Neutral Macroecology

Graham Bell

The central themes of community ecology—distribution, abundance, and diversity—display strongly marked and very general patterns. These include the log-normal distribution of abundance, the relation between range and abundance, the species-area law, and the turnover of species composition. Each pattern is the subject of a large literature that interprets it in terms of ecological processes, typically involving the sorting of differently specialized species onto heterogeneous landscapes. All of these patterns can be shown to arise, however, from neutral community models in which all individuals have identical properties, as the consequence of local dispersal alone. This implies, at the least, that functional interpretations of these patterns must be reevaluated. More fundamentally, neutral community models provide a general theory for biodiversity and conservation biology capable of predicting the fundamental processes and patterns of community ecology.

The linked themes of range, abundance, and diversity form the core of community ecology. Although they are very simple in themselves, they give rise to patterns that have engrossed the attention of ecologists for the last 50 years. The distribution of range and abundance among species, the relation between range and abundance, the variation of diversity among sites, and the increase of diversity with area, in particular, have been investigated many times. They form part of the field of macroecology (1), which is concerned with the description and interpretation of broad ecological patterns. Such patterns are usually held to be produced by the differing characteristics of species, which in a given environmental context cause one species to be common and another rare, or one species to be a specialist adapted to a narrow range of conditions, whereas another is a generalist that can be found everywhere. This being so, the contemplation of ecological patterns can be used to infer the nature of ecological processes. There is a deep flaw in this research agenda. It is now becoming clear that patterns indistinguishable from those generated by survey data emerge from community models in which all individuals have identical demographic properties. The behavior of these neutral community models will force us to revise the procedures of comparative analysis in ecology and, indeed, to reconsider large parts of the classical ecological curriculum.

Competition and Community Sorting

Consider a community of ecologically similar species, which will be defined for present purposes as a set of species, each interacting with at least one other, with all interactions being negative. This specifically excludes predator-prey, parasite-host, and mutualistic relations, and includes only those species that pursue similar ways of life and thus compete with one another for resources. Competition among species is the ecological equivalent of selection among genotypes, and is expected to have the same outcome—at equilibrium, the single best-adapted species will have replaced all others. This may not happen, however, if species are divergently adapted to different conditions of growth. There is then no unequivocally superior type, and a heterogeneous environment may support a diverse community. Each species will tend to predominate in the habitats where it grows better than any other and will tend to be eliminated from the rest. The spatial structure of the community will emerge in this way through a process of sorting, with the distribution of each species being the consequence of its unique combination of adaptations. The sites occupied by a species will then represent only a fraction of the environmental variance present in a region, and they will tend to be aggregated, because conditions of growth will tend to be similar in nearby sites. From these two properties flow all the familiar ecological relations, such as the distribution of abundance among species or the increase in species richness with area.

Neutral models refer to communities of ecological similar species in which individuals compete with one another and do not describe trophic interactions. A simple neutral community model (NCM) has five variables. These are the probabilities of birth b and death d for each individual, the probability of immigration m for each species, the number of individuals K in the community, and the number of species N in the external species pool. The model is set up by inoculating the community with a given number of individuals drawn at random from the external pool. It runs by iterating a series of four procedures. First, a single individual of each
species is added to the community with probability \( m \). Next, each resident individual gives birth with probability \( b \) and dies with probability \( d \). Finally, if the number of individuals in the community now exceeds \( K \), excess individuals are removed at random. To investigate species distributions, we need a spatial NCM. This comprises a network of sites, each supporting a simple NCM, connected by dispersal. Dispersal can be modeled by a random walk. With probability \( u \), a newborn individual moves to a random adjacent site and continues to move until the criterion fails, and it settles permanently in the site to which it moved last.

The NCM provides a simulacrum of an ecological community, within which the abundance and range of each species, and the diversity of each site, are completely known and can, if desired, be estimated by sampling. Beginning with an empty or randomly assembled community, the model usually approaches dynamic equilibrium after a few thousand cycles of birth-death-dispersal, after which there is little if any systematic change in patterns of abundance and diversity; the results presented here were all obtained after 2000 cycles. It can then be used to generate ecological patterns, which can be compared with those emerging from real biological surveys. The “normal configuration” of the model refers to a situation in which one new immigrant of each species is introduced every few generations, while each site is neither very isolated from its neighbors nor more or less completely connected with them; thus, \( m = 0.001 \) to 0.01 for a \( 50 \times 50 \) grid, and \( u = 0.01 \) to 0.1. These values generate realistic distributions of range and abundance, but more extreme cases can readily be considered.

The Distribution of Abundance

The first systematic attempts to account for the variation of abundance among species were based on the observation that abundance often seemed to follow some simple statistical distribution. Fisher et al. analyzed samples of insects and found that abundance was distributed geometrically (9). Preston noted that in other cases the most numerous category contained species of intermediate abundance, and proposed that abundance was distributed log-normally (10). In samples from log-normal communities, however, the distribution would appear to be skewed, because the rarest species in the community would be unlikely to be sampled. In thoroughly censused communities, the log-normal distribution does, in fact, seem to fit the survey data with remarkable precision (11). This might merely reflect the tendency of exponential processes influenced by many independent factors to lead to log-normal distributions (12). On the other hand, the form of the distribution of abundance might emerge from the nature of interactions between organisms and their environment, and this led to attempts to identify the ecological processes responsible for variation in abundance (13–15).

The distribution of abundance among species in simple communities has been described (16). With moderate rates of immigration, this resembles a log-normal distribution, skewed to the left to form a minor mode of rare species representing recent immigrants. As immigration increases, this mode becomes larger, until at very high immigration rates, it dominates the distribution, which now resembles a geometric or log-series distribution. The NCM thus explains both of the major patterns reported by previous authors. Hubbell has proven that both the skewed log-normal and the geometric are special cases of a single distribution, which he called the zero-sum multinomial (8).

One prominent feature of survey data is that abundance tends to be a consistent characteristic of species: If a particular species of understory herb is abundant in one patch of woodland, it is likely to be abundant also in another patch in the same region. This has led to strenuous attempts to identify the ecological characteristics responsible for the abundance or rarity of species. A consistent level of abundance, however, is characteristic of species in a neutral metacommunity. This can be evaluated by calculating the correlation of species abundance among sites. For moderate levels of local dispersal (\( u > 0.01 \)) this usually exceeds +0.8, and it falls to low values only when sites are almost completely isolated from one another. Thus, species that are abundant (or rare) in one part of the grid tend also to be abundant (or rare) in other parts. The reason is that a species that becomes abundant in any part of the grid will supply a stream of migrants to other parts, making it likely that the species will become established elsewhere. Species are thus expected to show consistent patterns of abundance and rarity except at very great spatial scales.

Distribution of Range

Geographical range can be expressed in several ways, but the simplest is the number of sites occupied by a species within a region. There has been general agreement that the distribution of spatial extent is a left-skewed
The Range-Abundance Relation

If each site were so small that it could support only a single individual of a given species, then range and abundance would be identical. As sites become larger and their populations increase, the two concepts become decoupled, but a correlation between range and abundance can be expected to persist and has often been observed (23) (Fig. 2A). Gaston lists nearly a hundred cases involving a variety of animals and plants; about 80% reported a significant positive correlation (24). The fundamental relation is between the number of species occupied and global abundance (total number of individuals occurring) within a region. Although the direction of the effect is well-established, the shape of the relation has aroused much less interest, in contrast to the species-area curve (but see (25)). For well-studied communities, however, it is often a power law. In British vertebrates, for example, power laws have exponents of 0.43 for birds and 0.37 for mammals (26). These relations are well-fitted, with up to about 80% of the variance of range explained by global abundance. A worldwide survey of wildfowl gave a similar value of 0.33, with 60% of the variance explained (27). The relation between range and local abundance (mean number of individuals per site) is also positive but is usually much weaker, with only about 10 to 20% of the variance in range explained. Consequently, even the shape of this relation is poorly documented, and its slope is unknown (although for British birds the data again suggest a value of 0.3 to 0.4 (28)). The ecological mechanisms responsible for these patterns have been the subject of much inconclusive debate. Gaston et al. identify eight hypotheses, including the connection between rarity and resource specialization, resource availability, habitat selection, and position within geographical range, but conclude that “no single mechanism has unequivocal support” (29).

The regression of range on global abundance in neutral community models is invariably positive. It is usually well-fitted by a power law, which explains about 90% of the variance. The exponent of this law depends primarily on the rate of local dispersal, and therefore also on the grain at which the analysis is conducted (Fig. 2B). In normal configuration, it is about 0.6 to 0.7 for fine-grained analyses based on about 1000 sites. At lower rates of dispersal, species are more highly aggregated, and any new individual added to a species population is likely to remain in its natal site; consequently, the exponent tends to be lower. The same reasoning applies to a coarse-grained analysis that combines adjacent sites into larger blocks; thus, with 50 blocks, the exponent falls to about 0.4. At very high rates of dispersal,
however, the most abundant species are present almost everywhere, so that the log range to log abundance plot is nonlinear, and the exponent of the power law is again relatively low. A similar pattern holds for the relation between range and local abundance, which is likewise positive although much less well fitted.

Geographical Variation in Abundance

Brown et al. identified a series of generalizations about the structure of species distributions (20). The abundance of a species tends to be similar in nearby sites; abundances do not usually change much over periods of 10 generations or so; and the patterns of abundance of closely related species are often quite dissimilar. They interpret these patterns in terms of “the influence of spatial and temporal variation in environmental variables on population dynamics.”

Although species tend to be consistently abundant or consistently rare, abundance is not a fixed property of a species. It is usually greatest near the center of the geographical distribution of a species; populations become fewer and smaller toward the edges of their range, until the species is eventually unable to maintain itself (30). Thus, local population density tends to decrease from the center of the range of a species outwards (Fig. 3A). Brown argues that the center of each species’ range is likely to provide the conditions to which it is best adapted, so that the tendency for the similarity of sites to decay with distance explains the observed pattern (1).

The distribution maps generated by the NCM share many properties with survey data. The abundance of a species at any given site is strongly correlated with its abundance at nearby sites, through local dispersal, and the spatial pattern of abundance is stable for tens or hundreds of cycles. Any two species may differ to any extent, paralleling the observation that closely related species often differ markedly in range and abundance. Species tend to be most abundant at or near the geographical center of their range, and mean density declines consistently away from this central region. This is the consequence of two phenomena: a weak tendency for abundance at occupied sites to decrease toward the edge of the range, and a strong tendency for the frequency of occupied sites to decrease (Fig. 3B).

The Species-Area Relation

All things being equal, diversity will increase with sampling effort. In most cases, the number of species recorded will increase steeply at first as more individuals are examined, but will then level off as a steadily decreasing number of rare species remain to be discovered. The exact shape of the curve depends on the distribution of abundance among species: It will be a negative exponential curve if abundance is log-series, and a power law...
phenomenon called "turnover" (change as one moves across a region, a
another. Thus, species composition will species found at one site do not occur at
ent, and it will become more likely that conditions of growth become more differ-
crease, as the distance between sites increases, the number of species recorded follows a power law with an exponent of about 0.25 (Fig. 4A). The slope varies with the extent of the survey and the type of area included (31), but values of between 0.1 and 0.4 are obtained in most cases (32).

Neutral community models give rise to a positive relation between species richness and area that is governed in most cases by a power law. At low rates of dispersal there are very few species per site, and the correlation between neighboring sites is low. Consequently, as neighboring sites are grouped into blocks, species diversity rises steeply, with an exponent of 0.6 to 0.7. As dispersal rates increase, the number of species per site increases, but the rate at which new species are added as area increases becomes less. This is partly because the total number of species in the region is fixed, and partly because increased dispersal causes neighboring sites to have more similar species composition. At a dispersal rate of $u = 0.01$, the exponent falls to about 0.4, and at $u = 0.1$, it falls further to values between 0.1 and 0.2 (Fig. 4B). At even higher rates of dispersal, the mean number of species in each unit site is greater, but the number in large blocks of sites may be less than at lower rates of dispersal. This is because extinction rates rise as the metacommunity becomes more highly integrated. The slope of the species-area curve continues to drop, however, and falls below 0.1 for $u = 0.5$. A power law fits the data very closely for all combinations of immigration and dispersal rates, except when both are high, and in consequence most species are found at most sites.

Turnover and Community Structure
As the distance between sites increases, conditions of growth become more different, and it will become more likely that species found at one site do not occur at another. Thus, species composition will change as one moves across a region, a phenomenon called "turnover" (33). It can be expressed in terms of the correlation of species occurrence or abundance between sites. This will tend to decay with distance at a rate characteristic of a particular kind of environment: A rapid rate of decay, for example, would signify a patchy, coarse-grained environment with distinct groups of specialists occupying each different kind of habitat. Now, the overall number of species in two (or more) sites is a total score that can be partitioned into the individual contributions of each site and their covariance (34). One consequence of turnover, therefore, is that the combined diversity of any two sites will tend to increase with the distance between them (Fig. 5A). There is thus a species-distance law, which, unlike the more familiar species-area law, is independent of the number of individuals sampled. The most thorough quantitative analysis of the species-distance relation to date concluded that pooled diversity generally increased with distance for most of 15 groups of plants and animals along a north-south transect in the British Isles (35).

In neutral community models, local dispersal creates correlation between nearby sites and thereby gives rise to patterns of species turnover. The specific correlation is large for adjacent sites, even when dispersal is low ($u = 0.001$). It decays rapidly, however, even for moderate rates of dispersal (up to $u = 0.1$), and reaches or approaches zero within the half-distance of the region. At very low dispersal rates there are 50% more species when adjacent sites are pooled, and species number is doubled for pairs of sites separated by the half-distance of the region. This implies complete replacement within the survey region, something observed only at large geographical scales. At very high dispersal, there are far more species, but turnover is very slight: Adjacent sites differ by only about 10% of species, and distant sites by scarcely more. At intermediate levels of dispersal, there are moderately large numbers of species and substantial turnover. Thus, at $u = 0.1$ species number per site is 0.4 to 0.5 $N$; it increases by about 20% when adjacent sites are pooled and by about 50% to 60% when more distant sites are pooled (Fig. 5B).

Specialization and Co-occurrence
The degree of specialization of a species can be evaluated from the environmental variance of the sites it occupies. Different species occupy different kinds of site, so that species within a clade diverge ecologically. The statistical properties of specialization and divergence in neutral communities seem to be surprisingly difficult to distinguish from real data, at least for small-scale surveys having extent about a thousand times larger than grain size (36). This strongly counterintuitive result is generated by the spatial autocorrelation created by local dispersal, and is not, of course, a property of random models.

If species occur at sites providing conditions to which they are adapted, species with similar adaptations will tend to occur together at the same sites. Thus, instead of a random collection of species, sites will tend to be occupied by one of a number of distinct assemblages each with its characteristic species composition. There are many ways of representing the tendency of species to occur together, but the simplest is just the distribution of a measure of correlation between all pairwise combinations of species. If distinct assemblages exist, there will be far more highly positive and highly negative correlations than would be expected by chance. The correlation between species occurrence in the NCM for moderate rates of dispersal is quite broadly distributed, with a standard deviation of about 0.2. A few percent of species pairs, therefore, have quite high correlation coefficients of $\pm 0.5$ or so. Randomized data, on the other hand, are much more narrowly distributed, with a standard deviation of about 0.05. The spatial autocorrelation of species occurrence thus gives rise to much stronger correlations between species than would be expected from a simple random model, just as it gives rise to unexpectedly strong correlations between species and conditions of growth. This shows, at the least, how random models do not provide appropriate null hypotheses for judging the spatial relations among species.
Conclusions
It has often been regretted that we lack a formal general theory of abundance and diversity that will account, in a simple and economical fashion, for the many patterns that ecologists have documented. The neutral community model provides such a theory. Because it is unfamiliar to most ecologists, and seems bizarre to many, it may be appropriate to consider some of the most frequently voiced objections to neutral theory in ecology.

The first is that it is contrary to fact: Reciprocal transplant experiments show that resident genotypes or species have greater fitness than incomers. This is certainly true of transplants involving large distances and different kinds of habitat, which can be expected to reveal some degree of local adaptation. Coconuts cannot be successfully established in boreal peat bogs. Transplants made over quite small distances have sometimes yielded the same result (37, 38). However, they have often failed to show any consistent superiority for residents (39, 40). Given the general reluctance to publish negative results, the available evidence does not strongly support a scenario of precise local adaptation over moderate distances within a single habitat.

Second, it is felt that neutral models take no account of the strong and complex interactions among organisms that we know to occur in nature. Where interactions such as predation, parasitism, or mutualism are concerned, this is perfectly true; the application of the theory is limited to ecologically similar organisms. With this reservation, however, the objection is unfounded. There are very strong interactions among organisms in neutral models, generated by the finite capacity of sites and the competition that this generates. These interactions are complex, insofar as one species may have an indirect effect on another by virtue of their mutual interaction with a third. The defining feature of neutral models is not that they lack interactions, but rather that these interactions occur among individuals with identical properties.

A third objection is that the models are too complex: The spatial NCM has at least six parameters, and with six parameters free to vary, any result could be obtained and any pattern generated. The answer to this criticism is that all community models of this general kind have the same number of parameters; they differ only in the number that can be tuned. A model in which the immigration rate (say) does not appear because it has been set at zero is not simpler than a model in which it is specified explicitly; it is merely less flexible. Seemingly simple equation-driven models such as the Lotka-Volterra systems commonly used in theoretical community ecology will continue to fill a useful heuristic role, but high-speed computing has made it possible to explore many important factors that they conceal. In this context, neutral models are not more complex, but actually much simpler than alternatives in which the distinctive properties of different species must be specified.

If such objections can be set aside, then the success of the NCM in predicting the major patterns of abundance and diversity has profound consequences for community ecology. These depend on whether the “weak” or the “strong” version of the neutral theory is adopted.

The weak version recognizes that the NCM is capable of generating patterns that resemble those arising from survey data, without acknowledging that it correctly identifies the underlying mechanism responsible for generating these patterns. The role of the NCM is then restricted to providing the appropriate null hypothesis when evaluating patterns of abundance and diversity. Even this relatively modest role, however, involves revising the comparative approach to ecology. Statistical null hypotheses based on randomization are not appropriate for evaluating ecological patterns that stem from species distributions, because local dispersal readily gives rise to spatial patterns. These patterns cannot be evaluated using standard statistical procedures, because of spatial covariance, and they often seem unexpected, perhaps because we are not accustomed to thinking in terms of spatially correlated phenomena. All the familiar patterns must be revisited, then, and their interpretations reviewed in the light of neutral theory. In my view, not many of these interpretations will survive this scrutiny. It is even possible that this exercise will lead to the same conclusion that was reached many years ago by population geneticists: that the contemplation of pattern is very unlikely to succeed in distinguishing between neutral and adaptationist theories of diversity.

The strong version is that the NCM is so successful precisely because it has correctly identified the principal mechanism underlying patterns of abundance and diversity. This has much more revolutionary consequences, because it involves accepting that neutral theory will provide a new conceptual foundation for community ecology and therefore for its applied arm, conservation biology. We shall have, for the first time, a general explanation for community composition and dynamics, as well as a synthetic account of a range of seemingly disparate phenomena. In practical terms, we shall be able to predict community processes such as the rate of local extinction, the flux of species through time and the turnover of species composition in space, in terms of simple parameters such as dispersal rates and local community sizes. The neutral theory of abundance and diversity will certainly have its limitations; adaptation is, after all, a fact, and the theory must fail at the taxonomic and geographical scales where specific adaptation has evolved. What these limitations are remains to be seen.

References and Notes
33. R. H. Whitaker, Communities and Ecosystems (Macmillan, New York, 1975).
42. This work was supported by grants from the Fonds pour les Chercheurs et à l’Aide de la Recherche (Quebec) and the Natural Sciences and Engineering Research Council of Canada.