

Island-sharing by archipelago species

Alan Roberts¹ and Lewis Stone²

¹ Graduate School of Environmental Science, Monash University, Clayton, Vic., 3168, Australia

² Department of Epidemiology, University of Melbourne, Parkville, Vic., 3052, Australia

Received February 20, 1990/Accepted March 7, 1990

Summary. Diamond (1975) formulated “assembly rules” for avian species on islands in an archipelago, which made a successful colonisation depend essentially on which other species were present. Critically examining these rules, Connor and Simberloff (1979) maintained that, in the Vanuatu (New Hebrides) archipelago, the field data on species distribution was quite compatible with a null hypothesis, in which species colonise at random with no species interaction. Their work was in turn criticised (Diamond and Gilpin (1982), Gilpin and Diamond (1982)) and a vigorous controversy has ensued.

Here we contribute a method in which a simple but hitherto neglected statistic is used as a probe: the number of islands shared by a pair of species, with its first and second moments. The matrix of these sharing values is given as a simple product of the incidence matrix, and its properties are examined – first, for the field data, and then in the random ensemble used by Connor and Simberloff (1979). It is shown that their constraints hold constant the mean number shared, so that any fall in the number that one pair of species share, due to their excluding each other, must imply a rise in the number shared by some other species pair – i.e., an aggregation.

Turning to the second moment of the numbers shared, it is shown that its value in the Vanuatu field data exceeds the largest value to be found in a sample of 1000 matrices, these latter being constructed so that they obey the Connor and Simberloff constraints but are otherwise random. This indicates that exclusion and/or aggregation effects are present in the actual distribution of species, which are not catered for by the null hypothesis.

The observed distribution thus emerges as much more exceptional than found by Connor and Simberloff (1979), and even by Diamond and Gilpin (1982), when examining the same ensemble. The reason for this disagreement are sought, and some cautions are offered, supported by numerical evidence, concerning the use of the chi-square test when the data points involved are mutually dependent.

Key words: Species co-occurrence – Bird distributions – Community structure – Testing significance

To what extent can data on the distribution of species, over the islands of an archipelago, be used to illuminate the processes responsible for that distribution? Since Diamond (1975) offered his rules of species assembly, and Connor and Simberloff (1979) critically analysed them, this question has been the focus of an intense and fruitful discussion, with little consensus emerging among the participants. (For recent contributions see e.g. Wilson (1987), Gilpin and Diamond (1987) and references therein).

The work below is not claimed to present a more sophisticated or mathematically deeper analysis of the data. On the contrary, it offers a simpler approach which nevertheless seems to give unambiguous results, in singling out the actual data in a much-discussed case (the Vanuatu, formerly New Hebrides, archipelago) as being quite untypical of the “random” colonisation processes modelled by Connor and Simberloff.

The reasons why the latter writers arrived at a contrary result are investigated at length, and an attempt made to examine some of the pitfalls when one uses the χ^2 test for goodness of fit, in cases where the “theoretical” values are mutually dependent.

A simpler statistic for testing interactions

As a central preoccupation in the discussion, statistics have been sought that can reliably indicate whether there are species subsets which tend to avoid each other – that is, which share an island less frequently than they would in “random” colonisation.

This concern with the number of islands shared by a pair of species – S_{ij} , say, for the i^{th} and j^{th} species – corresponds, as Simberloff and Connor noted at an early stage of the discussion, to R -mode analysis in numerical taxonomy. It might be thought surprising, then,

that the simplest statistics associated with the set $\{S_{ij}\}$ have not been directly examined – namely, its mean and mean-square value for a given system. We proceed to do this.

Define a “sharing matrix” \underline{S} , whose $(i, j)^{\text{th}}$ entry is the number of islands on which both the i^{th} and j^{th} species occur. A diagonal element S_{ii} is just the number of islands occupied by the i^{th} species.

An overhead bar indicates an arithmetic mean over the non-diagonal entries of a matrix. Thus, for example, the mean and mean-square values of these off-diagonal entries, for a particular matrix, will be denoted by \bar{S} and \bar{S}^2 respectively. Thus, if there are m species,

$$\bar{S} = (1/m(m-1)) \sum_{i \neq j} S_{ij},$$

$$\bar{S}^2 = (1/m(m-1)) \sum_{i \neq j} S_{ij}^2.$$

The incidence matrix $\underline{A} = (a_{ij})$ is defined in the usual way:

$$a_{ij} = 1 \quad \text{if species } i \text{ occurs on island } j,$$

$$= 0 \quad \text{otherwise.}$$

It is then easy to show (Appendix 1, (1)) that

$$\underline{S} = \underline{A}\underline{A}^T, \tag{1}$$

where \underline{A}^T denotes the transpose of \underline{A} . With m species and n islands, \underline{A} is a $m \times n$ matrix and \underline{S} is a $m \times m$ matrix.

This result has a “dual”, in which the roles of species and islands are exchanged. If we denote the number of species common to the islands i and j by S'_{ij} , the matrix \underline{S}' is given by (Appendix 1, (2)):

$$\underline{S}' = \underline{A}^T \underline{A}. \tag{2}$$

Thus, starting with the fundamental observations recorded in the incidence matrix \underline{A} , a single matrix multip-

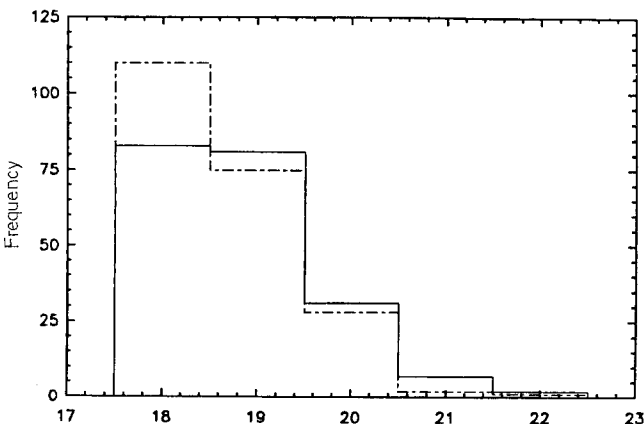


Fig. 1. The histograms show the number of random colonisation patterns, in a sample of 1000, in which P_{26} , the number of species pairs sharing 26 islands, has the value shown on the horizontal axis. The solid line is for the cases in which $P_{25} = 21$ (—), the dashed for $P_{25} = 20$ (---)

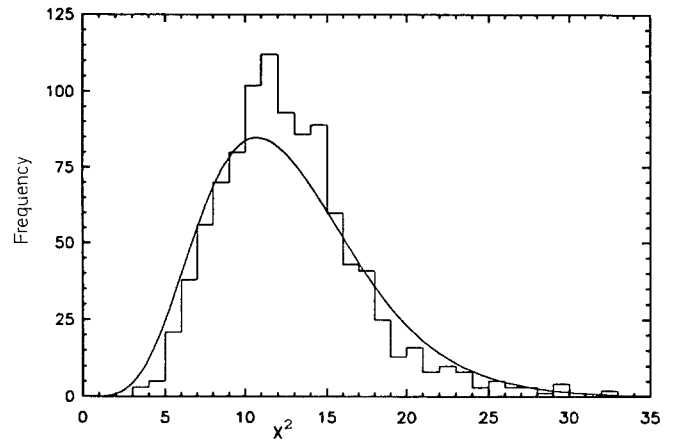


Fig. 2. Values on the horizontal axis are of the quantity X^2 as given by equation (7), in testing the distribution of the P_k 's. The histogram shows the number of random colonisation patterns, in a sample of 1000, giving one of these values. The curve shows the chi-square curve giving the best fit to the histogram, using 12.72 d.f.

lication gives us the data on island-sharing represented by \underline{S} .

In their original analysis, Connor and Simberloff (1979) examined a statistic we will here call P_k ; the number of species pairs sharing exactly k islands. This statistic has since received much attention in the literature, even though it is further removed from the observed data than S_{ij} itself, being in fact the number of times that the value k appears among the above-diagonal entries of \underline{S} . Here we will use the simpler statistic, S_{ij} itself, and its moments.

We first look at the field data, and extract the number of islands shared by a pair of species, for all possible pairs. We now follow the method most usual in the literature (but note the reservations in the next paragraph, and further below): that is to say, we compare this sharing data with that given by a random sample of incidence matrices, taken from an ensemble generated in accordance with an appropriate null hypothesis. Such an hypothesis would, of course, need to exclude any interaction between species.

There is no general agreement, however, on what other features should be built into this null hypothesis. What would make an “appropriate” null hypothesis, and thus lead to an appropriate ensemble to sample from, is itself a major point of dispute (see in particular the contributions by Gilpin and Diamond, and by Connor and Simberloff, in Strong et al. (1984), Loehle (1987), and the review by Harvey et al. (1983)).

Imposing the constraints

Particularly controversial are the three constraints originally used by Connor and Simberloff (1979). These require that, in the “random” occupation of islands envisaged, the following restrictions must apply to the notional species and notional islands (the translation into

properties of the incidence matrix is given in parenthesis):

(A). A notional species must occupy the same number of islands as does its corresponding real species. (Each row in a random matrix must sum to the same total – r_i , say, for the i^{th} row – as the corresponding row in the actual matrix.)

(B). A notional island must contain the same number of species as does its corresponding real island. (Each column in a random matrix must sum to the same total – c_j , say, for the j^{th} column – as the corresponding column in the actual matrix.)

(C). Suppose a species never occurs on islands containing less than s' or more than s'' species. Then its notional counterpart must likewise occur only on islands containing between s' and s'' species. (For each row of a random matrix, there is given a range of numbers. If there are columns whose sums c_j lie outside this range, the entries in these columns, for that row, must be zero.)

As noted above, these constraints have been criticised as inconsistent with an appropriate null hypothesis. It will soon be apparent that this objection cannot be brushed away; nevertheless, here these very constraints will be imposed as part of what might be called an a fortiori strategy. (The meaning of this cryptic remark will become clear in what follows.)

In each matrix of the random ensemble, then, the i^{th} row of the random incidence matrix A must have a constant row sum r_i , and the j^{th} column a constant sum c_j . The constraint (C) makes certain entries zero in all members of the ensemble.

Forming the product AA^T , we have the sharing matrix S . The most obvious statistic with which to characterise it would be, of course, \bar{S} , the first moment (arithmetic mean) of its entries; but this quantity cannot indicate whether or not a distribution is unusual, for the simple reason that, given the constraint (B) above, it can never vary. In fact (Appendix 1, (3)):

$$m(m-1)\bar{S} = \sum_k c_k^2 - N, \tag{3}$$

where $N = \sum_k c_k = \sum_i r_i$ is the (constant) total number of species occurrences.

From the “dual” viewpoint, constraint (A) gives a similar constancy for the mean number of species shared by a pair of islands (Appendix 1, (4)):

$$n(n-1)\bar{S}' = \sum_i r_i^2 - N. \tag{4}$$

These relations could be interpreted to favor the view expressed by Diamond and Gilpin (1982), when the row and column constraints were first used in a purportedly “random” ensemble: that they “already incorporate some effects of competition”. For they form part of a null hypothesis, in which species are distributed over islands in a way that (supposedly) owes nothing to species interactions; and yet, in every such sample, the constraints (A), (B), (C) force the species pairs to share islands in such a way that the overall mean number shared obeys equation (3).

Thus a prospective island colonisation by a species could be forbidden, because the biota it finds already present on that island would bring an addition to its sharing numbers that violated equation (3). The effect of the constraints is actually more far-reaching still, as will become clear below.

The information obtainable from the shared-island numbers

The discussion below basically depends on one simple fact: when a set of numbers add up to a constant, the sum Q of their squares (and so their mean square also) is least when they are all equal. Thus any change making for greater inequality will increase Q . (Consider the case where 3 numbers must add up to 3; the sets (1 1 1), (1 2 0), (3 0 0), becoming successively less equal, have Q values which are respectively 3, 5, 9.)

For a more precise study, it is useful to have a quantity that measures this inequality. Since the quantities of interest here are the entries of the sharing matrix S , an obvious step is to take the difference of a pair of these entries, square it and then form the sum D for all pairs:

$$D = \sum \sum (S_{ij} - S_{i'j'})^2. \tag{5}$$

The summation here is over all distinct non-diagonal pairs $ij, i'j'$; however, identical “pairs” (with both $i=i'$ and $j=j'$) can be included if we wish, since they contribute zero. Thus we can let the summation indices run over the whole range (1, m), and write

$$\begin{aligned} 2D &= \sum_{i \neq j} \sum_{i' \neq j'} (S_{ij}^2 + S_{i'j'}^2 - 2S_{ij}S_{i'j'}) \\ &= m^2(m-1)^2(\bar{S}^2 - \bar{S}'^2). \end{aligned}$$

The S_{ij} 's are the entries of an incidence matrix obeying the constraints above, which were generated by some process of “random colonisation”, in which species were assumed not to interact. Now, let an interactive process of (partial or complete) mutual exclusion come into play, involving a subset of the species. This subset will be distinguished by primes; thus, some or all of the $S_{i'j'}$ must fall below the values they achieved in random colonisation. Since the row and column constraints are still in force, the mean number of islands shared cannot change (Appendix 1, (3)); hence some of the other S_{ij} must increase in value, to compensate for the reduction in the $S_{i'j'}$.

Thus some of S 's entries are reduced, while others are increased. In the original colonisation process, nothing distinguished the primed species from the unprimed, so that there is no reason to believe that their original sharing numbers tended to exceed those of the remainder; thus, we would expect the process just described to widen the separation between S values measured by D in equation (5), and thus increase \bar{S}'^2 .

Alternatively, we might note that the bracketed terms on the right side of (5) are simply the variance of the S_{ij} , which of course increases when the exclusion process spreads out the S_{ij} values, as it generally would.

Since the mean \bar{S} is constant, an increase on the right of equation (5) can occur only if \bar{S}^2 increases. Thus the exclusion process described will generally lead to an increase in \bar{S}^2 .

The qualifying phrases above (“we would expect”, “generally” ...) are needed, for by specially choosing a species pair we could make them more exclusive (reduce their number of shared islands) so that the effect was to make \bar{S}^2 actually decrease.

For example, a primed-species pair might happen to share significantly more islands than the average; if so, reducing their shared number would narrow the spread and thus reduce \bar{S}^2 . Or, even if this were not so, it might have been predominantly the larger $S_{i,j}$, which fell, and/or the smaller $S_{i,j}$, which dropped in compensation – again decreasing \bar{S}^2 . But there are no obvious biological reasons to expect either of these effects.

Moreover, it is worth noting that the process described above is “stable”, in the following sense: if we imagine the exclusion taking place in several stages, so that the $S_{i,j}$ fall repeatedly, the previous falls (and compensating rises) make it more likely that a later fall will widen the spread and increase \bar{S}^2 . Even if the primed shared-island numbers were originally (just by chance) significantly greater than the average, the falls themselves go to rectify this, and create the more usual spread of $S_{i,j}$ that will ensure an increase in \bar{S}^2 .

We turn now to consider the contrary phenomenon: an “aggregation” process. Here the original random-colonisation incidence pattern is changed so that, for a pair of species belonging to a certain subset, the number of islands shared is liable to *increase*.

Once again, the $S_{i,j}$ values for the remaining (non-aggregating) species must compensate, to keep the mean \bar{S} constant; but in this case they must *fall*. These two effects generally spread out the $S_{i,j}$ values and hence, from the above, lead to an increase in \bar{S}^2 .

(Indeed, this second process becomes identical with one of the first type, if we simply take its non-aggregating species as the primed subset in an exclusion process.)

Thus, while a study of the numbers of islands shared can reveal whether one of these two processes is at work, it cannot tell us which. While the use of \bar{S}^2 is one of the approaches that are subject to this limitation, it nevertheless has the potential to detect whether or not an observed set of island-sharing numbers can be plausibly attributed to “random” colonisation, and its usefulness for this purpose will now be tested.

Generating the random ensemble

A random sample¹ from the ensemble described above will be generated by the method of interchanges (Brualdi 1980) as follows:

Take a pair of islands, and select any species which occurs on the first of them but not on the second. Then

¹ For reasons given later, it would be better to qualify the word “random” – by, for example, always enclosing it in quotes; but this would risk being a source of irritation, and instead we refer the reader to the relevant comments below.

find, if possible, a species which occurs on the second but not on the first. Then, if we interchange the species between islands, each still occurs on the same total number of islands, and each island still contains the same number of species; that is, such an interchange leaves the constraints (A) and (B) still obeyed. But, by performing an arbitrary number of such interchanges, we generally obtain a different species distribution over the islands – with, for example, different pair-sharings.

In terms of the incidence matrix, we start with a matrix A and look for submatrices of the form

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

An interchange consists of changing the first form to the second, or the second to the first. (Note that the rows can be anywhere in the original matrix, and not necessarily neighbours; this is true of the columns also.) After an arbitrary number of such interchanges, we get a new matrix which is generally different from the original.

It can in fact be proved (Brualdi 1980) that we thus obtain *all* matrices in the random ensemble obeying constraints (A) and (B). The constraint (C) reduces this ensemble to a subset, all of whose members will likewise be obtained by such interchanges, if we retain only those resulting matrices which satisfy constraint (C).

In order to assess how likely the observed distributions would be if the null hypothesis were true, we will generate a sample from the random ensemble and use it to estimate the chance of obtaining the observed \bar{S}^2 .

Island-sharing in the Vanuatu (New Hebrides) archipelago

Data was taken from Diamond and Marshall (1976) for the distribution of 56 avian species over the 28 islands of the Vanuatu (formerly New Hebrides) archipelago. Starting with this observed incidence matrix A , 100 000 interchanges were performed as an initial randomisation, to guard against the retention of any unusual qualities from A .

The resulting matrix was then subjected to J random interchanges, and the result accepted as the first random matrix A' ; this in turn was subjected to J'' random interchanges, giving another random matrix A'' . By continuing to iterate thus, a sample of 1000 from the random ensemble was generated. The numbers $J, J'' \dots$ were themselves chosen randomly from a distribution uniform in the interval $(0.95J, 1.05J)$.

Some care is needed to arrive at a suitable value for the mean number of interchanges J , one that ensures substantial difference between successive matrices and so gives a good approximation to truly random sampling from the ensemble. A useful guide here is $q(J)$, the chance that a given entry will be left untouched by J interchanges. This is simply $[1 - 4/(28 \times 56)]^J$, since each interchange alters the value of four entries in a total of (28×56) . We have that, for $J = 100, 1000, 2000$, q is respectively 0.775, 0.078, 0.006. It would obviously be unwise to use a J value much less than about 1000.

For safety, we let J vary uniformly in (1800, 2200), obtaining results also for the range (900, 1100) to serve as a comparison. We also compared, in each case, the results for the first 500 with those for the last 500 (to confirm that no effect from the original distribution remained). These checks revealed no significant differences in the estimates given below.

We have already used overhead bars above (as in \bar{S} , \bar{S}^2) for averages over the (non-diagonal) entries of a single matrix. Now we wish to average over a set of matrices also – usually, a random sample as described above. It is helpful to use notation which keeps these two kinds of averaging distinct, and so, given a (scalar) function Y of a matrix, we write the arithmetic mean of Y over a specified set of matrices as $\langle Y \rangle$. If the function is itself a matrix average, \bar{Y} say, its sample mean will be $\langle \bar{Y} \rangle$.

From the observed distribution and the computer-generated sample of 1000 matrices, the results were:

$$\begin{aligned}\bar{S} &= \langle \bar{S} \rangle = 9.57 && \text{(constant for all matrices),} \\ \bar{S}^2 &= 148.85 && \text{(observed distribution),} \\ \langle \bar{S}^2 \rangle &= 147.10 && \text{(random sample),} \\ \text{Sampling variance of } \bar{S}^2 &\equiv \langle (\bar{S}^2)^2 \rangle - \langle \bar{S}^2 \rangle^2, && \\ &= 0.0529, && \\ \text{S.D. of } \bar{S}^2 &= 0.230 && \text{(random sample)}^2.\end{aligned}$$

Thus the observed value of \bar{S}^2 differs from the random-sample mean by 1.75, or little over 1%. But any idea that the null hypothesis can therefore be accepted is quickly dispelled, when we note that this difference – tiny though it may appear – is 7.6 times the standard deviation (both mean and s.d. being estimated from the sample). However, since the distribution of \bar{S}^2 is unknown, it is preferable to give a more transparent estimate of significance:

In the whole random sample of 1000, the maximum value of \bar{S}^2 found was only 147.79. This provides a (conservative) Monte Carlo estimate for the probability p that the actual sharing variation, as measured by $\bar{S}^2 = 148.85$, would occur, if the null hypothesis were true:

$$P < 0.001. \quad (6)$$

Thus the species distribution in the archipelago cannot plausibly be regarded as arising only from the processes implied in the null hypothesis, even after incorporating the controversial constraints A, B, C above.

Disagreement with previous findings: the reasons

The improbability of the observed \bar{S}^2 , in (6) above, contrasts sharply with the finding in Connor and Simberloff

(1979), that χ^2 for the P_k 's is in the same range as for 90 to 95% of the random ensemble. Even when using a more adequate and representative Monte Carlo sample, Gilpin and Diamond (1984) found the probability reduced only from $p > 0.90$ to $0.10 < P < 0.25$ – still well above the 0.001 found above for \bar{S}^2 . Thus a large discrepancy remains.

Our findings on this point can be summed up briefly:

1. The sampling distributions of the P_k data are inappropriate to a chi-square analysis, and the latter gives misleading results.

2. Even if a chi-square analysis were valid, the number of degrees of freedom used by Simberloff and Connor is roughly double the “best fit” value found empirically.

These points have, we believe, some general interest and lessons, and so will now be considered more fully.

Limitations of the chi-square test

We are concerned here with the chi-square test when applied to a set of variate values (usually data points), to test whether their differences from a given set of constant values, calculated on the basis of a null hypothesis, can be plausibly regarded as an error normally-distributed about zero.

When the variates are integer-valued frequencies (as in the cases of interest here), these differences can of course be only approximately normal, and rely on the normal limiting form of the binomial (or multinomial) for large sample number. However, the approximation can be a satisfactory one even for small values of P_k (Lancaster (1969), page 175), and no serious problem arises here.

Half the standardised square of such an error – i.e., the error squared, divided by its variance – is a $\gamma(1/2)$ variate; the sum of n such quantities, if they are mutually independent, is a $\gamma(n/2)$ variate; the distribution of this variate is tabulated as the χ^2 distribution for n degrees of freedom (d.f.) (Lancaster (1969), page 19). It is the question of mutual dependence which is crucial here.

When the variates are not independent, the χ^2 form can still be correct, provided the dependence either arises from linear constraints, or is given by the distribution density $\{\text{const. } x \exp - [Q(P_k, R_k)]\}$ (“multivariate normal”), where Q is a positive definite quadratic form. (Then, by a linear transformation, the χ^2 can be exhibited as a sum of squares of independent variates, with the cross-products eliminated. See Lancaster (1969), chapter II.)

In the case of concern here, the data points are the observed P_k ; to check on the null hypothesis, they are to be compared with the mean $\langle P_k \rangle$ estimated from a sample of the random ensemble. Thus we form the sum X^2 , defined by

$$X^2 = \sum (P_k - \langle P_k \rangle)^2 / \langle P_k \rangle. \quad (7)$$

Now, the condition that $\sum P_k$ must equal the total number of species pairs is a simple linear constraint and offers no problem. It is very different, however, with the

² Note that 0.0529 is the variance over the sample of a quantity \bar{S}^2 which is the mean of $(55 \times 56/2) = 1540$ squared entries; these entries are, moreover, tightly correlated by the row and column constraints and so further reduced in variance. There is no basis for comparison of this variance with the (much larger) quantity $(147.01 - 9.57^2) = 55.43$, which is the sampling variance of a single matrix entry S_{ij} , this variance then being averaged over all $i \neq j$.

row and column constraints on the incidence matrix, which give rise to quite complex relations between the P_k . We are on shaky ground, if we assume that these variables must nevertheless be distributed in the multivariate-normal form required for the chi-square test.

We have in fact noted, surveying the actual distributions of P_k ($k=0$ to 28) in a random sample of 1000 matrices generated as described above, that they do not in general impress one with any close approximation to normality; however, in view of the discussion below, we will take up space here only for one striking example, presented in figure 1:

The histograms for P_{26} , either when $P_{25}=20$ (solid line) or when $P_{25}=21$ (dashed line), could hardly be further from normality. To say they are positively skewed is an understatement, since in fact they are J -curves, falling off from an initial peak.

Such a graph is sufficient to indicate why a χ^2 test of the observed P_k can fail to detect the full abnormality present. For the test assumes that the P_k are multivariate normal, and so credits them with a scope for fluctuation that is much greater than the constraints actually permit. Thus it finds unsurprising, and close to average behaviour, a deviation from mean values which in fact is extraordinary and well into the tail.

Essentially the same point may be made in a way that allows a quantitative measure of the size of the error involved here. For this, we regard the row and column restrictions as effectively correlating the P_k 's tightly with each other, and ask how many functional linear constraints would be needed to reduce the total variation as much as these restrictions do.

To obtain an empirical estimate, we examined a sample of 1000 random matrices generated as described earlier. We first calculated $\langle P_k \rangle$, the mean of P_k over the sample, and then, for each matrix, the value of X^2 (defined in equation (7)). The mean of these values was found to be 12.72; since the mean of a χ^2 distribution is equal to its d.f., this value was taken as the number of d.f. in the χ^2 curve to be fitted to the data points X^2 .

Comparing this with the number of possible island-sharing values (29), or the d.f. used by Connor and Simberloff (27), it appears that in fact the constraints effectively cut the variation in half. Obviously, if the mean variation is put at double what it should be, rare values may well be taken as simply average behaviour.

A histogram of these X^2 values is given in figure 2, together with a χ^2 curve for 12.72 d.f. It is clear that

even this "best-fit" curve is qualitatively unsuitable to represent the data points. We can in fact measure this unsuitability (with some degree of poetic justice) by applying a χ^2 test to the fit.

It may seem paradoxical or even perverse to use χ^2 in what might be called a "second-order" way, to test a fit in which the "theoretical" values themselves come from a χ^2 curve. But in fact it is quite appropriate here; the 1000 samples giving the data in figure 2 are mutually independent. To get the correct number of degrees of freedom for this "second-order" χ^2 distribution, we reduce the number of cells (33) by 1 to allow for the parameter (12.72) calculated from the sample, and by 11 more for the lumping together of the (sparse) extreme cells. We then find that χ^2 has the value 65.9 (21 d.f.), giving a probability $\cong 5 \times 10^{-6}$.

Note that here, in seeking a (first-order) χ^2 curve to fit the distribution of X^2 , we have been exceptionally generous. We have not required the number of degrees of freedom to be theoretically justified (a difficult if not impossible task), but treated it as a parameter to be adjusted so as to *improve* the fit. Thus we have simply fixed it at a value (which happens to be fractional!) suggested by hindsight as fitting well the empirical data. These concessions make the low probability just found even more striking, confirming our earlier findings: the P_k data is intrinsically unable to be fitted by a χ^2 curve.

A complementary test

The findings above are contrary evidence of some weight, needing to be explained if one wishes to contend still that the observed Vanuatu distribution is consistent with random colonisation. However, the generation method chosen (here, the method of interchanges), while more satisfactory than previous studies on these lines, shares with them the defect that it has not been *proved* to give unbiased samples from the ensemble – that is, to be a method of truly random selection. Even if there is no reason to doubt it, it is still desirable to check the findings by using a different method. With this in mind, we have proceeded as follows:

If a particular colonisation pattern has nothing exceptional about it, then it should not be greatly affected by carrying out a few interchanges. Recall that, in an interchange, two species on two different islands are simply swapped about. Let us make a few species pairs – n , say – swap islands in this way, and look at the resulting

Table 1

No. of interchanges	(0)	10	25	100	200	400
No. in sample $\langle \bar{S}^2 \rangle$	(1)	1000	1000	1000	1000	1000
Maximum \bar{S}^2 in sample	(148.85)	148.53	148.18	147.41	147.17	147.10
No. with $\bar{S}^2 \geq$ observed	(1)	9	0	0	0	0

pattern to see if it is significantly different from the original. Repeating this a large number of times – always going back to the observed pattern (this is where it differs from the method previously used) before each batch of interchanges – we will have a sample of “perturbed” patterns from which we can judge the degree of change that these n swaps have brought about.

The results yielded by this method are shown in Table 1 (where the observed data is bracketed). To appreciate their significance, recall that an interchange of a pair of species can be made for each pair of islands on which they occur separately; in the Vanuatu archipelago, there are over 14000 such occurrences. Table 1 shows that, when we carried out a mere ten of these swaps, choosing the pairs involved at random, less than 1% of the resulting patterns had an $\langle \bar{S}^2 \rangle$ as large as the observed value, and none at all did, in a sample of 1000 after only 25 swaps. This once again constitutes a serious difficulty for any contention that the observed pattern is not exceptional.

We may phrase this difficulty in the form of a challenge. If the observed pattern really has nothing exceptional about it, then it should not be hard to construct another pattern (obeying the constraints of part 4 but otherwise, of course, independent of the observed data) with the property noted above: that, when as few as 25 species swaps are carried out at random, *none* of the resulting patterns have a value of \bar{S}^2 as large as the original's. Until other matrices with such a strong “local maximum” property have been exhibited, we are justified in regarding the observed distribution as highly exceptional.

Conclusion

The quantity \bar{S}^2 has provided a test parameter indicating that the actual Vanuatu species distribution cannot plausibly be regarded as typical of “random” colonisation. But it emerges also that, so long as the random ensemble is made to satisfy the constraints on incidence and island-diversity, there can be no overall increase or decrease in the mean number of islands shared by species pairs; any exclusions must be matched by compensatory aggregations. We have also noted above that high values of \bar{S}^2 can follow from either mutual exclusion or habitat-seeking aggregation. These effects thus make it harder to establish the nature and even the existence of species interactions.

Obviously, measures already suggested in the literature could conceivably cope with this, and allow more sensitive probing of the actual mechanism at work. These measures include the restriction of the analysis to guilds or families, and the relaxation of the incidence constraints.

Alternatively, the ensemble first used here, or the “perturbed” patterns above, can be investigated, using as tools other parameters with possibly greater discriminatory power. For instance, we have found that the “checkerboardedness”, which may indicate the degree of mutual avoidance by species pairs, can be given a

quantitative measure, and results from this line of enquiry will shortly be reported.

Acknowledgements: We wish to thank Mr. Barry Milne, of the Monash University Computing Centre, for supplying the program and the random incidence matrices used above. Our thanks go also to Dr. Geoff Watterson, for pointing out the unsatisfactory character of our original random sample, and to Professor Chris Wallace for the test used in part 8 and for many other valuable insights.

Appendix 1.

Sharing and the incidence matrix

The k^{th} island is shared by the species pair (i, j) if and only if $a_{ik} = a_{jk} = 1$. Thus

$$\begin{aligned} \text{Number of islands shared by } (i, j) &= \sum_k a_{ik} a_{jk} \\ \text{— i.e.,} \quad S_{ij} &= (\underline{A} \underline{A}^T)_{ij}. \end{aligned} \quad (1)$$

Similarly, if $[k, m]$ denotes an island pair, which share S'_{km} species:

$$\begin{aligned} \text{Number of species shared by } [k, m] &= \sum_i a_{ik} a_{im} \\ \text{— i.e.,} \quad S'_{km} &= (\underline{A}^T \underline{A})_{km} \end{aligned} \quad (2)$$

Moments of the numbers shared

$$\begin{aligned} \sum_{i \neq j} S_{ij} &= \sum_{i,j} S_{ij} - \sum_j S_{ii} \\ &= \sum_{i,j} \sum_k a_{ik} a_{jk} - \sum_i r_i \\ &= \sum_k (\sum_i a_{ik}) (\sum_j a_{jk}) - N \\ &= \sum_k c_k^2 - N \\ \text{— i.e.,} \quad m(m-1) \bar{S} &= \sum_k c_k^2 - N, \end{aligned} \quad (3)$$

where N is the (constant) total number of species occurrences. Similarly,

$$n(n-1) \bar{S}' = \sum_i r_i^2 - N. \quad \dots (4)$$

References

- Brualdi RA (1980) Matrices of zeros and ones with fixed row and column sum vectors. *Lin Algebra Appl* 33:159–231
- Connor EF, Simberloff D (1979) The assembly of species communities: chance or competition? *Ecology* 60:1132–1140
- Diamond JM (1975) Assembly of species communities. In: Cody ML, Diamond JM (eds). *Ecology and evolution of communities*. Cambridge Mass: Harvard Univ Press, pp 342–444
- Diamond JM, Gilpin ME (1982) Examination of the “null” model of Connor and Simberloff for species co-occurrences on islands. *Oecologia* 52:64–74
- Diamond JM, Marshall AG (1976) Origin of the New Hebridean avifauna. *Emu* 76:187–200
- Gilpin ME, Diamond JM (1982) Factors contributing to non-randomness in species co-occurrences on islands. *Oecologia* 52:75–84
- Gilpin ME, Diamond JM (1987) Comments on Wilson's null model. *Oecologia* 74:159–160
- Harvey PH, Colwell RK, Silvertown JW, May RM (1983) Null models in ecology. *Ann Rev Ecol Syst* 14:189–211

- Lancaster HO (1969) *The Chi-squared Distribution*. John Wiley & Sons, New York 1969
- Loehle C (1987) Hypothesis testing in ecology: psychological aspects and the importance of theory maturation. *Quart Rev Biol* 62:397–409
- Strong DR, Simberloff D, Abele LG, Thistle AB (eds.) (1984) *Ecological communities: conceptual issues and the evidence*. Princeton University Press, Princeton, New Jersey, USA
- Wilson JB (1987) Methods for detecting non-randomness in species co-occurrences: a contribution. *Oecologia* 73:579–582

Comment

I had a look at Roberts and Stone (“Island-sharing by archipelago species”). The authors used the same null model to analyse New Hebrides bird distributions as Connor and Simberloff (1979), but they used a different statistic and, unlike C & S, obtained a significant result. The authors also show that C & S incorrectly compared simulation results to the χ^2 distribution, and this explains why C & S were unable to reject the null hypothesis. The Roberts and Stone statistic is far simpler than C & S’s, and it gives us much more insight into the properties of the null model. I think that the major points of the paper are noteworthy.

The manuscript contributes little to the controversy over appropriate null models, since the authors simply adopt the one used by C & S and proceed. They hardly address and do not improve upon any of the more fundamental flaws of the C & S null model, as summarized by Gilpin and Diamond (*in* Strong et al., 1984, *Ecological Communities: Conceptual Issues and the Evidence*). Given the problems of the null model, I wonder how the authors view the place of their statistic in future distributional studies.

The work is unrelated to the method of testing for species associations using the variance ratio (Schluter 1984; *Ecology* 65:998–1005), but the difference is instructive. In the null model for the variance ratio the number of islands per species is assumed to be fixed, but the number of species per island is free to vary. In the C & S null model (used by Roberts and Stone) both the # species/island and # island/species are fixed, and hence the variance ratio is also fixed. The net association among species is thus constant, and only the associations between species in individual pairs, trios, quartets, etc. is allowed to vary.

This could cause difficulties in interpreting the Roberts and Stone statistic, and to avoid problems a clearer formulation of their alternate hypothesis would be helpful. Consider below the four species *w–z* distributed across four islands. *A* is the incidence matrix, and *B* is the same as *A* except that the central submatrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ has been flipped:

$$A = \begin{matrix} w \\ x \\ y \\ z \end{matrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

The variance among species pairs in the number of islands shared is greater in *A* ($\overline{S^2} = 2$) than in *B* ($\overline{S^2} = 1.5$), and by Roberts and Stone’s criterion *B* would represent the less “interactive” situation. However, the variance ratio $V = 0$ in both circumstances – as negative an overall association as is possible. The pairwise associations are weaker in *B* than in *A* (*A* has two positively associated pairs and four negative ones, while *B* has two negatively associated pairs), but it could be argued that in *B* the higher-order interactions compensate.

Dolph Schluter
University British Columbia,
Vancouver, Canada

Comments on Dolph Schluter’s remarks

We think it would be outside the scope of our present enquiry to consider the null-model question, important though it certainly is. The null-model problem belongs to the class of questions: “Given that the hypothesis *H* explains the facts, what conclusions can be drawn?” Here we are concerned with a prior question: “Does the hypothesis *H* agree with the facts? How can this be established?”

Re the use of the variance ratio *V* vis-a-vis our $\overline{S^2}$. In this particular model, where the number of species per sample (i.e., per island) is fixed, *V* is of course always zero, as Dr. Schluter points out, and so is not a suitable statistic. If we look at $\overline{S^2}$ in the example cited, its value varies ($[2^2 + 2^2]/6 = 4/3$ for *A*, $[4 \times 1^2]/6 = 2/3$ for *B*); we find this reasonable as an indicator of association, given the constraints, since in *A* each pair of species occurs either *always* together or *always* apart. As explained in the paper (see the end of part 4), the net *direction* (positive or negative) of the association cannot be deduced from $\overline{S^2}$ alone.

The C & S constraints stand as formidable barriers in the way of attempts to get probability distributions associated with the colonisation patterns, and this applies no less to $\overline{S^2}$. Since the methodology of the paper is to accept these constraints, we can do no better than cite the Monte Carlo estimate of its probability ($P < 0.001$, for Vanuatu). Of course, to work within these constraints is by no means to endorse their validity, and we are quite interested in the question raised, of the distribution of $\overline{S^2}$ under more consensual constraints.

A difference in the way matrices are generated could not explain our disagreement with the test used by C & S, since it would not be relevant to the way they use χ^2 incorrectly. We do not know how this latter test could in fact be applied validly, when the quantities involved (the P_k) are mutually dependent in such a complex way that the relevant sampling statistic does not even have a χ^2 distribution, according to the evidence exhibited in figure 2 and discussed in part 8.

The authors