MAN AND THE BIOSPHERE SERIES

MAB

MAN AND THE BIOSPHERE SERIES

Series Editor J.N.R. Jeffers

VOLUME 20

FOREST BIODIVERSITY RESEARCH, MONITORING AND MODELING

Conceptual Background and Old World Case Studies

Edited by

F. Dallmeier

and

J.A. Comiskey

Smithsonian Institution
Washington DC, USA

PUBLISHED BY

UNESCO

PARIS

AND

The Parthenon Publishing Group

International Publishers in Science, Technology & Education

Published in 1998 by the United Nations Educational, Scientific and Cultural Organization,
7 Place de Fontenoy, 75700 Paris, France—UNESCO ISBN 92-3-103408-1
and
The Parthenon Publishing Group Inc.
One Blue Hill Plaza
PO Box 1564, Pearl River,
New York 10965, USA—ISBN 1-85070-963-7
and
The Parthenon Publishing Group Limited
Casterton Hall, Carnforth,
Lancs LA6 2LA, UK—ISBN 1-85070-963-7

British Library Cataloguing in Publication Data
Forest biodiversity research, monitoring and modeling : conceptual background and old world case studies. – (Man and the biosphere ; v. 20)
1. Biological diversity – Methodology 2. Forest ecology – Methodology
I. Dallmeier, Francisco II. Comiskey, James A.
577.3'072
ISBN 1-85070-963-7

Library of Congress Cataloging-in-Publication Data
Forest biodiversity research, monitoring and modeling : conceptual background and old world case studies / edited by Francisco Dallmeier and James A. Comiskey.
p. cm. — (Man and the biosphere series ; v. 20)
Includes bibliographical references and index.
ISBN 1-85070-963-7
1. Biological diversity—Research—Congresses. 2. Environmental monitoring—Congresses. 3. Forest ecology—Congresses.
I. Dallmeier, Francisco. II. Comiskey, James A. III. Series.
QH541.15.B56F67 1998
577.3'07'2—DC21 98-10706
CIP

Typeset by H & H Graphics, Blackburn, UK
Printed and bound by Butler and Tanner Ltd., Frome and London, UK

CHAPTER 14

ASSESSING FOREST DIVERSITY ON SMALL PLOTS: CALIBRATION USING SPECIES–INDIVIDUAL CURVES FROM 50-HA PLOTS

Richard Condit, Robin B. Foster, Stephen P. Hubbell, R. Sukumar, Egbert G. Leigh, N. Manokaran, Suzanne Loo de Lao, James V. LaFrankie and Peter S. Ashton

INTRODUCTION

One of the primary purposes of tree census plots in the tropics is to assess biodiversity with the intent of providing absolute and comparative estimates of species diversity. Understanding variation in diversity is basic to conservation biology. In particular, it is necessary for making sound decisions about protecting and managing forests. Diversity is also at the heart of theoretical ecology, especially in the diverse tropics. Theories about how species persist and coexist rely on comparative assessments of forest diversity (Ashton, 1977; Gentry, 1982, 1988; Wright, 1992).

Determining how much more diverse some forests are than others demands estimators of diversity. In working with trees, researchers are lucky – trees sit still and are easy to count. Tropical ecologists and foresters usually mark off an area of a certain size and identify every stem in the area above a certain size – but not always the same area and not always the same minimum size for stems (Gentry, 1982, 1988; Lott, 1987; Valencia *et al.*, 1994; Duivenvoorden, 1994; Phillips *et al.*, 1994). How can diversity be compared between two sites that are not the same size? If forest A has more species of large trees than forest B, can it also be concluded that forest A has more species of small stems? If forest A has more tree species in 1 ha than forest B has in 1 ha, does that also mean that forest A's diversity over wider areas is higher? These questions have yet to be evaluated.

The Center for Tropical Forest Science is a consortium of tropical biologists whose goal is to carry out thorough, long-term inventories of tropical forest trees at several key sites using a standardized approach (Condit, 1995). Among us – scientists with the Smithsonian Tropical Research Institute, the Indian Institute of Science, and the Forest Research Institute of Malaysia – we now have a unique opportunity to address questions about comparing forest diversity. We have fully mapped all stems ≥ 1 cm in diameter at breast height (dbh) in 50-ha plots at three sites that differ in flora and diversity. One is on Barro Colorado Island (BCI) in Panama, a tropical moist forest with a strong four-month dry

season. The second is in the Pasoh Forest Reserve in Peninsular Malaysia, also a moist forest but with no dry season (Kochummen *et al.*, 1990). The third is in the Mudumalai Game Sanctuary in Tamil Nadu, southern India, a dry, deciduous forest with a long, dry season (Sukumar *et al.*, 1992). The BCI forest has a canopy dominated by Leguminosae and Bombacaceae, the Pasoh forest by Dipterocarpaceae, and the Mudumalai forest by Verbenaceae, Lythraceae, and Combretaceae (Foster and Hubbell, 1990; Kochummen *et al.*, 1990; Sukumar *et al.*, 1992).

With large data sets, we can address questions about how well small samples predict diversity over large areas, while with data from three very different forests, we are in a position to suggest generalizations. In this paper, we review the more theoretical results on species accumulation curves presented in Condit *et al.* (1996b) from the perspective of how to interpret small forest inventories. We utilize the results to make several specific, quantitative recommendations about analyzing plot data.

MATERIALS AND METHODS

Study sites

The first 50-ha plot in tropical forest was established on Barro Colorado Island in the Panama Canal at the site of the Smithsonian Tropical Research Institute's field station. This semi-deciduous, moist forest has been the subject of detailed research for 70 years (Croat, 1978; Leigh *et al.*, 1982). The plot is in old-growth forest (> 500 years since human disturbance), except for 2 ha that were cleared less than 100 years ago (Condit *et al.*, 1992; Condit *et al.*, 1996a). The plot includes 230,000 individuals and 300 species ≥ 1 cm dbh.

The plot in Malaysia is in the 2,000-ha Pasoh Forest Reserve in the center of the Malay Peninsula, 140 km southeast of Kuala Lumpur. The research forest is operated by the Forest Research Institute of Malaysia and is connected to a larger production forest. Pasoh is old-growth dipterocarp forest, classified as red meranti-keruing by Wyatt-Smith (1987). Although structurally quite similar to BCI, Pasoh is evergreen and somewhat denser (335,000 stems ≥ 1 cm dbh in 50 ha). It is also much more diverse, with 817 species in 50 ha (Manokaran *et al.*, 1992).

The third plot is in the Mudumalai Sanctuary in the state of Tamil Nadu, about 250 km southwest of Bangalore, India. Although operated as a state game reserve, the Indian Institute of Science maintains a research headquarters next to the forest. It is a dry, teak forest (*Tectona grandis*), subjected to logging and clearing in the last two centuries. Structurally, it is quite different from the other two 50-ha plots; grasses dominate the understory, and there are just 26,000 woody stems in 50 ha. Large trees form a near-complete canopy, but most species are deciduous. Fires are common, and elephants play a major role as browsers on young trees and shrubs (Sukumar *et al.*, 1992). There are only 71 species in 50 ha, less than one-tenth the richness found at Pasoh.

Census methods

All free-standing, woody stems ≥ 1 cm dbh were measured, mapped, and identified in each plot (Manokaran *et al.*, 1990; Sukumar *et al.*, 1992; Condit *et al.*, 1993, 1995). BCI was censused from 1981 to 1983 and in 1985, 1990, and 1995 (Hubbell and Foster, 1983, 1986a, 1986b, 1990a, 1990b, 1992; Condit *et al.*, 1992). Here we use only 1990 data. Pasoh was fully censused from 1986 to 1989, then again in 1990 (Manokaran *et al.*, 1992); we use data from the first census here. Mudumalai was censused completely in 1987 and 1991, with incomplete censuses carried out each year in between and since (Sukumar *et al.*, 1992); we use the 1987 data here.

Species-accumulation curves

To generate species–area curves, the large plots were divided into non-overlapping square or rectangular quadrats, and the number of species present in each was calculated. The mean and standard deviation of species number in all square quadrats of any one area were tallied and used to construct a species–area curve; this was repeated for rectangular plots. Square quadrats with dimensions of 5, 10, 20, 25, 31.6, 40, 50, 100, 150, 200, 250, and 500 m were used as well as rectangular quadrats 1,000 × 1, 500 × 2, 500 × 20, 200 × 2, 400 × 4, 500 × 5, 1,000 × 10 m, and 1,000 × 500 (the whole plot; the long axis was parallel to the 1,000 m axis of the plot). Further details are given in Condit *et al.* (1996b).

Species–individual curves were constructed by first calculating the mean number of species and the mean number of stems in square quadrats of a given area (some individual trees had multiple stems, but in this paper the word 'stem' means one individual). The mean species count was plotted against the mean stem count across all quadrat sizes to form a species–individual curve. Condit *et al.* (1996b) give the entire data set of species and stem counts in all quadrats at all three plots.

We used three dbh classes in all analyses: ≥ 1 cm dbh, ≥ 10 cm dbh, and ≥ 30 cm dbh. These classes are obviously not independent of one another; however, open-ended dbh classes are standard in forest inventories. At both the BCI and Pasoh plot, there was such a vast excess of stems in small dbh classes that the three groupings actually were nearly independent of each other. In Condit *et al.* (1996b), we evaluated non-overlapping size classes (for example, 1 to 9.9 cm dbh) and found results similar to those presented here.

Diversity indices

Several standard diversity indices were assessed. These are listed in Table 14.1, along with formulae for calculating each. For one, Fisher's α, there is no explicit formula, and we used an implicit approach based on Newton's method; a

Table 14.1 Diversity indices evaluated for their usefulness in small tropical forest inventories. S equals species number, N stem number, ln the natural logarithm, f_i the relative abundance of species i, S_1 the number of species represented by one individual, and S_2 the number of species represented by two individuals. Stability refers to how constant the index is with changing sample size, and extrapolation refers to whether the index can be used to predict the number of species in larger samples. Sources: Berger and Parker, 1970 (1); Chao, 1984 (2); Colwell and Coddington, 1994 (3); Fisher *et al.*, 1943 (4); Kempton, 1979 (5); Pielou, 1966 and 1975 (6); and Simpson, 1949 (7)

Index	*Formula*	*Source*	*Stability*	*Extrapolation*
Shannon–Wiener	$\Sigma_i - (f_i \ln f_i)$	5, 6	good	no
Fisher's α	$S = \alpha \ln \left(1 + \frac{N}{\alpha}\right)$	4	good	yes
Dominance	maximum f_i	1	poor	no
Simpson's	$\Sigma_i f_i^2$	6, 7	poor	no
Chao	$S + \left(\frac{S_{12}}{2S_2}\right)$	2, 3	poor	yes
Jackknife (2nd order)	$S + \left\{\left(\frac{S_1(2N-3)}{N}\right)\left(\frac{S_2(N-2)^2}{N(N-1)}\right)\right\}$	2, 3	poor	yes
Richness	S	–	poor	sometimes

computer program illustrating the method is included as the appendix. Variation of each index with stem number was assessed using the method described for species–individual curves.

RESULTS AND DISCUSSION

Species–area *versus* species–individual curves

Species–area curves for the different dbh classes had different forms (Condit *et al.*, 1996b). On a log–log plot, the curves were not parallel, meaning that the ratio of species in one dbh class to species in another dbh class changed with the area (Figure 14.1). For example, at BCI in 20-m² quadrats (0.04 ha), there were nearly five times as many species with stems ≥ 1 cm dbh than there were with stems ≥ 10 cm (54 *versus* 12, mean species counts), whereas in 1 ha, there were fewer than twice as many (172 *versus* 91). This was a general pattern; species counts from different dbh classes differed greatly in small quadrats, but tended to converge in larger quadrats. Thus, we reject one approach for

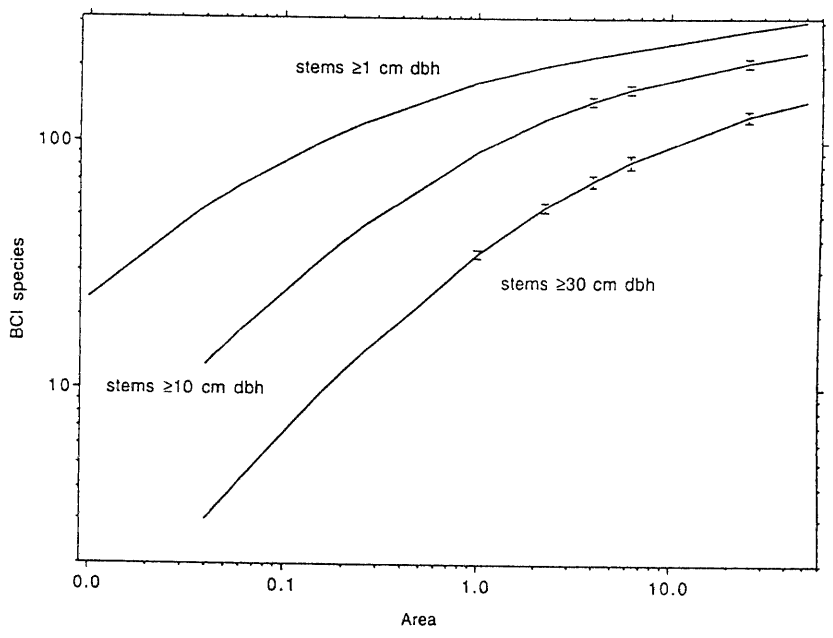

Figure 14.1 Species–area curve for three size classes of trees in the 50-ha BCI plot. Each point is enclosed by 95% confidence limits, but these are too small to be visible for most points (the last point to the right, representing the entire plot, has no confidence limits because there were no replicate quadrats). Curves from Pasoh and Mudumalai (not shown) were quite similar in form, although different in height

comparing diversity estimates based on two different dbh classes in plots of the same area – one cannot simply multiply by a constant.

Species–individual curves showed a different pattern. Instead of marked differences, essentially the same curve existed for all dbh classes at any one plot (Figure 14. 2). Species accumulated in a consistent manner as a function of individuals counted, not as a function of area. This simple but powerful result had not been appreciated before in forest inventories – the number of species in a fixed number of stems was almost independent of dbh. Moreover, the form of the species–individual curve was nearly identical at all three plots. This similarity has important practical implications for comparing and extrapolating species richness from small plots, which we evaluate below. Hubbell (1995, In press) provided a theoretical basis for the similarities.

Sampling implications of the species–individual curves

The similarity of species–individual curves for different dbh classes answered one of our primary questions; i.e. if forest A has more species of large trees than forest B, can it be concluded that forest A will have more species of small

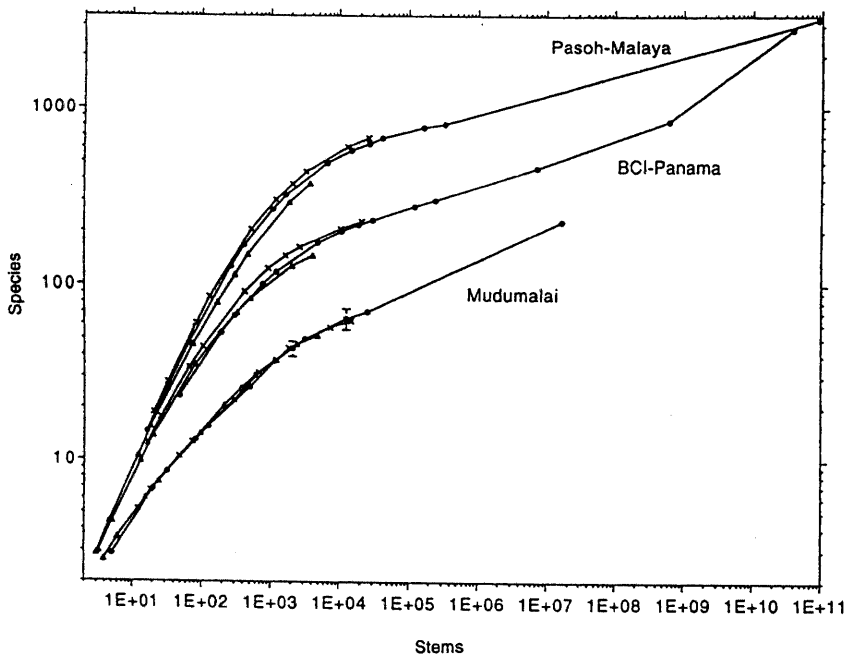

Figure 14.2 Species–individual curves for all three 50-ha plots, including extensions to species counts in larger regions. Closed circles indicate all stems ≥ 1 cm dbh, crosses all stems ≥ 10 cm dbh, and triangles all stems ≥ 30 cm dbh. The right-most point at Mudumalai represents the species count in the 32,000-ha Game Reserve, and the right-most point on the Pasoh curve represents the entire Malaysian portion of the Malay Peninsula (132,000 km²). The BCI-Panama curve has three points beyond 50 ha – the 1,500-ha island of Barro Colorado, the 120,000-ha Canal Area, and the nation of Panama (77,000 km²). Stem counts in these larger areas were estimated using the density of stems ≥ 1 cm dbh within 50 ha (note that the extended curves are connected to the curves for this size class). Confidence limits are as in Figure 14.1

stems? Yes, although slight differences among dbh classes must be considered. At BCI, samples of stems ≥10 cm dbh had about 15% more species than samples ≥ 1 cm dbh or ≥ 30 cm dbh (the latter two were essentially identical). At Pasoh, samples ≥10 cm had just 5% more species than samples ≥ 1 cm, but samples ≥ 30 cm had 20% to 25% fewer. At Mudumalai, there were essentially no differences among dbh classes (Condit *et al.*, 1996b). To be precise and quantitative, such differences should be considered. However, for assessing broad patterns marked by large differences, we conclude that one can compare diversity estimates from species counts using various dbh limits between 1 and 30 cm, as long as the stem number is held constant.

Another important result is that species–individual curves from the three different forests never crossed (Figure 14.2). This means that one can predict rank order in diversity from very small samples. Even with samples of just 50

stems, the Pasoh, BCI, and Mudumalai diversity curves were well differentiated. This answers another of the questions stated at the outset; i.e. if forest A has more tree species in 1 ha than forest B, can we conclude that its diversity over wider areas would also be higher? Yes, although as described shortly, this rule can break down in areas much larger than 50 ha.

The final important result illustrated in Figure 14.2 is that the curves from the three plots became nearly parallel on a log–log scale after sufficient stems were sampled. This means that the proportional difference in species richness among sites became constant. Pasoh had 2.7 times more species than BCI in all samples > 3,000 stems (excepting the furthest point to the right) and Mudumalai had 0.3 times as many species as BCI in samples > 300 stems (Condit *et al.*, 1996b). In smaller samples, the curves were not parallel, and in very small samples (< 30 stems), they barely differed. In samples of 500 stems, for example, Pasoh had 200 species, BCI 98, and Mudumalai 28. Thus, the richness of Pasoh relative to BCI would be underestimated (2.0 times instead of 2.7 times). In samples of 50 stems, the underestimate would be more severe; Pasoh had 41 species, BCI 29, and Mudumalai 11.

Diversity indices

The ideal measure of diversity would be independent of the stem number sampled. Then one could compare diversity of various sites using small samples, and samples could be of variable size. Species richness clearly depended on sample size (Figure 14.2), and, as explained above, the ratio of species richness among sites did not reach an asymptote until about 3,000 stems. Are any of the diversity indices more stable?

All indices were indeed less variable with sample size than species richness. However, none were independent of sample size (Table 14.2, Figure 14.3). All increased monotonically with stem count except Fisher's α, which increased until 2,000 to 12,000 stems and then decreased slightly. All indices showed similar patterns of change. There were rapid increases in small stem counts (< 500), then slower changes (Figure 14.3). The newer, non-parametric richness estimators were the most variable with sample size and greatly underestimated total species richness (Condit *et al.*, 1996b). The least variable indices were Fisher's α and the Shannon–Wiener index (Table 14.2); in samples of more than 1,000 stems, both varied by < 20% within a dbh class.

An index would also be useful if its dependence on sample size were similar across sites, because the between-site ratio would then be less dependent on sample size than the index itself. Both Fisher's α and the Shannon–Wiener index had this property (Figure 14.4; however, it was necessary to take the exponential of the Shannon–Wiener index – e^{index} – to generate a ratio similar to the richness ratio). This means that either α or the Shannon–Wiener index can be used to quantify diversity differences among sites using small stem samples as long as identical samples are used at all sites. Other indices also captured the

Table 14.2 Stability of diversity indices with sample size. Entries in the table give the ratio of the index calculated from samples of 10,000 stems to the index in 100 stems for the listed dbh classes at each plot. Estimates at exactly 100 and 10,000 stems were found from interpolation (Condit *et al.*, 1996b)

	Pasoh		*BCI*		*Mudumalai*	
Index	≥10 mm	≥100 mm	≥10 mm	≥100 mm	≥10 mm	≥100 mm
Shannon–Wiener (SW)	1.4	1.4	1.6	1.3	1.3	1.3
Fisher's α	1.4	1.2	1.5	1.2	1.9	1.7
Dominance	2.0	2.3	1.2	1.6	1.7	1.7
Simpson's	3.1	3.3	1.5	1.9	1.9	1.9
Chao	3.6	3.3	3.4	2.5	3.1	2.7
e^{SW}	3.9	3.9	3.5	3.5	3.8	3.7
Jackknife	4.4	4.4	3.6	2.5	3.2	3.0
Richness	7.6	8.2	5.1	4.8	4.4	4.1

relative diversity differences between sites (i.e. the approximately three-fold difference between Pasoh and BCI and between BCI and Mudumalai), but were considerably more variable in small stem samples (Figure 14.4).

There are interesting biological differences underlying the differences among the indices of diversity. For example, at BCI among stems ≥ 1 cm dbh, one shrub species, *Hybanthus prunifolius* (Violaceae), was numerically dominant to a remarkable extent. It included 15% of all stems in the plot (Condit *et al.*, 1996a). This shows up in Simpson's and the dominance indices, which are sensitive to the most abundant species. If we used these as our metrics, BCI would appear only slightly more diverse than Mudumalai. Another interesting contrast among forests is the relatively low diversity of trees in the large dbh class (≥ 30 cm) at Pasoh. This showed up slightly in species richness and in diversity indices, but was captured most strongly by Fisher's α in smaller quadrats (Figures 14.3 and 14.4). Large tree species were not particularly diverse at Pasoh and seemed to be more strongly aggregated than at BCI. This can be traced to the dominance of the Dipterocarpaceae among larger stems at Pasoh, many of which had strongly clustered distributions (Manokaran *et al.*, 1992).

Extrapolating beyond 50 ha

We used species counts from regional floras to extend the species–individual curves to much larger areas. There are 467 tree and shrub species on the entire island of Barro Colorado, 855 in the Canal Area and 2,870 in Panama; there are 229 species of trees and shrubs in the Mudumalai sanctuary; and there are 3,197 species in the Malay Peninsula (Condit *et al.*, 1996b). In all cases, we assumed

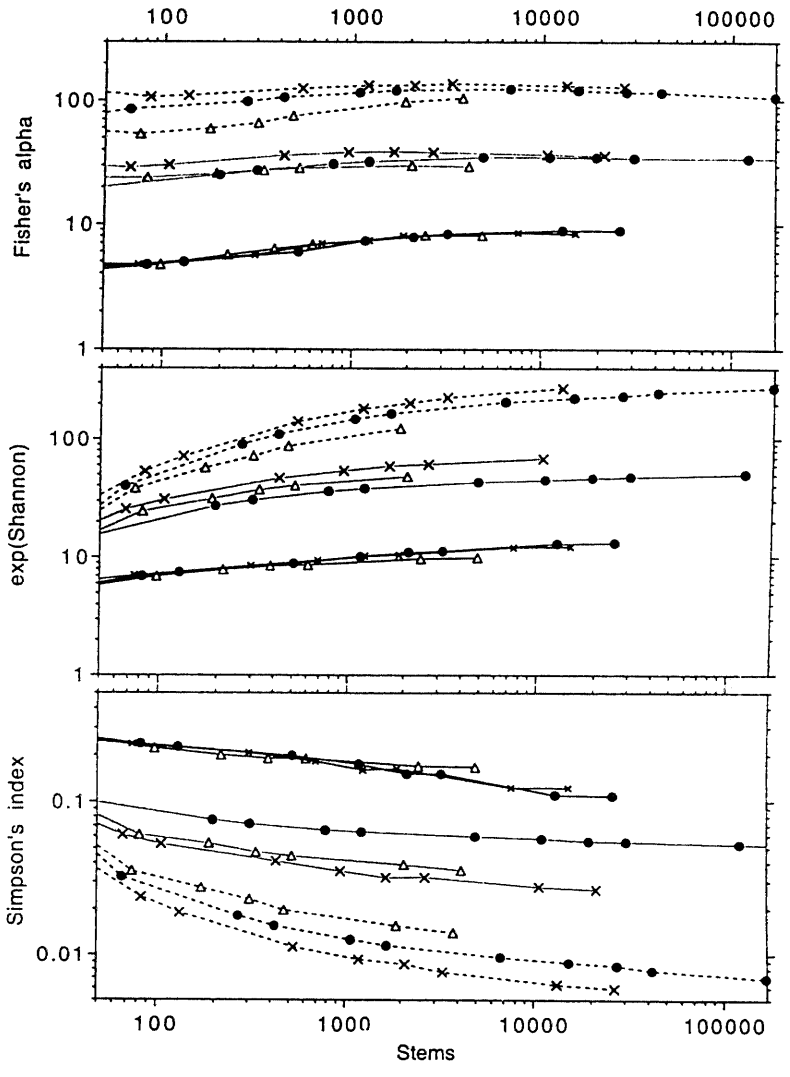

Figure 14.3 Variation in three diversity indices with sample size, plotted on a log–log scale. The exponential of the Shannon–Wiener index – e^{index} – is plotted, not the index itself, because the former gave a better quantitative diversity comparison. Simpson's index is shown as an example of the poorly performing indices; others were even more variable (Table 14.2)

these tallies of 'trees and shrubs' included all species ≥ 1 cm dbh. The counts were plotted on a log–log scale along with the data from the 50-ha plots (Figure 14.2), and the log–log slope (the z value) was calculated along the length of each curve (Condit *et al.*, 1996b). Within 50 ha, all three plots had declining z values until an asymptote of about 0.10 was reached at 10,000 stems at

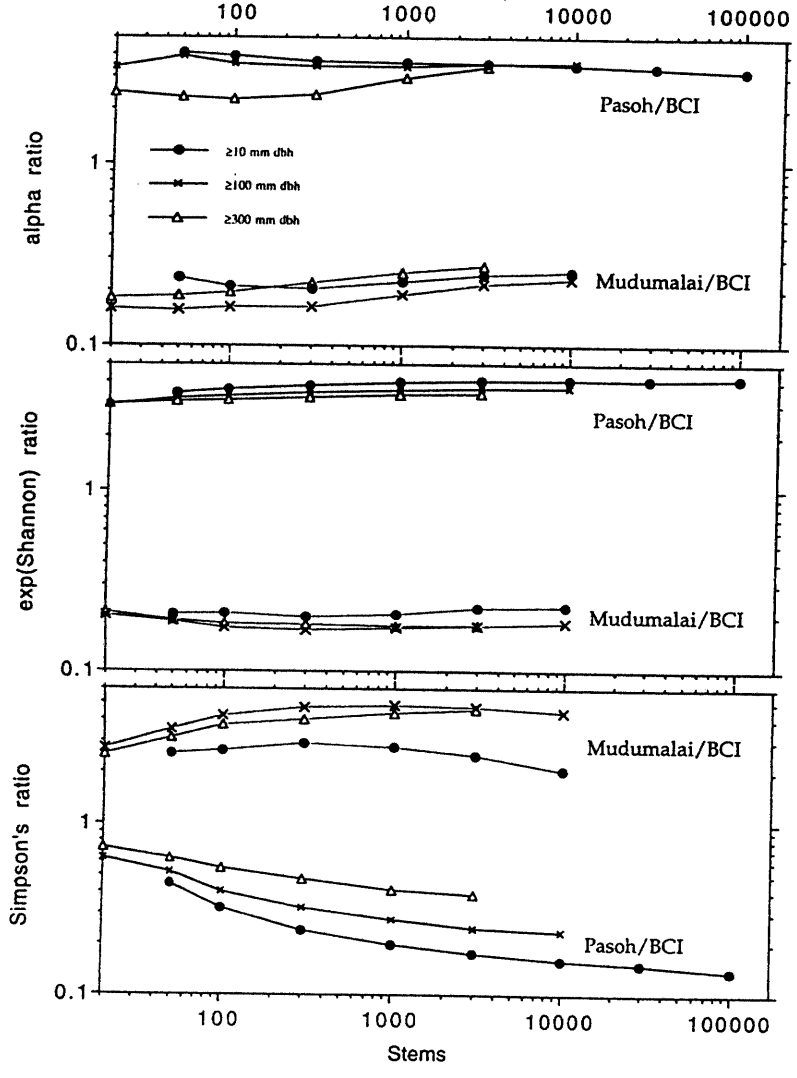

Figure 14.4 The ratio of diversity between sites as estimated by three different diversity indices. To calculate an index at, say, N = exactly 100 stems, linear interpolation was used from a graph of the index versus the log(N). The ratio between two plots of the index in question could then be calculated at exactly 20, 50, 100, 300, etc. stems. The ratio for Mudumalai/Pasoh is not shown, but could be calculated by dividing the other two ratios

Mudumalai, 20,000 at BCI, and 40,000 at Pasoh (Condit *et al.*, 1996b). The points for all of BCI, for the Canal Area, and for the Malay Peninsula fell close to these asymptotic lines, maintaining z values between 0.10 and 0.13 (Figure 14.2). The points for Panama and Mudumalai, however, required steeper slopes, 0.29 and 0.17 respectively; i.e. the species–individual curves had to turn upward

to meet these points (Figure 14.2). We interpreted the steeper slopes at the latter two sites as a reflection of the much greater climatic variability in Panama and Mudumalai, compared to the relatively uniform climate found within the Malay Peninsula (Condit *et al.*, 1996b).

The similarity of log–log slopes at all three sites suggested a way to extrapolate species richness from small plots to large regions. One can simply assume a slope of 0.10 to 0.13, which means that a 10-fold increase in area would lead to an additional 26% to 35% more species. Unfortunately, this estimate can only be applied with an initial species count from more than 40,000 stems, and it should only be extended to regions of fairly uniform climate or habitat (Condit *et al.*, 1996b).

Fisher's α provides an alternative for extrapolating because it is fairly stable with sample size and generates an estimate of species number given a stem count (from the equation in Table 14.1). (The Shannon–Wiener index, which was also fairly stable with respect to sample size, does not provide an equation relating total species number to stem number, so it cannot be used to extrapolate.) Since α was nearly constant for $N > 2{,}000$ stems, all that was needed was an estimate for α from 2,000 stems to estimate species number in any larger sample. Again, extrapolations should not go beyond regions of fairly uniform climate or habitat, as illustrated in Panama; α for the 50-ha plot was 34.1, for all BCI it was 38.4, for the Canal Area it was 52.5, and for all Panama it was 148.3.

To compare the approach based on Fisher's α with methods of extrapolation using z values, a simple mathematical result is useful: if α is constant, z declines as

$$\frac{1}{\ln N}\left(\text{or}\,\frac{1}{S}\right)$$

which is a slow decline when N is large. This means, of course, that if z is constant, α increases slowly. These are asymptotic results, applying when $N \gg \alpha$, or generally when $N > 1{,}000$.

With fewer than $N = 2{,}000$ stems, α varied rather sharply at all plots, but in a consistent way, suggesting that a correction factor could be applied to small samples and allow reasonably accurate extrapolations. We developed such a correction by modeling the relationship between Fisher's α and N with a log–log regression. This gives a function of the form:

$$\alpha = xN^y \quad \text{(Equation 1)}$$

where x and y are the regression parameters. For $50 \le N \le 2{,}000$, the regressions were very strong ($r^2 \ge 0.96$ for nine regressions calculated for each dbh class at all plots). With a single plot of N stems, one could fit this curve by dividing the plot into smaller quadrats of various sizes and calculating α in each. A log–log regression between the mean α and the mean N for each quadrat area could then be used to calculate x and y, and thus to extrapolate α to any larger N < 2,000.

As an alternative, we evaluated the relationship between the parameters x and y for the three 50-ha plots and three dbh classes. In as much as x and y vary together, it is possible to calculate them using only one data point without subsampling. We calculated a semi-log regression to describe their relationship: $y = -0.023 \ln x + 0.197$, with $r^2 = 0.53$. Combining this with Equation 1 generates two formulae:

$$y = \frac{8.45 - \ln \alpha}{43.0 - \ln N} \quad \text{(Equation 2)}$$

and

$$\ln x = 8.45 - 43.0y \quad \text{(Equation 3)}$$

We illustrate how the technique can correct for differences in sample size using published diversity inventories (Table 14.3). For example, a site at Boca de Uchire in Venezuela had only 53 tree species, substantially fewer than at San Carlos in Venezuela or Kade in Ghana. However, the Boca de Uchire inventory contained only 184 stems compared to > 500 at the other two sites. Which is more diverse? Fisher's α from the three inventories was similar, and adjusting α for sample size indicated that Boca de Uchire was more diverse than San Carlos or Kade. Therefore, we predicted that if samples of 1,000 stems were completed at the three sites, Boca de Uchire would have the most species.

As an illustration of the extrapolation technique, consider the Yanamono site in Peru. Using the x and y parameters from Table 14.3 with Equation 1, we would predict $\alpha = 244.7$ in $N = 2{,}000$ stems. Assuming a constant α in higher stem samples, we would estimate 1,249 species in 40,000 stems (80 ha), 1,696 species in 250,000 stems (500 ha), and 3,386 species in 2.5×10^8 stems (50,000 ha). Alternatively, if we assume $z = 0.10$ after 40,000 stems, there would be 1,500 species in 500 ha and 2,993 species in 50,000 ha.

These techniques are strictly empirical and must be used with caution. The qualitative relationship between α and N is probably a reasonable generalization, but the precise relationship between the parameters x and y may not be. Equations 2 and 3 show an inverse relationship – diverse sites with high x values are predicted to have low y values (Table 14.3). This means that α increases slowly with N at diverse sites, but more rapidly at species-poor sites, α conclusion based on just three forests. Not all species-poor sites are likely to have such high values of y. We recommend use of Equations 2 and 3 only at moderate or diverse tropical sites. Data from more sites is necessary to further evaluate how x and y are related. Better yet, theoretical understanding of how species accumulate may provide improved methods of extrapolating with α, z, or other diversity indices (Hubbell, In press).

Sampling error

The predictions above do not consider sampling error, but the 50-ha plots can provide estimates of sampling variance (P. Hall, Chapter 4). Figure 14.5 shows the coefficient of variation (CV) for species richness and the two recommended diversity indices as a function of stem sample. With 500 stems, typical of 1-ha plots of trees ≥ 10 cm dbh (Table 14.3), the CV for richness was 6% to 18% at the three plots, and the CV for α was 10% to 22%. In the absence of multiple plots at a single site, we suggest that a CV of 15% in 500 stems should be assumed for α, so that 95% confidence limits would extend > 30% above and below the estimate. This means, for example, that the Yanamono site in Peru is not significantly more diverse than the Manaus site in Brazil. However, three replicate 1-ha plots at each site would suffice to prove the difference, providing that the means hold.

Fisher's α had an extremely high CV in samples of fewer than 100 stems because of its extreme sensitivity to S when $S \geq N - 3$. Samples with < 100 stems in tropical forests should probably be avoided, and samples where $S \geq N - 3$ should definitely be expanded. This extreme sensitivity in small samples is a necessary property of a good diversity index, since highly diverse forests will seldom have more species than non-diverse forests in very small samples.

Plots of different shapes

At BCI and Pasoh, in samples >100 stems, plots 25 times longer than wide had 5% more species than square plots with the same number of stems and 12% to 13% more species than square plots in India (Condit *et al.*, 1996b). For plots 100 or 1,000 times longer than wide, differences were somewhat greater, and random collections of stems from throughout 50 ha had the most species (20% to 30% more than square quadrats with the same number of stems). With < 100 stems, all these differences declined. Thus, there is a consistent relationship between shape and diversity – the narrower the rectangle, the more the species, and rectangles 1,000 times longer than wide almost matched random collections (Condit *et al.*, 1996b). However, for most reasonable sampling regimes – Gentry (1982) used 25 × 2-m transects, the narrowest we know of – differences among sampling approaches were rather slight.

CONCLUSIONS AND RECOMMENDATIONS

Perhaps our most important finding was that there were near-constant differences in species richness among the forests. It required 3,000 stems to reach these asymptotic differences, but from there to > 250,000 stems, the richness ratio between two sites remained the same. If this were not true, comparative studies of species richness would be difficult. What if the Pasoh forest had more species

Table 14.3 Estimates of Fisher's α from published inventories, with corrections for sample size. The first column labelled α gives the estimate based on the published number of stems and species. *Y* and *X* are the parameters in Equation 1 relating α to stem number (calculated from Equations 2 and 3); α (1,000) is the estimate of α in 1,000 stems, and *S*(1,000) the number of species in 1,000 stems (found from the equation in Table 14.1). Some data taken from Gentry (1982) and Phillips *et al.* (1994) are from other sources; the original references are not given here. For BCI, Pasoh, and Mudumalai, the mean and standard deviation from each of the 50 1-ha² plots is given

Country site	*Reference*	*Area* (ha)	*Minimum dbh* (cm)	*Stems*	*Species*	α	*y*	*x*	α (1000)	*S* (1000)
Peru Yanamono	Phillips *et al.* (1994)	1.0	10	574	283	221.1	0.083	130.23	231.6	387
Ecuador Cuyabeno	Valencia *et al.* (1994)	1.0	5	1,561	473	230.8	0.084	124.00	222.3	379
Ecuador Cuyabeno	Valencia *et al.* (1994)	1.0	10	693	307	211.0	0.085	120.96	217.7	375
Brazil Manaus	Gentry (1982)	1.0	15	350	179	146.9	0.093	85.05	162.0	319
Malaysia Pasoh	This study	1.0	10	529 ± 38	206 ± 17	125.2 ± 18.6	0.099 ± 0.004	67.60 ± 11.8	133.3 ± 19.5	284 ± 25
Peru Tambopata	Phillips *et al.* (1994)	1.0	10	546	173	87.3	0 .109	44.03	93.2	230
Malaysia Sepilok	Phillips *et al.* (1994)	1.0	10	435	117	52.5	0.122	25.07	58.1	169
Panama BCI	This study	1.0	10	425 ± 44	91 ± 7	36.0 ± 4.6	0.132 ± 0004	16.30 ± 2.50	40.3 ± 5.2	131 ± 12
Brazil Belem	Gentry (1982)	2.0	10	423	87	33.2	0.134	14.76	37.3	124
Panama Curundu	Gentry (1982)	0.1	2.5	225	64	29.8	0.135	14.39	36.5	122
Venezuela Boca de Uchire	Gentry (1982)	0.1	2.5	184	53	24.9	0.139	12.10	31.5	110
Ghana Kade	Phillips *et al.* (1994)	1.0	10	562	86	28.3	0.139	11.71	30.7	108
Costa Rica La Selva	Gentry (1982)	0.8	10	328	65	24.3	0.141	10.71	28.5	102
Venezuela San Carlos	Phillips *et al.* (1994)	1.0	10	744	83	23.9	0.145	9.17	25.7	93
Puerto Rico Luquillo	Gentry (1982)	2.0	4	1,571	80	17.8	0.156	5.63	16.6	68
Brazil Mato Grosso	Gentry (1982)	1.0	15	274	36	11.1	0.162	4.47	13.7	59
USA Missouri	Gentry (1982)	0.1	2.5	165	23	7.3	0.171	3.04	9.9	46
India Mudumalai	This study	1.0	10	301 ± 67	22 ± 3	5.6 ± 1.2	0.181 ± 0.006	2.30 ± 0.50	7.0 ± 1.6	35 ± 6

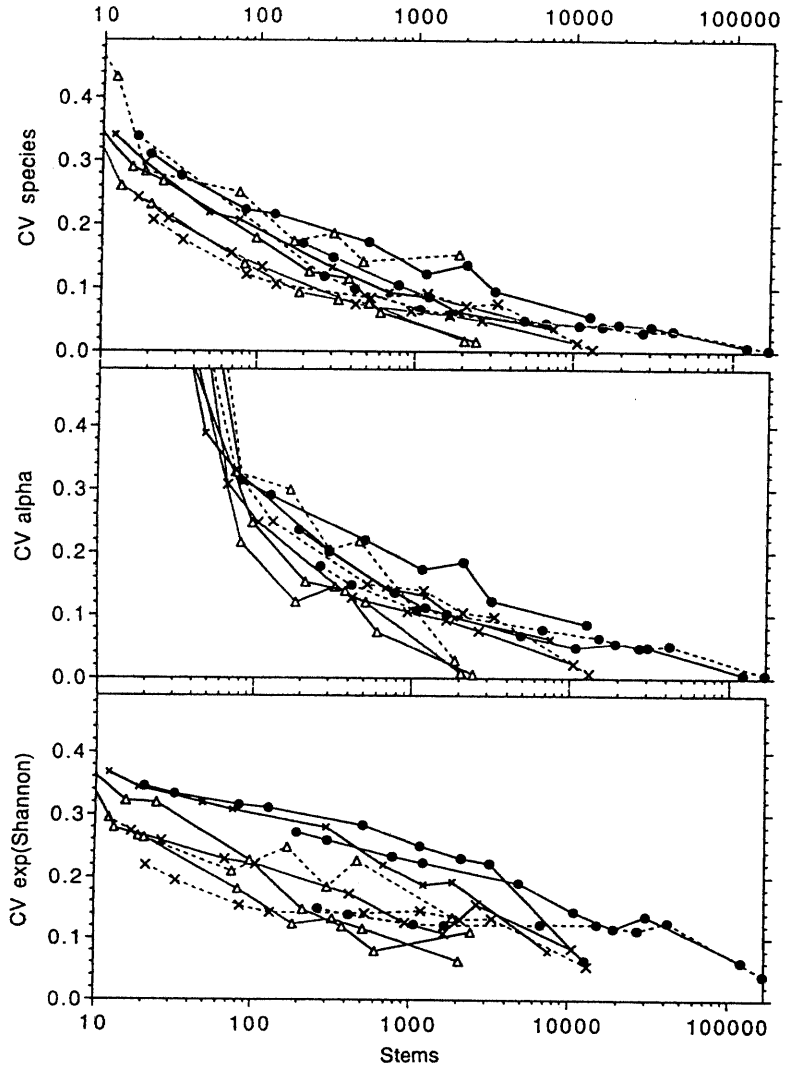

Figure 14.5 The coefficient of variation (standard deviation divided by mean) among square quadrats within a 50-ha plot for species richness and two diversity indices, as a function of the mean number of stems in the quadrat (created exactly as the species–individual curves in Figure 14.2). There are nine curves on each graph, representing three dbh classes from three 50-ha plots. The purposes of the graph are to illustrate the broad similarity among the plots and the dbh classes and give rough estimates of the magnitude of the CV, so individual curves are not identified

in 50 ha, but fewer in 1 ha, than BCI? At large enough scales, the pattern did break down; i.e. the entire nation of Panama 'caught up' with Malaya in terms of richness. Fortunately, local tree diversity can be compared across sites with one simple number.

Unfortunately, most diversity inventories in tropical forests are much smaller than 3,000 stems, and these will tend to underestimate differences among species-rich sites. We conclude with a series of recommendations for analyzing diversity in small inventories. First, to make the most rigorous quantitative comparisons of diversity at different sites, we suggest these rules:

(1) Compare samples with approximately equal numbers of stems, but ignore area.

(2) Use the same sampling protocol across sites; i.e. the same quadrat shape.

(3) Compare samples using the same dbh limit.

(4) With samples < 3,000 stems, do not use species richness as a diversity metric; instead, use Fisher's α or the Shannon–Wiener index.

(5) Be aware of sampling error because confidence limits on diversity in samples < 1,000 stems extend about 30% below and above the estimate.

We have suggested that rules 2 and 3 can be disregarded with only slight loss of accuracy, based on data from three forest sites. Rectangular quadrats were only slightly more diverse than squares, and diversity depended rather weakly on dbh class. Confidence limits on diversity estimates would generally be much higher than these differences, so it seems reasonable to compare samples from different dbh classes and from quadrats of different shapes.

It is more important to pay attention to differences in stem number. Fisher's α was almost independent of stem number beyond about 2,000 stems, but in smaller samples, variation in stem number must be considered. For this reason, we propose a series of specific rules about stem number for diversity estimates of tropical trees:

(1) Never use samples < 50 stems, and in very diverse forests, use 100 stems or more. These are the absolute minimums. Larger samples are clearly desirable.

(2) If each of two samples has > 2,000 stems, Fisher's α can be used for a comparison, even if the samples differ substantially in size.

(3) For samples between 50 and 2,000 stems, one option for drawing comparisons is to subsample stems to a common number at all sites.

(4) Alternatively, with samples between 50 and 2,000 stems, the correction for Fisher's α described above (based on a locally derived version of Equation 1 by subsampling) can be applied. Equations 2 and 3 can be used if subsampling is impossible. Recall, however, that confidence limits on α are already high, and Equations 1 through 3 would only add to the variance.

The advantage of approach 4 is that it allows a comparison among sites and extrapolation. One can either assume a constant α for samples > 2,000 stems or

a log–log slope of 0.10 to 0.13 > 40,000 stems. These approaches provide a range of estimates about the number of species that would be found in a 50-ha plot or in, say, an entire national park. Fairly large confidence limits should be attached to these estimates.

Because of the wide confidence limits on 1-ha surveys, we suggest that four plots with about 500 stems each are necessary to characterize the diversity at one site. This recommendation does not apply to other features of a forest. Much larger plots are necessary for demographic studies of individual species because most species are quite rare (Condit, 1995; Condit *et al.*, 1995). Hall *et al.* (chapter 4) deals with the issue of plot size for other parameters such as basal area, stem density, and mortality.

Uncovering the rules listed above that demonstrate the utility of small plots for diversity estimates required data from three 50-ha plots. One of our main goals within a network of large plots is to develop techniques for maximizing the utility of smaller plots, and the analysis of diversity presented here is an example of that approach.

ACKNOWLEDGMENTS

We thank the Forest Research Institute of Malaysia for sponsoring and supporting the work at Pasoh, the Smithsonian Tropical Research Institute in Panama for support at BCI, and the Indian Institute of Science for supporting the Mudumalai project. We thank I. Rubinoff and E. Losos for organizing the network of large plots. We also thank all of the workers who contributed to the three censuses, more than 100 people from 12 countries, especially E. S. Quah, S. Appanah, R. Pérez, and H. S. Suresh. The Pasoh project was also supported by the National Science Foundation (USA), the Rockefeller Foundation, and the John Merck Fund, and the BCI project by the National Science Foundation, the Smithsonian Scholarly Studies Program, the John D. and Catherine T. MacArthur Foundation, the World Wildlife Fund, the Earthwatch Center for Field Studies, the Geraldine R. Dodge Foundation, and the W. Alton Jones Foundation. This article is a scientific contribution from the Center for Tropical Forest Science, which is supported by the John D. and Catherine T. MacArthur Foundation.

REFERENCES

Ashton, P. (1977). A contribution of rainforest research to evolutionary theory. *Annals of the Missouri Botanical Garden*, **64**: 694–705

Berger, W.H. and Parker, F.L. (1970). Diversity of planktonic Foraminifera in deep-sea sediments. *Science*, **168**: 1345–7

Chao, A. (1984). Non-parametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, **11**: 265–70

Colwell, R.K. and Coddington, J.A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London* Series B, **345**: 101–18

Condit, R. (1995). Research in large, long-term tropical forest plots. *Trends in Ecology and Evolution*, **10**: 18–22

Condit, R., Hubbell, S.P. and Foster, R.B. (1992). Stability and change of a neotropical moist forest over a decade. *BioScience*, **42**: 822–8

Condit, R., Hubbell, S.P. and Foster, R.B. (1993). Identifying fast-growing native trees from the neotropics using data from a large, permanent census plot. *Forest Ecology and Management*, **62**: 123–43

Condit, R., Hubbell, S.P. and Foster, R.B. (1995). Mortality rates of 205 neotropical tree species and the responses to a severe drought. *Ecological Monographs*, **65**: 419–39

Condit, R., Hubbell, S.P. and Foster, R.B. (1996a). Changes in a tropical forest with a shifting climate: results from a 50-ha permanent census plot in Panama. *Journal of Tropical Ecology*, **12**: 231–56

Condit, R., Hibbell, S.P., LaFrankie, J.V., Sukumar, R., Manokaran, N., Foster, R.B. and Ashton, P.S. (1996b). Species–Area and Species–Individual Relationships for Tropical Trees: A Comparison of Three 50 Ha Plots. *Journal of Ecology*, **84**: 549–62

Croat, T.R. (1978). *Flora of Barro Colorado Island*. Stanford University Press, Stanford, CA

Duivenvoorden, J.F. (1994). Vascular plant species counts in the rain forests of the middle Caquetá area, Colombian Amazonia. *Biodiversity and Conservation*, **3**: 685–715

Fisher, R.A., Corbet, A.S. and Williams, C.B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Ecology*, **12**: 42–58

Foster, R.B. and Hubbell, S.P. (1990). The floristic composition of a 50-ha plot on Barro Colorado Island. In Gentry, A. (ed.) *Four Neotropical Forests*. Yale University Press, New Haven, CT, pp. 85–98

Gentry, A.H. (1982). Patterns of neotropical plant species diversity. *Evolutionary Biology*, **15**: 1–84

Gentry, A.H. (1988). Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Annals of the Missouri Botanical Gardens*, **75**: 1–34

Hubbell, S.P. (1995). Towards a theory of biodiversity and biogeography on continuous landscapes. In Carmichael, G.R., Folk, G.E. and Schnoor, J.L. (eds.) *Preparing for Global Change: A Midwestern Perspective*. SPB Academic Publishing bv, Amsterdam, pp. 171–99

Hubbell, S.P. In press. *A Theory of Biogeography and Relative Species Abundance*. Princeton Monographs in Population Biology. Princeton University Press, Princeton, NJ

Hubbell, S.P. and Foster, R.B. (1983). Diversity of canopy trees in a neotropical forest and implications for conservation. In Sutton, S.L., Whitmore, T.C. and Chadwick, A.C. (eds.) *Tropical Rain Forest: Ecology and Management*. Blackwell Scientific Publications, Oxford, pp. 25–41

Hubbell, S.P. and Foster, R.B. (1986a). Biology, chance, and the history and structure of tropical rain forest tree communities. In Diamond, J. and Case, T.J. (eds.) *Community Ecology*. Harper and Row, New York, pp. 314–29

Hubbell, S.P. and Foster, R.B. (1986b). Commonness and rarity in a neotropical forest: implications for tropical tree conservation. In Soulé, M. (ed.) *Conservation Biology:*

The Science of Scarcity and Diversity. Sinauer Associates, Inc., Sunderland, MA, pp. 205–31

Hubbell, S.P. and Foster, R.B. (1990a). Structure, dynamics, and equilibrium status of old-growth forest on Barro Colorado Island. In Gentry, A. (ed.) *Four Neotropical Rain Forests*. Yale University Press, New Haven, CT, pp. 522–41

Hubbell, S.P. and Foster, R.B. (1990b). The fate of juvenile trees in a neotropical forest: implications for the natural maintenance of tropical tree diversity. In Hadley, M. and Bawa, K.S. (eds.) *Reproductive Ecology of Tropical Forest Plants*. UNESCO and Parthenon Publishing, New York, pp. 317–41

Hubbell, S.P. and Foster, R.B. (1992). Short-term population dynamics of a neotropical forest: why ecological research matters to tropical conservation and management. *Oikos*, **63**: 48–61

Kempton, R.A. (1979). The structure of species abundance and measurement of diversity. *Biometrics*, **35**: 307–21

Kochummen, K.M., LaFrankie, J.V. and Manokaran, N. (1990). Floristic composition of Pasoh Forest Reserve, a lowland rain forest in Peninsular Malaysia. *Journal of Tropical Forest Science*, **3**: 1–13

Leigh, E.G., Jr., Rand, S.A. and Windsor, D.M. (eds.) (1982). *The Ecology of a Tropical Forest: Seasonal Rhythms and Long-Term Changes*. Smithsonian Institution Press, Washington, DC

Lott, E.J. (1987). Floristic diversity and structure of upland and arroyo forests of coastal Jalisco. *Biotropica*, **19**: 228–35

Manokaran, N., LaFrankie, J.V., Kochummen, K.M., Quah, E.S., Klahn, J., Ashton, P.S. and Hubbell, S.P. (1990). *Methodology for the 50-hectare research plot at Pasoh Forest Reserve*. Research Pamphlet No. 102. Forest Research Institute of Malaysia, Kepong

Manokaran, N., LaFrankie, J.V., Kochummen, K.M., Quah, E.S., Klahn, J., Ashton, P.S. and Hubbell, S.P. (1992). *Stand table and distribution of species in the 50-ha research plot at Pasoh Forest Reserve*. Research Data No. 1. Forest Research Institute of Malaysia, Kepong

Phillips, O.L., Hall, P., Gentry, A.H., Sawyer, S.A. and Vásquez, R. (1994). Dynamics and species richness of tropical rain forests. *Proceedings of the National Academy of Science*, **91**: 2805–9

Pielou, E.C. (1966). The measurement of diversity in different types of biological collections. *Journal of Theoretical Biology*, **13**: 131–44

Pielou, E.C. (1975). *Ecological Diversity*. John Wiley and Sons, New York

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

Sukumar, R., Dattaraja, H.S., Suresh, H.S., Radhakrishnan, J., Vasudeva, R., Nirmala, S. and Joshi, N.V. (1992). Long-term monitoring of vegetation in a tropical deciduous forest in Mudumalai, southern India. *Current Science*, **62**: 608–16

Valencia, R., Balslev, H. and Paz y Mino C.G. (1994). High tree alpha-diversity in Amazonian Ecuador. *Biodiversity and Conservation*, **3**: 21–8

Wright, S.J. (1992). Seasonal drought, soil fertility and the species diversity of tropical forest plant communities. *Trends in Ecology and Evolution*, **7**: 260–3

Wyatt-Smith, J. (1987). *Manual of Malayan silviculture for inland forest: red meranti-keruing forest*. Research Pamphlet No. 101. Forest Research Institute of Malaysia, Kepong

APPENDIX

Following is a C program for estimating Fisher's α from a species count and a stem count using Newton's method. The *main* routine prompts the user to enter a stem number and a species number, both floating point numbers, then calls the subroutine *calcalpha*. *Calcalpha* starts with an estimate of $\alpha = 20$ (α is a in the subroutine) and repeatedly recalculates it using Newton's method. The function f is

$$\alpha \ln\left(1 + \frac{N}{\alpha}\right) - S$$

from Table 14.1, and f' is its derivative. After applying Newton's method, a must be reset if it becomes negative, since $a < 0$ can cause f to blow up (negative values for a will only occur during the first few iterations if the initial estimate for a is very far off). After *calcalpha*, *main* prints the stem and species counts and the estimate of α to the screen. Fisher's α is infinite if $N = S$, and the program returns a value of –999 in such a case. If negative or zero values are entered for N or S (which are impossible and could only be errors), the program returns $\alpha = 0$. Since $\alpha = 0$ is impossible, this could not be confused with any valid calculation.

The program should work on any C compiler, either Macintosh or DOS. It could easily be adapted to read data out of a file instead of from the screen and could also be adapted to BASIC, FORTRAN, or other computer languages.

```
#include 'stdio.h'
#include 'stdlib.h'
#include 'string.h'
#include 'math.h'
float calcalpha(float stems, float species);
void main(void)
 {
  int i, j, pairs;
  char sppstr[10], stemstr[10];
  float alpha, stems, species=1;
  while(stems>0)
   {
    printf('Number of stems (enter 0 or less to end program): ');
    gets(stemstr);
    stems=atof(stemstr);
    if(stems<=0) break;
    printf('Number of species: ');
    gets(sppstr);
    species=atof(sppstr);
    alpha=calcalpha(stems,species);
```

```
      printf('%8.4f\t%8.4f\t%8.4f\n', species, stems, alpha);
    }
}
float calcalpha(float n, float s)
 {
  int i;
  float a=20;
  float f(float a, float n, float s);
  float fprime(float a, float n, float s);
  if(n<=0 || s<=0) return 0;
  if(n==s) return -999;
  while( fabs(f(a,n,s))>1e-2 )
   {
    a-=f(a,n,s)/fprime(a,n,s);
    if(a<=0) a=1;
   }
 return a;
}
float f(float a, float n, float s)
 { return a*log(1+n/a) - s ; }
float fprime(float a, float n, float s)
 { return log(1+n/a) - n/(a+n) ; }
```