INTRODUCTION

Natural or anthropogenic disturbance was traditionally viewed as an event that initiated primary or secondary succession, and succession explained the development of vegetation in the absence of disturbance. Thus, the concepts of disturbance and succession are inextricably linked in plant ecology.

Succession has been used in so many different ways and situations that it is almost useless as a precise idea. However, no matter whether succession has been considered a population (Peet and Christensen, 1980), community (Cooper, 1923a; 1923b; Clements, 1916), or ecosystem (Odum, 1969) phenomenon or process, it has contained certain common ideas. Succession is an orderly unidirectional process of community change in which communities replace each other sequentially until a stable (self-reproducing) community is reached (see definitions in Abercrombie et al. 1973; Small and Witherick, 1986; Allaby, 1994). The explanation of why and how succession is directed has changed over its more than hundred-year history, but most arguments share the notion that species are adapted to different stages in successions and in some way make the environment
unsuited for themselves and more suited for the species in the next stage. This group selection argument was first instilled into succession in the Lamarckian ideas of Warming, Cowles, and Clements.

Succession arose at the end of the 1800s and early 1900s out of a naturalist observation tradition when quantitative methods were almost non-existent, Aristotelean essentialism (Hull, 1965a, 1965b; Nordenskiöld, 1928) still had a firm grip on how nature should be understood, and meteorology, soil science, biology, and geology were very poorly developed. Further, and equally important, spatial and temporal scales of observation were limited to the scale of a naturalist’s sight.

DISTURBANCE AS THE NEMESIS OF SUCCESSION

By the beginning of 2000, most of the original classical examples of succession (e.g., Cowles, 1899; Shelford, 1911; Cooper, 1923b) given in textbooks had been restudied and found not to support the original arguments.

The first example in North America of primary succession was that on sand dunes (Cowles, 1899). The spatial sequence of plant communities as one moves away from the lake was interpreted by Cowles (1899; 1901) and Clements (1916) as representing a temporal succession of communities from dune grasses to cottonwoods, then pines and oaks to the climax beech-sugar maple forest (see Fig. 1 in Chapter 8). Olson’s (1958) study of the same dunes using techniques that allowed actual dating of the dunes produced a much more complex picture of community changes than the previously proposed simple sequence from grasses to mesophytic forest. Olson found that dunes of similar age supported a wide range of plant communities, depending on the location as well as the disturbance history of the site.

A second classic example of primary succession was that on glacial till left by the retreating glacier at Glacier Bay, Alaska (Cooper, 1923a; 1923b; 1926; 1931; 1939). Again, the spatial pattern of vegetation on areas deglaciated at varying times was interpreted as representing the temporal stages of communities through which each site would pass from herbaceous Dryas and Epilobium to shrubby willow and alder thickets, then Sitka spruce forest, and finally the spruce-hemlock climax forest. Subsequently, Crocker and Major’s (1955) study of soil properties at the different aged sites concluded that occupation of each site by the shrubs, particularly the nitrogen-fixing alders, allowed subsequent establishment of the later successional tree species through soil alteration (changes in pH and addition of
carbon and nitrogen). However, Cooper’s original study sites were reexamined by Fastie (1995), who found that the tree ring record from spruces in the oldest three sites did not indicate early suppression of growth with subsequent release once the spruces had exceeded the height of the alder or willow canopy. In other words, these oldest sites apparently had not experienced a succession from a community dominated by alders and willows to one dominated by spruce. Furthermore, the oldest sites showed a much more rapid colonization by a dense stand of trees soon after the sites were deglaciated compared to the younger sites. The differences between the different aged sites in their vegetation history (i.e., the order and rate of species establishment) as shown by Fastie’s reconstructions were explained primarily by the availability of propagules (distance to seed source) at the time the retreating ice exposed the bare substrate. Interestingly, Cooper (1923b) had also noted that “establishment of the climax does not depend upon previous dominance of alder, for in the areas of pure willow thicket the spruces were found to be invading with equal vigor,” “almost any plant of the region may be found among the vanguard,” and “even the climax trees make their first appearance with the pioneers.” Despite such observations, the lasting legacy of the early studies of primary succession at Glacier Bay has been the classic successional idea of sequential invasion and replacement of dominants driven by facilitation. As Colinvaux stated in his 1993 textbook *Ecology 2*: “The record from Glacier Bay shows that a spruce-hemlock forest cannot grow on the raw habitat left by the glacier, but that spruce trees and hemlock can claim habitats that have first been lived on by pioneer plants and alder bushes. . . . It is undeniable that primary succession on glacial till at Glacier Bay is driven by habitat modification.”

A third example of primary succession was the hydrarch succession of bogs and dune ponds. As with both of the previous examples, the spatial pattern of vegetation outward from the edge of bogs was interpreted as representing successional stages, leading to the conclusion by Clements (1916) that the open water would eventually become converted to a mesophytic forested site. However, the paleoecological reconstruction by Heinselman (1963) of the Myrtle Lake bog in Minnesota indicated that, despite deposition of organic matter and mineral sediments into the bog since deglaciation, the open water has persisted and has not been filled in and invaded by the surrounding forest because of the rising water table with the accumulation of peat.

Shelford (1911; 1913) used the spatial sequence of ponds in the Indiana Dunes of Lake Michigan to develop a model of temporal change in vegetation resulting from hydrarch succession. Jackson *et al.* (1988) tested
this classic hydrosere model by using paleoecological data spanning 3000 years and found no evidence of significant change in vegetation until the early 1800s, when rapid change occurred following European settlement. They concluded that the spatial differences in vegetation along the chronosequence reflected differential effects of disturbance rather than any temporal successional pattern.

A fourth example, this time of secondary succession in forests, by Stephens (1955) and Oliver and Stephens (1977) concerned whether forest canopy composition resulted from the continuous recruitment of new stems of more shade-tolerant species. What they found in the old, mixed-species, northern hardwoods Harvard Forest in Massachusetts was the overriding influence of small- and large-scale disturbances, both natural and anthropogenic. While small disturbances allowed release of suppressed understory trees that might otherwise never make it to the canopy, large disturbances resulted in seedling establishment of new trees. Thus, the canopy composition was determined by disturbance processes. Foster (1988) came to the same conclusions about the old-growth Pisgah Forest in New Hampshire.

Poulson and Platt’s (1996) long-term study of Warren Woods, the classic example of a climax beech-maple forest (Cain, 1935), led them to conclude that natural disturbances were chronic, occurring dependably on an ecological time scale and producing continual changes in light regimes. Because tree species respond differentially to the changing light conditions, different species are favored under different light regimes. Thus, the relative abundance of tree species in the understory at any given time cannot be used to predict the composition of the canopy at some later time. Poulson and Platt (1996) presented data to show that the relative abundance of beech and maple (as well as other species) in the canopy fluctuates in response to spatial and temporal fluctuations in frequency and sizes of treefall gaps. This is despite Cain’s (1935) tentative conclusion, based on the abundant maple reproduction he observed, that “maple seems destined to increase in importance.” In other words, the system is neither in, nor tending toward, an equilibrium climax community dominated by the most shade-tolerant species growing from the understory into the canopy.

Finally, the study by Forcier (1975) proposed a climax microsuccession with yellow birch replacing beech, sugar maple replacing yellow birch, and beech replacing sugar maple. This microsuccession model was based on a static study of trees less than 2.0 cm diameter at breast height (dbh) at Hubbard Brook Experimental Forest in New Hampshire to determine the pattern, structure, and population dynamics of the seedling layer. Major
problems with this study include the assumption that this forest is in a climax state of equilibrium, despite its history of logging (1906–1920) and hurricane disturbance in 1938 (Merrens and Peart, 1992), and the lack of aging of the canopy trees to show that they had not established concurrently following disturbance. In fact, Foster (1988) cites numerous studies that indicate repeated disturbances of the forests of central New England by windstorms, ice storms, pathogens, fire, and short-term climate changes. For example, the patterns of growth response and establishment of the canopy trees in the Pisgah Forest in New Hampshire showed the impact of 12 historically recorded storms between 1635 and 1938.

**THE CHRONOSEQUENCE BASIS OF SUCCESSION**

Despite the evidence presented in the preceding and other empirical studies that do not support the traditional ideas of succession, the tendency for ecologists to see vegetation changes as stages of succession has persisted (see Egler, 1981). As noted by Burrows (1990), “the basic concept of sequential development of vegetation on bare surfaces (first a colonizing phase, followed by immature ‘seral’ phases and culminating in a mature and stable ‘climax’ phase) is firmly embedded in the literature of vegetation ecology and in the minds of many plant ecologists.”

One reason for this persistence may lie in the chronosequence method typically used to study succession. This method involves a space-for-time substitution; that is, a chronosequence assumes that different sites, which are similar except in age since some initiating disturbance, can be considered a time sequence (Salisbury, 1952; Pickett, 1988). The key assumptions of chronosequences are that each of the sites representing different developmental stages had the same initial conditions and has traced the same sequence of changes. This assumption is rarely, if ever, carefully tested. In fact, the validity of this assumption is highly unlikely, given our increasing understanding of the temporal changes in environment and species availability over the time span represented by the chronosequences. As indicated in the previous section, studies that can “see” back in time through the pollen record or forest reconstructions (e.g., Stephens, 1955; Heinselman, 1963; Walker, 1970; Oliver and Stephens 1977; Jackson et al., 1988; Johnson et al., 1994; Fastie, 1995) have not shown the classic successional changes hypothesized from chronosequence studies. Instead, communities are found to be constantly changing, with species reassembling in different, often unfamiliar, combinations (Davis, 1981). These changes are often caused by changes in the physical environment.
COUPLING DISTURBANCE AND VEGETATION PROCESSES

Natural disturbances and often previously unappreciated human disturbances became a serious challenge to traditional succession beginning as early as the 1940s (e.g., Stearns, 1949; Raup, 1957) but really solidifying in the 1970s. Traditional succession viewed disturbances as infrequent and anomalous occurrences that initiated succession, which then proceeded in the absence of further disturbance. However, Raup (1957) commented that the ideas of succession and climax were “based largely upon the assumption of long-term stability in the physical habitat. Remove this assumption and the entire theoretical structure becomes a shambles.”

The increasing recognition of the pervasiveness of disturbances (White, 1979) led to the idea that ecological systems consisted of patches of different times since the last disturbance. This approach was reviewed in the patch dynamics book of Pickett and White (1985). The early development of patch dynamics focussed on wind-created gaps in deciduous forests (e.g., Barden, 1981; Runkle, 1981; 1982) but has since spread to almost all types of vegetation, whether appropriate and supported by evidence or not. Here again we often see the chronosequence approach being used without testing the assumptions. Remarkably, the notion of patch dynamics did not overthrow the traditional concept of succession because many ecologists simply saw patch dynamics as representing microscale successions (Forcier, 1975). Thus, the contemporary concept of succession has become a strange combination of traditional ideas of succession and patch dynamics.

Part of the reason for this strange, often inconsistent, idea of succession has been a rather poor understanding of the disturbance itself. One way to make progress in the study of dynamics of ecological systems and disturbance is to connect the disturbance processes to specific ecological processes. By process, we mean a natural phenomenon composed of a series of operations, actions, or mechanisms that explain (cause) a particular effect. This research approach (Fig. 1) has at least three parts:

**FIG. 1** Diagram illustrating the process-response model or approach to studying ecological effects of disturbance.
1. The ecological processes that will be affected by the disturbance must be precisely defined.
2. The parts of the disturbance processes that cause the ecological effect must be defined.
3. The ecological and disturbance processes must be brought together either as a coupling or a forcing.

Disturbance ecology has usually been approached in a much more informal manner than this program suggests. This has been due in part to the tools used by community ecologists (the main group of ecologists who have studied disturbances). Community ecologists in the last 50 years have taken a statistical–case study approach and have been less interested in physical environmental processes. The statistical–case study approach uses correlations or step-wise elimination curve-fitting (regressions) between variables that community ecologists hope are relevant. It is, in many ways, simply exploratory data analysis, although often used to develop predictive models. While such exploratory analysis plays a role in any research, the selection of variables is often haphazard, arbitrary, and guided by past usage or convenience (see example given by Miyanishi, 2001). No dimensional analysis is used to test either the selection of variables or the manner in which the variables are combined in the statistical model. Sometimes the variables are politically motivated (e.g., ecological integrity) and, after being chosen, the scientific models and units are then sought. Mechanisms by which cause and effect are defined are not explored formally. This is not to say that a statistical approach is not valid (e.g., population genetics) but that community ecologists have tended to use statistics to describe patterns. Rarely are statistical models seen as processes or mechanisms.

The study of earth surface interactions has been a flourishing research area that involves a wide range of disciplines, including atmospheric physics, hydrology, geomorphology, and biogeochemistry. The emerging field of biogeosciences has developed as the various disciplines have tried to integrate geophysical, geochemical, and biological (ecological) processes that are coupled to make up earth surface systems at all spatial and temporal scales (Hedin et al., 2002). These developments appear to have had little effect on community ecologists and their approach to studying disturbances and vegetation dynamics (but see Waring and Running, 1998).

Community ecologists have approached disturbances largely as a multivariate set of variables that describe a disturbance regime. The axes of the multivariate space consist of general descriptive variables, such as frequency of the disturbance and its severity, intensity, and size (Fig. 2).
This approach does not clearly define the disturbance or ecological processes of interest. The variables are often themselves vague in what they measure (e.g., frequency of what process of the disturbance) or about whether it is the disturbance or ecological effect that is being considered (e.g., severity or intensity). The coupling or forcing is almost never clearly defined. Finally, this multivariate regime model has been used more as an informal idea and rarely tested with empirical data.

We now give an example of connecting snow avalanches to tree populations (cf. Johnson, 1987) to illustrate how the approach of coupling disturbance processes and ecological processes might be used to study vegetation dynamics. Trees that grow on avalanche paths are subject to breakage and uprooting from the recurrent avalanches. Avalanche frequency changes down the slope following an extreme-value distribution whose slope depends on the tangent of the slope angle (Fig. 3). The impact pressure (k N m$^{-2}$) of avalanches also increases down slope (Fig. 4). The breakage of trees by avalanches can be determined by calculating the bending stress as the tree is deflected (Fig. 5). Bending stress (F) is determined by the applied load (P) and its lever arm (a), tree radius (r), and moment of inertia (I). Bending is determined for the deflection of the tree from its center of gravity by using a nonlinear differential equation for a tapered cantilever beam. Determining when
COUPLING DISTURBANCE AND VEGETATION PROCESSES

FIG. 3 Extreme-value distribution of avalanche frequency changes down slope.

FIG. 4 Changes in impact pressure (k Nm⁻²) of avalanches down slope.
the bending stress of the deflected beam at different avalanche loadings surpasses the modulus of rupture tells when the tree will break. Thus, by using the ages of trees at different diameters, one can determine the age-specific mortality of trees from avalanches at different avalanche frequencies. This mortality can then be compared to the age-specific mortality from other causes, such as thinning. Remember, trees must be above a certain size to break, which usually means they are big enough to start competing.

**CONCLUSION**

In organizing this book, it became apparent to us that there was abundant literature in the physical sciences (e.g., atmospheric physics, meteorology, hydrology, geomorphology) on the physical processes involved in natural disturbances, indicating a reasonably good understanding of these disturbance processes. However, the literature addressing the ecological effects of these disturbances generally showed a lack of awareness of this literature and hence made little or no attempt to couple any of the ecological processes to the disturbance processes. As a result, most ecologists have clung to the concept of succession, despite not only the lack of empirical validation (beyond the flawed chronosequence studies) but also the abundant empirical evidence disproving it as “a pervasive and fundamental phenomenon in nature” (Pickett *et al.*, 1992).
The problem is that succession is often not well defined by users of the concept, and its meaning has undergone a shift from the clear definition given by its original proponents (and as stated in numerous technical dictionaries and textbooks of biology, ecology, and geography through the 1990s) to a vague and all-inclusive concept of simply vegetation change over time. A good example of this shift can be seen in the glossary definition of succession between the 1998 and 2004 editions of Christopherson’s textbook *Elemental Geosystems*; the 1998 edition states that “changes apparently move toward a more stable and mature condition,” while the 2004 edition replaces this with “communities are in a constant state of change as each species adapts to conditions; ecosystems do not exhibit a stable point or successional climax condition as previously thought.”

We believe it is time we stop standing on our heads trying to make the concept fit our empirical observations and simply accept that the concept does not reflect a real phenomenon in nature and should therefore be abandoned. On one field trip in Colorado in the 1980s, after being informed by the trip leader that 224 different plant associations had been identified in the area, Grant Cottam wondered aloud, “In how many different ways can the number of species present be combined? I suspect it’s very close to 224.”

What we propose here is a change in viewpoint of vegetation dynamics that not only accepts the pervasive nature of disturbance in ecosystems but also incorporates the understanding that has developed of the physical disturbance processes. All disturbances have differential impacts on different populations within communities and also on different ecological processes. Therefore, to advance our understanding of the ecological effects of disturbances, we must couple the disturbance processes with ecological processes. In fact, as explained by Johnson *et al.* (2003), concepts such as metapopulations provide us with a way to incorporate disturbances into population processes.

The following chapters attempt to introduce this change in viewpoint by providing an introduction to the physical processes involved in a sampling of natural disturbances (wind, ice storms, hydrologic and fluvial disturbances, and so forth) and an attempt to couple these disturbance processes to their effects on individual plants, populations, and communities.

This book does not address all natural disturbances or all important ecological effects; for example, we do not discuss the effects of disturbances on nutrient cycling processes. However, we hope the book does provide a guide or blueprint for a different, more interdisciplinary, approach to plant disturbance ecology.
REFERENCES


REFERENCES


