

3

RELATIVE ABUNDANCE

In Chapter 2, we examined several sampling models for species diversity indices. These models began with an observed collection of individuals and species, and then sampled randomly to predict the diversity of smaller collections. These models accounted for differences in diversity between communities as sampling properties, so they were neutral with respect to species interactions. But what determines the species abundance distribution in the first place? A null model for species diversity must address the population processes that determine abundance, rather than just the sampling processes that might account for differences between collections.

A NULL MODEL FOR SPECIES DIVERSITY INDICES

Drawing inspiration from models of population genetics (Ewens 1972), Caswell (1976) pioneered an important approach to null models of species diversity. In models of population genetics, random extinction and mutation determine allelic frequencies in populations, and serve as benchmarks for gauging the effects of natural selection. In models of community assembly, random births and deaths within populations might determine relative abundances, and serve as benchmarks for gauging the effects of species interactions.

Caswell (1976) developed three such “neutral models,” all of which gave qualitatively similar results. In Model 1, there was no limit to population size. Species entered the community at a colonization rate v , and thereafter populations grew according to a stochastic model of exponential growth in which average birth and death rates were equal. In Models 2 and 3, total population size was bounded. In Model 2, a randomly chosen individual died at each time step and was replaced by a new individual, either of the same species or a “mutant” (i.e., a new species invading the community). Model 3 was similar. Each generation, a random sample of individuals was removed from the population and replaced with individuals from the previous generation, each of

which could mutate to a new species. Although Models 2 and 3 were somewhat constrained by the cap on total population size (Ugland and Gray 1982a), changes in population size of species were independent (Caswell 1983). Model 1 was computationally simpler, and was used in comparison with several empirical data sets.

Caswell (1976) first compared the shape of the species abundance curve produced by Model 1 to the log series, log normal, and broken-stick distributions. The curve for the neutral model followed a log series distribution, although its shape was determined somewhat by ν , the rate at which new species entered the community. At large values of ν , the neutral model curve tapered off and began to resemble a log normal curve, at least up to the mode of the distribution. Because it generated plausible species abundance curves, Model 1 was a reasonable null model for species diversity indices.

Caswell (1976) modified formulae in Ewens (1972) to generate the expected species diversity (H') and its variance for communities with a given number of individuals and species. A standardized deviate, V , quantified the extent to which observed diversity was above or below the predictions of the neutral model. Because the neutral model was predicated on the number of species in the collection, variation in V reflected changes in relative abundance that were not confounded with changes in species richness. V showed only modest dependence on sample size, probably because abundance and species richness were underlying components of the model. Calculation of expected diversity in the neutral model is cumbersome, but Goldman and Lamshead (1989) presented a modified algorithm that is suitable for personal computers.

Caswell (1976) used the neutral model to compare observed and expected diversity levels for published data sets that included complete counts of individuals (or pairs of breeding birds) in well-sampled communities. His purpose was to test whether species diversity differed systematically between simple and complex communities. If communities were self-regulated by internal dynamics, diversity should have been highest in stable, late successional, and undisturbed systems (Margalef 1968). For each assemblage, V was plotted against successional stage (or an ordinated continuum index) for different communities.

Caswell's (1976) results refuted conventional wisdom about the diversity of complex, species-rich communities. For the successional data, V fluctuated between positive and negative values. There was no tendency for diversity to increase with successional stage, and in many cases diversity was substantially reduced in "climax" communities (Figure 3.1). For temperate bird and tree communities, diversity was usually greater than that predicted by the neutral model ($V > 0$), whereas for tropical communities, diversity was less than

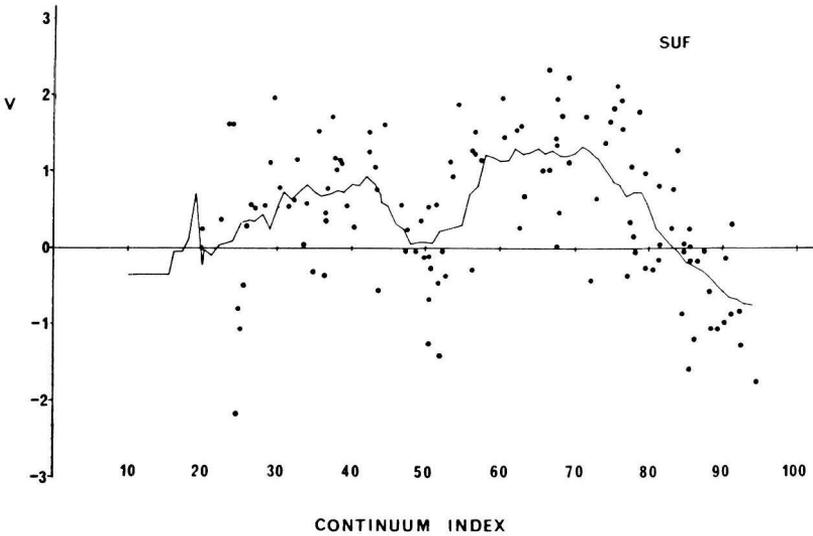


Figure 3.1. Expected and observed species diversity as a function of successional development in upland forests of southern Wisconsin (data from Auclair and Goff 1971). Each point is a different site, and the line is a 10-point running average. The expected diversity from the neutral model is $V = 0$. Note the downward trend for diversity of late-successional sites. From Caswell (1976), with permission.

expected ($V < 0$). For stream fish assemblages, diversity was lowest in stable, high-order streams.

It is important to note that these patterns were not detected in the original diversity indices— H' values were typically higher for late successional, tropical, and high-order stream assemblages. The large H' values reflected the greater species richness in these assemblages, but Caswell's (1976) analysis revealed that relative abundances were actually less even than expected. If a general inference can be drawn from these results, it is that species interactions tended to reduce diversity; species diversity may be highest in nonequilibrium assemblages that are frequently disturbed (Sousa 1984).

In other studies, Caswell's (1976) neutral model accurately described species diversity. For example, the diversity and species abundance curves for bacterivorous ciliates were similar to the neutral model predictions (Taylor 1979), suggesting that populations grew stochastically, exponentially, and independently of one another. For a community of tropical beetles sampled on an elevational gradient, Hanski (1983) found that evenness increased with species richness, and that relative abundances of some genera were more even than predicted by the log series and hence the neutral model. He claimed that

interspecific competition was important in this assemblage, and that species were “tightly packed” along an elevational gradient.

A more typical pattern is for diversity to fall short of the neutral model predictions, as exemplified by Platt and Lamshead’s (1985) analysis of 98 data sets for marine benthic assemblages. Diversity was usually higher in disturbed than undisturbed assemblages. Within an assemblage, diversity was usually higher after a disturbance than before. Other analyses of marine communities sampled along disturbance gradients also reveal higher diversity in more disturbed environments (Rainer 1981; Warwick and Gee 1984; Lamshead and Gooday 1990; Absalao 1991). These results suggest that it may be difficult to maintain high evenness in the absence of disturbance, and underscore the importance of analyzing abundance data with an appropriate null model.

SPECIES ABUNDANCE MODELS

If individuals in a community are randomly sampled, we usually find that most species are rare and a few species are common. What causes this pattern? That is, what forces determine the shape of the species abundance curve? MacArthur (1957, 1960) suggested the simplest null hypothesis: species abundances are entirely independent of one another and hence are sampled randomly from a uniform distribution (Pielou and Arnason 1966). Most real communities are not this even, but it has been difficult to understand why.

The study of species abundance distributions began as a statistical characterization of large samples of individuals. Fisher et al. (1943) successfully fit a log series distribution to collections of moths sampled at light traps. Preston (1948) found that a log normal distribution characterized large samples of bird and plant communities. The canonical log normal is a special form of this distribution in which the mean and variance are related to one another (Preston 1962); the canonical log normal distribution of species abundances forms the basis for quantitative predictions of the slope of species-area relationships (Preston 1948; MacArthur and Wilson 1967; see Chapter 8).

Neither the study of Fisher et al. (1943) nor that of Preston (1948) initially attached theoretical significance to these distributions; instead, they were simply used to characterize and describe species abundance patterns. Although the log series and the log normal distribution are common models of species abundance data (Figure 3.2), neither provides a perfect fit to data sets. For the log series, Fisher et al. (1943) did not count all individuals of the most common species, so the distribution was characteristic mostly of the rare species in the

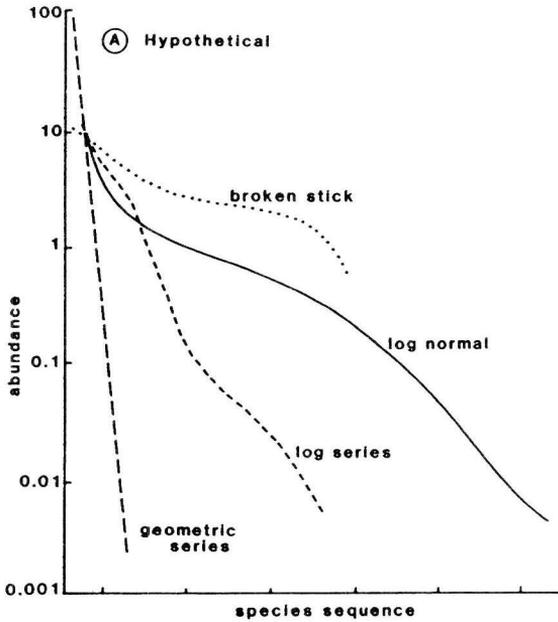


Figure 3.2. Hypothetical rank abundance plots for common species abundance models. From Magurran, A. E. *Ecological Diversity and Its Measurement*. Copyright © 1988 by Princeton University Press. Reprinted by permission of Princeton University Press.

assemblage (Hughes 1986). For the log normal, the fit is often biased in the tail of the distribution (Preston 1981).

In theory, it should not be difficult to distinguish between these distributions. The log series predicts that the rarest species should occur most frequently in a sample, whereas the log normal predicts that species of intermediate abundance should be most common. However, small samples from a log normal that fall to the right of Preston's (1948) "veil line" may be indistinguishable from the log series (Figure 3.3). The log normal is also difficult to distinguish from MacArthur's (1957) broken-stick distribution when sample sizes are small (Wilson 1993).

What factors might account for a log series or log normal distribution? If species are responding independently to different factors and their responses are expressed as differences in exponential growth, a log normal distribution will result (MacArthur 1957; May 1975a). For this same reason, the log normal distribution characterizes many nonbiological systems, such as the distribution of gross national products of different countries (May 1975a; Preston 1981). If

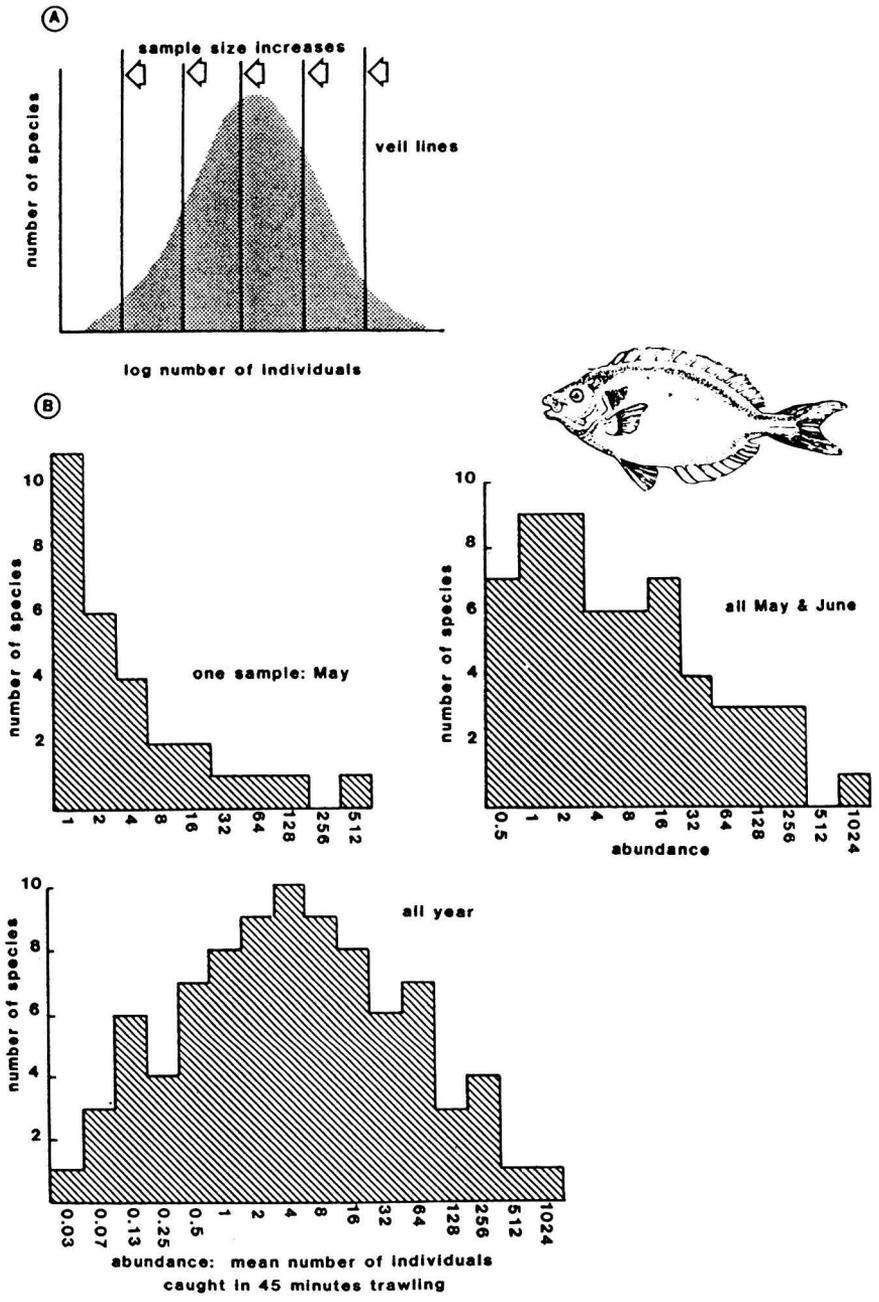


Figure 3.3. Sampling effects on the shape of the species abundance distribution. For the single (May) sample of fish diversity in the Arabian Gulf, the curve resembles a logarithmic series or a geometric series, but as more samples are added, the mode of the log normal distribution is revealed. From Magurran, A. E. *Ecological Diversity and Its Measurement*. Copyright © 1988 by Princeton University Press. Reprinted by permission of Princeton University Press.

species populations are at an equilibrium in small patches, the aggregate of their distribution may also follow a log normal (Ugland and Gray 1982b).

In contrast to these “statistical” explanations, Sugihara (1980) argued that the log normal distribution reflected hierarchical niche partitioning in natural assemblages. However, there are logical difficulties with this argument (Kolasa and Strayer 1988), and a set of species independently selecting hierarchical elements of habitat may lead to the same pattern (Kolasa and Biesiadka 1984). Similarly, Hughes’s (1984) “dynamics model” is based on stochastic processes of survivorship, recruitment, and species interaction, and seems to explain the concavity of some species abundance curves better than either the log normal or the log series (Hughes 1986). However, the parameters in the dynamics model have never been confirmed empirically, and the same distributions can be explained by sampling effects (Barangé and Campos 1991).

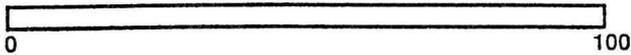
THE BROKEN STICK

MacArthur (1957, 1960) pioneered a novel approach to species abundance distributions. As a simple null model of species abundance, he imagined a one-dimensional resource distribution that was simultaneously fragmented at randomly located points. The length of each resulting piece was proportional to the abundance of a species. When the species were ranked from most common to least common, they formed the “broken-stick” distribution (Figure 3.4). Biologically, the broken stick corresponds to a community in which all species colonize simultaneously and partition a single resource axis randomly.

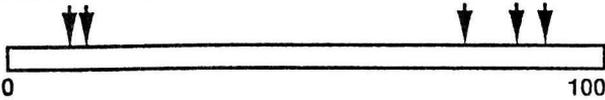
The broken stick can also be interpreted as a model of sequential colonization, with each species randomly invading some proportion of the niche of established species. For the resident species, the probability of invasion by a colonizer is proportional to the fraction of niche space the resident has sequestered. Any species can have its abundance reduced by an invader, but common species are more susceptible to invasion than rare species (Tokeshi 1990). Barton and David (1956) described the statistics used to calculate the expected segment lengths, and the broken-stick model has been used as the basis for statistical tests of overlap of species niches (Chapters 4 and 5), geographic ranges (Chapter 9), and body sizes (Chapter 6).

The broken stick incorporates competitive effects but is null with respect to how resources are partitioned. However, the same distribution can result from a model of incomplete niche partitioning or a model in which individuals are assigned randomly and equiprobably to different species (Cohen 1968). This

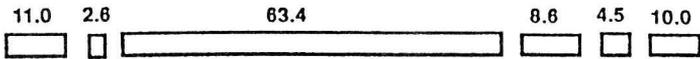
(a) A stick, here 100 units long, represent a resource gradient .



(b) For a 6-species (n) community, 5 ($n-1$) random throws are made at the stick .



(c) The stick is broken at each point that a throw landed. The 6 segments of stick represent the 6 species, the length of each segment represents the fraction of the resource used by that species, and hence its abundance.



(d) In a ranked-abundance plot, the 6 species are arranged in sequence of decreasing order of abundance, and the abundances plotted on a log scale.

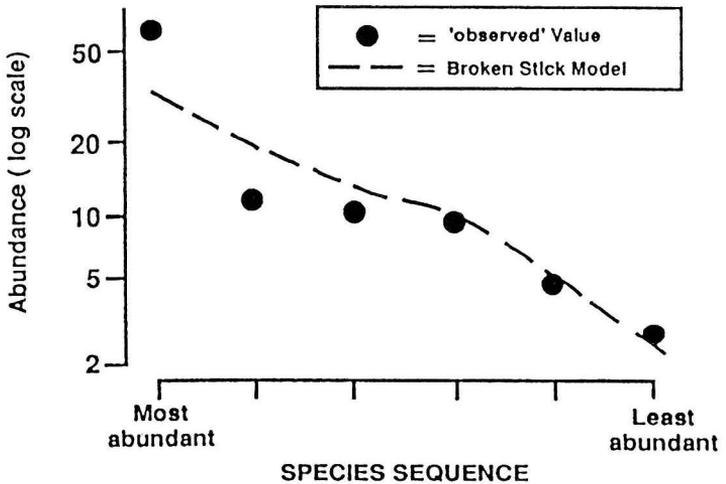


Figure 3.4. Simulation of MacArthur's (1957) broken-stick model. From Wilson (1993), with permission.

problem characterizes not just the broken stick but all the other models that have been fit to species abundance data. For each model in which biological effects can generate the pattern, there is a corresponding statistical or non-interactive interpretation (Table 3.1).

Table 3.1

Biological and statistical interpretations of common species abundance distributions

Distribution	Biological interpretation	Statistical interpretation
Broken stick	A one-dimensional resource axis is simultaneously and randomly broken (MacArthur 1957), or the breaks are sequential and proportional to segment length (Caswell 1976; Tokeshi 1990).	Individuals are randomly assigned to species (Cohen 1968).
Uniform	Species use resources independently of one another (MacArthur 1957).	Individuals of each species are sampled from an equiprobable, underlying distribution (Pielou and Amason 1966).
Log series	Each species arrives at random time intervals and preempts a constant fraction of the remaining resources (May 1975a). Characterizes samples from small, stressed, or pioneer communities (Whittaker 1972; May 1975a).	Sampling or stochastic effects (Boswell and Patil 1971), small samples from a log normal community (Preston 1948), or a noninteractive community with independent birth and death rates and a high rate of species immigration (Caswell 1976).
Geometric series	Each species arrives at regular time intervals and preempts a constant fraction of the remaining resources (Motomura 1932; May 1975a).	Same as for log series. Also, species abundances are sequentially ordered and each is a random fraction of the previous species (Tokeshi 1990).
Log normal	Hierarchical niche subdivision, in which the probability of breakage is independent of segment length (Sugihara 1980), or an assemblage of species that specialize on different elements of habitat, which is subdivided hierarchically (Kolasa and Biesiadka 1984; Kolasa and Strayer 1988). Characterizes large, stable assemblages at equilibrium (Whittaker 1972; May 1975a).	Species populations grow exponentially and respond independently to different factors (MacArthur 1960; May 1975a), or an aggregate of species populations that are at equilibrium in small patches (Ugland and Gray 1982b).
Dynamics	A dynamic model that incorporates parameters for survivorship, recruitment potential, and gregariousness (Hughes 1984).	Sampling effects and aggregation of heterogeneous samples (Barangé and Campos 1991).

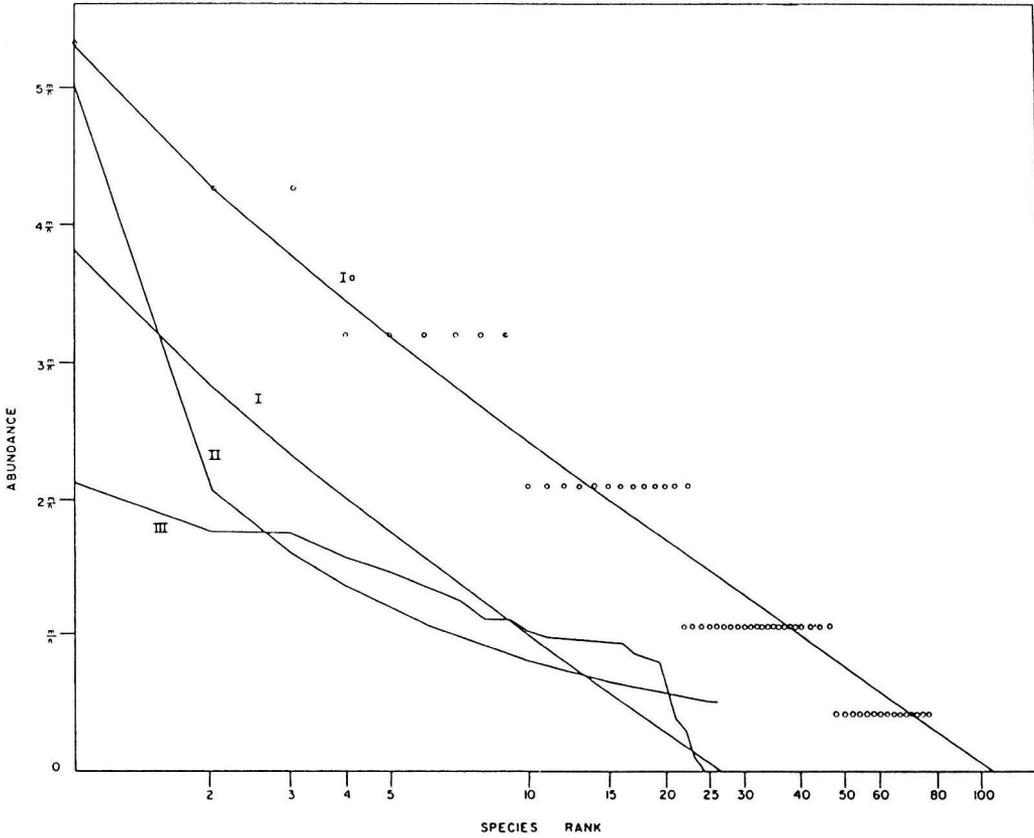


Figure 3.5. MacArthur's (1957) analysis of bird census data (106 species; open circles) from Pennsylvania. Curve Ia = broken stick expectation for an assemblage of 106 species. Curve I = broken stick (25 species). Curve II = expectation for a random uniform distribution (overlapping niches). Curve III = Monte Carlo simulation of random assignment of individuals to species (particulate niches). From MacArthur (1957).

MacArthur (1957) originally applied the broken-stick model to a census of temperate forest birds (Figure 3.5). The broken stick provided a better fit to the data than either a uniform distribution or a Monte Carlo simulation in which individuals were randomly apportioned to species (but see Cohen 1968). Subsequently, contradictory evidence for and against the broken stick accumulated (King 1964), and by 1966, MacArthur expressed a wish for the model to “die a natural death.”

A major stumbling block to adequately testing the broken stick is that the model specifies that all possible partitionings of relative abundance are equally likely (Pielou 1975). Consequently, it is not appropriate to test the species abundance curve of a single assemblage against the predictions of the broken-stick model (Smart 1976; Pielou 1981). However, it is possible to measure the relative fit of a single data set to several species abundance models (Wilson 1991b), as MacArthur (1957) did. If the log series or log normal is among the alternative models, it may not be possible to recognize the broken stick unless there are many species in the sample (Wilson 1993).

THE RELATIONSHIP BETWEEN RESOURCE USE AND RELATIVE ABUNDANCE

The broken stick is only one of several models of one-dimensional resource partitioning (Sugihara 1980; Tokeshi 1990). All such models make the critical assumption that resource consumption is proportional to relative abundance. However, because large-bodied animals are usually rare (Damuth 1981, 1987) and have high per capita resource requirements (Elgar and Harvey 1987), resource use may not be proportional to abundance. In particular, a log normal distribution of abundances need not reflect a log normal distribution of resource use (Harvey and Godfray 1987).

Sugihara (1989) showed for a few data sets that variances in biomass and abundance were approximately interchangeable. However, a more detailed allometric analysis revealed that variances in biomass were usually greater than variances in abundance, so that resource use may be even more inequitable than suggested by relative abundances (Pagel et al. 1991a). Size distributions in assemblages or guilds of related species are often complex and polymodal (Warwick 1984; Griffiths 1986), which may contribute to high variance. Null model analyses also confirm that results are often sensitive to whether biomass or abundance data are used (Rainer 1981; Tokeshi 1990). Consequently, explanations for species abundance data that are based on resource partitioning must be evaluated carefully, and supported by data that establish the relationships between relative abundance and resource use.

NULL MODELS FOR RESOURCE PARTITIONING

Given all these difficulties, what is the best strategy for analyzing resource-partitioning models such as the broken stick? We suggest three guidelines for analysis. First, analyses should be restricted to small guilds of common, potentially interacting species. Although the log normal and log series distributions may characterize large samples that include many rare species, it seems unlikely to us that rare species are often important in resource partitioning. Restricting the analysis to guilds may also ensure that the species are not grossly different in body mass and per capita resource consumption. Second, analyses should be based on replicated samples of an assemblage. These can be averaged so that variability in relative abundances can be estimated and sampling effects minimized. For the broken stick, in particular, it is important to compare the model with the average species abundance curve of several assemblages. Third, Wilson's (1991b) procedures for measuring relative fit of the data to several alternative models should be used.

An exemplary study by Tokeshi (1990) followed all three of these guidelines. Tokeshi (1990) examined the relative abundance patterns for six common, closely related species of epiphytic chironomids that comprised 95% of the individuals and the biomass in a typical sample. The study site was a small English river that was sampled repeatedly through the year. Tokeshi (1990) compared average abundance and biomass distributions with the predictions of seven resource-partitioning models:

1. *Geometric series.* This was the only strictly deterministic model in Tokeshi's (1990) analysis. The model assumes that species arrived at regular time intervals and sequestered a constant fraction k of the remaining resources. To fit the geometric series, Tokeshi (1990) chose values of k that minimized the squared deviations between observed and expected relative abundances.
2. *Dominance preemption.* This is a more general form of the geometric series. Each species entered the assemblage and took some random fraction of the remaining resources. Because species sequestered only unused resources, each invader was competitively inferior to all previous invaders. The fraction of resources used was a random proportion drawn from the interval 0.5 to 1.0, so the expected distribution converged on a geometric series with $k = 0.75$.
3. *Random fraction.* This model envisions a completely random partitioning of niche space. The line segment was randomly broken

and one of the two segments was chosen (with equal probability) for a second break. One of the three resulting segments was again chosen randomly and equiprobably for breakage, and so on. This model yields an approximate log normal distribution for large sample sizes (Pielou 1975), and corresponds to Sugihara's (1980) description of a community with hierarchical niche partitioning. Note that the dominance preemption represents a special case of this model in which the largest segment was always chosen for subdivision.

4. *Broken stick.* Each species that entered the assemblage took a random fraction of the resources of the previous invaders. The probability that the niche of a resident species was invaded was proportional to its resource utilization.
5. *Dominance decay.* Another special case of the random fraction model, the dominance decay model is the exact opposite of the dominance preemption model. In the dominance decay model, the largest existing segment was always chosen for random breakage. Ecologically, this means the most abundant species in the assemblage always had its niche invaded. In the broken-stick model, the most abundant species was more likely to be chosen, but this was not a fixed rule. As a consequence, the distribution of relative abundances under the dominance decay model is more equitable than under the broken-stick model.
6. *Random assortment.* Tokeshi (1990) envisioned random assortment as a null model for niche partitioning, because the abundances of the component species were unrelated to one another. This scenario might occur if resources were nonlimiting, or if resources were limiting but the assemblage was subject to frequent random disturbances so that it never reached an equilibrium. Tokeshi (1990) reasoned that if the species were ranked from most to least abundant, then the abundance of each species would be some arbitrary fraction of the abundance of the species that preceded it. Hence, the random assortment model is a stochastic analog of the geometric series with $k = 0.5$.
7. *Composite.* In this model, the first two species competitively divided the resources (dominance preemption), whereas the abundances of the remaining species were determined by random assortment.

These seven models generated a family of species abundance curves that differed in their equitabilities (Figure 3.6). The dominance decay and broken-

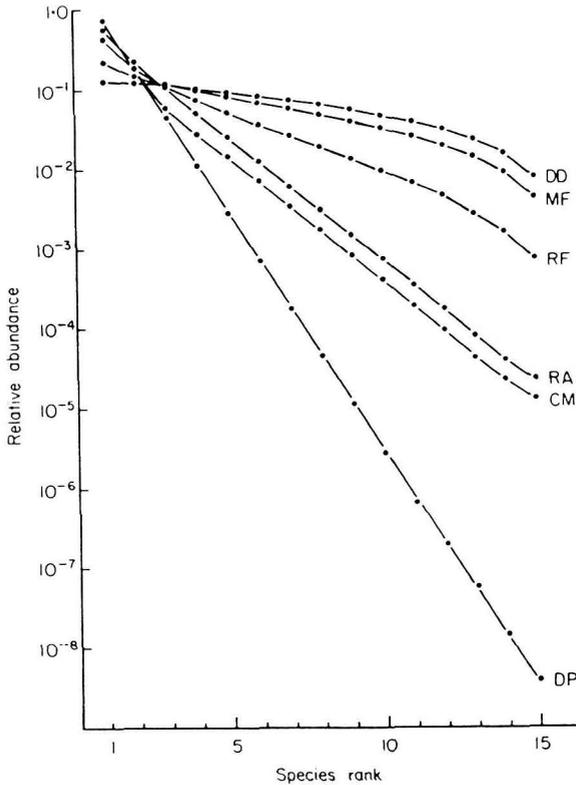


Figure 3.6. Expected rank abundance patterns for different models of resource partitioning. DD = Dominance Decay; MF = MacArthur Fraction (= Broken Stick); RF = Random Fraction; RA = Random Assortment; CM = Composite Model; DP = Dominance Preemption. Each point is the average from 1000 simulated distributions. From Tokeshi (1990), with permission.

stick models gave the most even distribution of expected abundances, whereas the dominance preemption model generated the steepest curve. With the exception of the geometric series, Tokeshi's (1990) models all included stochastic elements, so the expected abundances and confidence intervals were estimated with one thousand simulations.

Relative abundances of chironomids were more equitable than predicted by the dominance preemption model (2), but less equitable than predicted by the broken-stick model (4). Only the random fraction (3) and random assortment (6) models fit the data well, with the observed relative abundances falling within the 95% confidence interval for all six species (Figure 3.7). In contrast, the biomass data matched the predictions of only the random assortment model (6).

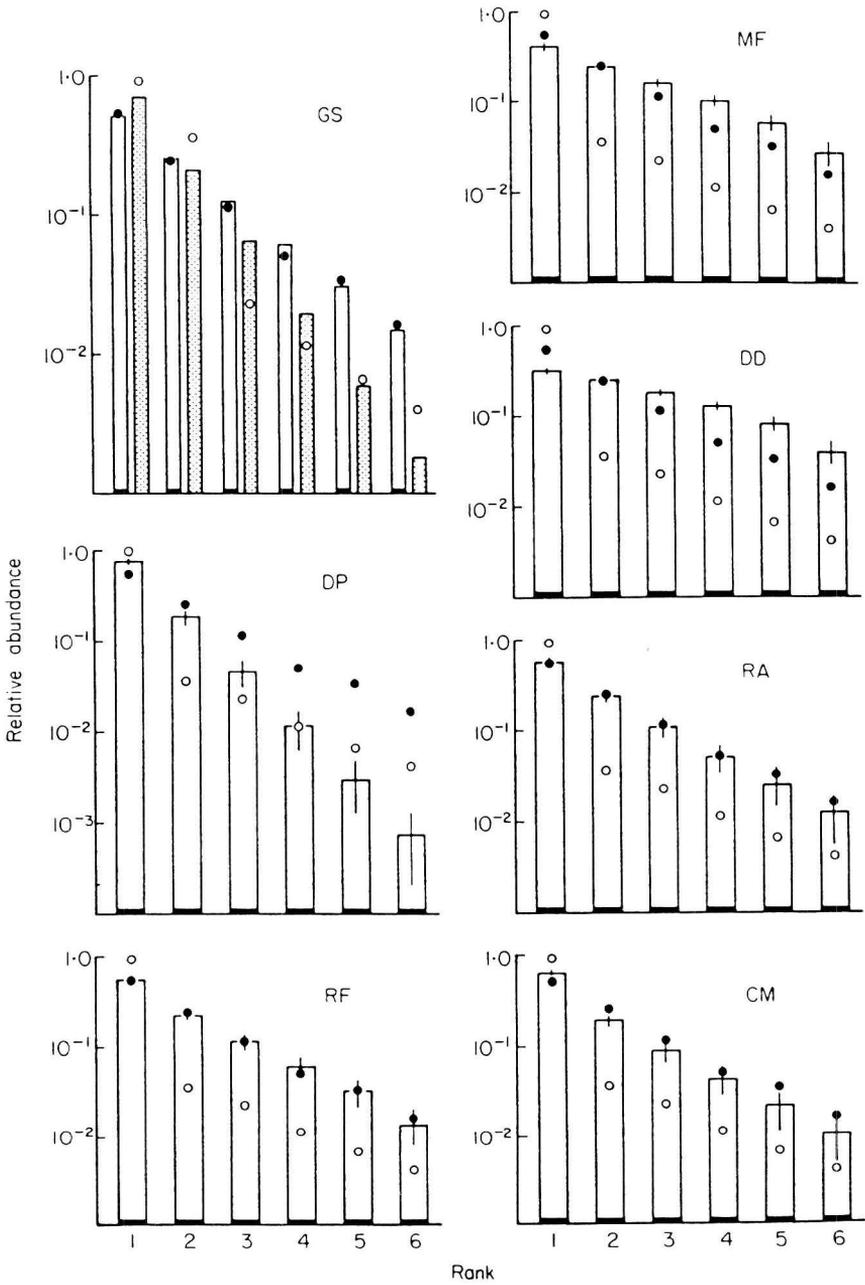


Figure 3.7. Fit of chironomid relative abundance data to the null models of Figure 3.6. Vertical bars = expected values; circles = observed values, averaged over several samples. ● = homogeneous summer data; ○ = heterogeneous summer/winter data, which was poorly fit by all models. Abbreviations as in Figure 3.6. For the Geometric Series (GS), a separate distribution was fit for each data set. Note that only the Random Fraction (RF) and Random Assortment (RA) models provide an adequate fit to the summer data. From Tokeshi (1990), with permission.

Tokeshi (1990) discussed the difficulties and ambiguities of interpreting species abundance data. He tentatively concluded that niche apportionment models did not account for relative abundance patterns in this chironomid community, because the data were best fit by the random assortment model. This interpretation is consistent with the fact that the chironomid community was dynamic, with considerable turnover in composition and substantial overlap in temporal activity and resource use (Tokeshi 1986; see Table 4.4).

However, the interpretation of Tokeshi's (1990) random assortment model is somewhat problematic. Because the random assortment model is a special case of the geometric series, data that fit the random assortment model could be described by a geometric series that reflects resource partitioning. More important, one could argue that a truly null distribution for competitive interactions would be that the abundance of each species was determined by a random draw from a uniform distribution, as MacArthur (1957) first suggested. This null distribution would not fit the chironomid data, which were less equitable than even the broken stick. Thus, some resource partitioning may have been important in determining relative abundances of chironomids, although the patterns did not conform to any of the models Tokeshi (1990) tested for this purpose. Tokeshi's (1990) results emphasize the importance of repeated sampling of small guilds of species, and of testing data against several alternative models.

Finally, we note that the most interesting questions about species abundance data may not be what determines the shape of the distribution, but what factors allow certain species to be abundant and cause others to be rare. In dynamic communities such as Tokeshi's (1990) chironomids, turnover is frequent, so the question of species identity is less important. But in most assemblages certain species are consistently abundant and others are consistently rare (Lawton and Gaston 1989; Ebeling et al. 1990; see Chapter 10). The study of factors contributing to commonness and rarity (e.g., Rabinowitz et al. 1984) may ultimately be more enlightening than further analyses of the form of the species abundance distribution.

RECOMMENDATIONS

We recommend Goldman and Lamshead's (1989) implementation of Caswell's (1976) neutral model as a benchmark for assessing species diversity. Tokeshi's (1990) models are worthwhile tests of resource partitioning, and these can be compared using the procedures of Wilson (1991b). Analyses of

species abundance patterns will be most informative when they are based on replicated samples of a small number of ecologically “similar” species. We suggest that further research on species abundance should focus on the relationship between abundance and resource consumption and the biological factors that allow some species to be persistently common and cause others to be persistently rare.

