

Measures of Niche Overlap, I

E. SCHATZMANN

Zoologisches Museum der Universität Zürich, Winterthurerstrasse 190,
CH-8057 Zürich

R. GERRARD AND A. D. BARBOUR

Institut für Angewandte Mathematik der Universität Zürich, Rämistrasse 74
CH-8001 Zürich

[Received 20 March 1986 and in revised form 11 July 1986]

Many methods of measuring niche overlap have been proposed, such as those of Renkonen, Morisita, and Horn. In this paper, conditions are put forward which overlap indices should reasonably be expected to satisfy. A general description of a good index is given, which covers as special cases those of the well-known indices which satisfy all the conditions. A mixture of the Renkonen and van Belle & Ahmad indices is suggested, as it has many desirable properties. The problem of estimating overlap from data is discussed briefly.

1. Introduction

MANY measures have been proposed to quantify the degree to which the distributions of two species overlap. In the usual setting of niche overlap, the numbers x_i and y_i of animals of species 1 and 2 in 'resource state' i are recorded, where i runs between 1 and n , say; other information, such as the abundance r_i of each resource state may also be known (Hurlbert, 1978). These data are then combined to give a measure of overlap. Writing $X = \sum_i x_i$, $Y = \sum_i y_i$, $R = \sum_i r_i$, $p_i = x_i/X$, and $q_i = y_i/Y$, common examples can be expressed as follows.

1. (Euclidean) $1 - [\sum_i \frac{1}{2}(p_i - q_i)^2]^{\frac{1}{2}}$.
2. (Renkonen, 1938) $\sum_i \min(p_i, q_i) = 1 - \frac{1}{2} \sum_i |p_i - q_i|$.
3. (Matusita, 1955) $\sum_i (p_i q_i)^{\frac{1}{2}}$.
4. (Morisita, 1959) $2 \sum_i p_i q_i / \left(\sum_i p_i \frac{x_i - 1}{X - 1} + \sum_i q_i \frac{y_i - 1}{Y - 1} \right)$.
5. (Horn, 1966) $\frac{\sum_i [(x_i + y_i) \log(x_i + y_i) - x_i \log x_i - y_i \log y_i]}{(X + Y) \log(X + Y) - X \log X - Y \log Y}$.
6. (Modified form of Horn's index as in Ricklefs & Lau, 1980)
 $(1/2 \log 2) \sum_i [(p_i + q_i) \log(p_i + q_i) - p_i \log p_i - q_i \log q_i]$.
7. (Lloyd's (1967) index of interspecies patchiness; Hurlbert, 1978) $n \sum_i p_i q_i$.
8. (Lance & Williams, 1967) $1 - (1/n) \sum_i |x_i - y_i| / (x_i + y_i)$.
9. (Pianka, 1973) $\sum_i x_i y_i / [(\sum_i x_i^2)(\sum_i y_i^2)]^{\frac{1}{2}}$.
10. (van Belle & Ahmad, 1974) $2 \sum_i [p_i q_i / (p_i + q_i)]$.

11. (Grassle & Smith, 1976)

$$\frac{2 \sum_i \left[1 - \binom{X-x_i}{m} / \binom{X}{m} \right] \left[1 - \binom{Y-y_i}{m} / \binom{Y}{m} \right]}{\sum_i \left[1 - 2 \binom{X-x_i}{m} / \binom{X}{m} + \binom{X-x_i}{2m} / \binom{X}{2m} \right] + \sum_i \left[1 - 2 \binom{Y-y_i}{m} / \binom{Y}{m} + \binom{Y-y_i}{2m} / \binom{Y}{2m} \right]}$$

where m may be any integer, $m \geq 1$: for $m = 1$, the formula yields Morisita's index.

12. (Hurlbert, 1978) $(R/XY) \sum_i (x_i y_i / r_i)$.

13. (Petraitis, 1979)

$$\exp \left\{ \frac{1}{X+Y} \sum_i \left[(x_i + y_i) \log \left(\frac{x_i + y_i}{X+Y} \right) - x_i \log \frac{x_i}{X} - y_i \log \frac{y_i}{Y} \right] \right\}.$$

Which of these, or other, indices to use is still much in debate (Abrams, 1980; Hurlbert, 1978; Ricklefs & Lau, 1980; Smith & Zaret, 1982; Wolda, 1981).

The problem may conveniently be divided into two parts. The first is to decide how, if perfect information about the distributions of the two species were available, an index of overlap should be defined. The second is to find a good way of estimating the chosen index of overlap, given the imperfect information actually available. In Sections 2-4 of this paper, the first part of the problem is addressed in detail. A rationale is put forward for assessing overlap indices, a general family of good indices is proposed, and a basis on which an appropriate index may be chosen is suggested. In Section 5, the problem of estimating overlap from sample data is discussed briefly. The more mathematical aspects of the argument are presented in Gerrard & Barbour (1986), referred to hereafter as Part II.

2. Underlying axioms

We assume that the two species are distributed over a region \mathcal{A} which consists of J biologically homogeneous sub-regions $\mathcal{A}_1, \dots, \mathcal{A}_J$. With perfect information one would know the population density of each species over each sub-region \mathcal{A}_j ($1 \leq j \leq J$). If L_j and M_j denote respectively the numbers of individuals of species 1 and 2 living in sub-region \mathcal{A}_j , the respective population densities over \mathcal{A}_j would most naturally be defined as L_j/a_j and M_j/a_j , where a_j is the area of \mathcal{A}_j ; however, this is by no means the only reasonable definition: a measure of volume may be more appropriate for the denominator than a measure of area, or, more significantly, the amount of a given resource in the region (Hurlbert, 1978), so we shall define the population densities in sub-region \mathcal{A}_j by

$$\lambda_j = L_j/\vartheta_j \quad \text{and} \quad \mu_j = M_j/\vartheta_j \quad (1 \leq j \leq J), \quad (2.1)$$

where the ϑ 's may be any positive constants, but usually correspond to some natural measure, such as area, evaluated over the sub-regions.

Note that the region \mathcal{A} need not represent a genuine geographical area under examination. For instance, there may be J distinguishable standard habitat types which are considered of importance with respect to the two species, which all occur in a geographical area \mathcal{G} . It is possible to let \mathcal{A} be the same as \mathcal{G} , and set \mathcal{A}_j to be that part of \mathcal{G} which has habitat type j , $1 \leq j \leq J$: in this case, supposing ϑ to be chosen as area measure, ϑ_j would be the area of habitat type \mathcal{A}_j . Alternatively, for comparability with other studies, it may be convenient to take \mathcal{A} to consist of sub-regions $\mathcal{A}_1, \dots, \mathcal{A}_J$ in \mathcal{G} of prescribed areas, \mathcal{A}_j again being of habitat type j , so that \mathcal{A} is only a part of \mathcal{G} . Overlap on \mathcal{A} then represents a kind of standardized measure of overlap for the species, as found in \mathcal{G} , in the calculation of which all habitat types are given pre-determined weights, corresponding to the areas of the sub-regions. Thus, if all the sub-regions are chosen to have the same area, each habitat type is given equal weight. The densities of each species per unit area would then in practice be estimated by sampling in each habitat type within \mathcal{G} , as discussed in Section 5. This second procedure has the advantage that it eliminates as far as possible the effects of variation in habitat proportions, when comparing the niche overlap of the two species between one geographical area and another.

A measure of overlap is to be constructed by comparing the densities λ_j and μ_j for $1 \leq j \leq J$. First define ℓ and m to be the average population densities over the whole space:

$$\ell = \left(\sum_j \lambda_j \vartheta_j \right) / \theta = \sum_j L_j / \theta, \quad (2.2)$$

$$m = \left(\sum_j \mu_j \vartheta_j \right) / \theta = \sum_j M_j / \theta, \quad (2.3)$$

where $\theta = \sum_j \vartheta_j$, and then define the standardized densities

$$f_j = \lambda_j / \ell, \quad g_j = \mu_j / m. \quad (2.4)$$

The f 's and g 's thus describe the distributions of the two species over the various sub-regions relative to their overall abundance, and can be seen in a mathematical sense as probability densities with respect to the measure ϑ_j / θ , in that they take only non-negative values and

$$\sum_j \vartheta_j f_j / \theta = \sum_j \vartheta_j g_j / \theta = 1.$$

The idea of overlap is associated naturally with the relative distributions given by the f 's and g 's, in the sense that overlap should remain unchanged if the f 's and g 's remain the same, even if either or both of the overall average population densities ℓ and m are changed. We shall therefore take as our first axiom:

AXIOM 1 [Relative abundances] A measure of overlap should depend at most on the f 's, the g 's, and the ϑ 's.

This axiom, like those that follow, is to be understood in the sense that it should be accepted unless there are strong reasons for rejecting it inherent in the situation under consideration; in other words, a measure of overlap designed for

general use should reasonably comply with the axiom. Axiom 1 reflects the spirit of the argument in, for example, Smith (1982).

Further axioms have been suggested, of which we take the following:

AXIOM 2 [Splitting of sub-regions] If the population densities are identical on two different sub-regions—that is, if $\lambda_1 = \lambda_2$ and $\mu_1 = \mu_2$ on sub-regions \mathcal{A}_1 and \mathcal{A}_2 —then the measure of overlap should be unchanged if we amalgamate \mathcal{A}_1 and \mathcal{A}_2 into a region \mathcal{A}^* over which $\vartheta^* = \vartheta_1 + \vartheta_2$, $\lambda^* = \lambda_1 = \lambda_2$, $\mu^* = \mu_1 = \mu_2$.

This property is suggested in, for example, Abrams (1980).

AXIOM 3 [Repeated observations] If a region \mathcal{A}' is adjoined to \mathcal{A} , consisting of sub-regions $\mathcal{A}'_1, \dots, \mathcal{A}'_J$ such that $\vartheta'_j = \alpha\vartheta_j$ for each j and for some constant $\alpha > 0$, and if $\lambda'_j = \lambda_j$ and $\mu'_j = \mu_j$ for each $1 \leq j \leq J$, then the overlap measure derived from the combined region $\mathcal{A} \cup \mathcal{A}'$ should be the same as that measured from \mathcal{A} alone.

In particular, corresponding to the case $\alpha = 1$, two separate and identical regions \mathcal{A} and \mathcal{A}' yielding exactly the same information about relative distribution should give the same overlap whether taken together or singly: that is, adding a second investigation exactly confirming the results of the first should not lead to a change in the degree of overlap.

AXIOM 4 [Empty sub-regions] If a sub-region \mathcal{A}_{J+1} is adjoined to \mathcal{A} over which $\mathcal{A}_{J+1} = M_{J+1} = 0$, then the overlap measured from $\mathcal{A} \cup \mathcal{A}_{J+1}$ should be the same as that measured from \mathcal{A} alone.

This property is advocated by Abrams (1982), whereas Hurlbert (1978) considers it undesirable.

AXIOM 5 [Calibration] If $\lambda_j = \mu_j$ for $1 \leq j \leq J$, then the overlap measure should be unity; if $\lambda_j \mu_j = 0$ for $1 \leq j \leq J$, the overlap should be 0; in any other circumstances the overlap should lie strictly between 0 and 1.

This last is a calibration axiom, making all states of perfect overlap equivalent, and all states of no overlap equivalent.

The first three axioms already limit the choice of measure of overlap strongly. Axiom 1 states that it must be a function of triples (f_1, g_1, ϑ_1) , $(f_2, g_2, \vartheta_2), \dots, (f_J, g_J, \vartheta_J)$, and the effect of Axioms 2 and 3 is to restrict the choice further to functions of the discrete probability distribution defined by assigning probability ϑ_j/θ to the point (f_j, g_j) , $1 \leq j \leq J$, with the understanding that when two or more of the points coincide, the corresponding probabilities are to be added. That is, mathematically speaking, if the distinct points generated by the set of points $(f_1, g_1), \dots, (f_J, g_J)$ —that is, ignoring repeats—are denoted by $(\bar{f}_1, \bar{g}_1), \dots, (\bar{f}_K, \bar{g}_K)$, where $K = J$ if all the original points were distinct, then the overlap measure must be a function of the discrete probability distribution P defined by

$$P[(\bar{f}_k, \bar{g}_k)] = \sum_{j \in \bar{f}_k} \vartheta_j / \sum_{j=1}^J \vartheta_j = \sum_{j \in \bar{f}_k} \vartheta_j / \theta \quad (1 \leq k \leq K)$$

where \mathcal{V}_k denotes the set of indices j for which $f_j = \bar{f}_k$ and $g_j = \bar{g}_k$. Such overlap measures are said to satisfy Condition C_1 ; it is exactly equivalent to Axioms 1-3 taken together.

Now if $f_j = g_j$ for all $1 \leq j \leq J$, the case of perfect overlap, the points (\bar{f}_k, \bar{g}_k) will all have $\bar{f}_k = \bar{g}_k$, so that the corresponding probability distribution concentrates all its probability on the diagonal $\mathcal{D} := \{(\omega, \omega), \omega \geq 0\}$. Thus, in view of Condition C_1 , we are led to considering ways of quantifying how close a distribution is to being concentrated on the diagonal \mathcal{D} .

Much the most natural way of attempting this is to choose a function $\phi(x, y)$ which, for each pair (x, y) , describes how near it is to being on \mathcal{D} , and then to take the expectation of ϕ with respect to the given distribution as a measure of how near the distribution itself is to being concentrated on \mathcal{D} , thus obtaining an intuitively appealing average degree of closeness. Note that such a function ϕ should reasonably be symmetric, satisfying $\phi(x, y) = \phi(y, x)$ for all x, y . Translated into our terms, this gives a candidate set of measures of overlap of the form

$$C_2: \sum_j \vartheta_j \phi(f_j, g_j) / \theta, \tag{2.5}$$

with ϕ a symmetric function. Then $\vartheta_j \phi(f_j, g_j) / \theta$ can be interpreted as the amount of overlap contributed by sub-region \mathcal{A}_j alone. These are not the only possibilities. For instance, one could take a measure such as

$$\min_{1 \leq j \leq J} \phi(f_j, g_j) \tag{2.6}$$

which does not involve the third argument of the triples, but such functions tend to be biologically unreasonable, for instance in laying too much weight on a single extreme sub-region, and we shall not consider them further. Or one could try taking, say, the median of the distribution of $\phi(x, y)$ when (x, y) has the distribution P . But combining comparisons on the individual sub-regions by an appropriate summation seems to be a universally accepted method, and so we shall restrict ourselves to measures based on the form C_2 , but perhaps extended by taking a function of one or more quantities of the form (2.5).

Axiom 4 has a further restrictive effect on the class C_2 . An expression of the form C_2 satisfies Axiom 4 if and only if the function ϕ can be written in the form

$$\phi(x, y) = \phi(0, 0) + \frac{1}{2}(x + y)\psi(|x - y|/(x + y)) \quad (x + y > 0, \quad x, y \geq 0) \tag{2.7}$$

for some function ψ whose single argument takes values in the interval $[0, 1]$ (see Part II). The constant $\phi(0, 0)$ may be taken without loss of generality to be 0. Thus, in view of Axiom 4, we confine attention to functions of quantities of the form

$$C_3: \sum_j \vartheta_j \frac{1}{2}(f_j + g_j)\psi(|f_j - g_j|/(f_j + g_j)) / \theta, \tag{2.8}$$

for some function $\psi(z)$ ($0 \leq z \leq 1$). Axiom 5, the calibration axiom, demands that $\psi(0) = 1$, $\psi(1) = 0$, and $0 < \psi(z) < 1$ for $0 < z < 1$. Further, the form (2.8) suggests that ψ should be decreasing: considering pairs (f_1, g_1) and (f_2, g_2) with

$f_1 + g_1 = f_2 + g_2$, the pair with the smaller value of $f-g$ should clearly give the greater contribution to overlap.

The form C_3 has an invariance property which appears stronger than Axiom 4 (but in fact merely emphasises the axiom's power) in that it depends only on the raw data L_j and M_j , and not on the θ_j . More explicitly, if $\pi_j = L_j / \sum_k L_k$ and $\rho_j = M_j / \sum_k M_k$, then C_3 can be rewritten, by replacing f_j and g_j with expressions in terms of L_j and M_j via (2.1)–(2.4), as

$$C'_3: \sum_j \frac{1}{2}(\pi_j + \rho_j)\psi(|\pi_j - \rho_j|/(\pi_j + \rho_j)). \quad (2.9)$$

Since it is difficult to suggest a methodical way of choosing the θ 's in many cases, and the choice actually made must be to some extent arbitrary, the invariance of C'_3 under changes in the θ_j is extremely useful. Notice also that C'_3 implies an extension of Axiom 2: the overlap measure remains unchanged when \mathcal{A}_1 and \mathcal{A}_2 are amalgamated not only if $\lambda_1 = \lambda_2$ and $\mu_1 = \mu_2$ but also if $\lambda_1/\mu_1 = \lambda_2/\mu_2$, or, equivalently, if $\pi_1/\rho_1 = \pi_2/\rho_2$.

At this point, it is interesting to observe to what extent the measures quoted in the introduction would be consistent with Axioms 1–5, in the presence of full information.

1. The Euclidean index yields

$$\begin{aligned} 1 - \left(\sum_j \frac{1}{2}(\pi_j - \rho_j)^2 \right)^{\frac{1}{2}} &= 1 - \left(\sum_j \frac{1}{2}\theta_j^2(f_j - g_j)^2 \right)^{\frac{1}{2}} / \theta \\ &= 1 - \left(\sum_k (\bar{f}_k - \bar{g}_k)^2 \sum_{j \in \mathcal{V}_k} \frac{1}{2}\theta_j^2 \right)^{\frac{1}{2}} / \theta, \end{aligned}$$

which is not of the form C_1 : indeed, Axiom 2 is not satisfied. It also violates Axiom 5.

2. The Renkonen index can be expressed in the form C'_3 , with $\psi(z) = 1 - z$, and so satisfies all of Axioms 1–5.

3. Matusita's index is of the form C'_3 , with $\psi(z) = (1 - z^2)^{\frac{1}{2}}$, and hence satisfies Axioms 1–5.

4. The fundamental comparison element of Morisita's index, the sum

$$\sum_j \pi_j \rho_j = \frac{1}{2} \sum_j (\pi_j + \rho_j) [2\pi_j \rho_j / (\pi_j + \rho_j)],$$

is not of the form C'_3 : indeed, writing it in terms of the f 's and g 's for a given underlying measure θ gives the formula

$$\sum_j \theta_j^2 f_j g_j / \left(\sum_j \theta_j \right)^2,$$

which is not of the form C_1 . The denominator, in the original or in Horn's (1966) form, is designed to satisfy Axiom 5, and Axiom 4 holds: but its failure to satisfy Axiom 2 leads to the observed drawbacks (under-weighting of small probabilities, Grassle & Smith, 1976), noted in the literature.

5. Horn's index fails to satisfy Axioms 1 and 5.

6. The modified Horn index satisfies C'_3 with

$$\psi(z) = 1 - [(1+z) \log(1+z) + (1-z) \log(1-z)] / 2 \log 2,$$

and hence satisfies Axioms 1-5.

7. Lloyd's index fails to satisfy Axioms 2, 4, and 5.

8. Lance & William's index fails to satisfy Axioms 1, 2, and 4.

9. Pianka's index fails to satisfy Axiom 2.

10. van Belle & Ahmad's index is of the form C'_3 with $\psi(z) = 1 - z^2$, and hence satisfies Axioms 1-5.

11. Grassle & Smith's index, based on the comparison $\sum_i [1 - (1 - \pi_i)^m] \times [1 - (1 - \rho_i)^m]$, fails like Morisita's to satisfy Axiom 2.

12. Hurlbert's index is of the form C_2 with $\vartheta_j = r_j$ and $\phi(x, y) = xy$: however, it fails to satisfy Axioms 4 and 5.

13. Petraitis's index fails to satisfy Axioms 1 and 5.

It should be stressed that indices based on the form C_3 are not the only ones satisfying Axioms 1-5, but derive also from the idea of averaging the closeness of f_j to g_j ($1 \leq j \leq J$). For instance, any function of the distribution on $[0, 1]$ which, for each $1 \leq j \leq J$, attaches mass $\frac{1}{2}(\pi_j + \rho_j)$ to the point $|\pi_j - \rho_j| / (\pi_j + \rho_j)$ (adding the masses in the case of coincident values) satisfies Axioms 1-4. Axiom 5 is satisfied if in addition the function takes values between 0 and 1, with the value 0 only for the distribution assigning all its mass to the point 1 and the value 1 only for the distribution assigning all its mass to the point 0.

The principal dissenting voice with regard to the treatment of overlap is that of Hurlbert (1978), in respect of Axiom 4. The difference is fundamental and hinges on different underlying motivations. Here, we have considered overlap as a property of the distributions of two species, which may be of use as a general descriptive aid. Hurlbert, on the other hand, argues that overlap is of interest as a tool to study competition for scarce resources. Although, in this regard, we might suppose that the amount of rivalry exhibited between two species per unit resource were $\eta := H \sum_j \lambda_j \mu_j \vartheta_j / \sum_j \vartheta_j$, the constant of proportionality H must unfortunately depend on factors such as aggressiveness, range of visibility of one species by the other, distribution of resources, and so on. Hurlbert's index is an attempt to derive a measure of overlap from η without knowledge of H , achieved only at the cost of sacrificing Axioms 4 and 5. It seems to us that competition is better analysed with reference to all parameters and by means of appropriate equations, and that overlap, for whatever reason it may occur, is of interest in its own right as a descriptive tool, for which it makes sense to use Axiom 4.

3. Further properties

In this section we consider two further axioms, which are not so obviously indispensable as Axioms 1-5. Both attempt to make the idea of overlap more precise—they are to be viewed in the same spirit as the other axioms, bearing in mind that the collection of possible indices of the form C'_3 is so extensive that we can afford to be selective.

The idea of measuring overlap is to quantify the similarity between two distributions. Similarity, though, is rather a vague concept. The complementary concept of difference, on the other hand, may be linked in with the precise mathematical notion of a distance. If a function $d(z_1, z_2)$ is to be a distance between the elements of a set Z , then for each $z_1, z_2, z_3 \in Z$ the following must be satisfied:

- (i) $d(z_1, z_2) \geq 0$, with $d(z_1, z_2) = 0$ if and only if $z = z_2$;
- (ii) $d(z_1, z_2) = d(z_2, z_1)$;
- (iii) $d(z_1, z_2) + d(z_2, z_3) \geq d(z_1, z_3)$.

Only the third of these requires comment: it stipulates that the distance from z_1 to z_3 via some other point z_2 must be at least as great as the direct distance from z_1 to z_3 , a most natural assumption which can be seen as a generalization of the old axiom that the shortest distance between two points is a straight line. It is by no means clear how one would go about constructing an equally appealing complementary condition which a similarity should satisfy.

We now introduce an axiom which fixes the idea that similarity and distance are to be in some sense complementary. A measure of overlap satisfying Condition C_3 can be viewed as a measure of similarity, S say, between pairs of probability distributions π and ρ over the natural numbers, because of the representation C'_3 . This suggests linking $S(\pi, \rho)$ with $d(\pi, \rho)$, for some distance d . Let Z be the collection of all probability distributions over the natural numbers.

AXIOM 6 [Distance] $S(\pi, \rho) = \sigma(d(\pi, \rho))$, where d is a distance function on Z satisfying $d(\pi, \rho) \leq 1$ with $d(\pi, \rho) = 1$ if and only if $\pi_j \rho_j = 0$ for all j , and where σ is some continuous, strictly decreasing function on $[0, 1]$, with $\sigma(0) = 1$ and $\sigma(1) = 0$.

Thus similarity is required to be a decreasing function of a distance. Or one could require more:

AXIOM 6⁺ [Strong distance] $S(\pi, \rho) = 1 - d(\pi, \rho)$, where d is as above.

This specifies the relationship between similarity and distance in the simplest possible way. It will be seen that Axiom 6⁺ is considerably more restrictive than Axiom 6.

Of the four indices Renkonen, Matusita, modified Horn, and van Belle & Ahmad which satisfy Axioms 1-5, all also satisfy Axiom 6; the Renkonen index alone satisfies Axiom 6⁺, and is associated with the total variation distance. The other three all satisfy Axiom 6 with $\sigma(d) = 1 - d^2$: the Matusita index is associated with the Hellinger distance, and it is proved in Part II that the other two are also associated with distances. Notice that if, for any of these three similarity indices, S , we define $\tilde{S} = 1 - (1 - S)^{\frac{1}{2}}$, then \tilde{S} satisfies Axiom 6⁺ as well as Axioms 1-5.

The second axiom qualifying similarity is essentially Wolda's (1981) property of linear increase. Let $v_{m,n}$ denote the uniform probability distribution on the integers $m + 1, m + 2, \dots, n$.

AXIOM 7 [Extended calibration] $S(v_{0,n}, v_{m,m+n}) = 1 - m/n$ for $1 \leq m \leq n$.

Thus, given two species, each having constant density over n habitats, Axiom 7 specifies a linear relationship between their overlap and the number $n - m$ of habitats the species have in common. The axiom attempts to standardize what is meant by values of overlap other than 0 and 1. It should, however, be stressed that different similarity measures S satisfying Axiom 7 may still produce different results, when the distributions being compared are not both of the form v .

Axiom 7 is satisfied by the four indices mentioned above, but not by the transformed indices \tilde{S} .

One use of measures of overlap is as a basis for automatic procedures, such as multi-dimensional scaling, which depict the relationships between species by representing each one as a point on a low- (usually two-) dimensional map, on which the closeness of a pair of points reflects as closely as possible the overlap between the corresponding species. Since the notion of a map as incorporated into such procedures is intrinsically connected with that of distance, in the mathematical sense, it would seem natural for such applications to use measures of overlap satisfying Axiom 6 at least. Conversely, in automatic procedures in which a distance measure is implicit, it should be consistent with Axioms 1-5 (with the obvious adjustments to Axiom 5), if the procedure is to be applied to data of the kind considered in this paper. In particular, correspondence analysis, as interpreted by the French school (see Greenacre, 1984), uses a distance function which depends on all the distributions in the species group under consideration, and not just on the two being compared, and, in the case of only two distributions, with the information of Section 2, reduces to

$$\left(\sum_j [(\pi_j - \rho_j)^2 / (L_j + M_j)] \sum_j (L_j + M_j) \right)^{\frac{1}{2}}$$

which fails to satisfy Axiom 1. Such an interpretation of correspondence analysis must therefore be treated with caution.

4. Construction of indices

In this section we outline how one would go about constructing an index satisfying Axioms 1-7 (though not necessarily 6⁺).

The four indices discussed at the end of the previous section are all of the form $1 - d^p$, where $p = 1$ for the Renkonen index and $p = 2$ for the other three, and where

$$d(\pi, \rho) = \left\{ \sum_j \frac{1}{2}(\pi_j + \rho_j) \left[1 - \psi \left(\frac{|\pi_j - \rho_j|}{\pi_j + \rho_j} \right) \right] \right\}^{1/p} \quad (\pi, \rho \in Z), \quad (4.1)$$

is a distance, ψ being a suitably chosen decreasing function satisfying $\psi(0) = 1$, $\psi(1) = 0$, $0 < \psi(z) < 1$ otherwise.

All indices of this form for some $p \geq 1$ are said to belong to Class C_4 . They have all the desirable properties dictated by Axioms 1-7 (though Axiom 6⁺ is only satisfied if $p = 1$), and provide a broad range of possible measures of overlap.

To begin with, let $u(z) = z^{-1}[1 - \psi(z)]$, so that (4.1) becomes

$$d(\pi, \rho) = \left[\sum_j \frac{1}{2} |\pi_j - \rho_j| u\left(\frac{|\pi_j - \rho_j|}{\pi_j + \rho_j}\right) \right]^{1/p}. \quad (4.2)$$

The selection problem thus consists of choosing a suitable function u on $[0, 1]$ with $u(1) = 1$, $u(z) \leq 1$, and $zu(z)$ non-decreasing. Axioms 1–5 and 7 are automatically satisfied by any choice of u , so the only restriction is provided by Axiom 6, and we may choose any u provided that d is then a distance. The question of how to fulfil the requirement is investigated in Part II, where various families of overlap indices are discussed.

u may be interpreted as follows. Consider the contribution to the similarity $1 - d^p$ from a sub-region \mathcal{A}_j in which $\pi_j = 0$, $\rho_j = \delta$, and compare it to the contribution from \mathcal{A}'_j , in which $\pi_j = \beta$ and $\rho_j = \beta + \delta$ for some $\beta > 0$. In the former case, the contribution is $-\frac{1}{2}\delta$, and in the latter $-\frac{1}{2}\delta u(z)$, where $z = \delta/(2\beta + \delta)$. The quantity

$$u(z) = u(|\pi_j - \rho_j|/(\pi_j + \rho_j))$$

thus represents the contribution to the similarity $1 - d^p$ of a difference of δ between π_j and ρ_j , expressed relative to the maximal possible contribution from a difference of δ . Thus, by choice of $u(z)$, we can weight the contribution of a difference $\delta = |\pi_j - \rho_j|$ in comparison to the contribution of such a difference when it represents the whole of $\pi_j + \rho_j$. Note that the contribution

$$[|\pi_j - \rho_j|/(\pi_j + \rho_j)]u(|\pi_j - \rho_j|/(\pi_j + \rho_j))$$

per individual to the similarity from sub-region \mathcal{A}_j remains constant whenever π_j/ρ_j is held constant.

The Renkonen index has the same value $u(z) = 1$ for all $0 \leq z \leq 1$, and it has been criticized for exactly this reason, in that it gives the same contribution to a difference of δ , whatever the value of $z = \delta/(\pi_j + \rho_j)$, instead of allowing the contribution to increase with z , as can be done within the class C_4 . In the spirit of this criticism, we propose the following two assumptions:

ASSUMPTION 8 The function u should be non-decreasing,

or (a strengthened version):

ASSUMPTION 8⁺ The function u should be increasing.

A biological argument to support Assumption 8⁺ is as follows. When comparing data from two apparently similar regions, or from the same region at different times, one would expect to find changes, even if sampling error were negligible, resulting from random environmental variation. Such variation would tend to cause greater changes in the differences $\pi_j - \rho_j$ corresponding to larger values of $\pi_j + \rho_j$ than in those corresponding to smaller values of $\pi_j + \rho_j$. Assumption 8⁺ can be interpreted in this context as asserting a preference for indices which are more stable with respect to random environmental variation than the Renkonen index.

Another biological reason for accepting Assumption 8⁺ is the following.

Habitats where both species are abundant can be thought of as being generally favourable, placing less stress on the organisms than those in which both species are rare. In the latter case, differences in abundance are more likely to correspond to differing abilities to adapt to the harsh environment, thus providing information of biological interest in the assessment of species overlap; whereas, in the former case, the same difference may represent a relatively trivial difference in competitiveness. Thus, for example, considering two sub-regions with equal θ -measure, the difference between $f_1 = 0.1$ and $g_1 = 0.0$ should be assessed as being more significant than the difference of the same magnitude between $f_2 = 10.1$ and $g_2 = 10.0$. Using (2.8) and the definition of $u(z)$ preceding (4.2), it follows that C_4 indices can be expressed in terms of u and the relative abundances as

$$1 - \sum_j \theta_j^{\frac{1}{2}} |f_j - g_j| u(|f_j - g_j| / (f_j + g_j)),$$

from which it can be seen that an increasing function u has precisely the effect argued for above.

The Renkonen index is shown in Part II to have the interesting property that it is the minimal similarity of class C_4 to satisfy Axiom 6^+ . Any other function u yielding a class C_4 similarity with $p = 1$ has $u(z) > 1$ for all $z < 1$. The Renkonen index is therefore the only similarity of class C_4 to satisfy Axiom 6^+ and Assumption 8. All other C_4 similarities satisfying Assumption 8 have the form $1 - d^p$ for some $p > 1$, and fail to satisfy Axiom 6^+ : in such cases, the alternative choice of $1 - d$ as similarity would retain Axiom 6^+ , but at the cost of Axiom 7. Hence, if Assumption 8^+ is considered to be biologically important, it is necessary to take $p > 1$ and a choice has to be made between Axioms 6^+ and 7.

Figure 1 shows the functions $u(z)$ for the four indices mentioned in the introduction which are of class C_4 . Except for that of the Renkonen index, all tend to 0 as $z \rightarrow 0$, which may seem to be a rather extreme downweighting of small differences between larger proportions. Indeed, the function u for the Matusita index is shown in Part II to give the greatest downweighting possible, at each value of z , of any C_4 similarity with $p \leq 2$. Both the modified Horn index and Matusita's index are very fast to downweight differences as soon as z falls below 1, which may also be considered rather undesirable, whereas that of van Belle & Ahmad is reasonable in this respect. Indeed, a linear function $u(z)$ has a certain simplicity and intuitive appeal, and we therefore propose the simple family of class C_4 indices based on the functions $u_\alpha(z) = \alpha + (1 - \alpha)z$ ($0 \leq \alpha \leq 1$) of which the Renkonen index ($\alpha = 1$) and the van Belle & Ahmad index ($\alpha = 0$) are extreme members. That these are indeed C_4 similarities for all α is shown in Part II. The explicit form of the combined indices is given by

$$I_\alpha := \alpha \sum_j \min(\pi_j, \rho_j) + 2(1 - \alpha) \sum_j [\pi_j \rho_j / (\pi_j + \rho_j)], \quad (4.3)$$

being simply a weighted average of the Renkonen and van Belle & Ahmad indices, and our subjective proposal for the choice of α would be $\alpha = 0.2$: no difference is then downweighted to less than a fifth of what it could be, but a given difference between small proportions is considered to be more important

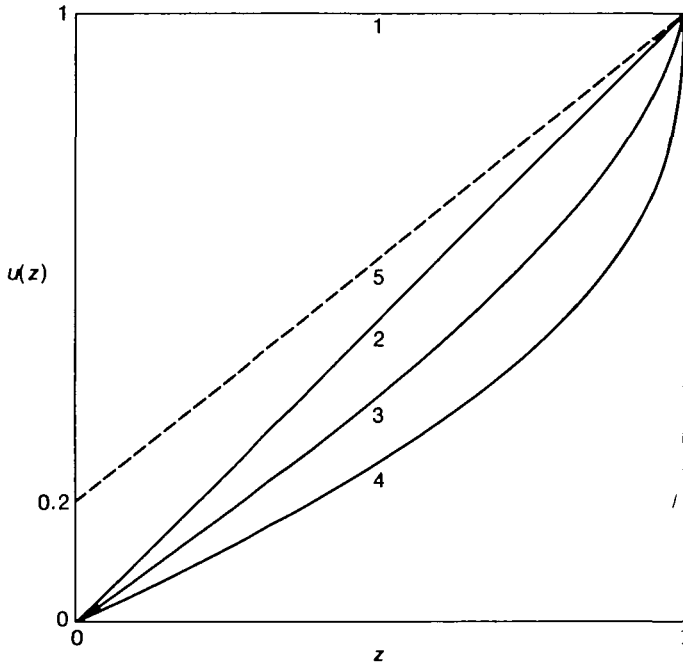


FIG. 1. The functions $u(z)$ for the C_4 indices of Renkonen (1), van Belle and Ahmad (2), Horn (modified) (3) and Matusita (4), together with that of $I_{0.2}$ (5).

than the same difference between large proportions. The corresponding function $u(z)$ is also shown in Fig. 1.

5. Estimating overlap

It is unrealistic to suppose that the information used in Section 2 is actually available. Instead, estimates of λ and μ are available at selected sites, and these are to be combined into an estimate of the true overlap. We distinguish two settings.

1. The region \mathcal{A} is known, but it is not clear how to split \mathcal{A} into homogeneous sub-regions.

In this case, the following procedure allows one to estimate overlap, when habitat types are weighted in proportion to their extent, defined by a chosen measure ϑ . Fix any measure ϑ over \mathcal{A} . Area (or volume) measure will do: but if, for instance, only water boatmen (Notonectidae) are to be considered, it may be more efficient to take ϑ to measure area of water. Then select n points within the area \mathcal{A} independently at random, from the probability distribution over \mathcal{A} which gives probability $\vartheta(\mathcal{B})/\vartheta(\mathcal{A})$ to any subset \mathcal{B} of \mathcal{A} (the larger n , the better the estimate obtained). For area measure, the distribution represents sticking a random pin n times into \mathcal{A} ; for area of water, do the same, throwing away all points falling on land, until n water points are obtained. It is implicit in this

procedure that the measure θ must be chosen in such a way as to be calculable for all known subsets \mathfrak{B} of \mathcal{A} , though, since the sub-regions \mathcal{A}_j are not known, the values of $\theta_j = \theta(\mathcal{A}_j)$ are also not known. Finally, at each point i , estimate the densities $\hat{\lambda}_i$ and $\hat{\mu}_i$ of the two species. This might well be done by taking a square of θ -measure t_i around the point i , counting the numbers x_i and y_i of individuals of the two species within the square, and setting

$$\hat{\lambda}_i = x_i/t_i \quad \text{and} \quad \hat{\mu}_i = y_i/t_i. \tag{5.1}$$

Then

$$\hat{l} = (1/n) \sum_i \hat{\lambda}_i \quad \text{and} \quad \hat{m} = (1/n) \sum_i \hat{\mu}_i \tag{5.2}$$

are consistent estimators of l and m respectively, and

$$(1/n) \sum_i \frac{1}{2}(\hat{\lambda}_i \hat{l} + \hat{\mu}_i \hat{m}) \psi(|\hat{\lambda}_i \hat{l} - \hat{\mu}_i \hat{m}| / (\hat{\lambda}_i \hat{l} + \hat{\mu}_i \hat{m})) \tag{5.3}$$

is a consistent measure of the C_3 -overlap

$$\sum_j \frac{1}{2}(f_j + g_j) \psi(|f_j - g_j| / (f_j + g_j)). \tag{5.4}$$

If $t_i = t$ for all $1 \leq i \leq n$, so that equal effort is expended at each sampling point, (5.3) reduces to

$$\sum_i \frac{1}{2}(p_i + q_i) \psi(|p_i - q_i| / (p_i + q_i)), \tag{5.5}$$

where $p_i = x_i / \sum_k x_k$ and $q_i = y_i / \sum_k y_k$. For the index I_α , the formulae corresponding to (5.3) and (5.5) are

$$\frac{1}{n} \sum_i \left(\alpha \min \left\{ \frac{\hat{\lambda}_i}{\hat{l}}, \frac{\hat{\mu}_i}{\hat{m}} \right\} + 2(1 - \alpha) \frac{\hat{\lambda}_i \hat{\mu}_i \hat{l} \hat{m}}{(\hat{\lambda}_i \hat{l} + \hat{\mu}_i \hat{m})} \right), \tag{5.6}$$

where $\hat{\lambda}_i, \hat{\mu}_i, \hat{l}$, and \hat{m} are defined in (5.1) and (5.2), and, in the case of equal effort at each sampling site,

$$\frac{1}{n} \sum_i \{ \alpha \min(p_i, q_i) + 2(1 - \alpha)[p_i q_i / (p_i + q_i)] \}. \tag{5.7}$$

It may often be reasonable to assume that the variance of the estimate (5.3) is proportional to $\sum_i (1/t_i)$. Were this exactly so, the choice $t_i = t, 1 \leq i \leq n$, would minimize the variance of estimate (5.3) subject to the constraint (on total effort) $\sum_i t_i = nt$.

2. The region \mathcal{A} consists of distinguishable homogeneous sub-regions $\mathcal{A}_1, \dots, \mathcal{A}_J$, and $\theta_j = \theta(\mathcal{A}_j)$ is known for each $1 \leq j \leq J$.

In this case, estimate the densities λ_j and μ_j in the region \mathcal{A}_j by counting the numbers x_j and y_j of the two species arising in a sampling area of θ -measure t_j , and take

$$\hat{\lambda}_j = x_j/t_j, \quad \hat{\mu}_j = y_j/t_j. \tag{5.8}$$

Then

$$\hat{l} = \sum_j \theta_j \hat{\lambda}_j / \sum_j \theta_j \quad \text{and} \quad \hat{m} = \sum_j \theta_j \hat{\mu}_j / \sum_j \theta_j \tag{5.9}$$

are consistent estimates of l and \bar{m} , and

$$\sum_j \frac{1}{2} \theta_j \left(\frac{\hat{\lambda}_j}{\hat{l}} + \frac{\hat{\mu}_j}{\hat{m}} \right) \psi \left(\left| \frac{\hat{\lambda}_j}{\hat{l}} - \frac{\hat{\mu}_j}{\hat{m}} \right| / \left(\frac{\hat{\lambda}_j}{\hat{l}} + \frac{\hat{\mu}_j}{\hat{m}} \right) \right) / \sum_j \theta_j \quad (5.10)$$

is a consistent estimate of the C_3 -measure (5.4). If $t_j = T\theta_j/\sum_k \theta_k$ ($1 \leq j \leq J$) (effort in area \mathcal{A}_j proportional to θ_j), then estimate (5.10) reduces to

$$\sum_j \frac{1}{2} (p_j + q_j) \psi(|p_j - q_j|/(p_j + q_j)), \quad (5.11)$$

where $p_j = x_j/\sum_k x_k$ and $q_j = y_j/\sum_k y_k$. For the index I_α of (4.3); the formulae corresponding to (5.10) and (5.11) are

$$\sum_j \theta_j \left(\alpha \min \left\{ \frac{\hat{\lambda}_j}{\hat{l}}, \frac{\hat{\mu}_j}{\hat{m}} \right\} + 2(1 - \alpha) \frac{\hat{\lambda}_j \hat{\mu}_j \hat{l} \hat{m}}{\hat{\lambda}_j \hat{l} + \hat{\mu}_j \hat{m}} \right) / \sum_j \theta_j \quad (5.12)$$

where $\hat{\lambda}_j$, $\hat{\mu}_j$, \hat{l} , and \hat{m} are given by (5.8) and (5.9), and, in the case of sampling effort in region \mathcal{A}_j being proportional to θ_j ,

$$\sum_j [\alpha \min(p_j, q_j) + 2(1 - \alpha)p_j q_j / (p_j + q_j)]. \quad (5.13)$$

If the variance of estimate (5.10) were proportional to $\sum_j (\theta_j^2/t_j)/(\sum_j \theta_j)^2$, as may often be approximately the case, the choice $t_j = T\theta_j/\sum_k \theta_k$ ($1 \leq j \leq J$) (proportional effort) minimizes the variance of (5.10) subject to the constraint (on total effort) $\sum_j t_j = T$.

It may sometimes be the case that the regions $\mathcal{A}_1, \dots, \mathcal{A}_J$ are distinguishable, but that the θ 's are fixed artificially, either because $\theta(\mathcal{A}_j)$ is unknown, or so as to give pre-determined weights to the various habitat types for reasons of comparability with other studies. In such circumstances, the second of the above estimation procedures is appropriate: however, the quantity being estimated now depends on the chosen values θ_j , which no longer necessarily agree with $\theta(\mathcal{A}_j)$, and overlap values can only properly be compared between investigations when the definitions of the sub-regions \mathcal{A}_j and their weights θ_j are in reasonable agreement.

The estimates (5.3) and (5.10), although consistent, need not be the best possible in any particular sampling scheme, and are frequently subject to bias (Mueller & Altenberg, 1985; Smith & Zaret, 1982). Improved estimators of the index I_α which show less bias than (5.13) are discussed in Schatzmann & Barbour (1986).

Acknowledgements

We would like to express our warm thanks to G. Bächli, M. Helg, and U. Pfändler for their helpful criticisms and suggestions.

REFERENCES

- ABRAMS, P. A. 1980 Some comments on measuring niche overlap. *Ecology* **61**, 44–49.
- ABRAMS, P. A. 1982 Reply to a comment by Hurlbert. *Ecology* **63**, 253–254.
- GERRARD, R., & BARBOUR, A. D. 1986 Measures of niche overlap, II. *IMAJ. Math. Appl. Med. Biol.*
- GRASSLE, J. F., & SMITH, W. 1976 A similarity measure sensitive to the contribution of rare species and its use in investigation of variation in marine benthic communities. *Oecologia* (Berl.) **25**, 13–22.
- GREENACRE, M. J. 1984 *Theory and Applications of Correspondence Analysis*. Academic Press, London.
- HORN, H. 1966 Measurements of overlap in comparative ecological studies. *Am. Nat.* **100**, 419–424.
- HURLBERT, S. H. 1978 The measurement of niche overlap and some relatives. *Ecology* **59**, 67–77.
- LANCE, G. N., & WILLIAMS, W. T. 1967 Mixed-data classificatory programs. I. Agglomerative systems. *Aust. Comput. J.* **1**, 15–20.
- LLOYD, M. 1967 Mean crowding. *J. Anim. Ecol.* **36**, 1–30.
- MATUSITA, K. 1955. Decision rules based on the distance, for problems of fit, two samples, and estimation. *Ann. Math. Statist.* **26**, 631–640.
- MORISITA, M. 1959. Measuring of interspecific association and similarity between communities. *Mem. Fac. Sci. Kyushu Univ.*, Ser. E (Biol.), **3**, 65–80.
- MUELLER, L. D., & ALTENBERG, L. 1985 Statistical inference on measures of niche overlap. *Ecology*, **66**, 1204–1210.
- PETRAITIS, P. S. 1979 Likelihood measures of niche breadth and overlap. *Ecology*, **60**, 703–710.
- PIANKA, E. R. 1973 The structure of lizard communities. *Ann. Rev. Ecol. Syst.* **4**, 53–74.
- RENKONEN, O. 1938 Statistisch ökologische Untersuchungen über die terrestrische Käferwelt der finnischen Bruchmoore. *Ann. Zool. Soc. Zool. Bot. Fenn. Vanamo* **6**, 1–231.
- RICKLEFS, R. E., & LAU, M. 1980 Bias and dispersion of overlap indices: Results of some Monte Carlo simulations. *Ecology* **61**, 1019–1024.
- SCHATZMANN, E., & BARBOUR, A. D. 1986 Estimating measures of overlap (to appear).
- SMITH, E. P. 1982 Niche breadth, resource availability, and inference. *Ecology* **63**, 1675–1681.
- SMITH, E. P., & ZARET, T. M. 1982 Bias in estimating niche overlap. *Ecology*, **63**, 1248–1253.
- VAN BELLE, G., & AHMAD, I. 1974 Measuring affinity of distributions. pp. 651–668 in F. Proschan & R. J. Serfling (Eds): *Reliability and Biometry*. SIAM Publications, Philadelphia, Pennsylvania, USA.
- WOLDA, H. 1981 Similarity indices, sample size and diversity. *Oecologia* (Berl.) **50**, 296–302.

