
Evaluating the relative strengths of biotic versus abiotic controls on ecosystem processes

M. A. Huston and A. C. McBride

5.1 Introduction

There is no doubt that the rates of ecosystem processes and the resulting properties and structure of ecosystems are influenced both by biotic factors, including species properties and interactions among multiple species, and by abiotic factors, such as climate, geology, and soils. The critical issue (and the focus of the debate about diversity and ecosystem function) is the relative contribution of each of these general factors to any particular ecosystem process, under any particular conditions, and at any particular scale.

While much remains to be learnt from future experiments, we can nonetheless evaluate the relative effects of biotic and abiotic factors in the results of current experiments, where they presumably result from processes operating in much the same way they operate under natural conditions. The detection of both types of factors in diversity experiments provides some insights into why these experiments have been so controversial. We will first review the patterns in natural systems that seem to contradict many of the experimental results, and then examine some of the experiments in more detail, in order to distinguish between biotic and abiotic effects.

5.2 Natural patterns of diversity and productivity

The conclusion of virtually all of the 'diversity-productivity' experiments that productivity

increases with increasing species diversity (Naeem *et al.* 1994a,b, 1995; Tilman 1996; Tilman *et al.* 1996, 2001; Hector *et al.* 1999; see however, Hooper and Vitousek 1997; Wardle *et al.* 1997b; Hooper 1998; Schwartz *et al.* 2000; Fridley in press) conflicts with the fact that most of the high productivity ecosystems found around the world have strikingly low plant diversity. This phenomenon was well known to early ecologists (Lawes *et al.* 1882) and was succinctly summarized by Rosenzweig as the 'paradox of enrichment' (1971), referring to the diversity-reducing effects of added nutrients (i.e. eutrophication).

In many regions, plant diversity decreases with increasing productivity over most of the range of productivity (Huston 1980, 1994). Naturally productive ecosystems with low species diversity include phytoplankton and algal blooms, salt marshes, freshwater marshes, riparian forests in the tropics and temperate zones (e.g. *Populus deltoides*, *Prioria copaifera*), bamboo forests, redwood forests, Douglas fir forests, eucalypt forests etc. High diversity plant communities are generally found on relatively unproductive sites, such as chalk grasslands, Mediterranean shrublands, and rainforests on oxisols and ultisols (Dawkins 1959, 1964; Grime 1979, 2001; Mahdi *et al.* 1989; Berendse 1994a; Huston 1979, 1980, 1993, 1994). If productivity in random-selection diversity experiments increases as a consequence of increasing diversity in experiments, then either (1) the processes involved in the experiments differ from the processes involved in natural patterns of diversity; or (2) the experiments

address a different portion of the complete range of productivity, where patterns may differ from those in other portions of the range.

Notwithstanding the inconsistency of the experimental results with high productivity–low diversity and low productivity–high diversity natural ecosystems, the experimental results are compatible with the observation that diversity increases with increasing productivity under low productivity conditions (Grime 1973a,b, 1979; Huston 1979, 1980, 1994; Huston and DeAngelis 1994). When put together, the naturally occurring increase in diversity with increasing productivity at low levels of productivity and the naturally occurring decrease in diversity with increasing productivity at high levels of productivity produce the unimodal ‘hump-backed’ diversity curve originally described by Grime (1973a,b, 1979; Al-Mufti *et al.* 1977).

However, mere consistency of the experimental pattern and the natural pattern under low productivity conditions does not mean that they have the same cause. The natural pattern of increasing plant diversity as productivity increases from very low levels can be understood most simply as resource levels rising above the minimum required for survival of various species with different minimum resource requirements (Grime 1973a,b, 1979; Huston 1979; Huston and DeAngelis 1994). This is a phenomenon observed along gradients of increasing resource supply in low resource environments (e.g. water in deserts), or potentially in experiments that use variation in resource availability as a treatment (e.g. Lawes *et al.* 1882; Huston 1979, 1980, 1994; Tilman 1987, 1988; Fridley 2001).

The increase in average productivity with increasing species richness that has been observed in many of the diversity–productivity experiments occurs under presumably constant and uniform resource conditions (i.e. soil nutrients and water), except to the degree that they are altered by the plants themselves. Thus, any increase in productivity with increasing diversity is expected to result from either (1) complementary resource use among multiple species (with no change in total resources, just a more complete use of them) or (2) facilitative interactions among species, in which one or more species improve conditions (resource availability or

a physical condition, such as temperature or pH) for other species, which then grow faster or larger than they would in the absence of the facilitator (Vandermeer 1989; Hooper 1998). The most widespread and important example of facilitation is nitrogen fixation by legumes, which has the effect of increasing nitrogen availability for other plant species that may eventually overgrow and eliminate the legumes (DeWit *et al.* 1966).

In summary, variation in species richness can play only a minor role in natural patterns of productivity, which are regulated primarily by environmental conditions such as climate and soils. Only when the dominant effects of environmental conditions on productivity are controlled experimentally can the subtle effects of species composition on productivity be detected. Consequently, the failure to adequately control environmental conditions is potentially a problem with any experiment that attempts to evaluate the effect of species richness, or any type of plant interactions, on productivity.

5.3 The influence of soil heterogeneity on diversity–productivity experiments

Small differences in soil conditions can have a large effect on plant growth and species diversity (Fig. 5.1). Natural variation in soil nutrients and water is associated with subtle differences in topography, with even slight depressions generally having higher nutrient and water availability than the surrounding area. Dealing with this natural variability is a major issue in the design of agricultural and ecological field experiments (Trenbath 1974; Vandermeer 1989, see Chapter 19), and has led to such designs as the ‘Latin Square’, replication by blocks for ANOVA etc.

Separation of ‘block effects’, some of which could be due to differences in soil conditions, from the main treatment effects requires replication of the full experiment over a set of blocks that may or may not have significant differences in soil conditions. Such replication adds cost and complexity to any experimental design, and may be infeasible for experiments with many different treatment levels. In random-selection diversity experiments, the

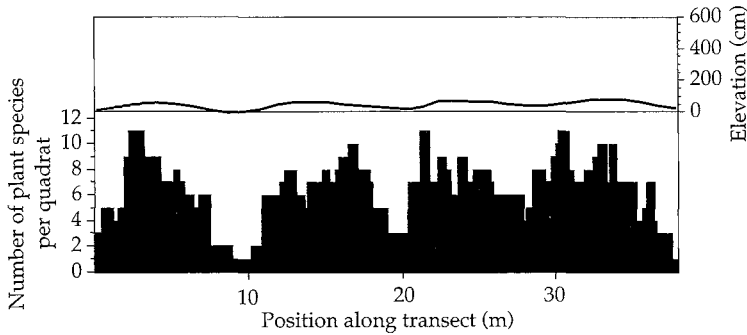


Figure 5.1 Natural variation in plant species richness across the repeating topographic pattern of 'ridge and furrow' grassland in Great Britain. Note that species richness is highest on the ridges, and lowest in the moister, more nutrient-rich furrows, where two clonal species, *Agrostis stolonifera* and *Ranunculus repens*, tend to exclude all the other species. (From unpublished data of G. R. Sagar in Harper 1977).

relatively large number of plots within any particular diversity level has generally precluded any replication by spatial blocks. Consequently, the performance of any treatment plot could potentially be influenced by three factors: (1) the number of species; (2) the particular species that are present; and/or (3) the soil conditions at the location of that plot. Separating the effects of these three factors is the primary challenge in the analysis of the experimental results, as well as the primary cause for disagreement about the interpretation of published results. Unfortunately, this is not a simple task, nor are standard statistical approaches well suited for the task (cf. Schmid *et al.*, Chapter 6).

The effect of soil heterogeneity on experimental results can be seen in the results of the BIODEPTH experiments (Hector *et al.* 1999, see Chapter 4), which were conducted at eight sites in seven European countries. The locations spanned a latitudinal gradient from Sweden to Greece, and included a range of soil conditions and climates. In addition to the variation among the sites at which the experiments were conducted (not replicated, since all sites did not use the same design), there was inevitably soil heterogeneity within the field in which each experiment was conducted.

A striking property of the BIODEPTH results, as well as the results of other experiments of this type, is the high variability in the responses at any given level of species richness (Fig. 5.2). However, this variability is not uniform across all sites, which provides a critical insight into the influence of soil factors on the results of 'pure diversity' treatments. A scattering of very high values, above the level at which most plots are clustered, can be seen at all of the sites but one, the site at Sheffield in UK.

Sheffield was one of three sites that used an experimental design suitable for distinguishing the effects of single species from multiple species using analysis of 'over-yielding' (the other two being Portugal and Sweden). However, Sheffield was the only site in the entire experiment in which the soil in the experimental field was artificially constructed by bringing in truckloads of soil, and mixing and spreading it uniformly over an existing field (A. Hector, personal communication). The exceptional homogeneity of the artificially constructed soil at Sheffield seems to be the explanation for the extremely low variability among the plots at each diversity level, particularly at the highest level where all the plots have identical species composition (Huston *et al.* 2000). This low variance within each treatment level leads to the highest proportion of variance explained by the experimental treatment at any of the eight sites ($r^2 = 0.49$).

The extremely low variance among plots at the same level of species richness on the homogeneous constructed soils of the Sheffield site suggests that the high variance observed at the other sites is at least partially due to variance in soil conditions across the experimental site. At Sheffield and the other two sites designed for over-yielding analysis, the variability at the highest level of species richness can be attributed primarily to soil variability because all of the replicates are identical in species composition. However, the other five sites include species in the high diversity treatments that were not planted in the monocultures. Consequently, some of the variance at these sites may be due to differences in species composition even at the high species richness levels, as well as to the inevitable heterogeneity in soil conditions across the sites, and differences in

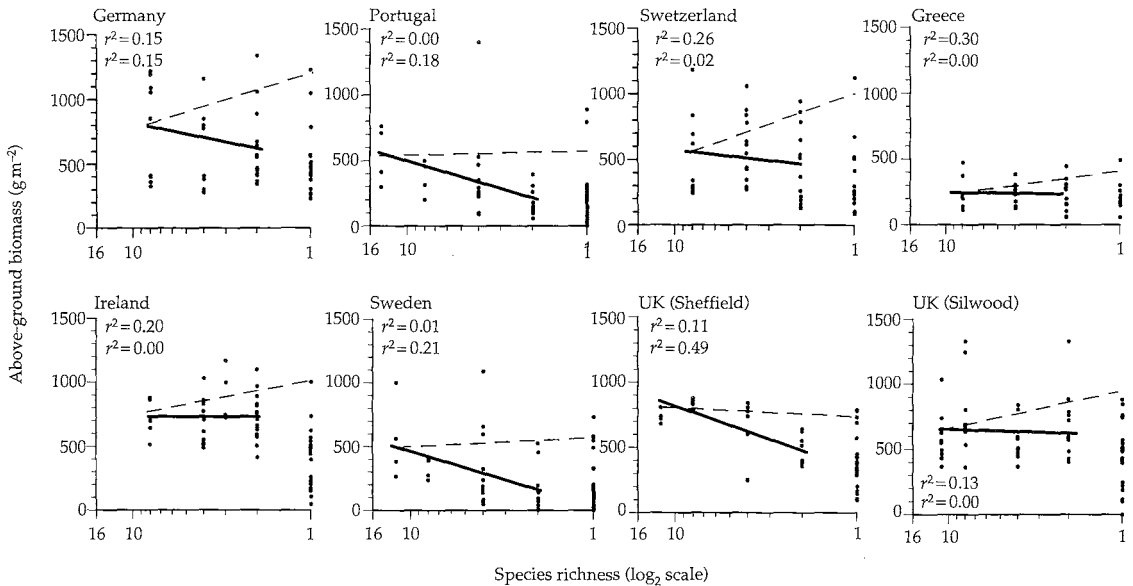


Figure 5.2 Results of the BIODEPTH experiment at eight sites across Europe. Only treatments with species richness less than or equal to the number of species used in monoculture treatment are shown. Solid line is log-linear regression through all plots with two or more species. Dotted line is the log-linear regression through plots expected to contain the species that was most productive in monoculture, for evaluation of the effect of species richness onoveryielding. Upper r^2 value refers to the dotted line; lower value refers to solid line (from Huston *et al.* 2000). Note that Sheffield, with artificially constructed soils, has the lowest within-treatment variance at high treatment levels.

species composition at the lower levels of species richness.

In summary, for diversity experiments in which soil conditions are not controlled by site preparation, variability in soil properties across an experimental site can have a strong effect on the magnitude and variability of the measured responses, such as above-ground biomass production. This introduces the possibility that soil heterogeneity can act as a 'hidden treatment' in these experiments, because soil variation can produce the same response, an increase in productivity, that is predicted to occur as a result of the diversity treatments.

A hidden treatment is any factor that is correlated with the treatments in an experiment, but not explicitly controlled, manipulated, measured, or considered in the experimental design and analysis of the experimental results (Huston 1997). If such factors are not recognized or are ignored in the analysis of data, they can lead to misinterpretation of causal relationships because the experimental treatment is assumed (in the context of regression analyses or ANOVA) to cause the observed

responses, rather than the hidden treatment factor that actually produced the response.

5.4 Quasi-replication, representation, similarity, and inhomogeneity of variance: a suite of hidden treatments in random-selection experiments

Given that there is inevitably some degree of soil heterogeneity in field experiments, the critical question becomes how inherent soil effects can be distinguished from the effects of the experimental treatment, specifically from the effects of species number or species composition. Obviously, experiments can be established using designs capable of distinguishing field effects from treatment effects. Alternatively, the actual soil properties in each plot can be measured and used as covariates in the analysis. However, neither of these options have been typically used in diversity–ecosystem function experiments, primarily because the logistics and cost of replicating species richness treatments is already quite daunting.

The issue of replication in random-selection diversity experiments is more complex than in standard experiments. Replication (or sample size) is a central issue in the design and analysis of all experiments. In general, higher replication increases the accuracy of estimates of mean response, and also increases the probability that any experimental response that is detected is statistically significant because it allows the response distribution to be determined more accurately. The pattern of variance among responses at different treatment levels is a critical issue in statistical analysis because 'homogeneity of variance' is a basic precondition for the validity of many statistical tests, including ANOVA. The critical assumption is that the variance among replicates in any sample (or treatment level) is the same as the variance in any other sample, i.e. variance is equal (or homogeneous) across all samples that are being compared (Sokal and Rohlf 1981).

Several inherent features in the design of random-selection diversity experiments cause the response variance to differ between treatment levels, compromising the validity of most statistical tests that have been used to analyse these experiments. Furthermore, the pattern of unequal variance is correlated with treatment levels because the cause of variance changes predictably with an increasing number of species. The assumption that variance is equal among treatments is invalid if the causes of variance differ among treatments. This correlation of treatment level with the causes of within-treatment variance increases the probability that the experimental results will be misinterpreted.

Understanding the patterns and causes of within-treatment variance in random-selection diversity experiments requires a careful evaluation of what replication means in these experiments. What is treated as replication in these experiments actually has two very different components, 'representation' and 'true replication', that change in their relative proportions along the treatment gradient from low species richness to high species richness.

Representation is the proportion of the total number of possible species compositions at a specific level of species richness (calculated using combinatorial probabilities) that is actually included in the experiment. The number of unique combinations of

x species drawn from a total species pool of N is calculated as $N!/(x!(N-x)!)$. Thus, for an experiment with a total species pool of 18, there are 153 different two-species combinations, and also 153 different 16-species combinations, but only one 18-species combination (Fig. 5.3).

True replication refers to the number of identical plots, i.e. plots with exactly the same set of planted species. For example, in an experiment with a total species pool of 18 that includes a monoculture of each species, representation at the one-species level would be 100%. If there were two monoculture

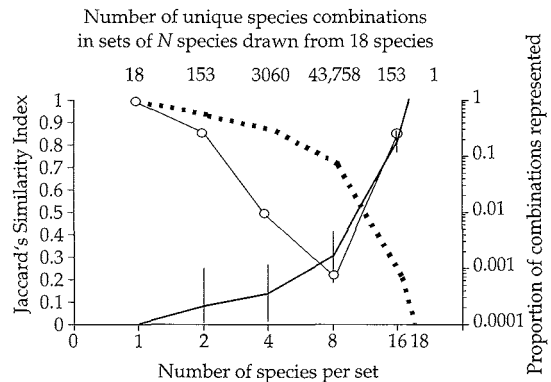


Figure 5.3 Statistical properties related to within-treatment variance that co-vary with treatment level (number of species per set) in random-selection diversity experiments. Jaccard's Similarity Index (left-hand scale and thick line with bars for ± 1 SD), illustrates the increasing similarity in species composition among randomly assembled sets of species (quasi-replicates) as the number of species in the set approaches the total number of species, caused by the variance reduction effect (Huston 1997). The proportion of variance ($\bullet\bullet\bullet\bullet$) at a specific treatment level due to differences in species composition is the inverse of the similarity among the quasi-replicates at that level. The proportion of combinations represented (right-hand scale and $\text{---}\circ\text{---}$) is the number of randomly drawn quasi-replicates at each treatment level (39, 35, 29, 30, 35 for the five treatment levels) divided by the total number of unique combinations of species at that treatment level (determined by combinatorial probabilities and listed at the top of the figure). The consequences of these co-varying properties are that (1) treatment responses are underestimated at low treatment levels because of under-sampling of the high variability in species composition (and therefore performance) among quasi-replicates, and (2) treatment effects are overestimated at high treatment levels because the low variability in species composition at high treatment levels allows experimental error (such as uncontrolled variability in soil properties) to become the primary source of variability.

plots of each species, the level of true replication would be 2, for a total of 36 plots.

In a 'well-replicated' diversity-productivity experiment, there may be 30 or more plots at each level of species richness (e.g. 1,2,4,8,16). However, these plots are not true replicates but rather 'quasi-replicates' whose statistical properties are correlated (and thus confounded) with the experimental diversity treatments. Note that 'quasi-replicates' represent a different phenomenon than 'pseudo-replicates', *sensu* Hurlbert (1984). The quasi-replicates at each treatment level are composed of a variable proportion of the total possible number of combinations, and, with a very low probability, some true replicates of a few of the combinations. For an experiment with 35 plots of two species drawn from a pool of 18 species, a maximum of 22.8% (35/153) of all possible combinations could be represented, and the probability that even a single combination is drawn twice (i.e. replicated) is quite low. Even with a number of quasi-replicates as large as 35, over 75% of the possible species combinations will not be represented, so there is a high probability that any specific pair of species (such as the two largest species, a particular legume and grass, or the most productive combination) is not even present in the experiment.

If all combinations are not evaluated (i.e. planted at least once in an experiment), then it is impossible to conclude that the observed maximum production is the actual maximum production, especially at low levels of species richness, because there is a high probability that the most productive combination of species was not even planted. Assuming that there is a similar number of plots at each level of species richness, the proportional representation of all possible combinations changes dramatically and non-linearly with increasing species richness in this type of experiment (Fig. 5.3). This pattern is symmetrical, with high representation at low levels of species richness, and also at high levels (near the size of the total pool of species), but with extremely low representation at intermediate levels of richness. For example, at a treatment level of eight species drawn out of a pool of 18, 35 replicates represents 0.08% (35/43,758) of the total number of combinations. Obviously, the results produced by such a small random sample could change

dramatically from one experiment to another, and would be unlikely to capture any valid statistical properties of the actual distribution of performance.

Interpretation of this symmetrical distribution of representation across the treatment levels in a random-selection diversity experiment is further complicated by the 'variance reduction effect' (Huston 1997), which is another consequence of the same probabilities for randomly selected sets that produce the 'selection probability' or 'sampling' effect (Aarssen 1997; Huston 1997; Tilman *et al.* 1997b). Just as the probability of drawing (without replacement) any particular species out of a fixed set of species increases as the number of species drawn increases, the probability of the same combination of species appearing in different draws increases as the number of species drawn increases. The probability of all draws having the same species composition reaches 100% when the number of species drawn (without replacement) equals the total number of species in the pool from which the draws are made.

The variance reduction effect complicates the interpretation of representation and replication in these experiments, because both of these properties reach a maximum in treatments with high species richness. Consider the cases of two-species and 16-species treatments in a random-selection experiment with a total species pool of 18 (Fig. 5.3). Both have the same number of unique species compositions, 153, so if both treatment levels have the same number of randomly assembled plots, both will have the same proportional representation of all possible combinations. However, as a consequence of the variance reduction effect, the 35 plots with 16 species are much more similar to one another than the 35 plots with two species. This occurs because each of the 16-species plots contains most of the 18 possible species. Each plot must share a minimum of 12 species with each other plot, and any individual species is expected to occur in 31 out of the 35 replicates. In contrast, any particular two-species plot does not necessarily share any species with another plot, and each individual species is expected to appear in only four out of the 35 plots.

This pattern of increasing similarity (i.e. reduction in variance of species composition and related

properties) with increasing levels of species richness can be simply illustrated using Jaccard's Index of similarity, a standard method for evaluating the similarity of species composition of pairs of ecological communities (i.e. species lists). Jaccard's index is equal to $a_{11}/(a_{11} + a_{10} + a_{01})$, where a_{11} is the number of species shared between list 1 and list 2, a_{10} is the number of species that occur only in list 1, and a_{01} is the number that occurs only in list 2. The index varies between 0 and 1.0.

Figure 5.3 shows the monotonic increase in pairwise similarity among randomly drawn sets of species as the sets include larger numbers of species. Similarity is 0 among unique single-species sets, and 1.0 among all 18-species sets. The critical distinction is between the two-species and 16-species sets. Mean pairwise similarity among two-species plots is 7.6%, while mean similarity among 16-species plots is 80%. Although there is the same number of compositionally distinct combinations (153) at both treatment levels, all 16-species plots are very similar to one another, while all two-species plots are very different from one another.

Consequently, the 16-species plots are almost complete replicates of the same species composition, and should have many of the same properties as true replicates in plant growth experiments, possibly including a normal distribution of values, depending on the factors determining the experimental error (i.e. unexplained variance in a regression or within-subgroup mean squares in ANOVA). In contrast, all of the two-species plots are quite distinct from one another, and are unlikely to show the same response to error factors that would be observed in the 16-species plots.

Each unique species composition should produce a characteristic (or 'true') response under constant or homogeneous conditions. This true response could be estimated as the mean of replicate responses under heterogeneous conditions. Assuming that there is some normally distributed error applied to all plots in the experiment, this error should shift the value of each plot away from its true value, but only a small percentage of the plots should be shifted far away from their true value. Because two-species plots are so different from one another, they would be expected to have true values that cover a broad range, e.g. from high biomass to low

biomass. So, a normally distributed error should have little effect on the overall distribution, since most values that are displaced far from their true values are still likely to be within the wide range of values produced by the different species combinations.

In contrast, because the 16-species plots are very similar in species composition (i.e. nearly true replicates), they would be expected to have similar true values of the observed response. Application of a normally distributed experimental error to these plots is very likely to produce values outside the relatively narrow range of the true values. For example, with a normally distributed error with a mean of 1.0, one might have 10% of values in each tail of the distribution, e.g. below 0.7 or above 1.3. If the error is applied as a multiplier of the true value, as might be the case for error associated with variations in soil properties across an experimental field, this would result in an average of 3.5 high outliers and 3.5 low outliers out of 35 plots.

This produces what is perhaps the most insidious hidden treatment in diversity/productivity research: the increasing probability of interpreting experimental error as a treatment effect at higher levels of species richness. The probability of high (or low) outliers that are outside the range of true values is much higher for high richness treatments than for low richness treatments. The 'experimental error' that produces these outliers may come from some undetermined and apparently random source of variability, or from an easily understood source, such as slight variation in soil properties over the experimental field. Combined with the higher mean value caused by the sampling effect at high species richness, this variability is more likely to produce very high outliers at high levels of species richness. This may lead to the incorrect conclusion that the treatment has produced an increase in response (e.g. biomass or productivity) at high levels of diversity. In reality, the presumed treatment effect is an error effect.

This interaction of 'variance reduction' with 'representation' and 'replication' produces reversed monotonic gradients in the causes of variation within treatment levels (i.e. experimental error) across the treatment gradient. This reversal in the causes of within-treatment variance confounds the

interpretation and statistical analysis of random-selection diversity experiments. At low treatment levels of species richness, most of the variance in response is the consequence of differences in species composition among the quasi-replicates, and a relatively small proportion of the variance is due to other sources of variability, such as soil heterogeneity. In contrast, at high levels of the diversity treatment, very little of the variance in response is due to differences in species composition, because the quasi-replicates have converged with true replicates because of their high similarity in species composition. Thus, the primary source of response variance at high levels of species richness is not species composition, but rather whatever sources of error are present in the experiment, such as soil heterogeneity.

This phenomenon is clearly illustrated by the BIODEPTH experiments, where all plots of the highest diversity treatment were identical in species composition at the three sites designed to test for over-yielding (Portugal, Sheffield, and Sweden). At these three sites, each of the high diversity plots contained all of the species in the total species pool (i.e. all of the species that were grown as monocultures). Consequently, the variability in biomass production among the plots of the high diversity treatment could not have been due to differences in species composition, since all were identical in composition. This variability must have been caused by some other factor, of which underlying soil heterogeneity is the most likely candidate. The role of soil heterogeneity as the primary cause of variability in biomass production is strongly supported by the dramatically lower variability at Sheffield, the only site at which the soil was trucked in and homogenized (compare Portugal, Sheffield, and Sweden in Fig. 5.2). Note that the variability of biomass at the lower levels of species richness (1,2,4) is relatively similar at the three sites. This is expected because the primary source of variability in randomly selected small numbers of species is random differences in species composition among the plots. It is primarily at the highest level of species richness in these experiments, where all plots are identical, that the outliers from the error caused by soil heterogeneity across the field can be detected.

In summary, not only are the high diversity treatments most suitable for distinguishing the effects of soil heterogeneity (or other spatial factors) from the effects of species composition, they are also most susceptible to the confounding of error effects with treatment effects. The laws of probability produce a continuous gradient of: (1) increasing similarity among the quasi-replicate sets of species (e.g. different combinations of N species from a pool of 18 species) drawn at any given diversity level, and (2) increasing number of replicates of each particular subset of species (i.e. how many times exactly the same set of two or three species are drawn from 18) at any given diversity level, as the treatment level approaches the total number of species. The fact that high diversity plots are more similar to one another than are low diversity plots (the variance reduction effect) means that there are effectively more replicates of highly productive species combinations, and a greater chance of some of the replicates falling on portions of the field with anomalous soils. Therefore, there is a higher probability of misinterpreting the effect of soil properties, or other sources of experimental error, as treatment effects at high levels of species richness.

5.5 Simulating hidden treatment effects in a diversity–productivity experiment

One of the most contentious issues in the diversity–productivity controversy is whether the widely observed increase in mean productivity (or biomass etc.) with an increasing number of randomly selected species is (1) the consequence of interactions among many species, i.e. a true diversity effect caused by processes such as facilitation or complementary interactions among many species; or (2) the result of the chance presence of one or a few species, i.e. a statistical artefact of random sampling, called the selection probability (Huston 1997) or sampling (Tilman *et al.* 1997b) effect.

The sampling effect is the ‘the increasing probability of selecting species with a specific property (e.g. large maximum height, stress tolerance, nitrogen-fixation ability, high seed germination rate) in samples of increasing number that are randomly selected from any group of species’ (Huston

1997). This purely statistical phenomenon results from the random selection of any one, two, or more objects (e.g. species) from a group of any number of objects, according to the laws of probability. It can produce a pattern of increasing average productivity of mixtures with an increasing number of species as the result of (1) the increased probability of the inclusion of a single large or fast-growing species (Aarssen 1997; Huston 1997; Tilman *et al.* 1997b); or (2) the inclusion of specific groups of two or more species that interact positively either through facilitation (e.g. the fertilization effect of a nitrogen-fixing legume on a grass, Huston 1997; Hooper and Vitousek 1997; Grime 1997; Hodgson *et al.* 1998; Wardle 1999; Huston *et al.* 2000; Loreau 2000a) or through complementarity among groups of two or more species (i.e. niche effects) to produce more biomass than a smaller number of species could produce (Huston 1997; Tilman *et al.* 1997b; Hooper 1998; Loreau 2000a; Loreau *et al.* 2001; Fridley 2001; Dukes 2001b).

The two primary alternative hypotheses to explain the observed increase in mean performance with increasing species richness in random-selection diversity experiments are (1) the sampling effect for a small number of species (e.g. a highly productive individual species, or a highly productive species pair, such as a legume and a grass); and (2) multi-species complementarity (niche effects) among a large number of species. The definitive pattern for distinguishing between these two alternatives is not the increase in mean production, which is expected to be the same both for interactions among many species (niche effects) and for interactions among only a few species (sampling effects), but rather the shape of the upper boundary of the response distribution.

'With niche complementarity, the upper bound increases with diversity because no monoculture is as productive as some combinations of two species, and *no combination of N species is as productive as some combinations of N + 1 species*' (Tilman *et al.* 2001, italics added). This definition requires a continuous increase in the maximum value, as well as of the mean of the measured property (e.g. biomass production). In contrast, the 'signature' of the sampling effect is a constant upper bound for maximum biomass (or any other measured property) that

remains unchanged while the mean response increases with the number of randomly selected species.

The most conservative criterion for distinguishing between these alternative hypotheses is a metric from agricultural research known as 'transgressive over-yielding'. Transgressive over-yielding occurs when the productivity of some mixture of species exceeds the maximum productivity of the monocultures of all of the component species or of mixtures with fewer species (Trenbath 1974; Mead and Riley 1981; Vandermeer 1989; Garnier *et al.* 1997; Fridley 2001). For example, transgressive over-yielding above the two-species level occurs when some plots with more than two species exceed the maximum production of all plots with two species.

In the case of multi-species niche complementarity, there is over-yielding above each sequential level of species richness, beginning with monocultures, and continuing to the maximum number of species at which niche complementarity has a significant effect on production. For few-species sampling effects, there is no over-yielding above the number of species for which the sampling effect occurs. If the increase in productivity is determined by a sampling effect for one or more pairs of species (e.g. one or more specific legume-grass combinations), no plots with more than two species will exceed the highest production level achieved by two-species combinations. The most productive two-species combination(s) is assumed to be present in the most productive plots at all higher levels of species diversity.

In most of the published experiments of this type (Naeem *et al.* 1995; Tilman *et al.* 1996; Hector *et al.* 1999), as well as the first several years of the long-term experiment reported by Tilman *et al.* (2001), none of the multi-species treatments exceeded the maximum productivity of monocultures, which demonstrated that there was no over-yielding above the single-species level. This result was quite surprising, since several of these experiments included nitrogen-fixing legumes, which are known to increase the productivity of plants grown with them as a result of the nitrogen fertilization effect. There were undoubtedly positive interactions between legumes and other species, but none that resulted in transgressive over-yielding.

However, in the last few years of the longest-running diversity–productivity experiment, there was clear transgressive over-yielding above the maximum production of monocultures (Tilman *et al.* 2001), and published analyses demonstrate the expected strong effect of legumes (particularly *Lupinus perennis*) growing with C₄ grasses. The critical question for interpreting these results is whether or not there was over-yielding above the maximum productivity at the two-species level. If there was no over-yielding above the level of two species, the observed increase in mean productivity with increasing species richness is most likely the result of the sampling effect for the

combination of a legume and a grass, a common and well-understood response, rather than the niche effects of complementary interactions among many (e.g. 9 to 13) species, as concluded by Tilman *et al.* (2001).

There is no doubt that the most recent experimental results show several plots with above-ground and total biomass that exceed the maximum amount at the two-species treatment level (Fig. 5.4). However, there are only a few plots (~10% of the 94 plots with four or more species) that exceed the two-species overyielding level. Does this represent the increasing upper bound expected for multi-species niche effects, or just scattered outliers above

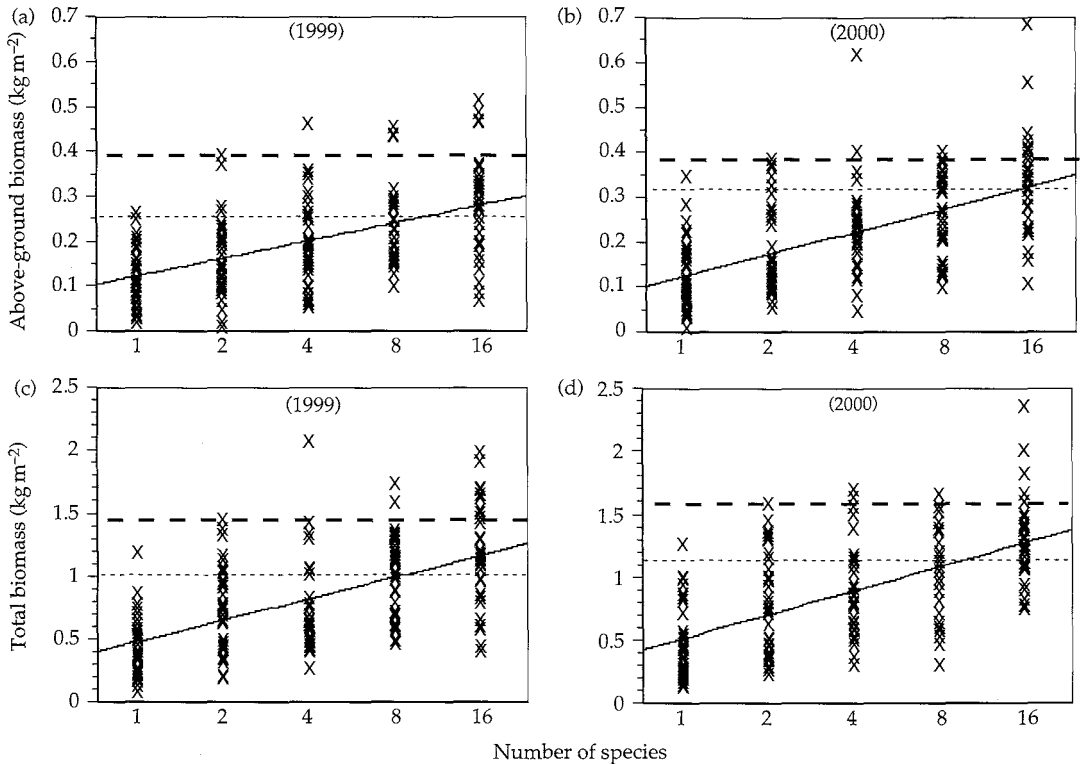


Figure 5.4 Biomass production of all replicates at five levels of species richness in a random-selection diversity experiment (Tilman *et al.* 2001). Data are shown for two years, and two measures of biomass: (a,b) above-ground biomass (approximating annual biomass production); (c,d) total biomass, including both above- and below-ground (the accumulation of roots which had grown over the 6–7 years of the experiment). Solid line indicates log-linear increase in average biomass with increasing number of species. Dotted line indicates the minimum biomass value for transgressive over-yielding above the level of one species (set at the mean of the two highest biomass values, because there are two replicates of each monoculture, and the mean is assumed to be the best estimate of the response of the most productive species). Upper dashed lines indicate the minimum biomass value for transgressive over-yielding above the level of two species (set at the level of the highest biomass value, because each of the two-species plots represents a unique species combination). The high biomass values in the 16-species treatments are most likely the result of experimental error caused by spatial variation in soil moisture or fertility across the experimental field (see text).

the maximum production level for the two-species mixtures as expected for the sampling effect?

We address this issue by simulating this experiment using the assumption that the *only interactions* among species are those between a legume and a C₄ grass. Of the 18 different species used in the experiment reported by Tilman *et al.* (2001), the only combinations that consistently produced over-yielding were those that included one or more of the four C₄ grasses plus one or more of the four legumes. To evaluate the consequences of the simplest possible explanation for the observed results, we restricted the mechanisms available to produce the treatment response to only the sampling effect for the presence of one legume plus one C₄ grass. Our question is whether this sampling effect alone can reproduce the observed responses, or whether some other mechanisms, such as multi-species niche effects involving more than two species, are necessary.

We assume that, in the absence of any positive or negative interactions with other species, each species has a characteristic maximum rate of productivity (or maximum accumulation of total biomass) that is equivalent to the maximum production of that species in monoculture (Table 5.1). This rate sets the maximum productivity of any mixture of non-interacting species to that of the species with the highest rate of productivity, which is the basic assumption for a single-species sampling effect. While we do not expect that the species with the highest monoculture productivity will dominate all mixtures or achieve the same biomass in mixtures as it did in monoculture, we assume

that a mixture will not be more productive than its most productive species (i.e. no over-yielding above the single-species maximum) *unless* facilitation or complementarity is occurring.

The only interaction among species in our model is the facilitation through nitrogen fertilization that occurs when a legume grows with a C₄ grass, so that the characteristic productivity of the grass is multiplied by some factor greater than 1.0 when a legume is present. Each of the four legumes species present in the experiment is assumed to have a different 'nitrogen fertilization multiplier', with the most abundant legume, *L. perennis*, having the highest multiplier (Table 5.1). If more than one legume and more than one C₄ grass species is present in a plot, we assume that the productivity of the plot is the rate of the most productive grass species multiplied by the highest fertilization factor among the legumes that are present. Thus, the productivity of all mixtures is constrained to result from a strictly two-species sampling effect.

Given these simple assumptions, we can ask what sorts of response patterns are likely to be observed. Each particular combination of species has a deterministic level of productivity, based on monoculture biomass that is augmented by the presence of legumes for C₄ grasses only (four species out of the total of 18). This mechanism produces the expected over-yielding above the single-species level and no over-yielding above the two-species level (Figs 5.5(a)–(c)). When there is a sufficiently high level of quasi-replication, the interaction of representation with variance reduction produces a maximum response that is identical at all treatment

Table 5.1 Parameter values used in simulation of diversity–productivity experiment. The total of 18 species includes four legumes and four C₄ grasses plus ten other species

Species type	Legume	C ₄ grass	Other species
Biomass (kg m ⁻²)	0.7 0.5 0.45 0.4	1.14, 0.9, 0.8, 0.4	0.45, 0.41, 0.37, 0.33 0.3, 0.28, 0.24 0.2, 0.16, 0.12
Legume multiplier	1.48 1.3 1.2 1.1		
Soil heterogeneity multiplier	Mean = 1.0, SD = 0.2		
Actual species number per treatment	16 species treatment: Mean = 11.5, SD = 1.4 8 species treatment: Mean = 7.0, SD = 1.4		

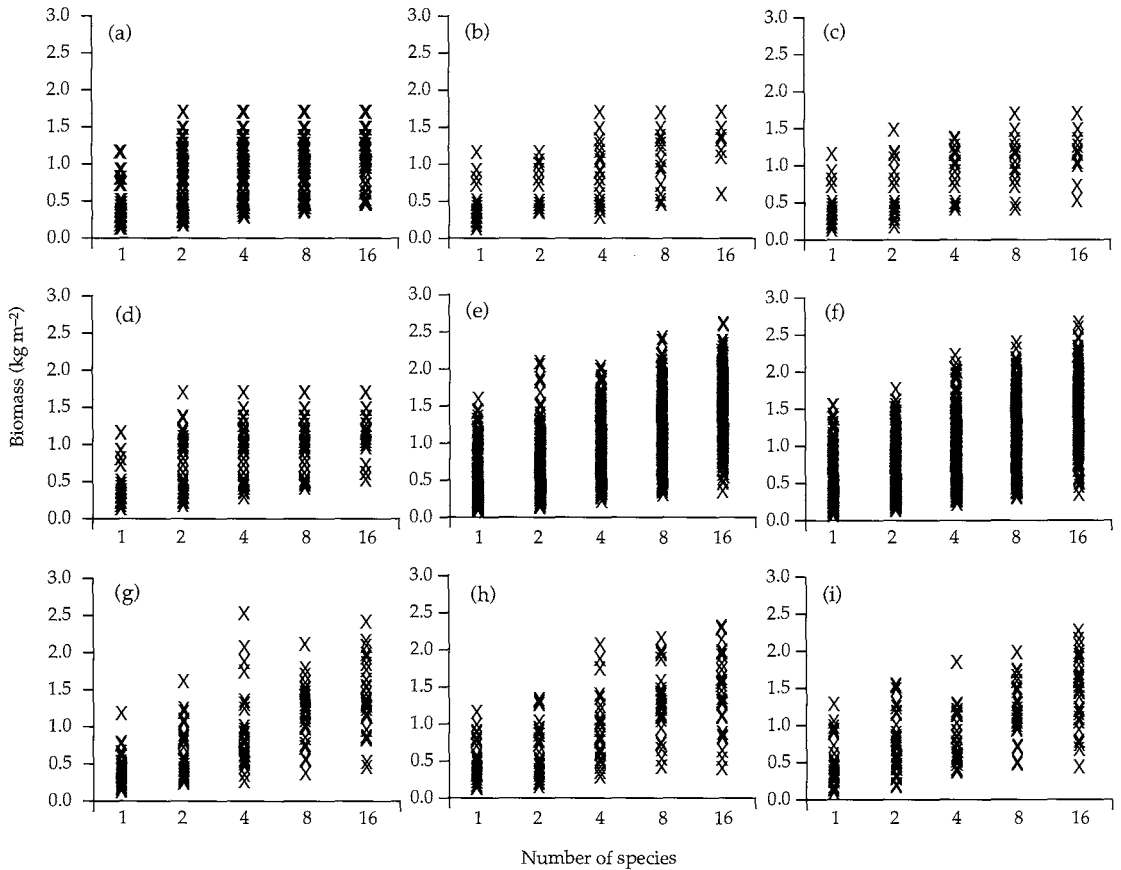


Figure 5.5 Results of simulation of a diversity–productivity experiment using the assumption that the most complex interaction is that between two species, specifically a C_4 grass and a legume. (a) 1000 quasi-replicates at each treatment level, note classic sampling effect of increasing mean with no values higher than the maximum production at the two-species level; (b) and (c) same model as (a), but with 36, 35, 30, and 35 quasi-replicates at treatment levels 1, 2, 4, 8, and 16, respectively; (d) same model as (a) with 300 quasi-replicates at each treatment level; (e) and (f) same model and number of quasi-replicates as (d) with multiplication by an error distribution with a mean of 1.0 and SD of 0.2 (simulating soil variation across experimental field), randomly applied to each quasi-replicate at each treatment level; ((g)–(i)) same model as (e) and (f), but the same number of quasi-replicates as in (b) and (c).

levels greater than one species (Fig. 5.5(a)). However, when quasi-replication is reduced to the low levels used in most experiments, the effects of random selection of species produce a highly variable treatment response (Figs 5.5(b,c)). The primary effect is that the most productive species compositions tend not to be represented in the small random samples (low levels of species richness). This generally leads to an increase in the number of high values (none of which exceed the true maximum productivity of two-species combinations) at high levels of species richness.

Addition of an error distribution, representing heterogeneity in soil properties across the experimental field, to the simulation produces a pattern that resembles true over-yielding, but is in reality the interaction of experimental error with the variance reduction effect. Figures 5.5(e) and (f) show the effect of random variation in soil conditions (field effects) added to the base case of deterministic two-species interaction, for 300 quasi-replicates (compare with Fig. 5.6(d), with 300 quasi-replicates but no error effect). Note that the number of high outliers increases at higher levels of species richness

because (1) the field effects are superimposed on a larger number of replicates because the quasi-replicates approach true replicates at the highest treatment levels as a consequence of the variance reduction effect; and (2) the replicates all have high rates of productivity because all contain at least one legume and at least one C_4 grass as a consequence of the sampling effect at the highest treatment levels. Note also that the actual pattern of outliers in any experiment will be determined by the actual distribution of soil properties across the experimental field, and will not necessarily match the normal distribution that we have used for generality. Because the pattern of soil properties is influenced by topography, soil depth, underlying geology, and other factors that have non-random spatial patterns (e.g. Robertson *et al.* 1988), the actual distribution is unlikely to be symmetrical.

The interaction of low levels of quasi-replication with field effects from soil heterogeneity produces great variety in the possible outcomes of the experiment, depending on which species combinations happen to be selected at low and intermediate levels of richness, and their random spatial locations in relation to soil properties (Figs 5.5 (g)–(i)). Many of these outcomes resemble the response observed in recent years in the Cedar Creek experiment, which shows only a few outliers that exceed the maximum productivity of two-species mixtures, most of which occur at the highest level of species richness (Fig. 5.4).

Thus, simple simulations based only on two-species interactions, combined with soil heterogeneity, reproduce the basic pattern of observed responses. This pattern occurs because the constant and low level of quasi-replication (e.g. ~ 35 plots per treatment level) results in very low representation of the most productive species combinations at low and intermediate levels of species richness, where similarity among the possible combinations is quite low. This correlation between treatment level and the causes of response variation is an inherent feature of random-selection diversity experiments. An increased number of plots (quasi-replication) at all levels of species richness, with highest numbers at the lower levels, would help solve the problem of inadequate representation. However, compensating for the increasing similarity among different

species combinations that is the inevitable consequence of the variance reduction effect would require sufficiently high levels of true replication at lower levels of species richness to sample soil heterogeneity and other sources of experimental error to the same extent that it is sampled by the high richness treatment, where all quasi-replicates are very similar in species composition. Such a design is likely to be logistically intractable. Alternative approaches include careful quantification of soil conditions in each plot for use as co-variables in analysis, and/or use of a large number of replicates of a few selected species compositions to quantify the spatial pattern of relevant soil properties across the experimental field.

5.6 Conclusions

The inherent statistical problems associated with random-selection diversity experiments make them highly susceptible to misinterpretation of cause-and-effect relationships. Hidden treatments that result from the combinatorial probabilities of random selection create internal patterns in the properties of the quasi-replicates that are correlated with treatment level, and thus can potentially be confused with treatment effects. The primary hidden treatments are: (1) the sampling effect which increases the probability of productive species combinations being present at high levels of species richness, and thus confuses cause-and-effect by misinterpreting the effect of a few species (the productive combinations of two or three species) as the effect of many species (the high diversity treatment level); (2) the quasi-replication effect, which results in under-representation of species combinations at low and intermediate treatment levels, and thus confuses cause-and-effect by underestimating the maximum productivity of low treatment levels because the most productive species combinations are more likely to be missing at low treatment levels than at high treatment levels; and (3) the variance reduction effect, which produces a continuous increase in similarity among quasi-replicates with increasing treatment level, causing the quasi-replicates to converge with true replicates and allowing the effects of experimental error, such as soil heterogeneity, to

appear most clearly at high treatment levels, thus potentially confusing cause-and-effect by interpreting experimental error as a treatment effect. None of these hidden treatments can be separated from presumed 'treatment effects' using multiple regression and ANOVA methods in the way that have been used in the past to analyse these experiments (cf. Schmid *et al.*, Chapter 6).

Given the inherent statistical problems with random-selection experiments that make them extremely difficult to analyse and interpret consistently, it does not seem advisable to proceed with more experiments of this design. Detailed retrospective analyses of the patterns of species abundance in these experiments are likely to lead to interesting insights (see Schmid *et al.*, Chapter 6). However, all conclusions and recommendations based on these experiments to date should be carefully re-evaluated.

Our understanding of the mechanistic basis of any relationships between species diversity and ecosystem function is likely to advance more rapidly using alternative experimental designs and analytic approaches. Experiments can be designed to

address specific mechanistic hypotheses using such approaches as single-species deletions (Fridley, in press), strategically nested designs (Dukes 2001b), and designs that explicitly manipulate the spatial and temporal variation in environmental conditions (Austin *et al.* 1988; Keddy *et al.* 1999; Dukes 2001b; Fridley, in press).

Maintaining appropriate levels of net primary productivity is essential for maintaining all life on Earth, and thus for preserving biodiversity. A better understanding of the global patterns and controls on net primary production in relation to the patterns of biodiversity can provide the basis for a scientific approach to conservation and sustainable resource management.

MAH was supported by the US Environmental Protection Agency's STAR Grant Program, and the US Department of Energy's Carbon Dioxide Research Program, Environmental Sciences Division, Office of Biological and Environmental Research. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the US Department of Energy under contract DE-AC05-00OR22725.

Biodiversity and Ecosystem Functioning

Synthesis and Perspectives

EDITED BY

Michel Loreau

Laboratoire d'Ecologie, Ecole Normale Supérieure, Paris, France

Shahid Naeem

Department of Zoology, University of Washington, USA

Pablo Inchausti

Laboratoire d'Ecologie, Ecole Normale Supérieure, Paris, France

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
Sao Paulo Shanghai Taipei Tokyo Toronto
with an associated company in Berlin

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Oxford University Press 2002

The moral rights of the author have been asserted

Database right Oxford University Press (maker)

First published 2002

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer

A Catalogue record for this title is available from the British Library

Library of Congress Cataloging in Publication Data

(Data available)

ISBN 0 19 851570 7 (Hbk)

ISBN 0 19 851571 5 (Pbk)

10 9 8 7 6 5 4 3 2 1

Typeset by Newgen Imaging Systems (P) Ltd, Chennai, India

Printed in Great Britain

on acid-free paper by The Bath Press, Avon

Contents

Contributors	x
PART I Introduction	
1 Biodiversity and ecosystem functioning: the emergence of a synthetic ecological framework	3
<i>S. Naeem, M. Loreau, and P. Inchausti</i>	
2 The debate on the role of biodiversity in ecosystem functioning	12
<i>H. A. Mooney</i>	
PART II Core areas of debate: biodiversity and ecosystem processes in grassland ecosystems	
3 Plant diversity and composition: effects on productivity and nutrient dynamics of experimental grasslands	21
<i>D. Tilman, J. Knops, D. Wedin, and P. Reich</i>	
4 Biodiversity manipulation experiments: studies replicated at multiple sites	36
<i>A. Hector, M. Loreau, B. Schmid, and the BIODEPTH project</i>	
5 Evaluating the relative strengths of biotic versus abiotic controls on ecosystem processes	47
<i>M. A. Huston and A. C. McBride</i>	
6 The design and analysis of biodiversity experiments	61
<i>B. Schmid, A. Hector, M. A. Huston, P. Inchausti, I. Nijs, P. W. Leadley, and D. Tilman</i>	
PART III New perspectives on ecosystem stability	
7 A new look at the relationship between diversity and stability	79
<i>M. Loreau, A. Downing, M. Emmerson, A. Gonzalez, J. Hughes, P. Inchausti, J. Joshi, J. Norberg, and O. Sala</i>	
8 Do species interactions buffer environmental variation (in theory)?	92
<i>J. B. Hughes, A. R. Ives, and J. Norberg</i>	

References

- Aarssen LW (1997). High productivity in grassland ecosystems: effected by species diversity or productive species? *Oikos* 80, 183-184.
- Al-Mufti MM, Sydes CL, Furness SB, Grime JP, Band SR (1977) A quantitative analysis of shoot phenology and dominance in herbaceous vegetation. *J. Ecology* 65, 759-91
- Austin MP, Fresco LFM, Nicholls AP, Groves RH, Kaye PE (1988). Competition and relative yield: estimation and interpretation at different densities and under various nutrient concentrations using *Silybum marianum* and *Cirsium vulgare*. *J. Ecol.* 76, 157 (1988).
- Berendse, F. 1994. Competition between plant populations at low and high nutrient supplies. *Oikos* 71, 253-260.
- Dawkins HC (1959). The volume increment of natural tropical high-forest and limitations on its improvement. *Empire Forestry Review* 38, 175-180.
- Dawkins HC (1964). The productivity of lowland tropical high forest and some comparisons with its competitors. *Oxford University Forestry Society Journal*, Fifth Series, 12, 1-8.
- deWit CT, Tow PG, Ennik GC (1966) Competition between legumes and grasses. *Versl. Landbouwk. Onderz.* 687, 3-30.
- Dukes, JS (2001). Productivity and complementarity in grassland microcosms of varying diversity. *Oikos* 94, 468-480.
- Fridley JD (2001). The influence of species diversity on ecosystem productivity: how, where, and why? *Oikos* 93, 514-526.
- Fridley JD (in press). Relative and interactive effects of resource availability and species diversity on ecosystem productivity in experimental plant communities. *Oikos*.
- Garnier E, Navas ML, Austin MP, et al. (1997). A problem for biodiversity-productivity studies: how to compare the productivity of multispecific plant mixtures to that of monocultures. *Acta Oecol.* 18, 657-670.
- Grime JP (1973a) Competitive exclusion in herbaceous vegetation. *Nature* 242, 344-347
- Grime JP (1973b) Control of species density in herbaceous vegetation. *J Environmental Management* 1, 151-167
- Grime JP (1979). Plant strategies and vegetation processes. Wiley, New York.
- Grime JP (1997). Biodiversity and ecosystem function. *Science* 277, 1260-1261.
- Grime JP (2001). Plant strategies, vegetation processes, and ecosystem processes. 2nd Ed., Wiley, New York.
- Harper JL. (1977). *Population Biology of Plants*. Academic Press, London.
- Hector A., Schmid B, Beierkuhnlein C et al. (1999). Plant diversity and productivity

- in European grasslands. *Science* 286, 1123-1127.
- Hodgson JG, Thompson K, Wilson PJ, Bogaard A (1998). Does biodiversity determine ecosystem function? The Ecotron experiment reconsidered. *Functional Ecology* 12, 843-856.
- Hooper DU (1998). The role of complementarity and competition in ecosystem responses to variation in plant diversity. *Ecology* 79, 704-719.
- Hooper DU, Vitousek PM (1997). The effects of plant composition and diversity on ecosystem processes. *Science* 277, 1302-1305.
- Hurlbert SH (1984) Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54, 187-211
- Huston MA (1979) A general hypothesis of species diversity. *American Naturalist* 113, 81-101
- Huston MA (1980) Soil nutrients and tree species richness in Costa Rican forests. *Journal of Biogeography* 7:147-157
- Huston MA (1993). Biological diversity, soils, and economics. *Science* 262, 1676-1680.
- Huston MA (1994). *Biological diversity: The coexistence of species on changing landscapes*. Cambridge University Press, Cambridge
- Huston MA, DeAngelis DL (1994). Competition and coexistence: the effects of resource transport and supply rates. *American Naturalist* 144, 954-977.
- Huston, MA, Aarssen, LW, Austin, MP, et al. (2000). No consistent effect of plant diversity on productivity. *Science* 289, 1255a.
- Keddy P, Twolan-Strutt L, Shipley B (1999). Experimental evidence that interspecific competitive asymmetry increases with soil productivity. *Oikos* 80, 253-256.
- Lawes JB, Gilbert JH, Masters MT (1882) Agricultural, chemical, and botanical results of experiments on the mixed herbage of permanent grasslands, conducted for more than twenty years in succession on the same land. Part II. The botanical results. *Phil Trans Royal Soc London A&B* 173, 1181-1423
- Loreau M (2000). Biodiversity and ecosystem function: recent theoretical advances. *Oikos* 91, 3-17.
- Loreau M, Naeem S, Inchausti P, Bengtsson J, et al. (2001). Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* 294, 804-808.
- Mahdi A, Law R, Willis AJ (1989). Large niche overlaps among coexisting plant species in a limestone grassland community. *J. Ecology* 77, 386-400.
- Mead R, Riley J (1981). A review of statistical ideas relevant to intercropping research. *J. Royal Stat. Soc. ser. A* 144, 462-509.
- Naeem S, Thompson LJ, Lawler SP, Lawton JH, Woodfin RM (1994a) Declining biodiversity can alter the

- performance of ecosystems. *Nature* 368, 734-737
- Naeem S, Thompson LJ, Lawler SP, Lawton JH, Woodfin RM (1994b) Biodiversity in model ecosystems (response). *Nature* 371, 565
- Naeem S, Thompson LJ, Lawler SP, Lawton JH, Woodfin RM (1995) Empirical evidence that declining species diversity may alter the performance of terrestrial ecosystems. *Proc Royal Soc Lond B* 347, 249-262
- Robertson GP, Huston MA, Evans FC, Tiedje JM (1988). Spatial variability in a successional plant community: Patterns of nitrogen mineralization, nitrification, and denitrification. *Ecology* 69, 1517- 1524.
- Rosenzweig ML (1971). The paradox of enrichment: destabilization of exploitation ecosystems in ecological time. *Science* 171, 385-387.
- Schwartz MW, Brigham CA, Hoeksema JD, Lyons KG, Mills MH, van Mantgem PJ (2000). Linking biodiversity to ecosystem function: implications for conservation ecology. *Oecologia* 122, 297-305.
- Sokol RR, Rohlf FJ (1981). *Biometry, The Principles and Practices of Statistics in Biological Research*, Second Edition, W.H. Freeman and Company, San Francisco.
- Tilman D (1987) Secondary succession and the pattern of plant dominance along experimental nitrogen gradients. *Ecological Monographs* 57, 189-214
- Tilman D (1988) *Plant strategies and the dynamics and structure of plant communities*. Princeton University Press, Princeton, NJ
- Tilman D (1996) Biodiversity: population versus ecosystem stability. *Ecology* 77, 350-363
- Tilman D, Wedin D, Knops J (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379, 718-720.
- Tilman D, Lehman C, Thompson K (1997). Plant diversity and ecosystem productivity: theoretical considerations. *Proc. Natl. Acad. Sci. USA* 94, 1857-1861.
- Tilman D, Reich PB, Knops J., Wedin D, Mielke T, Lehman C. (2001). Diversity and productivity in a long-term grassland experiment. *Science* 294, 843-845.
- Trenbath BR (1974). Biomass productivity of mixtures. *Adv. Agron.* 26, 177-210.
- Vandermeer J (1989) *The ecology of intercropping*. Cambridge University Press, Cambridge.
- Wardle DA (1999). Is "sampling effect" a problem for experiments investigating biodiversity-ecosystem function relationships. *Oikos* 87, 403-407.
- Wardle DA, Zackrisson O, Hornberg G, Gallet C (1997). The influence of island area on ecosystem properties. *Science* 277, 1296-1299.