Original article

Linking species richness curves from non-contiguous sampling to contiguous-nested SAR: An empirical study

CrossMark

Maria Lazarina [a], [*], Athanasios S. Kallimanis [b], John D. Pantis [a], Stefanos P. Sgardelis [a]

[a] Department of Ecology, School of Biology, Aristotle University, 54124 Thessaloniki, Greece
[b] Department of Environmental and Natural Resources Management, University of Patras, 30100 Agrinio, Greece

ABSTRACT

The species–area relationship (SAR) is one of the few generalizations in ecology. However, many different relationships are denoted as SARs. Here, we empirically evaluated the differences between SARs derived from nested-contiguous and non-contiguous sampling designs, using plants, birds and butterflies datasets from Great Britain, Greece, Massachusetts, New York and San Diego. The shape of SAR depends on the sampling scheme, but there is little empirical documentation on the magnitude of the deviation between different types of SARs and the factors affecting it. We implemented a strictly nested sampling design to construct nested-contiguous SAR (SA_CR), and systematic nested but non-contiguous, and random designs to construct non-contiguous species richness curves (SA_SRs for systematic and SACs for random designs) per dataset. The SA_CR lay below any SA_SR and most of the SACs. The deviation between them was related to the exponent f of the power law relationship between sampled area and extent. The lower the exponent f, the higher was the deviation between the curves. We linked SA_CR to SA_SR and SAC through the concept of "effective" area (A_e), i.e. the nested-contiguous area containing equal number of species with the accumulated sampled area (A_S) of a non-contiguous sampling. The relationship between effective and sampled area was modeled as $\log(A_e) = k\log(A_S)$. A Generalized Linear Model was used to estimate the values of k from sampling design and dataset properties. The parameter k increased with the average distance between samples and with beta diversity, while k decreased with f. For both systematic and random sampling, the model performed well in predicting effective area in both the training set and in the test set which was totally independent from the training one. Through effective area, we can link different types of species richness curves based on sampling design properties, sampling effort, spatial scale and beta diversity patterns.

1. Introduction

The species–area relationship (SAR) is the increase in species richness with increasing area and is one of the few generalizations in ecology. During the long and rich history of field measurements and theoretical analysis, a wide array of phenomena have been included and analyzed under the umbrella term "SAR". Several authors, tried to systematically classify these different types of SAR. Scheiner (2003) recognizes four types on the basis of sampling characteristics combined with the method of analysis: Type I-data from strictly nested blocks, Type II-data from contiguous grids, Type III-data from regular but not contiguous grids and Type IV-data from islands that differ in area. Dengler (2009) proposes the use of the term "species–area relationship" (SAR) when referring to contiguous area and the term "species sampling relationship" for any other sample-based curve. However, area has ecological meaning even if it is not continuous, since as Scheiner et al. (2011) argue some species use space discontinuously and therefore, non-contiguous sampling designs might be more efficient in capturing diversity patterns, for example through the use of stratified sampling among habitat types, when the habitat is fragmented.

Azovsky (2011) distinguishes the SAR, which describes the increase of heterogeneity with expansion of area, from the species-sampling effort relationship, which describes the heterogeneity of species composition within an area. We use the term "species richness curve", sensu Scheiner et al. (2011), to refer to any of the above relationships. We refer to relationship between species

* Corresponding author.
E-mail addresses: mlazarin@bio.auth.gr (M. Lazarina), akallim@upatras.gr (A.S. Kallimanis), pantis@bio.auth.gr (J.D. Pantis), sgardeli@bio.auth.gr (S.P. Sgardelis).

richness and nested-contiguous area as SA$_C$R, species richness and systematically sampled (non-contiguous) area as SA$_S$R and species richness and randomly sampled area as SAC. To avoid confusion with terminology, when we use the term "sampled area" we mean the sum of the area sampled within a given extent. "Extent" is defined as the total surface area within which the sampling took place. SA$_C$R is the only case where sampled area coincides with the spatial extent of the samples. In all other species richness curves, sampled area is less than the extent.

The characteristics of data collection (sampling design, extent and grain size) affect the shape of the species richness curve, making studies conducted at various scales incomparable, and may influence the conclusions about the spatial variation of species richness (Rahbek, 2005; Kéry et al., 2008; Albert et al., 2010; Dengler and Oldeland, 2010; Lazarina et al., 2013a). Dengler (2008) highlights many pitfalls encountered in small-scale sampling and argues that sampling design defines the type of species richness curve. Therefore, the species richness curves derived from different sampling designs deviate from each other, but there is little empirical documentation on the magnitude of this deviation and the factors affecting it (but see Ulrich and Buszko, 2007; Chiarucci et al., 2009). The species richness curves are usually constructed by spatially sparse ecological data due to the difficulty and cost of obtaining complete species list or performing exhaustive area biodiversity surveys (Chiarucci et al., 2003). Consequently, many of the resulting species richness curves are based on non-contiguous space, but are often handled as if they were SAR's referring to nested-contiguous area (Dengler, 2008). Given that the SAR is a frequently used tool for ecological biogeography and conservation biology e.g. for projecting biodiversity loss due to habitat loss (e.g. Pimm and Raven, 2000; Thomas, 2004; Triantis et al., 2010), such practices might lead to erroneous conclusions and there are many pitfalls to avoid when using non-contiguous space for species richness curves construction (see the review by Smith, 2010). When constructing a Type II or III SAR sensu Scheiner (2003), how can we treat the "area" axis in order to correspond to contiguous space? A possible solution is the concept of "effective area". We define the "effective area" as the nested-contiguous area containing a number of species equal to that obtained from a non-contiguous sampling. Conceptually, we use the term "effective area" in the way Veech et al. (2002) refer to what they call "true area" of community being sampled. The "effective area" lays in-between sampled area and extent due to the spatial autocorrelation of species distributions. So, for a given species distribution pattern and an accumulated sampled area, we expect to find more species when the samples are distributed over a larger extent. How many more species? The answer is likely to depend on the characteristics of the sampling design and on the spatial species turnover.

In the present study, we explicitly examined the deviation between species richness curves based on nested-contiguous and non-contiguous sampling designs. We used data from contiguous area surveys and simulated a number of non-contiguous nested and random sampling designs in order to: i. test the hypothesis that for a given accumulated sampling area we find more species when the samples are more distant than if the sample plots are contiguous (i.e. the "effective area" is greater than the accumulated sampled area), ii. estimate the "effective area" corresponding to a given sampled area, iii. explore the factors affecting the relationship between sampled area and effective area, and iv. built a model to predict the effective area from sampling design and the underlying species distribution properties (species turnover rate).

2. Material and methods

2.1. Data

We used seven datasets of different taxonomic groups (four for birds, two for plants and one for butterflies): the British Breeding Birds [BA birds I (Sharrock, 1976) and BA birds II (Gibbons et al., 1993)], the New York State Breeding Birds [NYI birds (Andrle and Carroll, 1988) and NYII birds (McGowan and Corwin, 2008)], the San Diego County Bird [SD birds (Unitt, 2004)], the Massachusetts Breeding Bird Data [MS birds (Petersen and Meservey, 2004)], the British Vascular Plant Species [BA plants (Preston et al., 2002)], the Woody Plants of Mt. Holomontas [HOL I and II plants (Kallimanis et al., 2006, 2008)] and the British Butterflies [BA butterflies (http://data.nbn.org.uk/)].

Datasets differed in extent, grain, total number of species and species richness per cell (Table 1). We restricted our analyses on contiguous square blocks of grid cells of each dataset. Given the wide differences in grain size among datasets, to provide

Table 1
a. Source and abbreviation of the datasets of plants of Great Britain, Mt. Holomontas (Greece), breeding birds of Great Britain, New York State, San Diego County and Massachusetts and British butterflies used in this work. b. Grain, extent, total species richness and mean species richness and also the number and size of each square block used of each dataset.

a. Datasets used

Abbreviation	Dataset	Source
BA plants	The British Vascular Plant Species	New Atlas of the British and Irish Flora (Preston et al., 2002)
HOL$_I$ plants, HOL$_{II}$ plants	The Woody Plants of Mt. Holomontas	Kallimanis et al., 2006, 2008
BA$_I$ birds	The British Breeding Birds[a]	The atlas of breeding birds in Britain and Ireland (Sharrock, 1976)
BA$_{II}$ birds		The new atlas of breeding birds in Britain and Ireland: 1988–1991 (Gibbons et al., 1993)
NY$_I$ birds	The New York State Breeding Birds	1st Breeding Bird Atlas Data (1980–1985) (Andrle and Carroll, 1988)
NY$_{II}$ birds		2nd Breeding Bird Atlas Data (2000–2005) (McGowan and Corwin, 2008)
SD birds	The San Diego County Bird	San Diego County Bird Atlas (Unitt, 2004)
MS birds	The Massachusetts Breeding Bird Data	Massachusetts Breeding Bird Atlas (Petersen and Meservey, 2004)
BA butterflies	The British Butterflies	Butterfly Distributions for Great Britain for the period 2005–2009 from Butterfly Conservation and the Biological Records Centre (http://data.nbn.org.uk/)

b. Properties of the datasets and of the square blocks used

Dataset abbreviation	Grain (km^2)	Extent (#grid cells)	Species richness	Mean species richness/ cell	Square blocks used (extent of each square block)
BA plants	10^2	2019	3200	683.02	3 (400, 225, 121)
HOL$_I$ plants	4×10^{-6}	1024	19	2.91	4 (1024)
HOL$_{II}$ plants	4×10^{-6}	1024	17	2.63	4 (1024)
BA$_I$ birds	10^2	2733	216	78.80	4 (529, 289, 169, 144)
BA$_{II}$ birds	10^2	2678	255	80.10	4 (529, 196, 144, 121)
NY$_I$ birds	25	5293	245	67.94	2 (529, 289)
NY$_{II}$ birds	25	5267	251	71.89	2 (289, 169)
SD birds	~4.83	478	433	100.03	1 (361)
MS birds	25	950	338	57.42	2 (256, 196)
BA butterflies	10^2	2538	57	19.38	3 (196, 169, 121)

[a] The British Breeding Birds datasets do not include the 15 or so species which are still considered to be sensitive, nor indeed some of the very rare species.