

Ecological and spatial processes

Introduction

Explicit considerations of spatial structure have come to play an increasingly important role in understanding ecological processes. The reason for this trend lies in the fact that every important ecological process takes place in a spatial context and may involve a spatial structure that affects or determines the detail of the outcome. To simplify our approach to an ecological process such as predation or eutrophication by averaging over locations (and times) does provide an easier path for such studies, but can, in the end, mislead by ignoring the heterogeneity and detailed circumstances that are crucial to the understanding we seek. Most organisms do not live in uniform environments; they encounter environmental gradients of abiotic and biotic origin, disturbance and the changes that follow it, and two- or three-dimensional mosaics with polygons and boundaries between. All of these represent spatial structure that affects the processes, and those effects may differ by the level of organization: individuals, populations, ecosystems, and so on. In this chapter, we provide an introductory discussion of the relationships between the spatial and ecological processes we study and wish to understand.

2.1 Ecological processes and spatial structure

In this section, we review the processes that are most directly relevant to the fundamentals of spatial analysis in ecology, admittedly with a bias toward plant

communities. Included are both biological and environmental processes, particularly different kinds of disturbance, but disturbance usually triggers a cascade of biological processes affecting spatial structure, such as mortality, dispersal, establishment, regeneration, growth, and competition. These biological processes interact with abiotic processes, both past (called spatial legacy) and present, to produce the current spatial pattern and spatial dynamics. While it is tempting to hope that we can deduce the past processes that gave rise to a currently observed spatial pattern, we must recognize that this is not always possible: the same process may result in different patterns, and different processes may give rise to indistinguishable patterns (cf. Figure 1.2). The patchiness of the environment or habitat can be the result of the growth pattern or architecture of organisms, such as clonal groves of aspen, or the result of topography, such as post-glacial knob-and-kettle landscapes with an alternation of ridges and ponds. In addition, even the same process may have several mechanisms contributing to it; therefore, we will focus on spatial aspects of process, not on the details of mechanisms.

A first step is to look at what should be considered spatial ecological responses; for example how organisms are arranged on one-dimensional environmental gradients. Here, the gradient is a monotonic change in a single environmental factor with physical distance, so that it is 'spatially continuous' (Keddy 1991) and it may result in obvious zonation in the community that develops on it. The spatial structure to be analysed is still nonrandomness in space with some predictability, and therefore falls into the general category of spatial pattern

although there are no repeating units of spatial structure. A gradient results in predictable and directional variability, which allows the appearance of species where they were previously absent and then their disappearance. The predictability, that is the 'pattern', is in how the species enter and leave along the gradient and in their ranges and densities where they are present.

On gradients, the observed spatial pattern can be the basis for *some* inferences about the processes that created it, and the current spatial pattern affects future processes. The potential location of any species is determined by its physiological responses to the gradient, but the observed locations result from the interaction of physiology with other ecological processes such as competition, facilitation and predation. The current arrangement of species on the gradient determines which species will be able to interact: the organisms that are located near each other on the gradient may be the ones competing most strongly, but they may also be the ones with the greatest potential for positive interactions (Bertness & Calloway 1994; Brooker *et al.* 2008).

The usual model of a species' response to an environmental gradient is a symmetric unimodal curve as a function of the controlling factor's intensity or of the physical distance along the gradient (Jongman *et al.* 1995; Legendre & Legendre 1998). These two may be different because the same physical distance along a gradient may not produce the same quantitative change in a controlling environmental factor, depending on position. The symmetric unimodal response curve may occur rarely in nature, with asymmetry being more common and some responses being bimodal (Austin & Austin 1980; Minchin 1989; Collins *et al.* 1993; Jongman *et al.* 1995; Legendre & Legendre 1998). The symmetry or skewness of a unimodal response also depends on the scaling of the axis representing the environmental factor. Furthermore, many landscapes are fragmented, so that species abundance fluctuates within the overall geographical range due to the availability of habitat (Opdam & Wascher 2004; Fortin *et al.* 2005).

Another consideration for the analysis of patterns on gradients is that the rate of change in a controlling factor may itself vary along a gradient, or organisms may respond more or less strongly to the same amount of change in the factor, depending on position. As a

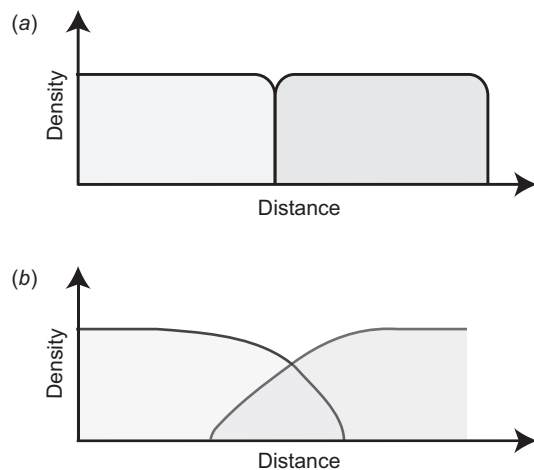


Figure 2.1 Species replacement allows a zone of competitor coexistence. (a) There is no spatial overlap between the two species. (b) Spatial overlap between the two species (different shades of grey).

consequence of this possibility, the 'critical tide level' hypothesis suggests that there are particular levels on a tidal shore where species replacement in communities of intertidal algae occurs more rapidly over small differences in height, because the responses of species to changes in the duration of desiccation is much stronger at that level (Doty & Archer 1950). Similar critical levels may exist in other kinds of environmental gradients, such as temperature or soil moisture, but there are few studies to confirm their existence.

Species can be arranged on environmental gradients in a variety of ways, and these can reveal characteristics of the organization of the communities. For instance, there may be critical levels in the controlling factor, as just described, or there may be evidence of biological interactions among the species. Interspecific competition is important in the development of spatial pattern on a gradient (Pielou 1977a), and the inability of two competitors to coexist can result in the beginning of one species' range following immediately after the ending of another's range. On the other hand, if species replacement allows a zone of competitor coexistence, the density of one species will decrease as the other's increases in that coexistence zone (cf. Dale 1999; Figure 2.1; Pielou 1978).

The arrangement of species in a community on a gradient (and how the individuals of different species occur with respect to each other in space) should provide insight as to how a community is structured (Shipley & Keddy 1987). One model is the 'community unit' view, which suggests that groups of species replace each other along the gradient so that there are clusters of upper and lower boundaries. Another model suggests that the species occur more or less independently of each other so that upper and lower boundaries occur independently and boundaries are therefore not clustered. These two models are not the only possibilities, and Whittaker (1975) described two others. As variants, where one species replaces another, there is a choice between sharp mutual exclusion and gradual replacement with zonal overlap. Distinguishing among these several possible arrangements and answering related questions can be achieved by spatial analysis in systems that are well-structured by gradients. These methods will be described in Chapter 9.

In addition to responding to environmental gradients, plants also seem to respond to the presence of plants of other species, that are in close proximity, in a number of different ways. In most situations, the plants of different species are neither randomly nor independently arranged. The term 'association' can refer to the tendency of the plants of different species to be found in close proximity more often than expected, 'positive association', or less often, 'negative association'. The association between species can be based on shared or divergent ecological requirements and capabilities or on the ability of one species to modify the environment to make it more (positive) or less (negative) suitable for the other. Examples of positive influence include facilitation by 'nurse plants' that enhance regeneration in stressful environments and such 'positive interactions during succession and recruitment . . . are unusually common characteristic forces in harsh environments' (Bertness & Calloway 1994). Examples of negative influence include situations in which plants are affected by allelopathy where chemicals from one plant reduce the growth of another, such as the effect of the shrub *Adenostoma fasciculatum* on annual herbs in California (cf. Crawley 1986). Negative influence would also include competition for resources such as light, soil moisture, or

nutrients, but the interaction might be more symmetric, depending on relative sizes, than in allelopathy which tends to be asymmetric.

The association of species in a plant community is usually treated as a pairwise phenomenon, and the resulting network of these pairwise associations (technically a graph, see Chapter 3) is sometimes referred to as the phytosociological structure of the community (Dale 1985). The presence or absence of a third (or fourth!) species may affect the relationship between any particular pair of species, and this possibility may lead us to consider investigating multiple species association, where the frequencies of various combinations of species presences and absences are examined (Dale *et al.* 1991). For the analysis of k species, this involves the analysis of 2^k contingency tables which summarize the combinations of presence and absence of all species, rather than examining a number of the more familiar 2×2 tables for pairwise analysis. This topic of multispecies combinations will be pursued in greater detail in Chapter 10. While this approach clearly involves greater complexities, in some circumstances, the greater insights it produces can often justify the effort involved.

Although in applying spatial analysis, as we might to communities on environmental gradients, we often focus on the pattern observed at a single time, yet we know that the processes are dynamic and that communities change through time. All of the processes leave spatial pattern and spatial structure behind, which will need to be considered in interpreting the results of spatial analysis. While the major disturbances may be rare or even non-recurring (at least within the time periods considered in ecology), ecologists are concerned with stochastic disturbances that do recur, like fire or infestation-induced mortality because they often give rise to recognizable patchiness in vegetation, as well as having large impacts on other processes. In some instances, the senescence of individual plants can have an effect similar to the externally induced mortality associated with abiotic disturbance or obvious pathogens or herbivores. In either case, the result is often a clear interaction between the ecological process and spatial pattern that arises, usually as patches of different composition, which can affect the dynamics of future processes.

An orthodox view of the development of vegetation by ecological succession begins with an intense and extensive disturbance, such as fire, which 'resets the clock'. When that reset is to the equivalent of zero, with all the vegetation and any remnants of it removed, leaving only bare (mineral) substrate, the subsequent development is called primary succession. When the vegetation is killed or removed but some residual (e.g. soil) remains after the disturbance, what follows is called secondary succession. This dichotomy sounds as though there is a clear distinction, but many forces of disturbance leave patchy areas of different severity, and different sizes and shapes on the landscape. For example, glaciation–deglaciation cycles tend to leave an interesting set of mainly linear features, like scrapes, moraines, eskers, and so on, as well as more isodiametric features like drumlins and knob-and-kettle topography. Forest fires are notoriously uneven in their effects, leaving areas more-or-less untouched (i.e. forest remnants) and areas of severe burn down to the mineral layer. Forest patches are usually elongated, often in the direction of the wind prevailing at the time of the burn, but large fires can burn for many days and may be influenced by winds from several directions, leaving a complicated spatial print on the forest landscape. Windstorms are also very uneven in their effects, with down-bursts knocking down groups of trees or single trees falling to leave small gaps in the canopy. Insect outbreaks can be heterogeneous even in single-species stands, and act selectively in mixed vegetation, again leaving a very patchy and complex structure in their wake. These insect outbreaks can be incomplete in their effects with some mortality and some recovery, and the outcomes are even more diverse with ongoing multi-year infestations. All these forces of disturbance affect the spatial pattern we observe in vegetation, usually in uneven and complicated ways.

Whatever the disturbance, a number of different processes contribute to the succession that follows, and it is impossible to separate completely the different views of the successional processes from concepts of the nature of the plant 'community', however broadly or narrowly defined. At one extreme is the view that the plant community is like an organism, and its development proceeds through a series of predictable phases until the final 'climax' self-replacing vegetation is established

(Clements 1916). The contrasting view (Gleason 1927) is that the plants of different species and their propagules act more-or-less independently in response to their availability and to the environment they encounter. This allows the possibility of different plant associations or combinations of plant species as the end-point of succession in similar regions or of different successional pathways to similar vegetation when the pathways converge (Glenn-Lewin *et al.* 1992). It is also possible that similar starting conditions can develop to different compositional end-points; there is a well-known example from Lake Michigan sand dunes (Olson 1958). For spatial analysis, the question arises whether there is a clear difference in spatial structure that might be seen in mature vegetation, depending on which view is the best description of real communities. As a first guess, we could suggest that the Gleasonian model might be expected to produce much more variability in species combinations and physical structure within any small area than would be predicted in the Clementsian model. Further consideration of this approach to examining the nature of succession and climax vegetation can wait for our discussion of spatial aspects of species diversity (Chapter 10).

The processes that may be invoked to explain some of these temporal patterns observed in successional sequences include the following:

'facilitation', when the plants early in the successional sequence modify the environment in such a way that the recruitment of later species is enhanced;

'inhibition', when the early plants' influence on the environment decreases its suitability of later species' recruitment, so that later species establish only when the first group have died out or have been reduced by disturbance; and

'tolerance', where the environmental modifications have little effect on subsequent recruitment (Connell & Slatyer 1977).

For completeness, we should include:

'self-inhibition' where the plants of early stages make the environment less suitable for themselves or at least reduce the rate of their own recruitment, as is often found with shade-intolerant trees that are typical of some early successional sequences.

All four of these may have a spatial effect, although 'tolerance' is clearly a kind of null model, and while the spatial effects may be primarily local, the local effects may have an influence over large areas (Solé 2007). The spatial version of facilitation is nucleation (Yarranton & Morrison 1974; Cutler *et al.* 2008), where the plants of one type act as the nuclei for the colonization and establishment by others, rather like 'nurse plants', see Munguia-Rosas & Sosa (2008); or plants as bird perches, see Pausas *et al.* (2006). Inhibition and self-inhibition may have very localized effects that could be detected by spatial analysis, just as nucleation may produce a clear spatial signature in the resulting vegetation.

Two other processes that are key factors in the development of vegetation are regeneration and competition. The regeneration of any species depends on the availability of propagules (e.g. a seed source close enough for a significant density of seeds to arrive), and of a substrate suitable for germination and subsequent development. These both clearly have spatial aspects that affect success, the number and distances of seed sources; and the number, locations and sizes of patches of suitable seedbed substrate. Competition, particularly from previously or simultaneously established plants, has a strong effect both numerically on population dynamics by affecting growth rates and survivorship, and spatially, particularly on the distances between plants and the relative sizes of neighbours as they grow. These various ecological processes all contribute to the spatial structure of vegetation and have, in turn, implications for the application of spatial analysis techniques and the interpretation of their results. The key is, of course, the relationship between these ecological processes and the spatial patterns we investigate.

We have already alluded to the presidential address by Watt (1947), which made explicit some basic concepts on the relationship between pattern and process in plant communities and laid the foundation for much of the research in plant ecology over the following decades. One of the major themes in that work is the concept of cyclic change at relatively small scales within reasonably well-defined plant communities. For example, the paper described in some detail

a number of plant systems in which aggregates of individuals of different species can be considered as different kinds of patches or phases forming a mosaic that is the community broadly considered. In many instances, there is a recognizable and predictable cycle of an upgrade period of regeneration and a downgrade period of degradation or senescence. Any instances of a particular phase may develop under somewhat different conditions with different juxtapositions, and at different rates and for different durations, but the whole community remains essentially the same with the processes that give rise to the sequence of phases unchanged.

The inference, based on reconstructing the sequence of the phases in any of Watt's examples, is that they form a cycle through time and that an understanding of the community 'as a working mechanism' should be based on that of the relationships among the phases. These sorts of example are often cited in relation to the concept of cyclic succession, a predictable cycle of species replacements in time, in contrast to directional succession which is seen as the orderly replacement of species or communities, like a 'compositional gradient in time' (Glenn-Lewin & van der Maarel 1992). However, in some examples, the cycle involves the replacement of phases of a single species, as it grows and senesces, rather than the more obvious replacement of one species with another.

In addition to the very general and pervasive influence of Watt's ideas on the thinking of plant ecologists on the relationship between pattern and process in the intervening decades, they have affected the development of at least two conceptual models of plant community dynamics.

The more specific of the two is the 'mosaic-cycle' concept of ecosystems (Remmert 1991, cf. Aubréville 1938, 1950–1951), which has been applied to a range of ecosystems, but chiefly to the more speciose forests of temperate and tropical systems. The model has a number of contributing sources, not only Watt and Aubréville, but also Jones (1945), Bray (1956), and Schulz (1960). The basic concept is a community of mosaic phases consisting of aggregations of species, with the phases developing in a fairly predictable but not completely orderly sequence in a cyclic fashion.

The idea of the mosaic partition is the partitioning of the two-dimensional space into polygons. Each polygon is a 'tile' of the mosaic, and each belongs to one of a limited number of phases that make up the cycle. What may be important is not just the phase to which a particular tile belongs, but also the phases to which neighbouring tiles belong. The effect of neighbouring tiles was not made explicit in the original description of the idea, but for spatial analysis, the identity and influence of neighbours may be important factors in the implications of this model.

The original description of this model provided many examples where the mosaic-cycle concept may be applied (cf. Remmert 1991), but the question remains as to how distinct and identifiable the mosaic phases in real plant communities are, forest or otherwise (more on this later!). Another important question about the functioning of a mosaic cycle concerns the mechanisms that determine the transitions. For example, to what extent is the completion of a full cycle of replacements dependent on extraordinary disturbance events, such as fire, insect outbreaks, or violent wind-storms? The related spatial question concerns the extent to which disturbances determine the physical sizes of the 'tiles' of the mosaic. For spatial analysis, the question is to what extent the spatial structure of different phases in a mosaic mimic the temporal stages in the cycle. In wave-generation cycles in forests, a spatial cross-section through the forest resembles a schematic diagram of the changes through time (cf. Sprugel 1976; Bekker & Malanson 2008).

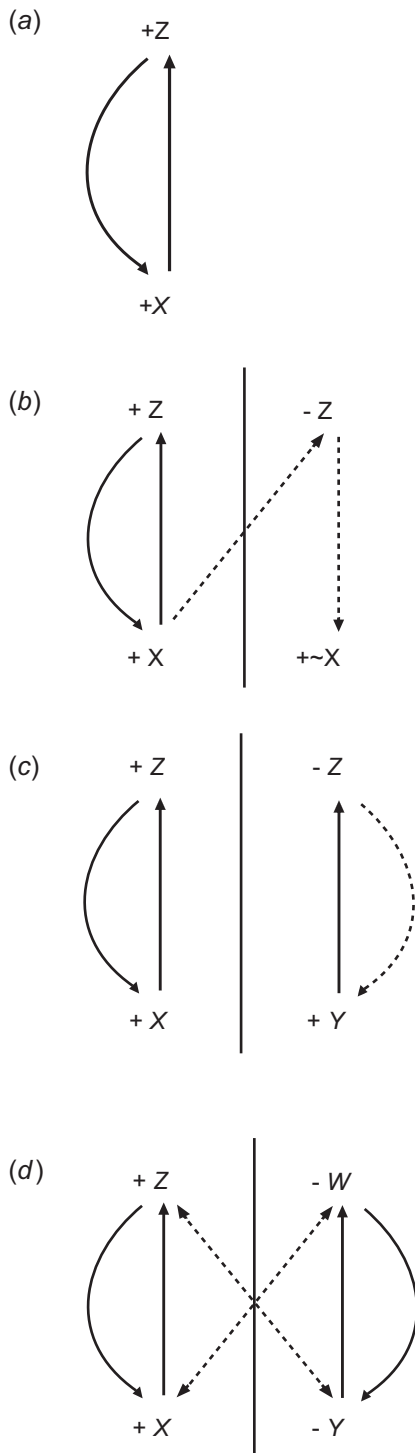
The more general concept related to the mosaic-cycle hypothesis is the 'patch-gap' model of vegetation development, particularly as applied to forests. In its simplest form, it merely suggests that significant recruitment takes place only in canopy gaps (supposing the canopy is that of forest trees), so that vegetation regeneration depends on the formation of gaps, whether by autogenic causes such as senescence and mortality or by allogenic forces such as disturbances in the form of wind-throw by storms (cf. Frelich 2002) or of fire-induced mortality (cf. Johnson 1992). This model, then, depends explicitly on some form of disturbance to create the gaps that drive or allow the next steps in the cycle of processes. The kind and rate of

regeneration will depend on the size of the gap created and the characteristics of the substrate and other conditions left behind.

Another factor in interpreting the effects of gap size is the possibly nonlinear effects of patch (or gap) size on the processes being studied. For example, doubling the size of a patch of vegetation in a mosaic consisting of vegetated patches alternating with bare ground may more than double the number of seeds and other propagules that are intercepted and retained in the patch. Then differences among patch sizes cannot be averaged and the distribution of sizes will markedly affect the outcome of the process of dispersal. The same is true of density effects, although the density of plant stems and the distance between neighbouring stems may be confounded, which may lead to nonlinear responses.

The relationship between patches and gaps or between patches of different kinds can be affected by feedback loops in ecological processes. Negative feedback is a common mechanism to maintain homeostasis (e.g. a thermostat), and positive feedback loops are found in self-reinforcing systems (e.g. lighting a fire). As an ecological example, in the competition between individual plants for light, any initial advantage of a plant being slightly larger than its competitors tends to be amplified as competition for light continues, with the advantage for the larger plants increasing as the plants grow (cf. Wilson 1988). Positive feedback loops can be found in vegetation, where any one might be called a 'switch' (Wilson & Agnew 1992). Switches represent situations in which the vegetation of one kind modifies the environment to make it more suitable for itself, in contrast to facilitated succession in which each stage makes the environment less suitable for itself and more suitable for the stage(s) that follow.

Wilson & Agnew (1992) provided a classification of these positive feedback switches, but most of the examples are of the 'one-sided' form where the plants of vegetation X change the environment where they are present, call it environmental factor Z, and thus enhance their ability to invade adjacent areas (Figure 2.2a). A familiar example is wetland invasion by *Sphagnum*, where the moss acidifies the



environment and raises the water table, allowing the sphagnum-dominated area to expand in a process known as paludification. The other common form of switch is a two-sided 'reaction' switch, where vegetation X increases variable Z in its own area, which also causes a decrease in Z elsewhere in the region (Figure 2.2*b*). For example, if the vegetation stripes in an arid region capture moisture as run-off at their upslope edges, there is less moisture available for the downslope areas between stripes. This process exaggerates the differences because the stripes capturing the moisture support plants that survive and grow, and the bare regions between, receiving less moisture, are unsuitable for establishment and growth.

There are two other, less common, two-sided switches possible: 'symmetric'; X increases or decreases factor Z in its own patches, and vegetation Y decreases or increases Z in its own patches (Figure 2.2*c*); and 'two factor': X changes factor Z in its patches, and vegetation Y changes another factor, W, in its patches as in Figure 2.2*d* (Wilson & Agnew 1992). An example of the former would be a mosaic of forest and savanna, in which the forest produces nonflammable litter, decreasing fire spread and increasing its own self-replacement, with the savanna producing flammable litter increasing fire spread and its own self-replacement. The latter

Figure 2.2 Examples of positive feedback switches. (a) Aspatial or within a patch. Simple positive feedback between vegetation X and factor Z: the presence of X increases Z; Z increases X. Solid arrows represent positive effects. (b) Spatial: two patches. Two-sided switch between vegetation X and factor Z: the presence of X increases Z in its own patch; Z increases X. Increasing Z in X's compartment decreases Z in a nearby compartment, reinforcing the absence of X there. Dashed arrows represent negative effects. (c) Spatial: two patches. Two-sided switch between vegetations X and Y, and factor Z: the presence of X increases Z in its own patch; Z increases X. The presence of Y decreases Z in its own patch; Z decreases Y, so that its decrease enhances the presence of Y, just as its increase enhances the presence of X in its patch. (d) Spatial: two patches. Two-sided switch between vegetations X and Y, and factors Z and W: The presence of X increases Z in its own patch; Z increases X. The presence of Y increases W in its own patch; W increases Y. The mutual interactions of X and W and of Y and Z are negative.

might have a similar mosaic example with deciduous forest interspersed with conifer stands. The deciduous trees cannot reproduce well in the acid soils produced by the conifers but can in their own shade, and the conifers can establish in their own acid soil, but not in the more complete shade of the deciduous stands (cf. Wilson & Agnew 1992).

The potential spatial effects of these switches can help interpret the results of spatial analysis of vegetation. In the last two examples, the positive feedback reinforces the mosaic system in which it occurs, thus maintaining differences. These switches can also increase the sharpness of a boundary between two vegetation types or can accelerate or delay the replacement of one community type with another (Wilson & Agnew 1992; Agnew *et al.* 1993). These switches are spatial analogies of chaos and catastrophes, which are usually described for dynamic systems with the changes occurring in time not space. Here chaos refers to the behaviour of a dynamic system that appears to be disorderly, but is not random, merely very difficult to predict. Classic examples are the three-body problem, weather, predator–prey models with ‘forcing’ (i.e. large population growth rates), and so on (see Schroeder 1991; Hastings 2009). For our purposes, there are two important features of chaos. First, chaos is not random; it does not result from a stochastic process, but it is deterministic. Second, an important characteristic is that small differences tend to be amplified, leading to divergence of conditions rather than converging to more-or-less equivalent behaviour (Gleick 1987). In spatial systems, the amplification of small differences through time can intensify the distinctness of the phases of a spatial mosaic. We will have more to say about spatial chaos later in this book (Chapter 11). For studies of the spatial dynamics of plant communities, the concept of chaos, just like that of positive feedback loops, forces us to consider the possibility of multiple pathways of successional sequences and the amplification of small differences, producing divergence rather than convergence of states.

Catastrophe theory (Jones 1977) studies abrupt changes in a system's dynamic behaviour that can depend on small differences in the conditions. The

concepts of stable and unstable equilibria, and of catastrophe theory can be applied to dynamic systems such as a simplified plant–herbivore interaction. For a more complicated example, in the multivariate system of a plant community, which can include a large number of biological and abiotic variables, the concept of several locally stable equilibria is presented in the form of ‘alternate stable states’ (May 1977; Beisner *et al.* 2003) resulting from different trajectories of temporal succession. Again, the concept may include the characteristic of sensitivity to small differences which can be amplified through time to produce large differences in the observed spatial structure.

The amplification or reinforcement of spatial differences is one general mechanism by which landscapes may be said to have memory. Landscape memory, or legacy, is the degree to which ecological processes are influenced by their past effects on a landscape (Hendry & McGlade 1995; Peterson 2002; James *et al.* 2007), and it can take the form of something that looks very much like spatially explicit temporal autocorrelation. However, that autocorrelation can be positive or negative. For example, if recently burned areas of a boreal forest tend not to burn again until substantial re-growth has occurred and fuel has built up over decades, the result is negative autocorrelation because fire will not soon recur. On the other hand, if there are areas of forest that are immune from fire because of structural or topographical characteristics (e.g. protected by rocky outcrops or lakes), whereas other areas (e.g. on south-facing slopes) tend to burn frequently, the result is positive autocorrelation because some areas tend to burn frequently and others not at all. In a study of the European beech forest mosaic cycle, Hendry & McGlade (1995) concluded that in such a system, the emergent spatial structure is dependent on the amplification of local interactions by the mechanism of ecological memory. If landscape memory is an important factor in the dynamics of a mosaic-cycle system, the effects of neighbouring tiles may be a critical, if neglected, factor in understanding the system.

The last concept in this ‘constellation’ of related ideas is that of self-organization or autogenic pattern formation. The phenomenon is that simple rules, acting very locally, can produce clear patterns over

large extents (see Solé & Bascompte 2006). Reaction-diffusion phenomena in chemical systems can produce spirals and rings that move in travelling waves (Volford *et al.* 2007), and simple spatial models of vegetation produce hexagonal arrays of spots, or rings, or the stripes often observed in arid tropical brushlands (brousse tigrée) (Thiéry *et al.* 1995). Colonial insects are often cited in this context as producing elaborate and robust structures from simple behavioural modules (Bonabeau *et al.* 1997; Halley & Winkler 2008). For plant-dominated systems, however, there are many examples similar to those described by Watt, with a strong spatial component to their processes. Bogs and other wetlands often have elevated strings of vegetation separated by long narrow pools or flarks, oriented across the direction of water flow through the complex (Rietkerk *et al.* 2004). In Watt's examples, *Calluna* and other dwarf shrubs produce stripes of vegetation (across the wind direction) and *Agrostis* forms ring structures, growing on the leading edge and dying in the centre (clumps of the dwarf shrub *Dryas*, as well as many other plants, also show this structure; Anderson 1967; Dale & MacIsaac 1989). While the development and orientation of all these structures may be driven by abiotic forces such as wind and water, the form produced (stripes, strings, rings) can be attributed, at least in part, to the growth form and mechanics and other characteristics of the plants themselves. In any case, self-organization usually involves a balance between positive feedback at some spatial scales (e.g. within a phase) and negative feedback at other scales and locations (e.g. between phases). There is evidence that in some semi-arid ecosystems, very local interactions such as plant-to-plant facilitation can affect significant transitions over large areas, as in the switch between vegetation and desert (Solé 2007).

We have described some of the processes in plant communities that affect spatial pattern and spatial dynamics, because these concepts provide the background for an informed interpretation of the results of spatial analysis. These apply most directly to studies of plants, but they also provide a conceptual framework for the organisms that depend on or inhabit a plant

community, whether ground beetles, insect pollinators, songbirds or large mammals.

2.2 Spatial processes by species level of organization

Spatial structure can vary according to the level of organization: individuals, populations or metapopulations, communities or metacommunities, and landscapes. In contrast to some of the examples in the preceding section, biological processes, particularly those in community succession, can act to obscure differences, blurring edges or spatial transitions, and enhancing homogeneity. We described phenomena which involved the amplification of existing differences, enhancing of boundaries between the elements of a spatial mosaic. This has at least two consequences for the spatial structure we are trying to analyse. The first is that small differences may not be negligible; they may be amplified through time to produce larger differences of ecological significance. The second is similar: boundaries between 'phases' of a spatial mosaic can turn out to be of great importance, even if they are not immediately obvious.

All these processes are dependent on scale and level of organization. The spatial scales that govern the data should be evaluated because the dynamics that follow disturbance or environmental change depend on the area affected and the duration (Levin 1992, 2000). Anything we might think of as 'cyclic succession' is noticeable because of the spatial grain and temporal pace at which it occurs. The scale of the observations and the scale of the observer both have effects. Cyclic species replacement in a grassland might be invisible because the spatial scale is too fine; cyclic replacement in a community of saxicolous lichens would be invisible because the temporal scale is too long.

In addition to the ecological processes at the population level (such as competition, predation, and facilitation) and the phenomena of switches, chaos, and self-organization, yet another level of organization is important to the spatial structure of natural systems: metapopulations and metacommunities (Hanski 2009; Leibold 2009). Populations and communities are not

usually isolated independent entities, but occur in the patchy environment as systems of metapopulations and metacommunities, linked by mutual dynamics and processes such as local extinction, dispersal, and re-colonization. Most simply, a metapopulation is a system of local populations, each occupying a patch in a heterogeneous environment, in which there is a dynamic balance caused by local extinctions in patches, and re-establishment due to re-colonization from surviving subpopulations in other patches. The balance is affected by risks and rates of extinction and the speed and distance range of dispersal, and the factors that affect the probability of re-establishment.

Similarly a metacommunity is made up of the component metapopulations where a set of local communities (possibly identifiable with distinct patches, but possibly not) are linked by some species dispersal between them. However, the spatially patchy structure makes the situation more complicated, and a number of different paradigms have been proposed for metacommunities (Leibold 2009). To paraphrase Leibold (2009), there are four different paradigms to be distinguished (although there are variants of each).

- (1) Patch dynamics (PD). This is a simple generalization of the metapopulation model, dependent mainly on the trade-off between extinction of local populations, a function of competitive ability, and colonization. This model often ignores the environmental heterogeneity of patch variability.
- (2) Species sorting (SS). This version focuses on community assembly and assumes that population dynamics once established are more important than the colonization process. Local interactions and environmental heterogeneity among patches are strong factors in determining the community structure.
- (3) Mass effects (ME). Here dispersal is taken to be of sufficient numerical intensity that it can affect local population numbers and dynamics. Populations can persist in poor environments because extinction is staved off by colonization, producing 'sink' populations. The result is that more species can exist in patches than would if extinctions were allowed to run their course, and species sorting

becomes important because there are more species in a patch.

- (4) Neutral dynamics (NM). This model is based on the assumptions that differences among species can be ignored and that the system behaviour is dominated by stochastic demographics within patches and migration among local communities. This provides a null model for comparison, and is distinguished by the importance of stochastic processes.

These four can be compared based on environmental heterogeneity, which is high for SS and ME, and low for PD and NM); and on the rate or range of dispersal: similar to extinction rate for PD, intermediate for NM and SS, and similar to demographic rates for ME (Leibold 2009).

How these various effects play out in real community systems is not fully resolved, but the concepts are important for the development of appropriate approaches to spatial analysis of communities and for the interpretation of the results of such analyses. They are especially important for a spatial analysis of species diversity in real systems, as we describe in Chapter 10, and how the differences between local communities may depend on physical distance. On an environmental gradient, compositional difference should increase linearly with geographic distance (Figure 2.3a), as with the 'isolation by distance' in population genetics, unless the mass effect is sufficiently rapid and strong to counteract environmental differences selecting out different species on different sections of the gradient. In that circumstance, or in an environment that is either uniform or just randomly heterogeneous, compositional difference should not exhibit a trend with distance, although actual values will depend on factors such as the degree of heterogeneity, the strength of the mass effect and the speed of species dispersal (Figure 2.3b). A third possibility is that dispersal, which is distance-limited in its effectiveness, can override the accumulation of compositional differences over shorter spatial ranges, but not over larger ranges, producing a curve that is flat initially, but then begins to rise. The long-distance end of the curve may or may not reach a maximum set by extrinsic factors (Figure 2.3c).

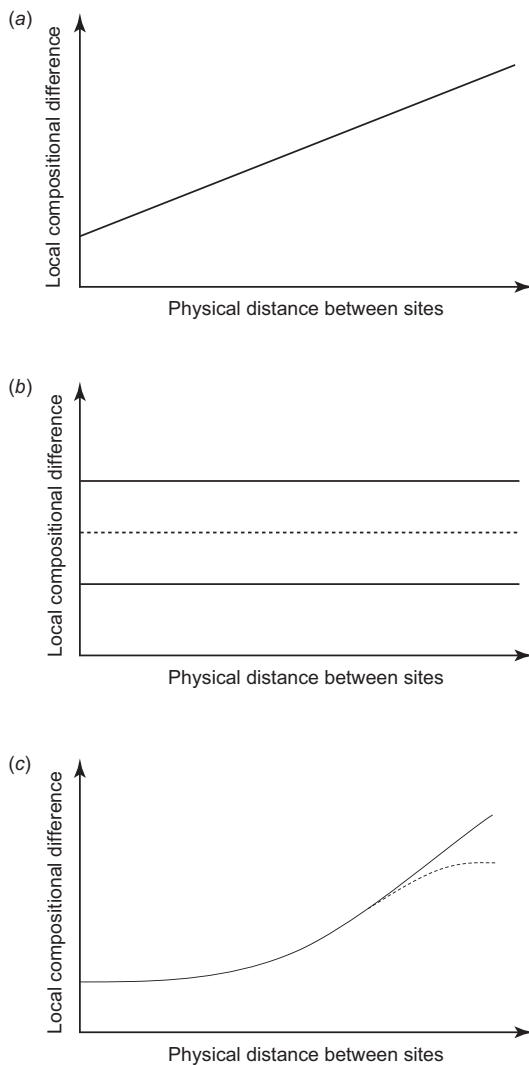


Figure 2.3 Predictions of communities' structure according to metacommunity models. (a) Communities' differences on environmental gradient increase. (b) Communities similar independent of distance: difference depends on heterogeneity. (c) Communities similar at short distances, but differences accumulate at larger distances: difference depends on processes and may reach a maximum.

Our understanding of the ecological processes acting at any particular site should influence the kind of data we collect and our choice of methods of

analysis. For example, in the establishment of plants of a single population, initial regeneration may be spatially clumped, but with growth and competition, size differences may result in asymmetric competition, nonrandom mortality and increasing distance between survivors, thus reducing clumping. The best choice of method may be therefore to map, categorize and measure all the plants in a given area, and then to use versions of point pattern analysis, such as Ripley's K for locations, marked point pattern analysis for living versus dead plants, and mark correlation analysis to investigate the sizes of neighbours as a function of distance (see Chapter 4). On the other hand, for inter-specific studies investigating localized facilitation or inhibition during succession, marked point pattern analysis methods can be used for mapped stem data (Chapter 4), or quadrat covariance methods for species density data from contiguous quadrats (Chapter 5). If the focus is on testing the mosaic-cycle hypothesis, methods for the detection of boundaries (Chapter 9) and for the analysis of neighbour networks (Chapter 3) or spatial transitions (Chapter 11), may be the most appropriate choices, with the mosaic tiles as the spatial units. One goal of this book is to provide commentary and advice on the range of methods and how they can be matched to the ecological hypotheses of interest. Those hypotheses intimately involve a number of potentially complex processes and phenomena, as have been introduced here.

For spatial analysis, knowing the ecological processes that may be active allows us to realize that the results can provide insights into the spatial processes, even if the relationship between pattern and process is not simple. Given the variety of such phenomena as feedback switches, chaos, and spatial self-organization, in addition to the more familiar processes like competition and facilitation, it is possible that several processes can produce patterns that cannot be distinguished, despite the strong relationship between pattern and process. Therefore, although spatial analysis can eliminate some hypotheses, it can seldom be used to confirm mechanisms definitively. On the other hand, knowledge of the kinds of processes that may be acting should affect the choice of data to be collected, the methods of

analysis used, and the interpretation of the results of that spatial analysis. The goal of the analysis, after all, is to develop a better understanding of this great range of ecological processes.

2.3 Spatial process

In mathematics, a stochastic process is a collection of random variables (a 'family' of random variables, in some versions, Doob 1996) governed by at least one parameter, such that each outcome has a distribution associated with it (because the variables are random the values are not uniquely determined). A spatial process is therefore a mathematical system with stochastic rules that generate events or values of variables in a spatially explicit framework, so that there is a location for each event or value. For example, a simple spatial process could create a set of 16 randomly located point events, each having an attribute labelled by an integer ranging between 1 and 5, in a square of 100 units by 100 units. The rules could be as follows.

Repeat the following steps (1 to 4) 16 times, starting with $i = 1$ and going to $i = 16$:

- (1) choose a random real number between 0 and 100 for the i th x -coordinate;
- (2) choose a random real number between 0 and 100 for the i th y -coordinate;
- (3) choose a random integer from 1 to 5 for the label of the i th event;
- (4) record [and plot on a diagram] the location and value of the i th event.

In this example, provided that the random numbers are generated correctly and independently, the events are independent of each other as are the events' labels. Depending on how the random numbers are created, a different set of numbers is generated every time the algorithm above is implemented with different starting conditions; and any one of these can be considered to be an individual realization of the same process. If a uniform distribution is used for the random coordinates, any point in the plane has an equal probability of being the location of an event. On the other hand, if both the variables for the spatial coordinates (x and y) follow bell-shaped distributions, the events will tend to

be more common in the centre of the square. In other cases, the rules can be structured so that spatial location has an effect on the label (e.g. higher label values may be more probable for higher valued coordinates) or the labels at different locations may depend on their relative positions (e.g. label values may tend to be the same as those of neighbouring events).

The resulting set of events or variable values has statistical properties that are determined by the rules of the process that generates them, although their observed magnitudes will vary from one realization of the process to another.

A spatial process is usually treated as a stochastic process like those that occur in time, but one that operates in a spatial context. Space is different from time in that it usually has no inherent directionality, where time has asymmetric, possibly causal, relations of before and after ('*post hoc, propter hoc*'), and space is usually, but not always, treated in two dimensions, not just one. Like a temporal process, a spatial process can be:

- (1) stationary, which means that the statistical characteristics of the process or of the variable of interest do not change with location (for *weak stationarity*, the mean and covariance structure are invariant; for *strong stationarity*, the distribution itself or all the moments of the distribution are invariant) (see Chapter 1); or
- (2) non-stationary, in which case, the mean or variance (or the covariance structure) of the values of the variable of interest varies with location. There are many different ways for a process and the realization of a process to exhibit non-stationarity: a trend in any direction, banding in any direction, or patchiness at any one (or more) scales and with or without directionality. That being said, it is also hard to imagine how spatial inference could proceed without an assumption of stationarity of some kind (Ripley 1988) because that is what allows prediction from one location to another.

As a simple illustration of a spatial process, consider a rectangular area A of a plane into which we are going to place n events, each consisting of a dimensionless point. The magnitude of the process is the mean number of events per unit area, which here is $\lambda = n/A$. If the events

are placed at random and independently of each other, every small sub-unit of the area will contain an event with a probability that is proportional to its area. The number of events per areal unit will follow a Poisson distribution defined by parameter λ , and any two sub-areas of the same size have equal probability of containing an event. The Poisson distribution has the characteristic that both the expected value (the overall mean) and the theoretical variance are equal to λ , and thus equal to each other. This is the 'null model' of complete spatial randomness (CSR), sometimes called a 'Poisson forest' for obvious reasons. It is important to realize that a completely random stationary process can give rise to a spatial arrangement of events that does not look particularly random, especially if the number of events is small. For example, in placing five events at random in a square, one in eight realizations (about 13%) will have all five events located in only one half of the square (see Figure 1.16). This may seem a bit contradictory or at least puzzling, because it does not look like the result of anything stationary.

This apparent non-uniformity resulting from a random process is different from the situation in which patchiness is created by the structure of the process itself. There are a number of different ways that this can happen, several based on the CSR process which gives rise to the Poisson pattern just described. For example, each event created by CSR could produce, in turn, a cluster of events, with its number and location governed by a second stochastic process. Another mechanism is to have a process very much like CSR with intensity parameter λ , but allowing the process intensity to vary with location \mathbf{s} , so that $\lambda(\mathbf{s})$ is itself the result of a second stochastic process. If the processes at both levels are Poisson, we end up with what is called a 'doubly stochastic Poisson process' (Cox 1955), now more commonly known as the 'Cox process' (Kingman 1993). Biologically, this situation can arise when a second generation is derived from a population of parent plants which were distributed following CSR, provided the offspring disperse from their own parent plants independently, but with their average location being the location of their parent. The positions of the daughter plants are then conditional upon the positions of the parental generation

(Kingman 1993; Felsenstein 1975). An extreme example would be to have the first generation of four plants all occurring in one half of the area available, just from randomness, each producing four offspring on average, with the mean of their locations being that of the original parent; clearly the second generation is unlikely to be evenly spread through the sample area. Processes like this result in pattern that has a distribution of event counts per areal unit for which the variance is greater than the mean, indicating patchiness of some kind. In that case, each realization of the process is not invariant under translation, and so not homogeneous, but heterogeneous, even though the whole process over many realizations is homogeneous and stationary because over all possible realizations, every small subarea still has an equal probability of being occupied and the expected value (mean) and variance are constant. (In a non-stationary process, the parameters of mean and variance are different at different locations.)

Heterogeneous and non-stationary processes can also be classified according to whether direction plays a role in affecting the probabilities:

- (1) isotropic (no directional effect); or
- (2) anisotropic (different characteristics in different directions, e.g. stripes)

where 'homogeneous' refers to a process that is invariant under translation, 'isotropic' refers to one that is invariant under rotation (Ripley 1988). Again, these terms refer to the underlying process and the characteristics may or may not be obvious in the pattern in any particular realization of the process.

For biological systems, as for mathematical ones, a spatial process is one that gives rise to events or to the values of variables with definite locations, but the activities that generate the patterns are now biological processes, like dispersal or mortality, rather than purely stochastic and abstract. In mathematics, both the generation of values and the values that result are referred to as a spatial process, but we will depart from this practice by distinguishing between the process that generates the values, and the set of values themselves, which we will continue to call the spatial pattern or a realization of the process. In the ecological context, the biological spatial process produces a result

that is like a single realization of some underlying probabilistic process. Even in artificial structures, it is difficult or almost impossible to resolve the true underlying structure of a model from a single realization unless the numbers are very large. In addition, what may seem to be a simple single spatial process may actually be a series of different processes. For example, within a conifer stand, the initial dispersal of pollen grains into the air may seem like a homogeneous process and their spatial dispersion in the air may be like a homogeneous pattern, but the process that determines their movements thereafter and the locations of landing is not homogeneous because of different probabilities associated with air currents and the structural characteristics of the female cones which increases the probability of pollen accumulation (see, for comparison, Niklas & Norstog 1984). Gravity

should be one of the most-nearly homogeneous processes acting in ecosystems, but the flow of glaciers that it drives tends to be both non-stationary (faster in the centre) and anisotropic as a physical process, and often creates anisotropic landscape structures even after retreat, which then produce an anisotropic pattern in ecological systems that follow. Because any observed biological pattern is like the result of a single realization of the underlying process, what seems to be heterogeneity or non-stationarity in the data does not guarantee that the underlying process shares the same characteristic. The concepts of spatial processes are very useful for our understanding of spatial analysis and its interpretation, and we have occasion to make use of them throughout the remainder of this book, but without the technical details that you can find in more mathematical texts on stochastic processes.