THE IMPORTANCE OF SCALE IN TESTING THE ORIGINS OF ALTERNATIVE COMMUNITY STATES

PETER S. PETRAITIS1 AND ROGER EARL LATHAM2

1Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6018 USA
2Department of Biology, Swarthmore College, Swarthmore, Pennsylvania 19081-1397 USA

Abstract. The possibility that different species assemblages may represent persistent alternative community states remains largely unexplored by experimental ecologists because of a variety of conceptual and experimental problems. We discuss some of the conceptual roadblocks to experimentation and propose several avenues for attacking the problem experimentally. We address the conceptual issues involved in (1) the blurring of the distinction between the processes that initiate the switch among alternative states and the positive-feedback processes that maintain those states, and (2) the role of spatial scale in initiating the switch.

We suggest that the switch between alternative states requires, first, a disturbance that removes species involved in the positive feedbacks needed for maintenance and, second, the arrival of other individuals that initiate the switch to the alternative assemblage. The removal of the species that maintain the system must be large enough and over a long enough time to allow the arrival and establishment of members of the alternative assemblage, and so we hypothesize that the switch among alternative states is scale dependent.

This scenario suggests that the switch among alternative states can be investigated experimentally through the manipulation of the scale of the disturbance and of the arrival of members of the alternative state. Small-scale disturbances should consistently fail to initiate a switch, while larger-scale events should initiate a switch at least part of the time. We also note that in some cases the scale of disturbance and/or the arrival of recruits cannot be manipulated or controlled and suggest that several approaches other than factorial experiments with ANOVA, such as spatial autocorrelation methods, may be useful.

We illustrate the potential and the difficulties of various approaches by discussing two systems in eastern North America that may contain alternative states. Mosaics of mussel beds and algal beds occupy rocky coasts from New England northward, and patchworks of forests and heathlands occur in eastern Canada and in the Appalachian highlands. While the study of alternative states in the marine system can be approached experimentally, the scale of disturbance required to switch forests to heathlands is too large for experimentation and must rely on the use of other approaches.

Key words: algae; alternative community states; Ascophyllum; community ecology; disturbance; Ericaceae; fire ecology; marine intertidal systems; mussels; scale-dependent processes; spatial autocorrelation; species assemblages, alternative.

INTRODUCTION

Models of multiple stable states or alternative community states (Noy-Meir 1975, Gilpin and Case 1976, May 1977) are often used to explain why assemblages of different species can occur side by side in apparently the same environment (e.g., Jackson 1968, Damman 1975, Kellman 1979, Streng and Harcombe 1982, Read 1984, Myers 1985, Berdowski and Zeilinga 1987, Ash 1988, Geldenhuys 1994). In theory, alternative assemblages of species arise from differences in starting conditions or from perturbations that push species densities off of one equilibrium point and into the basin of attraction of another equilibrium point (Lewontin 1969, Austin and Cook 1974, Noy-Meir 1975, Gilpin and Case 1976, May 1977, Knowlton 1992). Different initial densities dictate the outcome in these deterministic models. For experimentalists, this means that history matters (e.g., Hughes 1989).

While there have been attempts to test the theory in
the field (e.g., Sutherland 1974, Simenstad et al. 1978, Hatcher 1984, Dublin et al. 1990, D’Antonio and Vittousek 1992, Blindow et al. 1993), most have not tested the theoretical predictions in a rigorous fashion (Connell and Sousa 1983). Why is this so? Certainly part of the problem is the difficulty of taking mathematical theory into “the field.” In addition, we believe there are several conceptual obstacles to the development of experimental tests, and will address two. The first, noted by Drake (1991), but not widely appreciated, is a blurring of the distinction between the origin and the persistence (maintenance) of assemblages. The questions of origin vs. maintenance are quite different and, we believe, require different sorts of experiments. The second problem is the unappreciated role of spatial scale, and we suggest that good tests of the origins of alternative community states require scale-dependent experimentation. The role of scale has been widely discussed in other contexts (e.g., Paine and Levin 1981, O’Neill et al. 1986, Foster 1990, Holling 1992, Dunham and Beaupre 1998) and may sometimes be trivial and tautological (e.g., the notion that bigger experiments always capture nature better). However, the underlying ecological reasons for why a good test of alternative community states requires a large experimental scale seem not to be well understood.

We attempt to clarify these issues and suggest experimental protocols for testing the origins of alternative community states. We make no attempt to review the literature on alternative states since reviews exist (Laycock 1991, Knowlton 1992, Wilson and Agnew 1992). We do not try to resolve what is meant by “stable,” “permanent,” or “persistent” (see Frank 1968, Sutherland 1974, 1990, Sousa and Connell 1983) nor critique Connell and Sousa’s (1983) criteria for testing alternative states and the subsequent exchanges among Peterson (1984), Sousa and Connell (1985) and Sutherland (1990). We also do not see our thesis as being complete in every respect, rather this paper is an outline of what we believe are some of the major issues. The first section gives what we believe are the minimal criteria for testing the origin of alternative states. We then discuss origin vs. maintenance of alternative community states and the role of spatial scale, which we believe is the primary reason for why good tests in the field will usually require experiments involving the manipulation of scale. We also examine the statistical and logistical limitations of such experiments and suggest explicit experimental protocols for testing the origin of alternative community states.

We conclude with two examples—the rocky shores of the North Atlantic and the cool-temperate heathlands of northern North America. Historically, closely similar systems have been the source of many of the current ideas about alternative community states. Work by marine ecologists is widely cited in discussions of alternative states (e.g., Sutherland 1974, 1990, Connell and Sousa 1983, Peterson 1984, Sousa and Connell 1985, Knowlton 1992, Law and Morton 1993), but the idea that some plant assemblages may be dependent on wildfire and actually promote its spread was one of the earliest expressions of the concept of positive-feedback switching between alternative community states (Mount 1964, Jackson 1968, Mutch 1970). We believe these systems are fertile ground for testing the alternative-community-states hypothesis and they also provide good examples of the problems inherent in such tests.

Criteria for Experimental Tests

In practice, it is not clear when different assemblages of species can be considered alternative community states (Sutherland 1974, 1990, Connell and Sousa 1983, Peterson 1984, Sousa and Connell 1985). Theory addresses conditions at equilibrium in a constant environment that is closed to migration. Experimentalists, however, deal with natural populations that are not at equilibrium, not under constant environmental conditions, and open to migration. Theoretical treatments of alternative community states rely on traditional models of per capita rates of change as functions of densities. However, density may not always be the variable of interest, and experimentalists typically measure changes over time periods shorter than one generation. For example, alternative assemblages often appear to be structured by large, long-lived species (e.g., Dayton’s [1972] idea of foundation species) that may be present in all alternative states, but as small, suppressed individuals in some assemblages and large, dominant individuals in others. In this case, densities tell us little about the alternative states, and biomass far more. Here, as in other examples, there is not an explicit link between the theoretical parameters and the experimental measures, and thus, as May (1977:471) suggests, “the empirical observations [for alternative community states] remain largely anecdotal, and the theory remains largely metaphorical.”

Several authors have debated the appropriate criteria for testing the existence of alternative community states (Sutherland 1974, 1990, Connell and Sousa 1983, Peterson 1984, Sousa and Connell 1985). We agree with Peterson (1984:128), who concludes, “only by experiment could one convincingly demonstrate multiple stable states by showing that the very same site could come to be occupied by different, self-replicating communities.” The “very same site” condition is difficult to prove since feedback loops driven by some species alter the environment (Peterson 1984) and, in any case, the demonstration of “same environment” presents the difficult statistical problem of accepting a null hypothesis with a high level of confidence (e.g., see Cohen 1988). There also seems to be a consensus that the replacement of one alternative state by another requires a perturbation or “pulse” event (Bender et al. 1984) changing the species densities or environmental

430

PETER S. PETRAITIS AND ROGER EARL LATHAM

Ecology, Vol. 80, No. 2
TESTING ALTERNATIVE COMMUNITY STATES

Fig. 1. Switching from a stable state dominated by species A to an alternative state dominated by species B under two different scenarios. The left phase diagram shows only two species, A and B, but these may be foundation species (sensu Dayton 1972) for other species associated with the two alternative states. Solid circles mark the stable states, and open and dotted areas the two basins of attraction; in the dotted area the system returns to the lower equilibrium point where species A is most common, and in the white area it goes to the upper point where species B is most common. Note that the boundary between the two alternative states is shown as a smooth curve, although in theory the boundary may be very complex or even fractal (see Hastings 1993). In the case of fractal basin boundaries, it is not possible to predict which alternative state will prevail unless the initial conditions are known precisely. The first scenario involves recruitment of species B, which kills A and facilitates the arrival of other species (line 1). The second scenario requires a disturbance that removes A (line 2) and the proliferation of B from new arrivals or previously suppressed individuals already present (line 3). The right diagram illustrates a third scenario suggested by J. A. Drake (personal communication) showing a uniform decline in densities causing a switch between basins of attraction. It is perhaps easier to envision ecological conditions and populations that would play out the first two scenarios than those required for the third.

conditions, which in turn starts a set of positive feedback mechanisms resulting in a switch.

ORIGIN VS. MAINTENANCE OF SPECIES ASSEMBLAGES

We believe that sound tests of alternative community states can be based on Peterson’s dictum but that they require the treatment of origin and maintenance (self-sustainability) as separate issues. Plant ecologists have recognized the dichotomy between origin and maintenance for some time (e.g., Jackson 1968, Ugolini and Mann 1979, Westoby 1979, see review by Wilson and Agnew [1992]). They distinguish between the positive feedback mechanisms that sustain the alternative states (i.e., the problem of maintenance) and the initial environment (i.e., the problem of origin).

Tests of maintenance require showing that an established assemblage persists beyond the lifetime of individuals of species that contribute substantially to the biomass of the assemblage. To ensure that alternative states are self-sustaining, experiments and observations must be carried out over a sufficient time period (see Sutherland 1974, 1990, Sousa and Connell 1985). We concur with Sousa and Connell (1985) that the assemblages must persist for at least one generational turnover of dominant species. Assemblages dominated by long-lived plants or by clonal organisms (shrubs or corals) of unknown but very great longevity pose particular problems. We think indirect tests of self-replication can be done, but these must be viewed as separate from tests of origin and are beyond the scope of this paper.

Tests of origin require showing that pulse events can initiate the development of at least two assemblages of different organisms. These alternative states may or may not be stable, self-sustaining, or self-replicating. In theory, a switch between states occurs when there is a change in species densities that perturbs one assemblage far enough off its equilibrium point that the perturbed assemblage is now in the basin of attraction of the equilibrium point of an alternative assemblage (see Fig. 1). The question is what, in practice, pushes the species composition of one assemblage not only off its equilibrium point but also into the basin of attraction of another assemblage? We believe there are two prototypical scenarios (Fig. 1).

In the first situation, an influx of recruits initiates the switch by decimating some of the species in the original assemblage and facilitating the establishment of additional recruits of species that occur in the alternative assemblage. In the second scenario, a disturbance kills members of the original assemblage and opens the system to establishment of individuals of different species. The first individuals to establish facilitate the recruitment of other members of the alternative assemblage. This scenario includes cases where the disturbance selectively kills off the dominants and members of the alternative assemblage are already present in a suppressed state. All they need to do is grow. Suppressed individuals may be present in the original assemblage but account for a small fraction of the total biomass until a disturbance occurs. The occurrence of suppressed individuals and their release by disturbance is probably the norm in many terrestrial systems. In both...
cases, early colonizers may alter the environment in ways that favor the establishment of competitively superior species (negative feedback), and late-stage dominant species foster environmental changes that favor their own persistence (positive feedback).

The two scenarios are not mutually exclusive and differ in whether the initial recruits are agents of mortality. The first scenario has the initial recruits both killing and facilitating members of other species, while the second scenario uncouples these two events. To the extent that the initial arrival of invaders may be seen as a biotic disturbance, then the distinction between the two scenarios becomes blurred. It is possible, for example, that an assemblage could be invaded by a species that is not a member of the alternative assemblage but does kill members of the original assemblage. Here the initial recruits are nothing more than a biotic disturbance. It is also possible that a subtle disturbance (e.g., nutrient addition or fire suppression), which is not the key pulse event, may disrupt maintenance mechanisms and allow colonization by one or more species from the alternative assemblage. If the newcomers act as a biotic disturbance then they would magnify the effect of the original disturbance and the species composition could move into the basin of attraction of the alternative assemblage. For the remainder of the paper, we will focus on the second scenario, for ecological and practical reasons. On the ecological side, we believe the second scenario is more common, particularly in assemblages with size hierarchies, and it seems unlikely that new recruits can be agents of mortality and overwhelm large, established individuals, which is required in the first scenario. Also, many communities are altered by abiotic disturbances that kill members of resident species (e.g., fire and landslides in plant assemblages, ice and sand scour in marine intertidal systems); they have attracted most of the attention as candidates for alternative community states. It seems reasonable, therefore, to focus on constructing good experimental tests for these systems.

On the practical side, we can envision how to test experimentally the second but not the first scenario. Ecologists have a rich tradition of mimicking natural disturbances experimentally (e.g., Bormann et al. 1968, Dayton 1971, 1975, Christensen and Kimber 1975, Sousa 1979, MacGillivray et al. 1995; see also references in Petraitis et al. [1989]), and these methods can be easily exploited for the study of alternative community states.

**THE ROLE OF SCALE IN THE ORIGIN OF SPECIES ASSEMBLAGES**

The death or removal of dominant individuals must provide sufficient time and space for members of the alternative assemblage—whether new arrivals or previously suppressed individuals already present—to gain a foothold and initiate the pivotal positive feedback processes. These spatial and temporal windows of opportunity must occur, regardless of the agent of mortality. If there are too few deaths or removals, the assemblage may revert to the original state. We hypothesize that large-scale disturbances provide sufficient space and time by increasing the area immune to the edge effects that would tend to reestablish the original assemblage. If the spatial scale of the pulse event is large enough and post-disturbance recruitment is sufficiently rapid, then there will be enough space and time for the establishment of an alternative assemblage. The appropriate spatial or temporal scale of pulse events is, of course, relative and depends on the dynamics of recruitment and recovery within each particular ecosystem. There is an added requirement that assemblages switch abruptly at a critical threshold of disturbance (Knowlton 1992). If species composition changes gradually as the spatial extent or temporal duration of the disturbance lengthens, there would not be distinct patches of different species assemblages, but a gradient (e.g., see Whittaker 1967).

We predict that these thresholds will turn out to be spatially dependent in most ecosystems. Small disturbances that are below the required threshold will return to the original state. Large clearings or disturbances may switch to an alternative state because a large fraction of the disturbed area lies beyond edge effects from the surrounding assemblage. Thus “historical accidents” or contingencies (e.g., arrival of recruits that are not part of the surrounding assemblage) have more time and space in which to occur, and members of alternative assemblages are more likely to establish a foothold.

**Tests of Origin of Alternative Community States**

Our second scenario for the origin of alternative states requires two steps. First, a disturbance kills some portion of the resident species and, second, a recruitment event begins the establishment of the alternative state. Our ability to manipulate disturbance and recruitment depends on the natural scales of these processes (see Table 1). At one extreme are ecosystems in which natural disturbances are likely to be small in extent and frequent in occurrence, and recruits numerous and easily handled. It should be relatively simple to experimentally manipulate disturbances and recruitment on the scale of natural events. At the other ex-

<table>
<thead>
<tr>
<th>Size of disturbance can be manipulated?</th>
<th>Recruitment can be manipulated or is not limiting?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>ANOVA</td>
</tr>
<tr>
<td>No</td>
<td>ANCOVA</td>
</tr>
</tbody>
</table>

The death or removal of dominant individuals must provide sufficient time and space for members of the alternative assemblage—whether new arrivals or previously suppressed individuals already present—to gain a foothold and initiate the pivotal positive feedback processes. These spatial and temporal windows of opportunity must occur, regardless of the agent of mortality. If there are too few deaths or removals, the assemblage may revert to the original state. We hypothesize that large-scale disturbances provide sufficient space and time by increasing the area immune to the edge effects that would tend to reestablish the original assemblage. If the spatial scale of the pulse event is large enough and post-disturbance recruitment is sufficiently rapid, then there will be enough space and time for the establishment of an alternative assemblage. The appropriate spatial or temporal scale of pulse events is, of course, relative and depends on the dynamics of recruitment and recovery within each particular ecosystem. There is an added requirement that assemblages switch abruptly at a critical threshold of disturbance (Knowlton 1992). If species composition changes gradually as the spatial extent or temporal duration of the disturbance lengthens, there would not be distinct patches of different species assemblages, but a gradient (e.g., see Whittaker 1967).

We predict that these thresholds will turn out to be spatially dependent in most ecosystems. Small disturbances that are below the required threshold will return to the original state. Large clearings or disturbances may switch to an alternative state because a large fraction of the disturbed area lies beyond edge effects from the surrounding assemblage. Thus “historical accidents” or contingencies (e.g., arrival of recruits that are not part of the surrounding assemblage) have more time and space in which to occur, and members of alternative assemblages are more likely to establish a foothold.

**Tests of Origin of Alternative Community States**

Our second scenario for the origin of alternative states requires two steps. First, a disturbance kills some portion of the resident species and, second, a recruitment event begins the establishment of the alternative state. Our ability to manipulate disturbance and recruitment depends on the natural scales of these processes (see Table 1). At one extreme are ecosystems in which natural disturbances are likely to be small in extent and frequent in occurrence, and recruits numerous and easily handled. It should be relatively simple to experimentally manipulate disturbances and recruitment on the scale of natural events. At the other ex-
treme are ecosystems in which natural disturbances are large, severe, or infrequent, and recruits are rare, unpredictable in abundance, or difficult to handle. It is unlikely that disturbance or recruitment could be easily manipulated. Direct experimental tests would be very difficult if not impossible, and comparative approaches must be used to examine the origins of alternative states. In between are systems in which disturbance can be manipulated but recruitment cannot be, or vice versa. Tests of the origin of alternative states in these ecosystems will require approaches that combine experimental and observational studies.

**Situations in which disturbance and recruitment can be manipulated**

The core of the experimental design involves clearing areas of various sizes in established assemblages and ensuring a sufficient supply of recruits. If we can mimic the disturbance and control the number and kind of recruits following the disturbance, and if two assemblages are alternative states, then we should be able to shift between the assemblages under the appropriate conditions. We predict that we could shift between the assemblages if the disturbance was large enough to minimize edge effects and if a sufficient number of recruits is available (see Fig. 2). Thus the response depends on the size of the clearing. Small clearings would return to the original state, while larger clearings would switch to the alternative state if there were sufficient recruits of the right kind.

It is possible that clearings within the alternative assemblages will shift gradually from one assemblage to the other from the edge toward the center, with changes in the area of the clearing. Such gradational changes would not support the hypothesis of alternative states (Knowlton 1992). Rather, they suggest that localized, gap-dependent effects occur but these effects do not establish the required feedback loops. Examination of specific mechanisms may help to clarify such cases. For example, species composition in forest gaps often varies with the amount of sunlight reaching the ground. Since larger gaps receive more light, they may favor recruitment of more trees of light-demanding species and fewer of shade-tolerant species than smaller gaps. The shift in species composition, however, is mostly driven by light and positive feedback loops are not set up. As a result species composition rarely shifts at an abrupt threshold in gap size.

If all clearings return to the original state, then there are two possibilities. On the one hand, it is possible that the assemblages may be alternative states, but the largest clearing may be smaller than required to escape edge effects and/or the supply of recruits may be too small. On the other hand, the assemblages may not be alternative states, and site-specific differences maintain areas in either state.

If recruitment is consistent and sufficiently high then one could carry out a test by manipulating only the size of the disturbance. A three-factor design might be the most appropriate (Table 2). There should be two main factors that are completely crossed: the size of

<table>
<thead>
<tr>
<th>Analysis of variance design</th>
<th>Implication of ANOVA test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source of variation†</td>
<td>df‡</td>
</tr>
<tr>
<td>Size</td>
<td>$a - 1$</td>
</tr>
<tr>
<td>Site</td>
<td>$b - 1$</td>
</tr>
<tr>
<td>Size × Site</td>
<td>$(a - 1)(b - 1)$</td>
</tr>
<tr>
<td>Block(Site)</td>
<td>$b(c - 1)$</td>
</tr>
<tr>
<td>Size × Block(Site)</td>
<td>$(a - 1)b(c - 1)$</td>
</tr>
</tbody>
</table>

† Size = size of disturbed area; Block(Site) = block nested within site.
‡ $a =$ the number of levels in the Size treatment; $b =$ the number of Sites; $c =$ the number of Blocks within each Site.

![Diagram](image)
cleared areas as a fixed effect and different locations or sites as a random effect. The third factor should be blocks of the treatments, which are nested within locations (see Table 2). Testing for a threshold or step in the shape of the predicted curve (Fig. 2) requires a minimum of four different clearing sizes. This experimental design assures sufficient degrees of freedom for tests of the most important sources of variation (Underwood 1981, Underwood and Petraitis 1993).

The two most critical tests are the main effect of size and the interaction of size × site. The main effect of size tests the alternative hypotheses. It is also possible that assemblages are alternative states at one site but not another, and this will produce a significant interaction. This significant interaction may arise from variable recruitment, i.e., some locations received recruits while others did not.

**Situations in which only disturbance can be manipulated**

When recruitment events cannot be experimentally manipulated or when recruitment is highly variable, then we have no assurances that large clearings will switch to the alternative state. Some large clearings may switch because sufficient recruits arrive on their own, but others may not. Small clearings, on the other hand, should still show a consistent response because of edge effects. The uncertainty for large clearings, however, means the design suggested in Table 2 would not provide a good test.

The uncertainty of the outcome for large clearings but not small ones suggests that an examination of the variance may be a fruitful approach. Small clearings should revert to the original assemblage. We would expect the variation among small clearings to be moderate at first due to differences in initial colonization events but to decline over time as all small clearings return to the original state (see Fig. 3). We would also expect the variation among large clearings to be moderate at first. In contrast, we would expect the variation to increase over time as some return to the original state but others switch to the alternative state. The analysis would seem to be straightforward: compare the variance of small and large clearings early and late in the successional sequence.

There are several potential pitfalls. First, estimating variances requires a reasonably large sample, and so one would want at least a moderate number of replicate small and large clearings. A moderate number of replicates, however, may be impossible for more than a couple of different sizes. This is quite different from the design for tests of origin in which recruitment can be manipulated (Table 2), where the number of replicates (blocks within areas) is reduced in order to boost the number of treatment levels (clearings of different size). Second, many ecologists would be tempted to resample the fixed sites in the same clearings at successive intervals and examine the change in variance.

The variance estimates, however, are not independent in this case. Resampling randomly chosen sites within the same clearings might be slightly better, but rerandomizing plots may require considerably more intensive sampling effort within replicates than fixed plots, to be sure that micro- and meso-scale heterogeneity does not bias the results or swamp treatment effects in a sea of statistical "noise."

An alternative would be to use a repeated-measures design that takes advantage of the differences between small and large clearings early and late in the successional sequence. If the arrival of new recruits is highly variable across different locations, then changes across time should produce a significant treatment × location interaction. The cause of a significant interaction would be differences among locations in the number of new recruits and the resultant changes over time (see Fig. 3). All small clearings, regardless of the
number of recruits, would revert to the original state. In contrast, large clearings with large numbers of recruits may switch to the alternative state while others without recruits will revert to the original state.

Situations in which only recruitment can be manipulated

At times recruits may be abundant and easily transplanted or seeded, but natural disturbances may be too large or unusual to duplicate experimentally. For example, scrub or forest assemblages in which fire is an important agent of disturbance, or floodplain ecosystems that are prone to large, swift flows, may contain species that are easily manipulated and important in establishing the positive feedbacks leading to an alternative state, but the disturbance events cannot be mimicked or controlled.

If disturbances cannot be re-created experimentally, then natural events must be exploited whenever the opportunity arises. Naturally occurring disturbances, however, come in all sizes and shapes. They tend to span a range of size and severity, and are rarely randomly distributed in space or time. Continuous variation in size and severity means patches affected by disturbances may not be easily grouped or blocked into size or severity classes. Thus, statistical approaches, such as analysis of variance, that require distinct levels of treatment and replication within each level cannot be used. In addition, nonrandom distribution of disturbed patches in time and space raises the possibility that differences among patches are due to some unknown site-specific factor that is correlated with the occurrence, areal extent, or severity of disturbance. Experimental control is lacking. A trivial example would be differences in species assemblages on north- and south-facing slopes in a fire-maintained ecosystem. Fires may be more frequent and larger on south-facing slopes, which are drier in the Northern Hemisphere. Assemblages, however, may differ because of water availability rather than fire frequency. The correlations among slope aspect, water availability, and fire make it impossible to separate their effects on species assemblages (e.g., see Quinn and Dunham’s [1983] discussion of multiple causality).

In spite of these difficulties, it may be possible to use analysis of covariance to test the origin of alternative states. Recall that our scenario for the origin of the switch to an alternative state requires two events: a disturbance that kills members of the original assemblage and an invasion of new individuals (see Fig. 1). Within the context of analysis of covariance, the invasion or recruitment of new individuals, which can be manipulated, would be the independent variable, and the size (or severity) of the disturbance, which cannot be controlled, would be the covariate. The dependent variables would be the success of the manipulated recruits or of other members of the alternative assemblage whose arrival is facilitated by the manipulated recruits.

The analysis has four outcomes that are of interest, but only one that supports alternative states (Fig. 4), namely, significant heterogeneity of slopes. Heterogeneity of slopes arises because the switch to an alternative state occurs only when the patch opened by disturbance is large enough for recruits to gain a foothold. Small patches will revert to the original state
regardless of how many recruits have been introduced experimentally. Large patches will switch to the alternative state, but only if sufficient recruits are provided. Heterogeneity of slopes indicates an interaction between the scale of disturbance and the level of recruitment.

The other three interesting outcomes (Fig. 4) suggest other mechanisms at work. Significant main and covariate effects imply that the number of recruits and the level of disturbance are important, but their combined effect accumulates gradually. There is no threshold as required for alternative states. The same is true for the other two outcomes. A significant main effect with no effect of the covariate suggests only recruitment is important. A significant covariate effect with no main effect suggests only disturbance matters. Neither of these outcomes is consistent with our scenario, although knowledge of the mechanisms at work may suggest other explanations.

**Situations in which neither disturbance nor recruitment can be manipulated**

If neither disturbance nor recruitment can be manipulated, then we must rely on comparative approaches. This situation will occur when disturbances are too large or difficult to duplicate and the species involved too difficult to manipulate because of great longevity or size of individuals, or sporadic recruitment. Unfortunately some of the most interesting assemblages, e.g., coral reefs and forests, fall into this category, and causes of assemblage composition must be addressed by comparisons across patches in which neither recruitment nor disturbance can be controlled.

Variation in recruitment and disturbance and the underlying causes of the variation once again hamper our ability to make clear, unequivocal inferences about changes in species composition. Correlated responses with other site-dependent factors will make it difficult to assign specific causes to particular patterns (Quinn and Dunham 1983). Since disturbances are often clustered in space and time (e.g., occurrence of brush fires in summer months), responses may also show spatial and temporal autocorrelation.

To make matters worse, the alternative-states hypothesis predicts a threshold response, and so comparative analyses that assume linear responses may not be appropriate. For example, an ill-advised approach would be a correlational study. One might examine the correlation among species composition (the response), levels of recruitment and disturbance (the factors thought to be important), and other site-specific factors (“covariates” whose effect should be partialled out). Correlation analysis, however, assumes linear relationships among the variables, and we would not expect the effects of recruitment or disturbance to be linear if our scenario is correct. Yet, despite the problems, we believe this area is open to development of new approaches, and we offer one suggestion. That is to use the partial Mantel test (Smouse et al. 1986) to examine the relationship among ecological variables while controlling statistically for spatial autocorrelation (see Legendre 1993). Spatial autocorrelation of disturbance events and of other site-specific factors presents one of the biggest challenges to analyzing alternative species compositions.

**Examples and Conclusions**

We would like to emphasize that we do not see our suggestions as the final word on how to approach both theory and investigation of alternative community states. Rather, we hope this opens a wider discussion of the role of scale and of possible avenues for testing its importance. Rocky intertidal shores and North American heathlands may be ecosystems in which these sorts of approaches could be used.

Rocky intertidal shores may be an ideal system for experimental manipulations. Disturbances are often small enough that good experimental mimics can be made, and recruitment of many dominant species is heavy and predictable. For example, the rocky shores of sheltered bays throughout the North Atlantic are dominated either by stands of the macroalga *Ascophyllum nodosum* or beds of mussels and barnacles (Lewis 1964, Stephenson and Stephenson 1972). At first glance, it seems possible that these assemblages are alternative community states. Algal stands and barnacle–mussel beds in sheltered bays are persistent, have distinctly different species composition (e.g., see Coleman 1940, Wieser 1952, Lewis 1964, Stephenson and Stephenson 1972, Tsuchiya and Nishihira 1985), and occur in similar habitats.

If *Ascophyllum* stands and barnacle–mussel beds are alternative states, then specific “pulse” events should quickly initiate the “switch” and change *Ascophyllum* stands into barnacle–mussel beds or vice versa. Ice scour is the most probable pulse event since it often removes patches of *Ascophyllum*, mussels, and barnacles (e.g., for effects of ice on *Ascophyllum* and other fucoids, see Mathieson et al. [1982], McCook and Chapman [1991], Åberg [1990, 1992a, b]; for mussels and barnacles, P. S. Petraitis, personal observation). Once a large area is cleared, the switch from one assemblage to the other is likely to be driven by contingency (“historical accident”). Thus the cleared patch must be large enough so that edge effects do not extend across the entire patch. It has been suggested, but not directly tested, that commercial harvesting and other disturbances, which may denude large areas, can eliminate *Ascophyllum* and cause a switch to barnacle–mussel beds (e.g., Printz 1959, Vadas et al. 1990).

There are several positive feedback mechanisms that act in concert to return small patches to the original state or keep a particular site as either an *Ascophyllum* stand or a barnacle–mussel bed (see Fig. 5). On one hand, *Ascophyllum* can prevent establishment of mussels and barnacles by acting as a barrier to settlement...
(Menge 1976; P. S. Petraitis, unpublished data), and by providing a refuge for predators of mussels and barnacles (Menge 1976, Lubchenco and Menge 1978; P. S. Petraitis and P. G. Fairweather, unpublished data). *Ascophyllum*, even though it has very low recruitment rates itself, persists because of its extreme longevity and its ability to regenerate from holdfasts. *Ascophyllum* may live more than 100 yr (Aberg 1992a) and can regenerate from very small patches of holdfasts (P. S. Petraitis, personal observation).

Large clearings in *Ascophyllum* stands may become barnacle–mussel beds in the normal course of events because of the high recruitment of barnacles and mussels from year to year regardless of yearly variation in settlement. While yearly variation due to transport processes may be important (e.g., Gaines and Bertness [1992] showed that the recruitment of the barnacle *Semibalanus balanoides* was correlated with flushing time), average recruitment rates of mussels and barnacles in the North Atlantic usually exceed the amount needed for complete replacement, suggesting that post-recruitment processes cannot be discounted.

The switch to mussel–barnacle beds probably depends on “gregarious” settlement. Settlement and recruitment rates of mussels and barnacles are high, especially in the presence of conspecifics (Maas Gees-
In addition, mussels can move, and lateral movement quickly closes patches in mussel beds (see Paine and Levin [1981] for data on Mytilus californianus; P. S. Petraitis, personal observation for M. edulis in the Gulf of Maine). Taken together, these factors ensure that most open patches, which are well away from algal stands, quickly return to barnacle–mussel beds. Given the number of recruits and the fact that ice scour could be mimicked by manipulation, the switch from algae to mussels could be studied experimentally.

In contrast, the switch from barnacle–mussel beds to Ascophyllum stands in sheltered bays must be approached in a different way. This switch depends on two rare events—extraordinary recruitment of Ascophyllum and temporary reduction of grazers (see Vadas et al. [1990] for data on Ascophyllum recruitment, Lazo et al. [1994] for data on survival of germlings, and Sundene [1973], Vadas et al. [1977], Lubchenco [1978], and Petraitis [1987] for effects of grazers on fucoids). The density of recruiting plants must be large enough to overwhelm even reduced levels of the resident grazers (see Dayton [1971], Fairweather [1985, 1987], and Petraitis [1987] for importance of banner years of recruitment in other species, and by analogy, see Janzen [1970] and Connell [1970] for importance of temporary reduction of natural enemies in tropical rain forests). Yet, Ascophyllum, like many fucoids, has a very short dispersal distance and low recruitment, which prevents successful establishment of juveniles away from adult stands (Vadas et al. 1990, Arrontes 1993). This suggests successful establishment of new stands will be rare. A comparison of variances may be one of the few ways in which this switch could be studied.

If this scenario is plausible, then why don’t ecologists see areas with mussels becoming Ascophyllum stands more often and vice versa, particularly in experimental removals? The most plausible explanation is that the pulse event, such as ice scour, must occur on a large (10–100 m² or larger) spatial scale to initiate a “switch” because of edge effects. Unless this threshold is reached, patches will quickly close and return to their original species composition without allowing enough time for relatively rare events to play a role. Nearly all experimental work on rocky shores has been done on a much smaller scale; for example, the largest clearings done by Lubchenco and Menge (1978) were 3 m². Petraitis and Dudgeon (1999) suggest clearings may need to be between 12 and 50 m² for mussels to invade successfully.

At the other extreme are heathlands where the size and severity of disturbance preclude an entirely experimental approach. In North America, heathlands are often interspersed on the landscape with full-statured forests. Such juxtapositions appear to be examples of alternative community states. Heathlands dominated by Kalmia angustifolia (sheep-laurel) are widespread in central Newfoundland (Damman 1975, Meades 1983) and local elsewhere in eastern Canada (e.g., Strang 1971) and southward in the Appalachian Mountains (e.g., Latham et al. 1996). The extent of heathlands relative to neighboring forests has fluctuated widely in historic times (Weetman et al. 1990, Mallik 1995, Latham et al. 1996). The maintenance of cool-temperate heathlands has been attributed to a series of positive feedbacks fostered by the dominant ericaceous plants and similar shrubs, and facilitated by cool, moist climates that slow plant growth and microbial decomposition (Mallik 1995). The dominant plants acidify the soil, and the high content of phenolics, lignin, and waxes in their litter results in the accumulation of undecomposed organic matter. The recalcitrant litter inhibits nitrogen mineralization, resulting in low availability of nitrogen to most plants. Many ericaceous plants are highly tolerant of these conditions, partly because specialized ericoid mycorrhizae facilitate their hosts’ uptake of nitrogen in amino acids and other organic forms unavailable to most forest trees (Read 1983, Leake 1992). Some ericaceous plants also may inhibit forest tree establishment by allelopathy (Read 1984, Mallik 1995).

The switch from forest to heathland probably requires a pulse event that clears away not only the dominant plants of the forest—trees—but also a major fraction of the organic residues of these plants—the upper organic soil horizons. Such clearing occurs only with exceptionally severe disturbances, e.g., landslides, clearcutting, or intense windstorm followed by soil erosion, or widespread, severe burning under conditions of drought. Severe fire is probably the most likely event to initiate such a switch. Prolonged smoldering consumes soil organic matter and kills rootstocks, killed roots and denuded soil promote erosion, and exposed mineral soil provides a seedbed.

The switch from heathland to forest may occur following other, less catastrophic environmental influences. Nutrient addition from atmospheric pollution has been implicated in the decline of heaths in the Netherlands and their replacement by grasslands (Van Vuuren et al. 1992). If a tree species from the forest assemblage is somewhat tolerant of heathland soil conditions and lives in a suppressed state in the heathland assemblage, an interruption in a typical heathland regime of frequent, low-severity fires could allow it to persist and spread. Its readily decomposing litter, rapid recycling of nutrients, and the shade cast by its crown would favor the establishment of more forest trees (Zinke 1962, Kellman 1979, Parker and Muller 1982, Miles 1985, Archer et al. 1988, Belsky et al. 1989, Boettcher and Kalisz 1990, Kellman and Kading 1992, Vetaas 1992). If fire suppression continued long enough, the invaded heathland could cross a threshold into the basin of attraction of the forest assemblage. Fire suppression by humans and invasion by Acer rubrum (red maple) have been suggested as the causes of a drastic 30-yr decline.
in the area of heathlands on the Pocono Plateau in Pennsylvania (Latham et al. 1996).

Given the required conditions for the switch, i.e., extreme drought followed by high-severity burning, it seems unlikely these conditions could be experimentally manipulated. Prescribed burning could be used, but weather conditions under which prescribed burning may be conducted safely are conducive to fires of only low or moderate severity. Prescribed burning is banned by law during actual extreme droughts in most jurisdictions; in any case, it is impossible to schedule drought conditions for research purposes. Drought can be simulated (e.g., by temporarily covering research plots with polyethylene tents), but it is too difficult and expensive to be practical except in small areas and with little replication. Realistically, comparative methods may be the only ways of testing the alternative community states hypothesis in such systems.

Latham et al. (1996), for example, have assembled a set of computerized geographic information system (GIS) coverages of a 184-km² heathland–forest mosaic on the Pocono Plateau in Pennsylvania (USA). Data layers include surficial geology, topography, vegetation maps derived from 1938, 1963, and 1992 aerial photography, and 201 wildfires from the 1930s to the 1990s ranging in area from 0.5 to 2260 ha, mapped from government records and aerial photographs. The 1938 aerial photographs were the earliest ever taken at the site; 1963 was chosen because historical records show that wildfires burned essentially without control in the study area until then, but were rigorously suppressed from then on. Overlaying the vegetation maps yields a patchwork of vegetation transition trajectories (forest → forest, forest → heathland, heathland → forest, heathland → heathland) from a 25-yr period with no fire suppression and from a 29-yr period of rigorous fire suppression. Possible pathways of causality can be inferred using partial Mantel tests by comparing the spatially mapped variables considered most likely to influence vegetation change (surficial geology, slope and aspect) with each other and with spatial data on vegetation change, statistically controlling for spatial autocorrelation (Legendre 1993). For example, gradient effects would be considered as paramount and fire irrelevant if Mantel statistics showed a strong correlation between key environmental gradients and trajectories of vegetation change across the landscape mosaic, and no correlation between vegetation change and fire history.

If vegetation change correlates significantly with fire history but not with environmental gradients, then an alternative-states model would be supported. Other combinations may suggest various mixed models. For example, gradients and fire may have equal importance in determining pathways of vegetation change, or gradients may influence both fire and vegetation change but the two latter variables also could be found to have mutually reinforcing effects.

In closing, our purpose has been to highlight some conceptual issues concerning alternative community states and to suggest a variety of approaches that may be useful in testing the alternative-community-states hypothesis. Two conceptual issues seem to have impeded the development of definitive tests. First is the need to recognize that the causes that initiate the switch between alternative states need not be the same as the positive feedback mechanisms that maintain these states. Origin and maintenance are separate questions. We believe the question of origin is much less well understood in most systems. This leads us to the second issue—the role of scale. Origins of alternative states are not well studied because the initiation of the switch is scale dependent in most cases and most investigators have not incorporated scale as a factor in experimental designs. Disturbances or pulse events that start the necessary events must be large enough in areal extent and severe enough—as indicated by mortality of incumbent organisms or lag time to recovery—to provide a window of opportunity for establishment of individuals from the alternative assemblage. Tests of this idea require using patches of different sizes with the expectation that small patches would return to the original state and large patches would, in some cases, switch to the alternative state.

Clearly there are some situations in which both disturbance and recruitment can be manipulated, and good experimental tests of the role of scale in the origin of alternative states are possible. However, tests may be very difficult in some instances because both the extent of disturbance and the arrival of members of the alternative state must be manipulated. If the switch between alternative states depends on very large or unusual disturbances, then these events may be impossible to duplicate experimentally. In some cases, recruits may not be easily available or manipulated. We suggest that approaches combining observational studies with experimental manipulations could be used to address the more intractable systems.

Acknowledgments

During the preparation of this paper P. S. Petraitis received funding from the National Science Foundation (OCE 95-29564) and R. E. Latham received funding from the Andrew W. Mellon Foundation and The Nature Conservancy. We thank Jim Drake, Stephen Ellner, and an anonymous reviewer for helpful comments on an earlier version of the manuscript.

Literature Cited


of grassland to thorn woodland. Ecological Monographs **58**:111–127.


Christensen, P E., and P. C. Kiruber. 1975. Effect of pre-bolation experiments in community ecology: theory and prac-


———. 1990. Perturbations, resistance, and alternative


