1
PHILOSOPHY, HISTORY

Does competition lead to "forbidden combinations" of species that never coexist? Are species extinctions selective with respect to body size, geographic range, or trophic status? Has natural selection modified the body sizes of sympatric species? What factors are responsible for the relative abundance of species in an assemblage? These are the types of questions in community ecology that can be explored with null models.

Among ecologists, null models have gained popularity only in the last 20 years. The phrase "null models" was apparently coined by Robert K. Colwell and David W. Winkler at a 1981 conference at Wakulla Springs, Florida, devoted largely to the topic (Strong et al. 1984). Null models have been applied to a diverse set of questions and have yielded new insights into pattern and mechanism in community ecology. In this book, we survey both the theoretical framework and the empirical findings of null models in several disciplines of community ecology.

We restrict our focus to the community level of analysis, although we note that null models have also gained popularity in other areas of evolutionary biology (Nitecki and Hoffman 1987), including animal behavior (Aronson and Givnish 1983), population regulation (Crowley 1992), paleobiology (Raup et al. 1973; Gould et al. 1977), and phylogenetic reconstruction (Archie 1989). Our emphasis is not simply on randomization or Monte Carlo methods (Manly 1991), but on null models that have been tailored to address specific mechanisms and patterns in community ecology. Null models have been controversial in this context, and critics have raised important technical and philosophical objections (Harvey et al. 1983). In this chapter, we review the origins of these controversies and describe the earliest uses of null models in community ecology.

METHODS IN COMMUNITY ECOLOGY

Three empirical tools in ecology are laboratory, field, and "natural" experiments (Connell 1975). Diamond (1986) and Wiens (1989) have presented in-depth reviews of the strengths and weaknesses of these approaches. Laboratory experiments allow an investigator the greatest amount of control over independent variables, so that hypotheses can be rigorously tested (Mertz and McCauley 1980). Laboratory experiments have yielded data that have proved very difficult to gather in the field, such as parameter estimates for population growth models (Gause 1934; but see Feller 1940; Haefner 1980). In a simplified environment, laboratory experiments may be viewed as "simulations" that use real animals rather than a computer. The chief weakness of laboratory experiments is that biological realism is usually sacrificed for precision, so it is difficult to apply the experimental results to real communities.

By contrast, field experiments are widely regarded as the sine qua non for understanding nature (Underwood 1985)—the investigator manipulates variables in the field and then directly measures their effects. If the experiment is to be meaningful, the "signal" of manipulation must be detectable above the background "noise" of uncontrolled factors. The strength of this method is the realistic environment in which the experiment is embedded. In some systems, a long tradition of field experimentation has led to a detailed understanding of community structure (Paine 1977; Brown and Heske 1990).

Despite their importance, field experiments have several limitations. First, constraints of time, money, and resources severely limit replication and spatial dimensions of field experiments. Consequently, approximately 80% of all field manipulations have employed quadrats of only 1 m^2 (Kareiva and Anderson 1988). It is difficult to generalize the results of such small-scale experiments to larger spatial scales (Wiens et al. 1986). Second, it is often impossible to manipulate one, and only one, factor in a field experiment, and therefore it is difficult to establish appropriate controls (Hairston 1989). Third, the number of treatments necessary to reveal interactions at the community level may be prohibitive. For example, a minimum of five replicated treatment combinations is necessary to control for intraspecific effects in a two-species competition experiment (Underwood 1986). If there are age- or size-structured interactions among species (Polis et al. 1989), this number will increase severalfold. Finally, most field experiments have been conducted with small, relatively short-lived animals and plants and may have little relevance to the ecology of long-lived organisms. Some of the most important questions in ecology, such as the origin of latitudinal gradients in species diversity, may never be answered with field experiments. Another approach is needed.

The "natural experiment" (Cody 1974) overcomes some of the limitations of field and laboratory experiments, although it is not a true experiment because no variables are manipulated. Instead, an investigator compares two or more communities that are thought to differ mainly in the factor of interest. For example, to understand the effect of predators on prey abundance, one could compare prey distributions on islands with and without predators (e.g., Schoener and Toft 1983). Natural experiments form the basis of many studies in community ecology; their use is limited only by the imagination of the investigator, not by the constraints of experimental design.

The problem with natural experiments is not the discovery of pattern, but the inference of mechanism (James and McCulloch 1985; McGuinness 1988). Both predator and prey abundances may depend on the level of a third, unmeasured variable, such as the frequency of disturbance. Even when confounding variables are controlled statistically, the direction of cause and effect may not be obvious. Do predators control prey community structure, or do prey assemblages dictate predator community structure? Finally, the natural experiment begs the critical question: what predator and prey assemblages would exist in the absence of any trophic interactions?

Null models can address this last question through a statistical analysis of ecological pattern. A typical null model generates communities expected to occur in the absence of a particular mechanism. Patterns in these "pseudo-communities" (Pianka 1986) are then compared statistically to patterns in the real community. Finally, deviations from the null model can be compared to the predictions of ecological theory.

DEFINING THE NULL MODEL

The null model formalizes a particular null hypothesis in ecology: "Null hypotheses entertain the possibility that nothing has happened, that a process has not occurred, or that change has not been produced by a cause of interest. Null hypotheses are reference points against which alternatives should be contrasted" (Strong 1980). Although Strong's (1980) description of the null hypothesis captures its important features, we propose a more detailed working definition of an ecological null model:

A null model is a pattern-generating model that is based on randomization of ecological data or random sampling from a known or imagined distribution. The null model is designed with respect to some ecological or evolutionary process of interest. Certain elements of the data are held constant, and others are allowed to vary

stochastically to create new assemblage patterns. The randomization is designed to
produce a pattern that would be expected in the absence of a particular ecological
mechanism.

The null model is not a true experiment. Instead, it is a thought experiment that allows us to explore the range of possible worlds and patterns were certain ecological mechanisms not in operation. It is superior to the natural experiment because it incorporates stochastic effects and allows for a variety of possible outcomes, including that of "no effect."

There are at least two different interpretations of ecological null models. Simberloff (1983a) considered the null model as a complex form of statistical randomization. This interpretation places null models within the framework of statistical hypothesis testing. The statistical null hypothesis allows for simple pattern tests that may not require direct consideration of mechanisms. The second interpretation, epitomized by the simulations of Colwell and Winkler (1984), is that null models are explicit colonization scenarios used to test effects of biotic interactions in natural communities. Taken to their extreme, these scenarios do not qualify as null models because they are too complex and may incorporate effects of interest. We favor a more balanced view that null models describe the assembly of communities, but do not specify all the details of the colonization process. The analyses reveal community patterns that are relevant to tests of ecological theory.

NULL MODELS AND MATHEMATICAL MODELS

Although null models have a firm empirical base, there are many similarities between the construction of null models and mathematical models in ecology. Above all, they both benefit from simplicity and generality. Simplicity is important because it allows the investigator to highlight a few mechanisms without becoming entangled in complex interactions of many variables (Caswell 1988). For example, a mathematical model of predator-prey interaction may ignore the age structure of populations, whereas a null model of species co-occurrence may ignore habitat variation. Both factors can be, and have been, incorporated into more complex models. However, mathematical models become insoluble with too many parameters, and null models become difficult to generate and interpret with too many background factors. Thus, it would be a mistake to assume that a model is superior because it is more "realistic." The more detailed a model becomes, the more idiosyncratic its behavior, and the less applicable its results are to other systems (Levins 1966). Simple, well-

formulated models should be the goal of both mathematical ecology and null model testing.

Mathematical models and null models both emphasize a plurality of approaches to solving problems in community ecology. For example, a problem such as population growth in a resource-limited environment can be addressed with a variety of analytical tools, none of which provides a single "right" answer to the question (Woolhouse 1988). Similarly, many different null models can be constructed to generate community patterns in the absence of interspecific competition (Schluter and Grant 1984; Graves and Gotelli 1993). As Pianka (1994) notes, "The pseudocommunity approach is pregnant with potential and would seem to be limited only by our own ingenuity." The diversity of null models has not always been appreciated. Some critics have dismissed null models entirely without acknowledging that their criticisms are applicable to only one type of randomization algorithm.

The differences between null models and mathematical models are as important as the similarities. Mathematical models do not require empirical data for analyses, whereas null models are framed with reference to a real data set. Mathematical models highlight certain mechanisms by explicitly incorporating them (Caswell 1988). In contrast, null models deliberately *exclude* a mechanism in order to gauge its effect. Finally, null models are designed to test patterns in real data sets, whereas some mathematical models are constructed for heuristic purposes or for comparison with other models (Roughgarden 1983).

FEATURES OF NULL MODELS

Ecological null models have several important attributes:

1. Null models precipitate a sharp distinction between pattern and process. This dichotomy requires ecologists to distinguish between the patterns they observe and the different mechanisms that can produce them (Rathcke 1984). Null models force ecological theory to generate simple predictions of how nature is structured, and allow empiricists to test those predictions with real data. Much of the confusion and controversy in community ecology today can be traced to the complex predictions generated by much of theoretical ecology, and by the failure of ecologists to clearly distinguish pattern from process (Peters 1991).

2. Null models allow for the possibility of no effect. In an experimental study, one possible outcome is that the statistical null hypothesis of no differ-

ence among treatment means cannot be rejected. Similarly, one possible outcome of a null model analysis is that observed community patterns cannot be distinguished from those generated by the null model. If the null model has been *properly* constructed, we can conclude that the mechanism is not operating, or that the theory does not generate unique predictions. If the null model is flawed, it may not be rejected because it incorporates some of the processes it was designed to reveal, or because it simply has weak statistical power.

Alternatively, if the null model is rejected, and the pattern is consistent with predictions of theory, this provides some positive evidence in favor of the mechanism. However, positive evidence should not be construed as a definitive test (Brady 1979), because different mechanisms can generate similar ecological patterns. As before, the null model may also be rejected improperly if it is flawed. In particular, if the model is "too null" and does not incorporate realistic biological constraints, it may generate predictions that are very different from patterns seen in real communities. The controversy over null models is reflected in these alternative interpretations of the acceptance or rejection of a null model.

The strict falsificationist protocol embodied in null models is in contrast to a more inductive approach that was popularized in ecology by Robert H. MacArthur—ecologists searched for community patterns to corroborate a mechanism, but usually did not ask what the patterns would look like in the absence of the mechanism. Although MacArthur (1957, 1960) pioneered the "broken stick" as a null model for species abundance patterns (see Chapter 3), he also wrote that the group of researchers interested in making ecology a science "arranges ecological data as examples testing the proposed theories and spends most of its time patching up the theories to account for as many of the data as possible" (MacArthur 1962). The use of an explicit hypothesis test of expected patterns in the absence of a mechanism is what distinguishes the null model approach from other comparisons of model predictions with real data.

3. Null models rely on the principle of parsimony. Parsimony suggests that we favor simple explanations over complex ones (Brown 1950). Thus, if a null model that excludes predators successfully predicts prey community structure, parsimony dictates that we abandon the predation hypothesis—it adds unnecessary complexity. However, there is no guarantee that simple explanations are correct; parsimony is an empirical or aesthetic principle, not a logical one.

Either implicitly or explicitly, parsimony guides many of our attempts to explain patterns in nature. In evolutionary biology, for example, parsimony has guided systematists to search for the smallest amount of character change

necessary to reconstruct a phylogeny (Hennig 1966; Wiley 1981). The parsimony principle is unlikely to lead to the emergence of new hypotheses in science (Dunbar 1980; Loehle 1990a), but it is invaluable for testing and choosing among existing alternatives (Platt 1964; Loehle 1987).

4. Null models rely on the principle of falsification. Proponents of null models draw their inspiration from statistical hypothesis-testing and from the writings of Karl R. Popper (1959, 1965, 1972), who emphasized the importance of falsification of hypotheses and the asymmetry of scientific data: negative evidence can be used to refute an hypothesis, but confirmatory evidence cannot be viewed as a "test" of an hypothesis (Brady 1979).

Not all ecologists (or philosophers) agree with this view (itself unfalsifiable!). Some have pointed out that falsification is not the only way to make progress in evaluating ecological models (James and McCulloch 1985). Others have accepted the logic of comparing model predictions to data in order to detect anomalies, but not to reject hypotheses (Southwood 1980). Finally, some critics (e.g., Toft and Shea 1983) have suggested that incorrect acceptance of a null hypothesis (Type II error) is just as serious a mistake as incorrect rejection (Type I error). Conventional statistical tests emphasize the latter over the former (Shrader-Frechette and McCoy 1992), on the grounds that falsity (Type I error) is a more serious mistake than ignorance (Type II error).

Although null models rely on falsification, they do not conform precisely to the "strong inference" protocol of defining a set of mutually exclusive alternative outcomes (Platt 1964). Instead, the null hypothesis can be an aggregate of several possible mechanisms that is pitted against the predictions of a single alternative hypothesis. However, multiple alternatives can be tested with multiple null models (Schluter and Grant 1984; Graves and Gotelli 1993).

5. Null models emphasize the potential importance of stochastic mechanisms in producing natural patterns. Although much early ecological theory was based on deterministic mathematics, many patterns in nature have a stochastic component that reflects underlying environmental variability (Wiens 1977; den Boer 1981; Schmitt and Holbrook 1986). Because of this, ecologists have emphasized replication and repeatability of natural patterns (Strong et al. 1984). Null models reflect this natural variability in community structure and require that the "signal" of mechanism be stronger than the "noise" of natural variation (May 1974).

However, the inclusion of stochastic forces in null models does not imply that patterns in nature are random (Connor and Simberloff 1986). Some factors

that determine community structure are so variable in time and space that it is simpler to represent them as stochastic elements than to model them explicitly. As a familiar example, the trajectory of falling rain is determined by a complex interaction of wind speed, air temperature, and other factors. Nevertheless, the spatial pattern of drops is "random" (Simberloff 1980a) and can be described by a Poisson distribution. Simple community patterns, such as the canonical log normal distribution of species abundances, may also arise from complex interactions among factors (May 1975a).

THE CONTROVERSY SURROUNDING NULL MODELS

The use of null models in ecology has been highly contentious. Some proponents of null models (Connor and Simberloff 1979; Strong et al. 1979) have used this tool in a vigorous assault on the framework of competition theory, claiming, in essence, that "the emperor has no clothes." Defenders of competition theory have countered that null models are fatally flawed and inherently biased against detecting biotic interactions (Grant and Abbott 1980; Diamond and Gilpin 1982). Thus, the null model debate is embedded in two different controversies: (1) the importance of interspecific competition to community structure (Schoener 1982); (2) the way to go about testing theory in community ecology (Strong et al. 1984).

The controversy over competition theory calls for an historical perspective (see also Schoener 1982; Simberloff 1982; Wiens 1989). Ecology has always had an uneasy relationship with theoretical developments. McIntosh (1980) documents the change from the 1930s, when ecology had little theoretical framework, to the 1980s, when models had proliferated to the point where they could predict any possible ecological pattern (Pielou 1981a). Mathematical ecology flourished in the intervening years, owing, in part, to the influence of G. Evelyn Hutchinson and his student Robert H. MacArthur.

MacArthur synthesized the Hutchinsonian niche (Hutchinson 1957) with the competitive exclusion principle (Gause 1934; Hardin 1960). His models predicted a limit to the similarity of coexisting species, and a community packed like molecules of a crystal, with each species neatly fitting into the community niche space. In *Geographical Ecology* (1972), MacArthur extended these ideas to large geographic scales and suggested that the effects of competitive exclusion would be strong enough to control species coexistence on islands and entire continents. MacArthur's inductive approach of simple models supported by a few well-chosen empirical examples was rapidly incorporated into textbooks (Hall 1988). In a modified form, his ideas on species coexistence and

ecological niches continue to define much of the current research program for academic ecology.

PARADIGMS IN COMMUNITY ECOLOGY

In retrospect, MacArthur's approach constituted an important ecological paradigm (sensu Kuhn 1970), in which the primary focus was the coexistence of competing species in a community (Wiens 1989). A good indicator of paradigm status is that alternative viewpoints, and even conflicting data, are swept aside in favor of the reigning theory. Thus, in 1978, Diamond was able to advise ecologists to be suspicious of results suggesting that species were not directly or even currently competing with one another (Strong 1980)!

By the early 1980s, several important challenges to the MacArthurian paradigm had emerged (Schoener 1982): (1) Null model analyses indicated that some of the patterns ascribed to interspecific competition could arise from models that were competition-free; (2) mathematical analyses suggested that simple limits to similarity or exclusion of competitors were sensitive to model formulation (Abrams 1975; Turelli 1978a; Armstrong and McGhee 1980); (3) the "variable environments" hypothesis predicted that animal populations are rarely at carrying capacity (Andrewartha and Birch 1954; Wiens 1977; den Boer 1981), suggesting that competitive effects are important only during occasional "resource crunches" (Wiens 1977; but see Gotelli and Bossert 1991); (4) empirical and theoretical studies indicated that predation was often a stronger biotic interaction than competition (Connell 1975), particularly at lower trophic levels (Hairston et al. 1960). Although Schoener (1982) argued that MacArthurian competition was not a failed paradigm, competition theory is no longer accepted uncritically, as it often was before these challenges.

Not all ecologists accept the view of the dominance of a MacArthurian paradigm. Plant ecologists focused on studies of competition and the niche well before MacArthur's work became influential (Jackson 1981), and the MacArthurian approach may have been more of a preoccupation with American ecologists than with Europeans (May and Seger 1986). Wiens's (1989) analysis of literature citations over several decades also suggests that competitive explanations for avian community patterns were about as popular before MacArthur as during his reign. Nevertheless, MacArthur's blend of mathematical competition models and nonexperimental supporting data had a long-lasting influence on community ecology.

In a roundtable issue of *Synthese*, Simberloff (1980a) argued that ecology has passed through not one, but three successive paradigms. The first was one

of essentialism, in which populations and communities achieved an ideal, deterministic form. In community ecology, essentialism is embodied in Clements's (1904) idea of the community as a superorganism of tightly integrated species. The second paradigm was one of materialism, in which communities were not forced into a few monoclimax types; rather, the focus of study was the variation in composition and organization of communities, and the processes that lead to these differences. The MacArthurian approach, with its emphasis on geographic variation in community structure and complex interactions among species, seems to embody elements of both of these paradigms. Simberloff's (1980a) third paradigm was one of probabilism, in which variation in community structure is viewed as the outcome of a small set of probabilistic pathways. Although Simberloff (1980a) did not explicitly develop this point, it is clear that null models are central to testing mechanisms of community structure in a probabilistic framework.

In response to his essay, Levins and Lewontin (1980) argued that Simberloff (1980a) had mistaken the stochastic for the statistical and had confused the ideas of reductionism with materialism and idealism with abstraction. Levins and Lewontin's (1980) ecological criticism was that, in rejecting the Clementsian superorganism, Simberloff (1980a) "falls into the pit of obscurantist stochasticity and indeterminism." Levins and Lewontin (1980) believed the community to be a meaningful whole, with dynamics that are distinct from those levels below (populations) and above (ecosystems). In contrast, Simberloff (1980b) argued that community organization was a hypothesis to be tested. Simberloff's (1980b) null hypothesis was that emergent community properties are epiphenomena caused by patterns at the species level. That is, the synecological "whole" is primarily the sum of its autecological "parts."

CRITICISMS OF NULL MODELS

Roughgarden (1983) espoused a very different view of community structure, sparking another roundtable of essays in *American Naturalist* concerning competition, community structure, and null models in ecology. Prompted by challenges to competition theory (Connor and Simberloff 1979; Strong et al. 1979; Connell 1980), Roughgarden's (1983) essay was a defense of mathematical competition theory as well as a salient critique of null models. Roughgarden (1983) argued against null models and hypothesis testing. Instead, he suggested that we use common sense to establish facts in ecology and evaluate mechanisms on the basis of our ability to "build a convincing case. A convincing case should include on-site experi-

ments together with biogeographic and distributional data, and data addressing viable alternative hypotheses."

Roughgarden (1983) felt that null hypotheses were "empirically empty," because biological processes such as stochastic dispersal and population extinction were not incorporated in the randomization procedures (but see Caswell 1976). He argued that null models (as they had been developed by 1983) were not ecological analogs of the neutrality hypothesis in population genetics, because the underlying mechanisms had not been demonstrated. Even if the mechanisms could be established, null models would have no "logical primacy" over other types of models, in contrast to the claims of Strong et al. (1979). Whereas Strong et al. (1979) implied that pattern must be established before processes can be investigated, Roughgarden (1983) maintained that processes such as interspecific competition could be studied productively before establishment of pattern: "Sometimes it is obvious that a process is occurring. Knowledge of that process may aid in discovering its consequences." Finally, Roughgarden (1983) did not think that ecological null models were analogous to statistical null hypotheses: "A null hypothesis in statistics is a justified model of sampling procedure. It is not a hypothesis that the world has no structure."

In the same roundtable issue, Quinn and Dunham (1983) voiced similar concerns and raised three philosophical objections to the null model approach:

1. *Formal hypotheses about mechanisms cannot be stated in such a way as to allow meaningful disproof.* In other words, mechanisms in community ecology cannot generate simple predictions about patterns that can then be subject to a "litmus test" (Roughgarden 1983) for proof or disproof.
2. *Testing mechanisms as distinguishable hypotheses leads to univariate critical tests.* These tests are invalid if there are strong interactions among the mechanisms.
3. *In hypothetico-deductive formalism, understanding is only increased when a hypothesis is rejected.* Reliable null hypotheses may be impossible to construct, because we cannot generally deduce the nature of expected patterns that would emerge in the absence of any given biological process.

REPLIES TO THE CRITICS

These roundtable essays demonstrated that much of the debate involved personal styles of research and philosophy that could not be classified as "right" or

"wrong" (Salt 1983). Simberloff (1983a) was sympathetic to Roughgarden's (1983) common-sense approach, but considered it overly simplistic and therefore likely to lead us astray. Strong (1983) wrote, "Yes, common sense sometimes leads to sound judgment, but it is also ordinary, free from intellectual subtlety, not dependent on special or technical knowledge, and it is unreflective opinion." Strong (1983) also pointed out that a common-sense approach lacks two critical elements, imagination and testing, that feature prominently in the construction and interpretation of null models.

Our own opinion is that null models have been useful in ecology and that many of the flaws pointed to by critics can be addressed. Roughgarden's (1983) complaint that null models do not explicitly describe colonization processes is not necessarily damning if the randomization reflects the major patterns produced during colonization. Whether or not null models have "logical primacy" (Strong et al. 1979), they should be investigated first, so that stochastic and sampling effects can be distinguished from biologically meaningful patterns (Järvinen 1982). As a practical matter, null models should be tested first because they may save a huge amount of time that could be frittered away in search of a nonexistent process or phenomenon.

Null models do not portray the world as "having no structure" or even as being random. Rather, the null hypothesis (for community-wide competition) is that species occurrences are random *with respect to one another.* Moreover, randomization procedures are a well-established protocol in statistics for constructing null hypotheses; all of the conventional statistical tests, such as the F-ratio and the chi-squared distribution, have analogs based on a randomization test (Manly 1991). Indeed, some statisticians feel that randomization tests are actually preferable, because they are burdened by fewer assumptions (Edgington 1987).

Finally, we must take issue with Quinn and Dunham's (1983) argument that null models are not appropriate for ecology because ecological theory makes ambiguous predictions. This seems a damning criticism of theory rather than of null models! If we are to bridge the wide gulf between theoretical and empirical ecology, we must force ecological theory to give us simple predictions that we can test with data (Pielou 1981a; Simberloff 1982). This real-world portability, rather than the realism of underlying assumptions, may be the most useful criterion for judging the fitness of a model (Brown 1981). In its original, unadulterated form, MacArthurian competition theory and other simple ecological models do make predictions about species richness, co-occurrence, and niche overlap that can be tested with appropriate null models. We survey these tests in later chapters.

In summary, null models have been controversial in ecology for three reasons: (1) the null model principles of parsimony and falsification are not

accepted by some ecologists; (2) the findings of some null models directly contradicted predictions of orthodox competition theory; (3) the rhetoric of some of the early papers was contentious and often extreme, ensuring that a lively exchange on the issues would follow. In particular, the "assembly rules" debate has involved many players and has spanned over two decades, with no clear resolution in sight (Diamond 1975; Connor and Simberloff 1979; Alatalo 1982; Diamond and Gilpin 1982; Gilpin and Diamond 1982; Connor and Simberloff 1983, 1984; Gilpin and Diamond 1984; Gilpin et al. 1984; Gilpin and Diamond 1987; Wilson 1987; Roberts and Stone 1990; Stone and Roberts 1990, 1992).

If these exchanges were the sole impact of null models, they would remain only an interesting footnote in the history of ecology. But the influence of null models has been pervasive in ecology, and has spread far beyond the original focus on interspecific competition and species associations. Food web structure (Pimm 1980a,b; 1984), island extinction patterns (Gotelli and Graves 1990), and vertical (Underwood 1978a) and horizontal (Pielou 1977, 1978) zonation of intertidal communities represent just a few examples of problems that have been addressed with null models. Even landscape ecology, one of the more holistic subdisciplines of the field, has adopted null models of minimalist assumptions that describe random movement of individuals through a hetero-geneous landscape (Milne 1992). Null models have also had a positive influ-ence on experimental ecology, forcing investigators to articulate alternative mechanisms and make a priori predictions about pattern. Finally, the construc-tion of an appropriate null model and evaluation of its underlying assumptions is an important step toward rigor and clear thinking in community ecology (Platt 1964; Rathcke 1984; Peters 1991).

S/G RATIOS AND THE HISTORY OF NULL MODELS

Null models in ecology were first used in the analysis of species/genus (S/G) ratios. These ratios may reflect the intensity of competition, which Darwin (1859) and many others suggested was greatest within a genus: "As species of the same genus have usually, though by no means invariably, some similarity in habits and constitution, and always in structure, the struggle will generally be more severe between species of the same genus, when they come into competi-tion with each other, than between species of distinct genera."

Elton (1946) addressed the hypothesis with data on the number of species per genus, or the average S/G ratio, in a variety of animal and plant communities. He was impressed by the fact that most genera in insular or local assemblages

were represented by only a single species, with an average S/G ratio of 1.38 for animal communities and 1.22 for plant communities. In contrast, the S/G ratio was considerably larger for regional floras and faunas. For example, the average S/G ratio for British insects was 4.38, a value never reached in any of the local surveys.

What was the explanation for this pattern? Elton (1946) briefly considered the "null hypothesis" that S/G ratios represented sampling artifacts in small communities: "One possible explanation of the statistical relationships described above would be that the frequencies of species in genera simply reflect those of the fauna as a whole. For if, say 86% of species in the British Isles belonged to genera of which only one species was present in their region, the figures . . . would be the record of a faunistic distribution, rather than any peculiarity of homogeneous communities taken separately."

Elton (1946) rejected this null hypothesis and reasoned, incorrectly as it turned out, that disparity between regional and local S/G implied that competition within communities limited the coexistence of congeners. Citing Darwin's (1859) writings and Gause's (1934) laboratory studies, Elton (1946) argued that S/G ratios were lower for local communities because interspecific competition for limited resources precluded the coexistence of several species in the same genus.

Williams (1947a) took a very different approach and gave ecology one of its earliest null models. He asked what the S/G ratios of local communities would look like in the absence of competition. He emphasized that real communities must be compared to this null expectation in order to evaluate the role of competition: "It is, however, most important to consider in detail what exactly happens when a selection of a relatively small number of species is made from a larger fauna or flora, without reference to their generic relations . . . as a true interpretation can only be made by comparing the observed data with the results of a selection of the same size made at random."

Williams (1947a) assumed that the relative abundance of species followed a log series, based on his previous work on this problem (Fisher et al. 1943). From the log series, the expected S/G ratio in a random sample of species can be predicted. Williams (1947a) reanalyzed Elton's (1946) data, and found that in every case, the expected S/G ratio decreased as fewer species were sampled. Consequently, the match between expected and observed S/G ratios was quite good. When deviations did occur, S/G ratios were typically *larger* than expected by chance (Figure 1.1). Williams (1947a) argued that S/G ratios were influenced by two factors: (1) competition, which would be most severe among congeners and reduce the ratio, and (2) environmental suitability, which would be most similar among congeners and increase the ratio because similar species
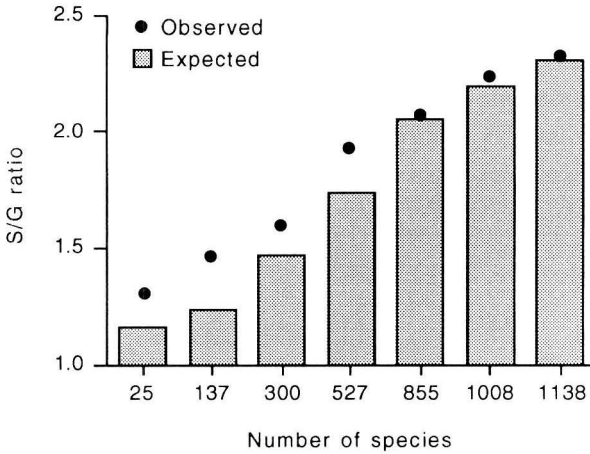
Figure 1.1. Expected and observed species/genus (*S/G*) ratios for vascular plants of the British Isles. Each point represents a single island. Expected values are determined by random sampling from the presumed source pool, the vascular plants of Britain (1,666 species; *S/G* ratio = 2.727). Elton (1946) believed that small *S/G* ratios on islands reflected strong competition between congeners and a limit to coexistence of closely related species. However, the expected *S/G* ratio is also small in randomly sampled communities that are not structured by competition. Observed *S/G* ratios are usually quite close to this expectation. Competition should depress the ratio below the expectation, but most communities show deviations in the positive direction. Data from Simberloff (1970).

would occur in the same habitat. Williams (1947a) argued that if competition was influencing *S/G* ratios in Elton's (1946) local communities, its effect was overshadowed by other factors operating in a reverse direction. In any case, no special significance could be attached to small *S/G* ratios per se.

Bagenal (1951) criticized Williams's (1947a) analysis and pointed out that some of Elton's (1946) local assemblages incorporated different habitats or were sampled at different times. Consequently, they could not be used to assess Gause's (1934) hypothesis of competitive exclusion, which Bagenal (1951) interpreted as applying only to species co-occurring in similar habitats. Moreau (1948) also noted this problem with Elton's (1946) analyses and offered more appropriate data for examining *S/G* ratios: detailed records on the occurrence and habitat associations of 172 species of African birds. Within the 92 genera represented, only 16% (excluding the Ploceidae) of the possible cases of habitat overlap among congeneric pairs were observed, and only a third of these apparently overlapped in diet. Moreau (1948) concluded that when habitat

affinities were properly defined, competitive exclusion was important in limiting coexistence of congeners and more distantly related species pairs as well.

Moreau (1948) assumed that 16% overlap was "low," but he did not consider sample size effects in his analysis. Williams (1951) reanalyzed Moreau's (1948) data, again using the log series to estimate the expected S/G ratios in small communities. As in some of Elton's (1946) assemblages, congeneric co-occurrence among Moreau's (1948) birds was many times greater than expected by chance, suggesting that competitive effects on the S/G ratio were swamped by similar habitat affinities of congeners. Williams (1951) also explored the statistical power of his test: if there were fewer than nine species in the assemblage, the expected S/G ratio was so low that competitive effects could never be detected. Finally, in *Patterns in the Balance of Nature* (1964), Williams again analyzed S/G ratios of British birds and plants, and Moreau's (1948) African bird communities. To relax the assumption of a log series distribution, Williams used numbered discs to represent species in different genera. He then took random samples to confirm that the S/G ratio declines with decreasing S, regardless of the underlying species abundance distribution.

In spite of Williams's (1947a, 1951, 1964) thorough treatment of this topic, S/G ratios and other taxonomic indices continued to be used as an index of competition without reference to an appropriate null hypothesis (Grant 1966; Moreau 1966; MacArthur and Wilson 1967; Cook 1969). However, Hairston (1964) did heed Williams's (1964) results and pointed out three biological assumptions implicit in his analyses:

1. The number of different habitats occupied by a species is unrelated to the size of the genus of that species.
2. The presence of a species in one habitat can be interpreted without reference to its presence or abundance in other types of habitat.
3. The human observer divides the total area into habitats in the same manner as do the species being studied.

Hairston (1964) felt Williams's (1947a) first assumption was violated and suggested that species belonging to species-rich genera usually occupied more habitats than species from monotypic genera. This sort of distribution would lead to lower S/G ratios within a single habitat type. Hairston (1964) also rejected assumptions (2) and (3), based on his own detailed studies of salamander ecology. He wrote that "it is not possible to draw valid conclusions about interspecific competition from an analysis of generic relations of species in a series of communities or habitats . . . it can be concluded that this approach is not a satisfactory substitute for more detailed studies."

Simberloff (1970) reviewed the controversy and analyzed *S/G* ratios for a number of data sets. Using computer simulations, he generated the expected *S/G* ratio for small communities by drawing from appropriate source pools. As Williams (1964) had shown earlier, the expected *S/G* ratio varied from a minimum of 1.0 (for a community of one species) to a maximum equal to that of the source pool. Between these limits, the expected curve was slightly convex; its shape depended on the underlying species abundance distribution in the source pool. The procedure is a computer simulation of rarefaction (Chapter 2) in which the expected species richness and its variance (Heck et al. 1975) can be calculated for a random sample of individuals from a collection.

Simberloff (1970) found that the expected *S/G* ratio was relatively insensitive to alterations in the source pool composition. For 180 island floras and faunas, 70% of the *S/G* ratios were larger than expected, contrary to the predictions of the competitive exclusion hypothesis (see McFarlane 1991 for a noteworthy exception). Three hypotheses might account for large *S/G* ratios:

1. Similar habitat affinities of congeneric species (Williams 1947a).
2. Restricted habitat affinities of monotypic genera (Hairston 1964).
3. A positive correlation between dispersal ability and taxonomic affinity. In other words, if species of certain genera are especially good at dispersal, they will tend to be overrepresented in island avifaunas.

Simberloff (1970) noted that it is not possible to distinguish among these hypotheses on the basis of species lists, although he did point to examples of widely distributed species of monotypic genera, which argues against hypothesis (2).

Several important points emerge from the history of *S/G* ratios. First, failure to consider an appropriate null hypothesis can lead investigators astray, causing them to attribute biological mechanisms to patterns that may represent sampling artifact. Second, when the null hypothesis is explicitly cast, it forms a reference point for examining patterns. In the case of the *S/G* ratio, observed ratios were frequently higher, not lower, than expected. Third, a good deal of controversy surrounds the construction of any null model. Important questions of source pool designation, taxonomic effects, and dispersal abilities were debated in this early literature, and ecologists continue to lock horns over these same issues today.

Finally, it is fascinating to see history repeat itself. Järvinen (1982) describes a controversy over *S/G* ratios among European plant ecologists in the 1920s, which predated the Eltonian controversy among English and American animal
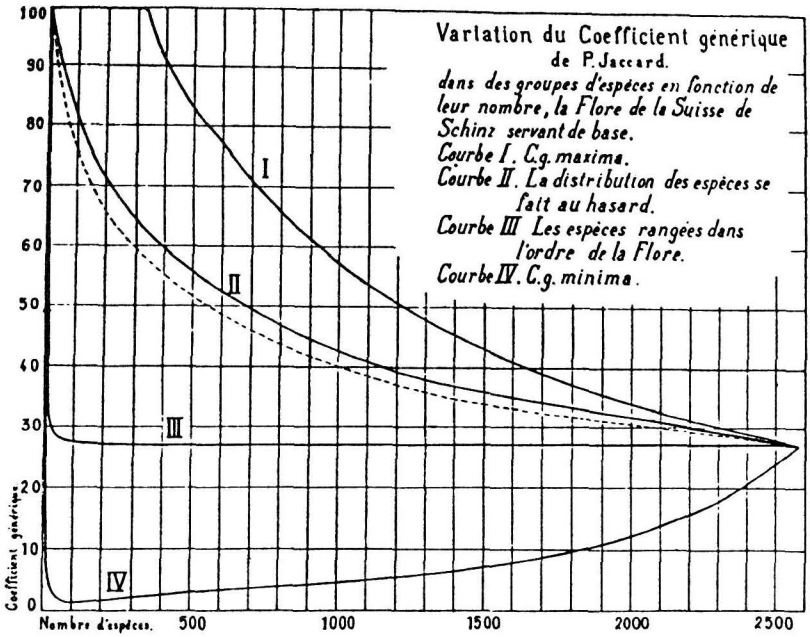
Figure 1.2. An early null model of taxonomic ratios for vascular plants of Switzerland by Maillefer (1929). The *y* axis is the genus/species ratio multiplied by 100, and the *x* axis is the number of species in the sample. Curves I and IV show the maximum and minimum values possible for the ratio, while Curve III is the expected value, based on random draws from a deck of marked cards. The dashed line is the standard deviation. Curve II is the expected value when species are sampled in systematic order. The controversy among European plant ecologists near the turn of the century was identical to the more familiar debate among animal ecologists over *S/G* ratios in the 1940s. Both groups used null models to understand the sampling properties of taxonomic ratios. From Järvinen (1982), with permission.

ecologists. The biogeographer Jaccard (1901) introduced the "generic coefficient" (*G/S* ratios, expressed as percentages) and argued that it reflected gradients of ecological diversity. Palmgren (1925) believed that the generic coefficient merely reflected the species richness of a region and said nothing about ecological diversity. Palmgren (1925) reasoned that if the number of species colonizing an area is small, chances are good that only a few genera will be represented. Maillefer (1929) discussed the statistical limitations of the generic coefficient and, like Williams (1964), used randomly shuffled cards to represent the Swiss flora and to derive expected generic coefficients (Figure 1.2). He also found that congeneric coexistence was higher than expected. Finally, the mathematician Pólya (1930) derived the expectation of the *G/S*

ratio, predating the work of Simberloff (1970), Hurlbert (1971), and Heck et al. (1975) by 40 years.

Why did most ecologists ignore or dismiss the works of Palmgren (1925) and Williams (1964)? Most likely, because probabilistic tests and notions of stochastic sampling effects simply did not fit the reigning ecological paradigms of the time (Simberloff 1970; Järvinen 1982). For similar reasons, most ecologists overlooked Munroe's (1948) independent discovery of the equilibrium theory of island biogeography (Brown and Lomolino 1989). Järvinen (1982) also noted the unfortunate tendency of (American) ecologists to ignore important literature not published in English. Finally, the lack of computers made it impractical to conduct randomization tests prior to the 1970s.

OTHER APPROACHES

After the *S/G* ratio debates, null model controversy arose again in the late 1970s over provocative papers from the ecology group at Florida State University (the so-called "Tallahassee Mafia"; Donald R. Strong, Jr., quoted in Lewin 1983). In the intervening years, null models were by no means ignored in the ecological literature. Like Elton (1946), some authors stated a null hypothesis but failed to test it. For example, in a review of competition and niche studies, Schoener (1974a) argued that observed utilization niches overlapped less than would be expected in the absence of interspecific competition. Other authors attempted to use conventional statistical tests to evaluate null hypotheses. Brown (1973), for example, used a chi-squared distribution to test the null hypothesis that coexistence is just as likely for species pairs with body size ratios greater or less than 1.5. The nonindependence of species pairs invalidates the chi-squared test (Kramer and Schmidhammer 1992), although in practice, results were identical to those generated by an appropriate null model analysis (Bowers and Brown 1982).

Some studies in this period made explicit use of null hypotheses. Sale (1974) and Inger and Colwell (1977) randomized resource utilization data to evaluate niche overlap in the absence of competition (see Chapters 4 and 5), providing a direct test of Schoener's (1974a) stated null hypothesis. Using a very different sort of null model, Caswell (1976) analyzed species abundance distributions (Chapter 3) and introduced the term "neutral model" into the ecological literature. In the study of species co-occurrence (Chapter 7), E. Chris Pielou developed elegant null models (Pielou and Pielou 1968; Pielou 1972a), but her work was virtually ignored for more than a decade (Simberloff and Connor 1981). Cohen (1978) proposed randomization algorithms for food web matrices, some

of which are identical to procedures used in competition analyses of presence-absence matrices (Simberloff 1978a). Finally, computer simulations of randomly connected food webs (Gardner and Ashby 1970; May 1972) produced surprising results (Chapter 10) and were in the spirit of later null models, although patterns were not compared with real data. Perhaps because some of these studies did not challenge the MacArthurian paradigm, they did not gain the notoriety of later null model papers, although in many cases the results were equally provocative.

RECOMMENDATIONS

Whether or not one accepts a strict falsificationist protocol, we recommend the use of null models for the analysis of nonexperimental data. Null models are particularly valuable for testing unique predictions of community theory, and for testing patterns when the assumptions of conventional statistical tests are violated. Chapters 2–10 illustrate the utility of null models for a variety of specific ecological questions.