

## Asymptotic models of species–area curve for measuring diversity of dry tropical forest tree species

R. Sagar, A. S. Raghubanshi\* and J. S. Singh

Department of Botany, Banaras Hindu University, Varanasi 221 005, India

In a dry tropical forest, we examine the fitness and predictability of two non-asymptotic models (log-linear and power) of species–area curve, and the effect of sample location and scale on their regression-derived coefficients ( $c$ ,  $z$ ) for measuring tree diversity. Results indicate that, the log-linear model relatively better fits the data set, and yields better prediction of number of species on a small scale (i.e. predicted number of species for 3 ha using an equation based on 1 ha data). On the other hand, predictions from power function model for a larger area (i.e. predicted number of species for 15 ha using 1 ha and 3 ha equations) were closer to the observed values. The suitability of the model to fit the data was strongly influenced by the site and the scale of the plot size. The equations for the two models derived from data of small area (1 ha plot size) yielded inconsistent results, but those derived from a larger plot size (3 ha) consistently underestimated the number of species for 15 ha. The underestimation by power function model was lower compared to that by log-normal model for predicting the number of tree species. The study also shows that the coefficient  $z$  is site- as well as scale-dependent. The coefficient  $c$  can be used to predict  $\alpha$ -diversity, and the number of species per individual can adequately describe the coefficient  $z$ . The results support discrete community concept for the dry tropical forests along a disturbance gradient and indicate that higher the  $z$ , greater would be the impact of harvest of individuals on biodiversity.

TROPICAL forests cover only 7% of the earth's land surface, but harbour more than half of the world's species<sup>1</sup>,

---

\*For correspondence. (e-mail: asr@bhu.ac.in)

and are currently disappearing at an overall rate of 0.8 to 2% per year<sup>2</sup>. Biodiversity of these forests has attracted much attention in recent years. Both the magnitude and the urgency of the task of assessing global biodiversity require that we make the most of what we know through the use of estimation and extrapolation<sup>3</sup>.

Local species richness can be estimated by extrapolating species–area curves. The species–area relationships arise partly from an increase in habitat diversity with increasing area sampled<sup>4</sup>. These relationships are important in ecological study because they provide insight into community structure<sup>5</sup>, and the mathematical expressions of the models are used for predicting species richness at larger scale, and extinction rates caused by habitat destruction<sup>6</sup>. The species–area relationship is a fundamental component of conservation biology, and is often used to assess the long-term effects of habitat fragmentation on biodiversity<sup>7</sup>.

Extrapolation using different models for the species–area relationships can yield different values of species richness for a given area<sup>7,8</sup>. According to Colwell and Coddington<sup>3</sup>, different models may be more effective for different environments as the shape of a species–area curve depends upon the pattern of relative abundances of the sampled species. Although Soberon and Llorente<sup>8</sup> argue for the a priori choice of models for species accumulation curves, Colwell and Coddington<sup>3</sup> believe the best approach is to test all reasonable models as rigorously as possible against known standards (complete or nearly complete inventories) for a wide variety of taxa and localities to avoid summary judgement based on a single data set.

In this study we examine the predictability of two non-asymptotic models (log-linear and power) and the effect of sample location and spatial scale of a dry tropical forest on the regression-derived coefficients of the species–area curves fitted by power function model. Relationships between these coefficients and other diversity measures are also examined.

The study was conducted on five sites (viz. Hathinala, Khatabaran, Majhauili, Bhawani Katariya and Kota) (24°6'52"–24°26'16"N and 83°1'86"–83°9'60"E) in a dry tropical forest region of India in the year 1998–1999. The sites were selected on the basis of satellite imagery and field observations to represent the entire range of conditions in terms of canopy cover and disturbance regimes. The area experiences tropical monsoon climate with a mean annual rainfall of 821 mm, of which about 86% is received from the southwest monsoon during June–August. Physiographically, the area is characterized by hillocks, escarpments, east-west trending gorge-like valleys, flat basins and flat-topped ridges. The topography is relatively flat on Kota and Khatabaran sites, gentle at Bhawani Katariya and undulating at Hathinala and Majhauili sites. The altitude varies from 313 to 483 m asl. The soils are Ultisols, sandy loam in texture and reddish to dark grey in colour.

At each of the five sites (viz. Hathinala, Khatabaran, Majhauili, Bhawani Katariya and Kota; the sequence followed the visually assessed order of disturbance from least to highest), three one-hectare contiguous permanent plots were established. Each one-hectare plot was gridded into 100 subplots, each 10 m × 10 m in size. All individual trees of ≥ 30 cm circumference over bark at breast height (1.37 m) were enumerated by species.

Species–area relationships were analysed by plotting cumulative number of species as a function of plot size. We used Arrhenius equation ( $S = cA^z$ , where  $S$  is the number of species,  $A$  is the area, and  $c$  and  $z$  are regression-derived coefficients), since it is the most frequently used model<sup>5,9–11</sup>. The log-linear model ( $S = a + b \ln A$ , where  $S$  is the number of species,  $A$  the area, and  $a$  and  $b$  are regression-derived coefficients) was used for comparison. The coefficients were calculated for each of the 1 ha plots, and for data pooled for 2 ha and 3 ha plots on each site. Examining the effect of scale was important from the point of view of predictability and interpretation of the exercise<sup>12</sup>.

In the species–area curves, the order in which samples are added affects the shape of the curve<sup>3</sup>. Sampling error as well as real heterogeneity among the sampling units are the two important reasons for variation in the shape of the species–area curve. To remove these kinds of sampling errors, the sample order is recommended to be randomized<sup>3</sup>. Species–area curve data were generated by randomized pooling of quadrat data using species–area curve option of PC-ORD<sup>13</sup>. Random subsampling for sequential accumulation was repeated 100 times. For obtaining the coefficients, power and log-linear equations were fitted using SPSS package<sup>14</sup>.  $\alpha$ -diversity (Shannon–Wiener index) was also obtained with the PC-ORD, using the base 2.718 (i.e. natural log). The site and scale differences were analysed by ANOVA using the SPSS package. One-ha values for  $c$  and  $z$  (3 plots per site) were used as replicates for examining site differences, and for the scale differences calculations from pooled data (1 ha, 2 ha, 3 ha) were used by taking the sites as replicates. The total number of species at each site was divided by the total number of individuals at the same site to represent the number of species per individual. The values of species per individual were regressed with coefficient  $z$  of the species–area curve parameter using the SPSS package<sup>14</sup>.

The total number of species, density and  $\alpha$ -diversity are presented in Table 1. Sites markedly differed in the presence and abundance of species.

Species–area relationship up to 1 ha has been shown to follow the log-linear model which may be the result of a random placement of species throughout the area<sup>15</sup>. However, if the area involved ranges from 0.01 to 10<sup>7</sup> km<sup>2</sup>, the species–area relationship best fits the power function model due to the addition of new ecological conditions and new habitats<sup>16</sup>. In our study, the log-linear

**Table 1.** Number of individuals for the main tree species at five sites of dry tropical forest. Only species with  $\geq 30$  individuals in any one site are included. Complete data are available from authors upon request. Nomenclature follows Verma *et al.*<sup>23</sup>

Species	Hathinala	Khatabaran	Majhauri	Bhawani Katariya	Kota
<i>Acacia catechu</i>	240	72	105	81	16
<i>Anogeissus latifolia</i>	46	63	27	15	12
<i>Briedelia retusa</i>	80	–	2	–	–
<i>Buchanania lanzan</i>	79	73	51	45	–
<i>Cassia fistula</i>	1	5	31	–	–
<i>Diospyros melanoxylon</i>	54	59	50	43	1
<i>Emblica officinalis</i>	43	19	36	1	–
<i>Hardwickia binata</i>	105	–	–	207	48
<i>Holarrhena antidysenterica</i>	3	125	28	5	–
<i>Lagerstroemia parviflora</i>	43	37	168	26	–
<i>Lannea coromandelica</i>	113	45	71	3	1
<i>Miliusa tomentosa</i>	9	–	38	9	–
<i>Shorea robusta</i>	188	45	317	105	–
<i>Tectona grandis</i>	–	165	–	–	–
<i>Terminalia tomentosa</i>	89	35	216	31	–
<i>Ziziphus glaberrima</i>	47	1	1	1	–
Other species	117	94	46	74	27
Total no. of individuals	1257	838	1187	646	105
Total no. of species	31	30	23	22	7
Shannon–Wiener index	2.629	2.617	2.285	2.235	1.398

model and the power function model both gave highly significant fits, but the log-linear model exhibited lower standard error of estimate and higher  $R^2$  values (Table 2). The standard error of estimate quantifies the spread of the real data points around the fitted regression curve. Interestingly, the standard error of estimate for both the models was lowest for the most species-poor site (Table 2), indicating the effect of the quantity and species composition on the fitness of the species–area curve models.

Utility of a model generally depends on its ability to predict the number of species for a given area. In a study by Palmer<sup>7</sup>, the power function model produced extreme overestimates of species richness, while the log-linear model performed much better. In this study, the 1 ha power equation overestimated the number of species for 3 ha area on all sites, except for the Kota site; but three out of five 1 ha equations underestimated the number of species for 15 ha area (Table 3). For all the five sites, 3 ha power equations underestimated the number of species for 15 ha area; the underestimation was relatively small by the Khatabaran and Hathinala power equations, and extremely large by the Kota power equation. In the case of the log-linear model, 1 ha equation overestimated the number of species for 3 ha area on Hathinala and Bhawani Katariya sites, but underestimated on the remaining three sites. Similar to the power equation, all the five 3 ha log-linear equations underestimated the number of species for 15 ha area, again discrepancy being highest in the case of the Kota site. Given the differences in species richness among the sites, it is not surprising that the

Kota site equation, based on the most species-poor site, underestimates the combined species richness. The fitness of the model to the data was evidently not related to its capability to predict the number of species. Overall, on a small scale, the log-linear model prediction was better, and on larger scale the power function model had better predictability (Table 3).

Predictions from 1 ha equations using the two models yielded inconsistent results, while those from the 3 ha equations were consistent (Table 3). At a small scale (1 ha plot size), local interspecific interactions can have a prominent effect on patterns of species diversity, while on relatively larger scales (3 ha plot size), the importance of local influences may decline in favour of the recruitment processes from a broader species pool and habitat heterogeneity. This might have a differential effect on the predictability of the two models. Evidently, the scale at which the equation has been developed as well as the scale of extrapolation are important, and may lead to different inferences.

The predictions from the 3 ha power equation were closer to the observed value than those from the log-linear equation. Therefore, further discussion is based on only the species–area relationship from 3 ha power equation. Sites significantly differed in  $z$  values ( $F_{4,10} = 3.588$ ,  $P = 0.046$ ) and  $c$  values ( $F_{4,10} = 33.102$ ,  $P = 0.000$ ). Further,  $z$  was significantly affected by scale ( $F_{2,12} = 4.469$ ,  $P = 0.035$ ). In most cases,  $z$  decreased with increasing sample size (scale), but  $c$  was not affected. The coefficients  $c$  and  $z$  were independent of each other, although

**Table 2.** Species–area relationship of tree species ( $\geq 30$  cm cbh) at different scales on five sites in a dry tropical forest region of India

Site	Log-linear model				Power function model			
	<i>a</i>	<i>b</i>	<i>R</i> <sup>2</sup>	<i>S</i>	<i>c</i>	<i>z</i>	<i>R</i> <sup>2</sup>	<i>S</i>
Hathinala (ha)								
One	27.530 (0.065)	5.652 (0.049)	0.993	0.448	28.609 (0.096)	0.303 (0.004)	0.990	0.537
Two	25.936 (0.025)	5.259 (0.025)	0.995	0.340	25.618 (0.041)	0.250 (0.002)	0.988	0.564
Three	24.803 (0.024)	5.263 (0.024)	0.994	0.404	24.143 (0.034)	0.242 (0.002)	0.989	0.533
Khatabaran (ha)								
One	21.789 (0.057)	4.927 (0.043)	0.993	0.396	22.788 (0.118)	0.344 (0.006)	0.980	0.646
Two	21.906 (0.021)	4.740 (0.021)	0.996	0.286	21.588 (0.067)	0.261 (0.005)	0.959	0.919
Three	23.188 (0.027)	5.297 (0.028)	0.992	0.470	22.472 (0.039)	0.263 (0.002)	0.987	0.598
Majhauri (ha)								
One	18.774 (0.021)	3.595 (0.016)	0.998	0.148	19.295 (0.099)	0.267 (0.006)	0.971	0.569
Two	17.984 (0.026)	3.027 (0.027)	0.985	0.357	17.828 (0.052)	0.192 (0.004)	0.941	0.711
Three	18.683 (0.018)	3.393 (0.018)	0.991	0.307	18.337 (0.033)	0.201 (0.002)	0.974	0.528
Bhawani Katariya (ha)								
One	18.171 (0.034)	3.865 (0.025)	0.996	0.232	18.795 (0.122)	0.307 (0.007)	0.965	0.680
Two	17.255 (0.014)	3.835 (0.014)	0.997	0.190	16.991 (0.045)	0.274 (0.004)	0.971	0.622
Three	17.687 (0.010)	3.937 (0.011)	0.998	0.175	17.216 (0.043)	0.248 (0.003)	0.970	0.664
Kota (ha)								
One	3.921 (0.014)	0.939 (0.011)	0.988	0.097	4.133 (0.023)	0.377 (0.007)	0.981	0.122
Two	5.837 (0.009)	1.421 (0.009)	0.992	0.120	5.727 (0.015)	0.310 (0.004)	0.976	0.211
Three	5.297 (0.007)	1.366 (0.008)	0.991	0.127	5.095 (0.013)	0.298 (0.003)	0.979	0.194

Values in parentheses are  $\pm 1$  SE (standard errors of regression coefficients).  $R^2$  (the square of correlation coefficient) values were significant at  $P < 0.0001$ . All values are rounded to three decimal places.

the trend of the relationship was negative. The coefficient  $c$  was linearly related with the total number of species (Figure 1a), and nonlinearly, but positively with  $\alpha$ -diversity (Figure 1b). The coefficient  $z$  was related with the number of species per individual (Figure 1c).

Crawley and Harral<sup>11</sup> reported that  $z$  varied systematically with spatial scale and from habitat to habitat at the same spatial scale. On the other hand, Weiher<sup>9</sup> suggested that  $z$  was independent of scale. Our study showed that  $z$  is site- as well as scale-dependent, and therefore, location of the sample (i.e. site) as well as the scale used need be

considered in extrapolations of species diversity from species–area curve parameters.

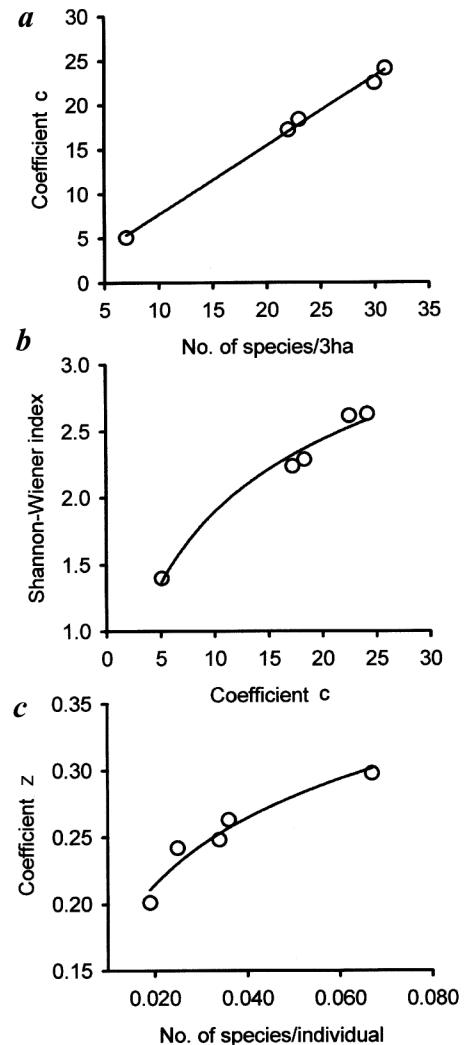
According to the species pool hypothesis<sup>17,18</sup>, a positive relationship between the coefficients  $c$  and  $z$  suggests that small-scale patterns are predictive of large-scale patterns. The no-interaction hypothesis<sup>19</sup> suggests that if there exists a negative relationship between the coefficients  $c$  and  $z$ , the small-scale patterns will not be predictive of large-scale patterns. Our study indicates that  $c$  and  $z$  are independent of each other, limiting their usefulness in predicting larger-scale patterns in species diversity.

**Table 3.** Prediction of number of species by power function and log-linear species–area equations. The observed number of species in 3 ha area on different sites is given in Table 1 and that in total 15 ha area was 49

Site	of species		Predicted number Percentage discrepancy in prediction	
	Power	Log-linear	Power	Log-linear
3 ha vs 1 ha equation				
Hathinala	40	34	+28.7	+8.8
Khatabaran	33	27	+10.8	–9.3
Majhauri	26	23	+12.3	–1.2
Bhawani Katariya	26	22	+19.7	+1.9
Kota	6	5	–10.7	–29.3
15 ha vs 1 ha equation				
Hathinala	65	43	+32.6	–12.6
Khatabaran	58	35	+18.1	–28.3
Majhauri	40	29	–18.9	–41.8
Bhawani Katariya	43	29	–11.9	–41.6
Kota	11	7	–76.6	–86.8
15 ha vs 3 ha equation				
Hathinala	46	39	–5.1	–20.3
Khatabaran	46	38	–6.5	–23.4
Majhauri	32	28	–35.5	–43.1
Bhawani Katariya	34	28	–31.2	–42.1
Kota	11	9	–76.7	–81.6

Wilson and Chiarucci<sup>10</sup> argued that according to the self-similarity concept, communities are not discrete, and the same pattern of heterogeneity is seen at all spatial scales<sup>20</sup>. Therefore, a large area will contain the number of species predicted from the small-scale species–area relation<sup>10</sup>. According to the discrete-community theory, plant communities have distinct boundaries within which species composition is homogenous, and the transition between them is relatively rapid. As such, the number of species would be underestimated when extrapolated from small-scale species–area curve to a larger scale due to the occurrence of new species across community boundaries<sup>10</sup>. The underestimation of number of species in 15 ha sampling plot by 3 ha equations seems to support the discrete-community concept. Hill<sup>21</sup> also pointed out that species–area relationships supporting Arrhenius law would not refute the existence of plant communities. An earlier study, on the basis of vegetation analysis on 42 sites using ten 10 m × 10 m quadrats per site, suggested that the dry tropical forest is a mosaic of communities, each of which is distributed in non-contiguous patches due to the heterogeneity of the environment and disturbance<sup>22</sup>.

Our study revealed that  $\alpha$ -diversity can be adequately described by the coefficient  $c$ , which itself can be accurately predicted from the total number of species; and the coefficient  $z$  can be used to predict the number of species



**Figure 1.** *a*, Relationship between total number of adult tree species  $S$  and power function coefficient  $c$ , according to  $c = -0.0922 + 0.776 S$  ( $R^2 = 0.996$ ,  $P = 0.0001$ ); *b*, Relationship between power function coefficient  $c$  and  $\alpha$ -diversity, according to  $\alpha = -0.216 + 0.816 \ln c$  ( $R^2 = 0.986$ ,  $P = 0.0007$ ); *c*, Relationship between number of species/individual ( $S_n$ ) and power function coefficient  $z$ , according to  $z = 0.497 + 0.072 \ln S_n$  ( $R^2 = 0.94$ ,  $P = 0.0065$ ).

per individual and vice versa. Thus these two coefficients are strong and complementary measures of species diversity. The higher the  $z$ , greater would be the impact of further harvest of individuals; in other words, the areas with high  $z$  would be more susceptible to selective or random felling of trees even if there is no contraction of area.

1. Wilson, E. O. (ed.), in *Biodiversity*, National Academy Press, Washington DC, 1988, pp. 3–18.
2. May, R. M. and Stumpf, M. P. H., *Science*, 2000, **290**, 2084–2086.
3. Colwell, R. K. and Coddington, J. A., in *Biodiversity: Measurement and Estimation* (ed. Hawksworth, D. J.), Chapman and Hall, London, 1995, pp. 101–118.
4. Diamond, J., *Ann. Mo. Bot. Gard.*, 1988, **75**, 117–129.
5. Leps, J. and Stursa, J., *Vegetatio*, 1989, **83**, 249–257.

## RESEARCH COMMUNICATIONS

---

6. Pimm, S. L., Russell, G. J., Gittleman, J. L. and Brooks, T. M., *Science*, 1995, **269**, 347–350.
7. Palmer, M. W., *Ecology*, 1990, **71**, 1195–1198.
8. Soberon, M. J. and Llorente, B. J., *Conserv. Biol.*, 1993, **7**, 480–488.
9. Weiher, E., *J. Ecol.*, 1999, **87**, 1005–1011.
10. Wilson, J. B. and Chiarucci, A., *J. Veg. Sci.*, 2000, **11**, 773–775.
11. Crawley, M. J. and Harral, J. E., *Science*, 2001, **291**, 864–868.
12. Huber, R., *Appl. Veg. Sci.*, 1999, **2**, 257–266.
13. McCune, B. and Mefford, M. J., PC-ORD Multivariate Analysis of Ecological Data, version 4, MjM Software Design, Oregon, USA, 1999.
14. SPSS, *SPSS Base 7.5 Applications Guide*, SPSS Inc, Chicago, 1997.
15. Williams, C. B., *Nature*, 1943, **152**, 264–267.
16. Connor, E. F. and McCoy, E. D., *Am. Nat.*, 1979, **113**, 791–833.
17. Taylor, D. R., Aarssen, L. W. and Loehle, C., *Oikos*, 1990, **58**, 239–250.
18. Ericksson, O., *Oikos*, 1993, **68**, 371–374.
19. Oksanen, J., *J. Ecol.*, 1996, **84**, 293–295.
20. Harte, J., Kinzig, A. and Green, J., *Science*, 1999, **284**, 334–336.
21. Hill, M. O., *J. Veg. Sci.*, 2001, **12**, 143–144.
22. Jha, C. S. and Singh, J. S., *J. Veg. Sci.*, 1990, **1**, 609–614.
23. Verma, D. M., Pant, P. C. and Hanfi, M. I., *Flora of Raipur, Durg and Rajnandgaon*, Botanical Survey of India, Howrah, 1985.

ACKNOWLEDGEMENT. Financial support from the Ministry of Environment and Forests, Govt. of India is acknowledged.

Received 5 December 2002; revised accepted 11 February 2003

---