be autocorrelated spatially or temporally. Models typically consider only one aspect—time or space—rarely both.

This proposition is part of all models of gradients in species richness, but it is often implicit. The environmental factor could be the concentration of one or more resources, or some condition such as stress or disturbance. For the purposes of our presentation, the exact mechanism creating the link between number of individuals and the factor(s) X does not matter and will differ for each particular situation. Importantly, not all environmental variation creates variation in numbers of individuals, thereby constraining the domain to which our theory applies. In particular, the modifier "limiting" implies that the value of X at a site determines, at least at equilibrium, the number of individuals present, $N(X)$. Most models assume that the system has approached some sort of long-term behavior, i.e., an asymptote or a dynamic equilibrium.

Careful consideration of this proposition helps to clarify limitations regarding the scope of particular models. If variation in X leads to variation in N, then we can write $N(X)$ as a function predicting the equilibrial or asymptotic number of individuals. This requires that the model be general—it predicts the long-term number of individuals that can persist at a particular level of X, not the number or identities of species found at a particular instance of X. The individuals are therefore assumed to be identical in key ecological respects, such as physical traits like body size or demographic traits, insofar as these reflect resource use. This assumption holds only for those key characteristics. For example, models that invoke niche partitioning (Hutchinson 1959; Schoener 1974; Chesson and Huntly 1988; Leibold 1995; Rosenzweig 1995; Chase and Leibold 2003; Kelly and Bowler 2005) assume that species are equal only in resource use.

Although the mathematics in the current literature use the equality assump-

tion, the models are universally interpreted as applying to cases in which the species under consideration are roughly similar but not identical. Under this relaxed assumption, species must be sufficiently similar to one another that, to a rough approximation, individuals in one species require the same amount of resource, or respond in a similar fashion to some condition, as do those of any other species in the species pool. This could easily be generalized to allow for an equivalence among species (e.g., 1 individual of species A equals 1.7 individuals of species B). Thus, a model might describe gradients in the richness of granivorous rodents, but cannot be expected to describe gradients for all vertebrates or even all mammals. It might be reasonable to develop a model that describes a richness gradient for herbaceous C_3 plants with respect to some environmental characteristic. But because "all plants" includes organisms with profoundly different metabolic pathways that span seven or eight orders of magnitude in dry weight, a single model likely will not describe the gradient in richness of all plants.

Model construction is simple enough when the environmental factor is a single resource, but becomes more complicated if multiple resources govern variation in abundance. If the same resource is limiting for all species at each location or time, the minimum (limiting) resource dominates (Liebig's law of the minimum; Sprengel 1839; van der Ploeg et al. 1999). If two or more resources (say, water and light availability for plants) are both limiting (either simultaneously or each one limiting at different times or sites), the combination can be quantified by the vector $\vec{X}$. If the resources affect abundance additively, then the vector $\vec{X}$ can be treated as a single resource X_c, a linear combination of the multiple resources.

Thus, gradients can be grouped into two general classes. In the first class are systems constrained by a single factor X or a combination of factors $\vec{X}$, and $N(\vec{X})$ increases monotonically with the linear combination X_c. In the second class are systems in which one or more pairs of limiting factors are negatively correlated. The negative correlation could be intrinsic (e.g., as soil water content increases, oxygen levels in the soil must decrease), or could be extrinsic to the factors themselves.

Our description of the first class as following a single constraint needs further explanation because circumstances can be more complicated. Multiple resources can interact, so that more than one is limiting at a particular time or place (Gleeson and Tilman 1992). In principle, the only difference is that instead of the number of individuals $N(\vec{X})$ being a curve, it will be a surface with the number of dimensions equal to the number of factors that are limiting at some point. Consider a system in which two resources—each limiting at different concentrations—vary in a nonlinear fashion with respect to each other.

Then it is possible for the system to switch between limitation by factor 1 and factor 2 more than once. Such is not a problem for the models, but in practice it may not be easy to determine what is limiting at each particular location. Indeed, most empirical studies do not address this concern. In practice, most studies examine only single factors (Scheiner and Willig 2005).

Finally, when gradients involve tradeoffs (proposition 4), it is generally useful to separately consider two different aspects of the environmental factors, one relating to resource attributes and the other to stress attributes. For example, consider communities of herbaceous plants arrayed from upland to wetland. Water may be a limiting resource at the higher elevations, but a stressor at lower elevations. It can be useful to model the responses of richness with the associated attributes as two separate factors because the biological responses to water as a resource and water as a stressor are different.

Many models assume that $N(\vec{X})$ is a linear function, although the critical nature of this assumption has not been explored in a comprehensive fashion. When a single factor is limiting, the important assumption is that $N(X)$ is monotonic. It is often possible to select a transformation (e.g., the log function) to linearize a monotonic pattern. However, multiple limiting factors that act singly or interactively create complications: if the factors are interactive, $N(\vec{X})$ may not be monotonic, and there may not be a transformation that will linearize the function.

One common instance of multiple factors is when richness is regulated by bottom-up and top-down interactions. If X is a resource (such as prey items) or an abiotic stressor, regulation is bottom-up; if X is predation, regulation is top-down. Because both top-down and bottom-up regulation occur in many systems, many models consider both kinds of factors in producing gradients of richness.

Individuals ∝ species

The proposition that the number of species increases with the number of individuals was developed by Fisher et al. (1943) and Preston (1962a). It has been termed the "more individuals hypothesis" (Srivastava and Lawton 1998), although it is not necessarily a hypothesis. Under random placement it is a simple sampling relationship, but if extinction or speciation mechanisms are involved in creating the richness gradient, it is indeed a hypothesis. Three mechanisms can account for this pattern: random placement, local extinction, and speciation. Random placement and local extinction are modeled using similar mathematical constructions but are distinct in their biological causations. Moreover, they operate at somewhat different scales of time and space.

Given an environmental gradient in the number of individuals, each of these mechanisms can lead to an environmental gradient in the number of species. All models of species richness gradients invoke at least one of these mechanisms, at least implicitly. In many cases, the models focus on causes of gradients in the number of individuals and assume a mechanism whereby more individuals give rise to more species.

Random placement (also called passive sampling) refers to the movement of individuals among patches or communities. It creates a relationship between the number of individuals and the number of species if local species richness is determined by random sampling of individuals from a regional species pool (Coleman 1981; Coleman et al. 1982) or metacommunity (Hubbell 2001). Biologically, this occurs as individuals move independently but tend to concentrate in areas of greatest resource or least stress. The assumptions behind random placement models are thus identical to those leading to an "ideal free distribution," an idea that has played a critical role in behavioral ecology (Fretwell and Lucas 1970; Sih Chapter 4). As the number of individuals in a local area increases, the number of species should increase because the likelihood of including a rare species increases due to chance. Under this model the species identity of each individual is random, but the number of individuals in a local assemblage is not—it is given by the function $N(X)$. That the relationship between local and regional richness is positive and monotonic does not depend on the abundance distribution locally or regionally, although those distributions determine the exact form of the relationship.

Local extinction is the mechanism invoked by Preston (1962a; 1962b) and highlighted most often as part of the MacArthur and Wilson (1967) theory of island biogeography (Sax and Gaines Chapter 10). This mechanism assumes that a local population will persist only above some minimum abundance. If an area holds more individuals, more populations can attain species-specific minimum viable sizes. Although details can differ about the exact form of the relationship between numbers of individuals and numbers of species, the core assumption is simply that the relationship is positive and monotonic.

Random placement and local extinction share some features: both operate through a balance between the entry of individuals into a site and their departure—by movement under random placement, and by death in extinction. That entry and exit could occur within the lifetimes of individuals through movement, or it could occur across generations through colonization and extinction. For convenience we divide this continuum into an individual-level mechanism (random placement) and a population-level mechanism (local extinction). Mathematically, they can be treated as equivalent for an equilibrial theory, although models will differ in detail depending on the particular set of

species under consideration. In general, random sampling operates at local to landscape scales over short time periods, whereas local extinction operates at landscape to regional scales over longer time periods, with the exact meaning of these differences determined by the species' biology.

This distinction between individual- and population-level processes can define the domain of a particular model. Previously, we asked whether seed-eating birds and seed-eating rodents are roughly equivalent. If the abundance of birds at sites is determined by the movement of individuals (e.g., Coleman et al. 1982), while the abundance of rodents is determined by population growth and extinction, then application of existing models to, say, richness gradients of granivorous vertebrates would be misleading. Because the entry and departure processes are quite different for granivorous birds and rodents, involving different parameter values, one would need to model the richness of the two groups separately. If interactions between the two groups could be ignored, then the predicted richness of the combined taxa would simply be the sum of the two predicted richnesses. However, if granivorous birds and rodents interact, studying the richness of both would require modeling their interactions as well.

Speciation operates at scales of time and space that are much greater than that of random placement or local extinction. It assumes a positive relationship between the number of individuals and the net rate of speciation (i.e., speciation minus extinction; VanderMeulen et al. 2001). This mechanism most appropriately deals with species richness patterns at large spatial scales and may provide an explanation for the richness of the regional species pool.

All extant models of richness gradients make another important assumption: *they do not consider species interactions*. To see this, consider a gradient model derived from one of the versions of neutral theory (Hubbell 2001; Chave 2004; Etienne and Olff 2004; Volkov et al. 2005). A model using either random placement or local extinction can logically find the expected number of species at any location along the gradient, using only propositions 1 and 2 (Table 13.1). Now consider a gradient model that concerns niche partitioning, using character displacement, microhabitat variation, or temporal niches (e.g., Hutchinson 1959; Schoener 1974; Chesson and Huntly 1988; Leibold 1995; Rosenzweig 1995; Chase and Leibold 2003; Kelly and Bowler 2005). This model can also find the expected number of species, given one of the following assumptions:

1. Local assemblages generally have reached persistent states. By persistent we include the textbook equilibria of Lotka-Volterra competition models, as well as the more complex kinds of persistence possible with multiple species and nonlinear interactions (Armstrong and McGehee 1980). This assumption

implies that local population and community dynamics dominate, so logically this assumption might hold under a local extinction model, but not under a random placement model.

2. The regional species pool contains only those species that can coexist with one another. This assumption could hold under random placement, which implies nothing about the long-term persistence of competitors. It could also hold under the (rather unlikely) assumption that the species in the regional pool have all coevolved to coexist.

3. Competitive coexistence of particular species in a given sample unit is not necessarily guaranteed, but sampling processes still guarantee that on average there will be $S[N(X)]$ species at X resource level.

Having said this, we hasten to add that it is not logically necessary for community theory (Holt Chapter 7; Leibold Chapter 8; Pickett et al. Chapter 9) to be external to gradient theory—just that this is presently the case. It is certainly possible for local community dynamics to interact with the factors determining the existence of a richness gradient; addressing this possibility is an open theoretical question.

Mean ∝ variance

The proposition that the mean and the variance of environmental characteristics are related positively is based, in part, on the recognition that most environmental factors are bounded by 0 (i.e., have a theoretical minimum). Such a bound can lead to a positive relationship between the mean and variance, although such a relationship need not exist empirically. If the magnitude of an environmental factor is 0 or close to 0, then perforce the variance initially will increase as the mean increases. A continued rise in the mean allows for the possible continued rise in the variance, unless an upper bound also exists (e.g., water saturation of soil). Thus, this proposition is limited to those environmental variables that have a lower but not an upper bound within the range of environmental conditions of the gradient. If the upper bounds on a limiting environmental factor also restrict the number of individuals, the theory as described in this chapter can be applied to that part of the environmental gradient where the variance does increase with the mean.

This mean-variance relationship is invoked in models that focus on patch dynamics (e.g., Abrams 1988). More specifically, species richness is measured in some area within which there are multiple patches. For some models, heterogeneity is generated by interactions among individuals (e.g., Tilman 1982; Huston 1994; Currie et al. 2004). Most commonly, the invoked mean-variance relationship is spatial (e.g., wet vs. dry, good vs. bad). A meta-analysis

(Lundholm 2009) found that plant species richness or diversity frequently increases with spatial heterogeneity, but does not always do so. Some models invoke temporal heterogeneity, considering specialization on different year-types as a mechanism that promotes coexistence of multiple species (Chesson and Huntly 1988; Rosenzweig 1995; Kelly and Bowler 2005). Regardless, if species specialize on combinations of environmental characteristics that occur in patches in which they can out-compete other species, then richness should increase as the number of patches (i.e., heterogeneity) in an area increases.

This proposition is explicitly scale-dependent (Lundholm 2009) as it deals with changes in variation within the grain of a particular model, the unit for which richness is measured. This grain is always dependent on the biology of the species under consideration. Thus proposition 2 also contains a hidden assumption that the species are equivalent in their use of space or time. At the lower end, the minimal grain size is that needed to hold one individual. At the upper end, the maximal grain size is such that all possible heterogeneity or habitat types are encompassed within a single grain.

The form of the relationship between mean patch characteristics and their variance is related to theories of species-area relationships (SARs; Fig. 13.1). SARs are determined by a variety of factors: more individuals are contained in larger areas, and environmental heterogeneity increases with greater area. Clearly, models of SARs share many features with models of species richness gradients. Models of SARs are currently being developed and debated (e.g., Scheiner 2003; Tjørve 2003; Maddux 2004; Ostling et al. 2004; Adler et al. 2005; Fridley et al. 2006; Chiarucci et al. 2009). Thus, we postpone any attempt to develop formal models of SARs specific to the context of species richness gradients until the more general forms of those models have been resolved more thoroughly.

Tradeoffs and hump-shaped curves

Many models of environmental gradients in species richness posit that a tradeoff leads to a hump-shaped pattern, with the maximum value of richness at some intermediate point along the axis of an environmental factor. The models differ with regard to the nature of the invoked tradeoff. Nonetheless, they share the basic proposition that a change in the sign of the slope arises as a consequence of two mechanisms acting in concert but in an opposite fashion on each species. Commonly invoked tradeoffs are competitive ability versus a variety of other abilities (e.g., stress tolerance, colonizing ability). The tradeoffs that matter in a particular instance depend on the species and type

of environmental variation, including its scale (grain and extent) in time and space.

Tradeoffs may occur at different levels in the biological hierarchy. For example, the tradeoff may involve the characteristics of individuals, such as competitive ability versus stress tolerance (Grime 1973). In other cases, the tradeoff may involve the characteristics of populations, such as the intensity of interspecific competition versus the intensity of predation (Oksanen et al. 1981). In yet other cases, the tradeoff may involve characteristics of species, such as speciation rates versus extinction rates (VanderMeulen et al. 2001). Scheiner and Willig (2005, Table 1) listed 17 different models of species richness gradients. In the conceptual scheme presented here (Fig. 13.2), we treat mechanisms that operate at the same level (i.e., individual, population or species) as mathematically equivalent. In doing so, we can unify some of those models, reducing the list of models from 17 to 9 (Table 13.2).

The maximum (or minimum) point in the curve describing an environmental gradient in species richness arises because of a change in the relative importance of factors that control the number of individuals. This tradeoff can be conceptualized as environmental variation in each of two factors that are negatively correlated. Along one portion of the environmental axis, the first factor limits the number of individuals; at some point a second factor becomes limiting. This switch results in the number of individuals increasing along one portion of the environmental axis and decreasing along another. For many models, this shift in importance is controlled by inherent properties of species. For example, Tilman (1988) theorized that in terrestrial plant communities increasing nitrogen availability causes an increase in numbers of individuals, until plant density is great enough that light becomes limiting and numbers of individuals begin to decrease. Although tradeoffs are invoked in models that produce a hump-shaped pattern, the mechanism can explain U-shaped patterns as well (Scheiner and Willig 2005).

In many models, the interacting mechanisms that determine the number of individuals are not stated explicitly. Similarly, the unique contributions of each mechanism to total abundance are rarely quantified with respect to variation in the environmental factors. As a result, the mechanistic tradeoff is neither emphasized in conceptual models nor detailed in quantitative models. The absence of mathematical or logical rigor enhances the likelihood that such concealment persists, diminishing an appreciation for the similarities of form that the details obscure. For example, various models posit tradeoffs between competition for different resources (e.g., Tilman 1982; 1988; Huston 1994) or competition vs. resistance to predation/herbivory (Leibold 1996; 1999).

Table 13.2 Models of diversity gradients and their components and mechanisms. Proposition 1: Type of gradient; Proposition 2: Mechanism linking the number of individuals and number of species; Proposition 3: Environmental heterogeneity; Proposition 4: Type of tradeoff. Previous number(s) refers to models listed in Table 1 of Scheiner and Willig (2005).

No.	Proposition 1	Proposition 2	Proposition 3	Proposition 4	Previous number(s)	Sources
1	Productivity or stress	Random placement	N/A	Individual characteristics	1	Oksanen (1996), Stevens (1999)
2	Productivity or stress	Local extinction	N/A	N/A	5	Connell (1964), Wright (1983)
3	Productivity or stress	Local extinction	N/A	Individual characteristics	11	Rosenzweig (1993), Tilman (1993)
4	Productivity or stress	Local extinction	Spatial	Individual characteristics	3, 7, 10	Tilman (1982; 1988), Abrams (1988), Huston (1994), Leibold (1996; 1999)
5	Productivity or stress	Local extinction	Temporal	Individual characteristics	4	Rosenzweig (1995)
6	Productivity or stress and disturbance	Local extinction	Temporal	Individual characteristics	6, 8	Grime (1973; 1979), Huston (1979), Huston and Smith (1987)
7	Productivity or stress	Local extinction	N/A	Population characteristics	2, 9	Rosenzweig (1971; 1995), Wollkind (1976), Oksanen et al. (1981)
8	Productivity or stress	Speciation	N/A	Individual characteristics	12, 15	Denslow (1980), Rosenzweig and Abramsky (1993), VanderMeulen et al. (2001)
9	Productivity or stress	Speciation	N/A	Species characteristics	13, 14	VanderMeulen et al. (2001)

Such models all have a similar mathematical form, but this similarity is not apparent until they are placed within a single framework.

Arguments for and against particular models often boil down to a personal preference for one tradeoff over another. We take a more catholic position by not advocating any one in particular. Rather, we embrace all of them as theoretical possibilities, although it remains to be seen whether some tradeoffs are more common than others. Perhaps most critically, the posited mechanisms often are not mutually exclusive. Tradeoffs may simultaneously exist between competition for two different resources and herbivory, for example. As with multiple environmental factors, it may be possible to model such multiple tradeoffs as an additive pair of tradeoffs. Otherwise, more complex models will be needed.

Although a specific tradeoff may exist for a particular set of species, we should not expect the same tradeoff to be ubiquitous for all species in a guild, trophic level, or community, thus limiting the scope of any particular model. It is possible that more closely related species will share a tradeoff, whereas more distantly related taxa will have different constraints, but this should not be assumed (Losos 2008). Thus, the type and form of tradeoffs sets another boundary on the conditions under which individuals of different species must be roughly equivalent. It is not known how rough this equivalence can be and still be consistent with the underlying models.

Relationship to the theory of ecology

The four propositions of the theory of environmental gradients of species richness (Table 13.1) derive from the fundamental principles of the theory of ecology (Table 1.3). Proposition 1 is a consequence of principles 4 or 5, depending on the nature of the environmental factor(s). The finite nature of resources (principle 5) creates the constraint that allows one or more resources to be limiting. The heterogeneity of environmental characteristics in space or time (principle 4) creates the potential for variation in resources or stressors. Environmental heterogeneity in time leads to the potential for variation among patches in the rate of disturbance. Proposition 2 is a consequence of principles 1, 2 or 7. The process of random sampling is one mechanism that creates the heterogeneous distribution of organisms (principle 1). Immigration-extinction balance comes about through the combination of processes that lead to heterogeneous distributions or organisms' and species' interactions (principle 2). Speciation is a suite of processes that derive from principle 7. Proposition 3 is a direct manifestation of environmental heterogeneity (principle 4). Finally, the tradeoffs embodied in proposition 4 derive from principles 6 and 7.

Applying the theory: the energy model

To see how the constitutive theory relates to current models of productivity-richness relationships, consider the energy model (Connell and Orias 1964; Wright 1983). We focus on this model for two reasons: (1) it has been very influential [we found 343 citations of Wright (1983) in the Web of Science database on February 16, 2009], and (2) it is one of the few that is written in explicit mathematical form. We follow the formal presentation of Wright (1983), which is couched in terms of the relationship of species richness and area. The model predicts the number of species in a sampling unit (Wright thought of these as islands) as a function of local energy availability. Wright considered energy input per unit area to be fixed so that his model predicts the consequences of variation in area on species richness. By contrast, the models considered in this chapter examine the consequences of variation in environmental resources or stressors among different locations while holding area constant. Thus, our explication of this model does not include terms for area as in Wright (1983).

Wright's model is $S = a(E\rho/m)^z$, where E is the amount of energy locally available for biosynthesis, ρ is an empirical constant for a given set of species describing the number of individuals that are supported per unit of available energy, and m is the population size of the smallest extant population. The terms a and z are empirical constants estimated from the data, although as we shall see, a, z, and m appear in this model because of some strong assumptions. We now examine how this model relates to our propositions, and consider some consequences of its assumptions.

Existence of a gradient and its consequences

Rewriting Wright's model in our more general terms, we begin with $N(X) = X\rho$. The use of the common term ρ means that the model describes richness when derived from a set of roughly similar species, as our explication of the general theory suggests it must. Wright further posited that species richness increases with decreasing latitude because available energy increases, a contention still advanced by many (e.g., Mittelbach et al. 2001; Hawkins et al. 2003a). This is a sensible model only if individuals of all species along the latitudinal gradient require about the same level of resources, which is certainly not true.

From individuals to species

Proposition 2 posits that more individuals lead to more species, and in Wright's model, most of the action is in proposition 2. To model species richness as a

function of X, we need to model $S = f[N(X)]$, where S is the number of species present and f is some function. Wright's choice of f is $f[N] = a(N/m)^z$, which comes from Preston (1962a). This equation, with $N = X\rho$ as above, produces a positive monotonic relationship between the resource X and species richness S. The exact shape of the relationship depends on a, z, and m. The model requires the first two propositions (Table 13.1), and nothing more. The energy model does not attempt to explain the source of available energy or its relationship to climate, which is the domain of other theories (e.g., O'Brien et al. 2000).

The Wright and Preston models rely on the local extinction mechanism, delimiting the temporal and spatial scales for which the model makes predictions. In particular, this model makes predictions about the equilibrial number of species at a location with resource concentration X, when individuals are drawn from a fixed regional pool of species. The model should not be interpreted as making predictions about variation in species richness over large spatial extents (e.g., across continents) because such variation cannot result from local extinctions from a single species pool. Such large-scale gradients must involve (at least) several species pools, and likely involve speciation processes as well. Thus, by its implementation of both propositions 1 and 2, Wright's model has a far more limited interpretation than stated by Wright or many subsequent authors (e.g., Currie 1991; Mönkkönen and Viro 1997; O'Brien 1998; Gaston 2000; Allen et al. 2002; Currie et al. 2004). Although some found apparently good fits of the model for continental- to global-scale data, because those data represent an inappropriate spatial domain, it is illogical to assign meaning to estimated parameters in terms of the Wright model.

Although the power-law function used by Preston and Wright is simple and familiar to several generations of ecologists, its derivation in this case rests on a complex and rather narrow argument concerning the distribution of species abundances and how population size relates to extinction probability. In particular, the power-law function depends on Preston's assumption that species-abundance curves are described by a form of the lognormal distribution that he termed canonical.

The division by m—not explained by Wright other than his citation of Preston (1962a)—seems odd. It is natural, albeit wrong, to assume that this division (N/m) is aimed at calculating the maximum number of species. Under Preston's (1962a) canonical lognormal distribution, one specifies the shape of the species-abundance distribution with any two of three quantities: the total number of species, the standard deviation of the lognormal distribution, and the number of species in the modal octave. The quantity m is required to specify the position of this distribution along the horizontal axis (the $\log_2$ of

abundance). Preston (1962a) calls m the size of the smallest population, but he also calls this a "tentative" definition (Preston 1962a, p. 190), and notes "in practice that m is less, even appreciably less, than unity, and the temporary interpretation we have given [as the size of the smallest population] then has no meaning." In other words, m is just a parameter that defines the location of the species-abundance distribution, in the same sense that statisticians speak of the mean as characterizing the location of the normal distribution.

Regardless of whether m is the size of the smallest population or an empirically estimated parameter, Wright's use of m, a, and z to defines his model links it intimately to the somewhat arbitrary assumptions of Preston's canonical lognormal distribution. Despite numerous criticisms of aspects of Preston's work (e.g., Pielou 1969; May 1975; Williamson and Gaston 2005), it has had remarkable staying power in the ecological literature: remarkable, because neither Preston nor subsequent researchers have linked the canonical lognormal to any underlying mechanisms. Preston himself (1962a) made it clear that he had none in mind. Unfortunately, ecologists are sometimes satisfied with curve-fitting exercises without concern with the underlying mechanisms. Such exercises teach us nothing beyond the narrow lesson that the particular data set is well-described by a particular curve, providing only a phenomenological description.

Most ecologists, trying to justify the use of the lognormal for species abundance distributions, do so with a vague and incorrect reference to the central limit theorem (Williamson and Gaston 2005). The central limit theorem predicts that each species' abundance will be lognormally distributed over time; unless the abundances are independently and identically distributed among species (i.e., the species are equivalent), this does not lead to a jointly lognormal distribution of abundances at a particular time. If the species are different from one another (i.e., they have a different means and variances of abundance), the joint distribution of abundances at a given time will not be lognormal. Šizling et al. (2009) proposed a more satisfying (and rigorous) explanation as to why species abundance distributions are often similar to the lognormal. Their derivation requires only that the abundance distribution be based on the combination of abundances in many nonoverlapping subplots.

None of this implies that the Wright energy model is wrong in some sense; rather, its basis is weaker than one might hope (given its influence), as it depends on the phenomenological assumption that species abundances are given by Preston's canonical lognormal. Other mechanisms could be invoked that yield the same qualitative relationship while differing in details (e.g., Hubbell 2001). Pueyo et al. (2007) showed that an infinite number of models varying between strict neutrality (all species identical) and strict idiosyncrasy (all

species unique) can generate identical abundance patterns. Although Wright's model has been interpreted as a predictor of continental to global patterns, it cannot logically be so as it is restricted to a set of roughly equivalent species (implementation of proposition 1) in a single regional pool (implementation of proposition 2).

The converse is also true. Many studies have shown a positive relationship between energy and species richness on a continental to global scale (e.g., Field et al. 2005; Rodriguez et al. 2005; Buckley and Jetz 2007; Davies et al. 2007; Kalmar and Currie 2007; Kreft and Jetz 2007; Woodward and Kelly 2008), and this has often been taken as support for Wright's model. However, those studies do not attempt to directly parameterize Wright's model and test whether the model is accurately predicting those relationships. Instead, we merely have a qualitative agreement between various empirical relationships and one particular model. Our dissection of that model suggests that it cannot be used as an explanation for those relationships because the mechanisms underlying that model operate at different scales. Given the generality of the observed relationships, further work is necessary to connect the mechanisms of Wright's model with global-scale mechanisms, or to develop new models with mechanisms operating at that scale.

Our explication of Wright's model suggests that it must be interpreted on a regional spatial scale with species that are roughly equivalent. That does not preclude the possibility that one might find that the model provides a good fit to data from much larger spatial or taxonomic scales. Indeed, if a model like Wright's provides a good prediction of the number of species, given a level of resources, it may be useful to managers even if its assumptions are violated severely. The only problem here is with interpretation: a good fit of a model to data that violate its assumptions cannot be interpreted as support for the concepts embodied in the model, but only as a useful description of data. Prediction and understanding are not always on the same footing.

Prospects

Model development

Theory unification is an iterative process that includes recognition of similarities among ostensibly competing models, development of a common framework, and construction of new overarching models within that framework. Additional effort is needed in domains, such as the one we consider, in which many of the models are verbal and even the analytic models have not been examined deeply. We are encouraged that our refinement of the conceptual

framework (Fig. 13.2) has led to further model unification (i.e., reducing the number of models from 17 to 9).

This is a step forward in model unification not simply because it reduces the number of models, but because it reveals their common bases, and because it points to some additional models that have not yet been studied (Fig. 13.2). The reduction in the number of models is a consequence of recognizing that the 14 different forms for proposition 4 listed in Scheiner and Willig (2005, Table 1) can be usefully placed in three categories: tradeoffs operating at the levels of individuals, populations, and species. For example, using the model numbers from Scheiner and Willig (2005, Table 1), we now treat models 3, 7, and 10 as equivalent because all assume that the gradient (proposition 1) is productivity or stress, the number of species (proposition 2) is generated by local extinction processes, heterogeneity (proposition 3) occurs over space, and tradeoffs (proposition 4) occur at the levels of individual characteristics. Similar reasoning leads to unifying other models.

Our approach has been to start with the simplest formal model, the Wright energy model, and carefully examine its assumptions and limitations. The challenge is to build a new, general and useful model that avoids the previously described limitations. The first limitation—restriction to a set of species with roughly equivalent requirements—is a hurdle only if one hopes to develop a model that explains richness in general. To the extent that progress can be made studying richness gradients of given taxa or guilds, there is no limitation. If interest lies in explaining more general gradients, however, it is not logically possible to follow the approach of first calculating the number of individuals (proposition 1) and then using a sampling argument (proposition 2)—either random placement or local extinction—to predict the number of species.

A more general model must incorporate the rules by which metacommunities are formed (Leibold Chapter 8). In other words, such a model would need our four propositions as well as propositions involving the way in which interactions among species determine numbers of individuals and species. One might argue that this is precisely what Preston (1948; 1962a; 1962b) attempted, but this is not the case. Preston's argument was couched entirely in terms of single species. We know of no persuasive models that jointly predict the numbers of species and the population sizes of multiple species.

How might we avoid the second limitation, being wed to a set of arbitrary assumptions necessary to go from $N(X)$, the number of individuals, to $S = f[N(X)]$, the number of species? Numerous models of species abundance distributions arise from quite different assumptions (Fisher et al. 1943; Preston 1948; Zipf 1965; Kempton and Taylor 1974; Pielou 1975; Mandelbrot 1977;

Engen and Lande 1996; Engen 2001; Hubbell 2001; Dewdney 2003; Lande et al. 2003; Williamson and Gaston 2005). At this point there is no basis for concluding that any particular model is either logically best or empirically most supported by available data. In the absence of such a model, assumptions about the form of $S = f[N(X)]$ are arbitrary. This does not necessarily mean that more progress in gradient theory must await developments in the theory of species abundance distributions. It is possible to make progress by using a number of different species abundance distributions and asking how the choice of distribution affects the model predictions about richness gradients. Many gradient models may be robust to such choices.

Further work is needed to relax the assumption of species equivalence. For example, for the random placement mechanism, one could substitute a distribution of body mass frequencies for the constant ρ. Such a model would still assume that the shape of the distribution is the same for all sites, but that is a much weaker assumption.

Linking models to data

More challenging than model development is linking models to data. Even for a model as simple as Wright's energy model, which does not invoke tradeoffs or spatial structure, the information necessary to estimate all of the parameters does not exist, as far as we are aware. When confronted by such challenges, ecologists often respond by questioning the utility of the model. Our reply is twofold. First, formalizing models makes data requirements clearer. Although many data have been gathered in the context of studying richness gradients (Mittelbach et al. 2001), those studies have not been guided by theory, thus the disconnect between the data and the models. For example, few studies collect data on richness, abundance, and the environmental variables thought to determine richness and abundance. It may be that sufficient data exist for some systems (e.g., Stiles and Scheiner 2010) and the challenge is to discover and assemble those data.

Second, only models can provide quantitative predictions. Enough may be known about processes such as herbivory or competition to permit a sufficiently constrained state-space within which a model can be explored. Given the growing urgency of understanding global change, these models, with their general parameters, may have to do while we work to collect more data. For example, our demonstration that as a mechanistic model, the Wright energy model should be restricted to local or regional gradients and limited sets of taxa or guilds suggests that it should not be combined with global change

models to predict changes in global species richness, or that any such model should be sharply delimited in its taxonomic or ecological scope (e.g., Field et al. 2005). Obviously, Wright's model can still be used on these scales as a phenomenological model, so long as interpretation of the fit and parameter estimates is restricted appropriately. Similar hidden limitations may be discovered as we explore the details of other models.

Linkages to other constitutive theories

The theory of gradients of species richness has direct linkages to many of the other constitutive theories presented in this book. Geographic gradients (Colwell Chapter 14) concern spatial gradients only; the models considered here may have a spatial component, but typically do not. Not surprisingly, the theories share points of contact concerning the multiplicity of causes that determine gradients and how variation in species ranges along a gradient determines the form of the species richness relationship. Island biogeography theory (Sax and Gaines Chapter 10) is another one with shared mechanisms concerning immigration and extinction (Table 10.1, propositions 1, 4, and 6). The Wright energy model was first developed within the context of island biogeography theory as a way of explaining the relationship between area and species richness. As we have discussed, metacommunity theory (Leibold Chapter 8) may provide important tools for linking species abundance and species richness. Similarly, in order to formalize models that invoke tradeoffs in competition or predation/herbivory will require and examination of niche theory (Chase Chapter 5) and enemy-victim theory (Holt Chapter 7). Thus, the entire processes of theory formalization represented by this book will be an important guide and useful tool for further model development.

Acknowledgments

U.S. National Science Foundation and on sabbatical at both the Center for Environmental Sciences and Engineering at the University of Connecticut and the University of South Florida. The views expressed in this paper do not necessarily reflect those of the National Science Foundation or the United States Government.

Environmental
Gradients

Geosciences   Geographical
Gradients   Niche

Global   Island
Change   Biogeography