that can be used by farmers to diagnose pest problems and suggest appropriate responses (Mahaman *et al.*, 2003).

IPM for the potato tuber moth

The caterpillar of the potato tuber moth (*Phthorimaea operculella*) commonly damages crops in New Zealand. An invader from a warm temperate subtropical country, it is most devastating when conditions are warm and dry (i.e. when the environment coincides closely with its optimal niche requirements – see Chapter 3). There can be as many as 6–8 generations per year and different generations mine leaves, stems and tubers. The caterpillars are protected both from natural enemies (parasitoids) and insecticides when in the tuber, so control must be applied to the leaf-mining generations. The IPM strategy for potato tuber moth (Herman, 2000) involves: (i) monitoring (female pheromone traps, set weekly from mid summer, are used to attract males, which are counted); (ii) cultural methods (the soil is cultivated to prevent soil cracking, soil ridges are molded up more than once and soil moisture is maintained); and (iii) the use of insecticides, but only when absolutely necessary (most commonly the organophosphate, methamidophos). Farmers follow the decision tree shown in Figure 15.5.

integration of IPM in sustainable farming systems

Implicit in the philosophy of IPM is the idea that pest control cannot be isolated from other aspects of food production and it is especially bound up with the means by which soil fertility is maintained and improved. These broader sustainable agricultural systems, including IFS (integrated farming systems) in the USA and LIFE (lower input farming and environment) in Europe (International Organisation for Biological Control, 1989; National Research Council, 1990), have advantages in terms of reduced environmental hazards. Even so, it is unreasonable to suppose that they will be adopted widely unless they are also sound in economic terms. In this context, Figure 15.6 shows the yields of apples from organic, conventional and integrated production systems in Washington State from 1994 to 1999 (Reganold *et al.*, 2001).

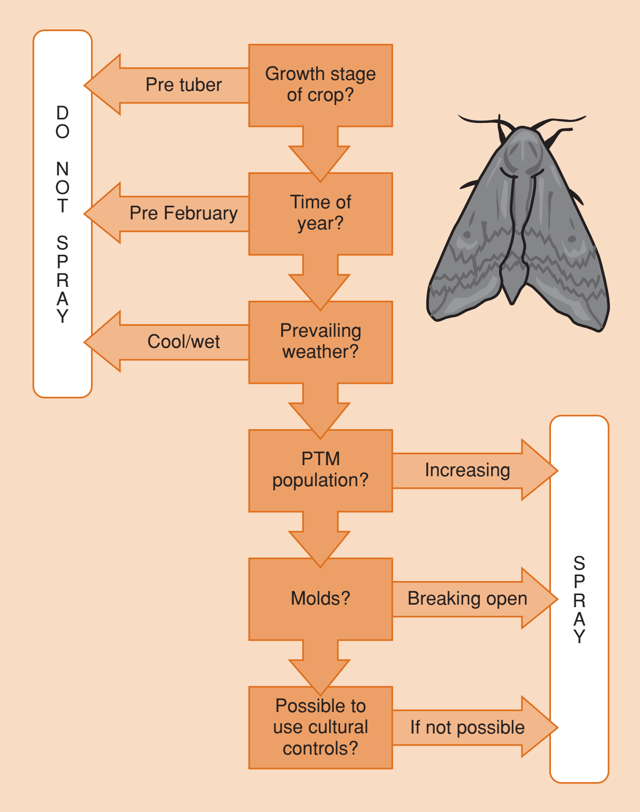


Figure 15.5 Decision flow chart for the integrated pest management of potato tuber moths (PTM) in New Zealand. Boxed phrases are questions (e.g. 'what is the growth stage of the crop?'), the words in the arrows are the farmer's answers to the questions (e.g. 'before the tuber has formed') and the recommended action is shown in the vertical box ('don't spray the crop'). Note that February is late summer in New Zealand. (After Herman, 2000.) Photograph © International Potato Center (CIP).

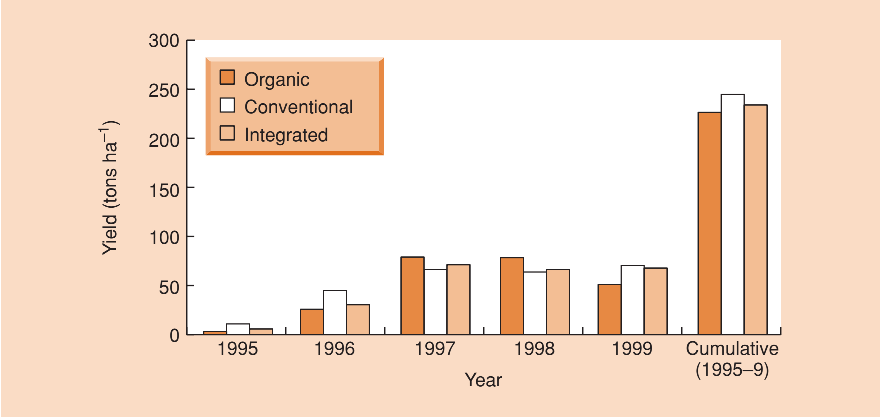


Figure 15.6 The fruit yields of three apple production systems. (From Reganold *et al.*, 2001.)

Organic management excludes such conventional inputs as synthetic pesticides and fertilizers whilst integrated farming uses reduced amounts of chemicals by integrating organic and conventional approaches. All three systems gave similar apple yields but the organic and integrated systems had higher soil quality and potentially lower environmental impacts. When compared with conventional and integrated systems, the organic system produced sweeter apples, higher profitability and greater energy efficiency. Note, however, that despite some widely held beliefs, organic farming is not totally free of adverse environmental consequences. For example, some approved pesticides are just as harmful as synthetic ones whilst the application of animal manure may lead to undesirable levels of nitrate runoff to streams just as synthetic fertilizers can (Trewavas, 2001). There is a need for research to compare the types and magnitudes of environmental consequences of the various approaches to agricultural management.

15.2.7 The importance of the early control of invaders

when a new pest invades . . .

Many pests begin life as exotic invaders. The best way to deal with the problem of potential invaders is to understand their immigration potential (see Section 7.4.2) and prevent their arrival by careful biosecurity processes at a nation's point of entry, or elsewhere on trade routes (Wittenberg & Cock, 2001). However, there are so many potential invaders that it is unrealistic to expect that they all will be prevented from arriving. Moreover, many arrivals will not establish, and many of those that do establish will do so without dramatic ecological consequences. Managers need to focus on the really problematic cases. Thus, the next step in an invader management strategy is to prioritize those that might arrive (or that have recently arrived) according to their likelihood of persisting, establishing large populations, spreading through the new area and causing significant problems. This is not an easy matter, but particular life history traits provide useful pointers (dealt with in Section 7.3.2). We will see in Chapter 22 that assessment of the potential to do harm at higher ecological levels (community/ecosystem) can also be helpful in prioritizing invaders for special attention (see Section 22.3.1).

. . . early control is best

The arrival of an exotic species with a high likelihood of becoming a significant invasive species should be a matter for urgent action, because this is the stage at which eradication is both feasible and easy to justify economically. Such campaigns sometimes rely on fundamental knowledge of population ecology. An example is the eradication of the South African sabellid polychaete worm, *Terebrasabella heterouncinata*, a parasite of abalone and other gastropods that became established near the outflow of an abalone aquaculture facility in California (Culver & Kuris, 2000). Its population biology was understood sufficiently to know it was specific to gastropods, that two species of *Tegula* were its principal hosts in the area, and that large snails were most susceptible to the parasite. Volunteers removed 1.6 million large hosts, thereby reducing the density of susceptible hosts below that needed for parasite transmission (see Chapter 12), which became extinct.

However, in the words of Simberloff (2003), rapid responses to recent invaders will often 'resemble a blunderbuss attack rather than a surgical strike'. He notes, for example, that a string of successful eradications of small populations of weeds such as pampas grass (*Cortaderia selloana*) and ragwort (*Senecio jacobaea*) on New Zealand's offshore islands (Timmins & Braithwaite, 2002) were effective because of early action using brute-force methods. Similarly, the white-spotted tussock moth (*Orygyia thyellina*), discovered in a suburban region of Auckland, New Zealand, was eradicated (at a cost of US$5 million) using *Bacillus thuringiensis* spray (Clearwater, 2001). The only population biological information to hand was that females attracted males by pheromone, knowledge that was used to trap males and determine areas that needed respraying. Eradication of a recently established species known to be invasive elsewhere usually cannot and should not wait for new population studies to be performed.

Once invaders have established and spread through a new area and are determined to be pests, they are just another species at which the pest manager's armory must be directed.

15.3 Harvest management

harvesting aims to avoid over- *and* underexploitation

Harvesting of populations by people is clearly in the realm of predator–prey interactions and harvest management relies on the theory of predator–prey dynamics (see Chapters 10 and 14). When a natural population is exploited by culling or harvesting – whether this involves the removal of whales or fish from the sea, the capture of 'bushmeat' in the African savanna or the removal of timber from a forest – it is much easier to say what we want to avoid than precisely what we might wish to achieve. On the one hand, we want to avoid overexploitation, where too many individuals are removed and the population is driven into biological jeopardy, or economic insignificance or perhaps even to extinction. But harvest managers also want to avoid underexploitation, where far fewer individuals are removed than the population can bear, and a crop of food, for example, is produced which is smaller than necessary, threatening both the health of potential consumers and the livelihood of all those employed in the harvesting operation. However, as we shall see, the best position to occupy between these two extremes is not easy to determine, since it needs to combine considerations that are not only biological (the well-being of the exploited population) and economic (the profits being made

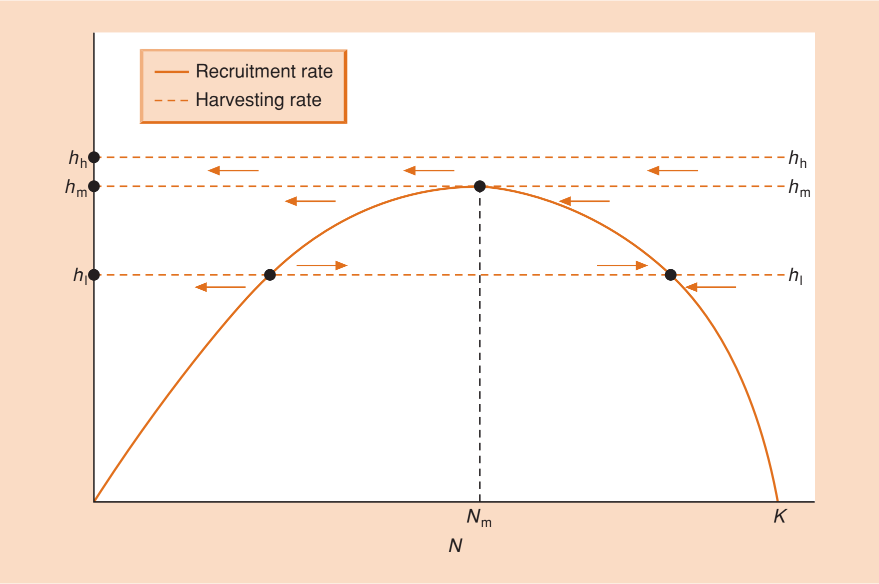


Figure 15.7 Fixed quota harvesting. The figure shows a single recruitment curve and three fixed quota harvesting curves: high quota (h_h), medium quota (h_m) and low quota (h_l). Arrows in the figure refer to changes to be expected in abundance under the influence of the harvesting rate to which the arrows are closest. ●, equilibria. At h_h the only 'equilibrium' is when the population is driven to extinction. At h_l there is a stable equilibrium at a relatively high density, and also an unstable breakpoint at a relatively low density. The MSY is obtained at h_m because it just touches the peak of the recruitment curve (at a density N_m): populations greater than N_m are reduced to N_m, but populations smaller than N_m are driven to extinction.

from the operation), but also social (local levels of employment and the maintenance of traditional lifestyles and human communities) (Hilborn & Walters, 1992; Milner-Gulland & Mace, 1998). We begin, though, with the biology.

15.3.1 Maximum sustainable yield

MSY: the peak of the net recruitment curve

The first point to grasp about harvesting theory is that high yields are obtained from populations held below, often well below, the carrying capacity. This fundamental pattern is captured by the model population in Figure 15.7. There, the natural net recruitment (or net productivity) of the population is described by an n-shaped curve (see Section 5.4.2). Recruitment rate is low when there are few individuals and low when there is intense intraspecific competition. It is zero at the carrying capacity (K). The density giving the highest net recruitment rate depends on the exact form of intraspecific competition. This density is $K/2$ in the logistic equation (see Section 5.9) but, for example, is only slightly less than K in many large mammals (see Figure 5.10d). Always, though, the rate of net recruitment is highest at an 'intermediate' density, less than K.

Figure 15.7 also illustrates three possible harvesting 'strategies', although in each case there is a fixed harvesting *rate*, i.e. a fixed number of individuals removed during a given period of time, or '*fixed quota*'. When the harvesting and recruitment lines cross, the harvesting and recruitment rates are equal and opposite; the number removed per unit time by the harvester equals the number recruited per unit time by the population. Of particular interest is the harvesting rate h_m, the line that crosses (or, in fact, just touches) the recruitment rate curve at its peak. This is the highest harvesting rate that the population can match with its own recruitment. It is known as the maximum sustainable yield (MSY), and as the name implies, it is the largest harvest that can be removed from the population on a regular and repeated (indeed indefinite) basis. It is equal to the maximum rate of recruitment, and it is obtained from the population by depressing it to the density at which the recruitment rate curve peaks.

MSY has severe shortcomings . . .

The MSY concept is central to much of the theory and practice of harvesting. This makes the recognition of the following shortcomings in the concept all the more essential.

1. By treating the population as a number of similar individuals, or as an undifferentiated biomass, it ignores all aspects of population structure such as size or age classes and their differential rates of growth, survival and reproduction. The alternatives that incorporate structure are considered below.
2. By being based on a single recruitment curve it treats the environment as unvarying.
3. In practice, it may be impossible to obtain a reliable estimate of the MSY.
4. Achieving an MSY is by no means the only, nor necessarily the best, criterion by which success in the management of a harvesting operation should be judged (see, for example, Section 15.3.9).

. . . but has been frequently used

Despite all these difficulties, the MSY concept dominated resource management for many years in fisheries, forestry and wildlife exploitation. Prior to 1980, for example, there were 39 agencies for the management of marine fisheries, every one of which was required by its establishing convention to manage on the basis of an MSY objective (Clark, 1981). In many other areas, the MSY concept is still the guiding principle. Moreover, by assuming that MSYs are both desirable and attainable, a number of the basic principles of harvesting can be explained. Therefore, we begin here by exploring what can be learnt from analyses based on the MSY, but then look more deeply at management strategies for exploited populations by examining the various shortcomings of MSY in more detail.

15.3.2 Simple MSY models of harvesting: fixed quotas

fixed-quota harvesting is extremely risky . . .

The MSY density (N_m) is an equilibrium (gains = losses), but when harvesting is based on the removal of a fixed quota, as it is in Figure 15.7, N_m is a very fragile equilibrium. If the density exceeds the MSY density, then h_m exceeds the recruitment rate and the population declines towards N_m. This, in itself, is satisfactory. But if, by chance, the density is even slightly less than N_m, then h_m will once again exceed the recruitment rate. Density will then decline even further, and if a fixed quota at the MSY level is maintained, the population will decline until it is extinct. Furthermore, if the MSY is even slightly overestimated, the harvesting rate will always exceed the recruitment rate (h_h in Figure 15.7). Extinction will then follow, whatever the initial density. In short, a fixed quota at the MSY level might be desirable and reasonable in a wholly predictable world about which we had perfect knowledge. But in the real world of fluctuating environments and imperfect data sets, these fixed quotas are open invitations to disaster.

. . . whose dangers are illustrated by the Peruvian anchovy fishery

Nevertheless, a fixed-quota strategy has frequently been used. On a specified day in the year, the fishery (or hunting season) is opened and the cumulative catch logged. Then, when the quota (estimated MSY) has been taken, the fishery is closed for the rest of the year. An example of the use of fixed quotas is provided by the Peruvian anchovy (*Engraulis ringens*) fishery (Figure 15.8). From 1960 to 1972 this was the world's largest single fishery, and it constituted a major sector of the Peruvian economy. Fisheries experts advised that the MSY was around 10 million tonnes annually, and catches were limited accordingly. But the fishing capacity of the fleet expanded, and in 1972 the catch crashed. Overfishing seems at least to have been a major cause of the collapse, although its effects were compounded with the influences of profound climatic fluctuations. A moratorium on fishing would have been an ecologically sensible step, but this was not politically feasible: 20,000 people were dependent on the anchovy industry for employment. The stock took more than 20 years to recover (Figure 15.8).

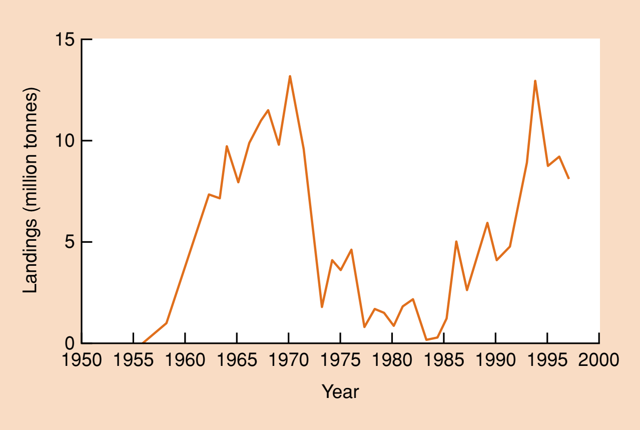


Figure 15.8 Landings of the Peruvian anchovy since 1950. (After Jennings *et al.*, 2001; data from FAO, 1995, 1998.)

15.3.3 A safer alternative: fixed harvesting effort

The risk associated with fixed quotas can be reduced if instead there is regulation of the harvesting effort. The yield from a harvest (H) can be thought of, simply, as being dependent on three things:

$$H = qEN. \tag{15.1}$$

regulating harvesting effort is less risky – but leads to a more variable catch

Yield, H, increases with the size of the harvested population, N; it increases with the level of harvesting effort, E (e.g. the number of 'trawler-days' in a fishery or the number of 'gun-days' with a hunted population); and it increases with harvesting efficiency, q. On the assumption that this efficiency remains constant, Figure 15.9a depicts an exploited population subjected to three potential harvesting strategies differing in harvesting effort. Figure 15.9b then illustrates the overall relationship to be expected, in a simple case like this, between effort and average yield: there is an apparently 'optimum' effort giving rise to the MSY, E_m, with efforts both greater and less than this giving rise to smaller yields.

Adopting E_m is a much safer strategy than fixing an MSY quota. Now, in contrast to Figure 15.7, if density drops below N_m (Figure 15.9a), recruitment exceeds the harvesting rate and the population recovers. In fact, there needs to be a considerable overestimate of E_m before the population is driven to extinction (E_0 in Figure 15.9a). However, because there is a fixed effort, the yield varies with population size. In particular, the yield will be less than the MSY whenever the population size, as a result of natural fluctuations, drops below N_m. The appropriate reaction would be to reduce effort slightly or at least hold it steady whilst the population recovers. But an understandable (albeit misguided) reaction

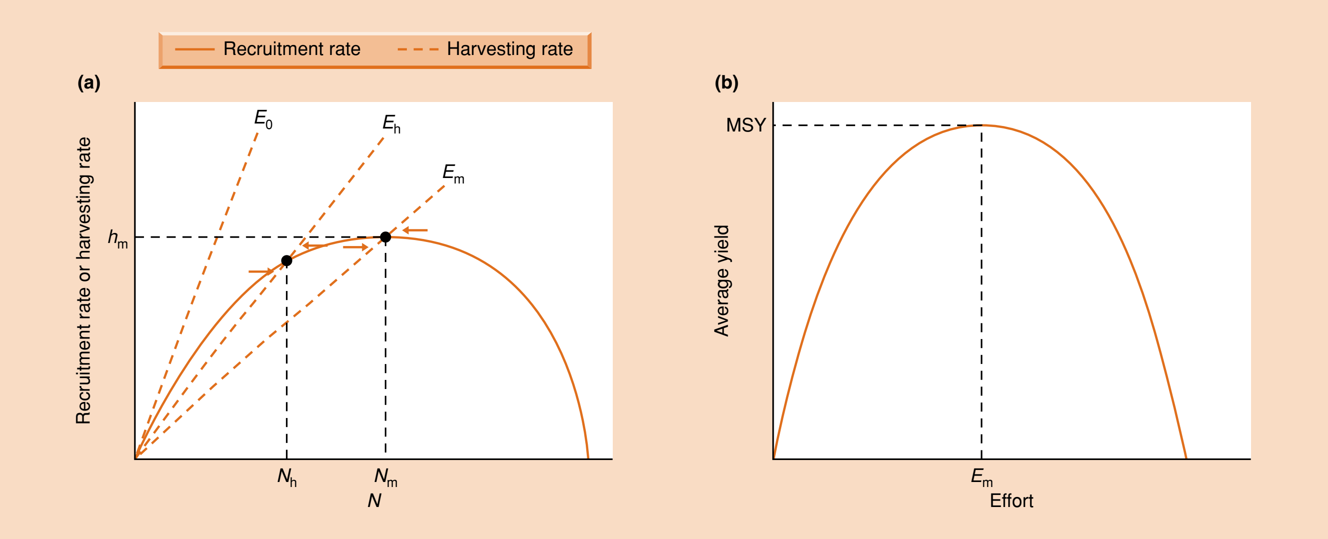


Figure 15.9 Fixed effort harvesting. (a) Curves, arrows and dots are as in Figure 15.7. The maximum sustainable yield (MSY) is obtained with an effort of E_m, leading to a stable equilibrium at a density of N_m with a yield of h_m. At a somewhat higher effort (E_h), the equilibrium density and the yield are both lower than with E_m but the equilibrium is still stable. Only at a much higher effort (E_0) is the population driven to extinction. (b) The overall relationship between the level of the fixed effort and average yield.

might be to compensate by increasing the effort. This, however, might depress population size further (E_h in Figure 15.9a); and it is therefore easy to imagine the population being driven to extinction as very gradual increases in effort chase an ever-diminishing yield.

There are many examples of harvests being managed by legislative regulation of effort, and this occurs in spite of the fact that effort usually defies precise measurement and control. For instance, issuing a number of gun licenses leaves the accuracy of the hunters uncontrolled; and regulating the size and composition of a fishing fleet leaves the weather to chance. Nevertheless, the harvesting of mule deer, pronghorn antelope and elks in Colorado was controlled by issuing a limited but varying number of hunting permits (Pojar, 1981). In the management of the important Pacific halibut stock, effort was limited by seasonal closures and sanctuary zones – although a heavy investment in fishery protection vessels was needed to make this work (Pitcher & Hart, 1982).

15.3.4 Other MSY approaches: harvesting a fixed proportion or allowing constant escapement

other MSY approaches: . . . . . . harvesting a constant proportion . . .

Two further management strategies are based on the simple idea of availability of a surplus yield. First, a constant proportion of the population can be harvested (this is equivalent to fixing a hunting mortality rate and should have the same effect as harvesting at constant effort) (Milner-Gulland & Mace, 1998). Thus, in the Northwest Territories of Canada, 3–5% of the caribou and muskox populations can be killed each year (Gunn, 1998), a strategy that involves the expense of preharvest censuses so that numbers to be harvested can be calculated.

. . . or leaving a constant 'escapement' of breeding individuals

Another strategy leaves a fixed number of breeding individuals at the end of each hunting season (constant escapement), an approach that involves the even greater expense of continuous monitoring through the hunting season. Constant escapement is a particularly safe option because it rules out the accidental removal of all the breeding individuals before breeding has occurred. Constant escapement is particularly useful for annual species because they lack the buffer provided by immature individuals in longer lived species (Milner-Gulland & Mace, 1998). The Falkland Islands government uses a constant escapement strategy for the annual *Loligo* squid. Stock sizes are assessed weekly from mid-season onwards and the fishery is closed when the ratio of stocks in the presence and absence of fishing falls to 0.3–0.4. After 10 years of this management regime the squid fishery shows good signs of sustainability (Figure 15.10).

constant escapement seems to work best for alpine marmot hunting

Stephens *et al.* (2002) used simulation models to compare the outcomes for a population of alpine marmots (*Marmota marmota*) of fixed-quota, fixed-effort and threshold harvesting. In the latter case,

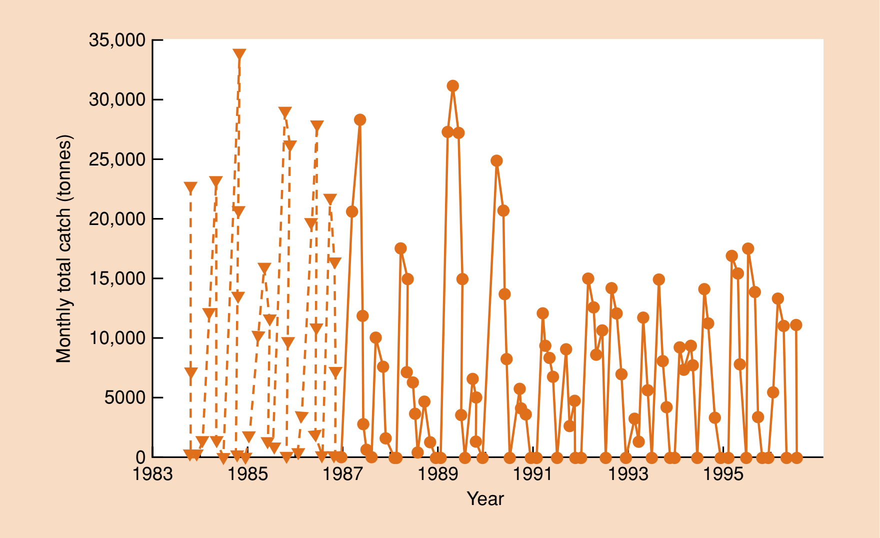


Figure 15.10 Monthly *Loligo* squid catches by licensed vessels in the Falkland Islands where a constant escapement management strategy is used. Note that there are two fishing seasons each year (February–May and August–October). The dotted lines (1984–86) represent estimated rather than actual catches. (After des Clers, 1998.)

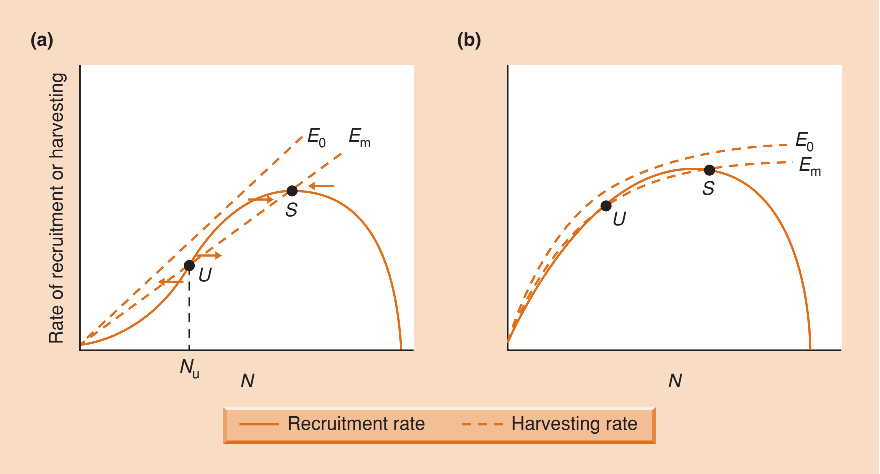


Figure 15.11 Multiple equilibria in harvesting. (a) When recruitment rate is particularly low at low densities, the harvesting effort giving the MSY (E_m) has not only a stable equilibrium (S) but also an unstable breakpoint (U) at a density below which the population declines to extinction. The population can also be driven to extinction by harvesting efforts (E_0) not much greater than E_m. (b) When harvesting efficiency declines at high densities, comments similar to those in (a) are appropriate.

harvesting only occurred during years in which the population exceeded a given threshold and exploitation continued until that threshold was reached (essentially a constant escapement approach). These social mammals are hunted in parts of Europe but the modeling was performed using extensive data available from a nonhunted population. They found that threshold harvesting provided the highest mean yields coupled with an acceptably low extinction risk. However, the introduction of error, associated with less frequent censuses (3-yearly rather than yearly), led to higher variance in yields and a much increased extinction probability (Stephens *et al.*, 2002). This emphasizes the importance of frequent censuses for constant escapement strategies to succeed.

15.3.5 Instability of harvested populations: multiple equilibria

the problem of 'depensation'

Even with regulation of effort, harvesting near the MSY level may be courting disaster. The recruitment rate may be particularly low in the smallest populations (a pattern known as *depensation*; Figure 15.11a); for instance, the recruitment of young salmon is low at low densities because of intense predation from larger fish, and the recruitment of young whales may be low at low densities simply because of the reduced chances of males and females meeting to mate. However, depensation is apparently quite rare; Myers *et al.* (1995) detected it in only three of 128 fish stock data sets with 15 or more years available for analysis. Alternatively, harvesting efficiency may increase in small populations (Figure 15.11b). For instance, many clupeids (sardines, anchovies, herring) are especially prone to capture at low densities, because they form a small number of large schools that follow stereotyped migratory paths that the trawlers can intercept. With either depensation or higher harvesting efficiency at low density, small overestimates of E_m are liable to lead to overexploitation or even eventual extinction.

harvesting operations with multiple equilibria are susceptible to dramatic irreversible crashes

Even more important, however, is the fact that these interactions may have crucial 'multiple equilibria' (see Section 10.6). Note the two points where the harvesting line crosses the recruitment curve in Figure 15.11a. The point S is a stable equilibrium but the point U is an unstable 'breakpoint'. If the population is driven slightly below the MSY density, or even to a level slightly above N_u, a breakpoint, it returns to the MSY density (Figure 15.11a). But a marginally increased depression in density, to a level slightly below N_u, perhaps resulting from only a very small increase in effort, would make the harvesting rate greater than the recruitment rate. The population would be *en route* to extinction. Moreover, once the population is on this slippery slope, much more than a marginal reduction in effort is required to reverse the process. This is the crucial, practical point about multiple equilibria: a very slight change in behavior can lead to a wholly disproportionate change in outcome as the point of attraction in the system shifts from one stable state to another. Drastic changes in stock abundance can result from only small changes in harvesting strategy or small changes in the environment.

15.3.6 Instability of harvested populations: environmental fluctuations

It is tempting to attribute all fisheries' collapses simply to overfishing and human greed. Doing so, however, would be an unhelpful oversimplification. There is no doubt that fishing pressure often exerts a great strain on the ability of natural populations to sustain levels of recruitment that counteract overall rates of loss. But the immediate cause of a collapse – in 1 year rather than any other – is often the occurrence of unusually unfavorable environmental conditions. Moreover, when this is the case, the population is more likely to recover (once conditions have returned to a more favorable state) than it would be if the crash was the result of overfishing alone.

the anchoveta and the El Niño

The Peruvian anchovy (see Figure 15.8), prior to its major collapse from 1972 to 1973, had already suffered a dip in the upward rise in catches in the mid-1960s as a result of an 'El Niño event': the incursion of warm tropical water from the north severely reducing ocean upwelling, and hence productivity, within the cold Peruvian current coming from the south (see Section 2.4.1). By 1973, however, because fishing intensity had so greatly increased, the effects of a subsequent El Niño event were much more severe. Moreover, whilst the fishery showed some signs of recovery from 1973 to 1982, in spite of largely unabated fishing pressure, a further collapse occurred in 1983 associated with yet another El Niño event. Clearly, it is unlikely that the consequences of these natural perturbations to the usual patterns of current flow would have been so severe if the anchovy had not been exploited or had been only lightly fished. It is equally clear, though, that the history of the Peruvian anchovy fishery cannot be understood properly in terms simply of fishing, as opposed to natural events.

herring and cold water

The three Norwegian and Icelandic herring fisheries also collapsed in the early 1970s and had certainly been subjected to increasing fishing intensities prior to that. Once again, however, an oceanic anomaly is implicated (Beverton, 1993). In the mid-1960s, a mass of cold, low-salinity water from the Arctic Basin formed north of Iceland. It drifted south until it became entrained in the Gulf Stream several years later, and then moved north again – although well to the east of its southward track. It eventually disappeared off Norway in 1982 (Figure 15.12a). Data for the number of 'recruits per spawner', essentially the birth rate, are illustrated in Figure 15.12b for the Norwegian springspawning and the Icelandic spring- and summer-spawning herring between 1947 and 1990, in terms of the difference each year between that year's value and the overall average. Also illustrated are the corresponding yearly temperature differentials in the Norwegian Sea, reflecting the southward and northward passage of the anomalous cold water body. There was a good correspondence between the cold water and poor recruitment in both the Icelandic and Norwegian stocks in the late 1960s and in the Norwegian stocks in 1979–81, the Icelandic stocks being then extinct (spring spawners) or too far west. It seems likely that the anomalous cold water led to unusually low recruitment, which was strongly instrumental in the crashes experienced by each of these fisheries.

This cannot, however, account for all the details in Figure 15.12b – especially the succession of poor recruitment years in the Norwegian stocks in the 1980s. For this, a more complex explanation is required, probably involving other species of fish and perhaps alternative stable states (Beverton, 1993). None the less, it remains clear that whilst the dangers of overfishing should not be denied, these must be seen within the context of marked and often unpredictable natural variations. Given the likely effects of environmental conditions on the vital rates of harvested populations, a reliance on models with constant vital rates is even more risky. Engen *et al.* (1997) argue that the best harvesting strategies for such highly variable populations involve constant escapement (see Section 15.3.4).

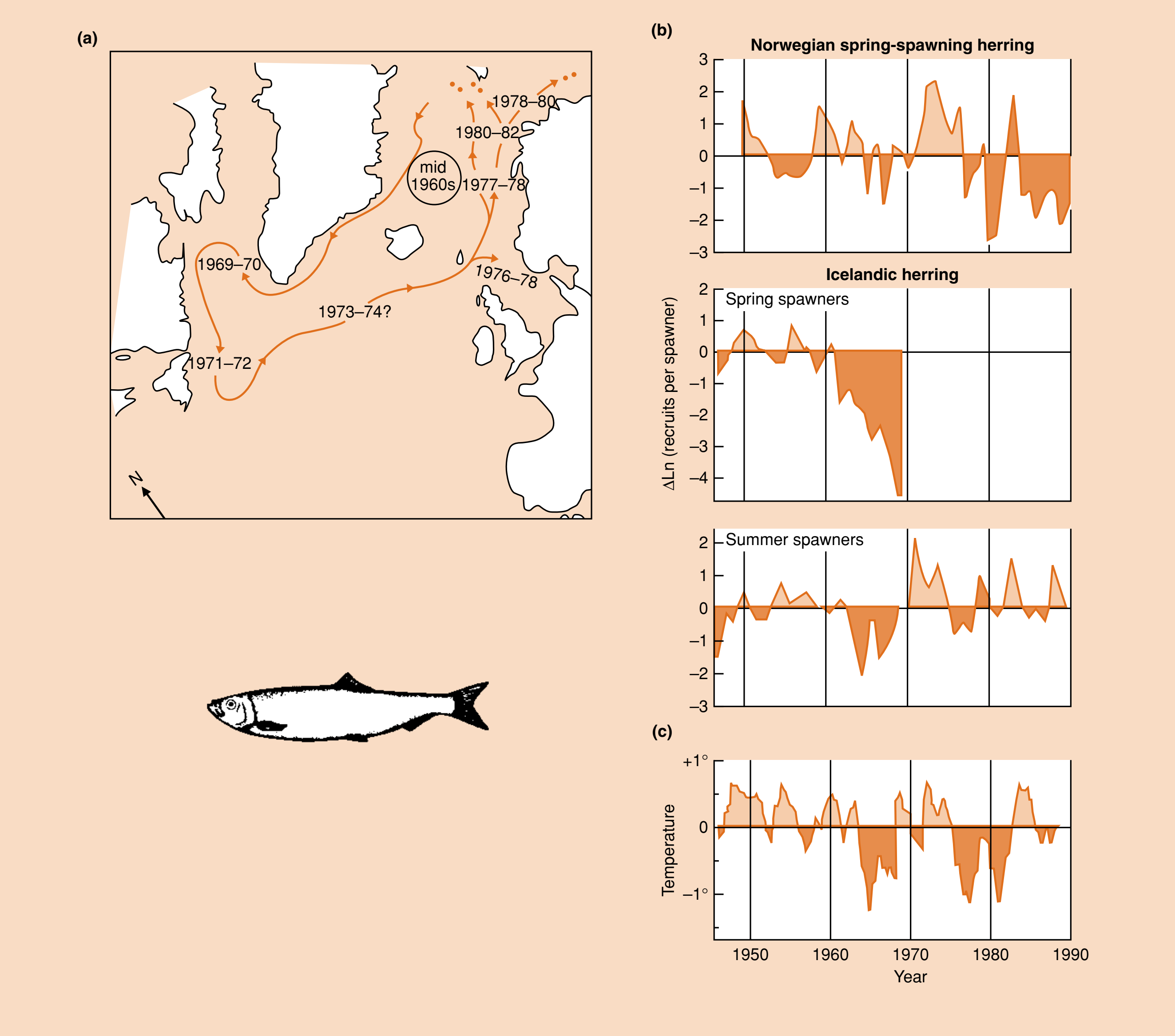


Figure 15.12 (a) The track of a large mass of cold, low-salinity water in the 1960s and 1970s, showing its presence in the Norwegian Sea both in the mid-1960s and the period 1977–82. (b) Annual differentials between overall averages and ln (recruits per spawner) for three herring stocks in the Norwegian Sea, and (c) the temperature in the Norwegian Sea. The Icelandic spring-spawning stock never recovered from its collapse in the early 1970s, preceded by low recruitment in the 1960s. (After Beverton, 1993.)

15.3.7 Recognizing structure in harvested populations: dynamic pool models

'dynamic pool' models recognize population structure

The simple models of harvesting that have been described so far are known as 'surplus yield' models. They are useful as a means of establishing some basic principles (like MSY), and they are good for investigating the possible consequences of different types of harvesting strategy. But they ignore population structure, and this is a bad fault for two reasons. The first is that 'recruitment' is, in practice, a complex process incorporating adult survival, adult fecundity, juvenile survival, juvenile growth, and so on, each of which may respond in its own way to changes in density and harvesting strategy. The second reason is that most harvesting practices are primarily interested in only a portion of the harvested population (e.g. mature trees, or fish that are large enough to be saleable). The approach that attempts to take these complications into

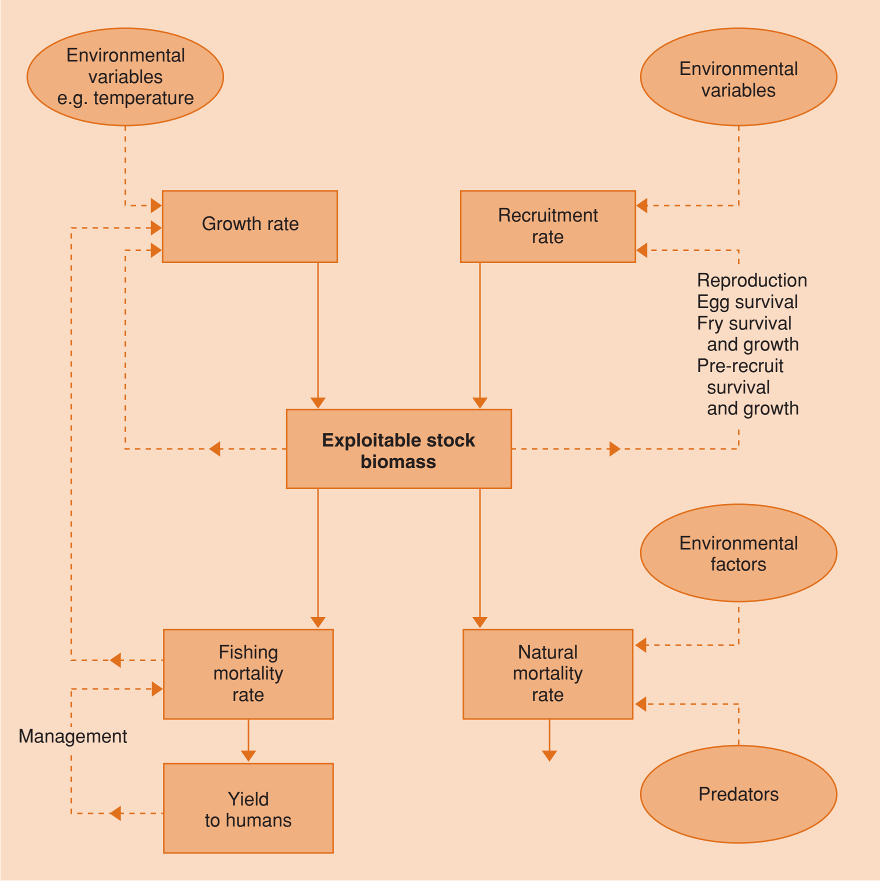


Figure 15.13 The dynamic pool approach to fishery harvesting and management, illustrated as a flow diagram. There are four main 'submodels': the growth rate of individuals and the recruitment rate into the population (which add to the exploitable biomass), and the natural mortality rate and the fishing mortality rate (which deplete the exploitable biomass). Solid lines and arrows refer to changes in biomass under the influence of these submodels. Dashed lines and arrows refer to influences either of one submodel on another, or of the level of biomass on a submodel or of environmental factors on a submodel. Each of the submodels can itself be broken down into more complex and realistic systems. Yield to humans is estimated under various regimes characterized by particular values inserted into the submodels. These values may be derived theoretically (in which case they are 'assumptions') or from field data. (After Pitcher & Hart, 1982.)

account involves the construction of what are called 'dynamic pool' models.

The general structure of a dynamic pool model is illustrated in Figure 15.13. The submodels (recruitment rate, growth rate, natural mortality rate and fishing rate of the exploited stock) combine to determine the exploitable biomass of the stock and the way this translates into a yield to the fishing community. In contrast to the surplus yield models, this biomass yield depends not only on the number of individuals caught but also on their size (past growth); whilst the quantity of exploitable (i.e. catchable) biomass depends not just on 'net recruitment' but on an explicit combination of natural mortality, harvesting mortality, individual growth and recruitment into catchable age classes.

There are many variants on the general theme (e.g. the submodels can be dealt with separately in each of the age classes and submodels can incorporate as much or as little information as is available or desirable). In all cases, though, the basic approach is the same. Available information (both theoretical and empirical) is incorporated into a form that reflects the dynamics of the structured population. This then allows the yield and the response of the population to different harvesting strategies to be estimated. This in turn should allow a recommendation to the stock-manager to be formulated. The crucial point is that in the case of the dynamic pool approach, a harvesting strategy can include not only a harvesting intensity, but also a decision as to how effort should be partitioned amongst the various age classes.

dynamic pool models can lead to valuable recommendations . . .

A classic example of a dynamic pool model in action concerned the Arcto-Norwegian cod fishery, the most northerly of the Atlantic stocks (Garrod & Jones, 1974). The age class structure of the late 1960s was used to predict the medium-term effects on yield of different fishing intensities and different mesh sizes in the trawl. Some of the results are shown in Figure 15.14. The temporary peak after 5 or so years is a result of the very large 1969 year-class working through the population. Overall, however, it is clear that the best longer term prospects were predicted for a low fishing intensity and a large mesh size. Both of these give the fish more opportunity to grow (and reproduce) before they are caught, which is important because yield is measured in biomass, not simply in numbers. Higher fishing intensities and mesh sizes of 130 mm were predicted to lead to overexploitation of the stock.

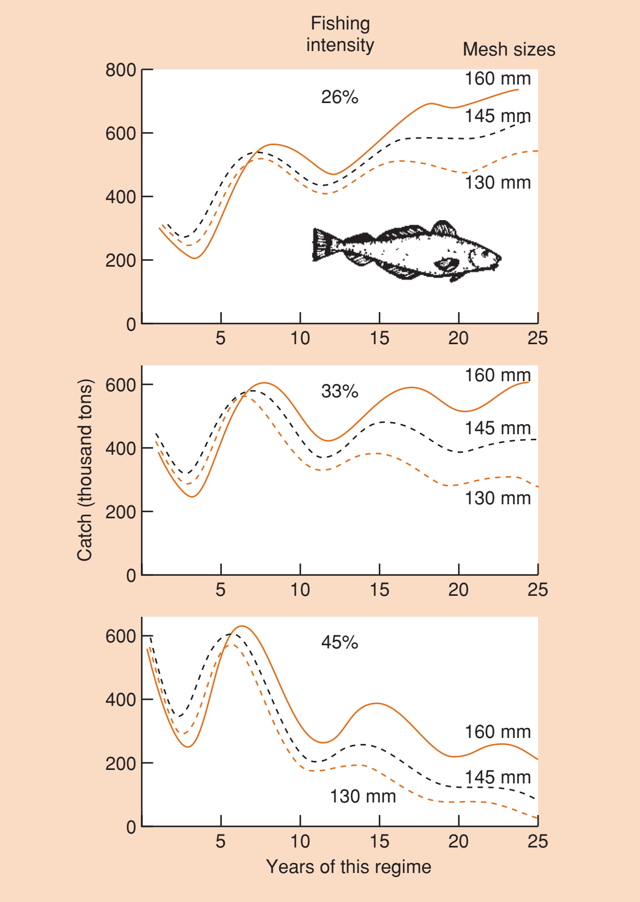


Figure 15.14 Garrod and Jones' (1974) predictions for the Arctic cod stock under three fishing intensities and with three different mesh sizes. (After Pitcher & Hart, 1982.)

. . . but these may still be ignored

Sadly, Garrod and Jones' recommendations were ignored by those with the power to determine fishing strategies. Mesh sizes were not increased until 1979, and then only from 120 to 125 mm. Fishing intensity never dropped below 45% and catches of 900,000 tonnes were taken in the late 1970s. Not surprisingly perhaps, surveys late in 1980 showed that these and other North Atlantic cod stocks were very seriously depleted as a result of overfishing. North Sea cod reach sexual maturity around the age of 4 years, but the species has been so heavily exploited that some 1 year olds are now harvested and 2 year olds are almost fully exploited, leaving only 4% of 1 year olds to survive to age 4 (Cook *et al.*, 1997).

Rattans (climbing spiny palms whose stems are used for weaving and furniture making in Southeast Asia) are threatened with overexploitation in a similar way, with harvesters cutting stems too young and reducing their ability to resprout (MacKinnon, 1998).

15.3.8 Objectives for managing harvestable resources

If we treat the Garrod and Jones example as typical, then we might conclude that the biologist proposes – but the manager disposes. This is therefore an appropriate point at which to reconsider not only the objectives of harvesting programs, but also the criteria by which successful management should be judged and the role of ecologists in management overall. As Hilborn and Walters (1992) have pointed out, there are three alternative attitudes that ecologists can take, each of which has been popular but only one of which is sensible. Indeed, these are increasingly important considerations that apply not just to fisheries management but to every entry of ecologists into the public arena.

three attitudes for ecologists towards managers in the real world . . .

The first is to claim that ecological interactions are too complex, and our understanding and our data too poor, for pronouncements of any kind to be made (for fear of being wrong). The problem with this is that if ecologists choose to remain silent because of some heightened sensitivity to the difficulties, there will always be some other, probably less qualified 'expert' ready to step in with straightforward, not to say glib, answers to probably inappropriate questions.

The second possibility is for ecologists to concentrate exclusively on ecology and arrive at a recommendation designed to satisfy purely ecological criteria. Any modification by managers or politicians of this recommendation is then ascribed to ignorance, inhumanity, political corruption or some other sin or human foible. The problem with this attitude is that it is simply unrealistic in any human activity to ignore social and economic factors.

. . . but only one of them is sensible

The third alternative, then, is for ecologists to make ecological assessments that are as accurate and realistic as possible, but to assume that these will be incorporated with a broader range of factors when management decisions are made. Moreover, these assessments should themselves take account of the fact that the ecological interactions they address include humans as one of the interacting species, and humans are subject to social and economic forces. Finally, since ecological, economic and social criteria must be set alongside one another, choosing a single, 'best' option is likely to be seen by some involved in the decision as an opinion based on the proponent's particular set of values. It follows that a single recommendation is, in practice, far less useful in this discourse than laying out a series of possible plans of action with their associated consequences.

In the present context, therefore, we develop this third alternative by first looking beyond MSY to criteria that incorporate risk, economics, social consequences, and so on (Hilborn & Waters, 1992). We then briefly examine the means by which crucial parameters and variables are estimated in natural populations,since these, by determining the quality of available information, determine the degree of confidence with which recommendations can be made.

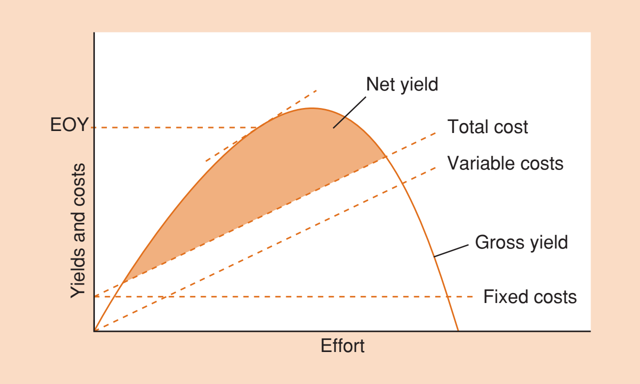


Figure 15.15 The economically optimum yield (EOY), that which maximizes 'profit', is obtained to the left of the peak of the yield-against-effort curve, where the difference between gross yield and total cost (fixed costs plus variable costs) is greatest. At this point, the gross yield and total cost lines have the same slope. (After Hilborn & Walters, 1992.)

15.3.9 Economic and social factors

the economically optimum yield – typically less than the MSY

Perhaps the most obvious shortcoming of a purely ecological approach is its failure to recognize that the exploitation of a natural resource is usually a business enterprise, in which the value of the harvest must be set against the costs of obtaining that harvest. Even if we distance ourselves from any preoccupation with 'profit', it makes no sense to struggle to obtain the last few tonnes of an MSY if the money spent in doing so could be much more effectively invested in some other means of food production. The basic idea is illustrated in Figure 15.15. We seek to maximize not total yield but net value – the difference between the gross value of the harvest and the sum of the fixed costs (interest payments on ships or factories, insurance, etc.) and the variable costs, which increase with harvesting effort (fuel, crew's expenses, etc.). This immediately suggests that the economically optimum yield (EOY) is less than the MSY, and is obtained through a smaller effort or quota. However, the difference between the EOY and the MSY is least in enterprises where most costs are fixed (the 'total cost' line is virtually flat). This is especially the case in high investment, highly technological operations such as deep-sea fisheries, which are therefore most prone to overfishing even with management aimed at economic optima.

discounting: liquidating stocks, or leaving them to grow?

A second important economic consideration concerns 'discounting'. This refers to the fact that in economic terms, each bird in the hand now (or each fish in the hold) is worth more than an equivalent bird or fish some time in the future. The reason is basically that the value of the current catch can be placed in the bank to accrue interest, so that its total value increases. In fact, a commonly used discount rate for natural resources is 10% per annum (90 fish now are as valuable as 100 fish in 1 year's time) despite the fact that the difference between the interest rates in the banks and the rate of inflation is usually only 2–5%. The economists' justification for this is a desire to incorporate 'risk'. A fish caught now has already been caught; one still in the water might or might not be caught – a bird in the hand really is worth two in the bush.

On the other hand, the caught fish is dead, whereas the fish still in the water can grow and breed (although it may also die). In a very real sense, therefore, each uncaught fish will be worth more than 'one fish' in the future. In particular, if the stock left in the water grows faster than the discount rate, as is commonly the case, then a fish put on deposit in the bank is not so sound an investment as a fish left on deposit in the sea. Nevertheless, even in cases like this, discounting provides an economic argument for taking larger harvests from a stock than would otherwise be desirable.

Moreover, in cases where the stock is less productive than the discount rate – for example, many whales and a number of long-lived fish – it seems to make sense, in purely economic terms, not only to overfish the stock, but actually to catch every fish ('liquidate the stock'). The reasons for not doing so are partly ethical – it would clearly be ecologically short sighted and a disdainful way of treating the hungry mouths to be fed in the future. But there are also practical reasons: jobs must be found for those previously employed in the fishery (or their families otherwise provided for), alternative sources of food must be found, and so on. This emphasizes, first, that a 'new economics' must be forged in which value is assigned not only to things that can be bought and sold – like fish and boats – but also to more abstract entities, like the continued existence of whales or other 'flagship species' (Hughey *et al.*, 2002). It also stresses the danger of an economic perspective that is too narrowly focused. The profitability of a fishery cannot sensibly be isolated from the implications that the management of the fishery has in a wider sphere.

social repercussions

'Social' factors enter in two rather separate ways into plans for the management of natural resources. First, practical politics might dictate, for instance, that a large fleet of small, individually inefficient boats is maintained in an area where there are no alternative means of employment. In addition, though, and of much more widespread importance, it is necessary for management plans to take full account of the way fishermen and harvesters will behave and respond to changing circumstances, rather than assuming that they will simply conform to the requirements for achieving either ecological or economic optima. Harvesting involves a predator–prey interaction: it makes no sense to base plans on the dynamics of the prey alone whilst simply ignoring those of the predator (us!).

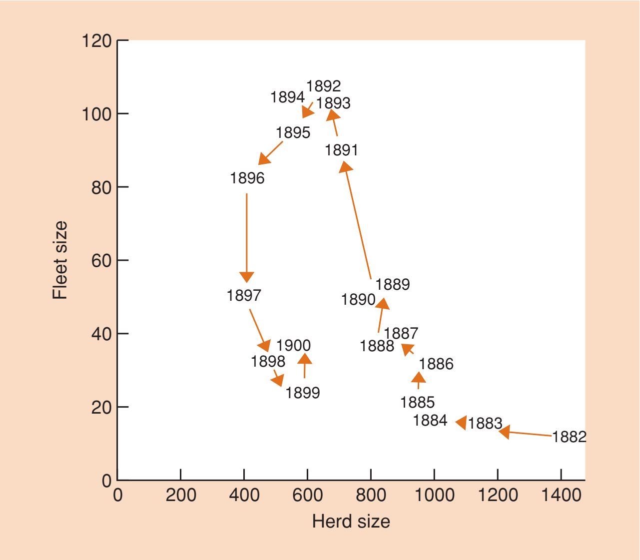


Figure 15.16 The fleet size of the North Pacific fur seal fishery (predators) responded to the size of the seal herd (prey) between 1882 and 1900 by exhibiting an anticlockwise predator–prey spiral. (After Hilborn & Walters, 1992; from data of Wilen, 1976, unpublished observations.)

harvester as predator: human behavior

The idea of the harvester as predator is reinforced in Figure 15.16, which shows a classic anticlockwise predator–prey spiral (see Chapter 10) for the North Pacific fur seal fishery in the last years of the 19th century. The figure illustrates a numerical response on the part of the predator – extra vessels enter the fleet when the stock is abundant, but leave when it is poor. But the figure also illustrates the inevitable time lag in this response. Thus, whatever a modeler or manager might propose, there is unlikely ever to be some perfect match, at an equilibrium, between stock size and effort. Moreover, whilst the sealers in the figure left the fishery as quickly as they had entered it, this is by no means a general rule. The sealers were able to switch to fishing for halibut, but such switches are often not easy to achieve, especially where there has been heavy investment in equipment or long-standing traditions are involved. As Hilborn and Walters (1992) put it, 'Principle: the hardest thing to do in fisheries management is reduce fishing pressure'.

Switching is one aspect of a harvester's predatory behavior – its functional response (see Chapter 10). Harvesters will also generally 'learn' as there is an inevitable trend towards technological improvement. Even without this, harvesters usually improve their efficiency as they learn more about their stock – notwithstanding the assumptions of simple fixed-effort models.

15.3.10 Estimates from data: putting management into practice

monitoring effort and yield: the difficulties of 'finding the top of the curve'

The role of the ecologist in the management of a natural resource is in *stock assessment*: making quantitative predictions about the response of the biological population to alternative management choices and addressing questions like whether a given fishing intensity will lead to a decline in the size of the stock, whether nets of a given mesh size will allow the recruitment rate of a stock to recover, and so on. In the past, it has often been assumed that this can be done simply by careful monitoring. For example, as effort and yield increase in an expanding fishery, both are monitored, and the relationship between the two is plotted until it seems that the top of a curve like that in Figure 15.7 has been reached or just exceeded, identifying the MSY. This approach, however, is deeply flawed, as can be seen from Figure 15.17. In 1975, the International Commission for the Conservation of Atlantic Tunas (ICCAT) used the available data (1964–73) to plot the yield–effort relationship for the yellowfin tuna (*Thunnus albacares*) in the eastern Atlantic. They felt that they

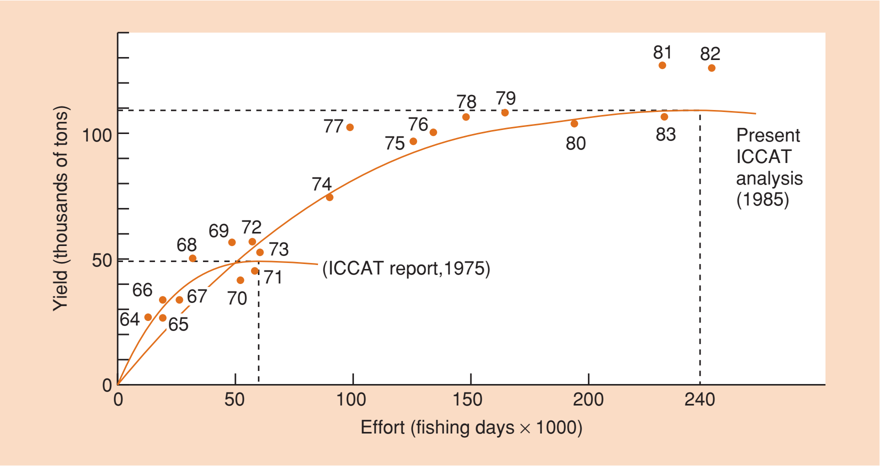


Figure 15.17 Estimated yield–effort relationships for the eastern Atlantic yellowfin tuna (*Thunnus albacares*) on the basis of the data for 1964–73 (ICCAT, 1975) and 1964–83 (ICCAT, 1985). (After Hunter *et al.*, 1986; Hilborn & Walters, 1992.)

had reached the top of the curve: a sustainable yield of around 50,000 tons (5.1×10^7 kg) and an optimum effort of about 60,000 fishing days. However, ICCAT were unable to prevent effort (and yield) rising further, and it soon became clear that the top of the curve had not been reached. A reanalysis using data up to 1983 suggested a sustainable yield of around 110,000 tons (1.1×10^8 kg) and an effort of 240,000 fishing days.

This illustrates what Hilborn and Walters (1992) describe as another principle: 'You cannot determine the potential yield from a fish stock without overexploiting it'. At least part of the reason for this is the tendency, already noted, for the variability in yield to increase as an MSY is approached. Furthermore, if we also recall the previously described difficulty in reducing fishing pressure, it is clear that in practice, managers are likely to have to wrestle with the combined challenge of estimation difficulties, ecological relationships (here, between yield and predictability) and socio-economic factors (here, concerning the regulation and reduction in effort). We have moved a long way from the simple fixed-effort models of Section 15.3.3.

The practical difficulties of parameter estimation are further illustrated in Figure 15.18, which displays the time series for total catch, fishing effort and catch per unit effort (CPUE) between 1969 and 1982 for yellowfin tuna for the whole Atlantic Ocean. As effort increased, CPUE declined – presumably, a reflection of a diminishing stock of fish. On the other hand, the catch continued to rise over this period, suggesting that perhaps the stock was not yet being overfished (i.e. the MSY had not yet been reached). These, then, are the data, and they come in probably the most commonly available form – a so-called 'one-way trip' time series. But can they suggest an MSY and can they suggest the effort required to achieve that MSY? Certainly, methods exist for performing the necessary calculations, but these methods require assumptions to be made about the underlying dynamics of the population.

estimates from catch and effort data: applying the Schaefer model

The most frequently used assumption describes the dynamics of the stock biomass, B, by:

$$\frac{dB}{dt} = rB\left(1 - \frac{B}{K}\right) - H \qquad (15.2)$$

(Schaefer, 1954), which is simply the logistic equation of Chapter 5 (intrinsic rate of increase, r, carrying capacity, K) with a harvesting rate incorporated. The latter may itself be given, following Equation 15.1 (see Section 15.3.3), by $H = qEB$, where q is harvesting efficiency and E the harvesting effort. By definition:

$$\text{CPUE} = H/E = qB. \qquad (15.3)$$

Hence:

$$B = \text{CPUE}/q \qquad (15.4)$$

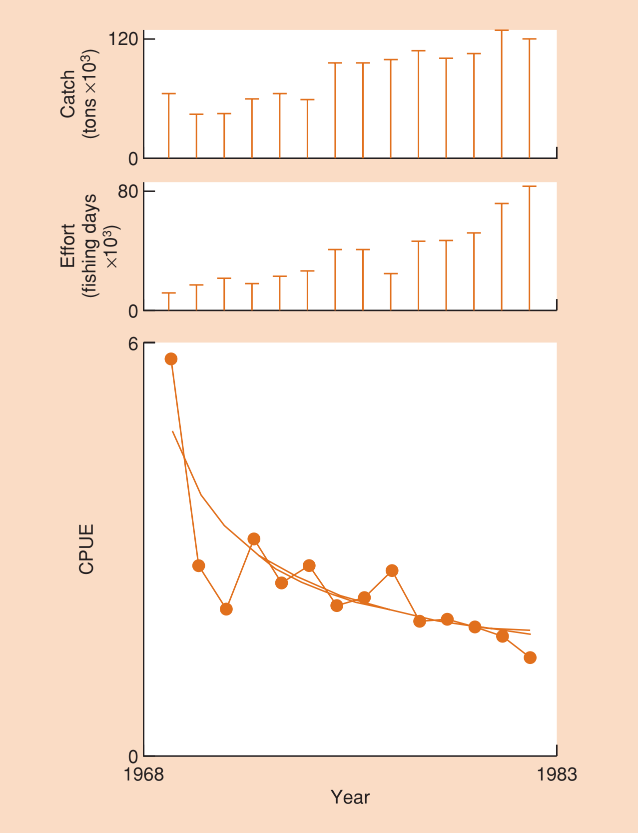


Figure 15.18 Changes in total catch, fishing effort and catch per unit effort (CPUE) between 1969 and 1982 for the yellowfin tuna (*Thussus albacares*) in the Atlantic Ocean. Also shown are three separate curves fitted to the CPUE time series by methods outlined in the text, the parameters of which are given in Table 15.4. (After Hilborn & Walters, 1992.)

and Equation 15.2 can be rewritten in terms of CPUE with either H or E as variables, and with r, q and K as parameters. For this model, the MSY is given by $rK/4$ and the effort required to achieve this by $r/2q$.

time series analysis is best – but answers are still equivocal

There are a number of methods of obtaining estimates of these parameters from field data, perhaps the best of which is the fitting of curves to time series (Hilborn & Walters, 1992). However, when the time series is a one-way trip, as we have noted it often is, there is no unique 'best' set of parameter values. Table 15.4, for instance, shows the parameters for three separate curves fitted to the data in Figure 15.18, providing equally good fits (the same sum of squares), but with widely differing parameter values. There are, in effect, a large number of equally good alternative explanations for the data in Figure 15.18, in some of which, for example, the population has a low carrying capacity but a high intrinsic rate of increase and is being harvested

Table 15.4 Parameter estimates from three fits to the catch per unit effort (CPUE) time series for yellowfin tuna shown in Figure 15.18. *r* is the intrinsic rate of increase, *K* is the carrying capacity (equilibrium abundance in the absence of harvesting) and *q* is the harvesting efficiency. Effort is measured in fishing days; *K* and maximum sustainable yield (MSY) in tons. (After Hilborn & Walters, 1992.)

Fit number	*r*	*K* (× *1000*)	*q* ($\times 10^{-7}$)	*MSY* (× *1000*)	*Effort at MSY* (× *1000*)	*Sum of squares*
1	0.18	2103	9.8	98	92	3.8
2	0.15	4000	4.5	148	167	3.8
3	0.13	8000	2.1	261	310	3.8

efficiently, whereas in others it has a high carrying capacity, a low rate of increase and is being harvested less efficiently. In the first case, the MSY had probably already been reached in 1980; in the second, catches could probably be doubled with impunity. Moreover, in each of these cases, the population is assumed to be behaving in conformity with Equation 15.2, which may itself be wide of the mark.

these uncertainties make ecologists all the more valuable

It is clear, therefore, even from this limited range of rather arbitrarily chosen examples, that there are immense limitations placed on stock assessments and management plans by inadequacies in both the available data and the means of analyzing them. This is not meant, though, to be a council of despair. Management decisions must be made, and the best possible stock assessments must form the basis – although not the sole basis – for these decisions. It is regrettable that we do not know more, but the problem would be compounded if we pretended that we did. Moreover, the ecological, economic and human behavioral analyses are important – as all analyses are – for identifying what we do *not* know, since, armed with this knowledge, we can set about obtaining whatever information is most useful. This has been formalized, in fact, in an 'adaptive management' approach, where, in an 'actively adaptive' strategy, a policy is sought which offers some balance between, on the one hand, probing for information (directed experimentation), and on the other, exercising caution about losses in short-term yield and long-term overfishing (Hilborn & Walters, 1992). Indeed, there is a strong argument that says that the inadequacies in data and theory make the need for ecologists all the more profound: who else can appreciate the uncertainties and provide appropriately enlightened interpretations?

'dataless management' where no estimates are available?

However, to be realistic, managing most marine fisheries to achieve optimum yields will be very difficult to achieve. There are generally too few researchers to do the work and, in many parts of the world, no researchers at all. In these situations, a precautionary approach to fisheries management might involve locking away a proportion of a coastal or coral community in marine protected areas (Hall, 1998). The term *dataless management* has been applied to situations where local villagers follow simple prescriptions to make sustainability more likely. For example locals on the Pacific island of Vanuatu were provided with some simple principles of management for their trochus (*Tectus niloticus*) shellfishery (stocks should be harvested every 3 years and left unfished in between) with an apparently successful outcome (Johannes, 1998).

15.4 The metapopulation perspective in management

A repeated theme in previous chapters has been the spatial patchiness upon which population interactions are often played out. Managers need to understand the implications of such heterogeneous landscape structure when making their decisions. Various approaches are available to improve our understanding of populations in complex landscapes and we consider two in the following sections. First, landscapes with different degrees of habitat loss and fragmentation can be artificially created at a scale appropriate to populations of interest and their behavior can then be assessed in carefully controlled experiments (see Section 15.4.1 – in the context of biological control of pests). Second, simple deterministic models can throw light on the factors that need to be taken into account when managing populations in a habitat patchwork (see Section 15.4.2 – in the context of creating protected areas for fisheries management). We also saw earlier (see Section 7.5.6 – in the context of a reserve patchwork for an endangered species) how stochastic simulation models can be used to compare management scenarios where subpopulations exist in a metapopulation.

15.4.1 Biological control in a fragmented landscape

natural enemy success may depend on predation efficiency in a patchy habitat

We know that spatial heterogeneity can stabilize predator–prey interactions (e.g. Chapter 10). However, the dynamics of pests and their biological control agents may become destabilized, resulting in pest outbreaks, if habitat change occurs at a scale that

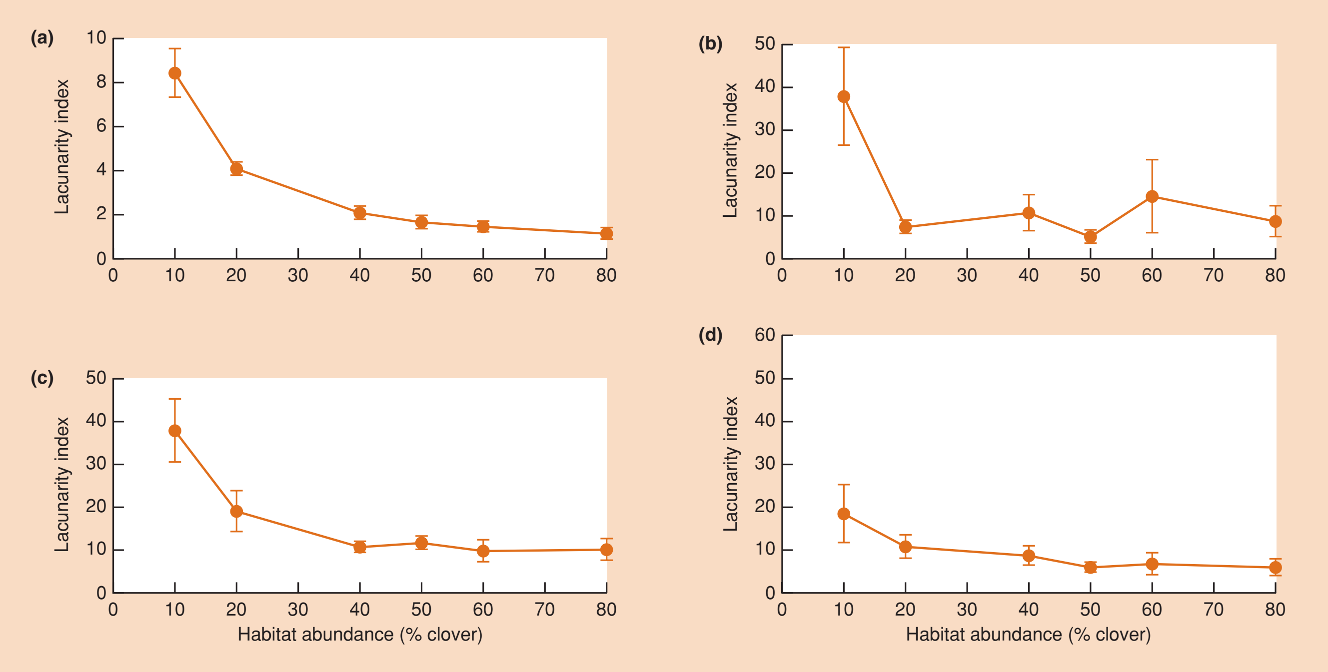


Figure 15.19 The distribution pattern (lacunarity index – a measure of aggregation) of (a) clover (i.e. habitat) and populations of (b) pest aphids, (c) an introduced ladybird beetle control agent (*Harmonia axyridis*) and (d) a native ladybird beetle (*Coleomegilla maculata*). In these experiments, clover plants were clumped together as opposed to being dispersed through the landscape. Error bars are ±1 SE. (After With *et al.*, 2002.)

interferes with the search behavior of a control agent (Kareiva, 1990).

With *et al.* (2002) created replicate landscapes (plots) of red clover (*Trifolium pratense*), each 16 × 16 m, that differed in terms of clover abundance (10, 20, 40, 50, 60 and 80% *T. pratense*). Their aim was to explore whether thresholds in landscape structure precipitate similar thresholds in the distribution of a pest aphid, *Acyrthosiphon pisum*, and to discover how landscape structure affects the search behavior of two ladybird beetle predators of aphids, one an introduced biocontrol agent, *Harmonia axyridis*, the other a native species, *Coleomegilla maculata*. Colonization by the aphids and beetles was by natural immigration to the outdoor plots.

Lacunarity is an index of aggregation derived from fractal geometry that quantifies the variability in the distribution of gap sizes (distances among clover patches in the landscape). The distribution of clover in the experimental landscapes showed a threshold at 20% habitat, indicating that gap sizes became greater and more variable below this level (Figure 15.19a). This threshold was mirrored by the aphids (Figure 15.19b) and was strongly tracked by the exotic control agent (*H. axyridis*) but not the native predator (*C. maculata*) (Figure 15.19c, d).

Although the native ladybird foraged more actively among stems within the clover cells, overall it was less mobile and moved less between clover cells in the landscape than the introduced ladybird, which showed a greater tendency to fly (Table15.5). With its greater mobility, the introduced species was more effective at tracking aphids when they occurred at low patch occupancy, a prerequisite for successful biological control (Murdoch & Briggs, 1996).

Findings such as these have implications both for the selection of effective biological control agents and for the design of agricultural systems, which may need to be managed to preserve habitat connectivity and thus enhance the efficiency of natural enemies and/or biological control agents (Barbosa, 1998).

15.4.2 Designing reserve networks for fisheries management

fishery management using no-take zones: metapopulation considerations

Over the last decade or so, coastal marine reserves or no-take zones have been promoted as a means of managing fisheries (e.g. Holland & Brazee, 1996). This is another example where an understanding of landscape structure, and metapopulation dynamics, will be necessary to devise management strategies. Probably the most fundamental questions of reserve design are the fraction of coastline that should be set aside and the appropriate size (and number) of reserves needed in relation to the

Scale and behavior measure	*Introduced Harmonia axyridis*	*Native Coleomegilla maculata*
Within clover cells		
Stems visited per minute	0.80 ± 0.05	1.20 ± 0.07
Between clover cells		
Cells visited per minute	0.22 ± 0.07	0.10 ± 0.04
Primary mode of movement	Fly	Crawl
Plot-wide movement		
Mean step length (m)	1.90 ± 0.21	1.10 ± 0.04
Displacement ratio	0.49 ± 0.05	0.19 ± 0.03

Table 15.5 Search behavior of introduced and native labybird beetles at different scales in experimental clover landscapes. Values are means ±1 SE. Each 16 × 16 m plot contains 256 cells (each 1 m^2); clover cells are those cells in which clover was present. For individual ladybirds that made at least five cell transitions, plot-wide movements were quantified in terms of mean step length and displacement ratio. Displacement ratio is net displacement (straight-line distance) divided by overall path length. (After With *et al.*, 2002.)

dispersal potential of the target species. Hastings and Botsford (2003) developed a simple deterministic model to answer these questions for a hypothetical species with characteristics that are most likely to benefit from no-take zones: one with sedentary adults and dispersing larvae. Their approach is based on the idea that altering the spacing and width of reserves changes the fraction of larvae that are retained within or exported from reserves (Figure 15.20). It is, of course, larval export that provides the basis for a sustainable yield from nonreserve areas.

The MSY problem can be stated as 'fix the level of larval retention within reserves, *F*, to preserve the species, and adjust the fraction of coastline in reserves, *c*, to maximize the number of larvae that settle outside the reserves (available as yield)'. Note that because *F* remains constant (something the modelers have chosen to assume), changing *c* means changing the width of reserves. Suppose that a value of *F* of 0.35 is deemed necessary to maintain the species. The solid line in Figure 15.20b shows how *c* and reserve width need to change to maintain an *F* of 0.35. The mathematical details of the model need not concern us but it turns out that although the largest yield is obtained when the reserves are as small as possible (the arrow in Figure 15.20b), so that larval export to fished areas is maximized, the yield is only slightly reduced as the reserve configuration moves away from this optimum. Thus, Hastings and Botsford (2003) argue that practical considerations, such as making reserves large enough to be enforced, can be allowed to play a major part in reserve design, as long as reserves are not so large (beyond the 'shoulder' of the curve in Figure 15.20b) as to significantly depress yield.

Although the model is a gross simplification, particularly in terms of the lack of any uncertainty or temporal or spatial heterogeneity, it usefully highlights some general considerations of importance and provides a starting point for more sophisticated and species-specific models to address the question of whether reserve networks will be useful for fisheries management.

In each of the sections of this chapter we have sought to build on relatively simple concepts by gradually adding more elements of realism. However, it should be remembered that even our most complex examples still lack realism in terms of the web of species interactions within which the target species are embedded. In fact, many management solutions have to be focused at a higher level of ecological organization – multispecies communities and whole ecosystems. We deal with the ecology of communities and ecosystems in Chapters 16–21 before considering ecological applications at this ecological level in Chapter 22.

Summary

Sustainability is a core concept in an ever-broadening concern for the fate of the earth and the ecological communities that occupy it. In this chapter we deal with two key aspects of ecological management – the control of pests and the management of harvests of wild populations. Each depends on an understanding of population interactions (discussed in Chapters 8–14) and each has sustainability as a primary aim.

One might imagine that the goal of pest control is total eradication but this is generally restricted to cases where a new pest has invaded a region and a rapid effort is made to completely eliminate it. Usually, the aim is to reduce the pest population to a level at which it does not pay to achieve yet more control (the economic injury level or EIL). In this way, we can see that economics and sustainability are intimately tied together. When a pest population has reached a density at which it is causing economic injury, however, it is generally too late to start controlling it. More important, then, is the economic threshold (ET): the density of the pest at which action should be taken to prevent it reaching the EIL.

We describe the tool kit of chemical pesticides and herbicides. These are a key part of the armory of pest managers but they

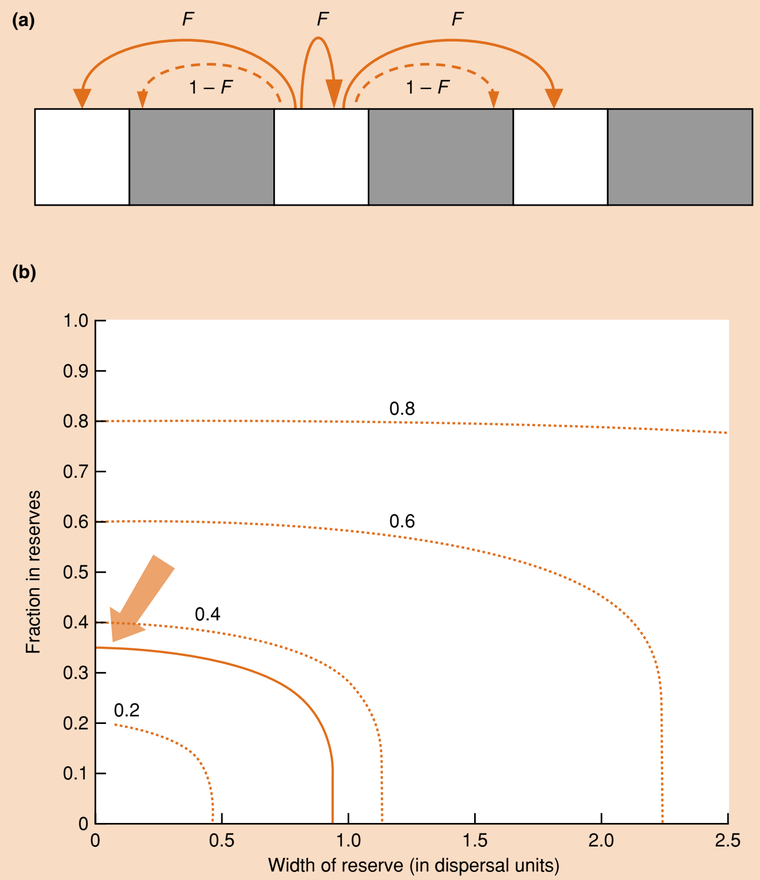


Figure 15.20 (a) Schematic representation of a network of marine reserves (white) and fished areas (gray). The fraction of coastline in the reserves is *c*, the fraction of larvae produced in the reserves is *F*, and the fraction of larvae produced in reserves that are exported is 1 – *F*. (b) The combination of values of the fraction of coastline in reserves, *c*, and mean width of reserve (in units of mean dispersal distance) that yield a value of 0.35 for the fraction of larvae that are retained within reserves, *F*, along with similar combinations for other values of *F*. The arrow indicates the configuration that produces the maximum fishing yield outside the reserves. (After Hastings & Botsford, 2003.)

have to be used with care because of the possibility of 'target pest resurgence' (when treatment affects the pest's natural enemies more than the pest itself) and 'secondary pest outbreaks' (when natural enemies of 'potential' pests are strongly affected, allowing potential pests to become actual pests). Pests are also adept at evolving resistance to pesticides.

An alternative to chemical pesticides is biological manipulation of the natural enemies of the pests. Biological control may involve: (i) 'introduction', with the expectation of long-term persistence, of a natural enemy from another geographic area (often the one from where the pest originated); (ii) the manipulation of natural predators already present ('conservation biological control'); (iii) the periodic release of an agent that is unable to persist through the year but provides control for one or a few pest generations ('inoculation'); or (iv) the release of large numbers of enemies, which will not persist, to kill only those pests present at the time ('inundation', sometimes called, by analogy, biological pesticides). Biological control is by no means always environmentally friendly. Examples are coming to light where even carefully chosen and apparently successful introductions of biological control agents have impacted on nontarget species, both by affecting nontarget species related to the pest and by affecting other species that interact in food webs with the nontarget species.

Integrated pest management (IPM) is a practical philosophy of pest management that is ecologically based but uses all methods of control, including chemicals, when appropriate. It relies heavily on natural mortality factors such as weather and natural enemies.

Whenever, a natural population is exploited by harvesting there is a risk of overexploitation. But harvesters also want to avoid underexploitation, where potential consumers are deprived and those who harvest are underemployed. Thus, as with many areas of applied ecology, there are important economic and sociopolitical perspectives to consider.

The concept of the maximum sustainable yield (MSY) has been a guiding principle in harvest management. We describe the different approaches to obtain an MSY – taking a fixed quota, regulating harvest effort, harvesting a constant proportion or allowing constant escapement – and we point out the shortcomings of each. More reliable approaches to sustainable harvesting are also discussed, including dynamic pool models (which recognize that all individuals in the harvested population are not equivalent and incorporate population structure into the population models) and approaches that explicitly incorporate economic factors (dealing with economically optimum yield, OEY, rather than simply MSY). We also note that no data are available for many of the world's fisheries, especially in developing areas of the world; in these cases, simple 'dataless' management principles may be the best that ecologists can propose.

Finally, many populations, including those of pests and harvested populations, exist in a heterogeneous environment, sometimes as metapopulations. Managers need to be aware of this possibility, for instance when determining which biological control agent to use in an agricultural landscape or when designing a network of 'no-take' zones as part of a fisheries management strategy.

Part 3

Communities and Ecosystems

Introduction

In nature, areas of land and volumes of water contain assemblages of different species, in different proportions and doing different things. These communities of organisms have properties that are the sum of the properties of the individual denizens plus their interactions. The interactions are what make the community more than the sum of its parts. Just as it is a reasonable aim for a physiologist to study the behavior of different sorts of cells and tissues and then attempt to use a knowledge of their interactions to explain the behavior of a whole organism, so ecologists may use their knowledge of interactions between organisms in an attempt to explain the behavior and structure of a whole community. Community ecology, then, is the study of patterns in the structure and behavior of multispecies assemblages. Ecosystem ecology, on the other hand, is concerned with the structure and behavior of the same systems but with a focus on the flux of energy and matter.

We consider first the nature of the community. Community ecologists are interested in how groupings of species are distributed, and the ways these groupings can be influenced by both abiotic and biotic environmental factors. In Chapter 16 we start by explaining how the structure of communities can be measured and described, before focusing on patterns in community structure in space, in time and finally in a more complex, but more realistic spatiotemporal setting.

Communities, like all biological entities, require matter for their construction and energy for their activities. We examine the ways in which arrays of feeders and their food bind the inhabitants of a community into a web of interacting elements, through which energy (Chapter 17) and matter (Chapter 18) are moved. This ecosystem approach involves primary producers, decomposers and detritivores, a pool of dead organic matter, herbivores, carnivores and parasites *plus* the physicochemical environment that provides living conditions and acts both as a source and a sink for energy and matter. In Chapter 17, we deal with large-scale patterns in primary productivity before turning to the factors that limit productivity, and its fate, in terrestrial and aquatic settings. In Chapter 18, we consider the ways in which the biota accumulates, transforms and moves matter between the various components of the ecosystem.

In Chapter 19 we return to some key population interactions dealt with earlier in the book, and consider the ways that competition, predation and parasitism can shape communities. Then in Chapter 20 we recognize that the influence of a particular species often ramifies beyond a particular competitor, prey or host population, through the whole food web. The study of food webs lies at the interface of community and ecosystem ecology and we focus both on the population dynamics of interacting species in the community and on the consequences for ecosystem processes such as productivity and nutrient flux.

In Chapter 21 we attempt an overall synthesis of the factors, both abiotic and biotic, that determine species richness. Why the number of species varies from place to place, and from time to time, are interesting questions in their own right as well as being questions of practical importance. We will see that a full understanding of patterns in species richness has to draw on an understanding of all the ecological topics dealt with in earlier chapters of the book.

Finally, in the last of our trilogy of chapters dealing with the application of ecological theory, we consider in Chapter 22 the application of theory related to succession, food web ecology, ecosystem functioning and biodiversity. We conclude by recognizing that the application of ecological theory never proceeds in isolation – the sustainable use of natural resources requires that we also incorporate economic and sociopolitical perspectives.

To pursue an analogy we introduced earlier, the study of ecology at the community/ecosystem level is a little like making a study of watches and clocks. A collection can be made and the contents of each timepiece classified. We can recognize characteristics that they have in common in the way they are constructed and patterns in the way they behave. But to understand how they work, they must be taken to pieces, studied and put back together again. We will have understood the nature of natural communities when we *know* how to recreate those that we have, often inadvertently, taken to pieces.

Chapter 16

The Nature of the Community: Patterns in Space and Time

16.1 Introduction

Physiological and behavioral ecologists are concerned primarily with individual *organisms*. Coexisting individuals of a single species possess characteristics – such as density, sex ratio, age-class structure, rates of natality and immigration, mortality and emigration – that are unique to *populations*. We explain the behavior of a population in terms of the behavior of the individuals that comprise it. In their turn, activities at the population level have consequences for the next level up – that of the *community*. The community is an assemblage of species populations that occur together in space and time. Community ecology seeks to understand the manner in which groupings of species are distributed in nature, and the ways these groupings can be influenced by their abiotic environment (Part 1 of this textbook) and by interactions among species populations (Part 2). One challenge for community ecologists is to discern and explain patterns arising from this multitude of influences.

the search for rules of community assembly

In very general terms, the species that assemble to make up a community are determined by: (i) dispersal constraints; (ii) environmental constraints; and (iii) internal dynamics (Figure 16.1) (Belyea & Lancaster, 1999). Ecologists search for rules of community assembly, and we discuss these in this chapter and a number of others (particularly Chapters 19–21).

communities have collective properties . . .

. . . and emergent properties not possessed by the individual populations that comprise them

A community is composed of individuals and populations, and we can identify and study straightforward *collective* properties, such as species diversity and community biomass. However, we have already seen that organisms of the same and different species interact with each other in processes of mutualism, parasitism, predation and competition. The nature of the community is obviously more than just the sum of its constituent species. There are *emergent* properties that appear when the community is the focus of attention, as there are in other cases where we are concerned with the behavior of complex mixtures. A cake has emergent properties of texture and flavor that are not apparent simply from a survey of the ingredients. In the case of ecological communities, the limits to similarity of competing species (see Chapter 19) and the stability of the food web in the face of disturbance (see Chapter 20) are examples of emergent properties.

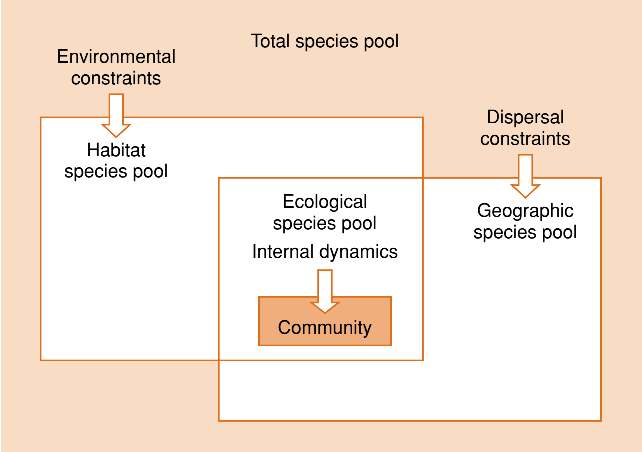


Figure 16.1 The relationships among five types of species pools: the total pool of species in a region, the geographic pool (species able to arrive at a site), the habitat pool (species able to persist under the abiotic conditions of the site), the ecological pool (the overlapping set of species that can both arrive and persist) and the community (the pool that remains in the face of biotic interactions). (Adapted from Belyea & Lancaster, 1999; Booth & Swanton, 2002.)

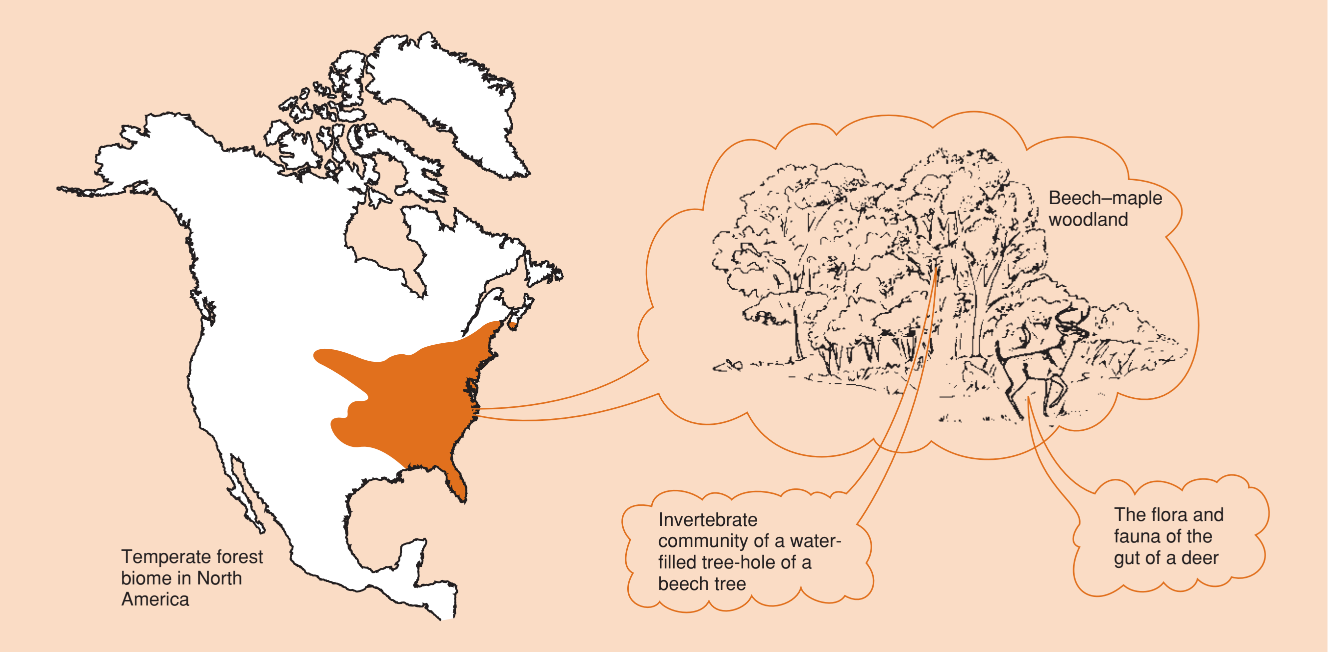


Figure 16.2 We can identify a hierarchy of habitats, nesting one into the other: a temperate forest biome in North America; a beech–maple woodland in New Jersey; a water-filled tree hole; or a mammalian gut. The ecologist may choose to study the community that exists on any of these scales.

Science at the community level poses daunting problems because the database may be enormous and complex. A first step is usually to search for patterns in the community's collective and emergent properties. Patterns are repeated consistencies, such as the repeated grouping of similar growth forms in different places, or repeated trends in species richness along different environmental gradients. Recognition of patterns leads, in turn, to the forming of hypotheses about the causes of these patterns. The hypotheses may then be tested by making further observations or by doing experiments.

A community can be defined at any scale within a hierarchy of habitats. At one extreme, broad patterns in the distribution of community types can be recognized on a global scale. The temperate forest biome is one example; its range in North America is shown in Figure 16.2. At this scale, ecologists usually recognize climate as the overwhelming factor that determines the limits of vegetation types. At a finer scale, the temperate forest biome in parts of New Jersey is represented by communities of two species of tree in particular, beech and maple, together with a very large number of other, less conspicuous species of plants, animals and microorganisms. Study of the community may be focused at this scale. On an even finer habitat scale, the characteristic invertebrate community that inhabits water-filled holes in beech trees may be studied, or the flora and fauna in the gut of a deer in the forest. Amongst these various scales of community study, no one is more legitimate than another. The scale appropriate for investigation depends on the sorts of questions that are being asked.

communities can be recognized at a variety of levels – all equally legitimate

Community ecologists sometimes consider all of the organisms existing together in one area, although it is rarely possible to do this without a large team of taxonomists. Others restrict their attention within the community to a single taxonomic group (e.g. birds, insects or trees), or a group with a particular activity (e.g. herbivores or detritivores).

The rest of this chapter is in six sections. We start by explaining how the structure of communities can be measured and described (Section 16.2). Then we focus on patterns in community structure: in space (Section 16.3), in time (Sections 16.4–16.6) and finally in a combined spatiotemporal setting (Section 16.7).

16.2 Description of community composition

species richness: the number of species present in a community

One way to characterize a community is simply to count or list the species that are present. This sounds a straightforward procedure that enables us to describe and compare communities by their species 'richness' (i.e. the number of species present). In practice, though, it is often surprisingly difficult, partly because

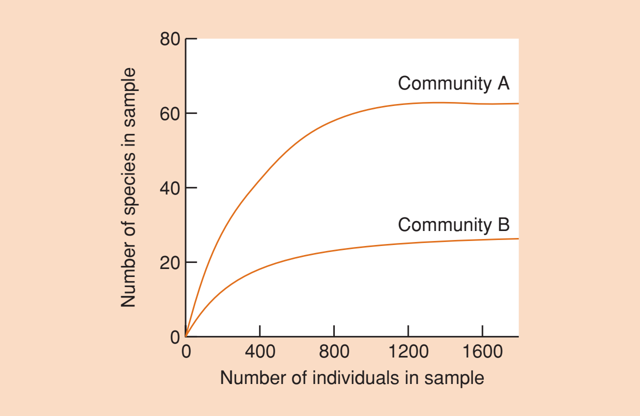


Figure 16.3 The relationship between species richness and the number of individual organisms from two contrasting hypothetical communities. Community A has a total species richness considerably in excess of community B.

of taxonomic problems, but also because only a subsample of the organisms in an area can usually be counted. The number of species recorded then depends on the number of samples that have been taken, or on the volume of the habitat that has been explored. The most common species are likely to be represented in the first few samples, and as more samples are taken, rarer species will be added to the list. At what point does one cease to take further samples? Ideally, the investigator should continue to sample until the number of species reaches a plateau (Figure 16.3). At the very least, the species richnesses of different communities should be compared on the basis of the same sample sizes (in terms of area of habitat explored, time devoted to sampling or, best of all, number of individuals or modules included in the samples). The analysis of species richness in contrasting situations figures prominently in Chapter 21.

16.2.1 Diversity indices

diversity incorporates richness, commonness and rarity

An important aspect of community structure is completely ignored, though, when the composition of the community is described simply in terms of the number of species present. It misses the information that some species are rare and others common. Consider a community of 10 species with equal numbers in each, and a second community, again consisting of 10 species, but with more than 50% of the individuals belonging to the most common species and less than 5% in each of the other nine. Each community has the same species richness, but the first, with a more 'equitable' distribution of abundances, is clearly more *diverse* than the second. Richness and equitablity combine to determine community diversity.

Knowing the numbers of individuals present in each species may not provide a full answer either. If the community is closely defined (e.g. the warbler community of a woodland), counts of the number of individuals in each species may suffice for many purposes. However, if we are interested in all the animals in the woodland, then their enormous disparity in size means that simple counts would be very misleading. There are also problems if we try to count plants (and other modular organisms). Do we count the number of shoots, leaves, stems, ramets or genets? One way round this problem is to describe the community in terms of the biomass per species per unit area.

Simpson's diversity index

The simplest measure of the character of a community that takes into account both the abundance (or biomass) patterns and the species richness, is Simpson's diversity index. This is calculated by determining, for each species, the proportion of individuals or biomass that it contributes to the total in the sample, i.e. the proportion is P_i for the ith species:

$$\text{Simpson's index, } D = \frac{1}{\sum_{i=1}^{S} P_i^2}, \qquad (16.1)$$

where S is the total number of species in the community (i.e. the richness). As required, for a given richness, D increases with equitability, and for a given equitability, D increases with richness.

'equitability' or 'evenness'

Equitability can itself be quantified (between 0 and 1) by expressing Simpson's index, D, as a proportion of the maximum possible value D would assume if individuals were completely evenly distributed amongst the species. In fact, $D_{\max} = S$. Thus:

$$\text{equitability, } E = \frac{D}{D_{\max}} = \frac{1}{\sum_{i=1}^{S} P_i^2} \times \frac{1}{S}. \qquad (16.2)$$

Shannon's diversity index

Another index that is frequently used and has essentially similar properties is the Shannon diversity index, H. This again depends on an array of P_i values. Thus:

$$\text{diversity, } H = -\sum_{i=1}^{S} P_i \ln P_i \qquad (16.3)$$

and:

$$\text{equitability, } J = \frac{H}{H_{\max}} = \frac{-\sum_{i=1}^{S} P_i \ln P_i}{\ln S}. \qquad (16.4)$$

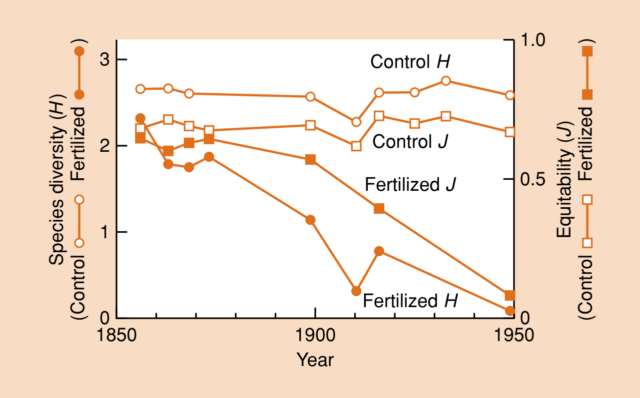


Figure 16.4 Species diversity (*H*) and equitability (*J*) of a control plot and a fertilized plot in the Rothamsteard 'Parkgrass' experiment. (After Tilman, 1982.)

An example of an analysis of diversity is provided by a uniquely long-term study that has been running since 1856 in an area of grassland at Rothamsted in England. Experimental plots have received a fertilizer treatment once every year, whilst control plots have not. Figure 16.4 shows how species diversity (*H*) and equitability (*J*) of the grass species changed between 1856 and 1949. Whilst the unfertilized area has remained essentially unchanged, the fertilized area has shown a progressive decline in diversity and equitability. One possible explanation may be that high nutrient availability leads to high rates of population growth and a greater chance of the most productive species coming to dominate and, perhaps, competitively exclude others.

16.2.2 Rank–abundance diagrams

Of course, attempts to describe a complex community structure by one single attribute, such as richness, diversity or equitability, can be criticized because so much valuable information is lost. A more complete picture of the distribution of species abundances in a community makes use of the full array of P_i values by plotting P_i against rank. Thus, the P_i for the most abundant species is plotted first, then the next most common, and so on until the array is completed by the rarest species of all. A rank–abundance diagram can be drawn for the number of individuals, or for the area of ground covered by different sessile species, or for the biomass contributed to a community by the various species.

rank–abundance models may be based on statistical or biological arguments

A range of the many equations that have been fitted to rank–abundance diagrams is shown in Figure 16.5. Two of these are statistical in origin (the log series and log-normal) with no foundation in any assumptions about how the species may interact with one another. The others take some account of the relationships between the conditions, resources and species-abundance patterns (niche-orientated models) and are more likely to help us understand the mechanisms underlying community organization (Tokeshi, 1993). We illustrate the diversity of approaches by describing the basis of four of Tokeshi's niche-orientated models (see Tokeshi, 1993, for a complete treatment). The *dominance–preemption model*, which produces the least equitable species distribution, has successive species preempting a dominant portion (50% or more) of the remaining niche space; the first, most dominant species takes more than 50% of the total niche space, the next more than 50% of what remains, and so on. A somewhat more equitable distribution is represented by the *random fraction model*, in which successive species invade and take over an arbitrary portion of the niche space of any species previously present. In this case, irrespective of their dominance status, all species are subjected to niche division with equal probability. The *MacArthur fraction model*, on the other hand, assumes that species with larger niches are more likely to be invaded by new species; this results in a more equitable distribution than the random fraction model. Finally, the *dominance–decay model* is the inverse of the dominance–preemption model, in that the largest niche in an existing assemblage is always subject to a subsequent (random) division. Thus, in this model the next invading species is supposed to colonize the niche space of the species currently most abundant, yielding the most equitable species abundances of all the models.

Rank–abundance diagrams, like indices of richness, diversity and equitability, should be viewed as abstractions of the highly complex structure of communities that may be useful when making comparisons. In principle, the idea is that finding the best fitting model should give us clues as to underlying processes, and perhaps as to how these vary from sample to sample. Progress so far, however, has been limited, both because of problems of interpretation and the practical difficulty of testing for the best fit between model and data (Tokeshi, 1993). However, some studies have successfully focused attention on a change in dominance/evenness relationships in relation to environmental change. Figure 16.5c shows how, assuming a geometric series can be appropriately applied, dominance steadily increased, whilst species richness decreased, during the Rothamsted long-term grassland experiment described above. Figure 16.5d shows how invertebrate species richness and equitability were both greater on an architecturally complex stream plant *Ranunculus yezoensis*, which provides more potential niches, than on a structurally simple plant *Sparganium emersum*. The rank–abundance diagrams of both are closer to the random fraction model than the MacArthur fraction model. Finally, Figure 16.5e shows how attached bacterial assemblages (biofilms), during colonization of

community indices are abstractions that may be useful when making comparisons

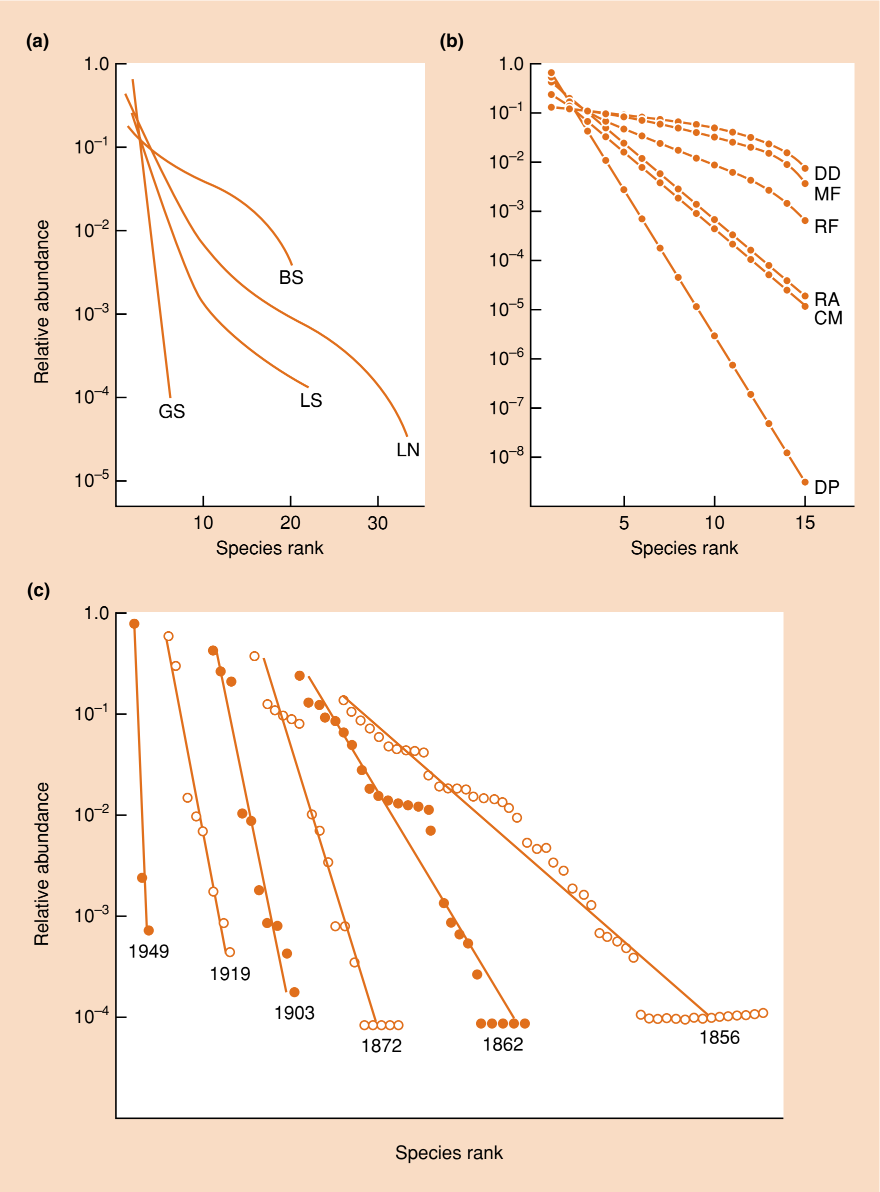


Figure 16.5 (a, b) Rank–abundance patterns of various models. Two are statistically orientated (LS and LN), whilst the rest can be described as niche orientated. (a) BS, broken stick; GS, geometric series; LN, log-normal; LS, log series. (b) CM, composite; DD, dominance decay; DP, dominance preemption; MF, MacArthur fraction; RA, random assortment; RF, random fraction. (c) Change in the relative abundance pattern (geometric series fitted) of plant species in an experimental grassland subjected to continuous fertilizer from 1856 to 1949. ((a–c) after Tokeshi, 1993.)

glass slides in a lake, change from a log-normal to a geometric pattern as the biofilm ages.

the energetics approach: an alternative to taxonomic description

Taxonomic composition and species diversity are just two of many possible ways of describing a community. Another alternative (not necessarily better but quite different) is to describe communities and ecosystems in terms of their standing crop and the rate of production of biomass by plants, and its use and conversion by heterotrophic microorganisms and animals. Studies that are orientated in this way may begin by describing the food web, and then define the biomasses at each trophic level and the flow of energy and matter from the physical environment through the living organisms and back to the physical environment. Such an approach can allow patterns to be detected amongst communities and ecosystems that may have no taxonomic features in common. This approach will be discussed in Chapters 17 and 18.

Much recent research effort has been devoted to understanding the link between species richness and ecosystem functioning (productivity, decomposition and nutrient dynamics). Understanding the role of species richness in ecosystem processes has particular significance for how humans respond to biodiversity loss. We discuss this important topic in Section 21.7.

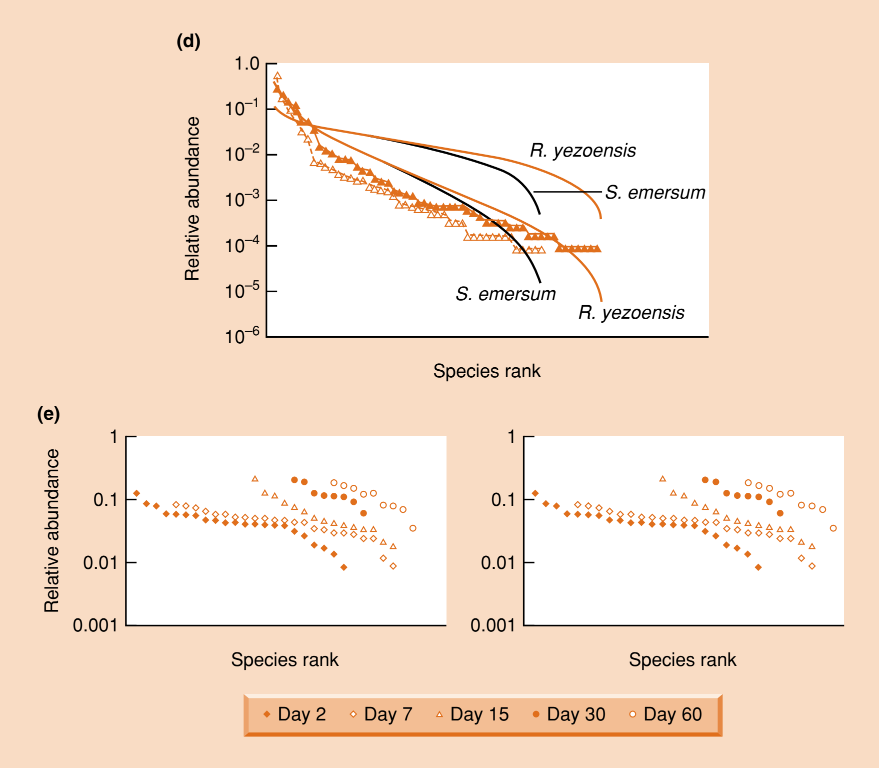


Figure 16.5 (*cont'd*) (d) Comparison of rank–abundance patterns for invertebrate species living on a structurally complex stream plant *Ranunculus yezoensis* (▲) and a simple plant *Sparganium emersum* (△); fitted lines represent the MacArthur fraction model (——, the upper one for *R. yezoensis* and the lower one for *S. emersum*) and the random fraction model (——, the upper one for *R. yezoensis* and the lower one for *S. emersum*). (After Taniguchi *et al.*, 2003.) (e) Rank–abundance patterns (based on a biomass index) for bacterial assemblages in lake biofilms of different ages (symbols from left to right represent days 2, 7, 15, 30, 60). (After Jackson *et al.*, 2001.)

16.3 Community patterns in space

16.3.1 Gradient analysis

Figure 16.6 shows a variety of ways of describing the distribution of vegetation used in a classic study in the Great Smoky Mountains (Tennessee), USA, where tree species give the vegetation its main character. Figure 16.6a shows the characteristic associations of the dominant trees on the mountainside, drawn as if the communities had sharp boundaries. The mountainside itself provides a range of conditions for plant growth, and two of these, altitude and moisture, may be particularly important in determining the distribution of the various tree species. Figure 16.6b shows the dominant associations graphed in terms of these two environmental dimensions. Finally, Figure 16.6c shows the abundance of each individual tree species (expressed as a percentage of all tree stems present) plotted against the single gradient of moisture.

species distributions along gradients end not with a bang but with a whimper

Figure 16.6a is a subjective analysis that acknowledges that the vegetation of particular areas differs in a characteristic way from that of other areas. It could be taken to imply that the various communities are sharply delimited. Figure 16.6b gives the same impression. Note that both Figure 16.6a and b are based on descriptions of the *vegetation*. However, Figure 16.6c sharpens the focus by concentrating on the pattern of distribution of the individual *species*. It is then immediately obvious that there is considerable overlap in their abundance – there are no sharp boundaries. The various tree species are now revealed as being strung out along the gradient with the tails of their distributions overlapping. The results of this 'gradient analysis' show that the limits of the distributions of each species 'end not with a bang but with a whimper'. Many other gradient studies have produced similar results.

choice of gradient is almost always subjective

Perhaps the major criticism of gradient analysis as a way of detecting pattern in communities is that the choice of the gradient is almost always subjective. The investigator searches for some feature of the environment that appears to matter to the organisms and then organizes the data about the species concerned along a gradient of that factor. It is not necessarily the most appropriate factor to have chosen. The fact that the species from a community can be arranged in a sequence along a gradient of some environmental factor does not prove that this factor is the most important one. It may only imply that the factor chosen is more or less loosely correlated with whatever really matters in the lives of the species involved. Gradient analysis is only a small step on the way to the objective description of communities.

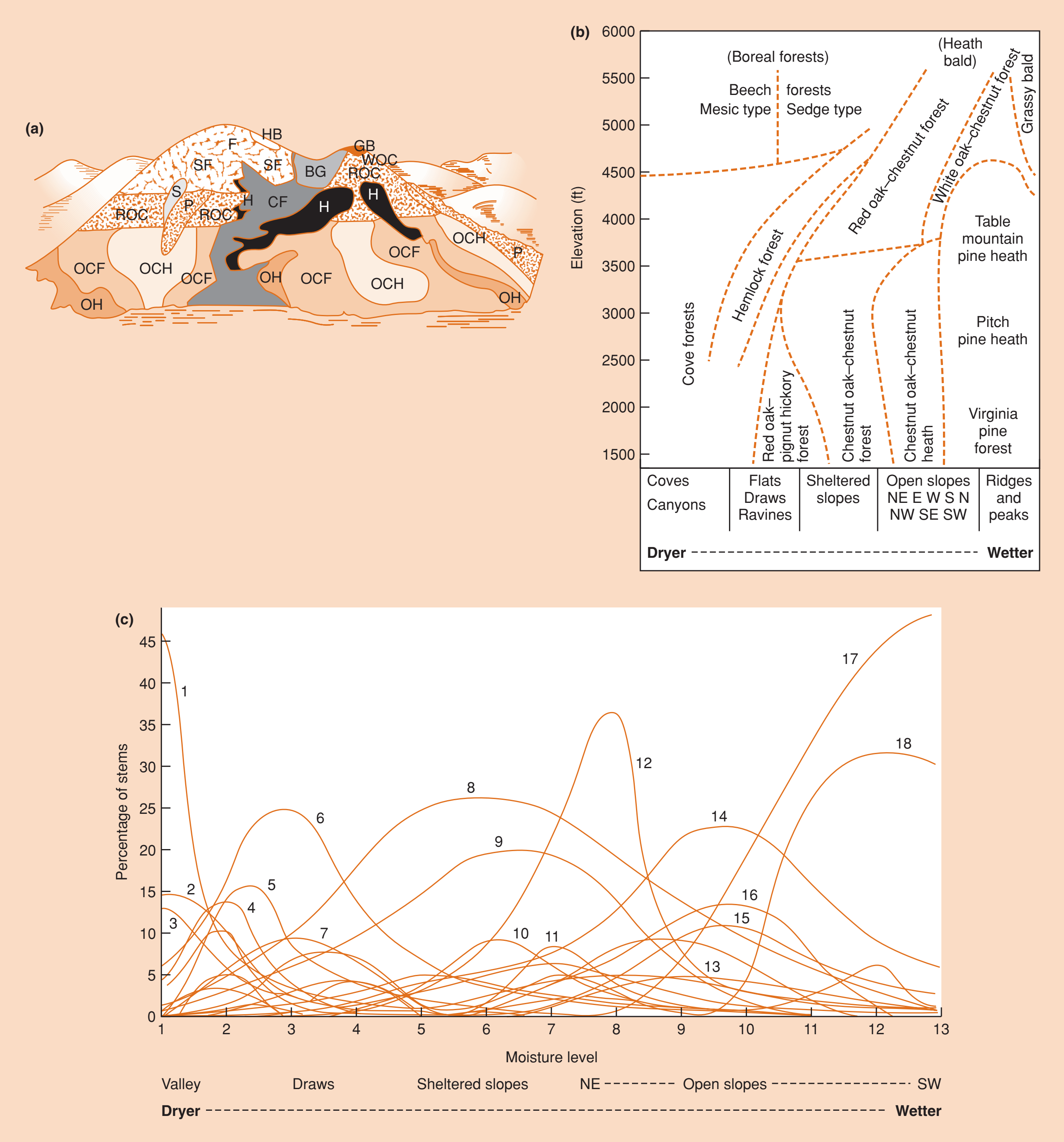
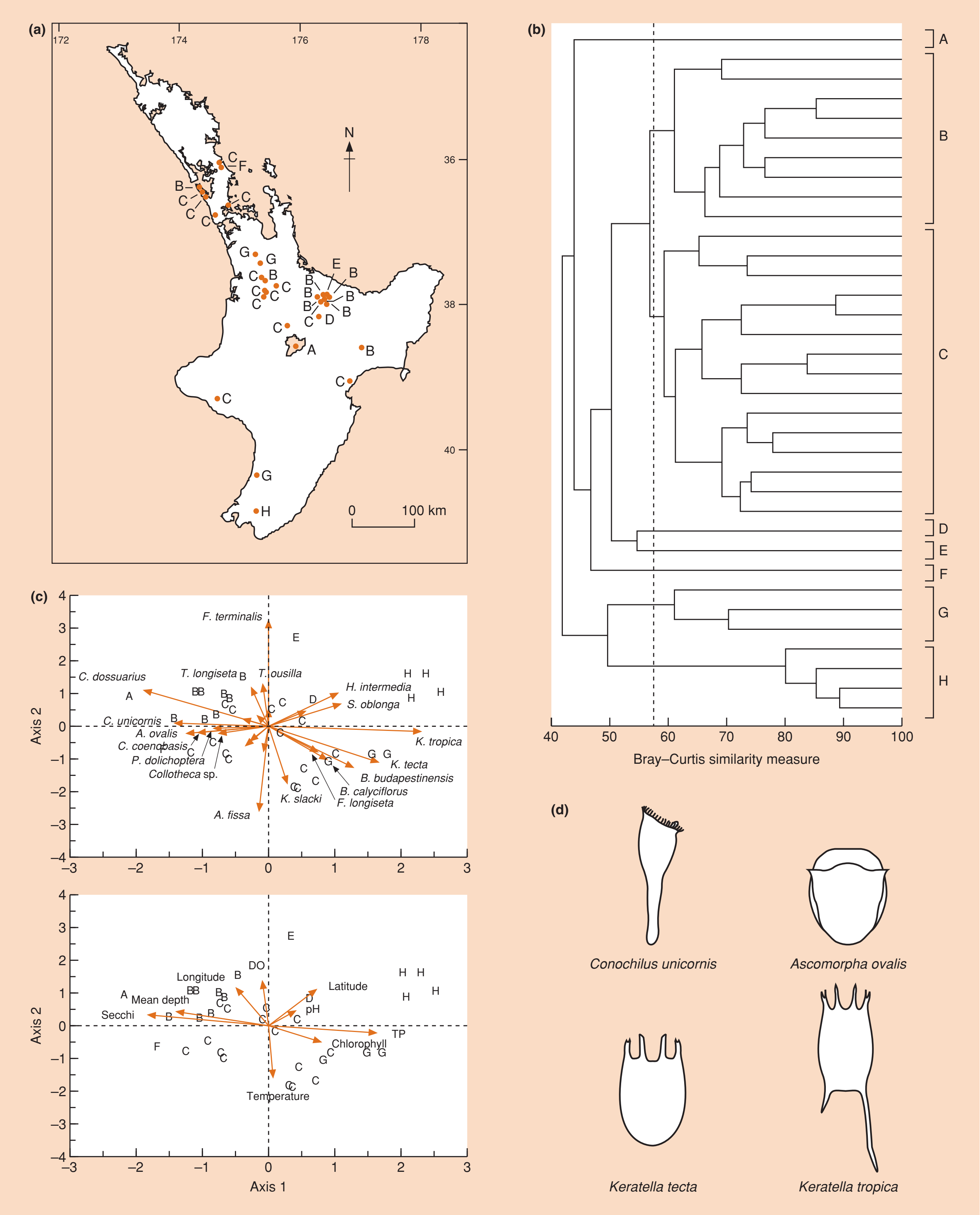


Figure 16.6 Three contrasting descriptions of distributions of the characteristic dominant tree species of the Great Smoky Mountains, Tennessee. (a) Topographic distribution of vegetation types on an idealized west-facing mountain and valley. (b) Idealized graphic arrangement of vegetation types according to elevation and aspect. (c) Distributions of individual tree populations (percentage of stems present) along the moisture gradient. Vegetation types: BG, beech gap; CF, cove forest; F, Fraser fir forest; GB, grassy bald; H, hemlock forest; HB, heath bald; OCF, chestnut oak–chestnut forest; OCH, chestnut oak–chestnut heath; OH, oak–hickory; P, pine forest and heath; ROC, red oak–chestnut forest; S, spruce forest; SF, spruce–fir forest; WOC, white oak–chestnut forest. Major species: 1, *Halesia monticola*; 2, *Aesculus octandra*; 3, *Tilia heterophylla*; 4, *Betula alleghaniensis*; 5, *Liriodendron tulipifera*; 6, *Tsuga canadensis*; 7, *B. lenta*; 8, *Acer rubrum*; 9, *Cornus florida*; 10, *Carya alba*; 11, *Hamamelis virginiana*; 12, *Quercus montana*; 13, *Q. alba*; 14, *Oxydendrum arboreum*; 15, *Pinus strobus*; 16, *Q. coccinea*; 17, *P. virginiana*; 18, *P. rigida*. (After Whittaker, 1956.)


(a)
172
174
176
178
36
38
40
N
0
100 km
(b)
A
B
C
D
E
F
G
H
40
50
60
70
80
90
100
Bray–Curtis similarity measure
(c)
F. terminalis
C. dossuarius
T. longiseta
T. pusilla
H. intermedia
S. oblonga
C. unicornis
A. ovalis
C. coenobasis
P. dolichoptera
Collotheca sp.
K. tropica
K. tecta
B. budapestinensis
B. calyciflorus
F. longiseta
K. slacki
A. fissa
Axis 2
Longitude
DO
Latitude
Mean depth
Secchi
pH
TP
Chlorophyll
Temperature
Axis 1
(d)
Conochilus unicornis
Ascomorpha ovalis
Keratella tecta
Keratella tropica

16.3.2 Classification and ordination of communities

Formal statistical techniques have been defined to take the subjectivity out of community description. These techniques allow the data from community studies to sort themselves, without the investigator putting in any preconceived ideas about which species tend to be associated with each other or which environmental variables correlate most strongly with the species distributions. One such technique is classification.

classification involves grouping similar communities together in clusters

Classification begins with the assumption that communities consist of relatively discrete entities. It produces groups of related communities by a process conceptually similar to taxonomic classification. In taxonomy, similar individuals are grouped together in species, similar species in genera, and so on. In community classification, communities with similar species compositions are grouped together in subsets, and similar subsets may be further combined if desired (see Ter Braak & Prentice, 1988, for details of the procedure).

The rotifer communities of a number of lakes in the North Island of New Zealand (Figure 16.7a) were subjected to a classification technique called cluster analysis (Duggan *et al.*, 2002). Eight clusters or classes were identified (Figure 16.7b), each based solely on the arrays of species present and their abundances. The spatial distribution of each class of rotifer community in the New Zealand lakes is shown in Figure 16.7a. Note that there is little consistent spatial relationship; communities in each class are dotted about the island. This illustrates one of the strengths of classification. Classification methods show the structure within a series of communities without the necessity of picking out some supposedly relevant environmental variable in advance, a procedure that is necessary for gradient analysis.

in ordination, communities are displayed on a graph so that those most similar in composition are closest together

Ordination is a mathematical treatment that allows communities to be organized on a graph so that those that are most similar in both species composition and relative abundance will appear closest together, whilst communities that differ greatly in the relative importance of a similar set of species, or that possess quite different species, appear far apart. Figure 16.7c shows the application of an ordination technique called canonical correspondence analysis (CCA) to the rotifer communities (Ter Braak & Smilauer 1998). CCA also allows the community patterns to be examined in terms of environmental variables. Obviously, the success of the method now depends on having sampled an appropriate variety of environmental variables. This is a major snag in the procedure – we may not have measured the qualities in the environment that are most relevant. The relationships between rotifer community composition and a variety of physicochemical factors are shown in Figure 16.7c. The link between classification and ordination can be gauged by noting that communities falling into classes A–H, derived from classification, are also fairly distinctly separated on the CCA ordination graph.

subsequently, it is necessary to ask what varies along the axes of the graph

Community classes A and B tend to be associated with high water transparency ('Secchi depth'), whereas those in classes G and H are associated with high total phosphorus and chlorophyll concentrations; the other lake classes take up intermediate positions. Lakes that have been subject to a greater level of runoff of agricultural fertilizers or input of sewage are described as eutrophic. These tend to have high phosphorus concentrations, leading to higher chlorophyll levels and lower transparency (a greater abundance of phytoplankton cells). Evidently, the rotifer communities are strongly influenced by the level of eutrophication to which the lakes are subject. Species of rotifer that are characteristic of particularly eutrophic conditions, such as *Keratella tecta* and *K. tropica* (Figure 16.7d), were strongly represented in classes G and H, while those associated with more pristine conditions, such as *Conochilus unicornis* and *Ascomorpha ovalis*, were common in classes A and B.

The level of eutrophication, however, is not the only significant factor in explaining rotifer community composition. Class C communities, for example, while characteristic of intermediate phosphorus concentrations, can be differentiated along axis 2 according to dissolved oxygen concentration and lake temperature (themselves negatively related because oxygen solubility declines with increasing temperature).

ordination can generate hypotheses for subsequent testing

What do these results tell us? First, and most specifically, the correlations with environmental factors, revealed by the analysis, give us some specific hypotheses to test about the relationship between community composition and underlying environmental factors. (Remember that correlation does not necessarily imply

Figure 16.7 (*opposite*) (a) Thirty-one lakes in the North Island of New Zealand where rotifer communities (78 species in total) were sampled and described. (b) Results of cluster analysis (classification) on species composition data from the 31 lakes (based on the Bray–Curtis similarity measure); lake communities that are most similar cluster together and eight clusters are identified (A–H). (c) Results of canonical correspondence analysis (ordination). The positions in ordination space are shown for lake sites (shown as letters A–H corresponding to their classification), individual rotifer species (orange arrows in top panel) and environmental factors (orange arrows in lower panel). (d) Silhouettes of four of the rotifer species. (After Duggan *et al.*, 2002.)

causation. For example, dissolved oxygen and community composition may vary together because of a common response to another environmental factor. A direct causal link can only be proved by controlled experimentation.)

A second, more general point is relevant to the discussion of the nature of the community. The results emphasize that under a particular set of environmental conditions, a predictable association of species is likely to occur. It shows that community ecologists have more than just a totally arbitrary and ill-defined set of species to study.

16.3.3 Problems of boundaries in community ecology

are communities discrete entities with sharp boundaries?

There may be communities that are separated by clear, sharp boundaries, where groups of species lie adjacent to, but do not intergrade into, each other. If they exist, they are exceptional. The meeting of terrestrial and aquatic environments might appear to be a sharp boundary but its ecological unreality is emphasized by the otters or frogs that regularly cross it and the many aquatic insects that spend their larval lives in the water but their adult lives as winged stages on land or in the air. On land, quite sharp boundaries occur between the vegetation types on acidic and basic rocks where outcrops meet, or where serpentine (a term applied to a mineral rich in magnesium silicate) and nonserpentine rocks are juxtaposed. However, even in such situations, minerals are leached across the boundaries, which become increasingly blurred. The safest statement we can make about community boundaries is probably that they do not exist, but that some communities are much more sharply defined than others. The ecologist is usually better employed looking at the ways in which communities grade into each other, than in searching for sharp cartographic boundaries.

the community: not so much a superorganism . . .

In the first quarter of the 20th century there was considerable debate about the nature of the community. Clements (1916) conceived of the community as a sort of *superorganism* whose member species were tightly bound together both now and in their common evolutionary history. Thus, individuals, populations and communities bore a relationship to each other resembling that between cells, tissues and organisms.

In contrast, the *individualistic* concept devised by Gleason (1926) and others saw the relationship of coexisting species as simply the results of similarities in their requirements and tolerances (and partly the result of chance). Taking this view, community boundaries need not be sharp, and associations of species would be much less predictable than one would expect from the superorganism concept.

The current view is close to the individualistic concept. Results of direct gradient analysis, ordination and classification all indicate that a given location, by virtue mainly of its physical characteristics, possesses a reasonably predictable association of species. However, a given species that occurs in one predictable association is also quite likely to occur with another group of species under different conditions elsewhere.

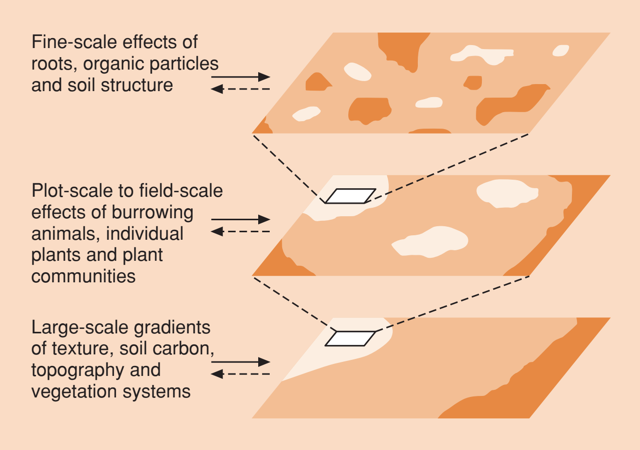


Figure 16.8 Determinants of spatial heterogeneity of communities of soil organisms including bacteria, fungi, nematodes, mites and collembolans. (After Ettema & Wardle, 2002.)

A further point needs to be born in mind when considering the question of environmental patchiness and boundaries. Spatial heterogeneity in the distribution of communities can be viewed within a series of nested scales. Figure 16.8, for example, shows patterns in spatial heterogeneity in communities of soil organisms operating at scales from hectares to square millimeters (Ettema & Wardle, 2002). At the largest scale, these reflect patterns in environmental factors related to topography and the distribution of different plant communities. But at the other extreme, fine-scale patterns may be present as a result of the location of individual plant roots or local soil structure. The boundaries of patterns at these various scale are also likely to be blurred.

. . . more a level of organization

Whether or not communities have more or less clear boundaries is an important question, but it is not the fundamental consideration. Community ecology is the study of the *community level of organization* rather than of a spatially and temporally definable unit. It is concerned with the structure and activities of the multispecies assemblage, usually at one point in space and time. It is not necessary to have discrete boundaries between communities to study community ecology.

16.4 Community patterns in time

Just as the relative importance of species varies in space, so their patterns of abundance may change with time. In either case, a species will occur only where and when: (i) it is capable of reaching a location; (ii) appropriate conditions and resources exist there; and (iii) competitors, predators and parasites do not preclude it. A temporal sequence in the appearance and disappearance of species therefore seems to require that conditions, resources and/or the influence of enemies themselves vary with time.

For many organisms, and particularly short-lived ones, their relative importance in the community changes with time of year as the individuals act out their life cycles against a background of seasonal change. Sometimes community composition shifts because of externally driven physical change, such as the build up of silt in a coastal salt marsh leading to its replacement by forest. In other cases, temporal patterns are simply a reflection of changes in key resources, as in the sequence of heterotrophic organisms associated with fecal deposits or dead bodies as they decompose (see Figure 11.2). The explanation for such temporal patterns is relatively straightforward and will not concern us here. Nor will we dwell on the variations in abundance of species in a community from year to year as individual populations respond to a multitude of factors that influence their reproduction and survival (dealt with in Chapters 5, 6 and 8–14).

Our focus will be on patterns of community change that follow a disturbance, defined as a relatively discrete event that removes organisms (Townsend & Hildrew, 1994) or otherwise disrupts the community by influencing the availability of space or food resources, or by changing the physical environment (Pickett & White, 1985). Such disturbances are common in all kinds of community. In forests, they may be caused by high winds, lightning, earthquakes, elephants, lumberjacks or simply by the death of a tree through disease or old age. Agents of disturbance in grassland include frost, burrowing animals and the teeth, feet, dung or dead bodies of grazers. On rocky shores or coral reefs, disturbances may result from severe wave action during hurricanes, tidal waves, battering by logs or moored boats or the fins of careless scuba divers.

16.4.1 Founder-controlled and dominance-controlled communities

founder control: many species are equivalent in their ability to colonize

In response to disturbances, we can postulate two fundamentally different kinds of community response according to the type of competitive relationships exhibited by the component species – founder controlled and dominance controlled (Yodzis, 1986). *Founder-controlled* communities will occur if a large number of species are approximately equivalent in their ability to colonize an opening left by a disturbance, are equally well fitted to the abiotic environment and can hold the location until they die. In this case, the result of the disturbance is essentially a lottery. The winner is the species that happens to reach and establish itself in the disturbed location first. The dynamics of founder-controlled communities are discussed in Section 16.7.4.

dominance control: some potential colonizers are competitively dominant

Dominance-controlled communities are those where some species are competitively superior to others so that an initial colonizer of an opening left by a disturbance cannot necessarily maintain its presence there. In these cases, disturbances lead to reasonably predictable sequences of species because different species have different strategies for exploiting resources – early species are good colonizers and fast growers, whereas later species can tolerate lower resource levels and grow to maturity in the presence of early species, eventually outcompeting them. These situations are more commonly known by the term *ecological succession*, defined as the *nonseasonal, directional and continuous pattern of colonization and extinction on a site by species populations*.

16.4.2 Primary and secondary successions

primary succession: an exposed landform uninfluenced by a previous community

Our focus is on successional patterns that occur on newly exposed landforms. If the exposed landform has not previously been influenced by a community, the sequence of species is referred to as a primary succession. Lava flows and pumice plains caused by volcanic eruptions (see Section 16.4.3), craters caused by the impact of meteors (Cockell & Lee, 2002), substrate exposed by the retreat of a glacier (Crocker & Major, 1955) and freshly formed sand dunes (see Section 16.4.4) are examples. In cases where the vegetation of an area has been partially or completely removed, but where well-developed soil and seeds and spores remain, the subsequent sequence of species is termed a secondary succession. The loss of trees locally as a result of disease, high winds, fire or felling may lead to secondary successions, as can cultivation followed by the abandonment of farmland (so-called old field successions – see Section 16.4.5).

secondary succession: vestiges of a previous community are still present

Successions on newly exposed landforms typically take several hundreds of years to run their course. However, a precisely analagous process occurs amongst the animals and algae on recently denuded rock walls in the marine subtidal zone, and this succession takes only a decade or so (Hill *et al*., 2002). The research life of an ecologist is sufficient to encompass a subtidal succession but not that following glacial retreat. Fortunately, however, information can sometimes be gained over the longer timescale. Often, successional stages in time are represented by community gradients in space. The use of historic maps, carbon dating or other techniques may enable the age of a community since exposure of the landform to be estimated. A series of communities currently in existence, but corresponding to different lengths of time since the onset of succession, can be inferred to reflect succession. However, whether or not different communities that are spread out in space really do represent various stages of succession must be judged with caution. We must remember, for example, that in northern temperate areas the vegetation we see may still be undergoing recolonization and responding to climatic change following the last ice age (see Chapter 1).

16.4.3 Primary succession on volcanic lava

facilitation: early successional species on volcanic lava pave the way for later ones

A primary succession on basaltic volcanic flows on Miyake-jima Island, Japan, was inferred from a known chronosequence (16, 37, 125 and >800 years old) (Figure 16.9a). In the 16-year-old flow, soil was very sparse and lacking in nitrogen; vegetation was absent except for a few small alder trees (*Alnus sieboldiana*). In the older plots, 113 taxa were recorded, including ferns, herbaceous perennials, lianas and trees. Of most significance in this primary succession were: (i) the successful colonization of the bare lava by the nitrogen-fixing alder; (ii) the facilitation (through improved nitrogen availability) of mid-successional *Prunus speciosa* and the late successional evergreen tree *Machilus thunbergii*; (iii) the formation of a mixed forest and the shading out of *Alnus* and *Prunus*; and (iv) finally, the replacement of *Machilus* by the longer lived *Castanopsis sieboldii* (Figure 16.9b).

16.4.4 Primary succession on coastal sand dunes

importance of seed availability rather than facilitation in sand dune succession

An extensive chronosequence of dune-capped beach ridges has been undertaken on the coast of Lake Michigan in the USA. Thirteen ridges of known

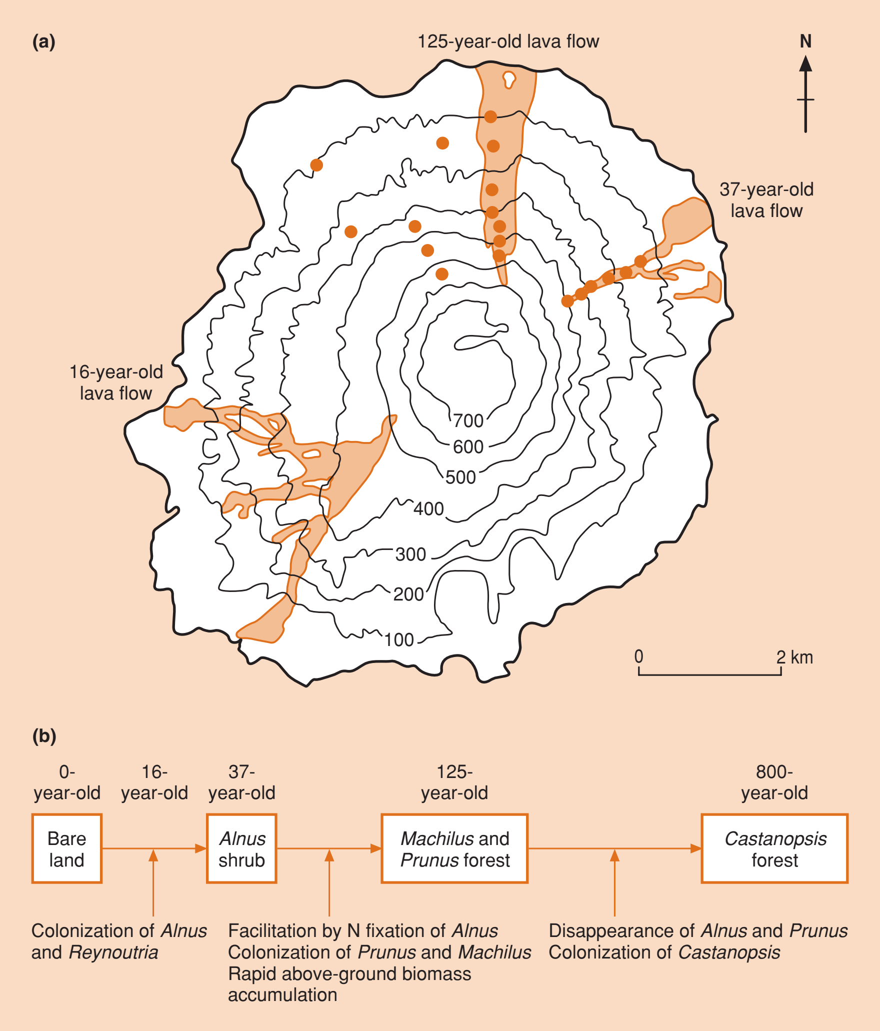


Figure 16.9 (a) Vegetation was described on 16-, 37- and 125-year-old lava flows on Miyake-jima Island, Japan. Analysis of the 16-year-old flow was nonquantitative (no sample sites shown). Sample sites on the other flows are shown as solid circles. Sites outside the three flows are at least 800 years old. (b) The main features of the primary succession in relation to lava age. (After Kamijo *et al.*, 2002.)

age (30–440 years old) show a clear pattern of primary succession to forest (Lichter, 2000). The dune grass *Ammophila breviligulata* dominates the youngest, still mobile dune ridge, but shrubby *Prunus pumila* and *Salix* spp. are also present. Within 100 years, these are replaced by evergreen shrubs such as *Juniperus communis* and by prairie bunch grass *Schizachrium scoparium*. Conifers such as *Pinus* spp., *Larix laricina*, *Picea strobus* and *Thuja occidentalis* begin colonizing the dune ridges after 150 years, and a mixed forest of *Pinus strobus* and *P. resinosa* develops between 225 and 400 years. Deciduous trees such as the oak *Quercus rubra* and the maple *Acer rubrum* do not become important components of the forest until 440 years.

It used to be thought that early successional dune species facilitated the later species by adding organic matter to the soil and increasing the availability of soil moisture and nitrogen (as in the volcanic primary succession). However, experimental seed addition and seedling transplant experiments have shown that later species are capable of germinating in young dunes (Figure 16.10a). While the more developed soil of older dunes may improve the performance of late successional species, their successful colonization of young dunes is mainly constrained by limited seed dispersal, together with seed predation by rodents (Figure 16.10b). *Ammophila* generally colonizes young, active dunes through horizontal vegetative growth. *Schizachrium*, one of the

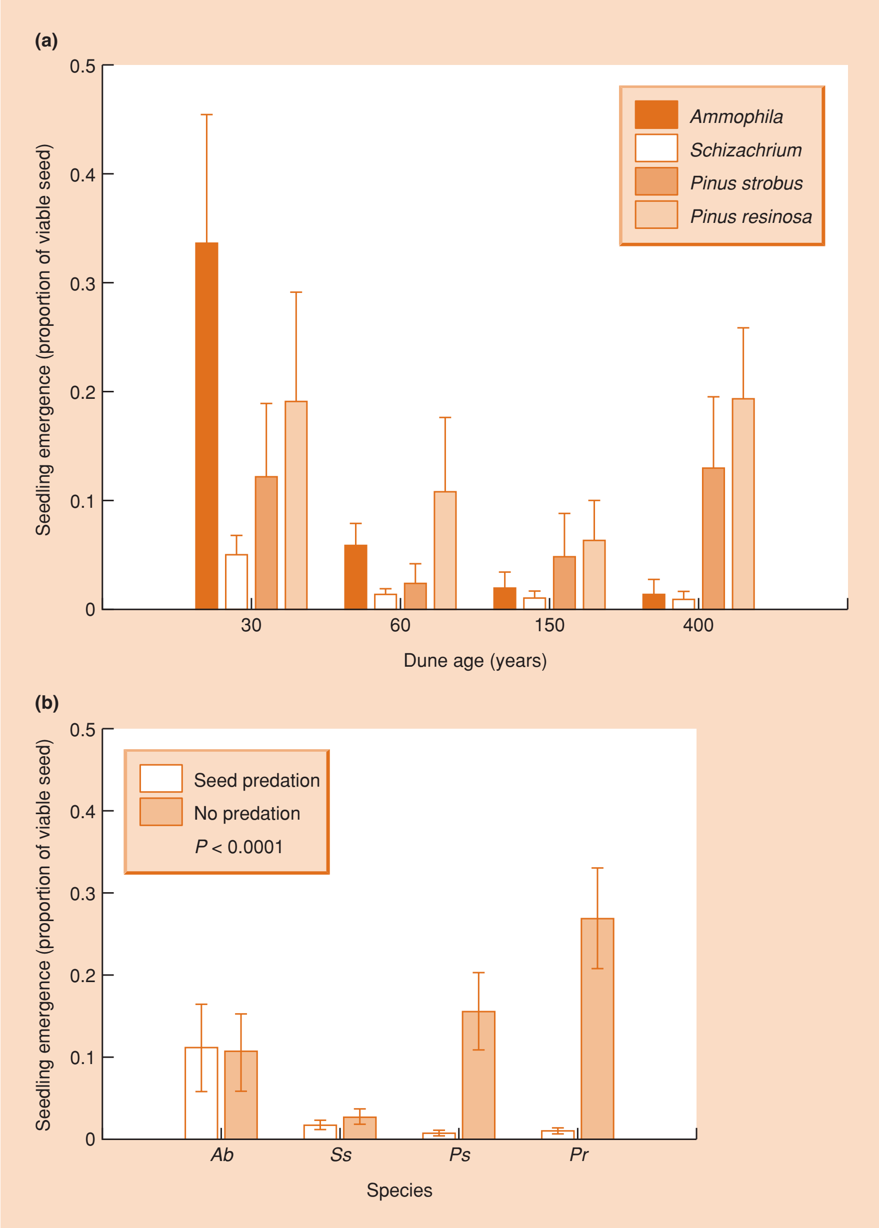


Figure 16.10 (a) Seedling emergence (means + SE) from added seeds of species typical of different successional stages on dunes of four ages. (b) Seedling emergence of the four species (*Ab*, *Ammophila breviligulata*, *Ss*, *Schizachrium scoparium*, *Ps*, *Pinus strobus*, *Pr*, *Pinus resinosa*) in the presence and absence of rodent predators of seeds (After Lichter, 2000.)

dominants of open dunes before forest development, has rates of germination and seedling establishment that are no better than *Pinus*, but its seeds are not preyed upon. Also, *Schizachrium* has the advantage of quickly reaching maturity and can continue to provide seeds at a high rate. These early species are eventually competitively excluded as trees establish and grow. Lichter (2000) considers that dune succession is better described in terms of the transient dynamics of colonization and competitive displacement, rather than the result of facilitation by early species (improving soil conditions) followed by competitive displacement.

16.4.5 Secondary successions in abandoned fields

abandoned old fields: succession to forest in North America . . .

Successions on old fields have been studied particularly along the eastern part of the USA where many farms were abandoned by farmers who moved west after the frontier was opened up in the 19th century (Tilman, 1987, 1988). Most of the precolonial mixed conifer–hardwood forest had been destroyed, but regeneration was swift. In many places, a series of sites that were abandoned for different, recorded periods of time are available for study. The typical sequence of dominant vegetation is: annual weeds, herbaceous perennials, shrubs, early successional trees and late successional trees.

. . . but to grassland in China

Old-field succession has also been studied in the productive Loess Plateau in China, which for millennia has been affected by human activities so that few areas of natural vegetation remain. The Chinese government has launched some conservation projects focused on the recovery of damaged ecosystems. A big question mark is whether the climax vegetation of the Plateau will prove to be grassland steppe or forest. Wang (2002) studied the vegetation at four plots abandoned by farmers for known periods of time (3, 26, 46 and 149 years). He was able to age some of his plots in an unusual manner. Graveyards in China are sacred and human activities are prohibited in their vicinity – gravestone records indicated how long ago the older areas had been taken out of agricultural production. Of a total of 40 plant species identified, several were considered dominant at the four successional stages (in terms of relative abundance and relative ground cover). In the first stage (recently abandoned farmland) *Artemesia scoparia* and *Seraria viridis* were most characteristic, at 26 years *Lespedeza davurica* and *S. viridis* dominated, at 46 years *Stipa bungeana*, *Bothriochloa ischaemun*, *A. gmelinii* and *L. davurica* were most important, while at 149 years *B. ischaemun* and *A. gmelinii* were dominant (Figure 16.11). The early successional species were annuals and biennials with high seed production. By 26 years, the perennial herb *L. davurica*, with its ability to spread laterally by vegetative means and a well-developed root system, had replaced *A. scoparia*.

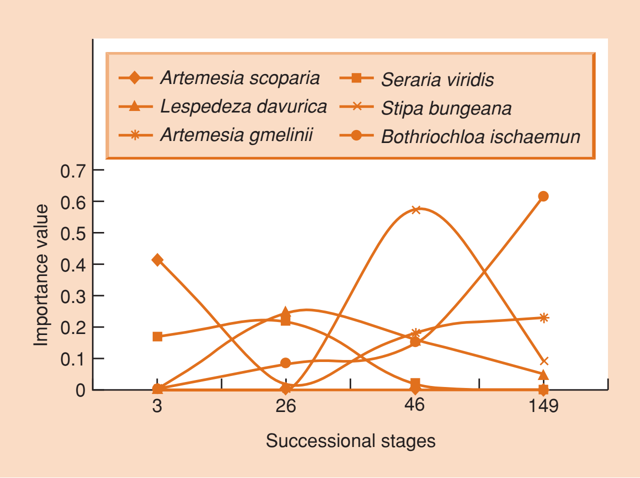


Figure 16.11 Variation in the relative importance of six species during an old-field succession on the Loess Plateau in China. (After Wang, 2002.)

The 46-year-old plot was characterized by the highest species richness and diverse life history strategies, dominated by perennial lifestyles. The dominance of *B. ischaemun* at 149 years was related to its perennial nature, ability to spread clonally and high competitive ability. As in Tilman's (1987, 1988) North American studies, soil nitrogen content increased during the succession and may have facilitated some species in the succession. Wang concludes that the grass *B. ischaemun* is the characteristic climax species in this Loess Plateau habitat, and thus the vegetation seems likely to succeed to steppe grassland rather than forest.

16.5 Species replacement probabilities during successions

forest succession can be represented as a tree-by-tree replacement model . . .

A model of succession developed by Horn (1981) sheds some light on the successional process. Horn recognized that in a hypothetical forest community it would be possible to predict changes in tree species composition given two things. First, one would need to know for each tree species the probability that, within a particular time interval, an individual would be replaced by another of the same species or of a different species. Second, an initial species composition would have to be assumed.

Horn considered that the proportional representation of various species of saplings established beneath an adult tree reflected the probability of an individual tree's replacement by each of those species. Using this information, he estimated the probability, after 50 years, that a site now occupied by a given species will be taken over by another species or will still be occupied by

Table 16.1 A 50-year tree-by-tree transition matrix from Horn (1981). The table shows the probability of replacement of one individual by another of the same or different species 50 years hence.

	Occupant 50 years hence			
Present occupant	*Grey birch*	*Blackgum*	*Red maple*	*Beech*
Grey birch	0.05	0.36	0.50	0.09
Blackgum	0.01	0.57	0.25	0.17
Red maple	0.0	0.14	0.55	0.31
Beech	0.0	0.01	0.03	0.96

the same species (Table 16.1). Thus, for example, there is a 5% chance that a location now occupied by grey birch will still support grey birch in 50 years' time, whereas there is a 36% chance that blackgum will take over, a 50% chance for red maple and 9% for beech.

Beginning with an observed distribution of the canopy species in a stand in New Jersey in the USA known to be 25 years old, Horn modeled the changes in species composition over several centuries. The process is illustrated in simplified form in Table 16.2 (which deals with only four species out of those present). The progress of this hypothetical succession allows several predictions to be made. Red maple should dominate quickly, whilst grey birch disappears. Beech should slowly increase to predominate later, with blackgum and red maple persisting at low abundance. All these predictions are borne out by what happens in the real succession (final column).

. . . that predicts a stable species composition and the time taken to reach it

The most interesting feature of Horn's so-called Markov chain model is that, given enough time, it converges on a stationary, stable composition that is independent of the initial composition of the forest. The outcome is inevitable (it depends only on the matrix of replacement probabilities) and will be achieved whether the starting point is 100% grey birch or 100% beech, 50% blackgum and 50% red maple, or any other combination (as long as adjacent areas provide a source of seeds of species not initially present). Korotkov *et al.* (2001) have used a similar Markov modeling approach to predict the time it should take to reach the climax state from any other stage in old-field successions culminating in mixed conifer–broadleaf forest in central Russia. From field abandonment to climax is predicted to take 480–540 years, whereas a mid-successional stage of birch forest with spruce undergrowth should take 320–370 years to reach the climax.

Since Markov models seem to be capable of generating quite accurate predictions, they may prove to be a useful tool in formulating plans for forest management. However, the models are simplistic and the assumption that transition probabilities remain constant in space and over time and are not affected by historic factors, such as initial biotic conditions and the order of arrival of species, are likely to be wrong in many cases (Facelli & Pickett, 1990). Hill *et al.* (2002) addressed the question of spatiotemporal variation in species replacement probabilities in a subtidal community succession including sponges, sea anenomes, polychaetes and encrusting algae. In this case, the predicted successions and endpoints were similar whether replacement probabilities were averaged or were subject to realistic spatial or temporal variation. And the outcomes of all three models were very similar to the observed community structure (Figure 16.12).

Table 16.2 The predicted percentage composition of a forest consisting initially of 100% grey birch. (After Horn, 1981.)

	Age of forest (years)						
Species	*0*	*50*	*100*	*150*	*200*	*∞*	*Data from old forest*
Grey birch	100	5	1	0	0	0	0
Blackgum	0	36	29	23	18	5	3
Red maple	0	50	39	30	24	9	4
Beech	0	9	31	47	58	86	93

16.6 Biological mechanisms underlying successions

an ideal theory of succession should predict *and* explain

Despite the advantages of simple Markov models, a theory of succession should ideally not only predict but also explain. To do this, we need to consider the *biological* basis for the replacement values in the model, and here we have to turn to alternative approaches.

16.6.1 Competition–colonization trade-off and successional niche mechanisms

a trade-off between colonization and competitive ability?

Rees *et al.* (2001) drew together a diversity of experimental, comparative and theoretical approaches to produce some generalizations about vegetation dynamics. Early successional plants have a series of correlated traits, including high fecundity, effective dispersal, rapid growth when resources are abundant, and poor growth and survival when resources are scarce. Late successional species usually have the opposite traits, including an ability to grow, survive and compete when resources are scarce. In the absence of disturbance, late successional species eventually outcompete early species, because they reduce resources beneath the levels required by the early successional species. Early species persist for two reasons: (i) because their dispersal ability and high fecundity permits them to colonize and establish in recently disturbed sites before late successional species can arrive; or (ii) because rapid growth under resource-rich conditions allows them to temporarily outcompete late successional species even if they arrive at the same time. Rees and his colleagues refer to the first mechanism as a *competition–colonization trade-off* and the second as the *successional niche* (early conditions suit early species because of their niche requirements). The competition–colonization trade-off is strengthened by a further physiological inevitability. Huge differences in per capita seed production among plant species are inversely correlated to equally large variations in seed size; plants producing tiny seeds tend to produce many more of them than plants producing large seeds (see Section 4.8.5). Thus, Rees *et al.* (2001) point out that small-seeded species are good colonists (many propagules) but poor competitors (small seed food reserves), and vice versa for large-seeded species.

16.6.2 Facilitation

the importance of facilitation – but not always

Cases of competition–colonization trade-offs and/or successional niche relations are prominent in virtually every succession that has been described, including all those in the previous section. In addition, we have seen cases where early species may change the abiotic environment in ways (e.g. increased soil nitrogen) that make it easier for later species to establish and thrive. Thus, *facilitation* has to be added to the list of phenomena underlying some successions. We cannot say how common this state of affairs is. However, the converse is by no means uncommon; thus, many plant species alter the environment in a way that makes it more, rather than less, suitable for themselves (Wilson & Agnew, 1992). Thus, for example, woody vegetation can trap water from fog or ameliorate frosts, improving the conditions for growth of the species concerned, whilst grassy swards can intercept surface flowing water and grow better in the moister soil that is created.

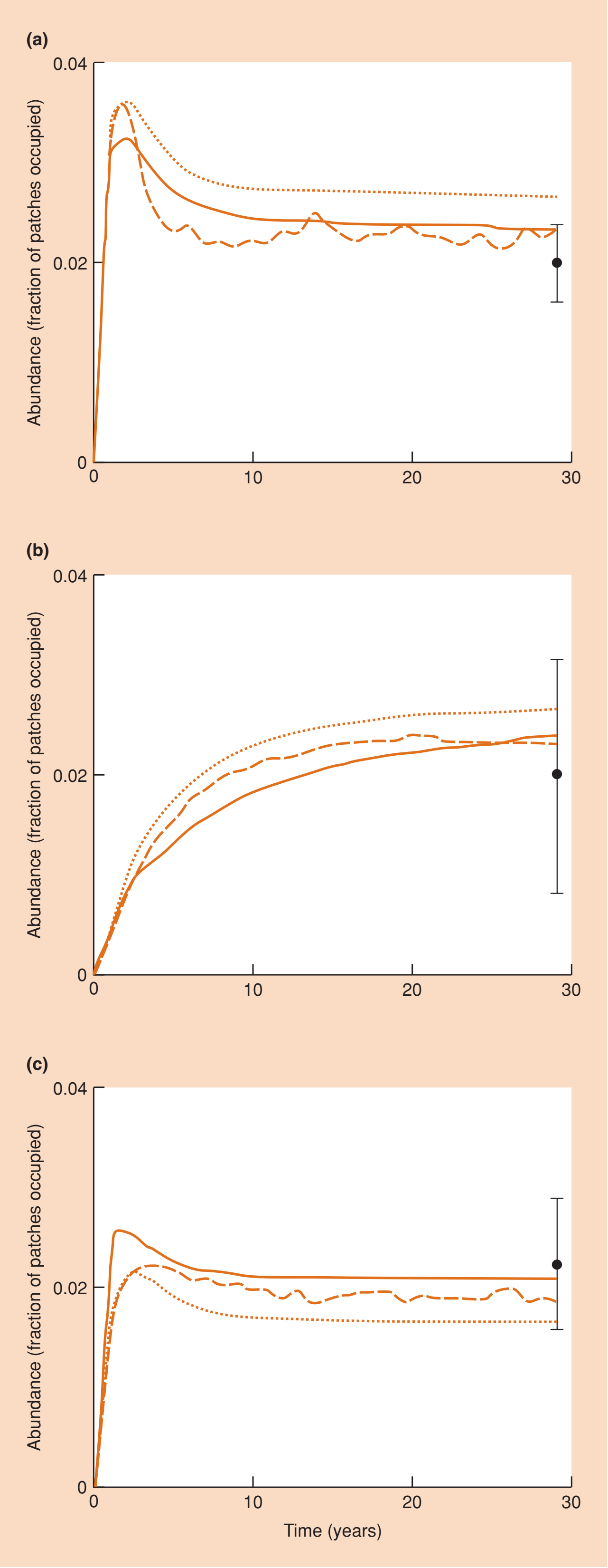


Figure 16.12 (*left*) Simulated recovery dynamics (Markov chain models) of three of the species that make up a subtidal community starting from 100% bare rock for spatially varying, time varying or homogeneous replacement probabilities: (a) the bryozoan *Crisia eburnea*, (b) the sea anenome *Metridium senile* and (c) encrusting coralline algae. The points at the end of each plot (±95% confidence intervals) are the observed abundances at a site in the Gulf of Maine, USA. (After Hill *et al.*, 2002.)

16.6.3 Interactions with enemies

an important role for seed predation?

Rees *et al.* (2001) point out that it follows from the competition–colonization trade-off that recruitment of competitively dominant plants should be determined largely by the rate of arrival of their seeds. This means that herbivores that reduce seed production are more likely to reduce the density of dominant competitors than of subordinates. Recall that this is just what happened in the sand-dune study described in Section 16.4.4. In a similar vein, Carson and Root (1999) showed that by removing insect predators of seeds, the meadow goldenrod (*Solidago altissima*), which normally appears about 5 years into an old-field succession, became dominant after only 3 years. This happened because release from seed predation allowed it to outcompete earlier colonists more quickly.

Thus, apart from competition–colonization trade-off, successional niche and facilitation, we have to add a fourth mechanism – interactions with enemies – if we are to fully understand plant successions. Experimental approaches, such as that employed to understand the role of seed predators, have also shown that the nature of soil food webs (Gange & Brown, 2002), the presence and disturbance of litter (Ganade & Brown, 2002), and the presence of mammals that consume vegetation (Cadenasso *et al.*, 2002) sometimes play roles in determining successional sequences.

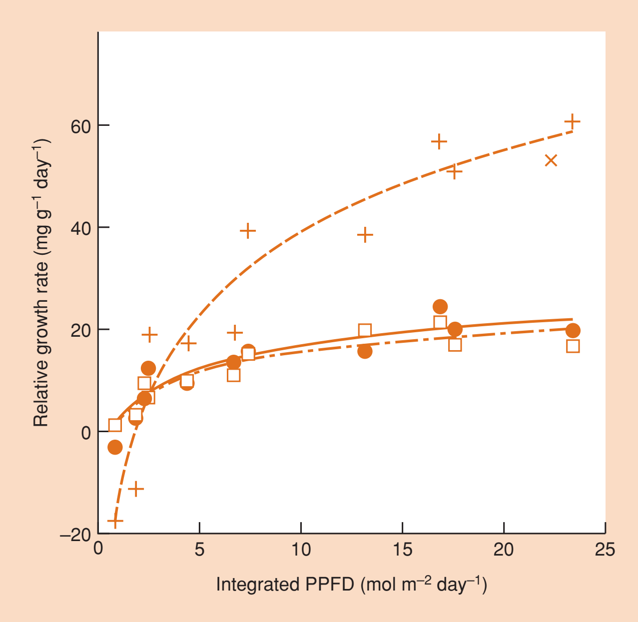


Figure 16.13 Relative growth rate (during the July–August 1994 growing season) of trembling aspen (+), northern red oak (●) and sugar maple (□) in relation to photosynthetic photon flux density (PPFD). (After Kaelke *et al.*, 2001.)

16.6.4 Resource-ratio hypothesis

Tilman's resource-ratio hypothesis emphasizes changing competitive abilities

A further example of a successional niche being responsible for species replacement is worth highlighting. Trembling aspen (*Populus tremuloides*) is a tree that appears earlier in successions in North America than northern red oak (*Quercus rubra*) or sugar maple (*Acer saccharum*). Kaelke *et al.* (2001) compared the growth of seedlings of all three species when planted along a gradient of light availability ranging from forest understory (2.6% of full light) to small clearings (69% of full light). The aspen outgrew the others when relative light availability exceeded 5%. However, there was a rank reversal in relative growth rate in deep shade; here the oak and maple, typical of later stages of succession, grew more strongly and survived better than aspen (Figure 16.13). In his *resource-ratio* hypothesis of succession, Tilman (1988) places strong emphasis on the role of changing relative competitive abilities of plant species as conditions slowly change with time. He hypothesized that species dominance at any point in a terrestrial succession is strongly influenced by the relative availability of two resources: not just by light (as demonstrated by Kaelke *et al.*, 2001) but also by a limiting soil nutrient (often nitrogen). Early in succession, the habitat experienced by seedlings has low nutrient but high light availability. As a result of litter input and the activities of decomposer organisms, nutrient availability increases with time – this can be expected to be particularly marked in primary successions that begin with a very poor soil (or no soil at all). But total plant biomass also increases with time and, in consequence, light penetration to the soil surface decreases. Tilman's ideas are illustrated in Figure 16.14 for five hypothetical species. Species A has the lowest requirement for the nutrient and the highest requirement for light at the soil surface. It has a short, prostrate growth form. Species E, which is the superior competitor in high-nutrient, low-light habitats, has the lowest requirement for light and the highest for the nutrient. It is a tall, erect species. Species B, C and D are intermediate in their requirements and each reaches its peak abundance at a different point along the soil nutrient–light gradient. There is scope for further experimental testing of Tilman's hypothesis.

16.6.5 Vital attributes

beyond just competitive ability: Noble and Slatyer's 'vital attributes'

Noble and Slatyer (1981) were also interested in defining the qualities that determine the place of a species in a succession. They called these properties *vital attributes*. The two most important relate to: (i) the method of recovery after

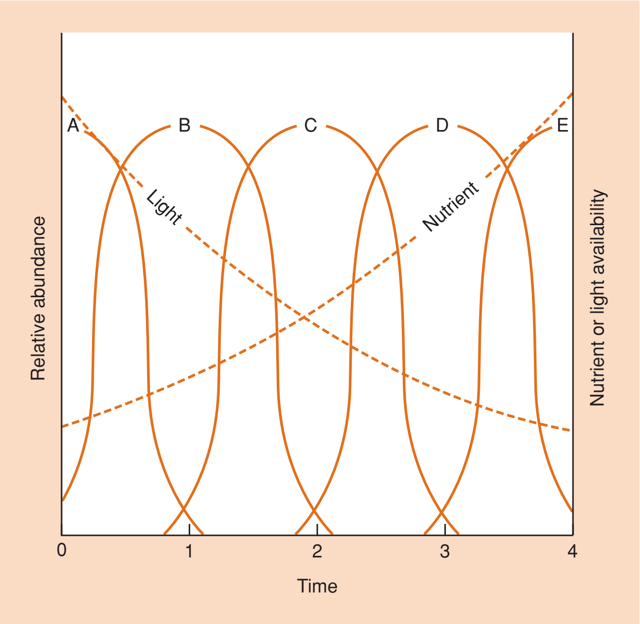


Figure 16.14 Tilman's (1988) resource-ratio hypothesis of succession. Five hypothetical plant species are assumed to be differentiated in their requirements for a limiting soil nutrient and light. During the succession, the habitat starts with a nutrient-poor soil but high light availability, changing gradually into a habitat with a rich soil but low availability of light at the soil surface. Relative competitive abilities change as conditions vary, and first one species and then another comes to dominate.

disturbance (four classes are defined: vegetative spread, V; seedling pulse from a seed bank, S; seedling pulse from abundant dispersal from the surrounding area, D; no special mechanism with just moderate dispersal from only a small seed bank, N); and (ii) the ability of individuals to reproduce in the face of competition (defined in terms of tolerance T at one extreme and intolerance I at the other). Thus, for example, a species may be classed as SI if disturbance releases a seedling pulse from a seed bank, and if the plants are intolerant of competition (being unable to germinate or grow in competition with older or more advanced individuals of either their own or another species). Seedlings of such a species could establish themselves only immediately after a disturbance, when competitors are rare. Of course, a seedling pulse fits well with such a pioneer existence. An example is the annual *Ambrosia artemisiifolia* which often figures early in old-field successions. In contrast, the American beech (*Fagus grandifolia*) could be classed as VT (being able to regenerate vegetatively from root stumps, and tolerant of competition since it is able to establish itself and reproduce in competition with older or more advanced individuals of either its own or another species) or NT (if no stumps remain, it would invade slowly via seed dispersal). In either case, it would eventually displace other species and form part of the 'climax' vegetation. Noble and Slatyer argue that it should be possible to classify all the species in an area according to these two vital attributes (to which relative longevity might be added as a third). Given this information, quite precise predictions about successional sequences should be possible.

Lightning-induced fires produce regular and natural disturbances in many ecosystems in arid parts of the world and two fire-response syndromes, analogous to two of Noble and Slatyer's disturbance recovery classes, can be identified. Resprouters have massive, deeply penetrating root systems, and survive fires as individuals, whereas reseeders are killed by the fire but re-establish through heat-stimulated germination and growth of seedlings (Bell, 2001). The proportion of species that can be classified as resprouters is higher in forest and shrubland vegetation of southwest Western Australia (Mediterranean-type climate) than in more arid areas of the continent. Bell suggests that this is because the Western Australian communities have been subject to more frequent fires than other areas, conforming to the hypothesis that short intervals between fires (averaging 20 years or less in many areas of Western Australia) promote the success of resprouters. Longer intervals between fires, on the other hand, allow fuel loads to build up so that fires are more intense, killing resprouters and favoring the reseeding strategy.

r and *K* species and succession

The consideration of vital attributes from an evolutionary point of view suggests that certain attributes are likely to occur together more often than by chance. We can envisage two alternatives that might increase the fitness of an organism in a succession (Harper, 1977), either: (i) the species reacts to the competitive selection pressures and evolves characteristics that enable it to persist longer in the succession, i.e. it responds to *K* selection; or (ii) it may develop more efficient mechanisms of escape from the succession, and discover and colonize suitable early stages of succession elsewhere, i.e. it responds to *r* selection (see Section 4.12). Thus, from an evolutionary point of view, good colonizers can be expected to be poor competitors and vice versa. This is evident in Table 16.3, which lists some physiological characteristics that tend to go together in early and late successional plants.

16.6.6 The role of animals in successions

necromass and the late successional role of trees

The structure of communities and the successions within them have most often been treated as essentially botanical matters. There are obvious reasons for this. Plants commonly provide most of the biomass and the physical structure of communities; moreover, plants do not hide or run away and this makes it rather easy to assemble species lists, determine abundances and detect change. The massive contribution that plants make to

Table 16.3 Physiological characteristics of early and late successional plants. (After Bazzaz, 1979.)

Attribute	*Early successional plants*	*Late successional plants*
Seed dispersal in time	Well dispersed	Poorly dispersed
Seed germination:		
enhanced by		
light	Yes	No
fluctuating temperatures	Yes	No
high NO_3^-	Yes	No
inhibited by		
far-red light	Yes	No
high CO_2 concentration	Yes	No?
Light saturation intensity	High	Low
Light compensation point	High	Low
Efficiency at low light	Low	High
Photosynthetic rates	High	Low
Respiration rates	High	Low
Transpiration rates	High	Low
Stomatal and mesophyll resistances	Low	High
Resistance to water transport	Low	High
Recovery from resource limitation	Fast	Slow
Resource acquisition rates	Fast	Slow?

determining the character of a community is not just a measure of their role as the primary producers, it is also a result of their slowness to decompose. The plant population not only contributes biomass to the community, but is also a major contributor of *necromass*. Thus, unless microbial and detritivore activity is fast, dead plant material accumulates as leaf litter or as peat. Moreover, the dominance of trees in so many communities comes about because they accumulate dead material; the greater part of a tree's trunk and branches is dead. The tendency in many habitats for shrubs and trees to succeed herbaceous vegetation comes largely from their ability to hold leaf canopies (and root systems) on an extending skeleton of predominantly dead support tissue (the heart wood).

animals are often affected by, but may also affect, successions

Animal bodies decompose much more quickly, but there are situations where animal remains, like those of plants, can determine the structure and succession of a community. This happens when the animal skeleton resists decomposition, as is the case in the accumulation of calcified skeletons during the growth of corals. A coral reef, like a forest or a peat bog, gains its structure, and drives its successions, by accumulating its dead past. Reef-forming corals, like forest trees, gain their dominance in their respective communities by holding their assimilating parts progressively higher on predominantly dead support. In both cases, the organisms have an almost overwhelming effect on the abiotic environment, and they 'control' the lives of other organisms within it. The coral reef community (dominated by an animal, albeit one with a plant symbiont) is as structured, diverse and dynamic as a tropical rainforest.

The fact that plants dominate most of the structure and succession of communities does not mean that animals always follow the communities that plants dictate. This will often be the case, of course, because the plants provide the starting point for all food webs and determine much of the character of the physical environment in which animals live. But it is also sometimes the animals that determine the nature of the plant community. We have already seen how seed-eating insects and rodents can slow successions in old fields and sand dunes by causing a higher seed mortalilty of later successional species. A particularly dramatic example of a role for animals, and on a much larger scale, comes from the savanna at Ndara in Kenya. The vegetation in savannas is often held in check by grazers. The experimental exclusion of elephants from a plot of savanna led to a more than threefold increase in the density of trees over a 10-year period (work by Oweyegha-Afundaduula, reported in Deshmukh, 1986).

More often though, animals are passive followers of successions amongst the plants. This is certainly the case for passerine bird species in an old-field succession (Figure 16.15). Arbuscular mycorrhizal fungi (see Section 13.8.2), which show a clear sequence of species replacement in the soils associated with an old-field succession (Johnson *et al.*, 1991), may also be passive followers of the plants. But this does not mean that the birds, which eat seeds, or the fungi, which affect plant growth and survival, do not influence the succession in its course. They probably do.

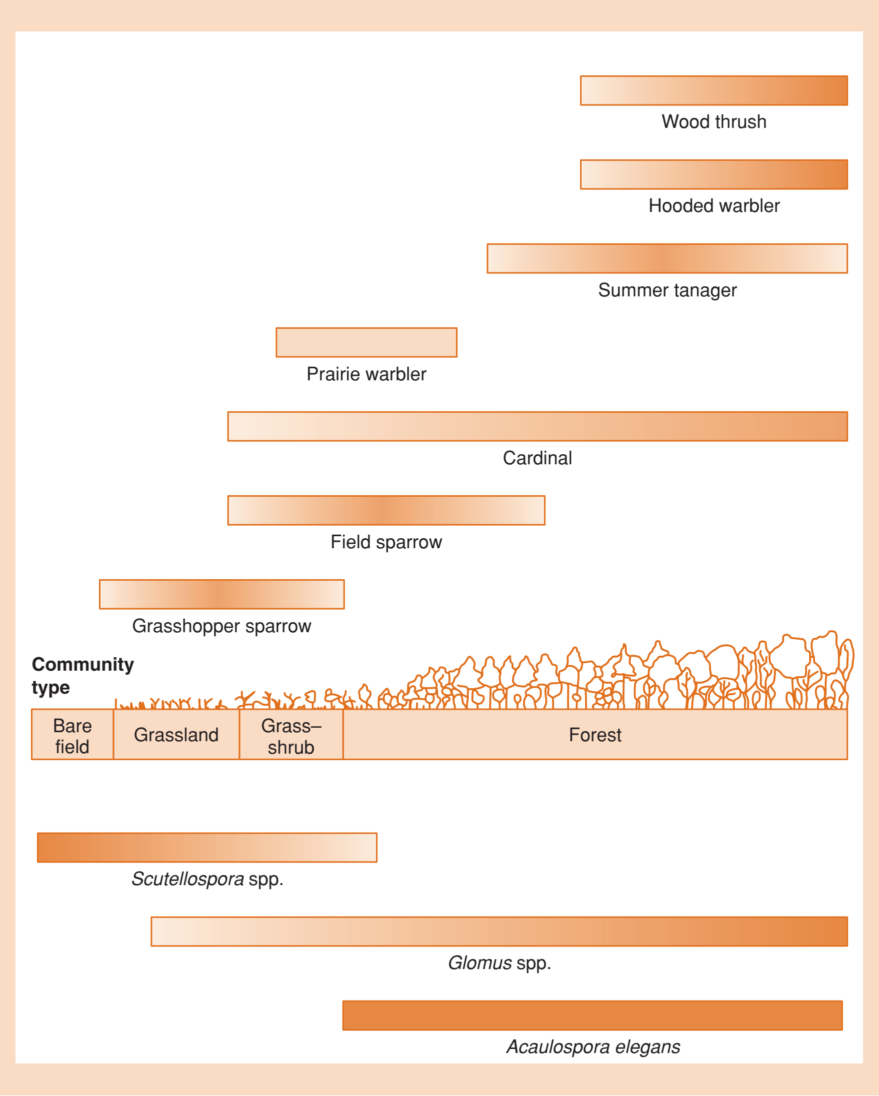


Figure 16.15 Top: bird species distributions along a plant succession gradient in the Piedmont region of Georgia, USA. Differential shading indicates relative abundance of the birds. (After Johnston & Odum, 1956; from Gathreaux, 1978.) Bottom: distributions of vesicular–arbuscular mycorrhizae in the soils associated with an old-field succession in Minnesota. Differential shading indicates relative abundance of spores of species in the genera *Scutellospora, Glomus* and *Acaulospora*. (After Johnson *et al.*, 1991).

16.6.7 Concept of the climax

Do successions come to an end? It is clear that a stable equilibrium will occur if individuals that die are replaced on a one-to-one basis by young of the same species. At a slightly more complex level, Markov models (see Section 16.5) tell us that a stationary species composition should, in theory, occur whenever the replacement probabilities (of one species by itself or by any one of several others) remain constant through time.

The concept of the climax has a long history. One of the earliest students of succession, Frederic Clements (1916), is associated with the idea that a single climax will dominate in any given climatic region, being the endpoint of all successions, whether they happened to start from a sand dune, an abandoned old field or even a pond filling in and progressing towards a terrestrial climax. This *monoclimax* view was challenged by many ecologists, amongst whom Tansley (1939) was prominent. The *polyclimax* school of thought recognized that a local climax may be governed by one factor or a combination of factors: climate, soil conditions, topography, fire and so on. Thus, a single climatic area could easily contain a number of specific climax types. Later still, Whittaker (1953) proposed his climax pattern hypothesis. This conceives a continuity of climax types, varying gradually along environmental gradients and not necessarily separable into discrete climaxes. (This is an extension of Whittaker's approach to gradient analysis of vegetation, discussed in Section 16.3.1.)

In fact, it is very difficult to identify a stable climax community in the field.

climaxes may be approached rapidly – or, so slowly that they are rarely ever reached

Usually, we can do no more than point out that the rate of change of succession slows down to the point where any change is imperceptible to us. In this context, the subtidal rockface succession illustrated in Figure 16.12 is unusual in that convergence to a climax took only a few years. Old-field successions might take 100–500 years to reach a 'climax', but in that time the probabilities of further fires or hurricanes are so high that a process of succession may rarely go to completion. If we bear in mind that forest communities in northern temperate regions, and probably also in the tropics, are still recovering from the last glaciation (see Chapter 1), it is questionable whether the idealized climax vegetation is often reached in nature.

16.7 Communities in a spatiotemporal context: the patch dynamics perspective

the idea of a successional mosaic

A forest, or a rangeland, that appears to have reached a stable community structure when studied on a scale of hectares, will always be a mosaic of miniature successions. Every time a tree falls or a grass tussock dies, an opening is created in which a new succession starts. One of the most seminal papers in the history of ecology was entitled 'Pattern and process in the plant community' (Watt, 1947). Part of the pattern of a community is caused by the dynamic processes of deaths, replacements and microsuccessions that the broad view may conceal. Thus, although we can point to patterns in community composition in space (see Section 16.3) and in time (see Section 16.4), it is often more meaningful to consider space and time together.

disturbance . . . gaps . . . dispersal . . . recruitment

We have already seen that disturbances that open up gaps are common in all kinds of community. The formation of gaps is obviously of considerable significance to sessile or sedentary species that have a requirement for open space, but gaps have also proved to be important for mobile species such as invertebrates on the beds of streams (Matthaei & Townsend, 2000). The patch dynamics concept of communities views the habitat as patchy, with patches being disturbed and recolonized by individuals of various species. Implicit in the patch dynamics view is a critical role for disturbance as a reset mechanism (Pickett & White, 1985). A single patch without migration is, by definition, a closed system, and any extinction caused by disturbance would be final. However, extinction within a patch in an open system is not necessarily the end of the story because of the possibility of reinvasion from other patches.

Fundamental to the patch dynamics perspective is recognition of the importance of migration between habitat patches. This may involve adult individuals, but very often the process of most significance is the dispersal of immature propagules (seeds, spores, larvae) and their recruitment to populations within habitat patches. The order of arrival and relative recruitment levels of individual species may determine or modify the nature and outcome of population interactions in the community (Booth & Brosnan, 1995).

In Section 16.4.1 we identified two fundamentally different kinds of situations within communities: those in which some species are strongly competitively superior are *dominance controlled* (equivalent to succession) and those in which all species have similar competitive abilities are *founder controlled*. Within the patch dynamics framework, the dynamics of these two situations are different and we deal with them in turn.

16.7.1 Dominance-controlled communities

dominance control and succession

In patch dynamics models where some species are competitively superior to others, the effect of the disturbance is to knock the community back to an earlier stage of succession (Figure 16.16). The open space is colonized by one or more of a group of opportunistic, early successional species (p_1, p_2, etc., in Figure 16.16). As time passes, more species invade, often those with poorer powers of dispersal. These eventually reach maturity, dominating mid-succession (m_1, m_2, etc.) and many or all of the pioneer species are driven to extinction. Later still, the community regains the climax stage when the most efficient competitors (c_1, c_2, etc.) oust their neighbors. In this sequence, diversity starts at a low level, increases at the mid-successional stage and usually declines again at the climax. The gap essentially undergoes a minisuccession.

disturbance scale and phasing

Some disturbances are synchronized, or phased, over extensive areas. A forest fire may destroy a large tract of a climax community. The whole area then proceeds through a more or less synchronous succession, with diversity increasing through the early colonization phase and falling again through competitive exclusion as the climax is approached. Other disturbances are much smaller and produce a patchwork of habitats. If these disturbances are unphased, the resulting community comprises a mosaic of patches at different stages of succession. A climax mosaic, produced by unphased disturbances, is much richer in species than an extensive area undisturbed for a very long period and occupied by just one or a few dominant climax species. Towne (2000) monitored the plant species that established in prairie grassland where large ungulates had died (mainly bison, *Bos bison*). Scavengers remove most of the body tissue but copious amounts of body fluids and decomposition products seep into the soil. The flush of nutrients combined with death of the previous vegetation produces a competitor-free, disturbed area where resources are unusually abundant. The patches are also exceptional because the soil has not been disturbed (as it would be after a ploughed field is abandoned or a badger makes a burrow); thus,

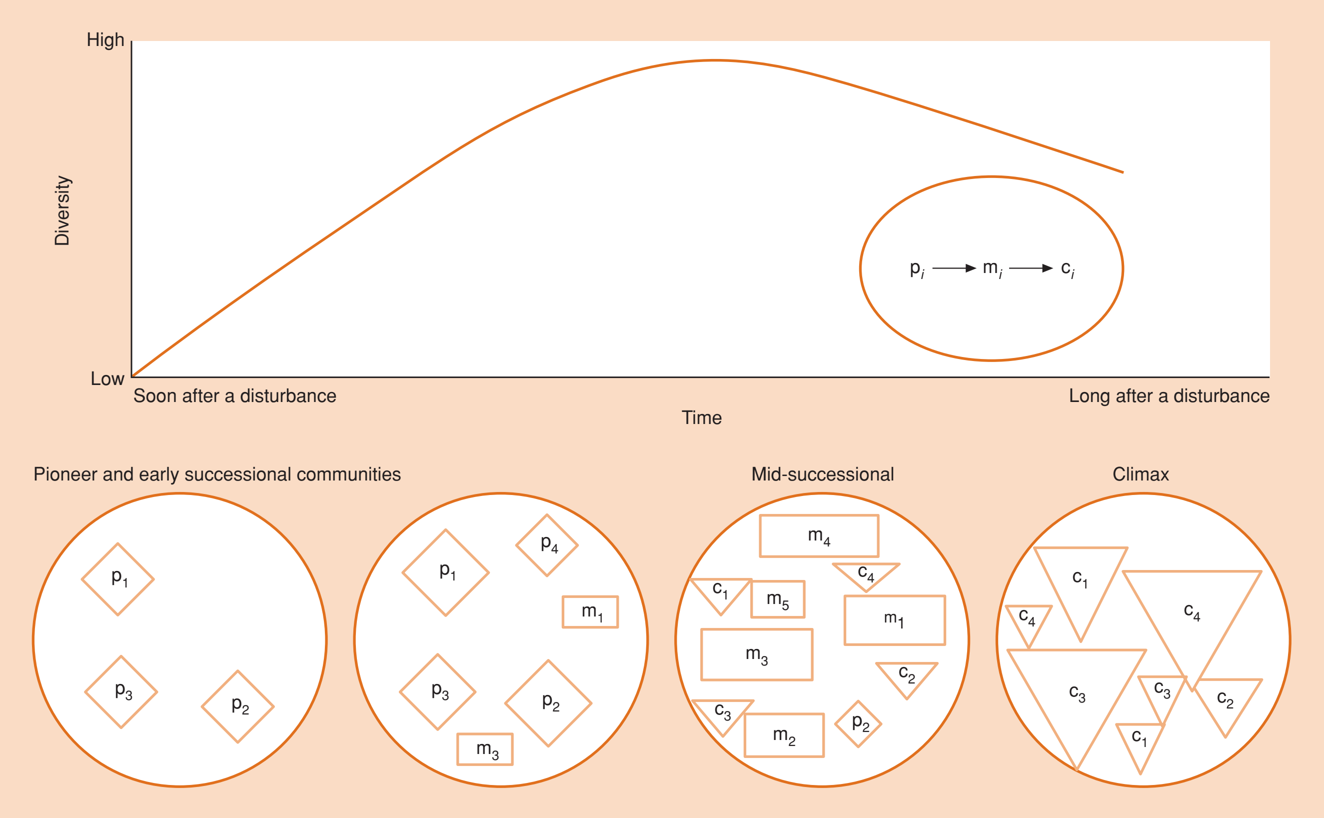


Figure 16.16 Hypothetical minisuccession in a gap. The occupancy of gaps is reasonably predictable. Diversity begins at a low level as a few pioneer (p_i) species arrive; reaches a maximum in mid-succession when a mixture of pioneer, mid-successional (m_i) and climax (c_i) species occur together; and drops again as competitive exclusion by the climax species takes place.

the colonizing plants do not derive from the local seed bank. The unusual nature of the disturbed patches means that many of the pioneer species are rare in the prairie as a whole, and carcass sites contribute to species diversity and community heterogeneity for many years.

16.7.2 Frequency of gap formation

Connell's 'intermediate disturbance hypothesis'

The influence that disturbances have on a community depends strongly on the frequency with which gaps are opened up. In this context, the intermediate disturbance hypothesis (Connell, 1978; see also the earlier account by Horn, 1975) proposes that the highest diversity is maintained at intermediate levels of disturbance. Soon after a severe disturbance, propagules of a few pioneer species arrive in the open space. If further disturbances occur frequently, gaps will not progress beyond the pioneer stage in Figure 16.16, and the diversity of the community as a whole will be low. As the interval between disturbances increases, the diversity will also increase because time is available for the invasion of more species. This is the situation at an intermediate frequency of disturbance. At very low frequencies of disturbance, most of the community for most of the time will reach and remain at the climax, with competitive exclusion having reduced diversity. This is shown diagrammatically in Figure 16.17, which plots the pattern of species richness to be expected as a result of unphased high, intermediate and low frequencies of gap formation, in separate patches and for the community as a whole.

boulders on a rocky shore that vary in disturbability . . .

The influence of the frequency of gap formation was studied in southern California by Sousa (1979a, 1979b), in an intertidal algal community associated with boulders of various sizes. Wave action disturbs small boulders more often than large ones. Using a sequence of photographs, Sousa estimated the probability that a given boulder would be moved during the course of 1 month. A class of mainly small boulders (which required a force of less than 49 Newtons to move them) had a monthly probability of movement of 42%. An intermediate class (which required a force of 50–294 N) had a much smaller monthly probability of movement,

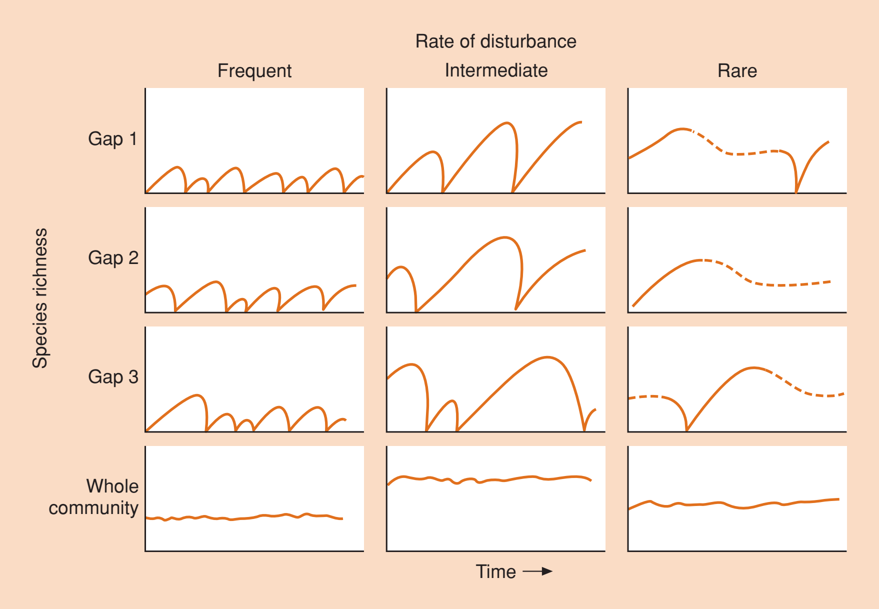


Figure 16.17 Diagrammatic representation of the time course of species richness in three gaps, and in the community as a whole, at three frequencies of disturbance. The disturbance is unphased. Dashed lines indicate the phase of competitive exclusion as the climax is approached.

9%. Finally, the class of mainly large boulders (which required a force >294 N) moved with a probability of only 0.1% per month. The 'disturbability' of the boulders had to be assessed in terms of the force required to move them, rather than simply in terms of top surface area, because some rocks which appeared to be small were actually stable portions of larger, buried boulders, and a few large boulders with irregular shapes moved when a relatively small force was applied. The three classes of boulder (<49, 50–294 and >294 N) can be viewed as patches exposed to a decreasing frequency of disturbance when waves caused by winter storms overturn them.

Species richness increased during early stages of succession through a process of colonization by the pioneer green alga *Ulva* spp. and various other algae, but declined again at the climax because of competitive exclusion by the perennial red alga *Gigartina canaliculata*. It is important to note that the same succession occurred on small boulders that had been artificially made stable. Thus, variations in the communities associated with the surfaces of boulders of different size were not simply an effect of size, but rather of differences in the frequency with which they were disturbed.

. . . provide support for the hypothesis

Communities on unmanipulated boulders in each of the three size/disturbability classes were assessed on four occasions. Table 16.4 shows that the percentage of bare space decreased from small to large boulders, indicating the effects of the greater frequency of disturbance of small boulders. Mean species richness was lowest on the regularly disturbed small boulders. These were dominated most commonly by *Ulva* spp. (and barnacles, *Chthamalus fissus*). The highest levels of species richness were consistently recorded on the intermediate boulder class. Most held mixtures of three to five abundant species from all successional stages. The largest boulders had a lower mean species richness than the intermediate class, although a monoculture was achieved on only a few boulders. *G. canaliculata* covered most of the rock surfaces.

These results offer strong support for the intermediate disturbance hypothesis as far as frequency of appearance of gaps is concerned. However, we must be careful not to lose sight of the fact that this is a highly stochastic process. By chance, some small boulders were not overturned during the period of study. These few were dominated by the climax species *G. canaliculata*. Conversely, two large boulders in the May census had been overturned, and these became dominated by the pioneer *Ulva*. On average, however, species richness and species composition followed the predicted pattern.

This study deals with a single community conveniently composed of identifiable patches (boulders) that become gaps (when overturned by waves) at short, intermediate or long intervals. Recolonization occurs mainly from propagules derived from other patches in the community. Because of the pattern of disturbance, this mixed boulder community is more diverse than would be one with only large boulders.

further support from a study of streams

Disturbances in small streams often take the form of bed movements during periods of high discharge. Because of differences in flow regimes and in the substrates of stream beds, some stream communities are disturbed more frequently and to a larger extent than others. This variation was assessed in 54 stream sites in the Taieri River in New Zealand (Townsend *et al.*, 1997) by recording the frequency at which at least 40% (chosen arbitrarily) of the bed moved and the average percentage that

			Species richness		
Census date	*Boulder class (N)*	*Percentage bare space*	*Mean*	*Standard error*	*Range*
November 1975	< 49	78.0	1.7	0.18	1–4
	50–294	26.5	3.7	0.28	2–7
	> 294	11.4	2.5	0.25	1–6
May 1976	< 49	66.5	1.9	0.19	1–5
	50–294	35.9	4.3	0.34	2–6
	> 294	4.7	3.5	0.26	1–6
October 1976	< 49	67.7	1.9	0.14	1–4
	50–294	32.2	3.4	0.40	2–7
	> 294	14.5	2.3	0.18	1–6
May 1977	< 49	49.9	1.4	0.16	1–4
	50–294	34.2	3.6	0.20	2–5
	> 294	6.1	3.2	0.21	1–5

Table 16.4 Seasonal patterns in bare space and species richness on boulders in each of three classes, categorized according to the force (in Newtons) required to move them. (After Sousa, 1979b.)

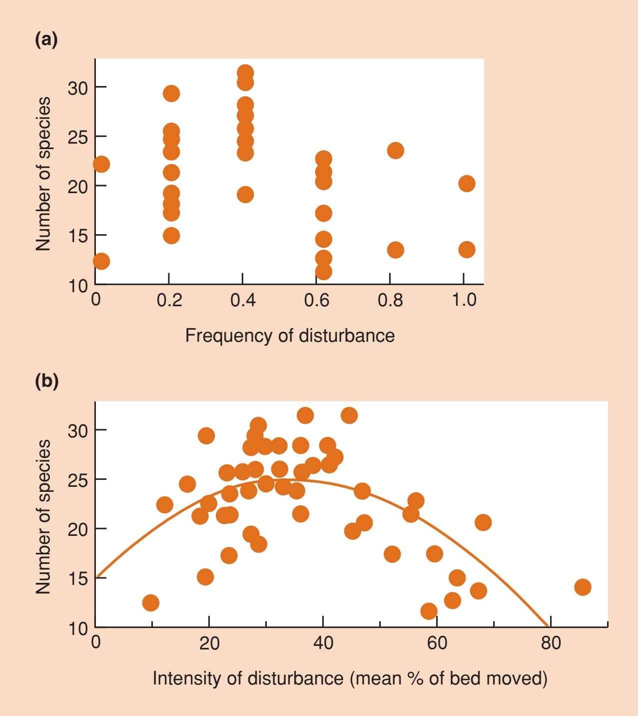


Figure 16.18 Relationship between invertebrate species richness and: (a) frequency of disturbance – assessed as the number of occasions in 1 year when more than 40% of the bed moved (analysis of variance significant at $P < 0.0001$), and (b) intensity of disturbance – average percentage of the bed that moved (polynomial regression fitted, relationship significant at $P < 0.001$) assessed at 54 stream sites in the Taieri River, New Zealand. The patterns are essentially the same; intensity and frequency of disturbance are strongly correlated. (After Townsend *et al*., 1997.)

moved (assessed on five occasions during 1 year, using painted particles of sizes characteristic of the stream bed in question). The pattern of richness of insect species conformed to the intermediate disturbance hypothesis (Figure 16.18). It is likely that low richness at high frequencies and intensities of disturbance reflects the inability of many species to persist in such situations. Whether low richness at low frequencies and intensities of disturbance is due to competitive exclusion, as proposed in the intermediate disturbance hypothesis, remains to be tested.

16.7.3 Formation and filling of gaps

influence of gap size . . .

Gaps of different sizes may influence community structure in different ways because of contrasting mechanisms of recolonization. The centers of very large gaps are most likely to be colonized by species producing propagules that travel relatively great distances. Such mobility is less important in small gaps, since most recolonizing propagules will be produced by adjacent established individuals. The smallest gaps of all may be filled simply by lateral movements of individuals around the periphery.

Intertidal beds of mussels provide excellent opportunities to study the processes of formation and filling-in of gaps. In the absence of disturbance, mussel beds may persist as extensive monocultures. More often, they are an ever-changing mosaic of many species that inhabit gaps formed by the action of waves. Gaps can appear virtually anywhere, and may exist for years as islands in a sea of mussels. The size of these gaps at the time of formation ranges from the dimensions of a single mussel to hundreds of square meters. In general, a mussel or group of mussels becomes infirm or damaged through disease, predation, old age or, most often, the effects of storm waves or battering by logs. Gaps begin to fill as soon as they are formed.

Table 16.5 Measures of area, perimeter and perimeter : area ratio for the experimental gaps created in two experiments on semiexposed shores in southeast Brazil. (From Tanaka & Magalhaes, 2002.)

	Area (cm^2)	*Perimeter (cm)*	*Perimeter : area ratio*
Patch size effects			
Square	25	20	0.8
Square	100	40	0.2
Square	400	80	0.2
Patch shape effects			
Square	100.0	40.0	0.4
Circle	78.5	31.4	0.4
Rectangle	112.5	45.0	0.4
Sector	190.1	78.6	0.4

. . . and gap shape

In their experimental study of mussel beds of *Brachidontes solisianus* and *B. darwinius* in Brazil, Tanaka and Magalhaes (2002) aimed to determine the differential effects of patch size and perimeter : area ratio on the dynamics of succession. In an experiment on one moderately exposed shoreline, they created square gaps with different areas (because of identical shapes, the bigger squares had smaller perimeter : area ratios) (Table 16.5). On a nearby and physically very similar shore, they created patches of four different shapes and chose areas for each that produced identical perimeter : area ratios (Figure 16.19a). Note that a circle has the most perimeter per unit area of any shape. The gap sizes were within the range observed for natural gaps, which did not differ on the two shores (Figure 16.19b).

colonization of gaps . . .
. . . in mussel beds, . . .

Higher densities of the herbivorous limpet *Collisella subrugosa* occurred in the small gaps in the first 6 months after gap formation (Figure 16.19c). Small gaps, compared to medium and large gaps, were also most quickly colonized by lateral migration of the two mussel species, but with *B. darwinius* predominating. The larger gaps had higher densities of the barnacle *Chthamalus bisinuatus* and sheltered more limpets at their edges, while central areas had more *Brachiodontes* recruited from larvae after 6 months (Figure 16.19d). The gaps with identical perimeter : area ratios showed very similar patterns of colonization despite their different sizes, emphasizing that colonization dynamics are mainly determined by distance from adjacent sources of colonists.

The limpet is probably associated with patch edges because here they are less vulnerable to visually hunting predators. The negative relationship between distributions of the limpet and the barnacle may be due to the former dislodging the latter from the substrate. Tanaka and Magalhaes conclude that the mussel *B. darwinius* is a more effective colonist of disturbed patches than *B. solisianus*, and suggest that *B. darwinius* would gradually come to dominate the whole of the shoreline if it were not for occasional massive recruitment events of *B. solisianus*.

. . . in grassland . . .

The pattern of colonization of gaps in mussel beds is repeated in almost every detail in the colonization of gaps in grassland caused by burrowing animals or patches killed by urine. Initially, leaves lean into the gap from plants outside it. Then colonization begins by clonal spread from the edges, and a very small gap may close up quickly. In larger gaps, new colonists may enter as dispersed seed, or germinate from the seed bank in the soil. Over 2–3 years the vegetation begins to acquire the character that it had before the gap was formed.

. . . and in mangrove forest

The gaps produced in forests vary greatly in size. Lightning-induced gaps in mangrove forest in the Dominican Republic, for example, range from 200 to 1600 m^2 or more (Figure 16.20). Lightning almost always kills groups of trees in a 20–30 m circle, and the trees remain as standing dead for several years. In a forest dominated by red mangrove *Rhizophora mangle* and white mangrove *Laguncularia racemosa*, and with some black mangrove *Avicennia germinans*, Sherman *et al.* (2000) compared the performance of the three species in lightning gaps and under forest canopy. Seedling density did not differ in gaps and intact forest, but sapling density and the growth rates of all three species were much higher in the gaps (Table 16.6). However, gap regeneration was dominated by *R. mangle* because its mortality rate was much lower in gaps than was the case for the other species. Sherman *et al.* (2000) note that the peat mat on the forest floor usually collapses after lightning damage, resulting in increased levels of standing water. They suggest that the success of *R. mangle* in gaps is due to their higher tolerance of flooding conditions.

Organisms other than plants can also be overrepresented in gaps. In a study of tropical rainforest in Costa Rica, Levey (1988) found that nectarivorous and frugivorous birds were much more abundant in treefall gaps, reflecting the fact that understory plants in gaps tend to produce more fruit over a longer period than conspecifics fruiting under a closed canopy.

16.7.4 Founder-controlled communities

founder-controlled communities: a competitive lottery not a predictable succession

In the dominance-controlled communities discussed in Section 16.7.1 there was the familiar *r* and *K* selection dichotomy in which colonizing ability and competitive status are inversely related. In founder-controlled communities, on the other hand, all species are both good colonists and essentially equal competitors; thus, within a patch opened by disturbance, a competitive lottery rather than a predictable succession is to

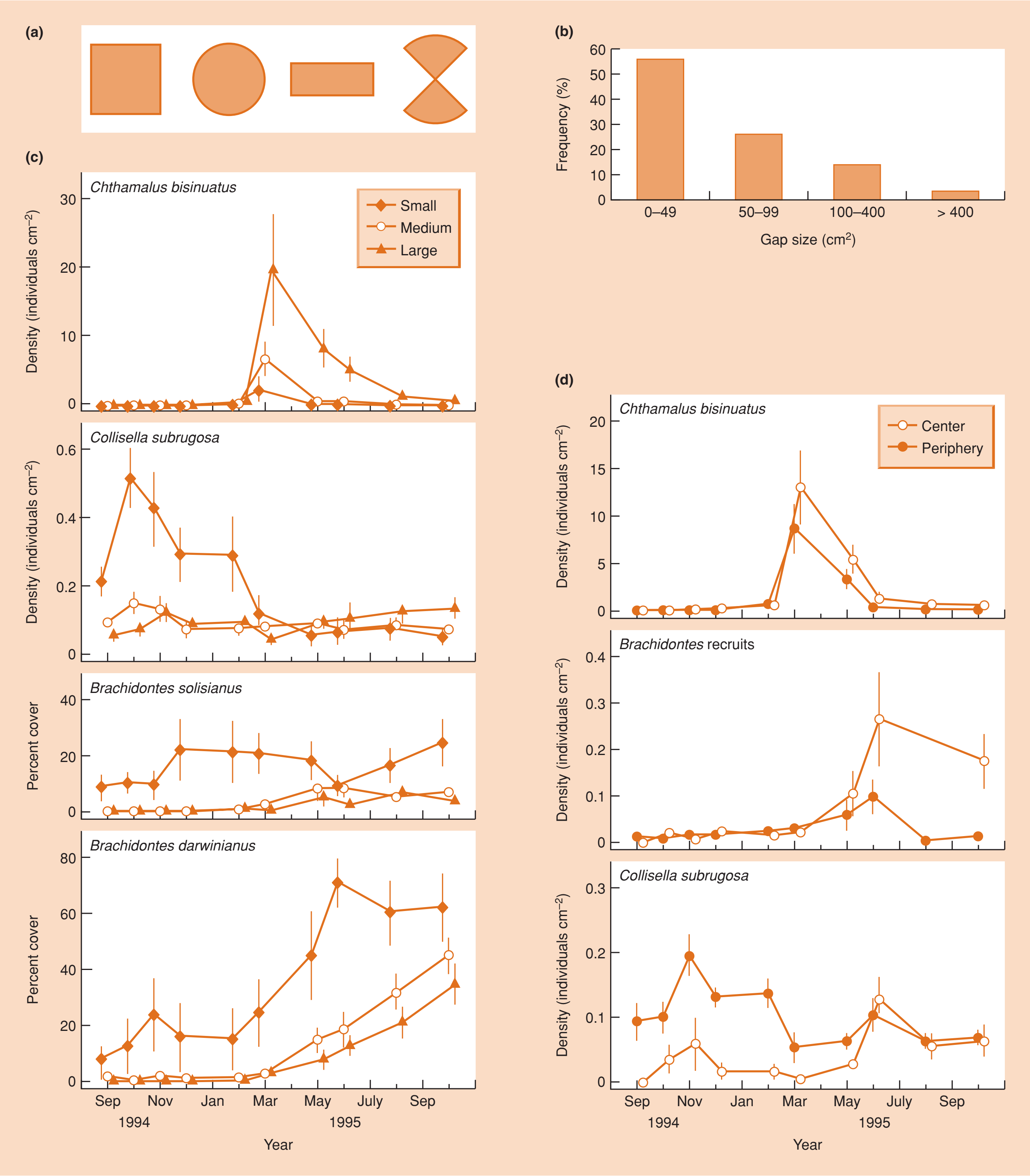


Figure 16.19 (a) The four shapes used in patch-shape experiments: square, circle, rectangle and 'sector' (see Table 16.5). (b) Size distribution of natural gaps in the mussel beds. (c) Mean abundances (±SE) of four colonizing species in experimentally cleared small, medium and large square gaps. (d) Recruitment of three species at the periphery (within 5 cm of the gap edge) and in the center of 400 cm^2 square gaps. (After Tanaka & Magalhaes, 2002.)

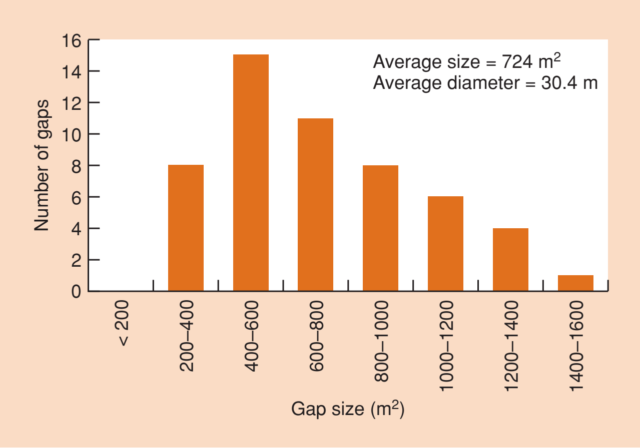


Figure 16.20 Frequency distribution of gaps created by lightning in a tropical mangrove forest in the Dominican Republic. (After Sherman *et al.*, 2000.)

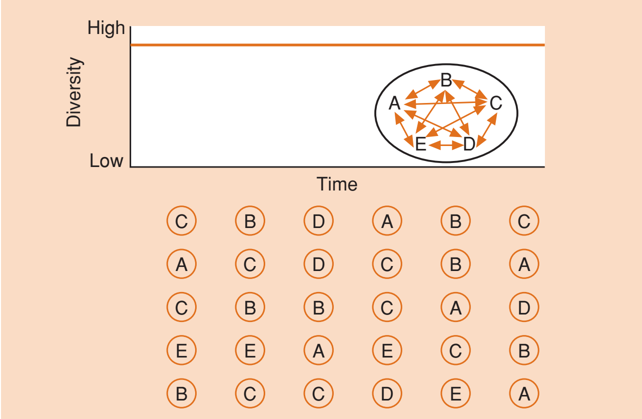


Figure 16.21 Hypothetical competitive lottery: occupancy of gaps which periodically become available. Each of species A–E is equally likely to fill a gap, regardless of the identity of its previous occupant. Species richness remains high and relatively constant.

be expected. If a large number of species are approximately equivalent in their ability to invade gaps, are equally tolerant of the abiotic conditions and can hold the gaps against all comers during their lifetime, then the probability of competitive exclusion may be much reduced in an environment where gaps are appearing continually and randomly. A further condition for coexistence is that the number of young that invade and occupy the gaps should not be consistently greater for parent populations that produce more offspring, otherwise the most productive species would come to monopolize space even in a continuously disturbed environment.

fish coexisting on coral reefs

If these idealized conditions are met, it is possible to envisage how the occupancy of a series of gaps will change through time (Figure 16.21). On each occasion that an organism dies (or is killed) the gap is reopened for invasion. All conceivable replacements are possible and species richness will be maintained at a high level. Some tropical reef communities of fish may conform to this model (Sale, 1977, 1979). They are extremely diverse. For example, the number of species of fish on the Great Barrier Reef ranges from 900 in the south to 1500 in the north, and more than 50 resident species may be recorded on a single patch of reef 3 m in diameter. Only a proportion of this diversity is likely to be attributable to resource partitioning of food and space – indeed, the diets of many of the coexisting species are very similar. In this community, vacant living space seems to be a crucial limiting factor, and it is generated unpredictably in space and time when a resident dies or is killed. The lifestyles of the species match this state of affairs. They breed often, sometimes year-round, and produce numerous clutches of dispersive eggs or larvae. It can be argued that the species compete in a lottery for living space in which larvae are the tickets, and the first arrival at the vacant space wins the site, matures quickly and holds the space for its lifetime.

Three species of herbivorous pomacentrid fish co-occur on the upper slope of Heron Reef, part of the Great Barrier Reef off eastern Australia. Within rubble patches, the available space is occupied by a series of contiguous and usually nonoverlapping territories, each up to 2 m² in area, held by individuals of *Eupomacentrus apicalis*, *Plectroglyphidodon lacrymatus* and *Pomacentrus wardi*. Individuals hold territories throughout their juvenile and adult life and defend them against a broad range of chiefly herbivorous species, including conspecifics. There seems to be no particular tendency for space initially held by one species to be

Table 16.6 Initial size, and growth and mortality rates over a 1-year period of saplings of three mangrove species in lightning-induced gaps and under intact forest canopy. (After Sherman *et al.*, 2000.)

	Initial sapling diameter (cm ± SE)		*Growth rate–diameter increment (cm ± SE)*		*Mortality (%)*	
	Gaps	*Canopy*	*Gaps*	*Canopy*	*Gaps*	*Canopy*
Rhizophora mangle	1.9 ± 0.06	2.3 ± 0.06	0.58 ± 0.03	0.09 ± 0.01	9	16
Laguncularia racemosa	1.7 ± 0.11	1.8 ± 0.84	0.46 ± 0.04	0.11 ± 0.06	32	40
Avicennia germinans	1.3 ± 0.25	1.7 ± 0.45	0.51 ± 0.04	–	56	88

Table 16.7 Numbers of individuals of each species observed occupying sites, or parts of sites, that had been vacated during the immediately prior interperiod between censuses through the loss of residents of each species. The sites vacated through loss of 120 residents have been reoccupied by 131 fish; the species of the new occupant is not dependent on the species of the previous resident.

	Reoccupied by:		
Resident lost	*E. apicalis*	*P. lacrymatus*	*P. wardi*
Eupomacentrus apicalis	9	3	19
Plectroglyphidodon lacrymatus	12	5	9
Pomacentrus wardi	27	18	29

taken up, following mortality, by the same species. Nor is any successional sequence of ownership evident (Table 16.7). *P. wardi* both recruited and lost individuals at a higher rate than the other two species, but all three species appear to have recruited at a sufficient level to balance their rates of loss and maintain a resident population of breeding individuals.

plants in grassland or forest

Thus, the maintenance of high reef diversity depends, at least in part, on the unpredictability of the supply of living space; and as long as all species win some of the time and in some places, they will continue to put larvae into the plankton, and hence, into the lottery for new sites. An analogous situation has been postulated for the highly diverse chalk grasslands of Great Britain (Grubb, 1977) and even for trees in temperate and tropical forest gaps (Busing & Brokaw, 2002). Any small gap that appears is rapidly exploited, by a seed in grassland and very often by a sapling in a forest gap. In these cases, the tickets in the lottery are saplings or seeds (either in the act of dispersal or as components of a persistent seed bank in the soil). Which seeds or saplings develop to established plants, and therefore which species comes to occupy the gap, may depend on a strong random element since many species overlap in their requirements for successful growth. The successful plant rapidly establishes itself and retains the patch for its lifetime, in a similar way to the reef fish described above.

16.8 Conclusions: the need for a landscape perspective

founder and dominance control as a continuum of possibilities

The lottery hypothesis and the notion of the founder-controlled community were important steps in the development of our understanding of the range of community dynamics that can occur. However, these should be viewed not as hard and fast rules to which some communities are subject, but rather as extremes on a continuum from dominance to founder control. Real communities may be closer to one or other end of this continuum, but in reality component species or component patches may be dominance controlled or founder controlled within the same community. Syms and Jones (2000), for example, acknowledge that more than half of within-reef variation in fish species composition in their study of patch reefs in the Great Barrier Reef was attributable to unexplained, and thus possibly stochastic, factors such as those emphasized in the lottery hypothesis. But a significant proportion of variation could be explained by specific habitat requirements of the constituent species.

importance of a 'landscape ecology' perspective

More generally, no community is truly the homogeneous, temporally invariant system described by simple Lotka–Volterra mathematics and exemplified by laboratory microcosms, although some are less variable than others. In most real communities, population dynamics will be spatially distributed and temporal variation will be present. In a closed system, composed of a single patch, species extinctions can occur for two very different reasons: (i) as a result of biotic instability caused by competitive exclusion, overexploitation and other strongly destabilizing species interactions; or (ii) as a result of environmental instability caused by unpredictable disturbances and changes in conditions. By integrating unstable patches of either of these types into the open system of a larger landscape (consisting of many patches out of phase with each other), persistent species-rich communities can result (DeAngelis & Waterhouse, 1987). This is the principal message to emerge from the patch dynamics perspective, and its larger scale counterpart, 'landscape ecology' (Wiens *et al.*, 1993), stressing the importance of the spatial scale at which we view communities and the open nature of most of them. Note the strong link between the patch dynamics view of community organization and metapopulation theory, which deals with the effects on the dynamics of populations of dividing them into fragments (see Section 6.9). In a model that combines extinction–colonization dynamics (the metapopulation approach) with the dynamics of patch succession, Amarasekare and Possingham (2001) show that persistence of a species in the landscape depends: (i) on the net rate at which suitable patches arise relative to the species' colonization ability; as well as (ii) the longevity of the dormant stages (e.g. seed bank) relative to disturbance frequency.

multiple classes of disturbance . . .

Future development of ideas about patch dynamics is likely to concern the consequences of multiple classes of disturbance. Steinauer and Collins (2001) have made a start by showing that disturbances caused by urine deposition and grazing by bison (*Bos bison*) interact with each other. The abundance of four common grass species, and of all of them combined, increased on urine patches in ungrazed prairie grassland. However, the abundance of the grass *Andropogon gerardii*, and all grasses combined, decreased on urine patches in grazed prairie. The changed dynamics reflect the fact that bison preferentially graze on urine patches. In addition, grazed areas initiated on urine patches tend to expand well beyond the area of urine deposition, increasing the size and severity of disturbance by grazing.

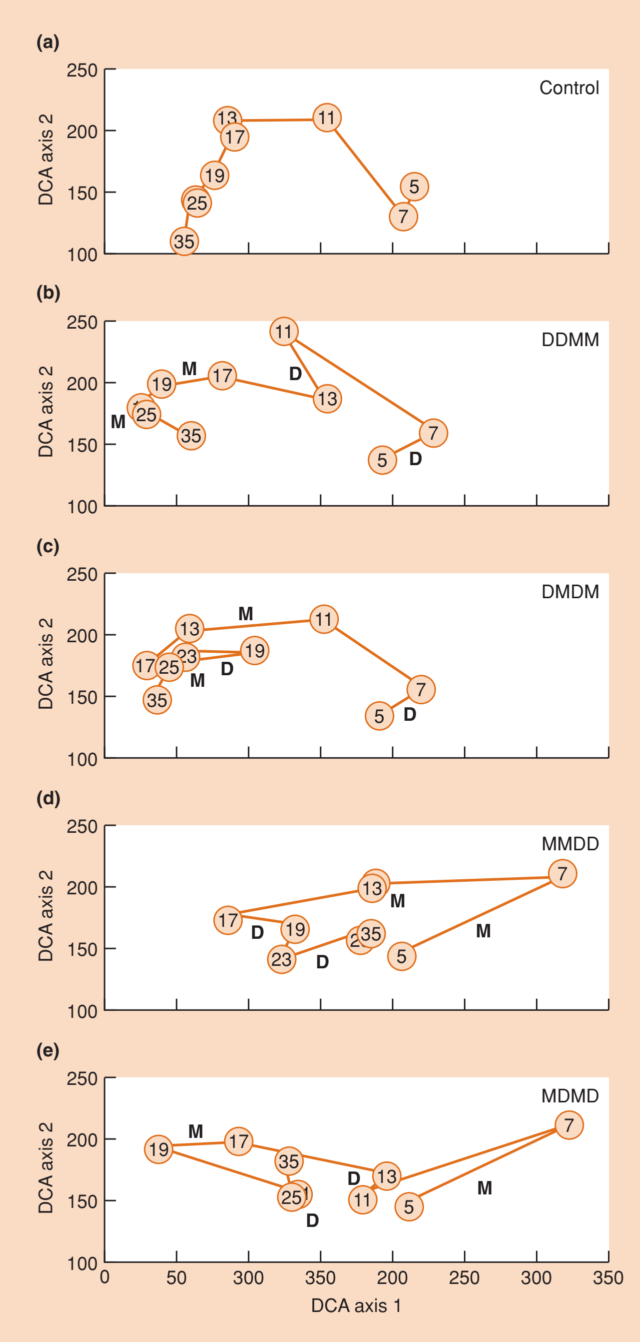


Figure 16.22 (*left*) Temporal changes in species composition and relative abundance of microcosms composed of a specific mix of protists and metazoans. The change is expressed in ordination plots based on a procedure called detrended correspondence analysis (DCA). (Recall that ordination is a mathematical treatment that allows communities to be organized on a graph so that those that are most similar in species composition and relative abundance appear closest together, whilst communities that differ greatly in the relative importance of a similar set of species, or that possess quite different species, appear far apart.) Data points are the mean ordination scores on different days in the experiment (from day 5 to day 35). The letter D indicates periods of drought disturbance, and the letter M, mosquito disturbance. (a–e) The results of the control and disturbances imposed in different sequences. (After Fukami, 2001.)

. . . may interact to determine community patterns

Finally, just as we can readily see how community dynamics may vary according to the order in which colonizing species happen to appear after a disturbance, it is equally the case that the order in which different kinds of disturbance occur may affect the outcome. Fukami (2001) addressed this issue by introducing two classes of disturbance (drought or the addition of predatory mosquito larvae) in various sequences to laboratory micrososms of protists and small metazoans (analogous to natural communities in water-filled bamboo stumps). Different disturbance sequences drove the microcosms into different successional trajectories, sometimes leading to divergence in final community composition (both in terms of species richness and relative abundance of the component species). This is illustrated graphically by ordination diagrams (see Section 16.3.2) that show the sequence of positions of communities in the same ordination space through experiments where disturbances were imposed in different sequences (Figure 16.22). It may often be the case that a knowledge of the disturbance history will be needed to predict the responses of communities to disturbances imposed in the future (such as global climate change).

Summary

The community is an assemblage of species populations that occur together in space and time. Community ecology seeks to understand the manner in which groupings of species are distributed in nature, and how they are influenced by their abiotic environment and by species interactions.

We begin by explaining how the structure of communities can be measured and described, in terms of species composition, species richness, diversity, equitability (evenness) and rank–abundance diagrams.

The assessment of community patterns in space has progressed from subjective 'gradient analysis' to objective mathematical approaches ('classification' and 'ordination') that permit relationships between community composition and abiotic factors to be systematically explored. We note that most communities are not delimited by sharp boundaries, where one group of species is abruptly replaced by another. Moreover a given species that occurs in one predictable association is also quite likely to occur with another group of species under different conditions elsewhere.

Just as the relative importance of species varies in space, so their patterns of abundance may change with time. A particular species can occur where it is capable of reaching a location, appropriate conditions and resources exist, and competitors, predators and parasites do not preclude it. A temporal sequence in the appearance and disappearance of species therefore requires that conditions, resources and/or the influence of enemies themselves vary with time. We emphasize and explain patterns of community change that follow a disturbance. Sometimes these patterns are predictable (succession; dominance control), in other cases highly stochastic (founder control).

Although we can discern and often explain patterns in community composition in space and in time, it is often more meaningful to consider space and time together. The patch dynamics concept of communities views the landscape as patchy, with patches being disturbed and recolonized by individuals of various species. Implicit in this view are critical roles for disturbance as a reset mechanism, and of migration between habitat patches. The community dynamics of patchy landscapes are strongly influenced by the frequency of gap formation and the sizes and shapes of these gaps in relation to the colonization and competitive properties of the species concerned.

Chapter 17

The Flux of Energy through Ecosystems

17.1 Introduction

All biological entities require matter for their construction and energy for their activities. This is true not only for individual organisms, but also for the populations and communities that they form in nature. The intrinsic importance of fluxes of energy (this chapter) and of matter (see Chapter 18) means that community processes are particularly strongly linked with the abiotic environment. The term *ecosystem* is used to denote the biological community *together with* the abiotic environment in which it is set. Thus, ecosystems normally include primary producers, decomposers and detritivores, a pool of dead organic matter, herbivores, carnivores and parasites *plus* the physicochemical environment that provides the living conditions and acts both as a source and a sink for energy and matter. Thus, as is the case with all chapters in Part 3 of this book, our treatment calls upon knowledge of individual organisms in relation to conditions and resources (Part 1) together with the diverse interactions that populations have with one another (Part 2).

Lindemann laid the foundations of ecological energetics

A classic paper by Lindemann (1942) laid the foundations of a science of ecological energetics. He attempted to quantify the concept of food chains and food webs by considering the efficiency of transfer between trophic levels – from incident radiation received by a community through its capture by green plants in photosynthesis to its subsequent use by herbivores, carnivores and decomposers. Lindemann's paper was a major catalyst for the International Biological Programme (IBP), which, with a view to human welfare, aimed to understand the biological basis of productivity of areas of land, fresh waters and the seas (Worthington, 1975). The IBP provided the first occasion on which biologists throughout the world were challenged to work together towards a common end. More recently, a further pressing issue has again galvanized the community of ecologists into action. Deforestation, the burning of fossil fuels and other pervasive human influences are causing dramatic changes to global climate and atmospheric composition, and can be expected in turn to influence patterns of productivity on a global scale. Much of the current work on productivity has a prime objective of providing the basis for predicting the effects of changes in climate, atmospheric composition and land use on terrestrial and aquatic ecosystems (aspects that will be dealt with in Chapter 22).

progressive improvements in technology to assess productivity

The decades since Lindemann's classic work have seen a progressive improvement in technology to assess productivity. Early calculations in terrestrial ecosystems involved sequential measurements of biomass of plants (usually just the aboveground parts) and estimates of energy transfer efficiency between trophic levels. In aquatic ecosystems, production estimates relied on changes in the concentrations of oxygen or carbon dioxide measured in experimental enclosures. Increasing sophistication in the measurement, *in situ*, of chlorophyll concentrations and of the gases involved in photosynthesis, coupled with the development of satellite remote-sensing techniques, now permit the extrapolation of local results to the global scale (Field *et al.*, 1998). Thus, satellite sensors can measure vegetation cover on land and chlorophyll concentrations in the sea, from which rates of light absorption are calculated and, based on our understanding of photosynthesis, these are converted to estimates of productivity (Geider *et al.*, 2001).

some definitions: standing crop and biomass, . . .

Before proceeding further it is necessary to define some new terms. The bodies of the living organisms within a unit area constitute a *standing crop* of biomass. By *biomass* we mean the mass of organisms per unit area of ground (or per unit area or unit volume of water) and this is usually expressed in units of energy (e.g. J m^{-2}) or dry organic matter (e.g. t ha^{-1}) or carbon (e.g. g C m^{-2}). The great bulk of the biomass in communities is almost always formed by plants, which are the primary producers of biomass because of their almost unique ability to fix carbon in photosynthesis. (We have to say 'almost unique' because bacterial photosynthesis and chemosynthesis may also contribute to forming new biomass.) Biomass includes the whole bodies of the organisms even though parts of them may be dead. This needs to be borne in mind, particularly when considering woodland and forest communities in which the bulk of the biomass is dead heartwood and bark. The living fraction of biomass represents active capital capable of generating interest in the form of new growth, whereas the dead fraction is incapable of new growth. In practice we include in biomass all those parts, living or dead, which are attached to the living organism. They cease to be biomass when they fall off and become litter, humus or peat.

. . . primary and secondary productivity, autotrophic respiration, . . .

The *primary productivity* of a community is the rate at which biomass is produced per unit area by plants, the primary producers. It can be expressed either in units of energy (e.g. J m^{-2} day^{-1}) or dry organic matter (e.g. kg ha^{-1} $year^{-1}$) or carbon (e.g. g C m^{-2} $year^{-1}$). The total fixation of energy by photosynthesis is referred to as *gross primary productivity* (GPP). A proportion of this is respired away by the plants (autotrophs) and is lost from the community as respiratory heat (RA – *autotrophic respiration*). The difference between GPP and RA is known as *net primary productivity* (NPP) and represents the actual rate of production of new biomass that is available for consumption by heterotrophic organisms (bacteria, fungi and animals). The rate of production of biomass by heterotrophs is called *secondary productivity*.

. . . net ecosystem productivity, and heterotrophic and ecosystem respiration

Another way to view energy flux in ecosystems involves the concept of *net ecosystem productivity* (NEP, using the same units as GPP or NPP). This acknowledges that the carbon fixed in GPP can leave the system as inorganic carbon (usually carbon dioxide) via either autotrophic respiration (RA) or, after consumption by heterotrophs, via *heterotrophic respiration* (RH)—the latter consisting of respiration by bacteria, fungi and animals. Total *ecosystem respiration* (RE) is the sum of RA and RH. NEP then is equal to GPP – RE. When GPP exceeds RE, the ecosystem is fixing carbon faster than it is being released and thus acts as a carbon sink. When RE exceeds GPP, carbon is being released faster than it is fixed and the ecosystem is a net carbon source. That the rate of ecosystem respiration can exceed GPP may seem paradoxical. However, it is important to note that an ecosystem can receive organic matter from sources other than its own photosynthesis – via the import of dead organic matter that has been produced elsewhere. Organic matter produced by photosynthesis within an ecosystem's boundaries is known as *autochthonous*, whereas that imported from elsewhere is called *allochthonous*.

In what follows we deal first with large-scale patterns in primary productivity (Section 17.2) before considering the factors that limit productivity in terrestrial (Section 17.3) and aquatic (Section 17.4) settings. We then turn to the fate of primary productivity and consider the flux of energy through food webs (Section 17.5), placing particular emphasis on the relative importance of grazer and decomposer systems (we return to food webs and their detailed population interactions in Chapter 20). We finally turn to seasonal and longer term variations in energy flux through ecosystems.

17.2 Patterns in primary productivity

primary productivity depends on, but is not solely determined by, solar radiation

The net primary production of the planet is estimated to be about 105 petagrams of carbon per year (1 Pg = 10^{15} g) (Geider *et al.*, 2001). Of this, 56.4 Pg C $year^{-1}$ is produced in terrestrial ecosystems and 48.3 Pg C $year^{-1}$ in aquatic ecosystems (Table 17.1). Thus, although oceans

Marine	*NPP*	*Terrestrial*	*NPP*
Tropical and subtropical oceans	13.0	Tropical rainforests	17.8
Temperate oceans	16.3	Broadleaf deciduous forests	1.5
Polar oceans	6.4	Mixed broad/needleleaf forests	3.1
Coastal	10.7	Needleleaf evergreen forests	3.1
Salt marsh/estuaries/seaweed	1.2	Needleleaf deciduous forests	1.4
Coral reefs	0.7	Savannas	16.8
		Perennial grasslands	2.4
		Broadleaf shrubs with bare soil	1.0
		Tundra	0.8
		Desert	0.5
		Cultivation	8.0
Total	48.3	Total	56.4

Table 17.1 Net primary production (NPP) per year for major biomes and for the planet in total (in units of petragrams of C). (From Geider *et al.*, 2001.)

cover about two-thirds of the world's surface, they account for less than half of its production. On the land, tropical rainforests and savannas account between them for about 60% of terrestrial NPP, reflecting the large areas covered by these biomes and their high levels of productivity. All biological activity is ultimately dependent on received solar radiation but solar radiation alone does not determine primary productivity. In very broad terms, the fit between solar radiation and productivity is far from perfect because incident radiation can be captured efficiently only when water and nutrients are available and when temperatures are in the range suitable for plant growth. Many areas of land receive abundant radiation but lack adequate water, and most areas of the oceans are deficient in mineral nutrients.

17.2.1 Latitudinal trends in productivity

the productivity of forests, grasslands and lakes follows a latitudinal pattern

In the forest biomes of the world a general latitudinal trend of increasing productivity can be seen from boreal, through temperate, to tropical conditions (Table 17.2). However, there is also considerable variation, much of it due to differences in water availability, local topography and associated variations in microclimate. The same latitudinal trend (and local variations) exists in the above-ground productivity of grassland communities (Figure 17.1). Note the considerable differences in the relative importance of above-ground and below-ground productivity in the different grassland biomes. It is technically difficult to estimate below-ground productivity and early reports of NPP often ignored or underestimated the true values. As far as aquatic communities are concerned, a latitudinal trend is clear in lakes (Brylinski & Mann, 1973) but not in the oceans, where productivity may more often be limited by a shortage of nutrients – very high productivity occurs in marine communities where there are upwellings of nutrient-rich waters, even at high latitudes and low temperatures.

Table 17.2 Gross primary productivity (GPP) of forests at various latitudes in Europe and North and South America, estimated as the sum of net ecosystem productivity and ecosystem respiration (calculated from CO_2 fluxes measured in the forest canopies – only one estimate for tropical forest was included by the reviewers). (From data in Falge *et al.*, 2002.)

Forest type	*Range of GPP estimates* (*g C* m^{-2} *year*$^{-1}$)	*Mean of estimates* (*g C* m^{-2} *year*$^{-1}$)
Tropical rainforest	3249	3249
Temperate deciduous	1122–1507	1327
Temperate coniferous	992–1924	1499
Cold temperate deciduous	903–1165	1034
Boreal coniferous	723–1691	1019

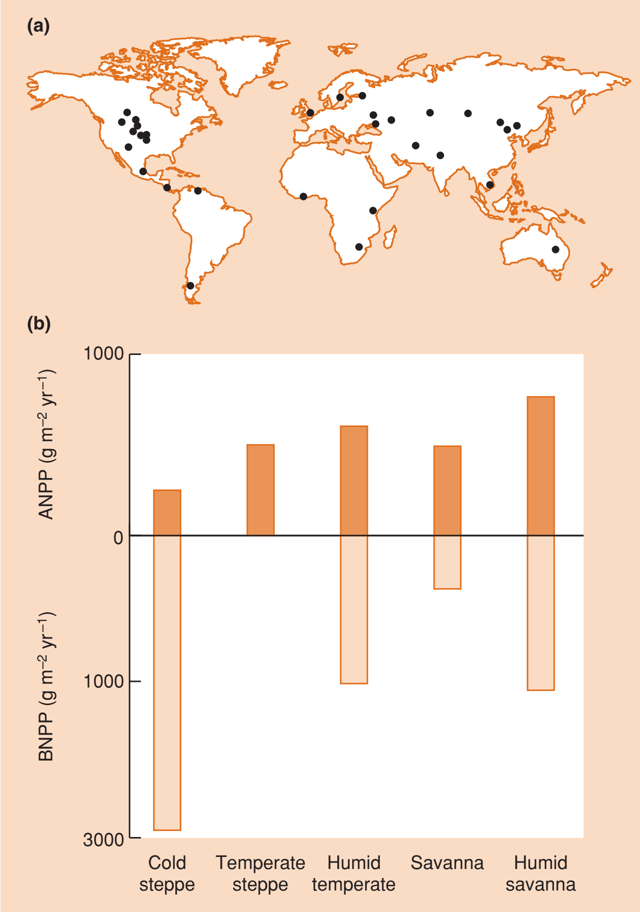


Figure 17.1 (a) The location of 31 grassland study sites included in this analysis. (b) Above-ground net primary productivity (ANPP) and below-ground net primary productivity (BNPP) for five categories of grassland biomes (BNPP not available for temperate steppe). The values in each case are averages for 4–8 grassland studies. The technique involved summing increments in the biomass of live plants, standing dead matter and litter between successive samples in the study period (average 6 years). (From Scurlock *et al.*, 2002.)

The overall trends with latitude suggest that radiation (a resource) and temperature (a condition) may often limit the productivity of communities. But other factors frequently constrain productivity within even narrower limits.

17.2.2 Seasonal and annual trends in primary productivity

productivity shows considerable temporal variation

The large ranges in productivity in Table 17.2 and the wide confidence intervals in Figure 17.1 emphasize the

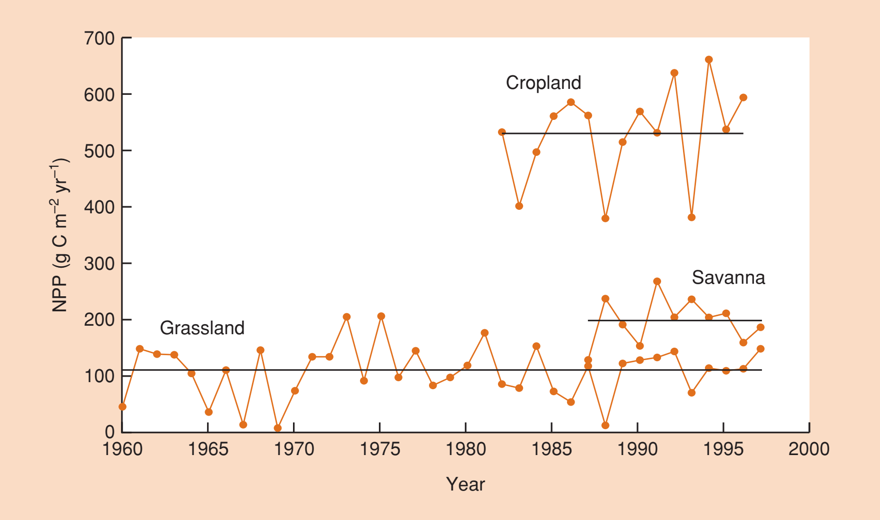


Figure 17.2 Interannual variation in net primary productivity (NPP) in a grassland in Queensland, Australia (above-ground NPP), a cropland in Iowa, USA (total above- and below-ground NPP) and a tropical savanna in Senegal (above-ground NPP). Black horizontal lines show the mean NPP for the whole study period. (After Zheng *et al.*, 2003.)

considerable variation that exists within a given class of ecosystems. It is important to note also that productivity varies from year to year in a single location (Knapp & Smith, 2001). This is illustrated for a temperate cropland, a tropical grassland and a tropical savanna in Figure 17.2. Such annual fluctuations no doubt reflect year-to-year variation in cloudless days, temperature and rainfall. At a smaller temporal scale, productivity reflects seasonal variations in conditions, particularly in relation to the consequences of temperature for the length of the growing season. For example, the period when daily GPP is high persists for longer in temperate than in boreal situations (Figure 17.3). Moreover, the growing season is more extended but the amplitude of seasonal change is smaller in evergreen coniferous forests than in their deciduous counterparts (where the growing season is curtailed by the shedding of leaves in the fall).

17.2.3 Autochthonous and allochthonous production

autochthonous and allochthonous production . . .

All biotic communities depend on a supply of energy for their activities. In most terrestrial systems this is contributed *in situ* by the photosynthesis of green plants – this is autochthonous production. Exceptions exist, however, particularly where colonial animals deposit feces derived from food consumed at a distance from the colony (e.g. bat colonies in caves, seabirds on coastland) – guano is an example of allochthonous organic matter (dead organic material formed outside the ecosystem).

. . . vary in systematic ways in lakes, rivers and estuaries

In aquatic communities, the autochthonous input is provided by the photosynthesis of large plants and attached algae in shallow waters (littoral zone) and by microscopic phytoplankton in the open water. However, a substantial proportion of the organic matter in aquatic communities comes from allochthonous material that arrives in rivers, via groundwater or is blown in by the wind. The relative importance of the two autochthonous sources (littoral and planktonic) and the allochthonous source of organic material in an aquatic system depends on the dimensions of the body of water and the types of terrestrial community that deposit organic material into it.

A small stream running through a wooded catchment derives most of its energy input from litter shed by surrounding vegetation (Figure 17.4). Shading from the trees prevents any significant growth of planktonic or attached algae or aquatic higher plants. As the stream widens further downstream, shading by trees is restricted to the margins and autochthonous primary production increases. Still further downstream, in deeper and more turbid waters, rooted higher plants contribute much less, and the role of the microscopic phytoplankton becomes more important. Where large river channels are characterized by a flood plain, with associated oxbow lakes, swamps and marshes, allochthonous dissolved and particulate organic may be carried to the river channel from its flood plain during episodes of flooding (Junk *et al.*, 1989; Townsend 1996).

The sequence from small, shallow lakes to large, deep ones shares some of the characteristics of the river continuum just discussed (Figure 17.5). A small lake is likely to derive quite a large proportion of its energy from the land because its periphery is large in relation to its area. Small lakes are also usually shallow, so internal littoral production is more important than that by phytoplankton. In contrast, a large, deep lake will derive only limited organic matter from outside (small periphery relative to lake surface area) and littoral production, limited to the shallow margins, may also be low. The organic inputs to the community may then be due almost entirely to photosynthesis by the phytoplankton.

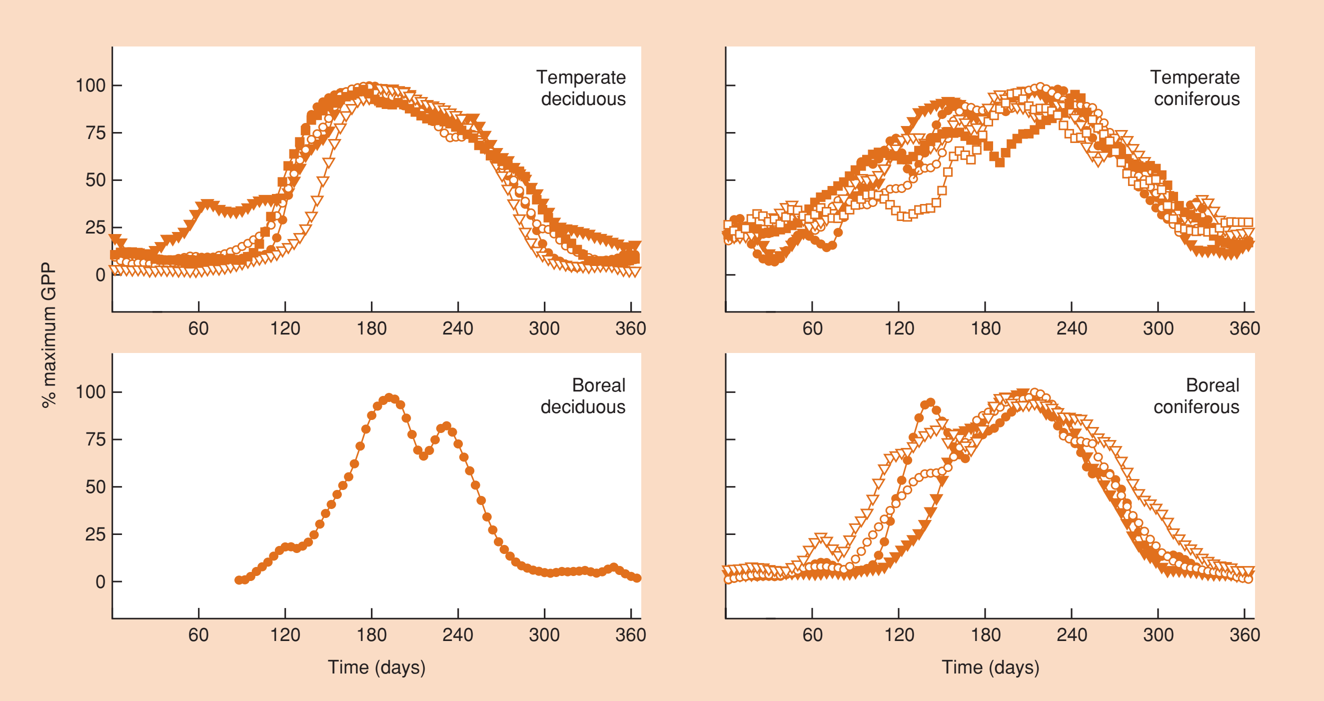


Figure 17.3 Seasonal development of maximum daily gross primary productivity (GPP) for deciduous and coniferous forests in temperate (Europe and North America) and boreal locations (Canada, Scandinavia and Iceland). The different symbols in each panel relate to different forests. Daily GPP is expressed as the percentage of the maximum achieved in each forest during 365 days of the year. (After Falge *et al*., 2002.)

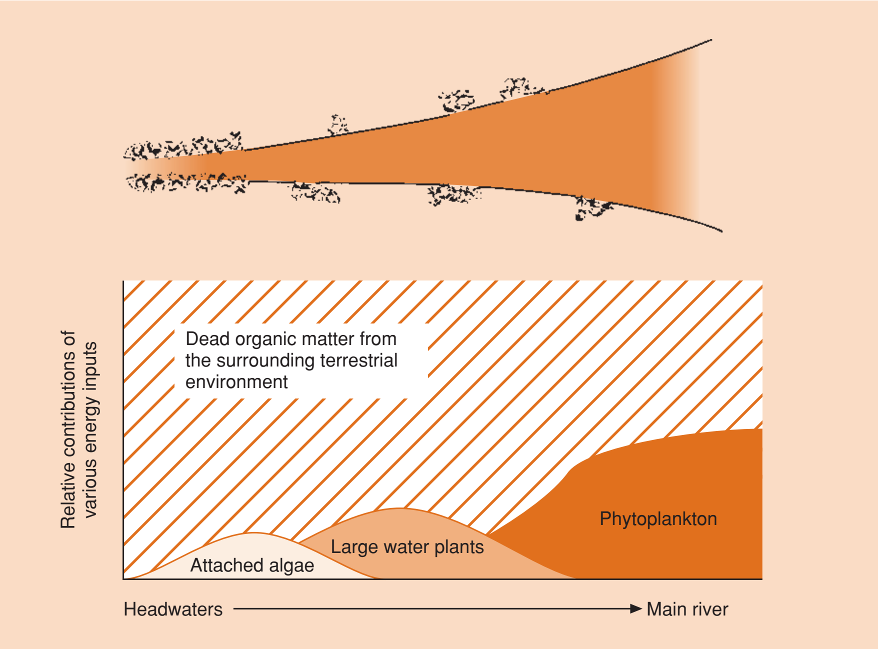


Figure 17.4 Longitudinal variation in the nature of the energy base in stream communities.

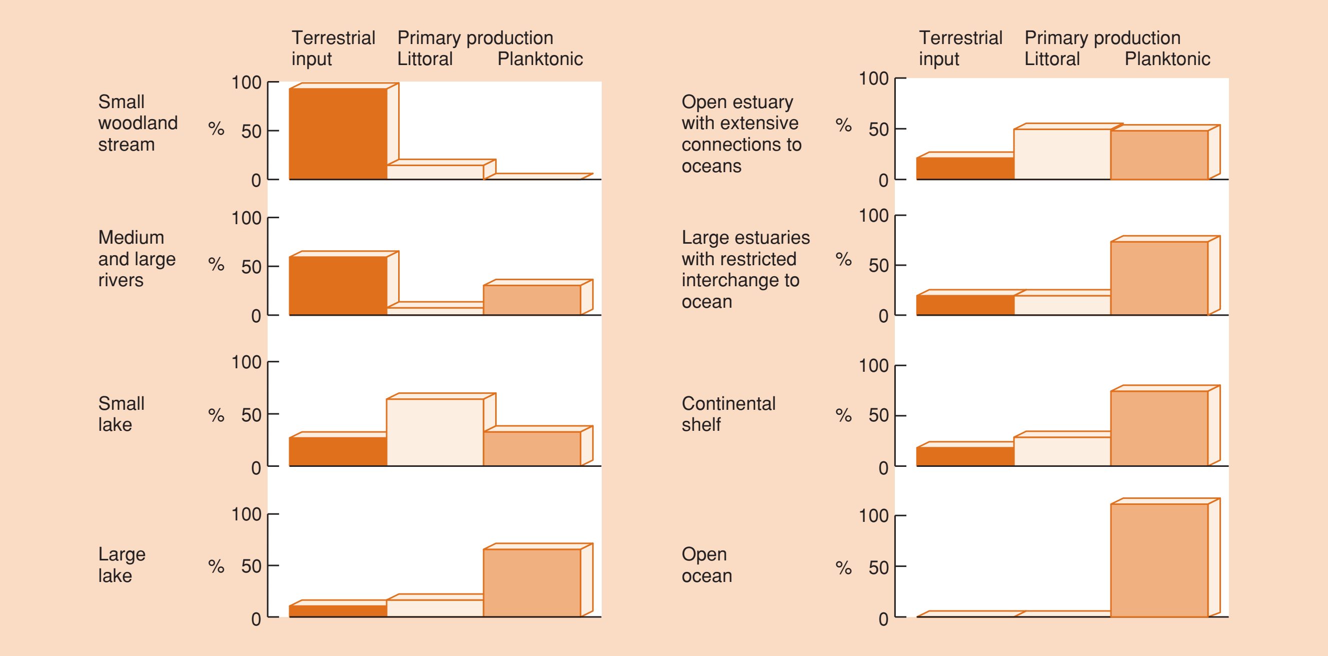


Figure 17.5 Variation in the importance of terrestrial input of organic matter and littoral and planktonic primary production in contrasting aquatic communities.

Estuaries are often highly productive systems, receiving allochthonous material and a rich supply of nutrients from the rivers that feed them. The most important autochthonous contribution to their energy base varies. In large estuarine basins, with restricted interchange with the open ocean and with small marsh peripheries relative to basin area, phytoplankton tend to dominate. By contrast, seaweeds dominate in some open basins with extensive connections to the sea. In turn, continental shelf communities derive a proportion of their energy from terrestrial sources (particularly via estuaries) and their shallowness often provides for significant production by littoral seaweed communities. Indeed, some of the most productive systems of all are to be found among seaweed beds and reefs.

Finally, the open ocean can be described in one sense as the largest, deepest 'lake' of all. The input of organic material from terrestrial communities is negligible, and the great depth precludes photosynthesis in the darkness of the sea bed. The phytoplankton are then all-important as primary producers.

17.2.4 Variations in the relationship of productivity to biomass

NPP : B ratios are very low in forests and very high in aquatic communities

We can relate the productivity of a community to the standing crop biomass that produces it (the interest rate on the capital). Alternatively, we can think of the standing crop as the biomass that is sustained by the productivity (the capital resource that is sustained by earnings). Overall, there is a dramatic difference in the total biomass that exists on land (800 Pg) compared to the oceans (2 Pg) and fresh water (< 0.1 Pg) (Geider *et al.*, 2001). On an areal basis, biomass on land ranges from 0.2 to 200 kg m^{-2}, in the oceans from less than 0.001 to 6 kg m^{-2} and in freshwater biomass is generally less than 0.1 kg m^{-2} (Geider *et al.*, 2001). The average values of net primary productivity (NPP) and standing crop biomass (B) for a range of community types are plotted against each other in Figure 17.6. It is evident that a given value of NPP is produced by a smaller biomass when nonforest terrestrial systems are compared with forests, and the biomass involved is smaller still when aquatic systems are considered. Thus NPP : B ratios (kilograms of dry matter produced per year per kilogram of standing crop) average 0.042 for forests, 0.29 for other terrestrial systems and 17 for aquatic communities. The major reason for this is almost certainly that a large proportion of forest biomass is dead (and has been so for a long time) and also that much of the living support tissue is not photosynthetic. In grassland and scrub, a greater proportion of the biomass is alive and involved in photosynthesis, though half or more of the biomass may be roots. In aquatic communities, particularly where productivity is due mainly to phytoplankton, there is no support tissue, there is no need for roots to absorb water and nutrients, dead cells do not accumulate (they are usually eaten before they die) and the photosynthetic output per kilogram of biomass is thus very high indeed. Another factor that helps to account for high NPP : B ratios in phytoplankton communities is

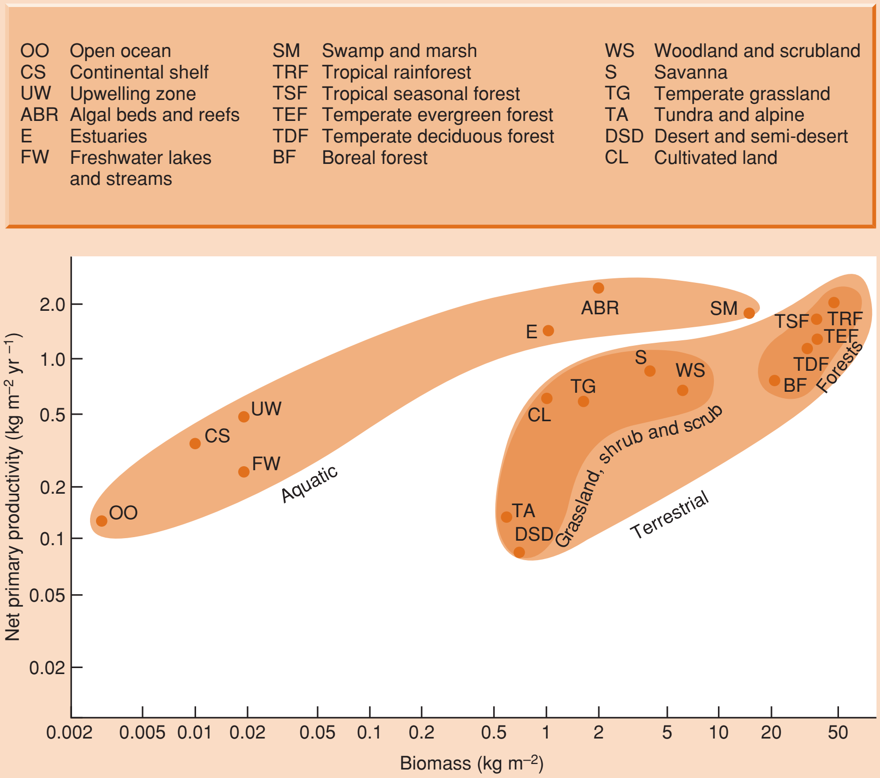


Figure 17.6 The relationship between average net primary productivity and average standing crop biomass for a range of ecosystems. (Based on data in Whittaker, 1975.)

the rapid turnover of biomass (turnover times of biomass in oceans and fresh waters average 0.02–0.06 years, compared to 1–20 years on land; Geider *et al.*, 2001). The annual NPP shown in the figure is actually produced by a number of overlapping phytoplankton generations, while the standing crop biomass is only the average present at an instant.

NPP : B ratios tend to decrease during successions

Ratios of NPP to biomass tend to decrease during successions. This is because the early successional pioneers are rapidly growing herbaceous species with relatively little support tissue (see Section 16.6). Thus, early in the succession the NPP : B ratio is high. However, the species that come to dominate later are generally slow growing, but eventually achieve a large size and come to monopolize the supply of space and light. Their structure involves considerable investment in nonphotosynthesizing and dead support tissues, and as a consequence their NPP : B ratio is low.

When attention is focused on trees, a common pattern is for above-ground NPP to reach a peak early in succession and then gradually decline by as much as 76%, with a mean reduction of 34% (Table 17.3). The reductions are no doubt partly due to a shift from photosynthesizing to respiring tissues. In addition, nutrient limitation may become more significant later in the succession or the longer branches and taller stems of older trees may increase resistance to the transpiration stream and thus limit photosynthesis (Gower *et al.*, 1996). Trees characteristic of different stages in succession show different patterns of NPP with stand age. In a subalpine coniferous forest, for example, the early successional whitebark pine (*Pinus albicaulis*) reached a peak above-ground NPP at about 250 years and then declined, whereas the late successional, shade-tolerant subalpine fir (*Abies lasiocarpa*) continued towards a maximum beyond 400 years (Figure 17.7). The late successional species allocated almost twice as much biomass to leaves as its early successional counterpart, and maintained a high photosynthesis : respiration ratio to a greater age (Callaway *et al.*, 2000).

17.3 Factors limiting primary productivity in terrestrial communities

Sunlight, carbon dioxide (CO_2), water and soil nutrients are the resources required for primary production on land, while temperature, a condition, has a strong influence on the rate

Table 17.3 Above-ground net primary productivity (ANPP) for forest age sequences in contrasting biomes. (After Gower *et al.*, 1996.)

Biome/species	*Location*	*Range of stand ages, in years (no. of stands shown in brackets)*	*ANPP (t dry mass ha^{-1} $year^{-1}$)* Peak	Oldest	% change
Boreal					
Larix gmelinii	Yakutsk, Siberia	50–380 (3)	4.9	2.4	−51
Picea abies	Russia	22–136 (10)	6.2	2.6	−58
Cold temperate					
Abies baisamea	New York, USA	0–60 (6)	3.2	1.1	−66
Pinus contorta	Colorado, USA	40–245 (3)	2.1	0.5	−76
Pinus densiflora	Mt Mino, Japan	16–390 (7)	16.1	7.4	−54
Populus tremuloides	Wisconsin, USA	8–83 (5)	11.1	10.7	−4
Populus grandidentata	Michigan, USA	10–70	4.6	3.5	−24
Pseudotsuga menziesii	Washington, USA	22–73 (4)	9.9	5.1	−45
Warm temperate					
Pinus elliottii	Florida, USA	2–34 (6)	13.2	8.7	−34
Pinus radiata	Puruki, NZ (Tahi)	2–6 (5)	28.5	28.5	0
	(Rue)	2–7 (6)	29.2	23.5	−20
	(Toru)	2–8 (7)	31.1	31.1	0
Tropical					
Pinus caribaea	Afaka, Nigeria	5–15 (4)	19.2	18.5	−4
Pinus kesiya	Meghalaya, India	1–22 (9)	30.1	20.1	−33
Tropical rainforest	Amazonia	1–200 (8)	13.2	7.2	−45

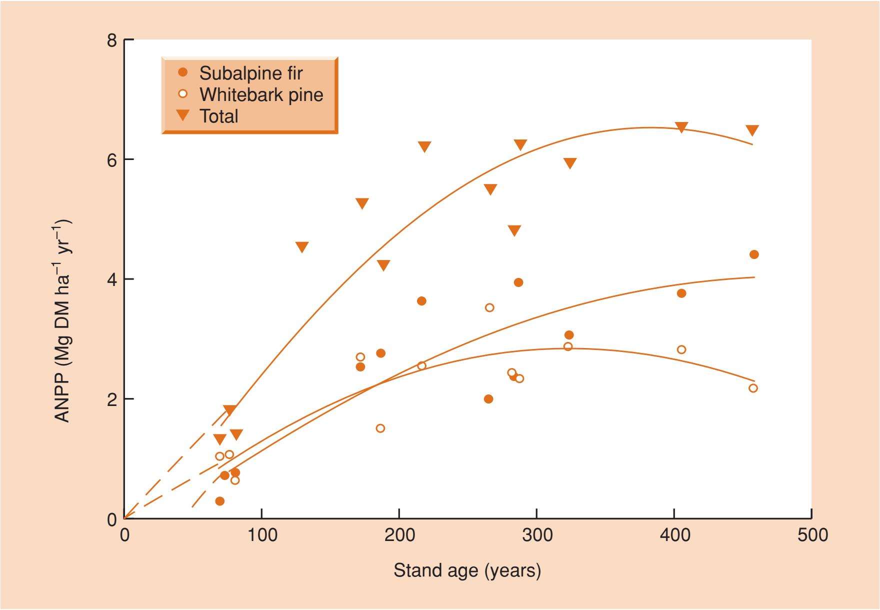


Figure 17.7 Annual above-ground net primary productivity (ANPP) (Mg dry matter ha^{-1} $year^{-1}$) in stands of different ages in a subalpine coniferous forest in Montana, USA: early successional whitebark pine, late successional subalpine fir, and total ANPP. (After Callaway *et al.*, 2000.)

of photosynthesis. CO_2 is normally present at a level of around 0.03% of atmospheric gases. Turbulent mixing and diffusion prevent the CO_2 concentration from varying much from place to place, except in the immediate neighborhood of a leaf, and CO_2 probably plays little role in determining differences between the productivities of different communities (although global increases in CO_2 concentration are expected to have profound effects (e.g. DeLucia *et al.*, 1999). On the other hand, the quality and quantity of light, the availability of water and nutrients, and temperature all vary dramatically from place to place. They are all candidates for the role of limiting factor. Which of them actually sets the limit to primary productivity?

17.3.1 Inefficient use of solar energy

terrestrial communities use radiation inefficiently

Depending on location, something between 0 and 5 joules of solar energy strikes each square meter of the earth's surface every minute. If all this were converted by photosynthesis to plant biomass (that is, if photosynthetic efficiency were 100%) there would be a prodigious generation of plant material, one or two orders of magnitude greater than recorded values. However, much of this solar energy is unavailable for use by plants. In particular, only about 44% of incident shortwave radiation occurs at wavelengths suitable for photosynthesis. Even when this is taken into account, though, productivity still falls well below the maximum possible. Photosynthetic efficiency has two components – the efficiency with which light is intercepted by leaves and the efficiency with which intercepted light is converted by photosynthesis to new biomass (Stenberg *et al.*, 2001). Figure 17.8 shows the range in overall net photosynthetic efficiencies (percentage of incoming photosynthetically active radiation (PAR) incorporated into above-ground NPP) in seven coniferous forests, seven deciduous forests and eight desert communities studied as part of the International Biological Programme (see Section 17.1). The conifer communities had the highest efficiencies, but these were only between 1 and 3%. For a similar level of incoming radiation, deciduous forests achieved 0.5–1%, and, despite their greater energy income, deserts were able to convert only 0.01–0.2% of PAR to biomass.

productivity may still be limited by a shortage of PAR

However, the fact that radiation is not used efficiently does not in itself imply that it does not limit community productivity. We would need to know whether at increased intensities of radiation the productivity increased or remained unchanged. Some of the evidence given in Chapter 3 shows that the intensity of light during part of the day is below the optimum for canopy photosynthesis. Moreover, at peak light intensities, most canopies still have their lower leaves in relative gloom, and would almost certainly photosynthesize faster if the light intensity were higher. For C_4 plants a saturating intensity of radiation never seems to be reached, and the implication is that productivity may in fact be limited by a shortage of PAR even under the brightest natural radiation.

There is no doubt, however, that what radiation is available would be used more efficiently if other resources were in abundant supply. The much higher values of community productivity recorded from agricultural systems bear witness to this.

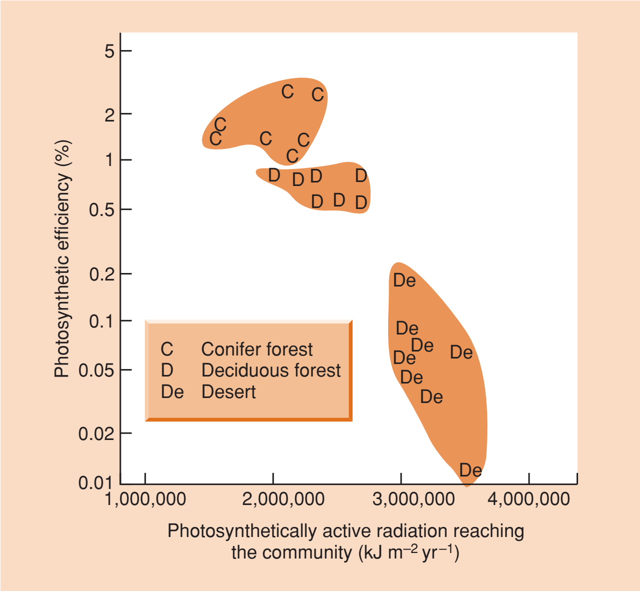


Figure 17.8 Photosynthetic efficiency (percentage of incoming photosynthetically active radiation converted to above-ground net primary productivity) for three sets of terrestrial communities in the USA. (After Webb *et al.*, 1983.)

17.3.2 Water and temperature as critical factors

shortage of water may be a critical factor

The relationship between the NPP of a wide range of ecosystems on the Tibetan Plateau and both precipitation and temperature is illustrated in Figure 17.9. Water is an essential resource both as a constituent of cells and for photosynthesis. Large quantities of water are lost in transpiration – particularly because the stomata need to be open for much of the time for CO_2 to enter. It is not surprising that the rainfall of a region is quite closely correlated with its productivity. In arid regions, there is an approximately linear increase in NPP with increase in precipitation, but in the more humid forest climates there is a plateau beyond which productivity does not continue to rise. Note that a large amount of precipitation is not necessarily equivalent to a large amount of water available for plants; all water in excess of field capacity will drain away if it can. A positive relationship between productivity and mean annual temperature can also be seen in Figure 17.9. However, the pattern can be expected to be complex because, for example, higher temperatures are associated with rapid water loss through evapotranspiration; water shortage may then become limiting more quickly.

interaction of temperature and precipitation

To unravel the relationships between productivity, rainfall and temperature, it is more instructive to concentrate on a single ecosystem

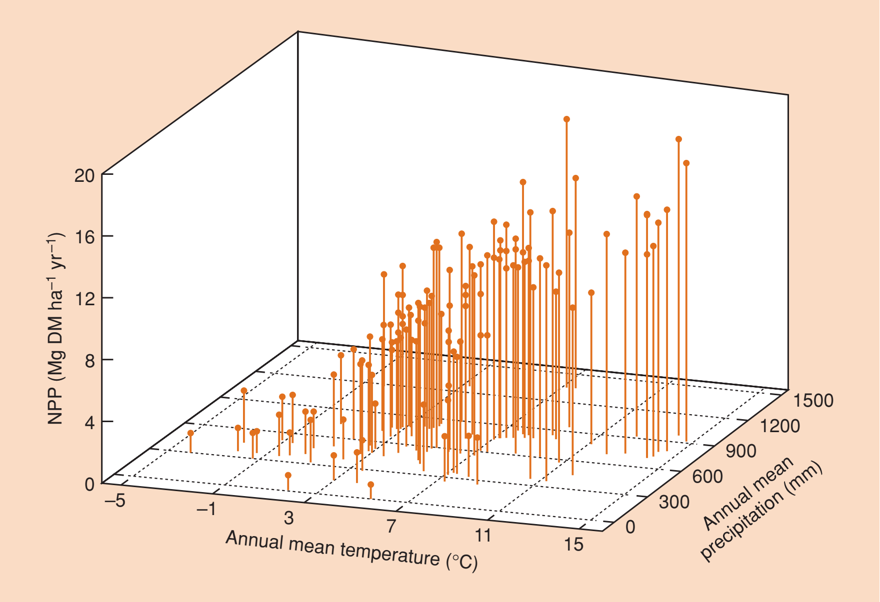


Figure 17.9 Relationship between total net primary productivity (Mg dry matter ha^{-1} year^{-1}) and annual precipitation and temperature for ecosystems on the Tibetan Plateau. The ecosystems include forests, woodlands, shrublands, grasslands and desert. (After Luo *et al.*, 2002.)

type. Above-ground NPP was estimated for a number of grassland sites along two west-to-east precipitation gradients in the Argentinian pampas. One of these gradients was in mountainous country and the other in the lowlands. Figure 17.10 shows the relationship between an index of above-ground NPP (ANPP) and precipitation and temperature for the two sets of sites. There are strong positive relationships between ANPP and precipitation but the slopes of the relationships differed between the two environmental gradients (Figure 17.10a).

The relationships between ANPP and temperature are similar for two further environmental gradients (both north-to-south elevation transects) in Figure 17.10b – both show a hump-shaped pattern. This probably results from the overlap of two effects of increasing temperature: a positive effect on the length of the

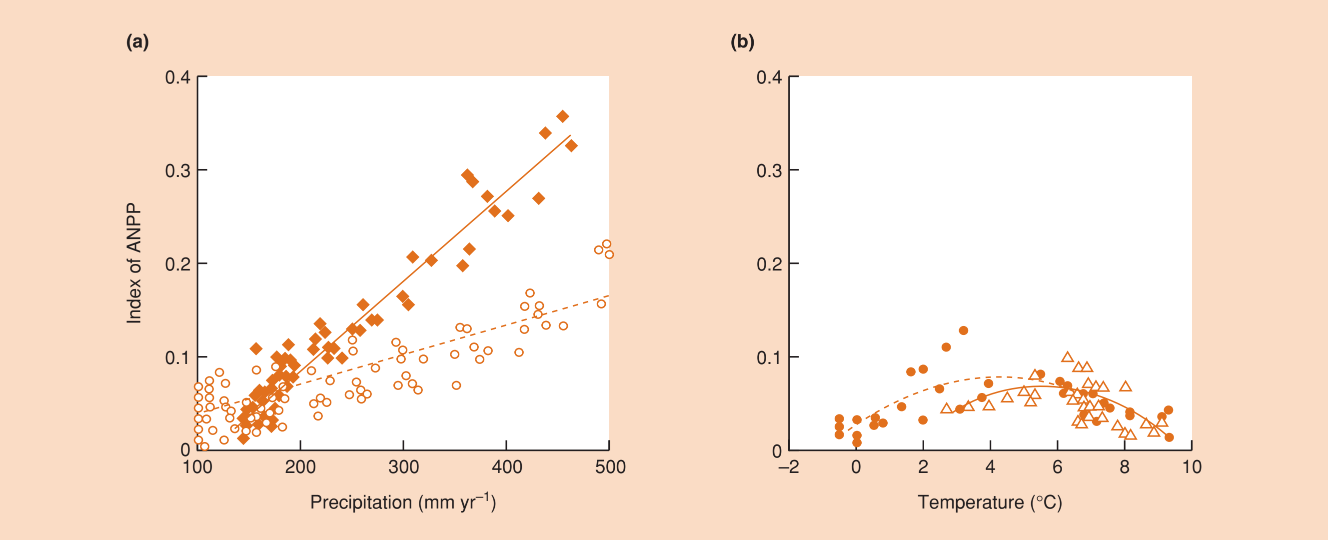


Figure 17.10 Annual above-ground net primary productivity (ANPP) of grasslands along two precipitation gradients in the Argentinian pampas. NPP is shown as an index based on satellite radiometric measurements with a known relationship to absorbed photosynthetically active radiation in plant canopies. (a) NPP in relation to annual precipitation. (b) NPP in relation to annual mean temperature. Open circles and diamonds represent sites along precipitation gradients in the lowland and mountainous regions respectively. Closed circles and triangles represent sites along two elevation transects. (After Jobbagy *et al.*, 2002.)

growing season and a negative effect through increased evapotranspiration at higher temperatures. Because temperature is the main constraint on productivity at the cool end of the gradients, an increase in NPP is observed as we move from the coolest to warmer sites. However, there is a temperature value above which the growing season does not lengthen and the dominating effect of increasing temperature is now to increase evapotranspiration, thus reducing water availability and curtailing NPP (Epstein *et al.*, 1997).

productivity and the structure of the canopy

Water shortage has direct effects on the rate of plant growth but also leads to the development of less dense vegetation. Vegetation that is sparse intercepts less light (much of which falls on bare ground). This wastage of solar radiation is the main cause of the low productivity in many arid areas, rather than the reduced photosynthetic rate of drought-affected plants. This point is made by comparing the productivity per unit weight of leaf biomass instead of per unit area of ground for the studies shown in Figure 17.8. Coniferous forest produced 1.64 g g^{-1} year^{-1}, deciduous forest 2.22 g g^{-1} year^{-1} and desert 2.33 g g^{-1} year^{-1}.

17.3.3 Drainage and soil texture can modify water availability and thus productivity

There was a notable difference in the slopes of the graphs of NPP against precipitation for the mountainous and lowland sites in Figure 17.10. The slope was much lower in the mountainous case and it seems likely that the steeper terrain in this region resulted in a higher rate of water runoff from the land and, thus, a lower efficiency in the use of precipitation (Jobbagy *et al.*, 2002).

soil texture can influence productivity

A related phenomenon has been observed when forest production on sandy, well-drained soils is compared with soils consisting of finer particle sizes. Data are available for the accumulation through time of forest biomass at a number of sites where all the trees had been removed by a natural disturbance or human clearance. For forests around the world, Johnson *et al.* (2000) have reported the relationship between above-ground biomass accumulation (a rough index of ANPP) and accumulated growing season degree-days (stand age in years × growing season temperature × growing season as a proportion of the year). In effect, 'growing season degree-days' combine the time for which the stand has been accumulating biomass with the average temperature at the site in question. Figure 17.11 shows that productivity of broadleaf forests is generally much lower, for a given value for growing season degree-days, when the forest is on sandy soil. Such soils have less favorable soil-moisture-holding capacities and this accounts in some measure for their poorer productivity. In addition, however, nutrient retention may be lower in coarse soils, further reducing productivity compared to soils with finer texture. This was confirmed by Reich *et al.* (1997) who, in their compilation of data for 50 North American forests, found that soil nitrogen availability (estimated as annual net nitrogen mineralization rate) was indeed lower in sandier soils and, moreover, that ANPP was lower per unit of available nitrogen in sandy situations.

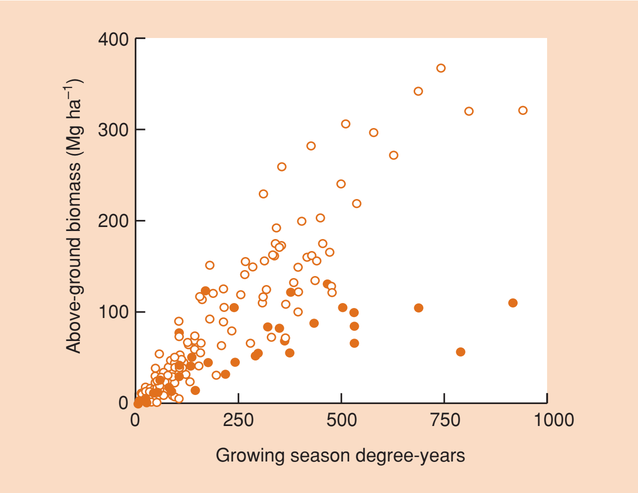


Figure 17.11 Above-ground biomass accumulation (a rough index of NPP) expressed as megagrams (= 10^6 g) per hectare in relation to accumulated growing season degree-days in broadleaf forest stands growing on sandy or nonsandy soils. ○, nonsandy soils; ●, sandy soils. (After Johnson *et al.*, 2000.)

17.3.4 Length of the growing season

The productivity of a community can be sustained only for that period of the year when the plants have photosynthetically active foliage. Deciduous trees have a self-imposed limit on the period when they bear foliage. In general, the leaves of deciduous species photosynthesize fast and die young, whereas evergreen species have leaves that photosynthesize slowly but for longer (Eamus, 1999). Evergreen trees hold a canopy throughout the year, but during some seasons they may barely photosynthesize at all or may even respire faster than they photosynthesize. Evergreen conifers tend to dominate in nutrient-poor and cold conditions, perhaps because in other situations their seedlings are outcompeted by their faster growing deciduous counterparts (Becker, 2000).

length of the growing season: a pervasive influence on productivity

The latitudinal patterns in forest productivity seen earlier (see Table 17.2) are largely the result of differences in the number of days when there is active photosynthesis. In this context, Black *et al.* (2000) measured net ecosystem productivity (NEP) in a boreal deciduous forest in Canada for 4 years. First leaf emergence occurred considerably earlier in 1998 when

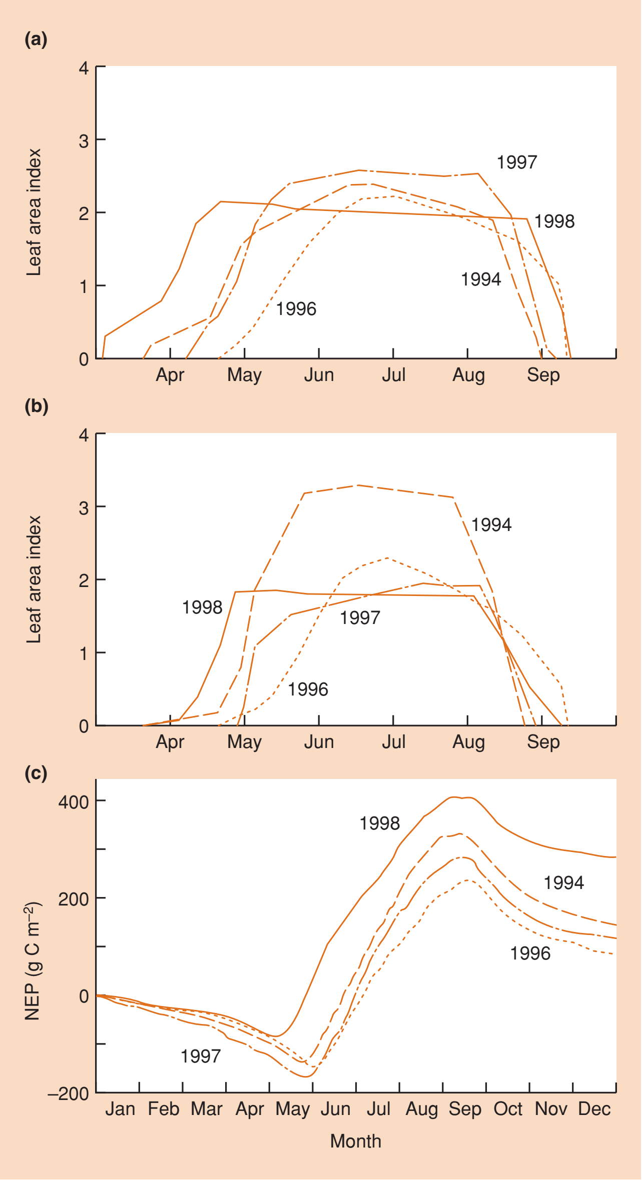


Figure 17.12 Seasonal patterns in leaf area index (area of leaves divided by ground area beneath the foliage) of (a) overstory aspen (*Populus tremuloides*) and (b) understory hazelnut (*Corylus cornuta*) in a boreal deciduous forest during four study years with contrasting spring temperatures. (c) Cumulative net ecosystem productivity (NEP). (After Black *et al.*, 2000.)

the April/May temperature was warmest (9.89°C) and a month later in 1996 when the April/May temperature was coldest (4.24°C) (Figure 17.12a, b). Equivalent spring temperatures in 1994 and 1997 were 6.67 and 5.93°C. The difference in the length of the growing season in the four study years can be gauged from the pattern of cumulative NEP (Figure 17.12c). During winter and early spring, NEP was negative because ecosystem respiration exceeded gross ecosystem productivity. NEP became positive earlier in warmer years (particularly 1998) so that overall total carbon sequestered by the ecosystem in the four years was 144, 80, 116 and 290 g C m^{-2} $year^{-1}$ for 1994, 1996, 1997 and 1998, respectively.

In our earlier discussion of the study of Argentinian pampas communities (see Figure 17.10) we noted that higher NPP was not only directly affected by precipitation and temperature but was partly determined by length of the growing season. Figure 17.13 shows that the start of the growing season was positively related to mean annual temperature (paralleling the boreal forest study above), whereas the end of the growing season was determined partly by temperature but also by precipitation (it ended earlier where temperatures were high and precipitation was low). Again we see a complex interaction between water availability and temperature.

17.3.5 Productivity may be low because mineral resources are deficient

the crucial importance of nutrient availability

No matter how brightly the sun shines and how often the rain falls, and no matter how equable the temperature is, productivity must be low if there is no soil in a terrestrial community, or if the soil is deficient in essential mineral nutrients. The geological conditions that determine slope and aspect also determine whether a soil forms, and they have a large, though not wholly dominant, influence on the mineral content of the soil. For this reason, a mosaic of different levels of community productivity develops within a particular climatic regime. Of all the mineral nutrients, the one that has the most pervasive influence on community productivity is fixed nitrogen (and this is invariably partly or mainly biological, not geological, in origin, as a result of nitrogen fixation by microorganisms). There is probably no agricultural system that does not respond to applied nitrogen by increased primary productivity, and this may well be true of natural vegetation as well. Nitrogen fertilizers added to forest soils almost always stimulate forest growth.

The deficiency of other elements can also hold the productivity of a community far below that of which it is theoretically capable. A classic example is deficiency of phosphate and zinc in South Australia, where the growth of commercial forest (Monterey pine, *Pinus radiata*) is made possible only when these nutrients are supplied artificially. In addition, many tropical systems are primarily limited by phosphorus.

17.3.6 Résumé of factors limiting terrestrial productivity

The ultimate limit on the productivity of a community is determined by the amount of incident radiation that it receives – without this, no photosynthesis can occur.

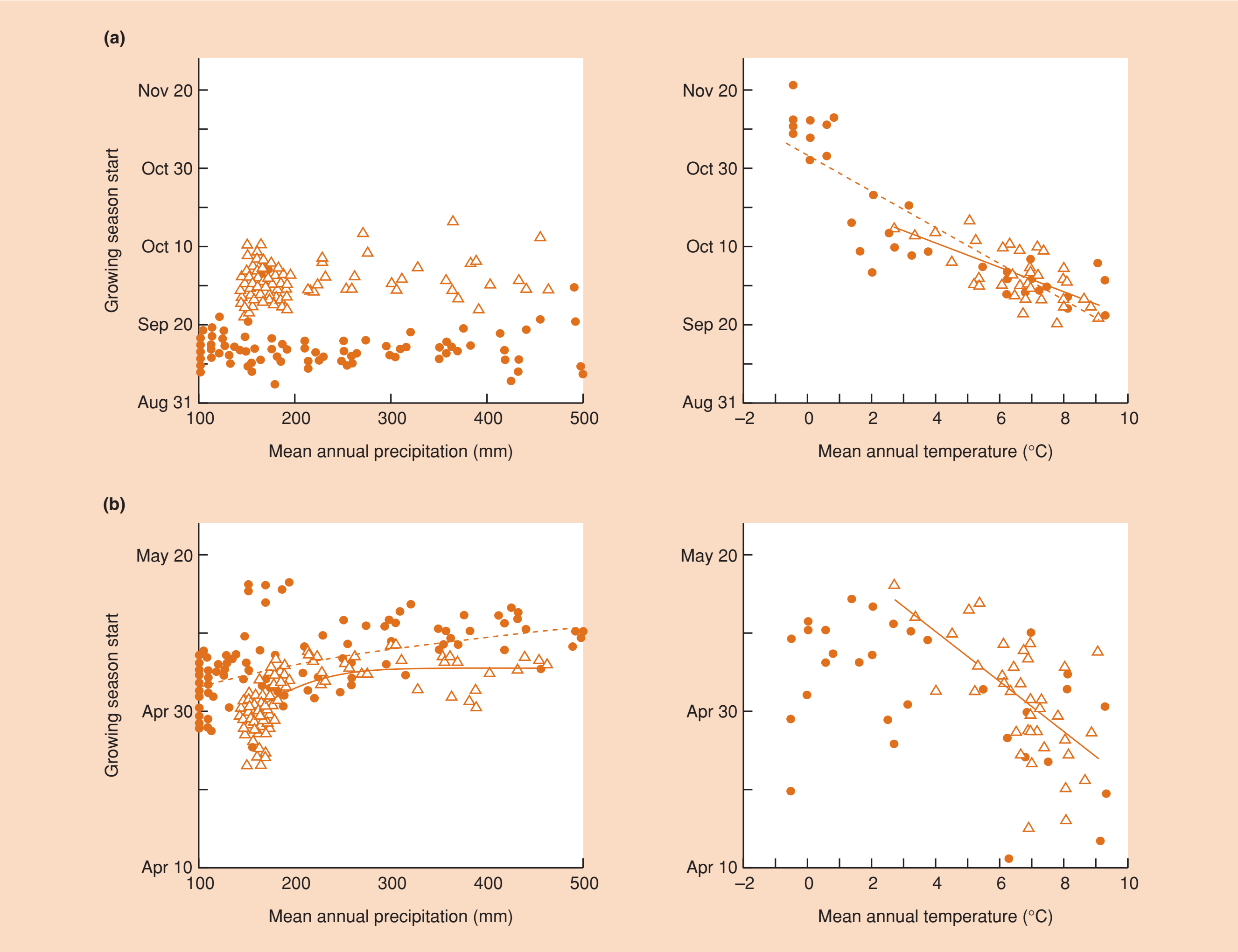


Figure 17.13 (a) Start and (b) end dates of the growing season for Argentinian pampas communities described in Section 17.3.2. Circles represent sites along the precipitation gradient in the mountainous region and triangles represent sites along the lowland gradient. (After Jobbagy *et al.*, 2002.)

Incident radiation is used inefficiently by all communities. The causes of this inefficiency can be traced to: (i) shortage of water restricting the rate of photosynthesis; (ii) shortage of essential mineral nutrients, which slows down the rate of production of photosynthetic tissue and its effectiveness in photosynthesis; (iii) temperatures that are lethal or too low for growth; (iv) an insufficient depth of soil; (v) incomplete canopy cover, so that much of the incident radiation lands on the ground instead of on foliage (this may be because of seasonality in leaf production and leaf shedding *or* because of defoliation by grazing animals, pests and diseases); and (vi) the low efficiency with which leaves photosynthesize – under ideal conditions, efficiencies of more than 10% (of PAR) are hard to achieve even in the most productive agricultural systems. However, most of the variation in primary productivity of world vegetation is due to factors (i) to (v), and relatively little is accounted for by intrinsic differences between the photosynthetic efficiencies of the leaves of the different species.

In the course of a year, the productivity of a community may (and probably usually will) be limited by a succession of the factors (i) to (v). In a grassland community, for instance, the primary productivity may be far below the theoretical maximum because the winters are too cold and light intensity is low, the summers are too dry, the rate of nitrogen mobilization is too slow, and for periods grazing animals may reduce the standing crop to a level at which much incident light falls on bare ground.