

On testing temporal niche differentiation in carabid beetles

M. Loreau

Laboratoire de Zoologie systématique et d'Ecologie animale, C.P. 160, Université Libre de Bruxelles, 50 av. Roosevelt, B-1050 Bruxelles, Belgium

Summary. It is first shown that den Boer's (1985) claims that competition should be rejected as a factor explaining community structure are not warranted by his tests on temporal niche differentiation in carabid beetles. The power of these tests is so low that they are unlikely to detect even the most extreme pattern of niche differentiation. A major factor that affects their power is the procedure of building up an artificial species pool from the local communities to which the test is to be applied. Alternative tests are then presented, and applied to data on annual activity cycles of carabids in three forest biotopes in Belgium. Three factors are shown to strongly affect their outcome: the unit into which the time scale is divided, the criteria used to incorporate constraints, and the set of species selected. Nevertheless, they provide some evidence that there is a growing niche differentiation from the successional to the climax forest, as predicted by competition theory. Other hypotheses, however, could explain the observed pattern.

Key words: Niche differentiation – Competition – Statistical inference – Phenology – Carabid beetles

There has been for a decade a growing interest in statistical testing of community-wide patterns in ecology, especially those interpreted as evidence for competition (Harvey et al. 1983; Strong et al. 1984). This interest shows a progress in community ecology, in the face of the danger that competition theory turns into a dogma supposed to explain any pattern in nature. However, there is equally a danger that failure to devise adequate tests leads to an unwarranted rejection of competition as a factor determining community structure. This has been pointed out on several occasions in connection with various kinds of tests (see e.g. Diamond and Gilpin 1982; Roughgarden 1983; Case and Sidell 1983; Colwell and Winkler 1984).

In an earlier paper (Loreau 1986), I arrived at the conclusion that there was a growing niche differentiation in forest carabid beetle communities as these develop towards the climax stage, and that competition was a likely process involved in this trend. Annual activity cycle was one of the niche components considered. Den Boer (1985) challenged this view on the grounds that his tests using data on reproductive periods of carabids did not reveal any significant temporal separation of species, and further used this as an example supporting his general ideas on the non-significance of competition in the course of evolution. In

this paper, I first show that – irrespective of whether his ideas are right or wrong – his tests are an illustration of the above-mentioned danger of improper rejection of competition. I then present alternative tests of temporal niche differentiation which provide some evidence for my earlier conclusion, while at the same time clearly pointing to their intrinsic limitations.

Why den Boer failed to detect temporal niche differentiation

1. Null hypotheses and their testing

Although the purpose of this paper is not to discuss the general issue of the so-called “null models” in ecology, a few methodological comments are necessary to begin with. Indeed, I argue that den Boer, like others before him, misinterprets the meaning of his null hypothesis and the failure to reject it.

The basic principle of a statistical test is, first to define a null hypothesis, and second to find the probability that the observed data are consistent with the null hypothesis. If this probability lies below a certain critical level, the null hypothesis is considered rejected, while if the reverse is true, the null hypothesis cannot be rejected. It is worth recalling first that statistics do not specify what the null hypothesis must be. For example, if competition could be shown to produce a specific distribution, then the adjustment to the latter could be used as a null hypothesis in tests of competitive patterns. However, because the effects of competition can be manifold, usually no such distribution is defined. What is generally used as a null hypothesis is the consistency of the data with a distribution based on random assembling of species according to some rules that are supposedly independent of some specific effect of competition one wishes to test for.

From this it should already be clear that such distributions are never true “null models” with respect to competition – a point that has been amply emphasized by Diamond and Gilpin (1982), Roughgarden (1983) or Colwell and Winkler (1984). Any such randomly constructed distribution has an underlying structure which may be determined by many factors, including competition in other forms and with other effects. Therefore the greatest caution is necessary in interpreting the tests: all one can test is specific – and often oversimplified – predictions of competition theory. A reasoned discussion on this point, with special reference to the patterns we are concerned with here, i.e. phenological patterns, is to be found in Rathcke (1984).

Den Boer (1985) misinterprets his tests when he claims that these fulfil the condition that "hypothesis + null hypothesis have completely to fill up the sample space of possibilities", and therefore that "we can safely state that interspecific competition cannot be a necessary condition for evolutionary processes" (p. 271). His null hypothesis is: "the reproductive periods of a number of coexisting carabid species will not be separated more in time than those of carabid species in samples (of the same size) that are randomly chosen from the carabid fauna as a whole" (p. 265). As any such hypothesis, this is built according to some rules which reflect specific assumptions on the way competition acts and imply specific constraints:

(1) The hypothesis-building procedure mimics a process of community formation based on selection of species from a regional pool ("the carabid fauna as a whole"). The action of competition in such a process is to direct selection towards exclusion of too similar species. But competition may act not only through species exclusion, but also through temporal niche displacement, and the two processes can lead to different null hypotheses, as pointed out by Case and Sidell (1983) and as will be shown in the second part of this paper. Den Boer's tests would be qualitatively similar to tests for temporal niche displacement only if the regional pool was very large and diversified, and if the following implicit assumptions were met.

(2) For the procedure of random assortment of species to produce a "null" distribution with respect to competition, the regional species pool must not itself be structured by competition. But if niche separation is achieved in each local community through competition, then it is also present in the regional pool. The latter is in no way a pre-competitive pool, but only the sum of the species as they have been locally selected through competition. This does not necessarily invalidate the tests, but clearly restricts their generality and power (see Diamond and Gilpin 1982; Case and Sidell 1983; Colwell and Winkler 1984; and below).

(3) The null hypothesis of den Boer assumes that every species (or its temporal niche) in the pool is equally likely to be present in any site. This is generally not true. In particular, it is well known in carabid beetles that there are more "autumn breeders" in forests and more "spring breeders" in open sites (see e.g. Thiele 1969). As will be shown below, this difference between sites has a dramatic effect on the power of the tests.

(4) Lastly, it is assumed that the reproductive period is sufficient to describe the temporal niche of a species. But within the reproductive period of a species, there are wide variations in activity, which allow for more niche differentiation.

Hence it is obvious that den Boer's tests are only tests of a specific prediction derived from competition theory – and rather weak ones as we shall see below. Therefore, whatever their outcome, they cannot lead to any "safe statement" about the importance of competition in evolutionary processes.

A second methodological error of den Boer relates to his failure to reject his null hypothesis. He claims that this failure means the rejection of the alternative hypothesis of niche separation, hence of competition (p. 268). Statistics do not warrant such a conclusion. When a test fails to reject the null hypothesis, it does not follow that the alternative hypothesis must be rejected, because only type-I error is usually minimized to some (arbitrary) critical level. By doing

so, type-II error (i.e. the probability to mistakenly accept a false null hypothesis) may be very high, depending on the power of the test (Toft and Shea 1983). Unfortunately, many tests of community-wide patterns are very weak and biased against detection of competitive effects (Diamond and Gilpin 1982; Roughgarden 1983; Colwell and Winkler 1984). This is also the case with den Boer's, as shown hereafter. Den Boer's bias are further shown by the fact that even when a test does detect significant effects, he argues that they are likely not to be explained by the alternative hypothesis (p. 269). He might be right in some of his explanations of the effects observed, but the point is this: why to bother to test predictions of a theory if you reject the theory anyway?

2. Power of the tests

2.1. Effect of pooling species from different sites

For a test to have any power at all, it must at least be able to allow rejection of the null hypothesis. But this elementary requirement is far from assured in den Boer's tests, because of their way of building up an artificial species pool from the species actually observed in the various sites investigated. To what extent will this procedure affect the power of the tests?

Let us take a simple and extreme case of niche separation (Fig. 1): in each site there is a perfect succession of equal periods of reproduction, the overlap between two successive periods being α (here, $\alpha = 0.5$). Obviously a test that could not detect niche differentiation in such a case would be of no use whatsoever. Note that Fig. 1 is not unlike simplified pictures of temporal niches in some real carabid communities (Loreau 1986).

Den Boer's tests rely on the mean niche overlap as a measure of temporal similarity among species. Throughout this paper, we shall use the following general formula for mean overlap:

$$\bar{\alpha} = \frac{T - S}{S(S - 1)} \quad (1)$$

where S is the number of species,

$T = \sum_{i,j} \alpha_{ij}$ is the total of all overlaps,

and the negative terms result from subtraction of the S intraspecific overlaps (equal to 1). Den Boer used a different, more specific formula, which however is strictly equivalent when applied to the examples in Fig. 1. Note that in the present case the overlaps are symmetric ($\alpha_{ij} = \alpha_{ji}$); but we shall have to use the whole matrix of overlaps for the sake of demonstration hereafter.

2.1.1. All sites similar. Consider first the case (Fig. 1 a) where the species are different but their distribution in time is identical in all sites.

Let us call $\bar{\alpha}_s$ and T_s respectively the mean overlap and the total of overlaps in each site, and $\bar{\alpha}_p$ and T_p the mean overlap and the total of overlaps in the pool that results from pooling m such sites. Let us further relate $\bar{\alpha}_s$ and $\bar{\alpha}_p$ by:

$$\bar{\alpha}_p = k \bar{\alpha}_s \quad (2)$$

Since the mean overlap in the pool will also be the expected value of the mean overlap in random samples from the pool, an absolute condition for the test to detect the niche differentiation present in each site is

$$\bar{\alpha}_s < \bar{\alpha}_p, \quad \text{that is, } k > 1 \quad (3)$$

Furthermore, k can be used as an index of the power of the tests, because the larger k , the easier the rejection of the null hypothesis.

Now, there is a straightforward relation between T_p and T_s : because of the identity of the temporal distributions in all m sites, the overlap matrix for the pool simply repeats m^2 times the overlap matrix for a site; hence $T_p = m^2 T_s$. Also, if n is the number of species in a site, the number of species in the pool $S_p = nm$. Using (1) and (2), we get:

$$\begin{aligned} \bar{\alpha}_s &= \frac{T_s - n}{n(n-1)} \\ \bar{\alpha}_p &= \frac{m^2 T_s - nm}{nm(nm-1)} \\ k &= \frac{\bar{\alpha}_p}{\bar{\alpha}_s} = \frac{(mT_s - n)(n-1)}{(T_s - n)(nm-1)} \end{aligned} \quad (4)$$

Condition (3) is fulfilled if $m > 1$ and $T_s < n^2$. But T_s is of course a monotonic increasing function of α , the overlap between two successive species in a site, and lies between n when $\alpha = 0$ (the overlap matrix for a site comprises only n non-zero elements, corresponding to the intraspecific overlaps equal to 1) and n^2 when $\alpha = 1$ (the overlap matrix is filled with 1's). Thus $T_s < n^2$ for any $\alpha < 1$, and one concludes that, even though the niche differentiation in each site is also reflected in the pool, its detection by the test is possible in principle when the distributions of reproductive periods are identical in all sites.

However, the power of the test tends to be nil when α tends to 1, and, more generally, decreases with increasing α , since

$$\frac{\delta k}{\delta T_s} = -\frac{n(n-1)(m-1)}{(T_s - n)^2(nm-1)} < 0 \quad \text{for } m > 1.$$

It is also easily seen that it increases with the number of sites, since

$$\frac{\delta k}{\delta m} = \frac{(n-1)(n^2 - T_s)}{(T_s - n)(nm-1)^2}$$

which is positive for $n < T_s < n^2$, i.e. $0 < \alpha < 1$.

2.1.2. Sites of two different types. As mentioned above, temporal distributions are not likely to be identical in all sites. In particular, when both forests and open biotopes are included in the tests, as in some of den Boer, two types of sites with rather different temporal distributions are mixed to build up the pool. How do the tests behave when this effect is taken into account?

Let us again take an extreme example (Fig. 1 b), in which there are two equally abundant types of sites characterized by completely separated breeding seasons.

In this case, the overlap matrix for the pool is composed of $2(m/2)^2$ times the overlap matrix for a site, corresponding to the overlaps between species from sites of the same type, and of zero elements corresponding to the overlaps between species from sites of different types. Hence $T_p = (m^2/2) T_s$.

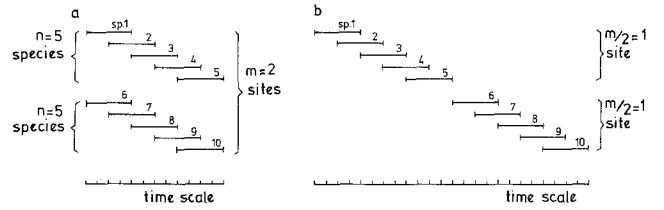


Fig. 1 a, b. Two hypothetical examples of extreme niche differentiation in local sites. **a** all sites identical; **b** two types of sites with completely separated breeding seasons

We now find:

$$k = \frac{[(m/2) T_s - n](n-1)}{(T_s - n)(nm-1)} \quad (5)$$

Here, condition (3) is fulfilled if

$$T_s < n^2 \left[\frac{2m-2}{(n+1)m-2} \right] \quad (6)$$

Since the bracketed expression in (6) is smaller than 1, there is an upper bound on T_s , hence on α , for the test to detect the pattern of niche separation present in each site.

Now, for $\alpha \leq 0.5$, it is straightforward to express T_s as a function of α . Only the overlaps between successive species are then non-zero, and there are $2(n-1)$ such overlaps. Thus $T_s = n + 2(n-1)\alpha$. Putting this into (6) leads to:

$$\alpha < \frac{n(m-2)}{2(nm+m-2)} \quad (7)$$

Since the right-hand side of (7) is always smaller than 0.5 and only approaches 0.5 as m and n tend to infinity, and since T_s increases with α , the upper bound on α imposed by (6) is lower than 0.5 and is given by (7).

Thus for large α – and always for $\alpha > 0.5$, which is the case with den Boer's data – not only would rejection of the null hypothesis be impossible, but comparison of $\bar{\alpha}_s$ and $\bar{\alpha}_p$ would even suggest that reproductive periods are more similar than predicted by chance, while their differentiation is the most extreme possible!

Of course the situation depicted in Fig. 1 b is not very realistic; all intermediate situations are possible between Fig. 1 a and 1 b. It is not difficult to show that k values then lie between (4) and (5), with increasing difficulties to satisfy condition (3) as the breeding seasons become more dissimilar. The lesson, however, is clear: random sampling from a pool that is artificially built up from different sites can reduce the power of detection of the patterns present in each site to an unknown degree, to the point of annihilating it.

2.2. Effect of shared species

In the above examples, all species were different between sites. What if species are shared by several sites? It is obvious that shared species can also reduce the power of the tests: one only needs to observe that when several sites have the same species in the first case above (Fig. 1 a), this simply amounts to reducing m , the number of sites, in (4), thereby reducing k .

But the effect of shared species is generally ambiguous. As an example, consider again the situation depicted in

Fig. 1 a, but now with every odd species shared by all sites, the other half of the species being different. It is possible to show that, when $\alpha \leq 0.5$, the k value in this case is greater than that in (4) if $m > n$ or if α is small enough. The demonstration is tedious and of little interest, but what stands out is that shared species may either increase or further decrease the power of the tests, depending on the relationship between the number of sites and the number of species per site.

2.3. Numbers of species and sites required for reaching the significance level

Although those of den Boer's tests which include all sites are clearly liable to the effect of pooling species from different sites as discussed above, his tests including only forest sites should be less so, given the generally greater homogeneity of forest communities. But even assuming that the extreme situation depicted in Fig. 1 a is then approached, this does not yet ensure that the power of the test is sufficient for rejection of the null hypothesis. Condition (3) is only a minimum condition; it is further necessary that the numbers of species and sites be large enough.

This can be investigated by deriving a direct expression for the probability P that a random sample from the species pool has a mean overlap $\bar{\alpha}_r$ lower than or equal to the observed mean overlap per site, $\bar{\alpha}_s$ – which is the probability computed in the test. This is obtained by dividing the number of possible species combinations that give $\bar{\alpha}_r = \bar{\alpha}_s$ (since the configuration in Fig. 1 is the most extreme one, $\bar{\alpha}_r$ cannot be lower than $\bar{\alpha}_s$ here), A , by the total number of possible species combinations, B : $P = A/B$. B is simply:

$$B = \binom{n}{nm} = \frac{nm!}{n!(nm-n)!}$$

The expression for A depends on α and n . For $\alpha < 0.5$, it is straightforward: only the uniform succession achieved in Fig. 1 leads to a minimum $\bar{\alpha}_r$, and there are m^n such combinations. Thus $A = m^n$. For P to lie below the critical level of 0.05 requires $n \geq 4$ for $m = 2$, $n \geq 6$ for $3 \leq m \leq 7$, and $n \geq 5$ for $m \geq 8$.

For $\alpha \geq 0.5$, the number of configurations leading to a minimum $\bar{\alpha}_r$ increases rapidly with α . Let us confine ourselves to the case $\alpha = 0.5$:

(1) For $n = 2$, again only the uniform succession leads to a minimum $\bar{\alpha}_r$, and $A = m^n$.

(2) For $n = 3$ or 4, quasi-uniform configurations with either the first two or the last two species possessing identical periods of activity also lead to a minimum $\bar{\alpha}_r$; there are $2m^{n-1}(m-1)$ such combinations, which makes a total of $A = m^n + 2m^{n-1}(m-1)$.

(3) For $n \geq 6$ and n even, n different configuration provide a minimum $\bar{\alpha}_r$:

(a) the above 3 configurations;

(b) 1 configuration with both the first two and the last two species possessing identical periods of activity;

(c) 2 configurations with either species 2 and 4 having their periods of activity both shifted to the left from the uniform configuration, or species $n-1$ and $n-3$ both shifted to the right; plus 2 configurations with 3, 4, ..., and up to $n/2 - 1$ species shifted together in the same way.

Noting that there are $m^{n-c}(m-1)^c$ possible combinations for each configuration with c species shifted together,

we have:

$$A = m^n + 2m^{n-1}(m-1) + 3m^{n-2}(m-1)^2 \quad \text{for } n=6;$$

$$A = m^n + 2m^{n-1}(m-1) + 3m^{n-2}(m-1)^2 + 2m^{n-3}(m-1)^3 \quad \text{for } n=8;$$

etc.

(4) for $n \geq 5$ and n odd, $2n-4$ configurations provide a minimum $\bar{\alpha}_r$:

(a) the same configurations as for $n+1$;

(b) 2 configurations with 1, 2, ..., and up to $(n-5)/2$ species shifted together to the left or to the right in the middle of the succession while both the first two and the last two species possess identical periods of activity. Thus:

$$A = m^n + 2m^{n-1}(m-1) + 3m^{n-2}(m-1)^2 \quad \text{for } n=5;$$

$$A = m^n + 2m^{n-1}(m-1) + 3m^{n-2}(m-1)^2 + 4m^{n-3}(m-1)^3 \quad \text{for } n=7;$$

etc.

Using these formulae, it turns out that P cannot lie below the critical level of 0.05 when $n < 8$ for any $m > 2$ (in the special case $m = 2$, $n \geq 4$ is enough). Since the number of species is probably slightly less than or close to 8 and moreover $\alpha > 0.5$ in den Boer's tests for forest sites (although he does not provide figures for n in those tests), it is no surprise that such tests systematically failed to reveal any significant effect.

In conclusion to this section, the combined effects of pooling species from different sites, of shared species and of low numbers of species make it exceedingly difficult for den Boer's tests to detect even the most extreme niche differentiation. This lack of power gives all their importance to the methodological criticisms developed earlier.

Alternative tests

1. Methods

Den Boer's tests were more specifically concerned with potential competitive exclusion occurring in each site among species from a common pool. But, as pointed out before, this is only one of the ways niche differentiation could occur. Another way is through niche shifts of the species present in each site. This opens up the possibility of formulating tests that are applicable to each site separately. Previous tests of temporal separation of species have usually been of that kind (see Fleming and Partridge 1984).

1.1. Basic principle

The basic principle of the tests devised here is the same as in Fleming and Partridge (1984). The tests rely on quantitative data of temporal activity rather than mere presence/absence of a phenological character (activity, reproduction, flowering, ...). As in den Boer's, competition is assumed to lead to a state of minimum overlap between species. The pattern thus generated should be detected by rejecting the null hypothesis that the observed overlap does not differ from the overlap in samples with randomly shifted temporal distributions.

The test is performed using a randomization procedure, which allows one to directly generate the relevant probability distribution of overlaps. In practice, the observed temporal distribution of each species in a site is allowed to shift

independently with a discrete step of a time unit and within certain time limits (see below) by computer simulation. For each possible arrangement of these distributions in a site (or for randomly drawn arrangements), the overlap statistic (see below) is computed and compared with the observed value of the statistic. The exact probability (or an estimate of this) that a value of the statistic is smaller than or equal to the observed value is then simply given by the relative frequency of the arrangements in which this event occurs. The null hypothesis is rejected if this probability lies under the significance level.

1.2. Statistics used

The most natural statistic to quantify community-wide patterns of niche differentiation is the mean pairwise overlap between species, as in den Boer's tests. Fleming and Partridge (1984) instead tested overlaps species by species, but this artificially isolates each species from the whole web of potential competitive interactions. A minimum overlap at the community level does not necessarily imply a minimum overlap for each species separately. Mean overlap is not a perfect index (Thomson and Rusterholz 1982); however, when comparisons involve constant numbers of species – as is the case here –, its drawbacks are limited and shared by other possible indices (Thomson and Rusterholz 1982). It was calculated according to Equ. (1) using Levins's (1968) index of overlap:

$$\alpha_{ij} = \frac{(\sum_h p_{ih} p_{jh})}{(\sum_h p_{ih}^2)}$$

where p_{ih} is the proportional activity during period h by species i .

Another statistic was developed following Thomson and Rusterholz's (1982) suggestion. If niche overlap is assumed to be related to the relative amount of competitive effects, the product of overlap α_{ij} by species j 's abundance, N_j (or another measure of its quantitative importance), provides a better measure of the possible competitive effects of population j on population i , because species can vary widely in quantitative importance, as is the case in carabid beetles. The mean potential competitive effect experienced by an individual from all other species in the community, which will be called the "mean competitive load", was thus estimated as:

$$\bar{c} = \frac{(\sum_{i,j} \alpha_{ij} N_j)}{(\sum_i N_i)} - 1 \quad (8)$$

Here again, as in (1), the negative term corresponds to subtraction of the intraspecific effect.

1.3. Data

The tests were performed using data from the three forest biotopes which were studied in Loreau (1986): a nearly climax beechwood, a pine plantation and a successional forest in Belgium. The data consisted in the mean annual activity cycles of the main carabid species over the 4 years 1978–81, standardized using a time unit of 1/24 year. Methods and data pertaining to the first two biotopes were described in Loreau (1985); data pertaining to the successional forest are unpublished. The total 4-year catch of a species was used as a rough measure of its quantitative importance, N , in Equ. (8). The rationale for this as well as the data are presented in Loreau (1984).

Given the "dilution effect" which may arise as a result of mixing too many species (Diamond and Gilpin 1982), the tests were repeated with three sets of species in each biotope:

- (1) all species (13 in the beechwood, 6 in the pinewood, 11 in the successional forest);
- (2) large species (≥ 10 mm: 7 in the beechwood, 3 in the pinewood and in the successional forest);
- (3) the 4 dominant species.

1.4. Constraints

Clearly, if cycles are allowed to move freely throughout the year in the simulations, no pattern of niche separation can ever be detected. Winter is a season of reduced activity for almost all species; allowing cycles to shift towards winter will necessarily result in a large number of arrangements with a niche separation greater than observed. Thus biological constraints have to be introduced. The need to define such constraints should not be regarded as a weakness of the tests. Constraints are always present; when they are made explicit, it is easier to keep in mind the limitations of the tests.

I shall make the conservative assumption that every period of activity can move within the limits set by the first and last periods of activity. It is conservative because (1) even so, some distributions will correspond to biologically unrealistic life cycles, and (2) in the absence of competition, periods of activity would likely – though not necessarily – converge towards the most favourable season. However, a further problem arises: when will a period of activity be regarded as starting and ending? Its limits are not clear-cut, and some species can even maintain a low activity throughout the year.

In order to avoid an arbitrary choice, I used several different criteria to impose constraints on the shifts of cycles. These were as follows, from the least to the more restricting one:

A: the activity peak of a species (which is generally well-marked) may not come before that of the first species or after that of the last species.

B: the period of activity of a species is regarded as starting and ending when activity reaches the level of 1% of its total activity.

C: the period of activity of a species is regarded as starting and ending when activity reaches the level of 0.25% of its total activity.

D: criteria A and B simultaneously fulfilled.

E: criteria A and C simultaneously fulfilled.

However, criteria C and E were only applied in the simulations involving the 4 dominant species, for the other species usually have too low an activity for a value of proportional activity to be as low as 0.25%.

1.5. Power

The power of the present tests is increased compared to den Boer's in three ways:

- (1) the negative effects of pooling species from different sites are eliminated;
- (2) quantitative data allow for more niche differences than qualitative ones;
- (3) there is no fixed limit to the number of possible shifts; this depends on the fineness of the subdivision of the time scale.

The latter factor plays an important role, which can be illustrated by deriving the relevant probability P that an arrangement of randomly shifted distributions has a mean overlap, $\bar{\alpha}_r$, lower than or equal to the observed mean overlap, $\bar{\alpha}_s$, in a site as described in Fig. 1. This will also allow a comparison between the powers of this test and den Boer's in the same situation, without consideration of the effect of using quantitative instead of qualitative data.

P is obtained by dividing the number of possible configurations of periods of activity that give $\bar{\alpha}_r = \bar{\alpha}_s$, C , by the total number of possible configurations, D : $P = C/D$. Let us define x as the number of time units in the period of activity of a species (e.g., $x=4$ in Fig. 1), and, as before, n as the number of species in the site and α as the overlap between two successive species. There are $(1-\alpha)(n-1)x+1$ possible positions for the period of activity of each species, which gives a total of $[(1-\alpha)(n-1)x+1]^n$ possible arrangements. But each configuration is realized by $n!$ arrangements which differ only in the identity of the successive species. Thus:

$$D = [(1-\alpha)(n-1)x+1]^n/n!$$

The expression for C here again depends on α and n . We shall confine ourselves to the two special cases $\alpha=0$ and $\alpha=0.5$ for comparison with the power of den Boer's tests.

For $\alpha=0$, $C=1$. It is readily found that P lies below the critical level of 0.05 for $x>5$ when $n=2$, and for any $x>0$ when $n\geq 3$.

For $\alpha=0.5$, applying the same line of reasoning as before with den Boer's tests and after some algebra, the expression for C is found to be:

$$C = 1 \quad \text{for } n=2;$$

$$C = x+1 \quad \text{for } n=3;$$

$$C = (x/2+1)(x/2+2)/2 \quad \text{for } n=4;$$

$$C = (x/2+1)(x+1) \quad \text{for } n=5 \text{ or } 6;$$

$$C = (x/2+1)(x^2/2+3x/2+1) - \sum_{y=1}^{x/2} y^2 \quad \text{for } n=7;$$

$$C = (x/2+1)(x^2/4+3x/2+1) - \sum_{y=1}^{x/2} y^2 \quad \text{for } n=8;$$

with x even. P can now lie below 0.5 provided that $x>10$, 8, 4, 2, and 0, respectively for $n=2, 3, 4, 5-7$, and ≥ 8 .

In all cases, P tends to zero as x tends to infinity. Thus the capacity of the test to detect the extreme niche separation involved in Fig. 1 can increase indefinitely with smaller time units. This does not mean, however, that the time scale should be divided in arbitrarily small units. A limit to this subdivision is imposed both by the degree of precision of the data and by constraints of a biological nature: for instance, it would make no sense to divide the time scale in days while year-to-year fluctuations in periods of activity are measured in weeks.

2. Results

Table 1 provides the relevant probability values as well as the number of possible arrangements of temporal distributions for each test. When this number was smaller than 4000, all arrangements were simulated; otherwise, 2000 randomly chosen arrangements were simulated.

The most striking feature of the results is their huge dependence on the set of species selected and on the criteria used to restrict the movement of the activity cycles. In the beechwood, probability values vary from 0.007 to 0.932, the conclusions of the test varying accordingly from a clear

Table 1. Results of the tests in the three biotopes

| Conditions | Beechwood | | | Pinewood | | | Successional forest | | |
|--------------------|-------------------|--------------|----------------------|-------------------|--------------|-------------------|---------------------|--------------|----------------------|
| | $P(\bar{\alpha})$ | $P(\bar{c})$ | NA | $P(\bar{\alpha})$ | $P(\bar{c})$ | NA | $P(\bar{\alpha})$ | $P(\bar{c})$ | NA |
| All species | | | | | | | | | |
| A | 0.685 | 0.729 | 9.8×10^{13} | 0.142 | 0.442 | 10^6 | 0.798 | 0.871 | 7.4×10^{11} |
| B | 0.932 | 0.918 | 1.1×10^{13} | 0.356 | 0.609 | 3.0×10^5 | 0.740 | 0.803 | 3.2×10^9 |
| D | 0.237 | 0.291 | 6.7×10^{10} | 0.029* | 0.248 | 8.1×10^4 | 0.380 | 0.593 | 5.1×10^8 |
| Large species | | | | | | | | | |
| A | 0.436 | 0.543 | 1.9×10^7 | 0.064 | 0.048* | 125 | 0.604 | 0.810 | 6561 |
| B | 0.534 | 0.601 | 2.4×10^6 | 0.104 | 0.111 | 144 | 0.147 | 0.400 | 150 |
| D | 0.083 | 0.148 | 1.5×10^6 | 0.080 | 0.060 | 100 | 0.078 | 0.300 | 90 |
| 4 dominant species | | | | | | | | | |
| A | 0.382 | 0.376 | 1.5×10^4 | 0.153 | 0.355 | 4096 | 0.891 | 0.870 | 625 |
| B | 0.167 | 0.059 | 1176 | 0.560 | 0.724 | 3960 | 0.880 | 0.920 | 100 |
| C | 0.204 | 0.042* | 432 | 0.658 | 0.793 | 8008 | 0.625 | 0.719 | 64 |
| D | 0.049* | 0.028* | 784 | 0.024* | 0.157 | 720 | 0.963 | 0.963 | 27 |
| E | 0.038* | 0.007** | 288 | 0.023* | 0.140 | 1050 | 0.938 | 0.938 | 16 |

$P(\bar{\alpha})$ =probability that a combination of randomly shifted temporal distributions has a mean overlap lower than or equal to the observed mean overlap

$P(\bar{c})$ =probability that a combination of randomly shifted temporal distributions has a mean competitive load lower than or equal to the observed mean competitive load

NA=number of possible arrangements of temporal distributions

Significances at the 0.05 and 0.01 levels are indicated respectively by * and **

See text for the definition of the conditions

Table 2. Rank correlations (Spearman coefficient) between $P(\bar{x})$, $P(\bar{c})$ and NA in Table 1

| | $P(\bar{x}) - P(\bar{c})$ | $P(\bar{x}) - NA$ | $P(\bar{c}) - NA$ |
|---------------------|---------------------------|-------------------|-------------------|
| beechwood | 0.964*** | 0.827*** | 0.900*** |
| pinewood | 0.745** | 0.418 | 0.791** |
| successional forest | 0.945*** | -0.309 | -0.264 |

** $P \leq 0.01$,

*** $P \leq 0.001$

rejection to a clear acceptance of the null hypothesis, depending on these factors. On the other hand, the type of statistic (mean overlap or mean competitive load) has a much less marked effect, as shown by the strong correlations between the probability values obtained with the two statistics (Table 2).

Despite their considerable disparity, the results are not inconsistent. In the beechwood, there is an overall decrease in probability values as one moves from the top to the bottom of Table 1, that is, from more to fewer species and from less to more severe constraints. A number that expresses both the effects of increased number of species and decreased constraints is the number of possible arrangements. This was found to be strongly correlated with probability values (Table 2). Such a trend is consistent with both the dilution effect, the greater niche complementarity among dominant species suggested in Loreau (1986), and the expectation that patterns should be better detected when biological constraints are added. Furthermore, the mean competitive load provides systematically lower probability values than does the mean overlap for the dominant species. Thus it seems reasonable to conclude that – at least among the latter species – a niche differentiation consistent with the competition hypothesis does occur in this biotope.

The results are more ambiguous in the pinewood. The significance level is reached in several tests, but the correlation between probability value and number of arrangements is weaker and significant only in the case of the mean competitive load.

The results are clearer again in the successional forest. The null hypothesis can never be rejected, and there is even a non-significant tendency to convergence of dominant species, as shown by the negative correlations between probability values and number of arrangements.

Discussion

In the first part of this paper, it was shown that den Boer's (1985) claims that competition should be rejected as a factor explaining community structure were not warranted by his tests of temporal niche differentiation. These, like any other tests, are only tests of a specific prediction derived from competition theory, and their power is so low that they are unlikely to detect even the most extreme pattern of niche differentiation. A major factor that affects their power is the procedure of building up an artificial species pool from the local communities to which the test is to be applied. When these are very different, it can even lead to the opposite conclusion that niches converge instead of diverging. Since the example used to bring out this effect is not specific to the tests examined but is of a general nature, it should serve as a warning against hasty conclusions from

tests that have resorted to this type of procedure, including previous tests by den Boer (1980) on carabid beetles.

The apparent paradox of niches being found more similar than expected by chance while their differentiation is maximum, and thus the apparent contradiction between den Boer's "coexistence principle" and the "competitive exclusion principle", are resolved by realizing that different spatial scales can be involved. On a large scale – geographical, regional, or even smaller, depending on the individuals' size –, the same main environmental factors will exert a prevailing influence on the distributions of ecological attributes of related species, and these will appear to be tightly associated. On a smaller scale, the influence of local factors such as interspecific interactions will show itself, and the same species may appear to diverge. Therefore it is essential to assess the scale on which a niche differentiation is expected.

The alternative tests performed in this paper shift the focus from the regional to the local scale. As a result, they are more powerful at detecting local niche differentiation. But they are no less specific than den Boer's, and rest on several assumptions, among which the following:

- (1) All positions of temporal distributions of activity are equally likely within a shared period; there are only constraints on the borders of this period.
- (2) The shape of the distributions is preserved while their position changes.
- (3) Niche overlap in some way expresses competitive effects.

Clearly, none of these assumptions is wholly valid. Fleming and Partridge (1984) already discussed the first two ones; the third one is implicit in all tests of niche differentiation but can only be regarded as an inevitable approximation (see e.g. Abrams 1980 for a discussion on the limitations of overlap measures).

However, other factors were shown to have a considerable effect on the outcome of the tests:

- (1) the units into which the time scale is divided: the smaller the unit, the greater the potential for detecting niche differentiation;
- (2) the criteria applied to define the constraints on the movement of the distributions: the more restricting the criteria, the easier apparently the detection of a pattern when it exists;
- (3) the set of species selected: the larger the set, the more likely the dilution of significant niche differences between dominant species.

Since these factors are either imposed by the data or chosen by the investigator, it is all the more imperative that one explores various conditions when one attempts to test for community-wide patterns, and that one remains cautious in the interpretation of the results. In the present case, radically opposite conclusions would have been drawn depending on the particular conditions of the tests.

Replication of the tests over a wide range of conditions can also offer a new perspective on the results. The correlations between probability values and the number of possible arrangements of temporal distributions give more confidence in the rejection of the null hypothesis for the dominant species in the climax beechwood. Similarly, the negative correlation found in the successional forest is an indication that temporal niches tend to converge rather than to diverge among dominant species in this biotope, and shows that failure to reject the null hypothesis is not a result of

a lack of power of the test. Therefore, the tests provide some evidence – albeit not unambiguous – for the conclusion of a growing niche differentiation from a successional to the climax stage in the carabid communities studied (Loreau 1986).

Whether this trend is really caused by competition is still another issue. All one may say is that it undoubtedly is consistent with the competition hypothesis. But other alternative hypotheses could explain it. Thus, the low productivity of the biocenoses due to the acid soil might prevent many species from becoming abundant or established. The niche differentiation between a few dominant species in the beechwood might then result from largely independent adaptations to different prey available in different seasons. Such adaptations – whether independent or not – were indeed shown in Loreau (1988). On the other hand, the weak development of the canopy and litter layers, which act as buffers against environmental variations in more mature forests, is likely to narrow the favourable season in the successional forest, while the greater heterogeneity of the latter allows presence of a more diversified fauna. This might explain the concurrent periods of activity of the dominant species in this biotope. Since all these factors are present in any case, it seems difficult to assess the specific influence of competition from tests of phenomenological patterns alone. The approach followed e.g. by Schluter and Grant (1984), based on specification of causal mechanisms leading to several alternative models, looks a more promising avenue in that direction.

Acknowledgements. I thank T.W. Schoener, who gave an impetus to this work, and P.J. den Boer, who provided critical comments on the manuscript. N. Van Mol kindly prepared the figure. This work was partly supported by a grant as senior research assistant of the National Fund for Scientific Research (Belgium), and by a NATO scientific research fellowship.

References

- Abrams P (1980) Some comments on measuring niche overlap. *Ecology* 61:44–49
- Boer PJ den (1980) Exclusion or coexistence and the taxonomic or ecological relationship between species. *Neth J Zool* 30:278–306
- Boer PJ den (1985) Exclusion, competition or coexistence? A question of testing the right hypotheses. *Z Zool Syst Evolut -forsch* 23:259–274
- Case TJ, Sidell R (1983) Pattern and chance in the structure of model and natural communities. *Evolution* 37:832–849
- Colwell RK, Winkler DW (1984) A null model for null models in biogeography. In: Strong DR, Simberloff D, Abele LG, Thistle A (eds) *Ecological communities: conceptual issues and the evidence*. Princeton University Press, Princeton NJ, pp 344–359
- Diamond JM, Gilpin ME (1982) Examination of the “null” model of Connor and Simberloff for species co-occurrences on islands. *Oecologia* 52:64–74
- Fleming TH, Partridge BL (1984) On the analysis of phenological overlap. *Oecologia* 62:344–350
- Harvey PH, Colwell RK, Silvertown JW, May RM (1983) Null models in ecology. *Ann Rev Ecol Syst* 14:189–211
- Levins R (1968) *Evolution in changing environments*. Princeton University Press, Princeton NJ
- Loreau M (1984) Composition et structure de trois peuplements forestiers de Carabides. *Acad r Belg, Bull Cl Sciences* 70:125–160
- Loreau M (1985) Annual activity and life cycles of carabid beetles in two forest communities. *Holarct Ecol* 8:228–235
- Loreau M (1986) Niche differentiation and community organization in forest carabid beetles. In: Boer PJ den, Luff ML, Mosakowski D, Weber F (eds) *Carabid beetles: their adaptations and dynamics*. Gustav Fisher, Stuttgart, pp 465–487
- Loreau M (1988) Determinants of the seasonal pattern in the niche structure of a forest carabid community. *Pedobiologia* 31:75–87
- Ratcke BJ (1984) Patterns of flowering phenologies: testability and causal inference using a random model. In: Strong DR, Simberloff D, Abele LG, Thistle A (eds) *Ecological communities: conceptual issues and the evidence*. Princeton University Press, Princeton NJ, pp 383–393
- Roughgarden J (1983) Competition and theory in community ecology. *Am Nat* 122:583–601
- Schluter D, Grant PR (1984) Determinants of morphological patterns in communities of Darwin’s finches. *Am Nat* 123:175–196
- Strong DR, Simberloff D, Abele LG, Thistle A (eds) (1984) *Ecological communities: conceptual issues and the evidence*. Princeton University Press, Princeton NJ
- Thiele HU (1969) Zusammenhänge zwischen Tagesrhythmik, Jahresrhythmik und Habitatbindung bei Carabiden. *Oecologia* 3:227–229
- Thomson JD, Rusterholz KA (1982) Overlap summary indices and the detection of community structure. *Ecology* 63:274–277
- Toft CA, Shea PJ (1983) Detecting community-wide patterns: estimating power strengthens statistical inference. *Am Nat* 122:618–625

Submitted April 9, 1988 / Accepted April 7, 1989