



**Self-Similarity and Clustering in the Spatial  
Distribution of Species**

Annette Ostling, *et al.*

*Science* **290**, 671a (2000);

DOI: 10.1126/science.290.5492.671a

***The following resources related to this article are available online at  
www.sciencemag.org (this information is current as of April 21, 2008 ):***

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/290/5492/671a>

This article **cites 6 articles**, 4 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/290/5492/671a#otherarticles>

This article appears in the following **subject collections**:

Ecology

<http://www.sciencemag.org/cgi/collection/ecology>

Technical Comments

[http://www.sciencemag.org/cgi/collection/tech\\_comment](http://www.sciencemag.org/cgi/collection/tech_comment)

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/about/permissions.dtl>

# Self-Similarity and Clustering in the Spatial Distribution of Species

Condit *et al.* (1) examined the spatial aggregation of individuals of tree species in six different tropical forest sites. They found that most species were aggregated, that rare species tended to be more aggregated than abundant ones, and that smaller individuals of a species tended to be more aggregated than larger individuals. The observed clustering pattern, they pointed out, was inconsistent with a random distribution of individuals. We show here that the clustering described in (1) is similar to that expected for a species with a self-similar spatial distribution. Such a distribution has been demonstrated to correctly predict range characteristics across a wide variety of taxa and spatial scales (2) and is the analog to the community-level self-similarity shown in (3) to be equivalent to the power-law species-area relationship (SAR).

To measure the clustering of a species, Condit *et al.* (1) used an index called the relative neighborhood density,  $\Omega_{x_1, x_2}$ . For a given species,  $\Omega_{x_1, x_2}$  is equivalent to the average density of conspecifics in the neighborhood of individuals, normalized by the density of individuals of the species in the entire plot:

$$\Omega_{x_1, x_2} = \frac{\langle n_{x_1, x_2} \rangle}{\frac{A_{x_1, x_2}}{A_0}} \quad (1)$$

where  $\langle n_{x_1, x_2} \rangle$  is the number of conspecifics located between a distance  $x_1$  and a distance  $x_2$  from each individual, averaged over all individuals of the species;  $A_{x_1, x_2}$  is the area of the annuli defined by the radii  $x_1$  and  $x_2$ ; and  $n$  is the number of individuals of the species in a plot of area  $A_0$  (4). If  $\Omega > 1$  at distances that are short relative to the plot size, the species is considered clustered, whereas  $\Omega < 1$  at short distances indicates spacing or dispersion of individuals.

Self-similarity in the distribution of a single species is defined as follows. Let  $A_0$  be a rectangular plot whose dimensions have the ratio  $\sqrt{2}$ , and let  $A_i = A_0/2^i$  be the size of areas obtained from  $A_0$  by  $i$  shape-preserving bisections (5). Given that a species is in a particular area of size  $A_i$ , let  $\alpha_i$  be the average probability that it is in at least a particular one of the two  $A_{i+1}$  contained in  $A_i$ . The distribution of the species is self-similar, or scale invariant, if  $\alpha_i = \alpha$  is independent of  $i$ .

Self-similarity relates properties of the distribution of a species at small scales to such properties at larger scales. For example,

self-similarity relates  $\bar{n}_i$ , here defined as the average abundance of a species in the  $A_i$  that it occupies, and  $n$ , the total abundance of the species in the plot  $A_0$ :

$$\begin{aligned} \bar{n}_i &\equiv \frac{n}{\# \text{ of } A_i \text{ occupied by species}} \\ &= \frac{n}{(2\alpha)^i} \end{aligned} \quad (2)$$

where  $\alpha$  depends on the species but not on  $i$ . Eq. 2 can be used to obtain an expression for  $\Omega_{x_1, x_2}$  in terms of the  $\alpha$  for each species. Let  $\Omega_i$  refer to  $\Omega_{x_1, x_2}$  in the case where  $x_1$  is the radius of a circle of size  $A_{i+1}$  and  $x_2$  is the radius of a circle of size  $A_i$ . For this case, the numerator in Eq. 1 can be closely approximated (6) by  $[1/(A_{i+1})](\bar{n}_i - \bar{n}_{i+1})$ , and hence

$$\Omega_i = \frac{2}{\alpha^i} - \frac{1}{\alpha^{i+1}} \quad (3)$$

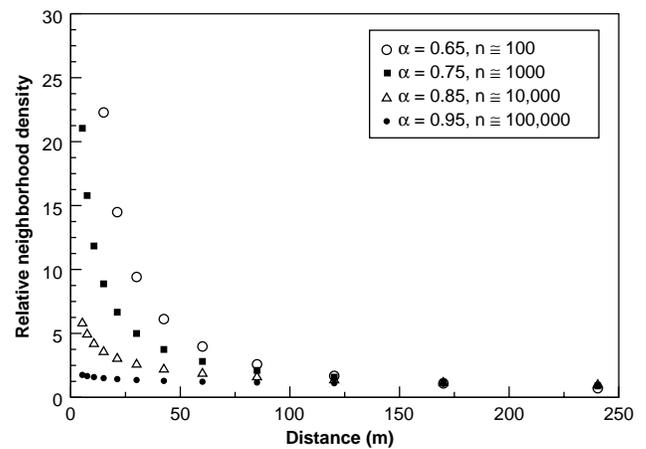
$\Omega_i$  can be expressed directly in terms of the distance  $r_i$ , defined as the average of the radii of  $A_i$  and  $A_{i+1}$ . Note that  $\Omega_i = (1/\alpha)^i \Omega_0$ , where  $\Omega_0 = 2 - (1/\alpha)$ , and that  $r_i = (1/\sqrt{2})^i r_0$ , where  $r_0 = (1/2)(1 + 1/\sqrt{2})(\sqrt{A_0}/\pi)$ . By defining  $(\sqrt{2})^w \equiv \alpha$ , we can write

$$\Omega_i = Cr_i^w \quad (4)$$

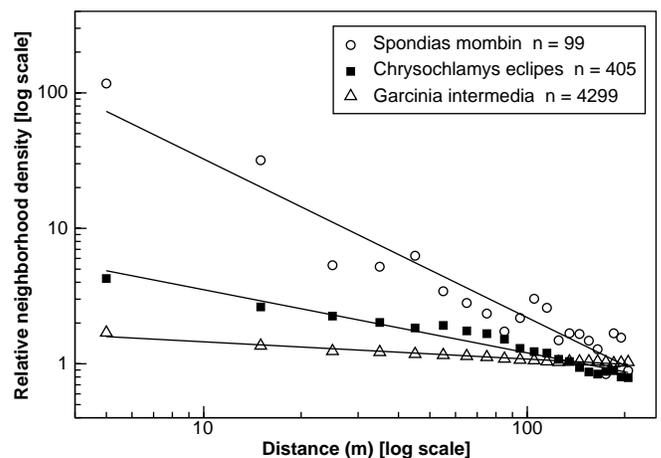
where  $C = \Omega_0/r_0^w$ .

The relative neighborhood density of self-similar distributions (Fig. 1) has characteristics in common with that of tropical forest plots described in (1). It is largest at the smallest scales and monotonically decreases with scale at a rate which is largest at small scales (7). Furthermore, since  $\alpha$  increases with abundance (8),  $\Omega$  at small distances will be largest for rare species. We also did a more quantitative comparison of the relative neighborhood density (as defined in Eq. 4) for 20 species chosen from (1) over a range of abundances. For each species, a linear regression was performed on the log-transformed data (9). The linear regression analysis yielded  $r^2$

**Fig. 1.** Relative neighborhood density,  $\Omega_r$ , as a function of distance,  $r$ , as expected for self-similarly distributed species with  $\alpha = 0.65, 0.75, 0.85$ , and  $0.95$ , in a plot of size  $A_0 = 50$  ha. These values of  $\alpha$  correspond to species with abundances  $n \approx 10^2, 10^3, 10^4$ , and  $10^5$ , respectively, if there are 300,000 individuals in the plot and a simplifying assumption is made (15). The smallest distance plotted is 5.3 m (this point is off the graph for  $\alpha = 0.65$ ).



**Fig. 2.** Relative neighborhood density,  $\Omega$ , versus distance,  $r$ , for three species in (1). The lines are the result of linear regression on the log-transformed data (to eliminate heteroscedasticity) (9):  $\ln \Omega = 480 - 1.17 \ln r, r^2 = 0.914, \alpha = 0.667 (\pm 0.077)$  for *Spondias mombin*;  $\ln \Omega = 10.3 - 0.466 \ln r, r^2 = 0.946, \alpha = 0.851 (\pm 0.022)$  for *Chrysochlamys eclipes*; and  $\ln \Omega = 1.95 - 0.128 \ln r, r^2 = 0.959, \alpha = 0.978 (\pm 0.008)$  for *Garcinia intermedia*, where number in parentheses are standard errors.



## TECHNICAL COMMENT

$> 0.9$  for half of the species, and  $r^2 > 0.8$  for all but two (10). Fig. 2 shows three of these analyses.

The exponent  $w$  in Eq. 4 is directly related to the exponent  $y'$  that appears in the relationship between the box-counting measurement of range size  $R$  of a species and the area  $A$  of the grid cell used to measure the range,  $R = A_0 (A/A_0)^{y'}$ . This range-area relationship was observed by Kunin for British floral census data (11) and derived from single-species self-similarity in (2). The clustering exponent  $w$  is related to the range-area exponent  $y'$  by

$$w = -2 y' \quad (5)$$

Condit *et al.* could test Eq. 5 using their tropical forest data.

The relative neighborhood density of tropical tree species as described in (1) has characteristics in common with the relative neighborhood density expected of a self-similarly distributed species, but community-level self-similarity does not hold at the tropical forest sites studied by Condit *et al.* (12, 13). However, although the notion is counterintuitive, a group of species whose distributions are individually self-similar is not expected to be self-similarly distributed at the community level (2). Therefore, the lack of community-level self-similarity in the sites studied by Condit *et al.* is not evidence against the possibility of species-level self-similarity in those sites.

**Annette Ostling**  
**John Harte**

*Energy and Resources Group  
University of California at Berkeley  
Berkeley, CA 94720, USA  
E-mail: aostling@socrates.berkeley.edu*

**Jessica Green**

*Department of Nuclear Engineering  
University of California at Berkeley*

### References and Notes

1. R. Condit *et al.*, *Science* **288**, 1414 (2000).
2. J. Harte, T. Blackburn, A. Ostling, in preparation.
3. J. Harte, A. Kinzig, J. Green, *Science* **284**, 334 (1999).
4. When calculating  $\Omega_{x_1, x_2}$  at large  $x_1, x_2$  from actual data, some of the area covered by the annuli  $A_{x_1, x_2}$  will fall outside of the plot area in which the location of individuals is known. In these cases, the average density of the species in the portions of the annuli inside of the plot is assumed to be representative of the average density of the species over the entire area covered by the annuli.
5. We take the set of areas  $A_i$  to all have the same shape because a self-similar distribution of a single species is expected to have shape-dependent properties in much the same way that the community-level version of self-similarity, the SAR, has shape dependence (14).
6. This expression is an approximation in the sense that the  $A_i$  and  $A_{i+1}$  relevant to  $\Omega_i$  are circles, rather than the same shape as the plot  $A_0$ . The difference between the average number of individuals in a circle and the average number of individuals in a rectangle of the same size is probably small, however, as long as the rectangle is not long and thin.
7. Some species distributions in (1) have a nonmono-

tonic decrease in  $\Omega$ . Any distribution, however, including one with a nonmonotonic decrease in  $\Omega$ , can be described by the probabilities  $\alpha_i$  if  $\alpha_i$  is allowed to vary with  $i$ .

8. For a given  $i$ ,  $\alpha$  increases with abundance across species that have the same  $\bar{n}_i$  (Eq. 2); hence,  $\Omega$  at small scales will be largest for the rarest of such species.
9. The linear regression was performed only over distances between 5 and 200 m, where most of the annuli used in calculating  $\Omega$  fall inside the plot.
10. For most of the 20 species examined, the value of  $\alpha$  calculated from the expression in (15) is not equal to that determined from the linear regression analysis, an indication that the simplifying assumption in (15) is not valid for these sites.
11. W. Kunin, *Science* **281**, 1513 (1998).
12. R. Condit *et al.*, *J. Ecol.* **84**, 549 (1996).
13. J. Plotkin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10850 (2000).
14. J. Harte, S. McCarthy, K. Taylor, A. Kinzig, M. L. Fisher, *Oikos* **86**, 45 (1999).
15. Under the simplifying assumption that each species has no more than one individual in areas of size  $A_0/N_0$ , where  $N_0$  is the total number of individuals across all tree species in the plot, the value of  $\alpha$  for a species is related to its abundance (2) by:

$$\alpha = \left( \frac{n}{N_0} \right)^{\frac{1}{\log_2(N_0)}}$$

16. We gratefully acknowledge financial support from the Class of 1935 Distinguished Professorship Fund of the University of California at Berkeley and from the American Association of University Women.

24 July 2000; accepted 1 September 2000

*Response:* The comment by Ostling *et al.* is elegant and interesting—and, indeed, precisely echoes the content of several paragraphs removed from the report by Condit *et al.* (1) during the editing process. In the following brief discussion, I paraphrase those omitted paragraphs for the present context, and offer several other observations on the Ostling *et al.* comment.

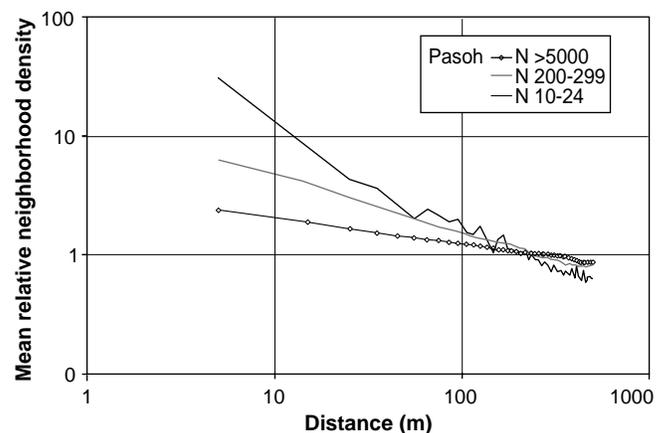
If the neighborhood density function for a species declined linearly on a log-log scale, then the species' distribution would be fractal and scale-invariant because the intensity of aggregation would decay similarly at all scales. Astronomers describe the distribution of galaxies as

being fractal in exactly this way. Most individual species in the forests examined by Condit *et al.* (1), however, did not display scale invariance across the plots. More typically for common species, neighborhood density declined at short distances more rapidly than log-linearly, and then leveled out. Other species showed more gentle declines initially, then rapid declines at greater distances.

Intriguingly, however, the aggregate behavior of the whole communities—the sum of relative neighborhood density across species—was indeed fractal, and showed very consistent patterns across forests (Fig. 1). The most abundant species had relatively gentle declines and large  $x$ -intercepts, while rare species had steep declines and smaller  $x$ -intercepts. (The  $x$ -intercept on a log-log scale is the distance at which  $\Omega = 1$ . Because  $\Omega > 1$  at short distances signifies at least some degree of aggregation, the  $x$ -intercept can thus be viewed as the clump radius, or the distance at which clumping ceases to be important.) The slope of these lines reflects the fractal dimension,  $D$ , because  $D$  is equivalent to the slope plus two:  $D = 2$  indicates spatial randomness;  $D = 0$  would be complete clumping, with all individuals concentrated at a single point (2).  $D$  for an aggregate of all common species varied from 1.65 to 1.83 in the six plots, and for aggregated rare species varied from 0.86 to 1.41.  $D$  declined smoothly with abundance at all plots, reflecting the tendency for rare species to be more clumped. Thus, in aggregate, the forests are scale invariant, and this should reflect scale-invariance in how species composition changes through space, although Condit *et al.* (1) did not investigate this.

Ostling *et al.* have cleverly shown how their description of self-similarity corresponds with the neighborhood function. This is useful, because the method based on quadrant occupancy that they have used can be associated with geographic ranges. Perhaps

**Fig. 1.** Aggregate neighborhood density functions from the Pasoh 50-ha plot. The steepest line, with wiggles, is the aggregate neighborhood function for all 89 species with 10 to 24 individuals in 50 ha. The aggregate function was calculated by taking the arithmetic average of all 89 individual neighborhood functions. The gray line, with intermediate slope, is the aggregate neighborhood function for all 73 species with 200 to 299 individuals, and the flattest line the aggregate for the seven species with  $\geq 5000$  individuals. At each of the plots, species were aggregated into abundance categories and the neighborhood functions were aggregated; in nearly all cases, the aggregate functions were very close to linear on the log-log scale, and always steeper for less common species.



## TECHNICAL COMMENT

they can make something of the observation in Fig. 1, that in an aggregate sense, the communities appear to be quite precisely self-similar.

Ostling *et al.* mention several tests that could be done using our distribution data for large forest plots; I would be happy to make data sets available if they would like to pursue the tests. And, finally, I present a challenge: Can the theories that Ostling *et al.* have put forth here predict ranges at much wider scales? The 50-ha plots have been excellent for testing predictions because distributions are completely known. But at larger scales, the data that I work with are far sparser—a few tens of plots, scattered over 1000 km<sup>2</sup>—and we don't

know the distributions of trees at these scales. I would like to draw conclusions, based on these sparse data, to questions such as, for example, how many species are widespread and how many occur in only one area. Can self-similarity suggest a way?

**Richard Condit**

*Center for Tropical Forest Science  
Smithsonian Tropical Research Institute  
Unit 0948  
APO AA 34002-0948, USA  
E-mail: [condit@ctfs.sri.si.edu](mailto:condit@ctfs.sri.si.edu)*

### References and Notes

1. R. Condit *et al.*, *Science* **288**, 1414 (2000).
2. This discussion follows the proofs and illustrations in (3), in which it is shown that the cluster dimension—

defined as the slope of log  $K$  versus log  $x$ , where  $K$  is Ripley's  $K$ —is equal to the fractal dimension. So  $K(x) = cx^D$ , where  $D$  is the fractal dimension and  $c$  is a constant.  $\Omega_x$ , the relative neighborhood density function as defined by Condit *et al.* (1), is

$$\Omega_x = \frac{1}{\bar{n}} \frac{K(\Delta x + x) - K(x)}{A(\Delta x + x) - A(x)}$$

where  $\bar{n}$  is the mean density across the plot and  $A(x)$  is the area of a circle with radius  $x$ . In the limit, this is equivalent to

$$\Omega_x = \frac{1}{2\pi\bar{n}x} \frac{dK}{dx},$$

or  $\Omega_x = c'x^{D-2}$ , so the fractal dimension is found by adding 2 to the slope of log  $\Omega_x$  versus log  $x$ .

3. H. M. Hastings, G. Sugihara, *Fractals: A User's Guide for the Natural Sciences* (Oxford Univ. Press, Oxford, 1993).