4
Thoughts on the Generation and Importance of Spatial Heterogeneity in Ecosystems and Landscapes

JOHN PASTOR

Abstract

Landscapes are spatially dynamic because materials and energy spread over them and change the distribution of ecosystem properties. This heterogeneity of the distribution of ecosystem properties can either be random or patterned. The landscape becomes patterned when the spread of materials and energy correlates an ecosystem property in one local neighborhood with that at another. When the spread of materials and energy does not correlate properties of different neighborhoods, then the landscape can still be heterogeneous but random. Various processes that result in spatial heterogeneity include physical disturbances (e.g., fire, erosion, etc.) that spread across neighborhoods and remove materials but whose spread is partly determined by previous disturbances; directional gradients in the flow of materials, energy, or information; and different diffusion rates of coupled ecosystem components combined with positive feedbacks, otherwise known as diffusive instability. Examples of these processes will be given from other papers in this conference and elsewhere.

Introduction

The living world is not all green slime or a big leaf; things are different from place to place. This variety of the living world is what makes it a stunningly beautiful and interesting place to live. It is also what makes understanding ecological systems difficult.

Spatial heterogeneity of the distribution of ecosystem processes across the landscape can be random or patterned (or a combination of both). A heterogeneous spatial distribution of ecosystem properties is random if, given the value of an ecosystem property at a point, the value of that property at adjacent points cannot be predicted. In contrast, a heterogeneous spatial distribution is patterned if, given the value of an ecosystem property at a point, the value at adjacent points and possibly points further away can be predicted with some confidence. Because the spread of materials and energy across the landscape correlates values of an ecosystem property between adjacent local neighborhoods, this spread can therefore result in patterned heterogeneity.

For the most part, we know how to analyze spatially homogeneous distributions through analysis of variance and general linear statistical models. We know how to model their dynamics through coupled ordinary differential equations that depict energy and material flows between ecosystem components and whose parameters do not depend on position in space. In contrast, we are only beginning to learn how to describe the origin and dynamics of spatial heterogeneity. These require new mathematical, experimental, and observational tools for their description and analysis.

Physical disturbances create and sustain heterogeneities by removing materials from ecosystems or transferring materials from one ecosystem or ecosystem component to another. Physical disturbances often have a large random element, but they also may depend on underlying heterogeneity, which is often caused by previous disturbances. The spread of a disturbance correlates values of an ecosystem property at a given point with those at its neighbors and beyond to the boundary of the patch created by the disturbance.

Transport of energy and materials along a directional gradient, such as movement of water and suspended sediments or dissolved compounds downhill, also creates patterned heterogeneity. The transport of energy and materials along a directional gradient correlates ecosystem properties along the gradient. Ecosystem properties will therefore be similar for long distances along transects in the direction of the gradient but become less similar more rapidly along transects perpendicular to the gradient.

Spatial heterogeneities can also be generated by positive feedbacks between ecosystem components, such as soil, vegetation, and higher trophic levels (Meinders and van Breemen this volume). Such patterned heterogeneity can arise even in the absence of gradients and physical disturbances and can create patterned heterogeneity from homogeneity or random heterogeneity. This generation of pattern from homogeneity or randomness in the environment via positive feedbacks between ecosystem components is sometimes called "self-organized complexity" (Kauffman 1993; Bak 1997; Meinders and van Breemen this volume).

If two interacting ecosystem components also diffuse or spread across the landscape, new and surprising heterogeneities can arise even without any underlying heterogeneity in the physical environment (Okubo and Levin 2002). Under some circumstances, such heterogeneities could be stable. This seems to be especially prevalent in herbivore-vegetation systems where both the herbivore populations and the plant species that support them are diffusing across the landscape. For example, the spatial dynamics of balsam fir is coupled to the spatial dynamics of spruce budworm populations during an outbreak. In turn, the changes in the spatial distribution of balsam fir affect the fate of the outbreak (Holling 1978).

In this paper, I wish to explore how the spread of physical disturbances, the directional flows of materials down a gradient, the positive feedbacks between ecosystem components, and the diffusion of interacting components across the landscape all generate spatial heterogeneity. I will use the papers in this volume and additional ones from the literature as examples. My purpose is to seek some general principles of the sources and consequences of spatial heterogeneity and attempt to reach broad conclusions about similarities and differences between major ecosystem types in order to offer approaches for organizing future research.

Physical Disturbances

The ecological literature on disturbance is vast, and it is not my intent to review it here. Instead, I wish to make a few remarks about some aspects of the nature of spatial heterogeneity caused by disturbances and why these might differ between terrestrial and aquatic ecosystems. By disturbance I mean some physical process that removes a fraction of an ecosystem component or adds to it. Thus, I exclude insect outbreaks, for example, because such biological processes (which are sometimes referred to as "disturbances") could be treated by other approaches involving diffusion of the population, which I discuss below. Physical disturbances, such as fire, erosion, landslides, avalanches, and so forth, are qualitatively different from "disturbances" initiated by growth of a population, because the physical disturbance itself is not a component or pool within an ecosystem but a process by which material is transferred spatially.

Disturbances have two aspects that are important for the generation of spatial heterogeneity. The first is where the disturbance is initiated, which has a large random component (e.g., where the lightening strikes) but also depends on the conditions in the initiation location (e.g., whether there is sufficient fuel of the right moisture content to ignite when struck by lightning). However, once initiated, the disturbance can and often does spread to adjacent locations whose conditions may not have been right for initiation but are sufficient for the spread (e.g., if your neighbor catches fire, you may burn, too). Thus, spatial heterogeneity caused by physical disturbances is partly random (through initiation) and partly patterned (through contagious spread).

Aquatic ecosystems, especially streams and rivers, are well mixed. Their components generally have rapid turnover because of short lifetimes of organisms and because currents break down structures by rolling and mixing of bedload and woody debris. Constant flux of water also dilutes the introduction of pollutants and contaminants at point sources. The spatial heterogeneities caused by many disturbances to streams, especially disturbances related to point-source pollution, are therefore quickly dissipated once the disturbance ends (Niemi et al. 1990).

In contrast, terrestrial ecosystems are not well mixed and often contain slow-growing perennial individuals. Therefore, the spatial pattern caused by a disturbance remains for long times. But if the recurrence interval of a disturbance is shorter than the recovery of a disturbed patch, then under some conditions the initiation and spread of any disturbance may partly depend on previous disturbances. How a disturbance moves through a landscape that previous disturbances have created is a major unanswered (and difficult) question of disturbance ecology.

From a modeling standpoint, this means that simple, first-order Markov chains, often used as a first approximation to modeling disturbance (see reviews by Baker 1989 and Pastor et al. 1993), will always be somewhat deficient. First-order Markov models assume that the probability of a transition in the system is constant and depends only on the current state of a system. But in fact, whether or not a disturbance happens at a point or propagates from it depends on disturbances back to some period in the past. Not only that, but the current state of a system (or local neighborhood) also includes the distribution of adjacent neighborhoods and their states (hence contagion). Cellular automata approaches are useful in dealing with these higher order effects because the change in a given cell depends in part on the state(s) of its neighbors (see review by Neuhauser 2001).

As Turetsky et al. (this volume) and Romme (this volume) show, disturbances are a particularly important source of heterogeneity in boreal regions and in coniferous forests of the arid West perhaps because of the slow recovery of vegetation owing to the slow growth rates of the species present (Chapin et al. 1986), because of the slow turnover rate of the soil N pool (Flanagan and Van Cleve 1983) that supplies the N required for plant growth (and hence recovery), and because of drought.

An excellent example of the importance of heterogeneity caused by a physical disturbance such as fire is the landscape of virgin forests of the Boundary Waters Canoe Area (BWCA) of northern Minnesota (Heinselman 1973). Virtually every stand in the BWCA originated from a fire, but the fire return intervals (which differ for different stands) are almost all less than the recovery time from the previous disturbance. Consequently, fires in the BWCA partly burn through previous burns. For example, some 44% of the BWCA burned during 1864, but only 20% of the current stands originated in the 1864 burns: the rest of the current landscape originated in fires that happened later but which spread to these burned areas from adjacent older stands that ignited first.

Romme (this volume) shows that the importance of spatial heterogeneity caused by previous burns in the arid West varies with forest type and climatic conditions. When the climate is dry and hot, everything burns and previous spatial heterogeneity is unimportant in fire spread. Thus, we get large-scale catastrophic fires as in Yellowstone during the late 1980s and in the southwest during 2001 and 2002. However, during more moderate years or in higher elevations where extended periods of hot and dry conditions are rare, the underlying spatial heterogeneity caused by previous burns is very important in determining initiation and spread of new fires.

These case studies raise several general questions. How do burns and other disturbances become overlaid on previous disturbances of the same type or of different types? Does it matter what the previous disturbance was, and if so, in what way does it matter? Is there a characteristic fractal or some other geometry of partly overlapping disturbances? If so, what determines it? Are these "geometries," if they exist, characteristic of a particular ecosystem or are there more general aspects common to two or more otherwise different ecosystems? These are some of the major questions, as I see them, which need to be answered to develop a more complete understanding of how disturbances produce and interact with spatial heterogeneities in any landscape.

Directional Gradients

Both terrestrial and aquatic ecosystem ecologists have long dealt with gradients in the vertical dimension and its effects on ecosystem properties. The premier example of such vertical spatial heterogeneity is the extinction of light through a canopy and water column, the fundamental starting point of much of forest ecology and limnology. If we assume that leaves are randomly distributed through the canopy or that the water column is homogeneous, then this light gradient can adequately be treated by means of a linear model whereby the change in light through a given layer at some depth d is some fraction k of the light entering that layer, leading to the familiar exponential extinction curve:

$$I_d = I_0 e^{-kd}. \tag{4.1}$$

It is a relatively simple matter to incorporate heterogeneities in the distribution of leaves through the canopy or vertical changes in water column transparency simply by replacing d with a function describing how leaf area or transparency change with depth and integrating down to depth d:

$$I_d = I_0 e^{-k\int_0^d LAI(\delta)d\delta}. \tag{4.2}$$

This vertical light gradient, I_d, often leads to a stratification of both terrestrial and aquatic communities according to the photosynthetic response curves of the constituent species (Shugart 1984; Tilman 1988). In terrestrial ecosystems, when the community is vertically stratified into shade-intolerant species above shade tolerants, then light-use efficiency and hence net primary production by the entire community may be maximized (Pastor and Bockheim 1984; Tilman 1988).

The most important horizontal directional gradient in landscape ecology may be topographic, causing water left after transpiration to flow transversely and downhill. Watershed studies have typically examined the mass balance of inputs to the watershed via precipitation and stream outputs, as demonstrated in an exemplary manner by the Hubbard Brook Ecosystem Study (Likens et al. 1970; Bormann et al. 1977). However, these watershed studies typically do not look in detail at the pathways and patterns of nutrients fluxes between stands within the watershed and how that affects the eventual transfer of nutrients to the streams (or lakes) at their base. Conversely, many detailed studies of nutrient cycles of ecosystems or stands within watersheds implicitly assume that the ecosystem sits on a flat table and leaching losses take place vertically rather than semihorizontally. To truly bridge watershed and stand-level approaches, we need to connect stands in the landscape by means of directional fluxes of nutrients down topographic gradients. Thus, the input-output balance of an ecosystem at a given point may depend as much on its position in the landscape and the delivery of nutrients to it from upslope as on the exchange of nutrients between components within it.

This has important consequences in both streams and the watersheds that surround them, perhaps especially so for the riparian zones. The riparian zone potentially receives nutrients from every stand above it, but the nutrients are delivered to it in sequence downslope. Therefore, the sequence of stands along a slope and their differing input-output balances may determine the loading of nutrients to the riparian zone.

Heterogeneity in riparian zones may also determine downstream flows of nutrients. Naiman et al. (this volume) review how sources of heterogeneity in riparian forests, such as coarse woody debris, denitrification hotspots, debris jams, formation of bars and side channels, and so forth, may mitigate large transfers of nutrients to aquatic ecosystems. The strong directional gradient that transfers nutrients along a topographic sequence may interact with fine-scale heterogeneity within the riparian zone to determine overall land-water material transfers. This fine-scale heterogeneity within the riparian zone may enhance nutrient retention if it increases the path length a molecule travels before it enters the stream channel, thus increasing its residence time within the riparian zone. The role of heterogeneity within the riparian zone must therefore be assessed in the context of the overall heterogeneity of the landscape and downslope transfers of nutrients to the riparian zone and how the heterogeneity of the riparian zone affects nutrient retention before the nutrient enters the stream channel. How heterogeneity of processes operating at different scales interacts to determine lateral transfers of material across landscapes is a difficult topic of great importance.

Urban ecosystems (Band et al. this volume) are distinguished partly by particular sorts of directional flows along the grid systems of streets. These directional flows can be parallel (one-way streets) or antiparallel (two-way streets) along two axes usually at 90 degrees to each other. This grid system is an attempt to impose some spatial order on travels of humans and commerce in a city, but it can have great consequences for the spatial dynamics of cities as landscapes. For example, city engineers must control downslope flows of water to prevent erosion and flooding of the roadbeds; these water diversions into storm sewers and along curbs have large effects on urban stream ecosystems (Band et al. this volume). Furthermore, pollutants from automobiles are dispersed from sources that move down streets and are dispersed further by wind tunnels or prevented from dispersing by wind-breaks caused by the buildings (Band et al. this volume). It would be interesting to learn how this grid system of directional flows of traffic, water, and wind disperses seeds of exotic plant species or diseases of boulevard trees.

When directional gradients of fluxes at boundaries of patches are very steep, the sign of the gradient can determine the degree of heterogeneity inside the boundary of a patch. Kratz and MacIntyre (this volume) remind us that there is a very important directional gradient at the surface of a lake, namely the heat flux gradient, which strongly determines the spatial heterogeneity within the lake. When the heat flux gradient at the lake surface is positive, heat flows out of the lake and the water column physically turns over, bringing nutrients from the sediment to the surface and oxygen from the surface to the lower depths. The lake is then also thermally homogeneous. But when the heat flux gradient at the lake surface is negative, heat flows into the lake and it becomes thermally stratified. This phenomenon, so important to aquatic ecosystems, depends on the fact that fluids such as water can be well mixed with fast time constants. Similar thermal stratification of the atmosphere over a city results in the formation of smog. Such spatial dynamics do not have any counterparts in terrestrial vegetation-soil systems because these systems cannot be well mixed over any reasonable ecological timescale.

Finally, positive feedbacks within ecosystems (Tongway this volume; Meinders and van Breemen this volume) can amplify the heterogeneity produced by directional gradients. Tongway shows how positive feedbacks between plants and soils in arid systems concentrate and retain soil moisture being delivered at a point such that water availability becomes raised above threshold levels required for plant growth, leading to the further development of patches of vegetation and high resource availability in a sea of low resource availability.

Such feedbacks and the spatial patterns that arise from them are not confined to arid systems. Peatlands are another excellent example of how plant-soil feedbacks lead to the formation of spatial patterns (Turetsky et al. this volume). Horizontal water flow patterns in peatlands are a result of microtopographic gradients and hydraulic permeability of the peat, both of which interact with the plant community. Broadly speaking, two different communities (bogs and fens) can be found in peatlands; these in turn appear to be related to hydrologic sources of nutrient inputs (Wright et al. 1992). In bogs, peat accumulation has raised the local water table above the regional water table; bogs therefore receive their exogeneous nutrient inputs solely from precipitation. Fens are in lower topographic positions or on the margins of peatlands and are not isolated from the regional groundwater table; they receive nutrient inputs from both precipitation and groundwater. *Sphagnum* mosses, ericaceous shrubs, and black spruce (*Picea mariana*) dominate the vegetation of bogs while sedges and other graminoids dominate fens.

These vegetation patterns are enhanced by positive feedbacks between the plant community and the type of peat formed from its litter (Glaser 1992). Sedges and other graminoids produce peat of high hydraulic permeability. Therefore, water preferentially flows through fens and maintains them. On the other hand, *Sphagnum*-derived peat has low permeability and water flow is diverted around it. *Sphagnum* mosses prefer these relatively drier conditions, and their continued dominance and production elevates the peat surface above the water table, leading to the formation of bogs (van Breemen 1995). These raised bogs shed precipitation to the surrounding wetter areas, further enhancing the dominance of graminoids there. Directional flows of water into peatlands from the upland is thus broken up into patterns of water tracks (occupied by fens) and raised bogs (occupied by *Sphagnum*), which are stabilized by these positive feedbacks between the plant community and the peat formed from it.

The positive feedbacks between peatland vegetation, peat formation, and hydrologic gradients and flows at local scales may have important implications for global carbon budgets. Although northern peatlands occupy less than 2% of the world's land surface (Post et al. 1982; Bridgham et al. 2001), they contain one third of the world's soil carbon and nitrogen pools (Post et al. 1982, 1985; Gorham 1991) and are the source for 6–9% of global methane emissions (Mathews and Fung 1987; Aselmann and Crutzen 1989; Bartlett and Harriss 1993). Carbon and nutrient budgets in bogs and fens are very different: bogs appear to accumulate more carbon and nutrients than fens (Glaser 1992; Bridgham et al. 1995, 2001). Therefore, the spatial distribution of bogs and fens and how that distribution arises from positive feedbacks between the plant community and water flow patterns may determine the pattern and degree of carbon balances of many northern regions.

Diffusion, Diffusive Instability, and Pattern Formation and Destruction

Mahadevan (this volume) points out that diffusion of an ecosystem component or property destroys heterogeneity by dispersing the property or agent responsible for it across the landscape or seascape. Thus, a plume of nutrients or pollutants introduced at a point into a fluid, an insect outbreak at a spot, or the aggregation behavior of some zooplankton are all dispersed as these entities diffuse through the landscape or fluid. This dispersal can, to a first approximation, be described by random Brownian motion, otherwise known as Fickian diffusion. Thus, under some circumstances, random spatial motion destroys spatial heterogeneity (Murray 1989; Okubo and Levin 2003; Mahadevan this volume).

However, dispersing species also interact with each other (through predator-prey interactions, for example). This trophic interaction of two dispersing species can create, rather than destroy, spatial heterogeneity under certain conditions. If the growth of the lowest trophic level involves a positive feedback (autocatalysis) with itself (e.g., population growth) or with some underlying environmental condition (e.g., enhancement of nutrient availability through litter feedbacks) or is sustained by inputs from the surrounding environment and if the populations of species in different and interacting trophic levels spread or diffuse at different rates, then conditions are ripe for creation of a rich variety of spatial heterogeneities and patterns. In this case, the diffusion causes spatial heterogeneity by modifying the trophic interactions as the interacting populations away from points at different rates. This heterogeneity can, under certain circumstances, then be amplified by the interactions between trophic levels or between species and their resources. This phenomenon, known as reaction-diffusion or diffusive instability, was first mathematically described by Turing (1952) and is often called a Turing mechanism in his honor. Excellent reviews of this theory rich with ecological examples are given by Edelstein-Keshet (1988), Murray (1989), Holmes et al. (1994), Okubo and Levin (2002), and Levin (2003). This theoretical approach gives explicit conditions for when either spatial heterogeneity or homogeneity is stable and, through numerical solutions or simulations, it can also give some predictions about the pattern of heterogeneity. These explicit conditions and solutions can then be tested in experiments or observations.

To see the conditions under which such spatial heterogeneities arise, consider first a set of coupled equations for the interactions of two species in an otherwise homogeneous environment:

$$\begin{cases} \dfrac{dS_1}{dt} = F_1(S_1,S_2) \\ \dfrac{dS_2}{dt} = F_2(S_1,S_2), \end{cases} \tag{4.3}$$

where S_1 and S_2 are prey and predator, respectively, and F_i are the differential equations (e.g., Lotka-Volterra predator-prey equations) describing their growth and interactions. "Predator" and "prey" are meant here in a general sense in that the predator "takes up" or consumes the prey. Thus, the "predator" can be a carnivore consuming an herbivore, an herbivore consuming a plant, or a plant species taking up a nutrient "prey" (an example of this will be given in a moment). For what follows, it is important to keep in mind that the growth of the "prey" population at the lowermost trophic level is either self-generating by means of autocatalysis, enhanced by positive feedbacks with some underlying environmental variable, or sustained by input from the outside environment.

Assume there is a spatially uniform (homogeneous) equilibrium in the absence of diffusion such that:

$$F_i(S_1^*, S_2^*) = 0, \tag{4.4}$$

where S_1^* and S_2^* represent equilibrium densities of S_1 and S_2. This equilibrium is spatially homogeneous and stable if small disturbances of size ΔS_i decay exponentially when the system is near equilibrium. Examples of disturbances of size ΔS_i could be harvesting or stocking of a population or enhancement of local nutrient availability by fertilization. The rates by which disturbances decay or grow are given by the eigenvalues of the Jacobian matrix J of partial derivatives (sometimes called the "community matrix" by ecologists):

$$J = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \tag{4.5}$$

where $a_{ij} = \partial F_i / \partial S_j$ and J is evaluated at the equilibrium points S_1^* and S_2^*. Analytical solutions of the eigenvalues near equilibrium are in terms of the parameters of the dynamical equations F_i; these parameters are usually the rate constants of fluxes between trophic levels of the system or the input-output terms.

The eigenvalues, λ_n, of J give the rates of growth or decay of the perturbations in n dimensions (where n is the number of compartments of the system):

$$\begin{bmatrix} S_1(t) \\ S_2(t) \end{bmatrix} = \sum_n \begin{bmatrix} S_1(t_0) - S_1^* \\ S_2(t_0) - S_2^* \end{bmatrix} c_n \omega_n e^{\lambda_n t}, \tag{4.6}$$

where $S_i(t_0) - S_i^* = \Delta S_i$ is the initial size of the perturbation to S_i, ω_n is the corresponding normalized eigenvector to λ_n, and c_n are constants that depend on initial conditions. Clearly, if all $\lambda_n < 0$, then the perturbation $S_i(t_0) - S_i^*$ decays exponentially, and the system returns to its homogeneous equilibrium state of S_1^* and S_2^*. Spatial homogeneity is then stable under these conditions. This happens when the trace of J is negative and the determinant is positive, or

$$\begin{aligned} & tr(J) = a_{11} + a_{22} < 0 \\ \text{and} \quad & \\ & \det(J) = a_{11} a_{22} - a_{12} a_{21} > 0. \end{aligned} \tag{4.7}$$

Recall that a_{11} represents the growth of S_1 with respect to itself, or the autocatalytic/positive feedback in the system, and a_{22} represents mortality of the predator (S_2) with respect to itself.

Now add diffusion terms to each equation (for simplicity, we will consider diffusion in only one lateral direction):

$$\begin{cases} \dfrac{\partial S_1}{\partial t} = F_1(S_1, S_2) + D_1 \dfrac{\partial^2 S_1}{\partial x^2} \\ \dfrac{\partial S_2}{\partial t} = F_2(S_1, S_2) + D_2 \dfrac{\partial^2 S_2}{\partial x^2}, \end{cases} \tag{4.8}$$

where D_1 and D_2 are rates of random spread or Fickian diffusion across space (x), and the partial derivatives with respect to x represent density or concentration gradients of S_1 and S_2 across space. (In much of the literature on reaction-diffusion equations, the prey is termed the "activator" because of the positive feedback, and the predator is termed the "inhibitor" because it consumes the prey, but I will continue to use the terms prey and predator in the general sense as defined above).

Perturbations to this spatially explicitly model (such as changing the population density of either species, corresponding, e.g., to an outbreak, an irruption, stocking, or harvesting) are introduced not simply at a point in time but at a point in both space and time. Furthermore, the perturbation propagates in space because the diffusion terms "spread" the perturbed population out in the x direction. The perturbation is further modified by the interactions between the two species who spread or diffuse at different rates. The Jacobian now becomes:

$$J_{spatial} = \begin{bmatrix} a_{11} - D_1\sigma^2 & a_{12} \\ a_{21} & a_{22} - D_2\sigma^2 \end{bmatrix}, \tag{4.9}$$

where σ is the wavenumber, or the number of a peak in population density assigned in increasing order away from the initial peak that was the perturbation. σ is proportional to 2π/distance between the peaks. The decay or growth of these perturbations is then necessarily a function of both space and time and is approximated by:

$$\begin{bmatrix} S_1(x,t) \\ S_2(x,t) \end{bmatrix} = \sum_n \begin{bmatrix} S_1(x_0,t_0) - S_1^* \\ S_2(x_0,t_0) - S_2^* \end{bmatrix} c_n \omega_n e^{\lambda_n t} \cos \sigma x. \tag{4.10}$$

Note the addition of the term $\cos \sigma x$ in comparison with Equation (4.6); this ensures that the fate of the disturbance depends both on space and on time. Again, the coexistence between S_1 and S_2 is stable and spatially homogeneous when the trace of J_{spatial} is negative, the determinant is positive, and hence the real parts of λ are all negative. These conditions obviously depend on the relative sizes of D_1 and D_2.

Assume that a perturbation is introduced at a point x_0, t_0. If $D_1 = D_2$, then some simple algebra shows that the heterogeneity introduced by the disturbance decays. Consequently, the spatially homogeneous distribution is stable with equilibrium values S_1^* and S_2^* [see Okubo and Levin (2002) for mathematical details and proofs].

But when $D_2 > D_1$ and D_2/D_1 is greater than some crucial value C, then the homogeneous steady-state distribution is not stable (the determinant becomes negative), and diffusive instability sets in. Spatial heterogeneity, rather than homogeneity, becomes the stable state of the system, and the disturbance propagates across space. Under these conditions, the two coexisting species are distributed heterogeneously across the landscape. Eventually, their distribution approaches a stable patterned heterogeneity [see Okubo and Levin (2002) for mathematical details and proofs].

The crucial value by which C must be exceeded for patterned heterogeneities to develop varies with functions F_1 and F_2, but in general $C = f(a_{22}/a_{11})$. Therefore, if

$$\frac{D_2}{D_1} > f\left(\frac{a_{22}}{a_{11}}\right) > 1, \tag{4.11}$$

then spatial homogeneity of two interacting populations of different trophic levels is unstable, and spatial heterogeneity of the two interacting populations is stable. In other words, for patterned heterogeneity to be stable: (1) the diffusion rate of the predator must be greater than that of the prey and greater than some function of the ratio of per capita mortality of the predator to per capita growth of the prey; (2) the growth of the prey (at least at low population densities) must involve a positive feedback within its own population or with some underlying ecosystem property [plant litter-nutrient availabilities discussed by Meinders and van Breemen (this volume) could be one such feedback]; and (3) an increase in predator densities decreases prey density through consumption, and therefore eventually predator densities as well.

To see how this works, first consider a stable predator-prey system without diffusion. A random increase in prey density at a point in a landscape results in a further increase in both its density and that of the prey, but increased predator density at the point of random increase in prey density reduces the prey and is also self-limiting through mortality. The system is thereby stabilized, and the perturbation in prey density at the point of the disturbance dies away exponentially.

Introduction of diffusion terms dissipates the negative effect of the predator. If the diffusion rate of the predator is sufficiently greater than that of the prey ($D_2 > D_1C$), then a local randomly introduced peak in prey density can grow because of autocatalysis or positive feedbacks to its population. The predator will be able to track the peak in prey density, causing "dents" to appear and separating the initial peak into two, which then grow by autocatalysis and the process repeats. Depending on the magnitudes of D_i in both x and y directions and the exact form and magnitude of C, a rich variety of patterned spatial heterogeneities can develop (Okubo 1978; Murray 1989; Holmes et al. 1994).

The most surprising aspect of this theory is that these spatial heterogeneities are due entirely to the interactions of the two components diffusing randomly at different rates and not necessarily due to any persistent heterogeneity in the underlying environment or preferred directional flow of one or both species. If there are positive feedbacks in the growth of the prey population and greater rates of diffusion of predator than prey, then spatial patterns (heterogeneities) are almost inevitable. Thus, neither nonrandom foraging of a predator nor underlying environmental heterogeneity is required to produce spatial patterns in generalized predator-prey systems. This is not to say that predators necessarily forage at random nor does it deny the existence of underlying environmental heterogeneities. Rather, such underlying heterogeneities, if present, can modify the patterns further, and the mere presence alone of a pattern is not sufficient to invoke them.

Further theoretical explorations of this mechanism of generating spatial heterogeneity have been developed. As opposed to predator-prey models, Levin (1974) showed that diffusive instability cannot occur in simple two-species Lotka-Volterra competition models with diffusion, but Evans (1980) showed that it happens in three-species Lotka-Volterra competition models. Powell and Richerson (1985) showed that diffusive instability and pattern formation can happen between two species competing for two resources if the dynamics of both species and their resources are all modeled.

This mathematical approach has found applications in various ecological settings, beginning with ocean systems. Malchow (2000) gives an extensive review of recent developments in the theory of pattern formation in aquatic systems. Diffusive instability was first proposed to explain fine-scale spatial heterogeneities of herbivorous zooplankton and phytoplankton in the oceans by Segal and Jackson (1972) and independently by Steele (1974) and developed further by DuBois (1975) and Levin and Segal (1976). Later observations showed that both fine- and coarse-scale patchiness of zooplankton and phytoplankton require not only diffusive instability but also directional gradients caused by currents and gyres (Weber et al. 1986; Mahadevan this volume).

Levin (1977) extended the development of this approach by showing that a positive feedback in the prey is not necessary if the predator consumes prey according to a saturating function, such as a Michaelis-Menten function. Okubo (1978) then showed that diffusive instability can occur between phytoplankton and the concentrations of limiting nutrients in the water column if one assumes that the phytoplankton take up nutrients in a Michaelis-Menten function and herbivores are a constant sink for the phytoplankton.

This mechanism of generating patterned spatial heterogeneity is probably not confined to aquatic systems, even though it has been more extensively investigated in such systems. One aspect of the above examples to notice is that a herbivore is present in all of them. Some recent studies also indicate that pattern formation through diffusive instability can arise in terrestrial systems with herbivores. Maron and Harrison (1997) showed that tussock moths attain stable, locally high densities even though they disperse faster than their host plants because of the even faster dispersal of a more mobile parasitoid, thus introducing the possibility of diffusive instability in a plant-herbivore-parasatoid system. Pastor et al. (1999) showed that foraging by mobile model moose in a model landscape that was initially random eventually produced spatial patterns characteristic of diffusive instability. These theoretical patterns also conformed to field measurements made on Isle Royale (Pastor et al. 1998). Because the patterns that develop affect the energy balance of the mobile moose, only certain foraging strategies produced landscapes in which food was distributed in such a pattern that the moose sustained positive energy balances and thereby survived. Therefore, diffusive instability can produce spatially heterogeneous landscapes that can either be detrimental or crucial to the energy balance of foraging animals and thus the survival of their populations.

Terrestrial herbivore populations almost always disperse faster than their forage species disperse seeds or propagules. If it is also common that a forage species is part of a positive feedback with soil properties (Meinders and van Breemen this volume), then spatial heterogeneity would seem to be common in terrestrial ecosystems where herbivores have strong effects on plant community composition and nutrient cycles. If one is working in an ecosystem in which herbivores exert strong control over species composition, nutrient cycling rates, or both, then one should immediately suspect diffusive instability as a possible source of any patterns one finds.

Diffusive instability and spatial pattern formation through trophic interactions is currently an area of theoretical research rich with nontrivial predictions that can be tested experimentally. Some of these experiments may involve long-term observations to determine the scales over which spatial patterns arise (e.g., Grünbaum 1992; Pastor et al. 1998) or to determine if spatial heterogeneities change with time (e.g., Pastor et al. 1999). Long Term Ecological Research (LTER) sites, the Joint Global Ocean Flux Study (JGOFS) sites, and other sites with repeatedly monitored observation grids are possible sites to gather data to refine and test these theories.

When Is Spatial Heterogeneity Important?

The above considerations beg the questions that the organizers of this conference have explicitly posed: When is spatial heterogeneity important? When is it not important?

These are difficult questions. In part, the answers depend on what is meant by "important." For example, to a moose walking across a landscape, the conditions in the next step may be important (e.g., whether or not there is edible food there). They may also be important to a behavioral ecologist trying to construct individual-based models of moose foraging. But whether or not they become important at population, ecosystem, and landscape levels depends on positive and negative feedbacks between the moose and plant growth and whether the recovery time of the browsed plant is longer or shorter than the average return time of a moose to each plant (Moen et al. 1998). Thus, the importance of spatial heterogeneity depends on the scale of the question being asked, a point made numerous times in the recent ecological literature and at this conference as well.

Disturbances create spatial heterogeneity in all systems almost by definition. To a crude first approximation, disturbances can be considered random losses of a certain percentage of biomass, easily modeled through stochastic linear processes such as Markov chains. Even when the dynamics are random and linear and therefore simple, they can be "important." Certainly, the loss of 75% of the biomass of an ecosystem over some mean recurrence interval and distributed more or less randomly over the landscape has large effects on ecosystem properties. We have gained a great deal of understanding of disturbances in ecosystems and landscapes through the application of linear data analysis and modeling techniques. Nonetheless, perhaps the more interesting and fruitful avenues for further exploration involve higher order effects of disturbances on landscape patterns, taking into account how and when heterogeneities created by one disturbance influence the spread and nature of future disturbances (see Romme this volume). Do these effects differ for different ecosystems? Do they depend on mean turnover rate of biomass or nutrient capital within the ecosystem, the rate of dispersal of component species, or the degree of mixing of materials, climatic conditions, or other forcing functions?

The spatial heterogeneity created by directional flows appears to be important when it affects the mass balance of materials in a local neighborhood: the position of the local neighborhood with respect to surrounding neighborhoods that deliver or receive materials from it must then be taken into account. This is particularly important when the materials limit growth, such as water, nitrogen, or photons of light, and especially when they are amplified by positive feedbacks within the local neighborhood, such as the formation of patterned communities in peatlands and arid lands. But we have much more to learn about this. When do lateral transfers become important and for what property or process? Is there a particular ratio of lateral inputs to internal rates of cycling above which we must consider position in the landscape and below which these lateral inputs can be ignored? Are there particular positions in the landscape such as riparian zones for which these lateral flows cannot be ignored? Do the importance of lateral flows increase "down gradient"?

The patterned heterogeneities created when positive feedbacks are coupled with different rates of diffusion between interacting trophic levels are important when they modify the success of individuals or populations of each trophic level in obtaining needed resources. This has obvious evolutionary implications, because it means that the landscape of selection pressures is dynamic precisely because of the interactions of individuals searching for food. Such dynamics may particularly be important in ecosystems in which herbivores control plant species composition and the cycling of nutrients and energy, but again we need to refine further these considerations. Does it matter how much the herbivore consumes? Or does the rate of recovery of plants from herbivore consumption matter even more? Or do both matter?

If nothing else, the papers of this conference show that we are only at the outset of being able to define the questions of how spatial heterogeneity is created in ecosystems and what are the consequences of it. Making cross-system comparisons will depend to what extent such questions can be more precisely defined so that experimental approaches can be brought to bear on them. The rich array of theoretical approaches to heterogeneity discussed above may prove useful in helping to define these questions.

References

Aselmann, I., and Crutzen, P.J. 1989. Global distribution of natural freshwater wetlands, their net primary productivity, seasonality, and possible methane emissions. J. Atmospheric Chem. 8: 307–359.

Bak, P. 1997. How nature works. New York: Springer-Verlag.

Baker, W. 1989. A review of models of landscape change. Landscape Ecol. 2: 111–133.

Bartlett, K.B., and Harriss, R.C. 1993. Review and assessment of methane emissions from wetlands. Chemosphere 26: 261–320.

Bormann, F.H., Likens, G.E., and Melillo, J.M. 1977. Nitrogen budget for an aggrading northern hardwood forest ecosystem. Science 196: 981–983.

Bridgham, S.D., Johnston, C.A., Pastor, J., and Updegraff, K. 1995. Potential feedbacks of northern wetlands on climate change. BioScience 45: 262–274.

Bridgham, S.D., Ping, C.-L., Richardson, J.L., and Updegraff, K. 2001. Soils of northern peatlands: histosols and gelisols. In Wetland soils: genesis, hydrology, landscapes, and classification, eds. J.L. Richardson and M.J. Vepraskas, pp. 343–370. Boca Raton, FL: CRC Press.

Chapin, F.S., Vitousek, P.M., and Van Cleve, K. 1986. The nature of nutrient limitation in plant communities. Am. Naturalist 127: 48–58.

DuBois, D.M. 1975. A model of patchiness for prey-predator phytoplankton populations. Ecol. Modelling 1: 67–80.

Edelstein-Keshet, L. 1988. Mathematical models in biology. New York: Random House.

Evans, G.T. 1980. Diffusive structure: counter examples to any explanation. J. Theor. Biol. 82: 313–315.

Flanagan, P.W., and Van Cleve, K. 1983. Nutrient cycling in relation to decomposition and organic-matter quality in taiga ecosystems. Can. J. Forest Res. 13: 795–817.

Glaser, P. 1992. Ecological development of patterned peatlands. In Patterned peatlands of minnesota, eds. H.E. Wright, B.A. Coffin, and N.E. Aaseng, pp. 27–42. Minneapolis: University of Minnesota Press.

Gorham, E. 1991. Northern peatlands: role in the global carbon cycle and possible responses to climatic warming. Ecol. Applications 1: 182–195.

Grünbaum, D. 1992. Local processes and global patterns: biomathematical models of bryozoan feeding currents and density dependent aggregation in Antarctic krill. Ph.D. Thesis, Ithaca, NY: Cornell University.

Heinselmann, M.L. 1973. Fire in the virgin forests of the Boundary Waters Canoe Area, Minnesota. Quaternary Res. 3: 329–382.

Holling, C.S., ed. 1978. Adaptive environmental assessment and management. New York: John Wiley & Sons.

Holmes, E.E., Lewis, M.A., Banks, J.E., and Veit, R.R. 1994. Partial differential equations in ecology: spatial interactions and population dynamics. Ecology 75: 17–29.

Kauffman, S. 1993. The origins of order. Oxford: Oxford University Press.

Levin, S.A. 1974. Dispersion and population interactions. Am. Naturalist 108: 207–228.

Levin, S.A. 1977. A more functional response to predator-prey stability. Am. Naturalist 111: 381–383.

Levin, S.A. 2003. Complex adaptive systems: exploring the known, the unknown, and the unknowable. Bull. Am. Math. Soc. 40: 3–20.

Levin, S.A., and Segal, L.A. 1976. An hypothesis for the origin of planktonic patchiness. Nature 259: 659.

Likens, G.E., Bormann, F.H., Johnson, N.M., Fisher, D.W., and Pierce, R.S. 1970. Effects of forest cutting and herbicide treatment on nutrient budgets in the Hubbard Brook Watershed-Ecosystem. Ecol. Monogr. 40: 23–47.

Malchow, H. 2000. Non-equilibrium spatio-temporal patterns in models of non-linear plankton dynamics. Freshwater Biol. 45: 239–251.

Maron, J.L., and Harrison, S. 1997. Spatial pattern formation in an insect host-parasitoid system. Science 278: 1619–1621.

Mathews, E., and Fung, I. 1987. Methane emissions from natural wetlands: global distribution, area, and environmental characteristics of sources. Global Biogeochem. Cycles 1: 61–86.

Moen, R., Cohen, Y., and Pastor, J. 1998. Evaluating foraging strategies with a moose energetics model. Ecosystems 1: 52–63.

Murray, J.D. 1989. Mathematical Biology. New York: Springer-Verlag.

Neuhauser, C. 2001. Mathematical challenges in spatial ecology. Notices Am. Math. Soc. 48: 1304–1314.

Niemi, G.J., DeVore, P., Detenbeck, N., Taylor, D., Yount, J.D., Lima, A., Pastor, J., and Naiman, R.J. 1990. Overview of case studies on recovery of aquatic ecosystems from disturbance. Environ. Manage. 14: 571–588.

Okubo, A. 1978. Horizontal dispersion and critical scales for phytoplankton patches. In Spatial pattern in plankton communities, ed. J.H. Steele, pp. 21–42. New York: Plenum.

Okubo, A., and Levin, S.A. 2002. Diffusion and ecological problems. New York: Springer-Verlag.

Pastor, J., and Bockheim, J.G. 1984. Distribution and cycling of nutrients in an aspen-mixed hardwood-spodosol ecosystem in northern Wisconsin. Ecology 65: 339–353.

Pastor, J., Bonde, J., Johnston, C.A., and Naiman, R.J. 1993. A Markovian analysis of the spatially dependent dynamics of beaver ponds. In Theoretical approaches for predicting spatial effects in ecological systems pp. 5–27 ed. R.H. Gardner. Lectures on Mathematics in the Life Sciences, Vol. 23. American Mathematical Society. Providence, RI.

Pastor, J., Dewey, B., Moen, R., White, M., Mladenoff, D., and Cohen. Y. 1998. Spatial patterns in the moose-forest-soil ecosystem on Isle Royale, Michigan, USA. Ecol. Applications 8: 411–424.

Pastor, J., Cohen, Y., and Moen, R. 1999. The generation of spatial patterns in boreal landscapes. Ecosystems 2: 439–450.

Post, W.M., Emanuel, W.R., Zinke, P.J., and Stangenberger, A.G. 1982. Soil carbon pools and world life zones. Nature 298: 156–159.

Post, W.M., Pastor, J., Zinke, P., and Stangenberger, A. 1985. Global patterns of soil nitrogen storage. Nature 317: 613–616.

Powell, T.M., and Richerson, P.J. 1985. Temporal variation, spatial heterogeneity, and competition for resources in plankton systems: a theoretical model. Am. Naturalist 125: 431–464.

Segal, L.A., and Jackson, J.L. 1972. Dissipative structure: an explanation and an ecological example. J. Theor. Biol. 37: 545–559.

Segal, L.A., and Levin, S.A. 1976. Applications of nonlinear stability theory to the study of the effects of diffusion on predator-prey interactions. In Topics in statistical mechanics and biophysics: a memorial to Julius L. Jackson, pp. 123–152. New York: American Institute of Physics.

Shugart, H.H. 1984. A theory of forest dynamics. New York: Springer-Verlag.

Steele, J.H. 1974. Spatial heterogeneity and population stability. Nature 248: 83.

Tilman, D. 1988. Dynamics and structure of plant communities. Princeton, NJ: Princeton University Press.

Turing, A.M. 1952. The chemical basis of morphogenesis. Philos. Trans. R. Soc. London Ser. B 237: 37–52.

van Breemen, N. 1995. How *Sphagnum* bogs down other plants. TREE 10: 270–275.

Weber, L.H., El-Sayed, S.Z., and Hampton, I. 1986. The variance spectra of phytoplankton, krill and water temperature in the Antarctic Ocean south of Africa. Deep-Sea Res. 33: 1327–1343.

Wright, H.E., Coffin, B.A., and Aaseng, N.E., eds. 1992. Patterned peatlands of Minnesota. Minneapolis: University of Minnesota Press.