

3 · Ecological understanding of species distributions

3.1 Introduction

Austin (2002) presented a framework for spatial prediction of species distributions (as outlined in Chapter 1) that links ecological theory to implementation (statistical modeling). This chapter discussed the ecological model portion of that framework – those ecological and biogeographical concepts and theories that are needed to frame the empirical modeling of species distributions. The ecological model is required in order to identify the characteristics of species occurrence data that are appropriate for modeling, select explanatory variables or their surrogates, specify appropriate scale(s) of analysis, hypothesize the nature or form of the species–environment relationship (the shape of the response curve), and select an effective modeling method. In this chapter, I will review the niche concept, and related to it, factors limiting species distributions, environmental gradients and species response functions. Finally, conceptual models of the environmental factors that control species distributions at hierarchical spatial and temporal scales that are particularly relevant to SDMs will be described.

3.2 The species niche concept

A number of ecological theories related to causes of species distributions, species diversity, and community structure may be relevant to SDM, but a lot of recent discussion has focused on the *species niche concept* as it relates to SDM (Austin & Smith, 1989; Austin, 2002; Guisan & Thuiller, 2005; Araújo & Guisan, 2006; Kearney, 2006; Soberón, 2007; Hirzel & Le Lay, 2008; Jiménez-Valverde *et al.*, 2008). The species niche concept is central to ecology, and the history of its development is discussed in most ecological textbooks and more specialized books (for example, Shugart, 1998; Chase & Leibold, 2003).

3.2.1 The species niche in environmental and geographical space

The concept of species niche has evolved over time and has several interpretations (Chase & Leibold, 2003). While Elton (1927) defined the niche as a species functional role in the biotic community, Grinnell (1917a) referred to those factors that influence where one would find a species (what it needs). Hutchinson, in a now-classic reference (1987), defined the niche as “... the hypervolume defined by the environmental dimensions within which that species can survive and reproduce.” Hutchinson further distinguished the *fundamental* (physiological or potential) niche, defined as the response of species to environment (resources) in absence of biotic interactions (competition, predation, facilitation), from the *realized* (ecological, actual) niche – the environmental dimensions in which species can survive and reproduce, including the effects of biotic interactions (Austin & Smith, 1989; Austin, 2002; Chase & Leibold, 2003), depicted in Fig. 3.1a. He defined niche as a property of a species, not the environment (as discussed by Pulliam, 2000). SDM has relied on the niche concept that emphasizes species requirements, particularly (but not exclusively) representations of abiotic factors controlling species distributions. Hutchinson’s hypervolume suggests multiple causal factors of species distributions and this has also been emphasized operationally in SDM, where multiple predictors are typically used. In contrast, theoretical treatments of the niche sometimes emphasize only a few resources, requirements, or limiting factors.

A recent synthesis links these concepts of species Eltonian impacts and Grinnellian requirements (Chase & Leibold, 2003, p. 15), and they define niche as “the joint description of the environmental conditions that allow a species to satisfy its minimum requirements so that...” there is net positive growth in the local population when the per capita effects of the species on the environment are taken into account. Pulliam (2000) also emphasized the importance of measuring fitness, specifically population growth rate, in order to identify a species niche (where growth rate is positive). He proposed that Hutchinson’s niche concept, metapopulation theory (Hanski, 1999), and source-sink theory (Pulliam, 1988) together can help explain the relationship between the distribution of species and suitable habitat.

Pulliam (1988) differentiated source habitat, where local reproduction exceeds mortality, from “sink” habitat where individuals are found but do not contribute to net population growth. The metapopulation framework (Hanski, 1999) is used to describe and model populations as

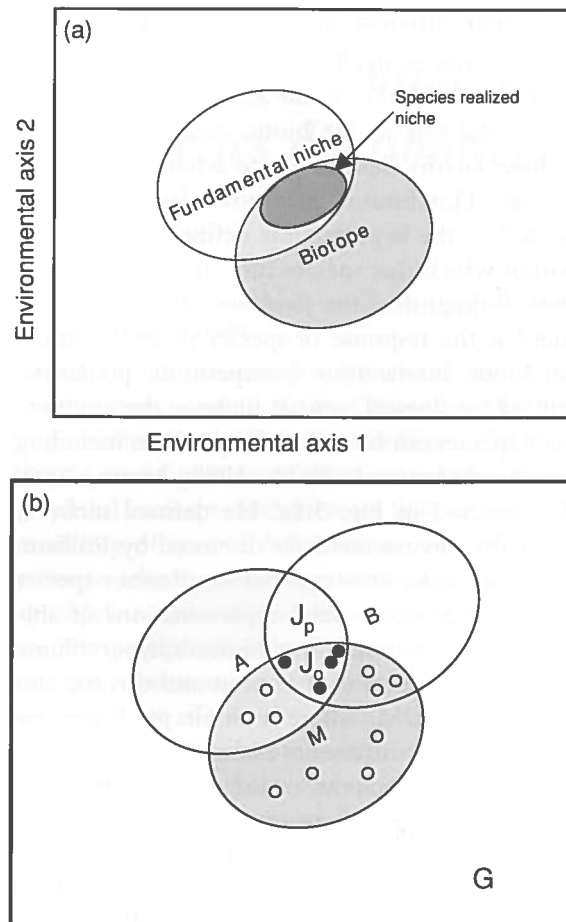


Fig. 3.1. The niche concept of Hutchinson. (a) The biotope encompasses the range of environmental conditions that occur in an area. The species fundamental niche represents its range of physiological tolerances in the absence of interactions with other species (competition, predation). The realized niche is the environmental space where a species actually occurs, accounting both for the availability of environmental conditions and for biotic interactions affecting a species distribution; after Shugart (1998) Fig. 5.7, with modifications. (b) representing geographical space, G, rather than environmental gradients, after Soberón (2007; Fig. 1). A = the geographical area where intrinsic growth rate is positive, B = the area where a species can coexist with competitors, M = the area that is accessible to a species within a given time frame given its dispersal ability. J_o is the occupied area and J_p is the potentially occupied area. Closed circles represent species occurrences in source habitat, open circles are occurrences in sink habitat (see text).

a group of spatially discrete subpopulations. A species can be absent from a patch of suitable habitat due to local extirpation resulting from population dynamics, or because of dispersal limitations. Taken together, the concepts show that a species might be found in unsuitable (sink) habitat, or might be absent from suitable habitat. This current understanding of the niche concept is important to keep in mind because most SDM studies rely on observations of species occurrence (Chapter 4), or sometimes abundance – fitness is rarely addressed directly, with a few exceptions (see Aldridge & Boyce, 2007). Equating habitat occupancy with habitat suitability may be an oversimplification.

The realized niche is often depicted as a subset of the fundamental niche (Fig. 3.1a), although it is not necessarily so, e.g., in the case of “sink” habitat or dispersal limitation, or in the case of positive biotic interactions (facilitation, symbiosis) where the occurrence of another species is a requirement. Figure 3.1b (modified from Soberón, 2007) depicts the concepts laid out by Pulliam in geographical space rather than environmental space. In this figure, A is the geographical area where intrinsic population growth rate is positive (corresponding to the fundamental niche); B is the area where a species can coexist with, or outcompete, competitors; M is the area that is accessible to a species within a given time frame, given its dispersal ability; J_o is the occupied “source” habitat and J_p is suitable habitat that is unoccupied due to dispersal limitations. If the species is found in portions of M and/or B that do not intersect A, or portions of A that do not intersect B, this would represent sink habitat.

Ackerly (2003) asserted that an assumption implicit in the Hutchinsonian niche concept is that all combinations of the relevant causal factors exist in the landscape. This is generally not the case, and he recounts the concept of “realized environment” proposed by Jackson and Overpeck (2000), who define “potential niche” as the intersection of the realized niche and the realized environment. This is an unnecessary elaboration on Hutchinson’s concepts. Hutchinson defined the “biotope” as the realized environment in the sense used above, and the realized niche as the intersection of the biotope and species fundamental niche (Fig. 3.1a), as described by Shugart (1998, p. 121).

There has been a great deal of recent discussion about the relationship between the species niche concepts described above and species distribution modeling. Returning to a question posed in Chapter 1, what is actually being modeled in SDM, the fundamental species niche, the realized niche, or the probability of habitat use? Most studies identify a description

of the realized niche as the outcome of SDM (Austin, 2002; Thuiller *et al.*, 2004b; Guisan & Thuiller, 2005), because data on actual species occurrence are usually used in modeling, and so the model extrapolates in geographical space those conditions associated with species abundance or occurrence (Chapter 4) in the environmental “hypervolume.” When this realized niche described by the statistical model is mapped in geographical space it represents the potential distribution or habitat suitability (Araújo & Guisan, 2006; Soberón, 2007).

Others strictly prefer the term habitat model or distribution model for correlative SDMs (Jiménez-Valverde *et al.*, 2008), both for the model itself and the spatial prediction, and reserve the term niche model for mechanistic analyses of factors limiting species fitness (Kearney, 2006). I have used the term distribution model throughout this book to reflect the fact that these models are conditioned on the distributions of species, and predict the potential geographical distribution of species or suitable habitat. However, in SDM the connection to an underlying species niche concept should be made as explicitly as possible in the choice of predictors, interactions between predictors, response functions, model type, and interpretation of the resulting predictions. For example, it has been suggested that environmental envelope-type models using presence-only data (Chapter 8) tend to depict potential distributions (suitable habitat), and are more suitable for extrapolation, while more complex models that discriminate presence from absence (Chapters 6–7) tend to predict realized distributions (occupied habitat), and are more suitable for interpolation (Jiménez-Valverde *et al.*, 2008; Hirzel & Le Lay, 2008). Alternatively, Soberón and Peterson (2005) argue that SDM based on coarse-scale climate variables (bioclimatic niche modeling) describes the species fundamental niche. This concept is elaborated by Hirzel and Le Lay (2008) who noted that biotic interactions (competition, predation) tend to occur at short distances, and that dispersal limitations and fine-scale environmental heterogeneity allow inferior competitors to evade negative interactions by persisting in competitor-free locations. Thus, they conclude, the realized and fundamental niche may not differ that much in practice, especially when predicted from coarser-scale environmental factors such as climate. We return to this issue in Section 3.5.

When applying the niche concept in static (statistical) species distribution modeling, that is, without including dynamic elements, an underlying assumption is that species are in (quasi-) equilibrium with contemporary environmental conditions, and that observed distribution

and abundance is indicative of environmental tolerances and resource requirements. The limitations of this assumption have been discussed extensively (Guisan & Zimmermann, 2000; Austin, 2002) and should be explicitly considered in each specific SDM study. For example, some species may still be spreading into suitable habitat following the last Glacial Maximum (Leathwick, 1998).

3.2.2 The species niche in evolutionary time

SDM is being increasingly used to test hypotheses in evolutionary biology regarding niche and geographic range as species traits (e.g., Ree *et al.*, 2005; Peterson & Nyari, 2008). Do closely related taxa share similar climatic tolerances even if their current distributions are disjunct (Kozak & Wiens, 2006)? This has been taken as evidence of so-called “phylogenetic niche conservatism” (or ecological niche conservatism) – the tendency of species to retain characteristics of their fundamental niche over evolutionary time via stabilizing natural selection (Wiens, 2004; Wiens & Graham, 2005). For example, Huntley *et al.* (1989) studied the distribution of pollen of the tree genus *Fagus* (beech) in Europe and North America with respect to climate and concluded that the distribution of that taxon is in equilibrium with current climate, and further, that their similar ranges of tolerance to climate factors on the two continents (in spite of long separation) implied that those physiological tolerances had been evolutionarily conserved. While similarity of climatic tolerances has been used as a line of evidence for niche conservatism in relation to physiological tolerances in this and other studies, the alternative scenario, adaptive evolutionary shifts in environmental tolerances, certainly also occurs, especially when the species ability to disperse or migrate (track environmental change) is limited (Ackerly, 2003).

Further, it is important to acknowledge that niche conservatism is an assumption, rather than a hypothesis, in a growing number of studies that reconstruct contemporary and paleo-distributions in order to delimit species, especially in the case of morphologically cryptic species (Pearson *et al.*, 2007; Rissler & Apodaca, 2007; Bond & Stockman, 2008). Niche conservatism is also assumed when using SDM to reconstruct paleo-distributions in order to examine other research questions in phylogeography (Hugall *et al.*, 2002; Carstens & Richards, 2007; Knowles *et al.*, 2007).

It is important to keep in mind that empirical SDM carried out to answer these phylogeographical questions has focused almost exclusively

on coarse-scale climate variables as the only predictors and has even been called phyoclimatic modeling – “combining phylogenetics and bioclimatic modeling” (Yesson & Culham, 2006). In fact, specialized modeling software aimed at this application includes built-in global climate datasets to be used as environmental predictors (Hijmans *et al.*, 2001). I argue that the “climatic niche” actually encompasses only a limited set of dimensions in Hutchinson’s “hypervolume” (see Chapter 5), and only at the broad scale described in the hierarchical framework in Section 3.5. If questions about speciation and phylogenetics are best addressed at the scale of species range limits, then broad-scale environmental predictors may be sufficient – but still, climate may not be the only factor circumscribing species ranges (e.g., Kozak *et al.*, 2008). These assumptions, and their limitations, should clearly be stated when using SDM in phylogenetic studies (Graham & Fine, in press).

While the growing use of SDM in phylogenetic research is an interesting development, the question of whether species niches are stable or not, over short (thousands of years) or long (millions of years) time periods, also has very practical implications for the use of SDM to predict the impacts of climate change on species. In particular, species whose climatic niche appears to have been conserved, at least over shorter time periods, may be the best candidates for reliable predictions of the impacts of future climate change (Pearman *et al.*, 2008).

3.2.3 Niche or resource selection function?

Niche is not the only ecological concept relating species to environment that has been used to frame SDM studies. The concept of habitat or resource selection in wildlife biology describes selection of resources by animals and is closely aligned to the quantitative methods developed to estimate it. Resource selection functions (see also Chapters 1, 4, and 8) are various types of statistical models that predict the probability of use of a resource unit, and RSF methods overlap substantially with SDM methods (Boyce *et al.*, 2002; Manly *et al.*, 2002; Boyce, 2006; Johnson *et al.*, 2006). The concepts of niche and resource selection are similar – sometimes niche is thought of in terms of the boundaries of the hypervolume where a species can exist, but these boundaries or limits simply represent habitat selection at the broadest scale – the species range. RSFs are thought of in terms of the probability of resource use relative to its availability, that is, the selection of home range or habitat patches by mobile organisms. But species response curves in the Hutchinsonian

hypervolume also suggest probability of resource use, so the concepts are not really that different.

3.3 Factors controlling species distributions

Austin (2002) described the types of factors that affect species distributions, and distinguished *proximal* (causal) factors from *distal* (proxy or surrogate) factors. Distal factors are related to resources or regulators (proximal factors), and therefore correlated with species distributions, but may be easier to measure or observe than the proximal factors themselves. The example Austin gives is that the availability of a certain mineral nutrient in the root zone of a plant might represent a proximal factor related to growth and survival, while distal factors related to the availability of that nutrient might include soil texture, pH or even general soil type. Austin (1980) further distinguished *direct* factor gradients, those with a direct physiological effect (temperature, pH), from *resource* gradients, variables that are consumed or used up (water, light, nutrients, prey, suitable nest sites). *Indirect* factor gradients have no direct effect on species distribution or abundance, and so always are distal variables. Examples are latitude, longitude, elevation, slope angle (steepness) and aspect (exposure). Ideally, variables describing direct and resource gradients would always be used as predictors in SDM. However, when only variables describing indirect gradients are available, it is important not to extrapolate the model results beyond the range of conditions used to develop the model. Further, there is no theoretical expectation for the shape of a species response curve (see next section) on an indirect gradient (Austin, 2007).

3.4 Environmental gradients and species response functions

Whittaker (1956, 1960, 1967), developed key ideas for the analysis and interpretation of the distribution of species abundance along environmental {factor}gradients. Those *response curves*, often depicted graphically, are often called species response functions (for plants) or resource selection functions (animals) (Austin, 2002, p. 103). They depict a function describing the relationship of species abundance (or fitness, occurrence, or resource selection) in relation to values of an environmental variable (Fig. 3.2).

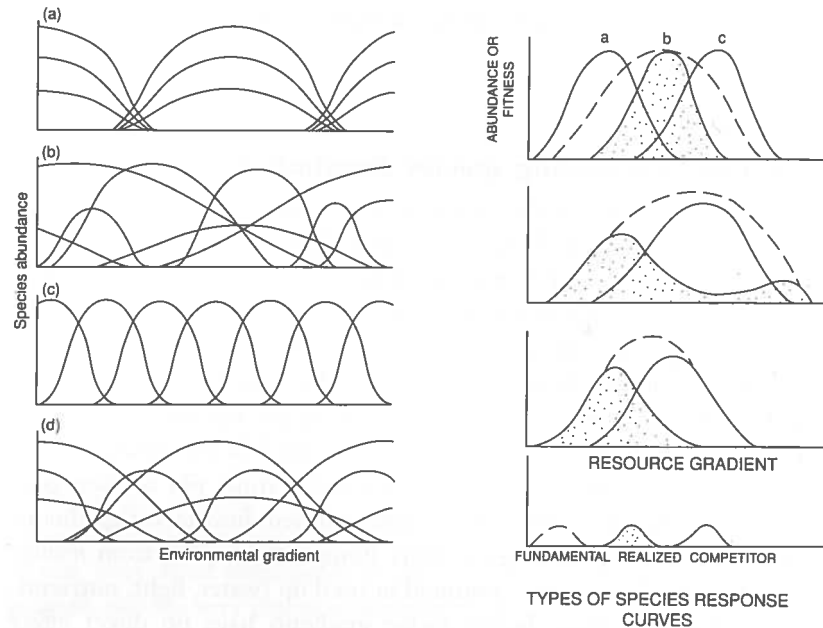


Fig. 3.2. Species distributions along an environmental gradient. Figures on the left (from Austin, 1985; Copyright © Annual Reviews; Used with permission): (a) community concept; (b) individualistic continuum; (c) competition and resource partitioning; (d) resource partitioning within strata or guilds, individualistic between strata. Figures on the right (Austin & Smith, 1989; Copyright © Kluwer Academic Publishers; Used with permission): top graph expresses same concept as (c) showing hypothetical response function in absence of competition (dashed line), middle graph showing Ellenberg's bimodal response with species restricted to extreme conditions under competition, next is species restricted to one extreme only (stippled areas). Bottom graph is the figure legend. These figures were themselves modified from those found in (Whittaker, 1970) and original sources quoted therein.

The individualistic continuum concept (Gleason, 1926) predicts species abundance optima and their limits to be independently distributed on environmental gradients (Fig. 3.2b). Alternatively, it has been hypothesized that species response functions in response to environmental factors are bell-shaped (Gaussian), equally spaced and of equal amplitude, with their width restricted by competition (e.g., Gause, 1934; Tilman, 1982), as in Fig. 3.2c (Gauch & Whittaker, 1972). However, symmetrical response functions are not usually found when they are carefully estimated using data. Unimodal, skewed responses to resource and direct

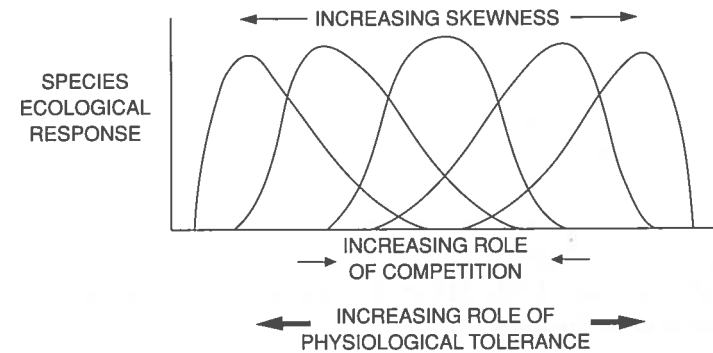


Fig. 3.3. Expected skewness of species response curves along environmental gradients. From Austin (1990). Copyright © Academic Press. Used with permission.

gradients are common. Skewed response curves may be expected (Austin & Smith, 1989), at least for certain resource gradients such as temperature, with physiological stress limiting species abundance or fitness at the "harsh" end of the gradient and competition limiting it at the benign end, as in Fig. 3.3 (Austin, 1990). Bimodal or multimodal response curves (Fig. 3.2) have also been hypothesized to result from competition (Whittaker, 1960, 1967; Mueller-Dombois & Ellenberg, 1974).

However (as reviewed by Austin 2007), Liebig's law of the minimum suggests that the true response of species to one (regulator or resource) factor can only be detected when all other factors occur at non-limiting levels (Huston, 2002, pp. 12–14). In reality, it is not usually possible to observe this without experimental data, and factors tend to covary in nature. Quantile regression, modeling the upper and lower bounds of the data, has been suggested and tested as an alternative approach to this problem (Huston, 2002; Guisan *et al.*, 2006; Austin, 2007; Vaz *et al.*, 2008), but has not been widely used in SDM (Chapter 6).

Mortality-causing disturbances (herbivory, fire) can create complex responses of species to environmental gradients (Huston, 2002), but responses to indirect gradients can also take complex forms. That form depends on the nature of the relationship between the indirect gradient and the "real" underlying causal factor (Austin, 2002). In fact, a number of alternative forms of the species response curves have been suggested that address the effects of interspecific competition, independent strata (layers or functional groups) in a plant community, and so forth,

as depicted in Fig. 3.2 (Mueller-Dombois & Ellenberg, 1974; Austin, 1985; Austin & Smith, 1989; Austin, 1999).

Finally, as Hirzel and Le Lay (2008) discussed, ecophysiological knowledge should also guide the modeler to characterize potential interactions among factors in determining species distributions. The example they give is that an organism may survive hot or dry conditions but not both. It has been found that plant distributions are affected by interactions among climate variables such as temperature and precipitation (for example, Prentice *et al.*, 1991; Miller & Franklin, 2002). Interactions among factors will be discussed further in the following section.

3.5 Conceptual models of environmental factors controlling species distributions

The previous sections described in general terms the kinds of factors (abiotic and biotic, resource, direct and indirect) affecting species distributions, and the form those response functions are expected to have along a resource gradient, but did not explicitly define *what* variables are expected to determine species distributions. “One of the first steps in building predictive distribution models is to assume a conceptual model of the expected species–environment relationships...” (Guisan *et al.*, 2006a, p. 387). Selection of candidate variables is often made based on expert knowledge (Hirzel & Le Lay, 2008). Must the factors or variables controlling species distribution and abundance always be defined on a case by case (species by species) basis? Or, are there some generalizations that can guide empirical SDM for all species, or for different taxonomic groups?

3.5.1 Heat, moisture, light, nutrients, and the distribution of plants

Climatic conditions have long been observed to play a primary role in limiting species distributions and vegetation patterns (Grinnell, 1917b; Köppen, 1923; Holdridge, 1947; Gaston, 2003), dating back to the work of von Humboldt and de Candolle in the eighteenth and nineteenth centuries. Woodward (1987; see also Woodward & Williams, 1987) discussed the physiological effects of climate on plant growth and survival via climatic controls on the light, moisture and temperature regimes. He emphasizes the lethal effects of low temperatures in setting range limits of plant species at higher latitudes and altitudes. But, as Gaston (2003) points out, there are hundreds of possible climate parameters that can

be derived from records of temperature, precipitation, humidity, evapotranspiration, solar radiation, cloudiness, etc., for any potential period of time, within and between years, for which records exist. So, connecting theory to data continues to be challenging (Chapter 5).

Several conceptual models describe the multiple environmental factors that control plant distributions. Mackey (1993a) described these factors as the *primary environmental regimes* of radiation, thermal, moisture and mineral nutrients. Most vascular plants are primary producers (photosynthetic autotrophs) and so perhaps developing a general model of environmental factors controlling plant distributions is easier than for animals of varying trophic levels, guilds, taxonomic groups, and so forth. Guisan and Zimmermann (2000) considered the resource gradients of nutrients, water, photosynthetically active radiation (light), and the direct gradient of heat sum to be the four gradients controlling plant distributions (see their Fig. 3). Franklin (1995) also described a conceptual model of plant distributions in which the environmental factors of climate, topography, and geologic substrate interact to control the primary environmental regimes in complex ways, for example, through evapotranspiration, soil development, and the subsurface moisture regime. I have modified that graphical model here to indicate that the four primary environmental regimes control the geographical distribution of the fundamental niche (potential distribution). Biotic processes and interactions – dispersal, competition, predation, facilitation, and mutualism – determine the realized niche, and additionally disturbance influences the actual distribution of species (Fig. 3.4).

Similar schemes have been developed in other fields such as ecohydrology (Wilby & Schimel, 1999) and vegetation science (Franklin & Woodcock, 1997). An important conceptual model of soil formation was developed by Jenny (1941), who proposed a functional factorial mode where soil properties in a location are a function of regional climate, parent material, topography, biota and time (for review see Scull *et al.*, 2003, in the context of predictive soil mapping). Austin (1999), in discussing the contributions of plant ecology to biodiversity research, praised the early work of Perring (1958, 1959, 1960), who showed that in chalk grasslands the distribution of plant species on topographic gradients were contingent on climate.

The graphical model shown in Fig. 3.4 also shares some elements with the conceptual model of SDM presented by Guisan and Thuiller (2005) (see Fig. 3.5) – the potential “bioclimatic” range of a species based on its requirements for, or tolerance to, heat, moisture and light

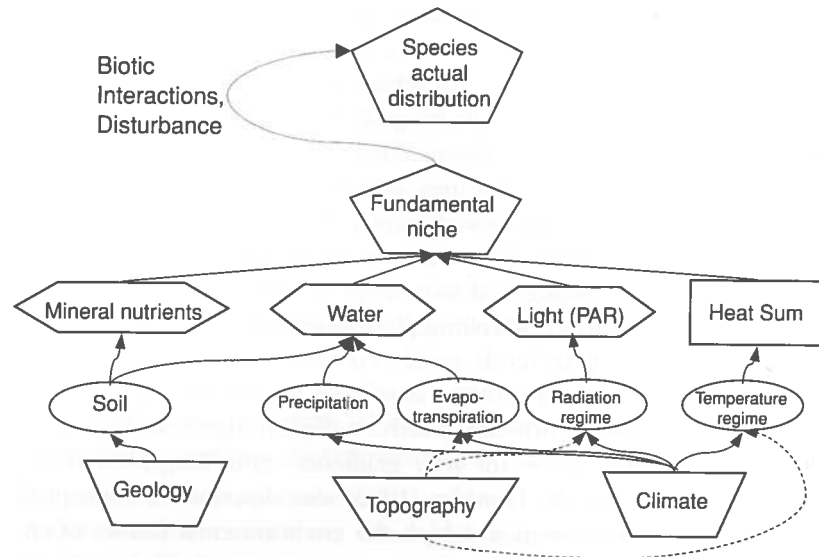


Fig. 3.4. Conceptual model of environmental factors controlling the primary environmental regimes circumscribing the fundamental species niche, further modified by biotic factors and temporal dynamics, resulting in observed species distribution. After Franklin (1995) (with inspiration from Guisan & Zimmermann, 2000).

regimes (controlled by climate at broad spatial scales) is then filtered by dispersal, disturbance and resource factors, and to tease apart their effects on the realized distribution requires dynamic modeling. This appears to limit the role of SDM to identifying the fundamental niche or potential distribution. However, SDMs, even though they are static and based on correlations, have been shown to be useful in identifying more than the “bioclimatic” niche or species range at broad spatial scales. I hope this book will convince you of this.

3.5.2 Hierarchical and nested scales of factors affecting species distributions

For animals, classically, wildlife habitat comprises cover, food and water (Morrison *et al.*, 1998, p. 41), but those authors emphasized the central role of vegetation in defining habitat conditions for wildlife. A more general, hierarchical framework for modeling the spatial distribution of animal species was presented by Mackey and Lindenmayer (2001). I think

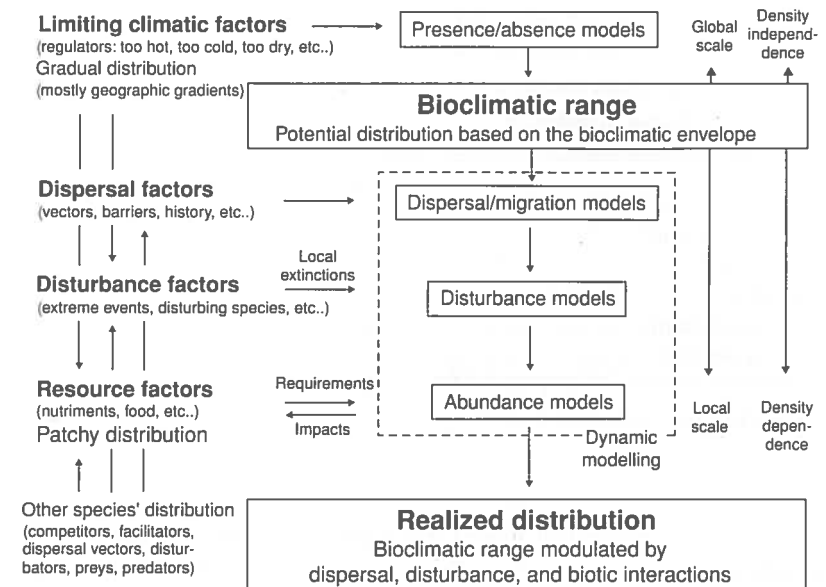


Fig. 3.5. A process hierarchy showing the relationship between static species distribution modeling and the dynamic population and community processes of dispersal, population dynamics, and disturbance. From Guisan and Thuiller (2005). Copyright © Wiley-Blackwell. Used with permission.

it is generally applicable to plants and animals and it directly and explicitly addresses spatial scale, which is important for SDM and lacking in the conceptual models just discussed.

Mackey and Lindenmayer built their framework on an ecological hierarchy of processes (Holling, 1992), and identified a five-level hierarchy of spatial scales (Table 3.1) defining natural breaks in the availability and distribution of the primary environmental resources (heat, light, water, and mineral nutrients, as discussed above). They define spatially nested global, topo-, meso-, micro-, and nanoscales. At the global scale, latitudinal and seasonal variation in incoming solar radiation drive atmospheric circulation and therefore regional weather and climate. At the meso-scale, the interaction of the synoptic weather patterns with the elevation of the land surface affects precipitation, temperature and radiation regimes, while underlying geology affects mineral nutrient availability. At the toposcale, local topography controls the distribution of water and radiation on the landscape via slope aspect, angle, slope length, and hillslope position. They define microscale as the scale at which patches

Table 3.1. *Spatial scales of environmental resources and response scale of biological organisms in a hierarchical framework linking species distributional behavior to models of their spatial distribution*

Environment	Organisms
Global (temperature)	Species
Meso (geology)	Species
Topography	Population
Micro (canopy/gap)	Group
Nano (soil patches)	Individual

From Mackey and Lindenmayer, 2001.

of vegetation or individual plants affect below-canopy moisture, heat, and light availability in a terrestrial setting. Nanoscale refers to processes occurring at finer scales within individual plant canopies. They match these five scales to appropriate biological levels of integration with regard to organisms – species, populations, groups, and individuals (Table 3.1). This conceptual model is similar in some ways to Fig. 3.4 but it describes an explicitly nested spatial hierarchy. The spatial occurrence of a focal taxon at each of these biological levels is referred to as its “distributional behavior.”

The concept of spatially nested hierarchical orders of habitat selection is well developed in the wildlife ecology literature. Johnson (1980) described four levels of habitat selection that share some characteristics with the system just outlined, and apply, at least, to vertebrates. The first level is the species geographic range, the second is selection of the home range, the third is selection of resource patches within the home range, and the fourth is selection of food items within the patches. Scale-dependent wildlife habitat selection has been examined in SDM through varying both the grain or resolution of environmental maps (Nams *et al.*, 2006; Lopez-Lopez *et al.*, 2007) and the extent (Ciarniello *et al.*, 2007) of the study area. Animals do not respond to different scales (Hobbs, 2003), but rather the ecological patterns perceived depend on the scale at which they are investigated (Weins, 1989; Levin, 1992).

There are two components of spatial scale, grain, and extent. Grain is defined as sample resolution or the size of single observation. Extent refers to the size of the study area. In the previous paragraphs, scale referred to grain (as in Table 3.1). Spatial scale and habitat selection

is discussed again in Chapter 4 with reference to collecting species data.

Pearson and Dawson (2003) presented another hierarchical framework for the scales at which environmental factors affect species distributions, and theirs specifically suggests spatial extents at which these factors operate with reference to SDM (Table 3.2). It is implicit that factors operating over larger extents typically vary slowly in space are therefore mapped at coarser grain, and vice versa. So, for example, they suggest that models based on climate variables may only be good at describing species distributions very generally at the macro-scale, and so broadly predict the potential future distribution when applied to climate change scenarios. But, these models based on macro-scale predictors do not perform well at reconstructing fine-scale distributions when those factors acting at the finer scale, such as land cover, are not included (Pearson *et al.*, 2002). For example, because of the importance of terrain in modifying the heat and moisture balance at a site, and the role of substrate (soil) in determining gradients of mineral nutrient availability, species distribution models that include edaphic and terrain factors reconstruct species distributions better than those based on climate factors alone (e.g., Coudun *et al.*, 2006; Coudun & Gégout, 2007).

Few studies have directly and comprehensively addressed the issue of selecting appropriate, spatially scaled and nested predictors of species habitat for SDM. However, one recent meta-analysis of 123 animal SDM studies did indeed find that those models that included environmental predictors from more than one hierarchical scale tended to yield more accurate predictions (Meyer & Thuiller, 2006). Specifically, they found that SDMs that included predictors at more than one grain size (local, landscape, and regional) were more accurate, even if the hierarchical relationship of those factors (finer nested within coarser) was not explicitly characterized in the model. They concluded that the majority of species were responding to habitat patterns at more than one grain (scale).

3.5.3 Environmental factors affecting species diversity and life form

Although the primary focus of this book is on modeling distributions of individual species, a number of related predictive mapping studies have used measures of species diversity, such as species richness (number of species in an area; for example, Pausas & Carreras, 1995; Waser *et al.*, 2004; Carlson *et al.*, 2007), or turnover (differences in species composition among sites; Ferrier *et al.*, 2007), as the response variable. Other studies have used the concept of species functional type (Smith *et al.*, 1997),

Table 3.2. Hierarchical framework of factors affecting species distributions in space and their scale domains, expressed as extent

Environmental driver	Extent range					
	Global > 10 000 km	Continental 2000–10 000 km	Regional 200–2000 km	Landscape 10–200 km	Local 1–10 km	Site 10–1000 m
Climate	x	x	x			
Topography				x	x	
Land use				x	x	
Soil type					x	
Biotic interaction					x	x

Based on Pearson & Dawson, 2003, Fig. 5.

such as plant life form, operationally in SDM studies (Pausas & Carreras, 1995; Thuiller *et al.*, 2006; Syphard & Franklin, in press-a). Predictive mapping of species richness often uses remotely sensed variables (Chapter 5) as predictors (Kerr *et al.*, 2001; Saatchi *et al.*, 2008). Ecological theories other than niche theory are relevant to these efforts.

Austin (1999) reviewed studies that addressed the hypotheses regarding the latitudinal gradient in species richness (Pielou, 1975), namely the energy diversity hypothesis, predicting that more species are found in lower latitudes due to greater available energy. Several studies have shown that broad-scale patterns of species richness are related to potential evapotranspiration or primary productivity, both indicators of energy availability (e.g., Currie, 1991). Global scale species richness has also been related to actual evapotranspiration, and to topographic heterogeneity (see also Pausas & Austin, 2001; Pausas *et al.*, 2003). At the regional scale, species diversity may be the results of biotic processes such as competition and predation (Huston, 2002), and historical processes including dispersal and speciation. However, as reviewed by Austin (1999), several studies have shown that regional-scale patterns in species richness are correlated with the same factors that control species distributions: climate (temperature, precipitation, and their interaction) and terrain (Margules *et al.*, 1987; Austin *et al.*, 1996).

Using an ecophysiological framework, the global distribution of plant life forms (Box, 1981) has been related to bioclimatic variables and soil properties, for example, minimum temperature, growing degree days, actual evapotranspiration, and potential evapotranspiration, which vary as a function of soil depth and texture (Prentice *et al.*, 1992). This framework forms the basis for modeling the response of vegetation to climate change, and describing feedbacks from vegetation to climate (global dynamic vegetation modeling; Neilson *et al.*, 1992; Neilson & Running, 1996; Cramer *et al.*, 2001; Sitch *et al.*, 2003).

Perhaps unsurprisingly, many of the same environmental factors that control species distributions also affect patterns of species richness and plant growth form. If an SDM framework and methods are used to model biotic or ecological response variables other than species, ecological theory apart from niche theory should be used to inform the formulation of these models.

3.6 Summary

Niche theory underpins predictive models of species distribution; the environmental factors assumed to be limiting, and represent niche

hyperspace for a species, as well as the form of the response function that is assumed or hypothesized, determine important choices about the data (sampling, variable selection, Chapters 4 and 5) and the selection and calibration of a model (Chapter 6–8). A hierarchical framework for modeling species distributions allows spatial scale to be explicitly considered when selecting environmental predictors for spatial prediction, and interpreting the resulting models and maps. Species distributions are determined by fine-scale as well as broad-scale environmental predictors, and SDMs that incorporate predictors at multiple scales, e.g., terrain, soil properties, or vegetation structure, as well as climate, more accurately describe species distributions.

General models of the factors affecting species distributions, as in Fig. 3.4, are useful as a starting point in any SDM study. However, it may be necessary to develop more detailed conceptual models linking species occurrence with environmental drivers on a case-by-case basis (for an example see Barrows *et al.*, 2005). However, even a well-developed conceptual model does not assure that data for the most important and proximal factors determining species distribution will be available, or in mapped form. But a conceptual model can help identify the most appropriate predictors, or their surrogates. For example, a model of topographically distributed solar radiation may be closer to a proximal driver of species distribution than are simple terrain variables such as slope and aspect (Chapter 5). Further, the expected form of the response function may be different depending on whether a measure of a direct or indirect gradient is used. Finally, the model may be more transferable in space or time if proximal predictors are used.

Part II

The data needed for modeling species distributions

This section addresses the data model – the second part of the framework presented by Austin (2002), outlined in Chapter 1 and used as an organizing principle for this book. Austin's data model encompasses theory and decisions about how the data are sampled, and measured or estimated. When making spatial prediction of species distributions, spatial aspects of both the species (Chapter 4) and environmental (Chapter 5) data are important.

Because species data (Chapter 4) are linked, in species distribution modeling, to digital maps of environmental variables (Chapter 5) used for spatial prediction, a conceptual model of geographical data will help to frame the material discussed in this section. Goodchild (1992, 1994) described two mental models of real geography relevant to species distribution mapping: the field and the entity. In the “field” view of the world, geographical variables can be categorical or continuous but they can be measured (have a value) at every location, so, mathematically, geography is a multivariate vector field. Examples are elevation (a continuous variable) and vegetation type (a categorical variable), which can be defined exhaustively for every point on the earth's surface. In the “entity” view, there are discrete geographic objects scattered in geographical space that is otherwise empty. Examples are records of species presence, or maps of trees, roads or fire perimeters.

In species distribution modeling we are frequently linking species location data (entities) with environmental maps (fields) to produce a new field-view of species occurrence, the distribution map. Often, the modeling techniques yield a predictive distribution map that depicts a continuous variable, the probability of occurrence or suitability of habitat (discussed in the next section), even if the response variable used