Overlap Summary Indices and the Detection of Community Structure
Author(s): James D. Thomson and Kurt A. Rusterholz
Source: *Ecology*, Vol. 63, No. 2 (Apr., 1982), pp. 274-277
Published by: Wiley
Stable URL: http://www.jstor.org/stable/1938941
Accessed: 26-06-2016 03:50 UTC

Ecology, 63(2), 1982, pp. 274–277

OVERLAP SUMMARY INDICES AND THE DETECTION OF COMMUNITY STRUCTURE[1]

JAMES D. THOMSON
Ecology and Evolution Department, State University of New York, Stony Brook, New York 11794 USA

AND

KURT A. RUSTERHOLZ
Department of Zoology, University of Wisconsin, Madison, Wisconsin 53706 USA

Abstract. Overlaps in resource use have been used to summarize community structure. It is often desirable to compare the amount or intensity of overlaps (or other pairwise measures such as competition coefficients) in different communities, especially in the context of comparing actual communities to "neutral model" simulations. We show that the most commonly used index for such comparisons, the mean pairwise overlap, is often unsuitable, and suggest alternate approaches.

Key words: community structure; competition; guild structure; neutral model; niche; overlap index; pairwise overlap; resource partitioning.

Although the precise implications of ecological overlap remain a subject for debate, many ecologists nevertheless use various overlap measures to summarize the degrees of ecological similarity within assemblages of species. A technique that is receiving increasing attention in this regard is the comparison of an observed set of overlaps with a "neutral model," i.e., a different set of overlaps generated by some sort of random algorithm. In the search for "community structure" in an assemblage, which in this context usually means significant interactions among species with respect to resource use, the neutral model serves as a null hypothesis because it gives the pattern of overlap expected from complete species independence of resource use. By generating large numbers of random assemblages, one can make a Monte Carlo estimate of the probability that some overlap summary statistic could be observed in an unstructured assemblage. Designing an algorithm that combines true species independence with ecological realism may not be a trivial problem, but we wish to consider a different problem regarding neutral-model comparisons as they have generally been used in the literature. This concerns the statistic or index that one calculates to make the comparisons of the real assemblage with the artificial ones. The point is important because the index contains all the information about structure that goes into the comparison. The index that has been used most often, both for neutral-model comparisons and for general use in loosely summarizing the "intensity of overlap" in an assemblage, is the arithmetic mean of the pairwise overlaps (e.g., Sale 1974, Brown 1975, Pianka 1975, Thomson 1978, Parrish and Bazzaz 1979, Pianka et al. 1979, Rusterholz 1979, Joern and Lawlor 1980, and others). This is a poor index because it is insensitive to certain changes in assemblage constitution that seem very important to assemblage structure. In certain realistic cases, the mean pairwise overlap may actually increase following changes in assemblage composition that most ecologists would intuitively consider reductions in overlap.

Consider the assemblage of plants whose time of flowering curves are displayed in Fig. 1. These curves summarize genuine census data from Rocky Mountain subalpine meadows (Thomson 1978), but the collection of plant species was chosen for clarity of illustration and represents no biologically defined guild. It corresponds to a common pattern for zoophilous plant communities noted by Thomson (1975, 1978), Parrish and Bazzaz (1979), D. Rabinowitz, *personal communication*, and B. J. Rathcke, *personal communication*; i.e., an aggregated distribution of species in time. (See also Stiles 1977, 1979, Poole and Rathcke 1979.) It is in such distributions that difficulties with mean pairwise overlap are most obvious. The assemblage in Fig. 1 includes two disparate aggregations (species A-B and D-E-F-G), within which pairwise overlaps are very high. It seems intuitively satisfying that the overall intensity of overlap would be higher in this assemblage than the assemblage in Fig. 2, which is identical except that species B is lacking. By removing a species that is contributing to a high-overlap cluster, but keeping other relationships unchanged, should not the result be an overall reduction in overlap? However, using Schoener's (1970) "proportional similarity" measure of overlap, mean pairwise overlap is greater in the second assemblage. This "counterintuitive" behavior

---

[1] Manuscript received 14 January 1980; revised 14 June 1981; accepted 3 July 1981.

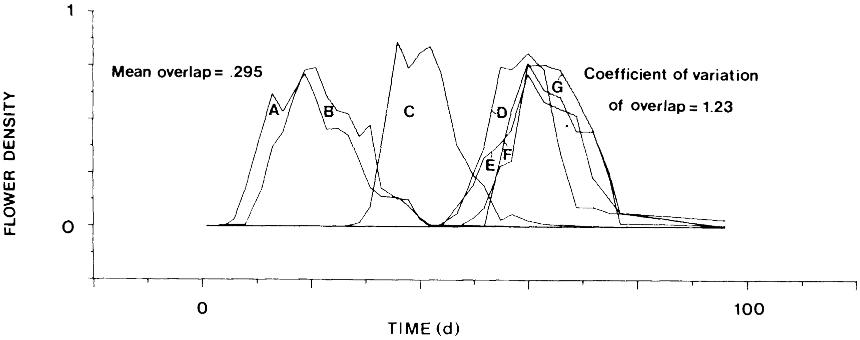

FIG. 1. The time-of-flowering curves for seven plant species, chosen for illustrative purposes. The vertical scale is arbitrary because the curves have been standardized.

is not a peculiarity of the overlap measure used, but a result of the insensitivity of the mean to changes in the shape of a distribution. It comes about because removing species B also removes the very small overlaps BD, BE, BF, and BG, and this loss of small overlaps more than offsets the loss of the large overlap AB in contributing to the mean. Thus, although the mean overlap is increased from .295 to .349, the coefficient of variation of the pairwise overlaps drops from 1.23 to 1.06. In this case, the "intuitive importance" of the change in assemblage composition lies mostly in its effect on the *distribution* of overlaps (see also Inger and Colwell 1977, Joern and Lawlor 1980).

The correspondence between verbal and mathematical hypotheses has often been rather loose in overlap ecology. We would be well within precedent if we concluded that assemblage 1 is more structured than assemblage 2 (as if by overlap-reducing competition) because the mean overlap is less. With equal respect for tradition, we could make the opposite claim because the overlaps are more even in assemblage 2.

To extend the comparison of mean and variance to a neutral-model analysis, we generated 100 "random" versions of the assemblage in Fig. 1 in which the shapes of the curves remained the same but their positions on the time axis were determined by random numbers. The mean of the mean pairwise overlaps for the set of random assemblages was essentially the same as that of the observed assemblage (3.00 to 2.95), while the variance of pairwise overlaps was much lower in the random constructions (.0896 to .1322). It appears worthwhile to examine both mean and variance of overlaps before making decisions about community structure vs. randomness (see Inger and Colwell 1977). However, neither index is universally suitable, which suggests a search for alternative methods.

Pleasants (1977) suggests that mean pairwise overlap behaves more acceptably if only overlaps with each species' first few nearest neighbors are considered. While this restriction will tend to prevent very small overlaps from dominating the average, as in our example, it entails both arbitrariness and loss of information. For example, Pleasants (1977:76) suggests that the first and second nearest neighbors are most important in one-dimensional systems such as that depicted in Fig. 1. Adopting this criterion for species C would include overlaps C-D and C-E, while ignoring C-A and C-B, which are nearly as large. Simply eliminating zero overlaps (cf. Pianka 1974) from the average also results in arbitrary loss of information.

Another alternative approach (P. Feinsinger, *personal communication*) is to deemphasize pairwise overlaps and compute instead a total overlap load for each species by summing its pairwise overlaps with

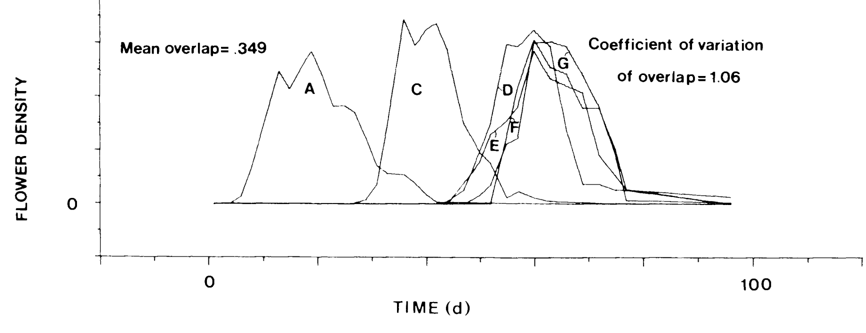

FIG. 2. The same set of curves as in Fig. 1, with species B removed. This change in composition increases the mean pairwise overlap.

each other species in the assemblage. By averaging these total overlaps one obtains the mean total overlap, which is another candidate for an index of overlap intensity. It would seem preferable to mean pairwise overlap because the use of total overlap loads focuses attention on the amount of overlap a species encounters, which is dependent on the number of species that it overlaps. The mean pairwise overlap on a species merely summarizes its average interaction without taking account of the number of such interactions. The total overlaps may also be more amenable to weighting by species abundance. However, in the simple case where abundances are not a consideration and where overlaps are symmetric (overlap A-B = overlap B-A), the mean total overlap is not much better than mean pairwise overlap as a way of summarizing overlap matrices. In fact, the two are directly related: (mean total overlap per species) = (mean pairwise overlap per pair) × (number of species − 1). Thus for our assemblages 1 and 2, the mean total overlaps are 1.77 and 1.75. The difference is now in the direction we intuitively expected, but appears trivial in magnitude.

Clearly, the situation is complex. What is required is a more precise connection between verbal statements about the "amount of overlap" and the mathematical indices used to approximate them, so that dependence on intuition can be eliminated. Hypotheses that are specifically based on particular indices are readily tested. For example, the hypothesis "the variance of the mean total overlap loads observed in assemblage X should be lower than that found in analogous random or pseudorandom assemblages generated by algorithm Y" is testable. (It may even have an explicable ecological basis, although that is the responsibility of its author). Hypotheses like "overlap should be lower in X than Y" are untestable without further qualification.

We concluded an earlier version of this note with a plea for ecologists to scrutinize their raw data in various ways before calculating any summary indices. Pianka (1980) has recently affirmed and extended this point in a useful and relevant paper partly devoted to the relative advantages of various data display and analysis techniques in uncovering "structure" in assemblages, especially "guild structure." Guild structure in this context means the subdivision of an assemblage into clusters of ecologically similar species (guilds). Overlaps are higher within guilds and lower between guilds; thus in our Fig. 1, a cluster analysis would probably identify two guilds consisting of species A and B, and D, E, F, and G. Species C would stand alone. Aside from cluster analysis for the detection of guilds, among the techniques Pianka suggests are examination of both the mean and the variation in pairwise overlap (following Inger and Colwell 1977) and the direct examination of the frequency distribution of overlaps. He points out (1980:196) that frequency distributions of overlaps will be affected by the extent of guild structure, citing this as a disadvantage that makes it difficult to interpret such frequency distributions. The point is similar to our demonstration that mean pairwise overlap behaves poorly when guild structure (i.e., aggregations) is present, as in Figs. 1 and 2. However, we take exception to the implication that one can avoid such problems by breaking assemblages down into guilds prior to analysis. The arbitrary nature of this procedure is fairly obvious. Suppose, for instance, that one wants to show that overlap has been minimized or that resource utilization curves are regularly spaced along a resource gradient. By using a cluster analysis to identify guilds, one can in principle keep splitting until all guilds have few enough species that species are necessarily regularly spaced within them. A good analogy is the well-known ability to demonstrate that a given spatial distribution of points is either regular or aggregated within sampling quadrats, simply by changing the size of the quadrats (Grieg-Smith 1957, Kershaw 1973). This sensitivity to quadrat size is exploited to detect spatial pattern in a Grieg-Smith (1961) pattern analysis, and perhaps an analogous technique could be developed for examining within-guild overlap patterns as the clustering level is changed. However, we think that it is more appropriate, and more in the spirit of Root's (1967) original definition, to delimit guilds by natural-historical criteria that are independent of the criteria used to investigate intraguild structure. Where the problem is to demonstrate regularity of overlap within a guild, we suggest that this amounts to using a different resource dimension to define guilds. In our example, which concerns the temporal overlaps of flowering curves, it would be appropriate to define a guild of plants as, for instance, those pollinated by bumblebees, but inappropriate to limit a guild's membership to those that bloom in May.

CONCLUSION

To some extent, the difficulty in finding good summary statistics for overlap matrices is simple loss of information about the distribution. When the goal is simply to translate a set of overlaps into a more comprehensible form with as little information loss as possible, one should probably not calculate indices at all, but use the multivariate techniques that have been developed precisely for summarizing this sort of matrix, e.g., principal components analysis, polar ordination, etc. In some cases where a very specific hypothesis can be stated, an index may be of value, but the hypothesis must be predicated on a firm understanding of the mathematical behavior and real implications of the index, whether it be the mean pairwise overlap, the variance of total species overlaps, or anything else. Because indices sometimes behave nonintuitively, selection of a plausible but inappropriate index can obscure real patterns. No index can replace the direct examination of data, and we join Connor and Simberloff (1979) in suggesting that such data be published whenever practical.

Acknowledgments

This paper is Contribution 393 from the Program in Ecology and Evolution, State University of New York at Stony Brook. We wish to thank P. Feinsinger, L. Lawlor, and P. Sale for valuable comments on previous versions.

Literature Cited

Brown, J. H. 1975. Geographical ecology of desert rodents. Pages 329–331 *in* M. Cody and J. Diamond, editors. Ecology and evolution of communities. Belknap Press, Cambridge, Massachusetts, USA.

Connor, E. F., and D. Simberloff. 1979. You can't falsify ecological hypotheses without data. Bulletin of the Ecological Society of America **60**:154–155.

Grieg-Smith, P. 1957. Quantitative plant ecology. Butterworths, London, England.

———. 1961. Data on pattern within plant communities. I. The analysis of pattern. Journal of Ecology **49**:695–702.

Inger, R., and R. B. Colwell. 1977. Organization of contiguous communities of amphibians and reptiles in Thailand. Ecological Monographs **46**:257–291.

Joern, A., and L. R. Lawlor. 1980. Food and microhabitat utilization by grasshoppers from arid grasslands: comparisons with neutral models. Ecology **61**:591–599.

Kershaw, K. A. 1973. Quantitative and dynamic plant ecology. American Elsevier, New York, New York, USA.

Parrish, J. A. D., and F. A. Bazzaz. 1979. Difference in pollination niche relationships in early and late successional plant communities. Ecology **60**:597–610.

Pianka, E. R. 1974. Niche overlap and diffuse competition. Proceedings of the National Academy of Sciences (USA) **71**:2141–2145.

———. 1975. Niche relations of dessert lizards. Pages 307–311 *in* M. Cody and J. Diamond, editors. Ecology and evolution of communities. Belknap Press, Cambridge, Massachusetts, USA.

———. 1980. Guild structure in lizards. Oikos **35**:194–201.

Pianka, E. R., R. B. Huey, and L. R. Lawlor. 1979. Niche segregation in desert lizards. *In* D. J. Horn, R. Mitchell, and G. Stairs, editors. Analysis of ecological systems. Ohio State University Press, Columbus, Ohio, USA.

Pleasants, J. M. 1977. Competition in plant-pollinator systems: an analysis of meadow communities in the Colorado Rocky Mountains. Dissertation. University of California at Los Angeles, Los Angeles, California, USA.

Poole, R. W., and B. J. Rathcke. 1979. Regularity, randomness, and aggregation in flowering phenologies. Science **203**:470–471.

Root, R. B. 1967. The niche exploitation pattern of the Bluegray Gnatcatcher. Ecological Monographs **37**:317–350.

Rusterholz, K. A. 1979. Niche relations among pine foliage-gleaning birds in different competitive regimes. Dissertation, University of Wisconsin, Madison, Wisconsin, USA.

Sale, P. 1974. Overlap in resource use and interspecific competition. Oecologia **17**:245–256.

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

Stiles, F. G. 1977. Coadapted competitors: the flowering seasons of hummingbird-pollinated plants in a tropical forest. Science **198**:1177–1178.

———. 1979. Regularity, randomness, and aggregation in flowering phenologies. Science **203**:471.

Thomson, J. D. 1975. Some community-level aspects of a bog pollination system. Thesis. University of Wisconsin, Madison, Wisconsin, USA.

———. 1978. Competition and cooperation in plant-pollinator systems. Dissertation. University of Wisconsin, Madison, Wisconsin, USA.