Basic Appl. Ecol. **3**, 15–18 (2002)
© Urban & Fischer Verlag
http://www.urbanfischer.de/journals/baecol

Basic and Applied Ecology

Computing β-diversity from species-area curves

Carlo Ricotta[1,*] Maria Laura Carranza[2], Giancarlo Avena[1]

[1]Department of Plant Biology, University of Rome “La Sapienza”, Rome, Italy
[2]Faculty of Mathematics, Physics and Natural Sciences, University of Molise, Isernia, Italy

Received April 4, 2001 · Accepted July 25, 2001

Abstract

A measure of β-diversity is proposed for equal-sized presence-absence vegetation sample plots. The method begins by computing the minimum variance unbiased estimator of the expected species richness $E[S_n]$ as a function of sample size n for a set of N plots. Using a semilogarithmic relation, β-diversity is then determined from the slope of the best fitting line of the $E[S_n]$ vs. log n plot. β-diversity is thus interpretable as the linear relation between expected species richness and log sample size. The method is illustrated using 15 square plots of 100 m^2 each from a Matorral community in the Arid Chaco of central Argentina. Further, some analogies between the proposed method and Whittaker's multiplicative model of β-diversity are discussed.

Ein Maß für die β-Diversität wird vorgeschlagen, das für „presence-absence“-Vegetationsaufnahmen auf gleichdimensionierten Probeflächen gilt. Nach der Methode wird zuerst der „minimum variance unbiased estimator“ des erwarteten Artenreichtums als eine Funktion der Probengröße n für eine Anzahl von N Flächen berechnet. Unter Verwendung der semilogarithmischen Beziehung wird die β-Diversität über die „best-fit“-Steigung des Graphen $E[S_n]$ vs. log n bestimmt. Die β-Diversität kann damit als die lineare Abhängigkeit des Artenreichtums von der logarithmierten Probengröße interpretiert werden. Die Methode wird anhand von 15 Probequadraten von jeweils 100 m^2 aus der Matorall-Lebensgemeinschaft des ariden Chaco Zentralargentiniens erläutert. Zudem werden einige Analogien zwischen der vorgeschlagenen Methode und Whittakers multiplikativen Modell der β-Diversität diskutiert.

Key words: β-diversity – species richness – species-area curves

Introduction

Measures of biological diversity have been one of the cornerstones of ecological research. Whittaker (1960, 1972) defined α-diversity as the species diversity within community sample plots and γ-diversity as the total species diversity in a set of sample plots. He also defined β-diversity as the amount of turnover in species composition from one location to another. In this sense, β-diversity is essentially the same as MacArthur's (1965) between habitat diversity.

In the ecological literature, the most widely used diversity indices are species richness (i.e., the number of species present in a given community sample) and

---

***Corresponding author:** Carlo Ricotta, Department of Plant Biology, University of Rome “La Sapienza”, Piazzale Aldo Moro 5, I - 00185 Rome, Italy, Phone: ++39-06-4991 2408, Fax: ++39-06-445 7540, E-mail: carlo.ricotta@uniroma1.it

1439-1791/02/03/01-015 $ 15.00/0

those computed from species relative abundances, such as Shannon's entropy (Shannon 1948) and Simpson's index (Simpson 1949). Wilson & Mohler (1983) and Tóthmérész (1998) address the problem of computing β-diversity with data on species relative abundances. However, with all the shortcomings of diversity indices, the least controversial way to quantify β-diversity is to use measures suitable for presence and absence data (Wilson & Shmida 1984, Johnson & Patil 1995, Tóthmérész 1998).

One common approach to β-diversity is to look at the degree to which species compositions of sample plots differ. Whittaker (1960) proposed to measure the degree to which habitats are partitioned among species by a multiplicative model so that β-diversity can be expressed as the ratio between γ-diversity and average α-diversity within the set of sample plots:

$$\beta_w = S_N / \bar{\alpha} - 1; \tag{1}$$

where β_W is Wittaker's index of β-diversity, S_N is the total number of species recorded within the set of N sample plots, and $\bar{\alpha}$ is average species richness in the sample plots. Notice that in this paper we will use the terms *sample*, *sample plot* and *sample size* in the statistical sense. From a statistical viewpoint, sample size refers to the number of sample plots, and sample will refer to the full set of such plots, whereas in the ecological literature a sample is commonly understood as a single sampling unit, and this is a common source of confusion (Podani et al. 1993).

Another way of viewing β-diversity is to look at how species diversity changes along an environmental gradient. Within this context, Wilson & Shmida (1984) proposed an index of β-diversity explicitly developed for quantifying the degree of species turnover along a transect that simply adds the number of new species encountered along the transect to the number of species that are lost over the same transect. The result is then standardized by average sample plot richness to render it independent from α-diversity:

$$\beta_T = [g(H) + l(H)] / 2\bar{\alpha}; \tag{2}$$

where g(H) is the number of species gained (i.e., newly encountered) along the habitat transect H, and l(H) is the number of species lost along H. Notice that if the separation between α-diversity and β-diversity is incomplete, so that β-diversity is affected by species richness, then differences in β-diversity values could reflect differences in the species count rather than any fundamental difference in the amount of species turnover within the sample.

Nevertheless, β-diversity may be also characterized by more sophisticated methods. For example, Tóthmérész (1998) suggests measuring β-diversity by estimating the distribution of plot-to-plot dissimilarities within a vegetation sample using a non-parametric density estimator based on kernel smoothing (for mathematical details, see Tóthmérész 1998).

In this paper, we propose to compute β-diversity as the slope of a linearized species-area curve obtained by sequentially adding equal-sized plots. Further, we shortly discuss some analogies between species-area curves and Whittaker's multiplicative model of β-diversity. The method is illustrated using 15 square plots of 100 m^2 each from a Matorral community in the Arid Chaco of central Argentina.

Methodology

The rule that one finds more species sampling a larger area has more evidence to support it that any other about species diversity (Rosenzweig 1995). Species-area curves plot the number of species sampled as a function of sample size (Palmer 1990). One of these curves is constructed by plotting the number of species S_n present when n sample plots are drawn at random from a sample of N equal-sized plots as a function of the cumulative area sampled. An important property of this curve is that a minimum variance unbiased estimator for the expected number of species $E[S_n]$ is known (Kobayashi 1974, Engen 1976). Out of the n! ways to randomly aggregate n equal-sized sample plots, the minimum variance estimator of the expected number of species is:

$$E[S_n] = \sum_{i=1}^{S_N} \left[1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}} \right]; \tag{3}$$

where N_i is the number of sample plots out of N total sample plots that contain species i (Johnson & Patil 1995). It is easily demonstrated that for a small sample size, the contribution of rare species to $E[S_n]$ is low. Conversely, for n = N, $E[S_n]$ equals S_N. Notice that Equation (3) is also used for the calculation of species-individuals diversity curves. In this case, n and N represent the number of individuals and the population size, respectively (Smith & Grassle 1977, Magurran 1988).

As an application for demonstration, we constructed a species-area curve using data on 15 plots from a *Larrea cuneifolia* and *Geoffroea decorticans* Matorral community in the Arid Chaco of central Argentina. The Arid Chaco occupies the southwestern extreme of the Chaco phytogeographic province (central Argentina) and extends through inter-mountain basins (locally known as bolsones) at elevations between 150 to 300 m above sea level. Annual precipitation is approximately 400 mm, with a maximum during summer.

Mean annual air temperature is approximately 18 °C. The *Larrea cuneifolia* and *Geoffroea decorticans* Matorral represents a discontinuous transitional formation mainly dominated by xerophylous nanophanerophytes. The characteristic species are *Larrea cuneifolia*, *Geoffroea decorticans* and *Mimozyganthus carinatus* with scattered individuals of *Aspidosperma quebracho-blanco* (Cabido et al. 1994).

Using presence-absence data, the species richness of the *Larrea cuneifolia* and *Geoffroea decorticans* Matorral was enumerated in 15 equal-sized 10 m × 10 m plots (see Cabido et al. 1994, p. 448). Next, the expected species richness $E[S_n]$ was computed for cumulative sample sizes n ranging from 1 to 15.

Discussion

In plant surveys, species-area curves can be generally linearized using the classical semilogarithmic model proposed by Gleason (1922):

$$E[S_n] = h\log(n) + k; \qquad (4)$$

where h and k are sample-specific constants that represent the slope and the y-intercept of the linearized semilogarithmic function, respectively (Rosenzweig 1995). In particular, the slope h measures the (linear) rate of accumulation in species richness as a function of sample size, thus representing an adequate measure of plot-to-plot species dissimilarity or β-diversity. The fewer species that sample plots share, the higher the β-diversity (i.e., the slope of the linearized species-area curve) will be (Magurran 1988). Other candidate models for linearizing species-area curves include the Arrhenius (1923) model, where the logarithm of $E[S_n]$ is proportional to the logarithm of sample size, and the model proposed by Lauga & Joachim (1987) in which species-area curves are best represented by a Monod function with an horizontal asymptote.

For our specific example, using natural logarithms, the resulting species-area accumulation curve obtained from the 15 sample plots of the Arid Chaco Matorral proves linear in the semilogarithmic space with a slope of the best-fitting line h = 15.557 (Fig. 1).

Notice that, for n = 1, $E[S_1]$ is the expected number of species of one sample plot that is drawn at random among the set of sample plots. Therefore, $E[S_1]$ is conceptually close to the term $\bar{\alpha}$ that represents α-diversity in Equations (1) and (2). Similarly, if all sample plots are included in the calculation of $E[S_n]$, the result is total species richness S_N, also defined as γ-diversity. Therefore, in analogy to Whittaker's multiplicative model that explicitly relates the α and β components of diversity to overall γ-diversity, species accumulation curves represent a graphical model that connects α-diversity to γ-diversity in the species richness-sample size space. In this view, the slope of the linearized species-area curve adequately summarizes plot-to-plot species dissimilarity expressed as the (linear) dependence of expected species richness on log sample size.

Regarding index properties, personal communications with colleagues have elicited the hypothesis that, unlike many of the existing β-diversity indices, the method of computing β-diversity from the slope of species-area curves seems to be rather independent of the sample size. In addition, jackknifed estimates of the slope of species-area curves (Efron 1982) might provide information about the contribution of single plots to β-diversity. For instance, an index such as $(\beta_{-1} - \beta)/\mathrm{sd}[\beta_{-1}]$, where β_{-1} is the slope of the species-area curve with a given plot removed, β is the slope using all plots, and $\mathrm{sd}[\beta_{-1}]$ is the standard deviation of the jackknifed estimates, might provide a measure of each plot's relative contribution to β-diversity that could be correlated with relevant environmental variables.

Finally, despite the fact that species-area curves have been proposed by a number of authors to quantify habitat heterogeneity (Williams 1964, Whittaker et al. 1979) and β-diversity (Connor & McCoy 1979),

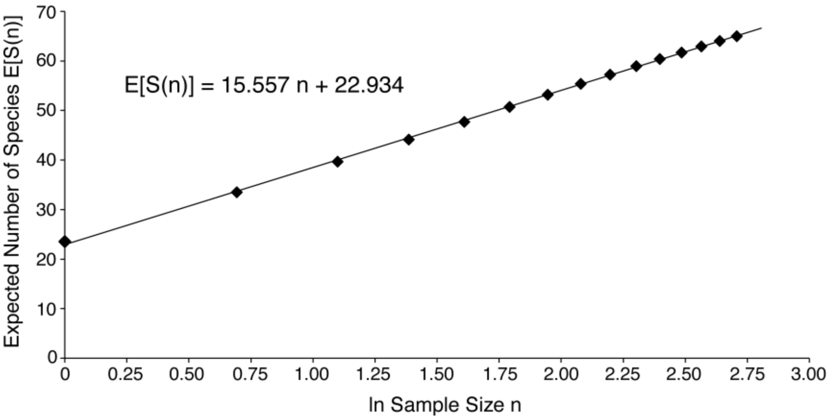

Fig. 1. Scatterplot of expected species richness $E[S_n]$ vs. cumulative sample size n for the 15 plots of the *Larrea cuneifolia* and *Geoffroea decorticans* Matorral.

the suggested measure should not be accepted without criticism. For instance, Wilson & Shmida (1984) note that it is both difficult statistically and inappropriate ecologically to select among competing mathematical models for representation of species-area curves. However, we suspect that this concern originates from the confusion between the mathematical formulation of β-diversity and its ecological meaning.

As ecological data are usually multivariate, there is a need for summarization. However, from a statistical viewpoint, condensing a large data set into one single index will always result in a loss of information, and there is no ideal summary statistics capable of uniquely characterizing all aspects of plot-to-plot species dissimilarity. Further, it is again worth noting that diversity as an unequivocal and inherent property of a biological system does not exist. Instead, as it is the case with any other statistics, such as the mean or the standard deviation, diversity values are merely numbers and their relevance to ecological problems must be judged on the basis of observed correlations with other environmental variables (Molinari 1989). This observation has led to the well-known comments by Hurlbert (1971) on the "nonconcept of species diversity" and by Poole (1974) that diversity measures are "answers to which questions have not yet been found". Nevertheless, despite this undue pessimism, we agree with Magurran (1988) that "diversity measures are valuable, but are only a means to an end. That end is that ecologists should be able to ask the questions and formulate the hypotheses to help them understand, and sensibly manage, the natural world".

Acknowledgements. The authors wish to thank two anonymous referees for the stimulating comments that substantially contributed improving this work.

References

Arrhenius O (1923) Statistical investigation in the constitution of plant associations. Ecology 4: 68–73.

Cabido M, Manzur A, Carranza L, Gonzalez Albarracín C (1994) La vegetación y el medio físico del Chaco Arido en la provincia de Córdoba, Argentina Central. Phytocoenologia 24: 423–460.

Connor EF, McCoy ED (1979) The statistics and biology of the species-area relationship. American Naturalist 113: 791–833.

Efron B (1982) The Jackknife, the Bootstrap, and other Resampling Plans. SIAM, Philadelphia.

Engen S (1976) A note on the estimation of the species-area curve. Journal du Conseil International pour l'Exploration de la Mer 36: 286–288.

Gleason HA (1922) On the relation between species and area. Ecology 3: 158–162.

Hurlbert SH (1971) The nonconcept of species diversity: a critique and alternative parameters. Ecology 52: 577–586.

Johnson GD, Patil GP (1995) Estimating statewide species richness of breeding birds in Pennsylvania. Coenoses 10: 81–87.

Kobayashi S (1974) The species-area relation I. A model for discrete sampling. Researches on Population Ecology 15: 223–237.

Lauga J, Joachim J (1987) L'échantillonage des populations d'oiseaux par la méthode des E.F.P.: interêt d'une étude mathématique de la courbe de richesse cumulée. Acta Oecologica. Oecologia Generalis 8: 117–124.

MacArthur RH (1965) Patterns of species diversity. Biological Reviews 40: 510–533.

Magurran AE (1988) Ecological Diversity and its Measurement. Princeton University Press, Princeton, NJ.

Molinari J (1989) A calibrated index for the measurement of evenness. Oikos 56: 319–326.

Palmer MW (1990) The estimation of species richness by extrapolation. Ecology 71: 1195–1198.

Podani J, Czárán T, Bartha S, (1993) Pattern, area and diversity: the importance of spatial scale in species assemblages. Abstracta Botanica 17: 37–51.

Poole RW (1974) An Intoduction to Quantitative Ecology. McGraw-Hill, New York.

Rosenzweig ML (1995) Species diversity in space and time. Cambridge University Press, Cambridge.

Shannon C (1948) A mathematical theory of communication. Bell System Technical Journal 27: 379–423.

Simpson EH (1949) Measurement of diversity. Nature 163: 688.

Smith W, Grassle FJ (1977) Sampling properties of a family of diversity measures. Biometrics 33: 283–292.

Tóthmérész B (1998) On the characterization of scale-dependent diversity. Abstracta Botanica 22: 149–156.

Whittaker RH (1960) Vegetation of the Siskiyou Mountains, Oregon and California. Ecological Monographs 30: 279–338.

Whittaker RH (1972) Evolution and measurement of species diversity. Taxon 21: 213–251.

Whittaker RH, Niering WA, Crisp MD (1979) Structure, pattern and diversity of a mallee community in New South Wales. Vegetatio 39: 65–76.

Williams CB (1964) Patterns in the Balance of Nature. Academic Press, New York.

Wilson MW, Mohler CL (1983) Measuring compositional change along gradients. Vegetatio 54: 129–141.

Wilson MW, Shmida A (1984) Measuring beta diversity with presence-absence data. Journal of Ecology 72: 1055–1064.