There has been considerable controversy recently on whether the effects of competition can be seen in species co-occurrences, especially in comparison with a null model. Diamond (1975) claimed evidence for community structure in avifaunal data from the Bismark islands. Connor and Simberloff (1979) analysed Diamond and Marshall’s (1976) data for Vanuatu birds, and data for West Indies birds and bats, and concluded there was little evidence for departure from a null model. (Vanuatu was known prior to 1980 as the New Hebrides.) Diamond and Gilpin (1982) criticised some aspects of Connor and Simberloff’s methods, and produced their own method (Gilpin and Diamond 1982) which gave highly significant departures from a null model for the Vanuatu data and for the Bismark data.

It is far from trivial to decide on the appropriate null model. Should it be assumed that all species are equally likely to occur on an island? Surely not, for we know there are common and rare species. Should the observed commonness be used as a probability of occurrence? If so, some species may fail to occur in null-model patterns. Should they be included? The original data excluded species never found, but to omit them would give a null-model matrix smaller than the observed one. The same problems arise with island richness. We know that some islands (large ones usually) have more species; are we to include this in the null model? Should other features of the observed data be included in the null model?

**Gilpin and Diamond’s method**

I checked Gilpin and Diamond’s method by using it on a matrix of random numbers, representing 28 islands and 56 species, the same size as the Vanuatu dataset. Occurrence/absence of each species on each island was determined by a random number generator, with an arbitrary probability of occurrence of 0.3333. This removes all possible structure from the data. Even with ‘species occurrences’ determined thus, Gilpin and Diamond’s method indicates significant co-occurrence between species (Fig. 1; chi-squared = 224.44 \( P<0.001 \)). Figure 1 shows the results for one random pattern; other random patterns gave essentially the same result (chi-squared varied from 187.3 to 274.5, highly significant in every case). A method that finds significant co-occurrence structure in random numbers is inappropriate.

There are several possible reasons for the problem:

1. Gilpin and Diamond’s formula for calculating \( p_{ij} \) (probability of the \( i^{th} \) species occurring on the \( j^{th} \) island) must be incorrect since it can yield a value greater than 1.0 (Connor and Simberloff 1984). However, this cannot be the primary problem since the random dataset used in Fig. 1 contained no \( p_{ij} \) values above 1.0, and yet is incorrectly analysed.

2. Gilpin and Diamond (1982, 1984) relied on the Central Limit Theorem to ensure a “quick” approach of the distribution of \( E_{ij} \) values (the expected number of islands shared by a species pair) to a normal distribution. Their Vanuatu bird dataset may be too small for this to occur.

3. Their formula for S.D. might be inappropriate (Connor and Simberloff 1983). Figure 1 suggests this to be the case.

4. The degrees of freedom for the chi-squared test are calculated on the basis of 1540 species pairs. However, these pairs are based on only 56 species so the true degrees of freedom are far fewer.

The pattern of deviation from their null model (in Fig. 1) is similar to that shown by the Vanuatu data (Fig. 2; cf Fig. 3 of Gilpin and Diamond [1982]). This suggests that Gilpin and Diamond’s result was caused by bias in their analysis.
Fig. 1. Distributions of the deviation \( d_{ik} \) calculated from a 28 \times 56 matrix of presences determined at random with a probability of 0.333 (histogram), and that predicted by Gilpin and Diamond's "null model" (spots)

Fig. 2. Distributions of the deviation \( d_{ik} \) calculated from the 28 \times 56 matrix of Vanuatu bird presences (histogram), and that predicted by Gilpin and Diamond's "null model" (spots)

Fig. 3. Convergence of the null model predictions, by the method suggested here, expressed as the discrepancy between the values after 1000 simulations and those after various numbers of simulations, meaned over those categories with final frequencies greater than 1%

Suggestion of an alternative test

The appropriate null model

The difficulty of deciding on the appropriate null model was mentioned above. Wright and Biehl (1982) assumed a null model with all islands equivalent. Departure from such a model may reflect only island size effects that are already well accepted. Connor and Simberloff's (1979) null model assumed, in addition to island and species totals fixed at those observed, that species were limited to the observed incidence ranges (Diamond 1975). A null model constrained this closely to the original data may include "hidden structure" (Gilpin and Diamond 1984). Gilpin and Diamond (1982) used a null model with the observed island and species totals taken as probabilities. This has the disadvantage of including "degenerate arrangements" with empty islands or non-occurring species which would not have been included in the observed data (Connor and Simberloff 1984).

Here, I take a null model with the island and species totals fixed at the observed values, without an incidence constraint. It is clear that island richness will vary with differences in the physical environment including island size, and species frequencies because of their autecology/physiology/behaviour and biogeography. The observed totals are the best available estimates of these tendencies. The danger of including "hidden structure" makes tests based on the model somewhat conservative. There seems to be no reasonable alternative; taking row and column totals as fixed is not only the obvious model to choose, but the one that is accepted without question in, e.g., chi-squared tests. Connor and Simberloff (1984) discuss this question.

Nesting and chequerboarding

From the exchanges between Gilpin and Diamond, Connor and Simberloff, it has become clear that for some combinations of row and column totals there is a quite limited number of possible matrices. For a matrix of the type shown