ON THE MEASUREMENT OF NICHE BREADTH AND OVERLAP[1]

ROBERT K. COLWELL

Department of Zoology, University of California, Berkeley

AND

DOUGLAS J. FUTUYMA

Department of Ecology and Evolution, State University of New York, Stony Brook

Abstract. Measures of niche breadth and overlap that depend on the distribution of individuals among resource states (ecological categories) should be independent of the relative abundance of the species and of the number of resource states considered. Such measures should also take into account the degree of distinctness of the resource states from the point of view of the organisms concerned. An ecoassay of the distinctness of resource states may well be easier and more meaningful than measurements of physical and chemical factors. We propose that the species composition of communities utilizing different resource states may be used to develop weighting factors with which each state may be weighted in proportion to its degree of distinctness. The weighting factors are used in the development of indices of niche breadth and overlap that correct for variation in the range and distinctness of resource states and that suffer less from human subjectivity than do the measures used to date. The use of such indices and the relationship of niche overlap to competition are discussed.

The analysis of niche relationships in natural communities by painstaking observation of natural history (e.g., MacArthur 1958, Cody 1968) is both valuable and necessary, but methods capable of revealing greater generality at the expense of detail are also desirable. The methods of niche analysis we propose are designed to be used with numerical data arranged in ecological categories. The data will usually be animal or plant abundances, in number-of-individuals or equivalent units. The categories need not be orderable, and they may differ one from another in known ways (food types, habitat types, substrates, lures, climatic regimes, etc.) or in unspecified ways (randomly or systematically spaced quadrats, times of day or year, traps, etc.). We will refer to such categories as "resource states."

In this paper we shall deal only with niche breadth and overlap, although there are other niche metrics of interest (dimensionality, topological deformation, fitness density and center of mass, dispersion). In terms of the spatial model of the niche, as formalized by Hutchinson (1958) and expanded by Slobodkin (1962), Levins (1968), and MacArthur (1968), niche breadth is the "distance through" a niche along some particular line in niche space. Other terms have been used for niche breadth, including "niche width" (Van Valen 1965, McNaughton and Wolf 1970), "niche size" (Klopfer and MacArthur 1960, Willson 1969), and "versatility" (Maguire 1967). In all these cases, as in this paper, the property referred to is essentially the inverse of ecological specialization, a term which has been used in a quantitative sense by Kohn (1968). With respect to food preferences, for example, a koala is more specialized (has a smaller

trophic niche breadth) than a Virginia opossum. Less extreme cases are more debatable, of course, and generalizations about relative niche breadth among taxonomic groups or in different communities are not even worth discussing with only anecdotal evidence as data. Some method of quantifying niche breadth is required for any worthwhile study of the property.

Niche overlap is simply the joint use of a resource, or resources, by two or more species. In other words, it is the region of niche space (in the sense of Hutchinson 1958) shared by two or more contiguous niches. The difficult question of the relationship of niche overlap to interspecific competition will be treated in the final section. As with niche breadth, any legitimate comparison of niche overlap among the species of a community, and especially among communities, requires careful quantification.

THE RESOURCE MATRIX AND SOME SIMPLE MEASURES

The measures of niche breadth and overlap we will discuss are all based on the distribution of individual organisms, by species, within a set of resource states. The table formed by using species as rows, and resource states as columns will be called the "resource matrix." A heterogeneous habitat, for example, might be subdivided into sunny-dry, sunny-wet, shady-dry, and shady-wet resource states, or alternatively, into unnamed random quadrats, which would then be considered the resource states. The resource matrix for vascular plants in this habitat would consist of a table with the species of vascular plants as rows, and the named habitat subdivisions or the unnamed random quadrats as columns. The typical cell would contain the number of individuals (N_{ij})

of species i found to be associated with resource state j. The total number of individuals of species i (the row total) will be called Y_i, and the total number of individuals of all species combined for resource state j (the column total) will be referred to as X_j. The total number of individuals in the matrix (the grand total) will be called Z. The number of species (rows) is s, and the number of resource states (columns) is r. We summarize this notation in Table 1.

TABLE 1. Notation for the resource matrix

	Resource states					
S	N_{11}	...	N_{1j}	...	N_{1r}	Y_1
p	.		.		.	.
e	.		.		.	.
c	.		.		.	.
i	N_{i1}	...	N_{ij}	...	N_{ir}	Y_i
e	.		.		.	.
s	.		.		.	.
	.		.		.	.
	N_{s1}	...	N_{sj}	...	N_{sr}	Y_s
	X_1	...	X_j	...	X_r	Z

The niche breadth of a species in the resource matrix can be estimated by measuring the uniformity of the distribution of individuals of that species among the resource states of the matrix. Levins (1968) suggests two simple measures of uniformity. In our notation, for the niche breadth of the ith species, these are

$$B_i = \frac{1}{\sum_j p_{ij}^2} = \frac{(\sum_j N_{ij})^2}{\sum_j N_{ij}^2} = \frac{Y_i^2}{\sum_j N_{ij}^2} \quad (1)$$

and

$$B'_i = -\sum_j p_{ij} \log p_{ij} \quad (2)$$

where $p_{ij} = N_{ij}/Y_i$, the proportion of the individuals of species i which is associated with resource state j. The measure B_i, equation (1), is closely related to the coefficient of variation, and is simply the inverse of Simpson's (1949) "measure of concentration." The measure B'_i, equation (2), is the Shannon-Wiener formula for information or uncertainty (Shannon and Weaver 1949). Both B_i and B'_i are maximized when an equal number of individuals of species i are associated with each resource state. This would imply that species i does not discriminate among the resource states (or discriminates, but requires or prefers them equally), and therefore has the broadest possible niche, with respect to those resource states. The measures are both minimized when all individuals of species i are associated with one resource state (minimum niche breadth, maximum specialization). The problem of standardizing the range of B_i and B'_i will be taken up in a later section.

Using the data of a resource matrix, niche overlap

between two species can be estimated by comparing the distribution of the individuals of the two species among the resource states of the matrix. If the distributions are identical, the two niches overlap completely, with respect to the resource states in the matrix. If the two species share no resource states, their niches do not overlap at all, with respect to these states. Perhaps the simplest measure of proportional overlap is

$$C_{ih} = 1 - \tfrac{1}{2} \sum_j | p_{ij} - p_{hj} | \quad (3)$$

where p_{ij} is the same as before, and $p_{hj} = N_{hj}/Y_h$, where h is a second species in the matrix. The measure C takes its minimum value of 0 when species i and species h share no resource states, and its maximum value of 1 when the proportional distributions of the two species among the resource states are the same. Schoener (1970) uses this measure of overlap in the form of a percentage to measure ecological similarity.

WEAKNESSES OF THE SIMPLE MEASURES

There are two situations in which one might wish to compare niche metrics: among species within a community, and between species of different communities. In both cases, the paramount difficulty is the standardization of the procedure so that measurements are comparable for different species. The problem is related to the familiar sophistry that the koala is really a generalist on *Eucalyptus* leaves, while the Virginia opossum is specialized for eating garbage. The difficulty, of course, arises in our definition of categories, since "specialized" means "covering few categories" and "generalized" means "covering many categories."

The question of defining appropriate categories can be divided into three parts: (i) the problem of range, (ii) the problem of spacing, and (iii) the problem of nonlinearity. To illustrate these, suppose someone wished to compare niche breadth and overlap, with respect to soil moisture, among species of soil mites at several study sites. At each site, he samples from each of 10 quadrats spaced along a transect from stream bank to dry hillside. Now, for simplicity, suppose that in fact the same five species are present at each study site, and that their true niches can be represented as in Fig. 1 (top set of axes), with respect to soil moisture measured as percentage saturation. Let the 10 quadrats at each of four different sites be arranged with respect to soil moisture as are the dots along the lines in the center of Fig. 1. Finally, assume that the abundance of each species in a quadrat is proportional to its fitness in that quadrat, so that, for example, the abundance of Species 2 at Site I would be maximal in the fourth quadrat (counting from the dry end), much less in the third
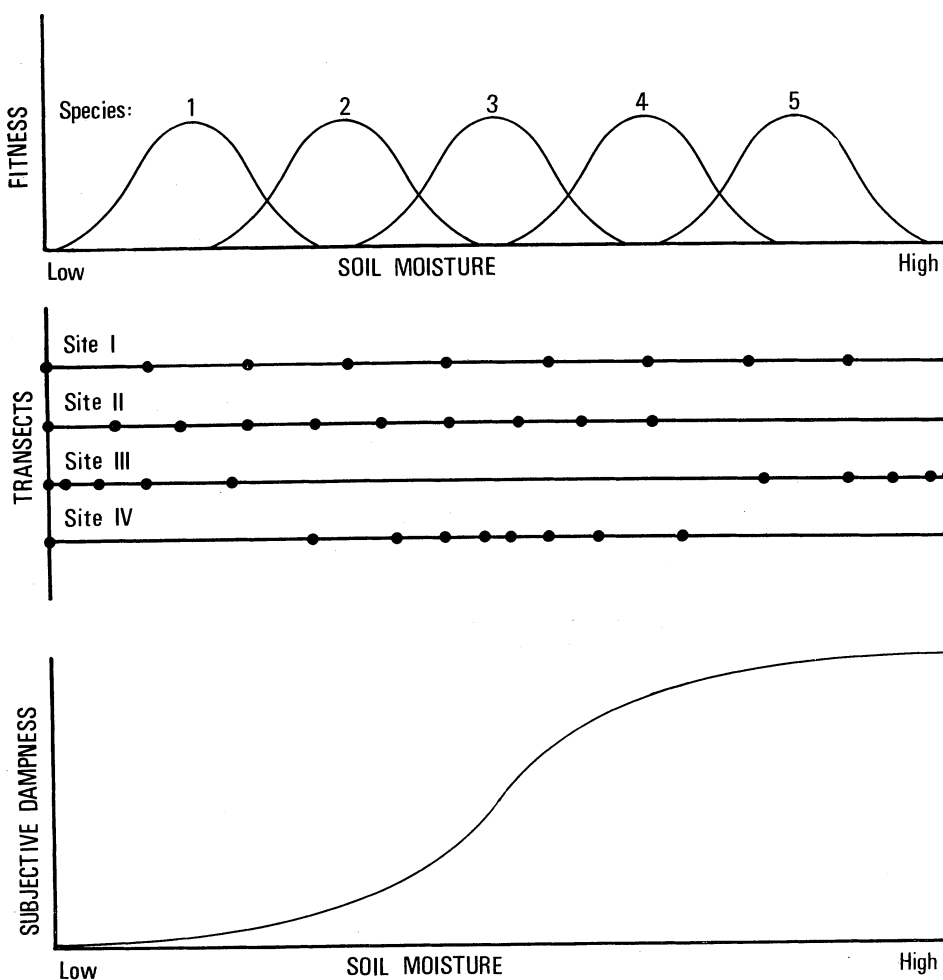
FIG. 1. A hypothetical example demonstrating the problems of range, spacing, and non-linearity in the measurement of niche breadth and overlap. The horizontal axis represents a gradient of soil moisture, from dry hillside (left) to moist stream bank (right). We consider 5 species of soil mites, each maximally fit at a different soil moisture (top graph). For simplicity we assume that, wherever they are found, the abundance of each mite species depends only on its fitness with respect to soil moisture. The same moisture gradient exists at each of four different study sites (Sites I, II, III, IV), but the sampling quadrats are placed in different patterns relative to soil moisture—shown by the large dots in the middle part of the figure. Niche breadth, being an intrinsic property of the species of mite, is the same wherever that species occurs. However, simple estimates of niche breadth will differ at different study sites, owing to differences in the patterns of spacing and total range of soil moisture covered by the quadrats. Estimates of niche overlap will be biased also. See text for a full explanation.

and fifth quadrats, and zero in all the other quadrats at that site.

The problem of range is demonstrated by a comparison of Site I with Site II. At Site II, Species 1 through 3 will be found in a larger number of quadrats than they are at Site I, so that the distribution of each of these species over the 10 quadrats (resource states) appears more uniform at Site II than at Site I. This uniformity would be reflected in higher estimates of niche breadth by equation (1) or (2) for the first three species at Site II than at Site I, even though they have the same tolerance for variation in

soil moisture at both sites. Differences in range, in themselves, will not affect proportional measures of niche overlap, such as equation (3), as long as the entire niche of each species is sampled (as is the case for Species 1 and 2 at Sites I and II).

The problem of range in the measurement of niche breadth is recognized by Ricklefs (1966), Cody (1968), and Maguire (1967), each of whom constructs an absolute scale by using as upper and lower bounds the extreme values found for all species combined, and then adjusts niche breadth estimates to that scale. McNaughton and Wolf (1970), on the

other hand, make no attempt to standardize ranges in calculating niche breadths from data on vegetational succession and on the distribution of organisms along "intuitive gradients." Consequently, the negative correlation they find between the number of species on a transect and the mean niche breadth of those species may be entirely spurious. Consider again Fig. 1, Sites I and II. Site I includes all five species, while the quadrats at Site II reveal fewer species, and yield a higher average niche breadth, as shown earlier, thus giving a negative correlation between niche breadth and diversity—for the same set of species.

The problem of spacing can be seen by comparing the distribution of quadrats at Sites I, III, and IV in Fig. 1. All three sets of quadrats cover the same range of soil moisture but differ in spacing along the moisture gradient. If niche breadths were calculated for the five species at each of these sites by equation (1) or (2), the data for Site III would yield higher estimates of niche breadth for Species 1 and 5 than for Species 2, 3, and 4, while just the reverse would result from the data for Site IV, with Site I giving approximately equal estimates for all five species. Thus spacing alone can easily bias estimates of niche breadth, and the problem will often arise even when samples are taken at even intervals of elevation, geological time, surface distance, and so forth, since the resource may not vary linearly with physical parameters. Unless resource states have ecologically equivalent degrees of distinctness among them, comparisons between communities, and particularly within communities, are perilous.

The problem of spacing also affects estimates of niche overlap. If equation (3) were calculated for Species 1 and 2 in the figure, a considerably higher estimate of niche overlap would be given for Site II than for Site III. This is a result of the concentration of a disproportionate number of quadrats in the exclusive portion of the niche of Species 1 at Site III. There is probably a general tendency toward the underestimation of niche overlap from abundance data due to the abruptness, and sometimes even the inaccessibility, of many ecotonal areas.

The problem of nonlinearity arises when an ecological variable is measured on a physical or chemical scale. Within any given range of total variation, resource states should be equally distinct, or equally spaced, not in relation to some physical or chemical variable measured in ordinary units, but in relation to an ecological variable, ideally measured in units of "subjective" effect on the organisms in question. Returning to the example in Fig. 1, the quadrats of Site I, which are equally spaced with respect to soil moisture measured as percentage saturation, will correctly estimate the niche breadths and overlaps of the five species of soil mites only if the effect of soil

moisture on the mites is linearly related to percentage saturation within the range measured. In fact, for most ecological variables, this kind of linearity seems rather unlikely. Nevertheless, the methods of Cody (1968), Maguire (1967), McNaughton and Wolf (1970), and Ricklefs (1966) for measuring niche breadths all assume a linear relationship between physical parameters and ecological variables, although the assumption is not explicitly stated.

By analogy to such measures as "subjective loudness" in human psychophysics (see Stevens 1959), we might refer in the example to "subjective dampness." If subjective dampness is related to soil moisture (measured as percentage saturation) as on the bottom set of coordinates of Fig. 1, then Species 2, 3, and 4 actually have broader niches than Species 1 or 5, since the middle species are able to survive over a much broader range of subjective dampness. Thus Site IV would yield the best estimates of niche breadth and overlap, since its quadrats are placed at equal increments of subjective dampness. The problem of nonlinearity is a subtle one, partially because of the difficulty of avoiding circularity in measuring its effect. But the data of human psychometrics and the response curves of physiological ecology are compelling evidence of the need to take the problem seriously in the estimation of niche metrics.

RESOURCE STATE WEIGHTING FACTORS

The amount of error caused by nonlinearity and ecological inequality of spacing among resource states would be considerably reduced if we could discover a way of weighting each resource state by its degree of distinctness from the other resource states in the resource matrix. The problem of range could be alleviated by correcting for the total or "collective" heterogeneity of the set of resource states in each estimation of a niche metric.

Although physical and chemical parameters may certainly be used to provide information on the distinctness and heterogeneity of resource states, we contend that biological parameters provide the best estimate of the degree of ecologically important difference among resource states. In particular, the relative abundance among resource states of a sufficiently wide variety of other species in the community should provide adequate information on the ecological distinctness of resource states, in most cases. This might be called an "eco-assay" of the resource states.

Thus in the example of the moisture gradient, we would argue that a census of the plant and animal species in each quadrat contains more relevant information about the degree of distinction of that quadrat from other quadrats, from a mite's point of view, than does a series of physical and chemical measure-

ments obtained with the same amount of effort. Not only are other species important to a mite directly— as competitors, predators, prey, and perhaps substrate and habitat—but the entire fauna and flora must reflect, indirectly, all relevant physical and chemical parameters.

We shall now develop a method for weighting resource states by their distinctness, and for estimating their collective heterogeneity. The method applies equally well, in theory, to chemical and physical parameters, although it is designed for biological weighting data.

Whether physical, chemical, or biological parameters are used, each resource state is scored for each parameter, as well as for the abundance of the species whose weighted niche breadth is to be calculated. The environmental parameters are cast into a resource matrix, as defined above and in Table 1, but one in which the rows may represent either species or physical-chemical parameters. In the latter case, the N_{ij}'s are scores or measurements, which need not be integers. (The possibility of weighting rows will be discussed later.) For simplicity, it will be assumed in what follows that the rows are all species, and that the abundance data for the species whose niche breadth is to be calculated is also in the matrix.

Now we define a new matrix in which each element π_{ij} is the proportion of the grand total represented by each N_{ij}, so that

$$\pi_{ij} = \frac{N_{ij}}{Z} . \qquad (4a)$$

If we call the row (species) classification the "Y-classification," and the column (resource state) classification the "X-classification," then the π matrix is the joint probability distribution of the X and Y classifications, and π_{ij} is the probability (actually an estimate, of course) that a randomly chosen individual from the pooled species will be an individual of species i associated with resource state j. We next define the conditional probabilities

$$p_{ij} = \frac{N_{ij}}{Y_i} , \qquad (4b)$$

$$P_j = \frac{X_j}{Z} , \qquad (4c)$$

$$q_{ij} = \frac{N_{ij}}{X_j} , \qquad (4d)$$

and

$$Q_i = \frac{Y_i}{Z} . \qquad (4e)$$

Thus p_{ij} is the probability that an individual is associated with resource state j, given that it is of species i, and q_{ij} is the probability that an individual is of

species i, given that it is associated with resource state j. Finally, Q_i is the probability that an individual is of species i, disregarding resource states, and P_j is the probability that an individual is associated with resource state j, disregarding species identity. Therefore,

$$\sum_i \sum_j \pi_{ij} = \sum_j p_{ij} = \sum_i q_{ij} = \sum_j P_j = \sum_i Q_i = 1 . \qquad (4f)$$

To obtain weighting factors for the resource states, we begin by defining the following information functions. The uncertainty with respect to resource state of one of the Z individuals in the resource matrix is

$$H(X) = - \sum_{j=1}^{r} P_j \log P_j , \qquad (5)$$

and the uncertainty of an individual with respect to species is

$$H(Y) = - \sum_{i=1}^{s} Q_i \log Q_i . \qquad (6)$$

Given an individual of species i, its uncertainty with respect to resource state is

$$H_i(X) = - \sum_j p_{ij} \log p_{ij} . \qquad (7)$$

Given an individual associated with resource state j, its uncertainty with respect to species is

$$H_j(Y) = - \sum_i q_{ij} \log q_{ij} . \qquad (8)$$

The uncertainty of an individual with respect to both species and resource state is

$$H(XY) = - \sum_i \sum_j \pi_{ij} \log \pi_{ij} . \qquad (9)$$

We wish to determine the total heterogeneity of the matrix with respect to resource states, and the contribution of each resource state to this total, which will be the basis for a weighting factor for that state. If there is complete homogeneity of resource states (i.e., if $q_{ij} = Q_i$ for all i and j), then the X and Y classifications are independent, and $\pi_{ij} = Q_i P_j$. In this limiting case,

$$H(XY) = H(Y) + H(X)$$

(see Appendix A for proof). If, however, the X and Y classifications are not independent, the relation

$$H(XY) < H(Y) + H(X) \qquad (10)$$

is always true, since the total uncertainty $H(XY)$ is reduced by any dependence between the two classifications [Khinchin (1957) gives a formal proof].

If we know the relative abundance of species (the Y_i's) then the total uncertainty $H(XY)$ is reduced by the amount $H(Y)$, equations (9) and (6). The remainder is defined as

$$H_Y(X) = H(XY) - H(Y) , \qquad (11)$$

which can be shown (see Appendix B) to be

$$H_Y(X) = -\sum_i \sum_j \pi_{ij} \log p_{ij}, \quad (12)$$

or alternatively, the weighted average

$$H_Y(X) = -\sum_i Q_i H_i(X), \quad (13)$$

where $H_i(X)$ is defined by equation (7). [See Appendix B and Pielou (1969).]

It is clear that when the distribution of individuals over resource states differs among species, the X and Y classifications are not independent, and, from equations (10) and (11), $H_Y(X) < H(X)$. Thus a measure of the departure from independence is given by

$$m(X) = H(X) - H_Y(X), \quad (14)$$

which is sometimes called "mutual information" (Abramson 1963), since

$$m(X) = m(Y) = H(Y) - H_X(Y).$$

The limits of $m(X)$ are zero, for complete homogeneity of resource states, and $H(X)$, for complete heterogeneity of resource states [see Appendix A and Hays (1964)]. Thus a standardized measure of resource state heterogeneity is

$$M(X) = \frac{m(X)}{H(X)} = \frac{H(X) - H_Y(X)}{H(X)} = 1 - \frac{H_Y(X)}{H(X)} \cdot \quad (15)$$

$M(X)$ may be called "relative mutual information"; it has the range 0 to 1, and does not depend upon either the number of resource states or the magnitude and evenness of column (resource state) totals. Partitioning $M(X)$ by resource states, we obtain

$$M_j(X) = (\sum_i \pi_{ij} \log p_{ij} - P_j \log P_j)/H(X), \quad (16)$$

or, for computation,

$$M_j(X) = \frac{X_j(\log X_j - \log Z) - \sum_i N_{ij} \log (N_{ij}/Y_i)}{\sum_j X_j \log X_j - Z \log Z}.$$

The quantity $M_j(X)$ is the distinctness of the jth resource state in the resource matrix. The collective heterogeneity of the resource states is therefore $\sum_j M_j(X) = M(X)$, with range $(0, 1)$. We now define the "absolute" weighting factor for the jth resource state as

$$\delta_j = M_j(X), \quad (17)$$

and the "relative" weighting factor for the jth resource state as

$$d_j = \frac{M_j(X)}{\sum_j M_j(X)} = \frac{\delta_j}{M(X)}, \quad (18)$$

so that $\sum_j d_j = 1$.

WEIGHTED EXPANSION OF THE RESOURCE MATRIX AND WEIGHTED NICHE METRICS

The simple niche measures, equations (1–3), will be good estimators of the true niche breadth and overlap of species in a resource matrix only if the resource states are all equally distinct, one from another. Conceptually, this might be accomplished by adding new columns (resource states) to the matrix that are identical to the more-distinct original resource states, which would then be less distinct in the expanded matrix. For example, in the pattern "abaa," the letter "b" is more unusual than "a," but if we expand the pattern to "abbaa," and then to "abbbaa," the letter "b" becomes less and less unusual. Theoretically, such an expansion of the resource matrix could be carried out until the distinctness of all resource states, as measured by equation (17) or (18), was the same.

An equivalent procedure, which is much more tractable, is to choose a large number k (say 10,000), and "expand" the original matrix of r resource states to a new matrix having k resource states, with each of the original resource states represented $d_j k$ times in the expanded matrix $(\sum_j d_j k = k\sum_j d_j = k)$. Suppose, for example, that the original matrix has three resource states, and the abundance N_{ij} of species i on these states and the calculated values of d_j are as follows:

j :	1	2	3
N_{ij}:	2	5	1
d_j :	.5	.1	.4

Now if we set $k = 10$, we obtain the expanded abundance vector:

j :	1	2	3	4	5	6	7	8	9	10
N_{ij}:	2	2	2	2	2	5	1	1	1	1

in which each of the original resource states is represented $d_j k$ times. In practice, there would of course be more resource states, and a much larger k. Even though it is not possible to represent $d_j k$ resource states when that product is not an integer, the use of the weighting factors in the weighted niche metrics is legitimate mathematically, as can be seen from what follows.

The application of the simple formulas for niche breadth and overlap to such an expanded matrix will yield measures corrected for the original variation in distinctness and for any ecological nonlinearity in the original resource states. To calculate the weighted niche measures directly from the original matrix and the vector of weighting factors, we first define Y^*_i as the total number of individuals of species i in the expanded matrix. Since the jth column of the original matrix is reproduced $d_j k$ times in the expanded matrix,

$$Y^*_i = \sum_j d_j k \, N_{ij} . \qquad (19a)$$

Now define

$$p^*_{ij} = \frac{N_{ij}}{Y^*_i} . \qquad (19b)$$

Then for niche breadth, by equation (1), we have

$$B_i = 1/\sum_j (d_j k \, p^{*2}_{ij}) .$$

This formulation has range $(1,k)$, so to standardize the range to $(0,1)$ we must calculate

$$\beta_i = \frac{B_{i(\text{observed})} - B_{(\min)}}{B_{i(\max)} - B_{(\min)}} ,$$

which is

$$\beta_i = \left[\frac{1}{\sum_j (d_j k \, p^{*2}_{ij})} - 1 \right] \left[\frac{1}{k-1} \right] . \qquad (20)$$

Niche breadth by equation (2) for the expanded matrix becomes

$$B'_i = -\sum_j d_j k \, (p^*_{ij} \log p^*_{ij}) .$$

This expression has the range $(0, \log k)$, so to standardize the range to $(0, 1)$, we must calculate

$$\beta'_i = -\frac{k}{\log k} \sum_j d_j \, (p^*_{ij} \log p^*_{ij}) . \qquad (21)$$

For niche overlap, the modification for C_{ih} [equation (3)], for the expanded resource matrix is

$$\gamma_{ih} = 1 - \tfrac{1}{2} \sum_j d_j k \, | \, p^*_{ij} - p^*_{hj} \, | . \qquad (22)$$

A second measure of niche overlap follows directly from the definition of "relative mutual information" given in the section on weighting factors, equation (15). We will first define a new measure of simple overlap, and then add the resource state weighting factors. Now let $t_j = p_{ij} + p_{hj}$, where $p_{ij} = N_{ij}/Y_i$, as previously defined, and $p_{hj} = N_{hj}/Y_h$ for the hth species in the original resource matrix. To simplify notation, define the function $I(x) = x \log x$. Then an index of the difference in proportional distribution of species i and h over the r resource states of the original matrix is the relative mutual information of the vectors $(p_{i1} \ldots p_{ij} \ldots p_{ir})$ and $(p_{h1} \ldots p_{hj} \ldots p_{hr})$, which is

$$M_{ih} = 1 - \frac{\sum_j [(p_{ij}/2) \log (p_{ij}/t_j) + (p_{hj}/2) \log (p_{hj}/t_j)]}{2(\tfrac{1}{2} \log \tfrac{1}{2})} .$$

The complement of M_{ih} (that is, $1 - M_{ih}$), an index of niche overlap, reduces to

$$C'_{ih} = -\frac{1}{2 \log 2} \sum_j \left[I(p_{ij}) + I(p_{hj}) - I(t_j) \right] , \qquad (23)$$

which ranges from 0, when species i and h do not co-occur on any resource state, to 1, when the proportional distribution of the two species over resource states is the same. Equation 23 is formally similar to the "faunal overlap" index of Horn (1966), which is a two-sample formula for relative mutual information. The weighted form of C' is simply equation (23) calculated for the expanded matrix, or

$$\gamma'_{ih} = -\frac{1}{2 \log 2} \sum_j d_j k \left[I(p^*_{ij}) + I(p^*_{hj}) - I(t^*_j) \right] , \qquad (24)$$

where p^*_{ij} is as defined in equation (19b), $p^*_{hj} = N_{hj}/Y^*_h$, and $t^*_j = p^*_{ij} + p^*_{hj}$.

So far we have only considered weighted niche metrics which utilize the "relative" weighting factors (d_j's) defined by equation (18). All four of these measures—β, β', γ, and γ' [equations (20), (21), (22), and (24)]—have the range $(0, 1)$, regardless of the total heterogeneity of the resource states in the original resource matrix, although they do give relative weight to each resource state proportional to its degree of distinctness. These measures are therefore appropriate only for comparisons among species in the same resource matrix, or among matrices of equal total heterogeneity.

To correct for differences in ecological range among matrices, we must take into account the total heterogeneity of resource states in each matrix by using the "absolute" weighting factors (δ_j's) defined by equation (17). The expressions for absolute niche breadth are the same as equations (20) and (21), but with δ_j substituted for d_j in the summations, as well as in the calculation of p^*_{ij} by equations (19a) and (19b). In the expressions for absolute niche overlap which correspond to equations (22) and (24), δ_j is substituted for d_j in the summations, but the values of p^*_{ij} and p^*_{hj} are calculated as in the original equations, using d_j. In addition, the summation must be subtracted from $\sum_j \delta_j$, instead of from 1, in the expression for absolute niche overlap corresponding to equation (22). The absolute measures will always be less than the corresponding relative measures, unless the resource matrix is maximally heterogeneous, in which case $\delta_j = d_j$ for all j.

The actual value chosen for the constant k is irrelevant for the niche overlap measures, but the measures of niche breadth depend upon the value of k used, so that the same value should be used consistently in any particular study. We have found 10,000 to be an adequate standard for entomological field data. In any case, k must be substantially larger than r to avoid positive logarithms in equation (21). The logarithms used in calculating the weighting factors and in equations (21) and (24) may be taken to any convenient base (as long as the same base is

TABLE 2. Hypothetical resource matrix and weighted niche metrics

A. Raw data, marginal totals, and resource state weighting factors

Species	Resource states												Y_i
1	1	1	1	1	1	1	1	1	1	1	1	1	12
2	1	1	1	1	0	0	0	0	1	1	1	1	8
3	0	0	0	0	1	1	1	1	1	1	1	1	8
4	0	0	1	1	1	1	0	0	1	1	1	1	8
5	0	0	0	0	2	2	2	2	2	2	2	2	16
6	7	0	0	0	0	0	0	0	0	0	0	0	7
7	1	6	0	0	0	0	0	0	0	0	0	0	7
8	0	1	6	0	0	0	0	0	0	0	0	0	7
9	0	0	1	6	0	0	0	0	0	0	0	0	7
10	0	0	0	7	0	0	0	0	0	0	0	0	7
X_j :	10	9	10	16	5	5	4	4	6	6	6	6	$Z=87$
													Sum:
d_j :	.175	.150	.140	.226	.040	.040	.042	.042	.036	.036	.036	.036	.999[a]
δ_j :	.073	.062	.058	.094	.017	.017	.017	.017	.015	.015	.015	.015	.415

B. Niche breadth

Species	Unweighted	Relative	Absolute
1	1.000	1.000	.683
2	.824	.931	.577
3	.824	.490	.126
4	.824	.781	.425
5	.824	.481	.135

C. Niche overlap

Species		Unweighted	Relative	Absolute
1	2	.809	.919	.414
1	3	.809	.564	.255
1	4	.809	.770	.347
1	5	.809	.553	.250
2	3	.500	.308	.139
2	4	.750	.745	.336
2	5	.500	.310	.140
3	4	.750	.538	.243
3	5	1.000	1.000	.451
4	5	.750	.538	.243

[a]Differs from 1.000 by rounding error.

used consistently) without affecting the resulting values. To avoid circularity, the species for which niche breadth is being calculated, or the two species for which niche overlap is being calculated, should be excluded from the computation of weighting factors. This means there is actually a different set of weighting factors for each calculation of a niche metric.

To illustrate the properties of weighted niche metrics, unweighted, relative, and absolute niche breadth [by equation (21)] and niche overlap [by equation (24)] were computed for the first five species of the hypothetical resource matrix in Table 2. Although the resource state weighting factors given in the table were calculated for the entire set of 10 species, the weighting factors used to compute the weighted niche metrics were derived anew for each value by skipping the row (for niche breadth) or rows (for niche overlap) for which the metric was to be calculated.

It may be desirable to alter the weighting of the rows in the resource matrix, particularly if physical or chemical factors are included, since each row has an equal influence on the resource state weighting factors only if $Q_i = 1/s$ (where s is the number of rows or species) for all i. Thus if biological abundance data are used, commoner species have more effect than less common species on the values of the weighting factors. If some other relative weighting is desired (such as logarithm of relative abundance), define a vector of row weights $(w_1 \ldots w_i \ldots w_s)$, where w_i is the desired weight of the ith row in the resource matrix. Then the typical element in the row-weighted matrix is

$$N'_{ij} = w_i \, p_{ij} = \frac{w_i \, N_{ij}}{Y_i}.$$

For equal weighting of all rows, $N'_{ij} = p_{ij}$.

Finally, we wish to emphasize strongly that the methods we have developed will do nothing to improve a poor research design or meaningless data.

It is essential that ecological data to be used for comparing niche metrics within, and especially among, communities be taken in a standardized and meaningful way. The corrective measures we propose are simply the "fine adjustment"; the coarse adjustment must be present in the experimental design. It is still not possible to compare the food niche of the Virginia opossum and the koala.

NICHE OVERLAP AND COMPETITION

Because of the widespread use of measures of overlap as estimates of competition for resources, we find it necessary to clarify the possible use of the measures we have proposed in studies of competition. To begin with, let us distinguish between the "actual" and "virtual" niches of a population. These terms are approximately equivalent to the "realized" and "fundamental" niches of Hutchinson (1958), but are operationally defined, and pertain specifically to the effects of competition on a local population level. Niche breadth and overlap, when measured under "natural" conditions, are "actual" metrics, while "virtual" niche breadth and overlap are the corresponding values measured in the absence of competition among species. Thus Levins (1968) refers to the virtual niche as "pre-competitive." Most commonly, competitive displacement or exclusion tends to reduce niche breadth and overlap among competing species, so that the actual niche is a proper subset of the virtual niche in environment space.

The realized niche of Hutchinson is defined for an entire species, and for species with a wide geographic range, or with clinical variation, it may be considerably larger than the actual niche of a local population. Likewise, Hutchinson's fundamental niche includes all regions of niche space in which a species has positive fitness, whether or not all such conditions would exist in nature, even if competitors were removed. Thus the fundamental niche is larger than the virtual niche for a local population, and usually even for an entire species.

We must emphasize that it is the conditions under which data are collected, rather than the method of calculation, that determine whether actual or virtual niches are measured. In general, natural history data (transects, quadrats, etc.) are not capable of yielding measurements of the virtual niche. Virtual niche measurements can be obtained either by actually removing competitors (e.g., Connell 1961, Culver 1970, Hespenheide 1971), by exploiting natural situations in which competitors are absent (e.g., Van Valen 1965, Culver 1970), or in some situations by creating an oversupply of a scarce resource in order to minimize competition (Levins 1968, Colwell 1969).

The distinction between actual and virtual niche measurements is particularly important when niche overlap is taken as a measure of competition. Paradoxically, simply demonstrating an overlap in resource use by two species in nature can be evidence either for or against the existence of competition between them. Competition may be operating, but exclusion or displacement may be incomplete, or even impossible (as with competition among plants for carbon dioxide in short supply). In this case, observed niche overlap is evidence of the existence of competition. However, overlap may be evidence of a lack of competition if the resource under consideration is in oversupply or is irrelevant to one or both species.

By similar arguments, it can be shown that lack of demonstrable overlap may also be evidence either for or against the existence of competition. The only way to demonstrate the existence of competition, and to measure its intensity, is by comparing actual to virtual niche overlap between suspected competitors: if actual and virtual overlap are both zero, or if they are equal and no change in the population of either species occurs in the absence of its putative competitor, then there is no evidence for competition at that time and place. If, on the other hand, virtual overlap is shown to exceed actual overlap, the existence of competition has been demonstrated.

Finally, we do not find it useful to speak of the niche of an individual, although that is conceptually possible. The dimensions of such an individual's niche might just as well be treated as characters of individual phenotype. A niche, then, like a gene pool, is a property of a set of individuals, usually a biological population. (In fact, it could be argued that the concept of niche should be replaced by some notion of "phene pool.") Consequently, the niche of a species is a statistical entity which changes whenever its constituents change.

ACKNOWLEDGMENTS

LITERATURE CITED

Abramson, N. 1963. Information theory and coding. McGraw-Hill, New York. 201 p.

Cody, M. L. 1968. On the methods of resource division in grassland bird communities. Amer. Natur. **102**: 107-147.

Colwell, R. K. 1969. Ecological specialization and species diversity of tropical and temperate arthropods. Doctoral dissertation, University of Michigan, Ann Arbor.

Connell, J. H. 1961. The influence of interspecific com-

petition and other factors on the distribution of the barnacle *Chthalamus stellatus.* Ecology 42: 710–723.

Culver, D. H. 1970. Analysis of simple cave communities: niche separation and species packing. Ecology 51: 949–958.

Hays, W. L. 1964. Statistics for psychologists. Holt, Rinehart, and Winston, New York. 719 p.

Hespenheide, H. A. 1971. Flycatcher habitat selection in the eastern deciduous forest. Auk 88: 61–74.

Horn, H. S. 1966. Measurement of "overlap" in comparative ecological studies. Amer. Natur. 100: 419–424.

Hutchinson, G. E. 1958. Concluding remarks. Cold Spring Harbor Symp. Quant. Biol. 22: 415–427.

Khinchin, A. I. 1957. Mathematical foundations of information theory. Dover, New York. 120 p.

Klopfer, P. H., and R. H. MacArthur. 1960. Niche size and faunal diversity. Amer. Natur. 94: 293–300.

Kohn, A. J. 1968. Microhabitats, abundance, and food of *Conus* in the Maldive and Chagos Islands. Ecology 49: 1046–1061.

Levins, R. 1968. Evolution in changing environments. Princeton Univ. Press, Princeton, N. J. 120 p.

MacArthur, R. H. 1958. Population ecology of some warblers of northeastern coniferous forests. Ecology 39: 599–619.

———. 1968. The theory of the niche, p. 159–176. *In* R. C. Lewontin [ed.] Population biology and evolution. Syracuse Univ. Press, Syracuse, New York.

MacArthur, R. H., and R. Levins. 1967. The limiting similarity, convergence, and divergence of coexisting species. Amer. Natur. 102: 377–385.

Maguire, B. 1967. A partial analysis of the niche. Amer. Natur. 101: 515–523.

McNaughton, S. J., and L. L. Wolf. 1970. Dominance and the niche in ecological systems. Science 167: 131–139.

Pielou, E. C. 1969. An introduction to mathematical ecology. Wiley Interscience, New York. 286 p.

Ricklefs, R. E. 1966. The temporal component of diversity among species of birds. Evolution 20: 235–242.

Schoener, T. W. 1970. Non-synchronous spatial overlap of lizards in patchy habitats. Ecology 51: 408–418.

Shannon, C. E., and W. Weaver. 1949. The mathematical theory of communication. Univ. of Illinois Press, Urbana. 117 p.

Simpson, E. H. 1949. Measurement of diversity. Nature 163: 688.

Slobodkin, L. B. 1962. Growth and regulation of animal populations. Holt, Rinehart, and Winston, New York. 184 p.

Stevens, S. S. 1959. Measurement, psychophysics, and utility, p. 18–63. *In* C. W. Churchman and P. Ratoosh [ed.], Measurement: definitions and theories. Wiley, New York.

Van Valen, L. 1965. Morphological variation and width of ecological niche. Amer. Natur. 99: 377–390.

Willson, M. F. 1969. Avian niche size and morphological variation. Amer. Natur. 103: 531–542.

APPENDIX A

The limits of $H(XY)$ and $m(X)$

From equation (9), $H(XY) = \Sigma\Sigma_{ij} \pi_{ij} \log \pi_{ij}$. When the resource states are completely homogeneous, $q_{ij} = Q_i$, and thus $p_{ij} = P_j$ necessarily, for all i and j. Therefore

$$P_j Q_i = P_j q_{ij} = \frac{X_j}{Z} \frac{N_{ij}}{X_j} = \pi.$$

Substituting, we have

$$
\begin{aligned}
H(XY) &= -\Sigma\Sigma_{ij} Q_i P_j \log Q_i P_j \\
&= -\Sigma\Sigma_{ij} Q_i P_j (\log Q_i + \log P_j) \\
&= -\Sigma_j P_j (\Sigma_i Q_i \log Q_i) - \Sigma_i Q_i (\Sigma_j P_j \log P_j).
\end{aligned}
$$

Thus, from equations (5) and (6), $H(XY) = H(Y) + H(X)$. Since $H_Y(X) = H(XY) - H(Y)$, equation (11), we have $H_Y(X) = H(X)$, so that $m(X) = 0$, equation (14).

When the resource states are completely heterogeneous, with no species in common among them, each row in the resource matrix may have only one non-zero N_{ij} (this will not necessarily be true for the columns), call it $N_{ig(i)}$. Then $N_{ig(i)} = Y_i$, and

$$
\begin{aligned}
\pi_{ij} &= 0, \text{ for } j \neq g(i) \\
\pi_{ij} &= N_{ij}/Z = Y_i/Z = Q_i, \text{ for } j = g(i).
\end{aligned}
$$

Therefore $H(XY) = H(Y)$, where $H(Y)$ is given by equation (6). Likewise, since

$$p_{ij} = N_{ij}/Y_i = 0, \text{ for } j \neq g(i),$$

and

$$p_{ij} = 1, \text{ for } j = g(i),$$

for all i, we have $H_Y(X) = -\Sigma\Sigma_{ij} \pi_{ij} \log p_{ij} = 0$, from equation (12), and thus $m(X) = H(X)$.

Note that $H(XY)$ has another "minimum" [i.e., $H(X)$] when there exists complete heterogeneity of species distribution (zero overlap for all pairs of species).

APPENDIX B

Derivation of expressions for $H_Y(X)$

From equation (11), we define

$$H_Y(X) = H(XY) - H(Y).$$

Substituting from equations (9) and (6), we have

$$
\begin{aligned}
H_Y(X) &= -\Sigma\Sigma_{ij} \pi_{ij} \log \pi_{ij} + \Sigma_i Q_i \log Q_i \\
&= -\Sigma\Sigma_{ij} \pi_{ij} \log \pi_{ij} + \Sigma_j p_{ij} \Sigma_i Q_i \log Q_i,
\end{aligned}
$$

since $\Sigma_j p_{ij} = 1$; thus

$$
\begin{aligned}
H_Y(X) &= -\Sigma\Sigma_{ij} \left[\frac{N_{ij}}{Z} \log \frac{N_{ij}}{Z} - \frac{N_{ij}}{Y_i} \frac{Y_i}{Z} \log \frac{Y_i}{Z} \right] \\
&= -\Sigma\Sigma_{ij} \left[\frac{N_{ij}}{Z} \left(\log \frac{N_{ij}}{Z} - \log \frac{Y_i}{Z} \right) \right] \\
&= -\Sigma\Sigma_{ij} \frac{N_{ij}}{Z} \log \frac{N_{ij}}{Y_i}.
\end{aligned}
$$

Thus we obtain equation (12):

$$H_Y(X) = \Sigma\Sigma_{ij} \pi_{ij} \log p_{ij}.$$

The alternative form, equation (13), is derived from the above form:

$$
\begin{aligned}
H_Y(X) &= -\Sigma\Sigma_{ij} \frac{Y_i}{Y_i} \frac{N_{ij}}{Z} \log \frac{N_{ij}}{Y_i} \\
&= -\Sigma_i Q_i \Sigma_j p_{ij} \log p_{ij} \\
H_Y(X) &= -\Sigma_i Q_i H_i(X)
\end{aligned}
$$

where $H_i(X)$ is defined by equation (7). Precisely symmetrical derivations can be given for the expressions

$$H_X(Y) = -\Sigma\Sigma_{ij} \pi_{ij} \log q_{ij} = -\Sigma_j P_j H_j(Y),$$

where

$$H(XY) = H(X) + H_X(Y).$$