Shuffled Sticks: On Calculating Nonrandom Niche Overlaps

George Sugihara

The American Naturalist, Vol. 127, No. 4. (Apr., 1986), pp. 554-560.

Stable URL:
http://links.jstor.org/sici?sici=0003-0147%28198604%29127%3A4%3C554%3ASSOCNN%3E2.0.CO%3B2-I

The American Naturalist is currently published by The University of Chicago Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ucpress.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org
Mon Feb 25 20:19:01 2008

SHUFFLED STICKS: ON CALCULATING NONRANDOM NICHE OVERLAPS

George Sugihara*

Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

Submitted April 21, 1983; Revised September 10, 1984; Accepted July 10, 1985

A topic of recent interest concerns how to distinguish true patterns in community organization from those generated by random or null processes (e.g., Heatwole and Levins 1972; Diamond 1975; Connor and Simberloff 1979; Toft and Shea 1983). This controversy, whose philosophical underpinnings harken back to the debate between the biotic and climate schools, centers on the question of whether communities should be viewed as random conglomerates of species that passively match the habitat or as systems internally structured by biological interactions. One vagile though currently well-defended front of this controversy advocates the importance of competition in structuring communities (MacArthur 1972; Pianka 1974; Schoener 1974; Wiens 1977; Gatz 1979; Ricklefs and Travis 1980; Connell 1983; Roughgarden 1983). Central to this debate is the question of whether niche overlaps among species in nature are minimized by competition in a statistically nonrandom way. This paper presents calculations for determining explicitly whether the extent of niche overlaps along some resource dimension (e.g., space, time) is minimized in a statistically significant manner.

Recently, De Vita (1979) proposed an interesting quantitative technique for studying this problem, which asks whether distances between the means of species' utilization spectra along a single resource continuum (e.g., space or time) are more uniform than would be expected with random placement. This is essentially a retooled broken-stick model wherein utilization means are thrown down randomly to subdivide a unit interval (MacArthur 1957). This competition-free null hypothesis asks whether the observed distribution of intermean distances can be distinguished from a broken-stick distribution. In application, data for stem-boring insects, tropical hummingbirds, predatory cone snails (De Vita 1979), and flowering phenologies (Poole and Rathke 1980) all appear to agree with the null broken-stick pattern, thereby arguing against the importance of competition as a force in structuring these communities.

* Present address (effective May 1986): Scripps Institution of Oceanography, Graduate Department, A-008, La Jolla, California 92093.

Pielou (1981) pointed out, however, that such agreement is not surprising because, according to the model, all ranked-size (intermean-distance) lists should be equally probable (see also Eberhardt 1969; Sugihara 1980; Shelly and Christensen 1982). This property of the broken-stick model results in a null hypothesis that is difficult to falsify (large Type II error; Sokal and Rohlf 1981). Furthermore, restricting attention to the means of the utilization spectra may be misleading for the problem of detecting minimal overlap. One can easily imagine a situation in which the utilization spectra are arranged end to end with no overlap at all, but the intermean distances fit the expectation of the broken-stick pattern. It would be preferable, therefore, to direct attention to the extent of the overlaps themselves. One could then ask whether the amount of overlap observed is less than would be expected if the utilization spectra were randomly placed on the resource continuum.

THE PROBLEM

Here, I present calculations for determining whether utilization spectra overlap in a manner that is consistent with random placement (see also Pielou 1977). These results are the analytical extension of the computer simulations used by Cole (1981) to measure overlap in flowering intervals.

In abstract terms, I ask what the complete distribution of pairwise overlaps is when sticks (utilization spectra) of known length are tossed randomly into an interval of known size. In this way, the resource continuum is likened to a long narrow pencil box and the utilization spectra (ranges) to sticks of various lengths, which are randomly and independently shuffled in the box. Our first objective will be to determine the distribution of the extent of overlap between any chosen pair of sticks. These results will then be gathered into a single statistic for all independent pairs to determine the probability that the entire collection represents a randomly shuffled array. This is intended to complement Pielou's elegant method of sheaves (1977), in which the extent of pairwise overlap is not explicitly taken into account.

COMPUTING THE DISTRIBUTION OF PAIRWISE OVERLAP

Suppose we have two sticks of length l_1 and l_2, with $l_1 \geq l_2$, thrown randomly and independently into an interval of size L spanning 0 to L, $L \geq l_1, l_2$. The overlap is

$$W = (l_1 \cap l_2) l_2^{-1} .$$

We solve for the cumulative density of W in terms of the distance between the midpoints of each stick, y_2 (see fig. 1). In particular, if x_1 and x_2 are the positions of the midpoints of the long and short sticks, respectively, then the distance separating them, y_2, is

$$y_2 = |x_1 - x_2| ,$$

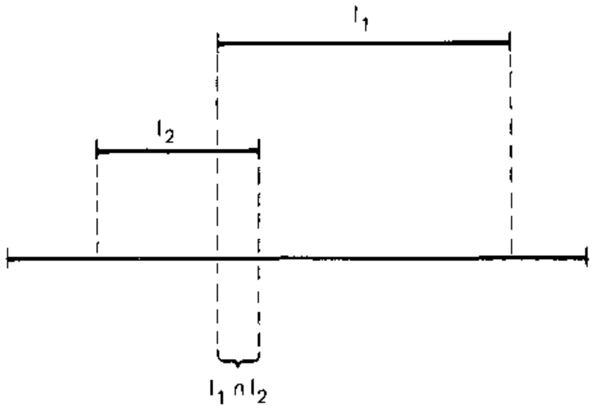

FIG. 1.—Two overlapping linear niches of size l_1 and l_2 within a larger resource interval.

and the overlap is

$$W = \begin{cases} 0 & \text{for } \bar{l} \leq y_2 \quad (1) \\ (\bar{l} - y_2)l_2^{-1} & \text{for } l^* \leq y_2 \leq \bar{l} \quad (2) \\ 1 & \text{for } 0 \leq y_2 \leq l^* \quad (3) \end{cases}$$

where $\bar{l} = ½(l_1 + l_2)$ and $l^* = ½(l_1 - l_2)$. As calculated in the Appendix, the probability density of y_2 is

$$g(y_2) = \begin{cases} 2\Delta_2^{-1} & \text{for } 0 \leq y_2 \leq l^* \quad (4) \\ 2(L - \bar{l} - y_2)\Delta_1^{-1}\Delta_2^{-1} & \text{for } l^* \leq y_2 \leq L - \bar{l} \quad (5) \end{cases}$$

where $\Delta_1 = L - l_1$ and $\Delta_2 = L - l_2$. Notice that the limits in equation (3) match those in equation (4). Hence, the probability of complete overlap, $P(W = 1)$, can be calculated from $g(y_2)$ in equation (4) as follows:

$$P(W = 1) = \int_0^{l^*} 2\Delta_2 \, dy_2 = 2l^*\Delta_2^{-1}. \quad (6)$$

Likewise, from (1), (2), and (5), the cumulative density for $0 \leq W < 1$ is

$$P(W \leq x) = \int_{\bar{l} - l_2 x}^{L - \bar{l}} 2(L - \bar{l} - y_2)\Delta_1^{-1}\Delta_2^{-1} \, dy_2, \quad (7)$$

where x is a number on the interval $0 \leq x < 1$. Therefore, the complete cumulative density function (cdf) for $0 \leq W < 1$ may be obtained by piecing together (6) and (7). Equation (6), $P(W = 1)$, is figured separately because the cdf may show a spike at $W = 1$. The size of this spike depends on the degree to which the two sticks differ in length; the spike is higher if they differ greatly.

A straightforward though tedious calculation (see the Appendix) shows that the expected overlap for the two sticks of length l_1 and l_2 thrown randomly into an interval of size L, for $l_1 + l_2 \leq L$, is

$$E(W) = (Ll_1 - l_1^2 - ⅓l_2^2)\Delta_1^{-1}\Delta_2^{-1}.$$

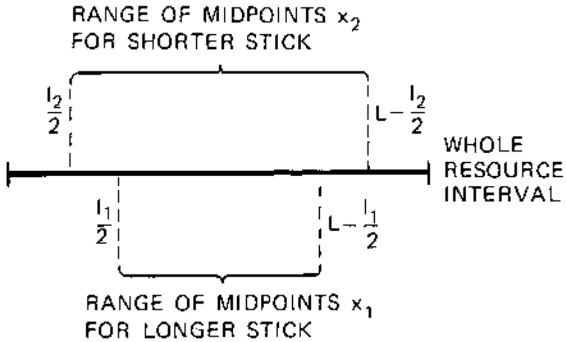

FIG. 2.—The problem of computing the overlap between two sticks of size l_1 and l_2 thrown down randomly into an interval of size L is transformed into a problem of finding the distance between the midpoints, x_1 and x_2, for each pair of sticks. The figure illustrates the ranges of the midpoints for the shorter and longer sticks in relation to the whole resource interval.

Therefore, if the two sticks are of equal size and exactly cover the interval ($l_1 = l_2 = L/2$), then the expected overlap is $E(W) = 2/3$.

One can use equation (7) to test whether the overlap of a given pair of species (sticks) is significantly less than expected at random. Plugging in observed values for x (observed overlap between the two species), l_1 and l_2 (the observed ranges of the utilization spectra), and L (the size of the resource continuum) yields a P value for the probability that two sticks overlap less than x. These P values may be obtained for any pair of species.

TESTING NONRANDOM OVERLAP FOR AN ASSEMBLAGE, $n > 2$

We will now outline a procedure for combining P values for each pairwise overlap into a single test statistic for judging the likelihood that a collection of n species (sticks) overlaps less than would be expected at random. This is the relevant test for measuring the role of competition in minimizing overlap for the whole assemblage.

As seen in figure 2, a given array of n sticks produces only $n - 1$ independent overlaps, and hence $n - 1$ independent P values. The $n - 1$ independent pairs may be chosen randomly as follows. With uniform probability, choose one of the n sticks at random without replacement. Choose the other member of this first pair from the $n - 1$ remaining. To form the $n - 2$ subsequent pairs, randomly choose one stick from the set of those already chosen and one stick from those not yet sampled. This procedure amounts to constructing a random tree graph whose $n - 1$ random edges or vertex pairs correspond to random-stick pairs.

For each of these $n - 1$ pairs compute a P value using equation (7). Using Fisher's method for combining independent tests on the independent P values gives a χ^2 value (with $n - 1$ df):

$$\chi^2 = \sum_{i=1}^{n-1} -2 \log P_i .$$

This χ^2 value is an aggregated statistic describing the likelihood that an array of n sticks (species' utilization spectra) overlaps less than at random.

LIMITATIONS

Perhaps the most severe limitation to the above method is obtaining reliable measurements of x (observed pairwise overlap), l_1 and l_2 (species ranges), and L (size of ensemble resource continuum). In computing the first three quantities, it is assumed that one can reasonably reduce a complete utilization spectrum to an interval (range) over the resource continuum. A judgment needs to be made about where the endpoints fall for each species. Similarly, the boundaries for L need to be chosen carefully. For example, in looking at the distribution of different species of microflora in a gut, where does the gut begin and end? Likewise, in examining flowering phenologies in relation to competition for pollinators, when does the flowering season begin and end? A large L decreases the expected overlap under the null (random) hypothesis, whereas a small L increases the expected overlap. The investigator needs to be careful, therefore, in choosing a conservatively large resource interval to demonstrate competitive displacement.

It is worth mentioning that Cole (1981) used a computer simulation of randomly placed line segments to reanalyze earlier results on flowering phenologies studied by Stiles (1982). Using De Vita's (1979) reinterpretation of the broken-stick model for intermean distances, Poole and Rathke (1980) reported no evidence of competition for pollinators. Using the shuffled-sticks model, however, Cole found that the interspecies overlap was significantly less than would be expected with random placement, thereby arguing for the importance of competition in setting flowering times.

SUMMARY

An analytical method is presented for detecting nonrandom niche displacement that might result from competition by examining the degree to which niche overlap is minimized along a single niche dimension. Species' utilization spectra for different resources are viewed as intervals, or "sticks," arranged over a finite linear resource continuum. The method tests the actual extent to which the observed configuration of sticks overlaps less than would be expected if the sticks were cast down at random. Data previously analyzed to refute the importance of niche displacement (Poole and Rathke 1980) have been found to support it when reanalyzed with a computer simulation of the shuffled-sticks model (Cole 1981).

ACKNOWLEDGMENTS

The author would like to acknowledge the following people for their assistance in this work: B. Cole, S. Christensen, R. Gardner, R. V. O'Neill, D. Downing, and D. DeAngelis. S. Christensen suggested publishing these results, which would otherwise remain as notes in my file cabinet. The research was supported by a Eugene P. Wigner Fellowship from Oak Ridge National Laboratory, National Science Foundation Ecosystems Studies Program, under Interagency Agreement

DEB 77-25781 with the U.S. Department of Energy under contract DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc.

APPENDIX

Herein we will derive a cumulative distribution function of the overlap W between two sticks of length l_1 and l_2 ($l_1 \geq l_2$) thrown randomly and independently into the interval 0 to L, $L \geq l_1, l_2$.

To account for upper (L) and lower (0) limits of the interval, I recast the problem in terms of the position of the midpoints for each stick x_i. Specifically, the midpoint of the longer stick ranges from $l_1/2$ to $L - l_1/2$, and the shorter stick ranges from $l_2/2$ to $L - l_2/2$. Randomly shuffling sticks in [0, L] translates to throwing down midpoints with uniform probability in each of the above intervals. The distance between midpoints, $y_2 = |x_1 - x_2|$, can then be translated into the overlap W:

$$W = \begin{cases} 0 & \text{for } \bar{l} \leq y_2 \\ (\bar{l} - y_2)l_2^{-1} & \text{for } l^* \leq y_2 \leq \bar{l} \\ 1 & \text{for } 0 \leq y_2 \leq l^*, \end{cases} \tag{A1}$$

where $\bar{l} = \frac{1}{2}(l_1 + l_2)$ and $l^* = \frac{1}{2}(l_1 - l_2)$.

Let x_1 be a uniform random variable in the interval $[l_1/2, L - l_1/2]$ and x_2 a uniform random variable in $[l_2/2, L - l_2/2]$. The cumulative densities are

$$P(x_1 \leq x) = x\Delta_1^{-1}$$
$$P(x_2 \leq x) = x\Delta_2^{-1},$$

and the probability densities are

$$fx_1 = \Delta_1^{-1}$$
$$fx_2 = \Delta_2^{-1},$$

where $\Delta_1 = L - l_1$ and $\Delta_2 = L - l_2$. Hence, the joint density for x_1 and x_2 is

$$\phi(x_1, x_2) = \Delta_1^{-1}\Delta_2^{-1}.$$

The two-dimensional set x_1, x_2 will now be transformed into the new planar set y_1, y_2^*, where

$$y_1 = x_1 \quad \text{and} \quad y_2^* = x_1 - x_2.$$

Because the Jacobian for this transformation is one, the new joint density is simply $g(y_1, y_2^*) = \Delta_1^{-1}\Delta_2^{-1}$. Considering only positive y_2^*, the probability density for $y_2 = |x_1 - x_2|$ is simply

$$g(y_2) = \begin{cases} 2\int_{l_1/2}^{L-l_1/2} \dfrac{dy_2}{\Delta_1\Delta_2} & \text{for } 0 \leq y_2 \leq l^* \\ 2\int_{y_2+l_1/2}^{L-l_2/2} \dfrac{dy_2}{\Delta_1\Delta_2} & \text{for } l^* \leq y_2 \leq L - \bar{l}, \end{cases}$$

which reduces to

$$g(y_2) = \begin{cases} 2\Delta_2^{-1} & \text{for } 0 \leq y_2 \leq l^* \\ 2(L - \bar{l} - y_2)\Delta_1^{-1}\Delta_2^{-1} & \text{for } l^* \leq y_2 \leq L - \bar{l}. \end{cases} \tag{A2}$$

Comparing limits in equations (A1) and (A2), the cumulative density for W is

$$P(W = 1) = \int_0^{l^*} \frac{2dy_2}{\Delta_2} = \frac{2l^*}{\Delta_2}, \tag{A3}$$

and for $x < 1$,

$$P(W \le x) = \int_{\bar{l} - l_2 x}^{L - l^*} \frac{2(L - \bar{l} - y_2)}{\Delta_1 \Delta_2} \, dy_2 . \tag{A4}$$

Notice that the cumulative distribution function may not be continuous and, because of equation (A3), may spike at $W = 1$.

The expected overlap, $E(W)$, may be calculated from equations (A1) and (A2) in a similar fashion for $L \ge l_1 + l_2$.

$$E(W) = \int_0^{l^*} \frac{2dy_2}{\Delta_2} + \int_{l^*}^{\bar{l}} \frac{\bar{l} - y_2}{l_2} \left[\frac{2(L - \bar{l} - y_2)}{\Delta_1 \Delta_2} \right] dy_2 ,$$

which reduces to

$$E(W) = (Ll_1 - l_1^2 - \tfrac{1}{3} l_2^2) \, \Delta_1^{-1} \Delta_2^{-1} .$$

LITERATURE CITED

Cole, B. 1981. Overlap regularity and flowering phenology. Am. Nat. 117:993–997.

Connell, J. H. 1983. On the prevalence and relative importance of interspecific competition: evidence from field experiments. Am. Nat. 122:661–696.

Connor, E. F., and D. Simberloff. 1979. The assembly of species communities: chance or competition? Ecology 60:1132–1140.

De Vita, J. 1979. Niche separation and the broken-stick model. Am. Nat. 114:171–178.

Diamond, J. M. 1975. Assembly of species communities. Pages 342–444 *in* M. L. Cody and J. M. Diamond, eds. Ecology and evolution of communities. Belknap Press of Harvard University Press, Cambridge, Mass.

Eberhardt, L. L. 1969. Some aspects of species diversity models. Ecology 50:503–505.

Gatz, A. J. 1979. Community organization in fishes as indicated by morphological features. Ecology 60:711–718.

Heatwole, H., and R. Levins. 1972. Trophic structure stability and faunal change during recolonization. Ecology 53:531–534.

MacArthur, R. H. 1957. On the relative abundance of bird species. Proc. Natl. Acad. Sci. USA 43:293–295.

———. 1972. Geographical ecology. Harper & Row, New York.

Pianka, E. R. 1974. Niche overlap and diffuse competition. Proc. Natl. Acad. Sci. USA 71:2141–2145.

Pielou, E. C. 1977. The statistics of biogeographic range maps: sheaves of one-dimensional ranges. Bull. Int. Stat. Inst. 47(2):111–122.

———. 1981. The broken-stick model: a common misunderstanding. Am. Nat. 117:609–610.

Poole, R. W., and B. J. Rathke. 1980. Regularity, randomness and aggregation in flowering phenologies. Science (Wash., D.C.) 203:407–471.

Ricklefs, R. E., and J. Travis. 1980. A morphological approach to the study of avian community organization. Auk 97:321–338.

Roughgarden, J. 1983. Competition and theory in community ecology. Am. Nat. 122:583–601.

Schoener, T. W. 1974. Resource partitioning in ecological communities. Science (Wash., D.C.) 185: 27–39.

Shelly, T. E., and S. W. Christensen. 1982. Testing for competition: a critique of De Vita's use of the broken-stick model. Am. Nat. 119:435–437.

Sokal, R. R., and F. J. Rohlf. 1981. Biometry. 2d ed. Freeman, San Francisco.

Stiles, E. W. 1982. Fruit flags: two hypotheses. Am. Nat. 120:500–509.

Sugihara, G. 1980. Minimal community structure: an explanation of species abundance patterns. Am. Nat. 116:770–787.

Toft, C. A., and P. J. Shea. 1983. Detecting community-wide patterns: estimating power strengthens statistical inference. Am. Nat. 122:618–625.

Wiens, J. A. 1977. On competition and variable environments. Am. Sci. 65:590–597.