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, $\Omega_{r}$. For a given species, $\Omega_{r}$ 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:

$$\Omega_{r} = \left\langle n_{i} \right\rangle_{x_1,x_2} = \frac{\left\langle n_{i} \right\rangle_{x_1,x_2}}{A_{i}}$$

where $\left\langle n_{i} \right\rangle_{x_1,x_2}$ 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; $A_{i}$ 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 $A_i$ (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 $A_i$ be a rectangular plot whose dimensions have the ratio $\sqrt{2}$, and let $A_{i}' = A_{i}/\sqrt{2}$ be the size of areas obtained from $A_i$ 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,

$$\Omega_{r} = \frac{2}{\alpha^2} - \frac{1}{\alpha^{2+1}}$$

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

$$\Omega_{r} = C r_{i}^{-\epsilon}$$

where $C = \Omega_{r}/r_{i}^\epsilon$.

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 $r^2$
tonic decrease in $\Omega$. Any distribution, however, including one with a nonmonotonic decrease in $\Omega$, can be described by the probabilities $a_i$, if $a_i$ is allowed to vary with $i$.

8. For a given $i$, $a_i$ increases with abundance across species that have the same $n_i$ (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.


15. Under the simplifying assumption that each species has no more than one individual in areas of size $A_c/N_0$, 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 by:

$$\alpha = \frac{a_i}{n_i}$$

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.

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. 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...
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 km²—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.stri.si.edu

References and Notes
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 $K(x) = cD$, 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

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

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

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

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