

1 **A new method to analyze species abundances in space using gen-**
2 **eralized dimensions**

3 **Leonardo A. Saravia**, Ph.D.

4 Instituto de Ciencias Básicas

5 Universidad Nacional de General Sarmiento

6 J.M. Gutierrez 1159 (1613), Los Polvorines

7 Buenos Aires, Argentina.

8 lsaravia@ungs.edu.ar

Abstract

- Species-area relationships (SAR) and species abundance distributions (SAD) are among the most studied patterns in ecology, due to their application in both theoretical and conservation issues. One problem with these general patterns is that different theories can generate the same predictions, and for this reason they can not be used to detect different mechanisms.
- A solution for this is to search for more sensitive patterns. One possibility is to extend the SAR to the whole species abundance distribution. A generalized dimension (D_q) approach has been proposed to study the scaling of SAD, but there has been no evaluation of the ability of this pattern to detect different mechanisms.
- An equivalent way to express SAD is the rank abundance distribution (RAD). Here I introduce a new way to study scaling of SAD using a spatial version of RAD: the species-rank surface (SRS), which can be analyzed using D_q . Thus there is an old D_q based on SAR (D_q^{SAD}), and a new one based on SRS (D_q^{SRS}). I perform spatial simulations to relate both D_q with SAD, spatial patterns and number of species. Finally I compare the power of both D_q , SAD, SAR exponent, and the fractal information dimension to detect different community patterns using a continuum of hierarchical and neutral spatially explicit models.
- The SAD, D_q^{SAD} and D_q^{SRS} all had good performance in detecting models with contrasting mechanisms. D_q^{SRS} had a better fit to data and a strong ability to compare between hierarchical communities where the other methods failed. The SAR exponent and information dimension had low power and should not be used.
- SRS and D_q^{SRS} could be an interesting addition to study community or macroecological patterns.

Keywords: multifractals, species-rank surface, species-area relationship, multi-species spatial pattern.

1 Introduction

2 The species-area relationship (SAR) is considered one of oldest and best-documented patterns, and one of
3 a few fundamental generalizations, in ecology (Crawley & Harral 2001; Šizling *et al.* 2011). The SAR is
4 often characterized through a triphasic curve, with a range of intermediate scales corresponding to power
5 law relationship between the number of species and the area (Preston 1960; Hubbell 2001). Although other
6 quantitative forms could also be appropriate (Tjørve 2003; White *et al.* 2010) this power-law is more widely
7 accepted (Rosindell & Cornell 2007). This relationship implies a self-similar or fractal structure of species
8 distributions for a community across spatial scales (Šizling & Storch 2004).

9 SARs only give information about the changes of richness with scale, but can be extended by incorporating
10 the species abundances distribution (SAD), using Hill's generalized diversity indexes (Hill 1973). These
11 indexes follow from the definition of generalized entropies used in statistical physics, called Renyi's entropies
12 (Renyi 1970). The scaling of Renyi's entropies is called generalized dimensions, and is used in physics to
13 characterize multifractals (Beck 1990). Multifractals and fractals are related techniques first used in physics
14 to characterize scaling behavior of complex structures (Stanley & Meakin 1988); the difference is that fractals
15 look at the geometry of presence/absence patterns, while multifractals describe the arrangement of quantities
16 such as population density or biomass (Saravia *et al.* 2012a). Multifractal analysis has been applied to
17 ecology in different areas: metapopulation models (Gamarra 2005), analysis of natural landscapes (Kirkpatrick
18 & Weishampel 2005), search patterns in copepods (Seuront & Stanley 2014), and biomass dynamics in
19 microalgae (Seuront & Spilmont 2002; Saravia *et al.* 2012a; Dal Bello *et al.* 2014).

20 The application of generalized dimensions to extend SAR was first suggested by Ricotta (2000), and the
21 methodology was later developed and applied to Barro Colorado Island forest plots by Borda-de-Água *et al.*
22 (2002), who estimated generalized dimensions of SAD. Since then, several field studies have characterized
23 species abundance scaling—also called the species diversity-area relationship (DAR)—using generalized
24 dimensions and other multifractal techniques (Zhang *et al.* 2006; Yakimov *et al.* 2008; Wei *et al.* 2013). In
25 addition, generalized dimensions have been applied to a spatially explicit neutral model (Yakimov *et al.* 2013)
26 and used in open source software for ecological multifractal analysis (Saravia 2014).

27 The species abundance distribution (SAD) is another fundamental pattern in ecological communities, and
28 play a major role in ecology and conservation (McGill *et al.* 2007). SADs have been used to compare different
29 communities and to compare models and data, but different mechanisms can produce nearly identical SADs
30 (Chave *et al.* 2002; Rosindell & Cornell 2013). SADs are often presented using rank-abundance diagrams
31 (RADs) in which the log-abundance is plotted against the rank of the species (McGill *et al.* 2007). RADs

are equivalent to cumulative distributions and thus are a robust way to visualize the SAD without losing information (Newman 2005).

Here I propose a new way to analyze the relation of SAD with spatial scale, by attaching the rank of each species to its spatial distribution; in this way the multivariate spatial distribution of species is summarized into a univariate two-dimensional distribution. I call this spatial distribution the species-rank surface (SRS), and it can be analyzed and compared using generalized dimensions. This paper uses simulations of spatial patterns to compare the behavior of generalized dimensions calculated from SRS and SAD. First I use simple artificial spatial patterns: regular and randomized, combined with a uniform and logseries SAD, to observe the behavior of generalized dimensions. Then I use a continuum of neutral and hierarchical models to test if generalized dimensions can detect different communities, estimating statistical power and type I error rate. I also compare the performance of generalized dimensions with single dimensional indicators: the SAR exponent and the information dimension. A summary of the methods used to compare communities is presented in table 1.

Methods

Multifractal analysis

Extensive reviews of generalized dimensions and multifractal methods applied to ecology are available (Seuront 2009) and some good introductions have also been published (Scheuring & Riedi 1994; Borda-de-Água *et al.* 2007). Thus I will only give a brief description. The generalized dimensions technique analyzes the scaling properties of quantities distributed in a space that we assume to be two dimensional (i.e., a plane). This distribution should be self-similar across some range of scales. This is called being multifractal, which can be mathematically represented in different ways (Harte 2001), of which the closest to ecology are the generalized dimensions D_q (Grassberger 1983), also called Renyi dimensions (Renyi 1970). D_q has been used to characterize the probabilistic structure of attractors derived from dynamical systems (Hentschel & Procaccia 1983).

I will analyze two kinds of D_q : the standard one used to analyze any quantity distributed in space, and another based on the SAD of a community in an area. The first kind is used to analyze the species rank surface SRS (described later) so I call it D_q^{SRS} . The second kind is based on SAD so I named it D_q^{SAD} . To estimate generalized dimensions I used in both cases the method of moments based on box-counting (Evertsz & Mandelbrot 1992).

1 The mathematical representation for D_q^{SRS} and D_q^{SAD} is slightly different. For D_q^{SRS} the spatial distribution
 2 of quantities μ is covered with a grid, dividing it into $N(\epsilon)$ boxes of side ϵ , allowing us to calculate the value
 3 $\mu_i(\epsilon)$ in each. Then the so-called partition function is computed as:

$$4 \quad (1) \quad Z_q(\epsilon) = \sum_i^{N(\epsilon)} (\mu_i(\epsilon))^q$$

5 Where q can be any real number and is called *moment order*. The operation is performed for different values
 6 of ϵ and q , within a predetermined range. The generalized dimension is then calculated as:

$$7 \quad (2) \quad D_q^{SRS} = \frac{1}{q-1} \lim_{\epsilon \rightarrow 0} \frac{\log(Z_q(\epsilon))}{\log \epsilon}$$

8 When $q = 1$, the denominator of the first term in D_q^{SRS} is undefined, so it must be replaced by the following
 9 expression:

$$10 \quad (3) \quad D_q^{SRS} = \lim_{\epsilon \rightarrow 0} \frac{\sum_i^{N(\epsilon)} \mu_i(\epsilon) \log(\mu_i(\epsilon))}{\log \epsilon}$$

11 In practical cases as the limit can not be assessed, and the dimensions are estimated as the slope of the
 12 $\log(Z_q)$ versus $\log(\epsilon)$ in equation (1) replacing by the numerator in equation (3). This is done for different
 13 values of q , provided that it is a real number, which yields a graphs of D_q^{SRS} in terms of q . This is called the
 14 spectrum of generalized dimensions.

15 To be an approximate multifractal, the relationship $\log(Z_q)$ versus $\log(\epsilon)$ should be well described by a linear
 16 relationship, but a linear relationship with superimposed oscillations is also acceptable (Borda-de-Água *et al.*
 17 2007). A range of q and ϵ values must be established, and then D_q^{SRS} is estimated using linear regression.
 18 Note that D_q^{SRS} is defined as the limit $\epsilon \rightarrow 0$ (equations 2 and 3), and thus to use the method it is sufficient
 19 that a scale exists below which a linear relationship applies (Hentschel & Procaccia 1983).

20 To analyze species-abundance-area relationships with multifractals as Borda-de-Água *et al.* (2002), the boxes
 21 are replaced by species. Thus at each spatial scale ϵ each species holds the quantity of interest: its own
 22 abundance. Then the partition function is defined as a sum over the species present $S(A)$ in an area A and
 23 the side of the box ϵ is replaced by the area:

$$24 \quad (4) \quad Z_q(A) = \sum_i^{S(A)} (\mu_i(A))^q$$

25 where $\mu_i(A)$ is the abundance of species i in an area A . D_q^{SAD} is defined as:

$$(5) D_q^{SAD} = \frac{1}{1 - q} \lim_{A \rightarrow \infty} \frac{\log(\langle Z_q(A) \rangle)}{\log A}$$

The angle brackets indicate averages of $Z_q(A)$ over subplots with area A . Thus when I calculate D_q based on species abundances distribution, I refer to it as D_q^{SAD} , and when I calculate D_q from the species rank surface, D_q^{SRS} ; if I mention D_q without superscript I refer to both.

D_q^{SAD} represents the scaling of the Hill's generalized diversity index (Hill 1973): when the moment order is $q = 0$, then D_q^{SAD} becomes the exponent of the SAR power-law scaling; when $q = 1$, D_q^{SAD} represent the scaling of Shannon diversity index; and when $q = 2$, D_q^{SAD} becomes the scaling of Simpson's index. This is why D_q^{SAD} can characterize diversity-area relationships.

Theoretically D_q must be a non-increasing function of q (Hentschel & Procaccia 1983), which means that if $q_1 \geq q_2$ then $D_{q_1} \leq D_{q_2}$. Some studies have shown small violations of this property for D_q^{SAD} (Borda-de-Água *et al.* 2002; Zhang *et al.* 2006). These violations are related to the way that D_q^{SAD} is defined: the summation of equation 4 is over species, while the summation of the original definition, equation 1, is over boxes, and this changes the way in which the mathematical limits are taken and also the computation method of D_q^{SAD} . A partial solution has been proposed (Yakimov *et al.* 2014), but the anomalies observed may be related to the mathematical assumptions needed for D_q to be non-increasing, in which case a new mathematical proof should be developed for D_q^{SAD} . Thus as long as the linear relationship is reasonable I take D_q^{SAD} as a useful technique of analysis.

In a previous work I proposed a new way to analyze species-abundance-area using multifractals, one that fits more closely to the original definitions of equations 1 - 3: the species-rank surface (Saravia 2014) or SRS. To construct the SRS the spatial distribution of species has to be transformed by assigning to each species position its rank. First I use the species abundances, at the whole plot level, to calculate the species' rank ordering from highest to lowest, assigning a number starting with one. If there are species with the same abundance the ranks are assigned at random, in this way all species are present in the surface. Then the rank is assigned to the spatial position of the individuals of each species, forming a surface. This (mathematical) landscape has valleys formed by the most abundant species and peaks determined by the rarest. Finally the standard multifractal analysis is applied. If sampling was performed using quadrats, without taking the spatial position of individuals, the sum of the ranks of the species in the smallest quadrats can be used to form the SRS.

The fact that oscillations should be present and the data still accepted as self-similar (Liebovitch 1998) and by extension as multifractal, makes difficult the rejection of a data set based on tests of quadratic trends as

proposed by Yakimov et al.(2008), the curvature of a quadratic function can be fitted by a periodic function with the right amplitude. The coefficient of determination (R^2) was used as a descriptive measure of goodness of fit for multifractals (Borda-de-Água *et al.* 2002); in general this is not a good indicator of linearity but since linearity is not exclusively required I use it here in a descriptive way. The C++ source code to perform multifractal analysis is available at <https://github.com/lisaravia/mfsba>.

We don't have a theoretical shape of D_q derived for spatial distributions of species, all we know is that D_q is bounded such as D_∞ and $D_{-\infty}$ exist, that D_q is a non-increasing function of q —that means $D_{-\infty} \leq D_\infty$. Generalized dimensions (D_q) can be interpreted like a SAR power law exponent: with larger values, the change in the number of species is greater when the scale of observation changes to a larger area. D_q express the change of the quantity under study when scale changes, but is modulated by q . When q is positive the terms of the sums (equations 2 & 5) with more abundant species have more weight, and become even more important when q is greater. When q is negative we have the opposite pattern: less abundant species have more weight in the sum, and so D_q reflects the change of rare species. When q is larger in its absolute value, D_q is driven by more and more extreme values, and thus D_q will have higher variance because extreme values are rarer. Here I present most figures with a range of q from -24 to 24, but for statistical comparisons use a smaller range (from -10 to 10) to avoid large variances.

Thus D_q^{SAD} for the negative q (left side of the curves in Figure 2) characterize the scaling of less abundant species but we don't know exactly which species, the curvature of D_q^{SAD} when it is approaching $D_{-\infty}$ depends on the species abundances of the less abundant species but we don't know exactly how the abundances and spatial patterns of these will influence the shape of D_q^{SAD} . Also D_q^{SRS} reflect the spatial patterns of the less abundant species, the quantity $D_{-\infty}$ can be independently calculated (Borda-de-Água *et al.* 2002) and used to study rare species from the point of view of SAD and SRS, this is a different possible application of generalized dimensions. The positive part of the D_q spectra characterize the most abundant species for both SAD and SRS, but again we don't know exactly how the shape of SAD or the spatial pattern will change the curvature of D_q . The quantity D_∞ characterize the most abundant species, its relationship with $D_{-\infty}$ could give us the information about the relation of dominant and rare species. Thus the range of D_q gives us important information but also the shape of the spectrum is necessary to characterize the scaling of SAD or SRS.

Generalized dimension relationship with spatial patterns and SADs

I simulated species' spatial patterns with different SADs to demonstrate how D_q is related to them. First I used a uniform SAD, in which all species have approximately the same densities. To generate it I take the

number of individuals of each species from a Poisson distribution with the same mean. I distributed them in bands over a spatial grid so they form a regular spatial pattern, in which each grid position is occupied by exactly one individual. I chose the number of species to exactly divide the side of the grid so all species are strips with approximately the same width (Figure 1). I used square grids with sides of 256 and 512 sites which contain 65536 and 262144 individuals respectively, and 8, 64 and 256 species, and then calculated D_q for the regular pattern, randomizing the positions of species to compare D_q obtained with these two extreme cases. The second SAD I used is a Logseries (Fisher *et al.* 1943) with the same number of species and the same sides as previously. I used the R package `untb` (Hankin 2007) to calculate the density for each species; this basically uses a Poisson distribution with the expected Logseries abundances as means. I then built the regular pattern with strips of species, but as species have different abundances the widths for each species are different (Figure 1). I then estimated D_q for the regular and randomized patterns, simulated 10 spatial patterns for each case, and calculated the mean and standard deviation of D_q .

Spatially explicit model

To simulate more realistic patterns of species-abundance-area relationships I used a stochastic spatially explicit model. I developed a stochastic cellular automata (Molofsky & Bever 2004) model that can switch between neutral or hierarchical competition, representing a continuum between niche and neutral communities (Gravel *et al.* 2006). Under neutral competition individuals do not interact, and all have the same mortality, colonization rates, and dispersal distances; in spite of these gross simplifications neutral models are capable of predicting several real community patterns (Rosindell *et al.* 2011). At the other end of the continuum are niche communities represented by hierarchical competition models (Tilman 1994). In this case species have differences that imply a competitive hierarchy, in which some species are always better than others, producing competitive exclusion (Chave *et al.* 2002). I added a probability of replacement ρ to the neutral model: when $\rho = 1$ more competitive species always replace less competitive and the model behaves as a pure hierarchical one, and when $\rho = 0$, there is no replacement of species and the model is completely neutral. A more thorough description of the model is given in appendix A, and its C++ source code is available at <https://github.com/lasaravia/neutral> and figshare <http://dx.doi.org/10.6084/m9.figshare.969692>.

Following a classical neutral scheme the model has a metacommunity: a regional collection of communities. With probability m an individual of a species i can migrate from the metacommunity at a rate proportional to its frequency X_i in the metacommunity. Species can also disperse locally, and I assume an exponential dispersal kernel with average dispersal distance d . Other model parameters are the mortality rate μ , the number of species in the metacommunity and also the size of the community, represented as the *side* of

the grid used in the simulations. I use a logseries SAD for the metacommunity, defined by the maximum number of individuals (*side x side*) and the number of species (Fisher *et al.* 1943). To represent a competition colonization trade-off I choose the most competitive species to be the less abundant in the metacommunity. The values of the parameters were in the range estimated for BCI from the existing literature (Condit *et al.* 2002; Etienne 2007; Anand & Langille 2010). I performed 50 simulations for each combination of parameters given in Table 2. To compute the statistical power I made comparisons of communities with different levels of ρ , representing more neutral or hierarchical communities, in which the other parameters were kept constant. I also made comparisons between repetitions with the same ρ to calculate the type I error.

Statistical comparison of methods

I analyzed the performance of two kinds of methods to differentiate communities. The first consists of a set of points or curves: species abundance distributions (SAD), generalized dimensions D_q^{SAD} and D_q^{SRS} . I am testing here for differences between two communities, thus I do not have information about the variability in D_q for each q or in abundance of each species, so I assume that the sample comes from an unspecified univariate distribution. For these I used a permutation procedure with the k-sample Anderson-Darling (AD) statistic (Scholz & Stephens 1987). The AD statistic measures the differences between the empirical distribution functions (EDF) of two datasets as a weighted sum of square deviations between the EDFs (Feigelson & Babu 2012). This means that the information about q is lost, but as D_q spectra is continuous and non-increasing—with the mentioned anomalies in D_q^{SAD} — the shape of the curve is mirrored in the EDF and this is why the method can be used. For species abundance distributions the EDF is equivalent to the rank abundance distribution (RAD) (Newman 2005), thus I am really comparing RADs with this method. I calculated p-values using 1000 permutations, in all cases I use the package kSamples (Scholz & Zhu 2012) in the R statistical language (R Core Team 2014). Scripts for all analyses are available at github <https://github.com/lisaravia/SpeciesRankSurface> and figshare <http://dx.doi.org/10.6084/m9.figshare.1276105>. The second kind is based on a single dimension or power exponent: the SAR exponent and the information dimension. The SAR exponent is part of the D_q^{SAD} spectra when $q = 0$ (Borda-de-Água *et al.* 2002); an equivalent single number measure from D_q^{SRS} is the information dimension (Ricotta 2000; Chappell & Scalo 2001), that is the D_q^{SRS} when $q = 1$. I calculated the power of these with a T-test using the standard deviation (SD) obtained from the box-counting regressions. These SD are obtained with autocorrelated data because small squares are nested within big squares (see Multifractal Analysis). The consequence is that the SD may be underestimated, but the slopes estimates are still unbiased (Kutner *et al.* 2005). This should result in an increased type I error rate and also in a spurious increase in power. Besides power calculated for

one dimension should be lower than the power for the complete spectra I would like to show if this simpler and less computationally intensive method could give good results to compare different communities. Other studies have shown that autocorrelation don't have in some cases significant consequences (Legendre *et al.* 2002) so I would also like to investigate the influence of autocorrelation in these dimensions.

Calculation of power and type I error

I simulated communities with different degrees of neutral/hierarchical structure, given by the parameter ρ of the model. The power of a test is the ability to reject the null hypothesis (H_0) when it is false. The significance level to reject H_0 was set *a priori* at $\alpha = 0.05$ in all cases, and the rejection rate of each test was calculated as the proportion of P values that less than or equal to α . To estimate power I used independent simulations of communities (50 repetitions) with the same parameters except ρ .

The type I error is the probability of rejecting H_0 when it is true (false positive). In our simulations, H_0 is true if two simulated communities have the same ρ (and also are equal in the other parameters). To estimate type I error I compared independent simulations of communities with the same set of parameters (50 repetitions) and computed the proportion of rejection.

Results

Simple spatial patterns and SAD

I calculated two versions of D_q : a) the original definition due to Borda-de-Água et al.(2002) where D_q measures the change in SAD as we change scale (D_q^{SAD}), and b) D_q based on SRS, which measures the change in the spatial distribution of species' ranks as scale changes (D_q^{SRS}). D_q measures the rate of change with scale from a baseline that is defined by D_0 . When we study SAD, D_0^{SAD} is the SAR exponent and its value is around 0.5. A spatial distribution of species that duplicates its number with a duplication of the side of the area studied has a value of exactly 0.5. When we study SRS the D_0^{SRS} is the fractal dimension of the spatial distribution of species. Note that in the simulations here, the individuals completely fill the available space, and thus D_0^{SRS} is equal to 2.

For the uniform SAD we expected D_0^{SAD} to be around 0.5, following a symmetric pattern around $q = 0$, as all species have the same abundance and occupy the same area. The symmetric pattern was not observed in the regular cases (Figure 2) because the negative part ($q < 0$) analyzes numbers close to 0 and the

logarithm enhances the differences between small numbers (Laurie & Perrier 2011). Thus the difference $\Delta D_q = |D_q - D_0|$ is greater for $q < 0$.

Theoretically D_q should be decreasing or constant, but this was not observed in D_q^{SAD} for the randomized spatial patterns with fewer species. This is because when changing scales, there is a point at which no new species are found, and the scaling relationship breaks. Figure 3 shows an example of D_q fitted using linear relationships for 64 species and a side=256 sites. The scaling for a randomized pattern D_q^{SAD} breaks at 1.2, equivalent to box with side 16 or an area of 256 units. In contrast, the scaling for the regular pattern D_q^{SAD} shows oscillations around the fitted line but no evidence of breaks. When the number of species is higher (256) the D_q^{SAD} is similar to the randomized one (Figure 2); this happens because new species appear in the whole range of scales used.

The R^2 values (Appendix table 1) indicate the presence of poor fits or a scaling break. The D_q^{SAD} for randomized patterns and uniform SAD have the lowest R^2 of all cases. Based on all simulations I observe that the cases with the following conditions do not present anomalies: 90% of D_q should have an R^2 of 0.6 or greater, and 50% should have an R^2 of 0.9 or greater (Appendix table 1); but one always should check the plots of the fits (Figure 3) and eventually change the ϵ range. Several patterns fail to comply these conditions: for example, all the uniform randomized patterns, and the logseries randomized with 8 species (Appendix table 1 and Appendix figures 4-8) and these can be used as a guide to determine when the method can be applied.

The D_q^{SAD} for logseries had a more symmetric pattern than for uniform SAD (Figure 2), and exhibited better fits with higher R^2 (Figure 3). Comparing regular and randomized spatial patterns, the D_q^{SAD} curves were superposed or inside the SD of the other. Thus it seems that D_q^{SAD} cannot distinguish between such patterns (only considering the cases where the fits are good). Moreover the range of D_q^{SAD} did not change very much with the number of species, as D_q^{SAD} seems to depend mostly on the SAD used to generate the spatial pattern.

For D_q^{SRS} the theoretical decreasing pattern was fulfilled in all cases, and no anomalies were observed (Figure 2). As in the previous case, an asymmetric pattern, was observed with D_q^{SRS} around 2, an asymmetry more pronounced for patterns with uniform SAD than for logseries SAD. This is because logseries SAD have one very abundant species, several less abundant and rare species scattered through the pattern (Figure 1). Thus the abundant species dominated the spatial pattern and in some cases produces a greater $\Delta D_q = |D_q - D_0|$ in the positive side of the plot (Figure 2, 8 Species).

The uniform SAD produced D_q^{SRS} with higher ΔD_q values for regular patterns in the $q < 0$ side. This is

because in the regular pattern the species are aggregated, whereas in the randomized pattern there is no aggregation so D_q^{SRS} is closer to two. Thus D_q^{SRS} for regular and randomized are more different on the negative side, and more similar on the positive side. For logseries SAD, the differences in D_q^{SRS} are similar at negative or positive sides of q . In general D_q^{SRS} curves for different spatial patterns and different SADs are distinct, except in some cases for 8 species the curves are inside the SD of a different pattern.

The R^2 values for D_q^{SRS} were all >0.9 , are higher than D_q^{SAD} , and all complied with the conditions described above (Appendix table 1). Their linear trends were also better (Figure 3). An example of linear trends for different number of species and different SADs is shown in the appendix (Appendix figures 4-8). The same qualitative patterns of D_q^{SAD} and D_q^{SRS} are observed for simulations with side=512 (Appendix figure 1).

Simulated Neutral communities

Examples of the patterns simulated by the Neutral/hierarchical model are shown in Figure 4. By definition, hierarchical communities have more competitive species with lower index numbers, and neutral communities have more abundant species with higher index numbers, as determined by metacommunity abundance (see appendix model description). With a greater degree of competitive hierarchy, one or few species dominate and several rare species are scattered over the landscape (Figure 5). This produces a mostly uniform pattern of dominant species with rare species distributed at random. In neutral communities the most abundant species are not so dominant (Figure 5), and leave space for species with intermediate abundances, producing a pattern of several aggregated species. Aggregation is produced in this model only because dispersal is mainly near the parent.

For both estimated D_q the R^2 values were very good: D_q^{SRS} was always $R^2 > 0.9$ and D_q^{SAD} had in almost all cases $R^2 > 0.6$ and a 50% or more of the cases greater than 0.9 (Appendix table 2). Figure 6 shows an example of D_q fitted using linear relationships for a metacommunity of 86 species and a side of 256 sites, examples with different metacommunity species and side are given in the appendix (appendix figures 9-13).

There are two groups of D_q^{SAD} (Figure 7): one composed of neutral like communities for $\rho < 0.1$ and another composed of more hierarchical ones for $\rho > 0.1$. The curves for hierarchical communities were more separated for negative q than for positive q . In neutral communities this pattern was inverted, with positive q having more different curves. This reflects the patterns in SAD: hierarchical communities have one or few relatively abundant species, resulting in D_q^{SAD} reaching 0 quickly, and no new abundant species are found when changing scale. Neutral communities have more species with intermediate densities, producing $D_q^{SAD} > 0$ on the positive side.

1 In theory D_q have a constant value when q tends to infinity (negative or positive). Here, D_q^{SAD} spectra
 2 quickly reached a constant maximum for negative q and a minimum for positive q , and this pattern was more
 3 pronounced with hierarchical communities because they tended to have two types of species: dominant ones
 4 reflected on the positive side, and rare species on the negative. When communities are more neutral ($\rho < 0.1$)
 5 and there are more species with intermediate densities, D_q^{SAD} tended to reach the asymptotic values more
 6 slowly in the negative side.

7 For D_q^{SRS} a similar groups of neutral or hierarchical communities are also present (Figure 7). We previously
 8 saw that D_q^{SRS} is more related to the spatial pattern than D_q^{SAD} , and thus we can interpret D_q^{SRS} in
 9 terms of randomness and aggregation of species. For hierarchical communities, negative-side D_q^{SRS} is very
 10 close to 2, that is the dimension of a uniform surface, with rare species exerting a very low influence on
 11 uniformly-distributed dominants. For neutral communities there are more species with low to medium
 12 densities, and they have greater aggregation, and thus D_q^{SRS} is higher.

13 When q is positive, lower values of D_q^{SRS} mean more-intense spatial patterns. Communities with $\rho = 1$
 14 are the most hierarchical, with one dominant species and a few very rare species (Figure 6). For these
 15 communities D_q^{SRS} is closer to 2, representing the uniform spatial distribution of dominant species. When
 16 the metacommunity has more species the local community also has more species (Appendix table 3) and
 17 D_q^{SRS} starts to deviate from 2 at lower q . D_q^{SRS} for the intermediate hierarchical case ($\rho = 0.1$) starts higher
 18 than neutral at q near 0, but crosses neutral curves and ends in the lowest place. The communities have
 19 more species that also are more abundant. but still have few individuals; this forms very sharp peaks in
 20 the SRS and produces a D_q^{SRS} farther from 2. The curvature of D_q^{SRS} is thus more pronounced when there
 21 are more species. For ρ less than 0.1 communities are more neutral and have more species with similar
 22 densities, forming softer valleys and peaks that result in a D_q^{SRS} intermediate between the two hierarchical
 23 cases. Simulations with side=512 exhibited similar patterns for D_q (Appendix figure 14).

24 Statistical Power and type I errors

25 To calculate the power of the methods I compared communities with different ρ values; in this comparison
 26 the alternative hypothesis is true. Instead for estimating type I error, we need to compare different runs of
 27 communities simulated with identical parameters. I talk of high power when its value is 0.75 or higher, and
 28 low power when it is 0.5 or lower.

29 For D_q^{SAD} and D_q^{SRS} different ranges of q can be used. High values of q in absolute terms should produce D_q
 30 with high variances, resulting in a higher spread of values obtained in different simulation runs. Ranges of q

between -10 and 10 or narrower are generally used (Yakimov *et al.* 2008; Laurie & Perrier 2011; Saravia *et al.* 2012b; Wei *et al.* 2013) but sometimes the applied range has been wider (Saravia *et al.* 2012a). I started using a q range of -24 to 24, and found that for this range type I error rates were, in all cases, higher than the nominal significance level $\alpha = 0.5$ (Appendix table 4). As a statistical test is valid if the type I error is lower or equal to α (Edgington 1995), to assure the validity for these methods a narrower range should be used. I thus used a q range between -10 and 10.

Using only one dimension of the spectra (D_0^{SAD} and D_1^{SRS}) resulted in a power generally below 0.5 (Table 4) and the type I error around 0.4, much greater than α . These high type I error values were expected due to the presence of spatial autocorrelations in the dependent variable (Legendre *et al.* 2002). Parameter estimates can be corrected in different ways (Legendre *et al.* 2002), but these procedures should not increase the power of D_0^{SAD} and D_1^{SRS} .

For communities with lower species numbers (11 species in the metacommunity) the comparisons made with SAD had a constant low power (Table 4) and also have low power comparing communities with different degree of neutrality (Figure 8), so no matter how different the communities are as the points used in the test are the number of species the power is low. In contrast, the generalized dimensions D_q^{SAD} and D_q^{SRS} had a high power but Type I error also greater than α . One way to alleviate this problem is to check for a coincidence of the two methods SAD & D_q ; another would be to increase the number of points used inside the q range, because D_q could be calculated for any real number. I used 21 points (Table 4) but that could be increased, as the only restriction is the additional computational time required. In simulated communities with more species (86 & 341 species metacommunity) the type I error fell below α for all the methods and the overall SAD was slightly more powerful (table 4).

Differences between communities influence power (Figure 8). Note that with $\rho < 0.1$, the communities compared are more neutral with a similar number of species and SADs. These cases correspond to the first two rows of Figure 9, the power in most cases was below 0.5, and thus D_q^{SAD} and D_q^{SRS} could not discriminate communities. The exception was SAD for neutral communities, when the the metacommunity had 341 species: in this case, the power was near or greater than 0.5. These are comparisons with a higher number of points (circa 100) so this results in a greater power.

Communities with $\rho \geq 0.1$ are more hierarchical and have different numbers of species and SAD than communities with $\rho < 0.1$. For these cases (last 2 rows of figure 8), the power was high (over 0.75) in most cases. The comparison between more hierarchical communities $\rho = 0.1 \geq 0.1$, is different: SAD and D_q^{SAD} , had less power (below 0.25 in some cases), and D_q^{SRS} had the highest power.

Discussion

In this paper I present a new macroecological metric D_q^{RS} based on generalized dimensions, and use model simulations to compare it with other similar metrics: D_q^{AD} , SAD, SAR exponent and information dimension.

While D_q^{AD} measures the change in species abundance distribution with scale, D_q^{RS} represents the change in the spatial distribution of ranks of species. Thus D_q^{RS} is related to the spatial pattern of species and to its abundance distribution. D_q^{AD} also reflects changes in spatial pattern; but my results suggest that it cannot distinguish between regular and randomized spatial patterns. In contrast, D_q^{RS} curves differ clearly between these patterns.

All D_q curves can be interpreted in terms of q , a parameter that modulates the weight of abundant and rare species in the distribution. D_q for positive q reflects more abundant species or dominance patterns in SAD, while D_q for negative q represents rare species patterns. An alternative way to analyze D_q would be to split species into ranges of abundances and calculate D_0^{AD} or D_1^{RS} . This was done for biomass and forest height spatial analysis (Seuront & Spilmont 2002; Kellner & Asner 2009), but for species distributions it has several drawbacks. First, the species' spatial distribution is analyzed as a whole, and it is quite possible that the complete set of species fits very well but one or more single species do not (Šizling & Storch 2004). Second, rare species represent a few points in space, and thus the estimation of D_q will have a high uncertainty. And third, the theory developed for D_q would not be valid (Harte 2001).

In neutral models, the SAR exponent depends on speciation rate (in this case migration from a metacommunity), dispersal distance, and local community size (Chave *et al.* 2002; Rosindell & Cornell 2007; Cencini *et al.* 2012). I did not expect to find high statistical power using the SAR exponent (D_0^{AD}) because I did not vary migration, dispersal and did not made comparisons between different community sizes. But I found high type I error rates for D_0^{AD} and the information dimension D_1^{RS} . This means that the statistical methods should be improved, applying a correction for autocorrelation to lower type I errors, and also a greater number of boxes should be used to increase power. In most cases, a range of different D_q values exists, meaning that the distribution is a multifractal (Stanley & Meakin 1988) and thus will not be well described by only one generalized dimension. To compare communities, D_q^{RS} and D_q^{AD} represent an improvement over comparisons made with only one dimension like SAR exponent or information dimension.

The species abundance distribution SAD is the most studied biodiversity pattern in ecology, but it is generally studied at one scale. Here I used the whole simulation area, and at this scale the power of SAD is comparable to that of generalized dimensions. Several studies regard SAD as not very informative because many different models can produce the same patterns, but in my simulations SAD could differentiate models quite well,

except for low species communities where its power was low. Generally, the performance of SAD depends on the number of species used in the comparison. When species are around 100, SAD comparison is the only method that can detect differences between very similar neutral communities.

In comparing between competitive hierarchical communities, the number of species was relatively low, and SAD and D_q^{SAD} had a low power, but D_q^{SRS} retained a high power. This highlights the ability of D_q^{SRS} to detect differences in spatial patterns of rare species. Spatial pattern is interdependent with the shape of SAD; for hierarchical communities there are few dominant species that form patches with size similar to the simulation area, and rare species are scattered. This pattern is enhanced by SRS, and thus different communities can be detected with high power. For neutral communities the SAD is more equitable and there are more species with enough abundances to form species clusters (I do not call them patches because species are intermingled). D_q^{SAD} and D_q^{SRS} thus have a high power to detect differences between neutral communities except where they are very similar. The advantage of D_q over SAD it is that the power should be improved by using a greater number of q values, and this possibility should be the subject of future studies.

When the communities compared had between 3 and 11 species, SAD had low power, D_q^{SRS} had a type I error slightly higher than α , and the D_q^{SAD} type I error was higher. To improve this the number of q used for comparison should be greater than that used in this work ($n=21$).

In summary, D_q^{SRS} always had better fits than D_q^{SAD} and can be applied in all the cases simulated here. It maintained a high power comparing hierarchical communities when the other methods failed. SADs also exhibit good performance with the exceptions already mentioned, although a better approach could be to use both D_q^{SRS} , SAD, and perhaps add other patterns (Münkemüller *et al.* 2012). This new macroecological metric could be a valuable addition to the already established ones and should be used in the study of the scaling of SAD (Borda-de-Água *et al.* 2012; Rosindell & Cornell 2013).

Acknowledgments

I am grateful to the National University of General Sarmiento for financial support, to Graeme Ruxton and Ben Bond-Lamberty for their English revisions.

References

- Anand, M. & Langille, A. (2010). A model-based method for estimating effective dispersal distance in tropical plant populations. *Theoretical Population Biology*, **77**, 219–226. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0040580910000158>
- Beck, C. (1990). Upper and lower bounds on the Renyi dimensions and the uniformity of multifractals. *Physica D: Nonlinear Phenomena*, **41**, 67–78. Retrieved from <http://www.sciencedirect.com/science/article/pii/016727899090028N>
- Borda-de-Água, L., Borges, P.A.V., Hubbell, S.P. & Pereira, H.M. (2012). Spatial scaling of species abundance distributions. *Ecography*, **35**, 549–556. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0587.2011.07128.x/abstract>
- Borda-de-Água, L., Hubbell, S.P. & He, F. (2007). Scaling biodiversity under neutrality. *Scaling biodiversity*, 347–375.
- Borda-de-Água, L., Hubbell, S.P. & McAllister, M. (2002). Species-Area Curves, Diversity Indices, and Species Abundance Distributions: A Multifractal Analysis. *American Naturalist*, **159**, 138–155.
- Cencini, M., Pigolotti, S. & Muñoz, M.A. (2012). What Ecological Factors Shape Species-Area Curves in Neutral Models? *PLoS ONE*, **7**, e38232. Retrieved from <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0038232>
- Chappell, D. & Scalo, J. (2001). Multifractal Scaling, Geometrical Diversity, and Hierarchical Structure in the Cool Interstellar Medium. *Astrophysical Journal*, **551**, 712–729.
- Chave, J., Muller-Landau, H.C. & Levin, S.A. (2002). Comparing Classical Community Models: Theoretical Consequences for Patterns of Diversity. *American Naturalist*, **159**, 1–23.
- Condit, R., Pitman, N., Jr., E.G.L., Chave, J., Terborgh, J., Foster, R.B., Núñez, P., Aguilar, S., Valencia, R., Villa, G., Muller-Landau, H.C., Losos, E. & Hubbell, S.P. (2002). Beta-Diversity in Tropical Forest Trees. *Science*, **295**, 666–669. Retrieved from <http://www.sciencemag.org/content/295/5555/666>
- Crawley, M.J. & Harral, J.E. (2001). Scale dependence in plant biodiversity. *Science*, **291**, 864–868.
- Dal Bello, M., Maggi, E., Rindi, L., Capocchi, A., Fontanini, D., Sanz-Lazaro, C. & Benedetti-Cecchi, L. (2014). Multifractal spatial distribution of epilithic microphytobenthos on a Mediterranean rocky shore. *Oikos*, n/a—n/a. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/oik.01503/abstract>
- Edgington, E. (1995). *Randomization Tests*, 3rd. Edition, 3rd.n. CRC Press, New York.

- 1 Etienne, R.S. (2007). A neutral sampling formula for multiple samples and an ‘exact’ test of neutrality.
2 *Ecology Letters*, **10**, 608–618. Retrieved from [http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2007.](http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2007.01052.x/abstract)
3 [01052.x/abstract](http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2007.01052.x/abstract)
- 4 Evertsz, C.J.G. & Mandelbrot, B.B. (1992). Multifractal measures. *Chaos and Fractals: New Frontiers of*
5 *Science* (eds H.O. Peitgen, H. Jurgens & D. Saupe), pp. 921–953. Springer, New York.
- 6 Feigelson, E.D. & Babu, G.J. (2012). *Modern statistical methods for astronomy: with r applications*.
7 Cambridge University Press.
- 8 Fisher, R.A., Corbet, A.S. & Williams, C.B. (1943). The Relation Between the Number of Species and the
9 Number of Individuals in a Random Sample of an Animal Population. *Journal of Animal Ecology*, **12**, 42–58.
10 Retrieved from <http://www.jstor.org/stable/1411>
- 11 Gamarra, J.G.P. (2005). Metapopulations in multifractal landscapes: on the role of spatial aggrega-
12 tion. *Proceedings of the Royal Society London Series B*, **272**, 1815–1822. Retrieved from [http://rsph.](http://rsph.royalsocietypublishing.org/content/272/1574/1815)
13 [royalsocietypublishing.org/content/272/1574/1815](http://rsph.royalsocietypublishing.org/content/272/1574/1815)
- 14 Grassberger, P. (1983). Generalized dimensions of strange attractors. *Physics Letters A*, **97**, 227–230.
- 15 Gravel, D., Canham, C.D., Beaudet, M. & Messier, C. (2006). Reconciling niche and neutrality: the continuum
16 hypothesis. *Ecology letters*, **9**, 399–409. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16623725>
- 17 Hankin, R.K.S. (2007). Introducing untb, an R package for simulating ecological drift under the unified
18 neutral theory of biodiversity. *Journal of Statistical Software*, **22**.
- 19 Harte, D. (2001). *Multifractals: Theory and Applications*. Chapman & Hall/CRC.
- 20 Hentschel, H.G.E. & Procaccia, I. (1983). The infinite number of generalized dimensions of fractals and
21 strange attractors. *Physica D*, **8**, 435–444.
- 22 Hill, M.O. (1973). Diversity and Evenness: A Unifying Notation and Its Consequences. *Ecology*, **54**, 427–432.
23 Retrieved from <http://www.jstor.org/stable/1934352>
- 24 Hubbell, S.P. (2001). *The unified neutral theory of biodiversity and biogeography*. Princeton University Press.
25 Retrieved from <http://books.google.com.ar/books?id=EIQpFBu84NoC>
- 26 Kellner, J.R. & Asner, G.P. (2009). Convergent structural responses of tropical forests to diverse disturbance
27 regimes. *Ecology Letters*, **12**, 887–897. Retrieved from [http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.](http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2009.01345.x/abstract)
28 [2009.01345.x/abstract](http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2009.01345.x/abstract)
- 29 Kirkpatrick, L.A. & Weishampel, J.F. (2005). Quantifying spatial structure of volumetric neutral

- 1 models. *Ecological Modelling*, **186**, 312–325. Retrieved from [http://linkinghub.elsevier.com/retrieve/pii/](http://linkinghub.elsevier.com/retrieve/pii/S0304380005000554)
- 2 [S0304380005000554](http://linkinghub.elsevier.com/retrieve/pii/S0304380005000554)
- 3 Kutner, M., Nachtsheim, C., Neter, J. & Li, W. (2005). *Applied Linear Statistical Models*, 5th ed.n.
- 4 McGraw-Hill/Irwin.
- 5 Laurie, H. & Perrier, E. (2011). Beyond species area curves: application of a scale-free measure for spatial
- 6 variability of species richness. *Oikos*, **120**, 966–978. Retrieved from [http://dx.doi.org/10.1111/j.1600-0706.](http://dx.doi.org/10.1111/j.1600-0706.2010.19134.x)
- 7 [2010.19134.x](http://dx.doi.org/10.1111/j.1600-0706.2010.19134.x)
- 8 Legendre, P., Dale, M.R.T., Fortin, M.J., Gurevitch, J., Hohn, M. & Myers, D. (2002). The consequences of
- 9 spatial structure for the design and analysis of ecological field surveys. *Ecography*, **25**, 601–615.
- 10 Liebovitch, L.S. (1998). *Fractals and Chaos Simplified for the Life Sciences*. Oxford University Press, Oxford.
- 11 McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benetcha, H.K., Dornelas, M., Enquist,
- 12 B.J., Green, J.L., He, F., Hurlbert, A.H., Magurran, A.E., Marquet, P.A., Maurer, B.A., Ostling, A., Soykan,
- 13 C.U., Ugland, K.I. & White, E.P. (2007). Species abundance distributions: moving beyond single prediction
- 14 theories to integration within an ecological framework. *Ecology Letters*, **10**, 995–1015. Retrieved from
- 15 <http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2007.01094.x/abstract>
- 16 Molofsky, J. & Bever, J.D. (2004). A New Kind of Ecology? *BioScience*, **54**, 440–446. Retrieved from
- 17 <http://bioscience.oxfordjournals.org/content/54/5/440.short>
- 18 Münkemüller, T., Bello, F. de, Meynard, C.N., Gravel, D., Lavergne, S., Mouillot, D., Mouquet, N. & Thuiller,
- 19 W. (2012). From diversity indices to community assembly processes: a test with simulated data. *Ecography*,
- 20 **35**, 468–480. Retrieved from <http://dx.doi.org/10.1111/j.1600-0587.2011.07259.x>
- 21 Newman, M.E.J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, **46**, 323–351.
- 22 Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/00107510500052444>
- 23 Preston, F.W. (1960). Time and Space and the Variation of Species. *Ecology*, **41**, pp. 612–627. Retrieved
- 24 from <http://www.jstor.org/stable/1931793>
- 25 R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for
- 26 Statistical Computing, Vienna, Austria. Retrieved from <http://www.r-project.org/>
- 27 Renyi, A. (1970). *Probability Theory*. North-Holland, Amsterdam.
- 28 Ricotta, C. (2000). From theoretical ecology to statistical physics and back: self-similar landscape metrics as
- 29 a synthesis of ecological diversity and geometrical complexity. *Ecological Modelling*, **125**, 245–253.

- 1 Rosindell, J. & Cornell, S.J. (2007). Species–area relationships from a spatially explicit neutral model in an
2 infinite landscape. *Ecology Letters*, **10**, 586–595. Retrieved from [http://onlinelibrary.wiley.com/doi/10.1111/
3 j.1461-0248.2007.01050.x/abstract](http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2007.01050.x/abstract)
- 4 Rosindell, J. & Cornell, S.J. (2013). Universal scaling of species-abundance distributions across multiple
5 scales. *Oikos*, **122**, 1101–1111. Retrieved from [http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0706.2012.
6 20751.x/abstract](http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0706.2012.20751.x/abstract)
- 7 Rosindell, J., Hubbell, S.P. & Etienne, R.S. (2011). The Unified Neutral Theory of Biodiversity and
8 Biogeography at Age Ten. *Trends in ecology & evolution*, **26**, 340–348. Retrieved from [http://www.ncbi.nlm.
9 nih.gov/pubmed/21561679](http://www.ncbi.nlm.nih.gov/pubmed/21561679)
- 10 Saravia, L.A. (2014). mfSBA: Multifractal analysis of spatial patterns in ecological communities [v2; ref
11 status: indexed, <http://f1000r.es/347>]. *F1000Research*, **3**, 14. Retrieved from [http://f1000research.com/
12 articles/3-14/v2](http://f1000research.com/articles/3-14/v2)
- 13 Saravia, L.A., Giorgi, A. & Momo, F. (2012a). Multifractal growth in periphyton communities. *Oikos*, **121**,
14 1810–1820. Retrieved from <http://doi.wiley.com/10.1111/j.1600-0706.2011.20423.x>
- 15 Saravia, L.A., Giorgi, A. & Momo, F.R. (2012b). Multifractal Spatial Patterns and Diversity in an Ecological
16 Succession. *PLoS ONE*, **7**, e34096. Retrieved from <http://dx.plos.org/10.1371/journal.pone.0034096>
- 17 Scheuring, I. & Riedi, R.H. (1994). Application of multifractals to the analysis of vegetation pattern. *Journal
18 of Vegetation Science*, **5**, 489–469.
- 19 Scholz, F.W. & Stephens, M.A. (1987). K-Sample Anderson–Darling Tests. *Journal of the American Statistical
20 Association*, **82**, 918–924. Retrieved from <http://dx.doi.org/10.1080/01621459.1987.10478517>
- 21 Scholz, F. & Zhu, A. (2012). *kSamples: K-Sample Rank Tests and their Combinations*. Retrieved from
22 <http://cran.r-project.org/package=kSamples>
- 23 Seuront, L. (2009). *Fractals and Multifractals in Ecology and Aquatic Science*. Taylor & Francis.
- 24 Seuront, L. & Spilmont, N. (2002). Self-organized criticality in intertidal microphytobenthos patch patterns.
25 *Physica A*, **313**, 513–539.
- 26 Seuront, L. & Stanley, H.E. (2014). Anomalous diffusion and multifractality enhance mating encounters
27 in the ocean. *Proceedings of the National Academy of Sciences*, **111**, 2206–2211. Retrieved from [http:
28 //www.pnas.org/content/111/6/2206](http://www.pnas.org/content/111/6/2206)
- 29 Stanley, H. & Meakin, P. (1988). Multifractal phenomena in physics and chemistry. *Nature*, **335**, 405–409.

- 1 Šizling, A.L. & Storch, D. (2004). Power-law species–area relationships and self-similar species distributions
2 within finite areas. *Ecology Letters*, **7**, 60–68. Retrieved from [http://onlinelibrary.wiley.com/doi/10.1046/j.](http://onlinelibrary.wiley.com/doi/10.1046/j.1461-0248.2003.00549.x/abstract)
3 [1461-0248.2003.00549.x/abstract](http://onlinelibrary.wiley.com/doi/10.1046/j.1461-0248.2003.00549.x/abstract)
- 4 Šizling, A.L., Kunin, W.E., Šizlingová, E., Reif, J. & Storch, D. (2011). Between Geometry and Biology: The
5 Problem of Universality of the Species–Area Relationship. *The American Naturalist*, **178**, 602–611. Retrieved
6 from <http://www.jstor.org/stable/10.1086/662176>
- 7 Tilman, D. (1994). Competition and biodiversity in spatially structured habitats. *Ecology*, **75**, 2–16.
- 8 Tjørve, E. (2003). Shapes and functions of species–area curves: a review of possible models. *Journal of*
9 *Biogeography*, **30**, 827–835. Retrieved from <http://dx.doi.org/10.1046/j.1365-2699.2003.00877.x>
- 10 Wei, S.-G., Li, L., Huang, Z.-L., Ye, W.-H., Gong, G.-Q., Zhou, X.-Y. & Lian, J.-Y. (2013). Multifractal
11 analysis of diversity scaling laws in a subtropical forest. *Ecological Complexity*, **13**, 1–7. Retrieved from
12 <http://www.sciencedirect.com/science/article/pii/S1476945X11000729>
- 13 White, E.P., Ernest, S.K.M., Adler, P.B., Hurlbert, A.H. & Lyons, S.K. (2010). Integrating spatial and
14 temporal approaches to understanding species richness. *Philosophical Transactions of the Royal Society*
15 *of London. Series B*, **365**, 3633–43. Retrieved from [http://www.pubmedcentral.nih.gov/articlerender.fcgi?](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2982007/&tool=pmcentrez/&rendertype=abstract)
16 [artid=2982007/&tool=pmcentrez/&rendertype=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2982007/&tool=pmcentrez/&rendertype=abstract)
- 17 Yakimov, B.N., Bossuyt, B., Iudin, D.I. & Gelashvili, D.B. (2008). Multifractal diversity–area relationship at
18 small scales in dune slack plant communities. *Oikos*, **117**, 33–39. Retrieved from [http://dx.doi.org/10.1111/j.](http://dx.doi.org/10.1111/j.2007.0030-1299.16121.x)
19 [2007.0030-1299.16121.x](http://dx.doi.org/10.1111/j.2007.0030-1299.16121.x)
- 20 Yakimov, B.N., Gelashvili, D.B., Solntsev, L.A., Iudin, D.I. & Rozenberg, G.S. (2014). Nonconcavity of
21 mass exponents’ spectrum in multifractal analysis of community spatial structure: The problem and possible
22 solutions. *Ecological Complexity*, **20**, 11–22. Retrieved from [http://www.sciencedirect.com/science/article/](http://www.sciencedirect.com/science/article/pii/S1476945X14000737)
23 [pii/S1476945X14000737](http://www.sciencedirect.com/science/article/pii/S1476945X14000737)
- 24 Yakimov, B.N., Iudin, D.I., Solntsev, L.A. & Gelashvili, D.B. (2013). Multifractal analysis of neutral
25 community spatial structure. *Journal of Theoretical Biology*. Retrieved from [http://www.sciencedirect.com/](http://www.sciencedirect.com/science/article/pii/S0022519313005092)
26 [science/article/pii/S0022519313005092](http://www.sciencedirect.com/science/article/pii/S0022519313005092)
- 27 Zhang, Y., Ma, K., Anand, M. & Fu, B. (2006). Do generalized scaling laws exist for species abundance
28 distribution in mountains? *Oikos*, **115**, 81–88. Retrieved from [http://dx.doi.org/10.1111/j.2006.0030-1299.](http://dx.doi.org/10.1111/j.2006.0030-1299.15021.x)
29 [15021.x](http://dx.doi.org/10.1111/j.2006.0030-1299.15021.x)

Manuscript tables

Table 1: Methods used to calculate the power to compare simulated communities with different degree of neutrality.

Method	Name - Meaning
SAD	Species abundance distribution.
D_q^{SAD}	Generalized dimension spectra based on SAD: Characterize the scaling of species abundances in space.
D_q^{SRS}	Generalized dimension spectra based on SRS: Characterize the scaling of the spatial distribution of the ranks of species (SRS) derived from SAD.
D_0^{SAD}	The power exponent of the species-area relationship. This is part of D_q^{SAD} and characterize the scaling of richness.
D_1^{SRS}	Information dimension of SRS. This is part of D_q^{SRS} and characterize the scaling of Shannon diversity index calculated on the spatial distribution of species ranks.

Table 2: Parameters values used in the simulations of the neutral-hierarchical model

Side	No. Species	μ	d	m	ρ
256	11	0.2	25	0.001	1
512	86				0.1
	341				0.01
					0.001
					0

Table 3: Power and Type I error rate for T-test comparison of a single dimension of the generalized spectra: the SAR exponent (D_0^{SAD}) and information dimension of the species rank surface (D_1^{SRS}). The test use the standard deviation obtained in the regressions to fit generalized dimensions. The number of comparisons to calculate the power is n=25000, and for type I error n=6125.

Side	Metacommunity No. Species	Mean No. Species	Type	Power	Type I Error
512	11	5.96	D_1^{SRS}	0.512	0.434
			D_0^{SAD}	0.498	0.494
	86	36.54	D_1^{SRS}	0.521	0.430
			D_0^{SAD}	0.445	0.426
	341	111.31	D_1^{SRS}	0.497	0.342
			D_0^{SAD}	0.494	0.436
256	11	5.90	D_1^{SRS}	0.491	0.408
			D_0^{SAD}	0.471	0.424
	86	32.27	D_1^{SRS}	0.501	0.447
			D_0^{SAD}	0.474	0.388
	341	76.57	D_1^{SRS}	0.490	0.389
			D_0^{SAD}	0.443	0.363

Table 4: Power and Type I error rate of Anderson-Darling statistic to test hypothesis of differences in: species abundance distributions (SAD), generalized dimension based on SAD (D_q^{SAD}) and generalized dimension based on SRS (D_q^{SRS}). The power is calculated testing communities with different ρ and type I error is calculated for communities with the same ρ . The p-values were estimated using 1000 randomizations. The number of points used for SAD comparisons is the number of species found in the communities. The number of points used for multifractal spectra correspond to the q in the range -10 to 10 ($n=21$), according to the following set $q=\{-10,-8,-6,-4,-3,-2.5,-2,-1.5,-1,-0.5,0,0.5,1,1.5,2,2.5,3,4,6,8,10\}$. The number of comparisons for the power calculations were $n=25000$ except for SAD with $side=500$ & metacommunity species=11, where some comparison with less than 3 species were skipped ($n=23800$). For type I error the comparisons were $n=6125$, and the same exception applies ($n=5846$).

Side	Metacommunity No. Species	Mean No. Species	Type	Power	Type I Error
512	11	5.96	SAD	0.115	0.025
			D_q^{SRS}	0.720	0.102
			D_q^{SAD}	0.568	0.212
	86	36.54	SAD	0.697	0.009
			D_q^{SRS}	0.680	0.014
			D_q^{SAD}	0.616	0.011
	341	111.31	SAD	0.830	0.039
			D_q^{SRS}	0.688	0.000
			D_q^{SAD}	0.609	0.017
256	11	5.90	SAD	0.175	0.000
			D_q^{SRS}	0.654	0.068
			D_q^{SAD}	0.704	0.204
	86	32.27	SAD	0.675	0.019
			D_q^{SRS}	0.657	0.025
			D_q^{SAD}	0.613	0.027
	341	76.57	SAD	0.799	0.035
			D_q^{SRS}	0.670	0.030
			D_q^{SAD}	0.610	0.048

Manuscript Figures

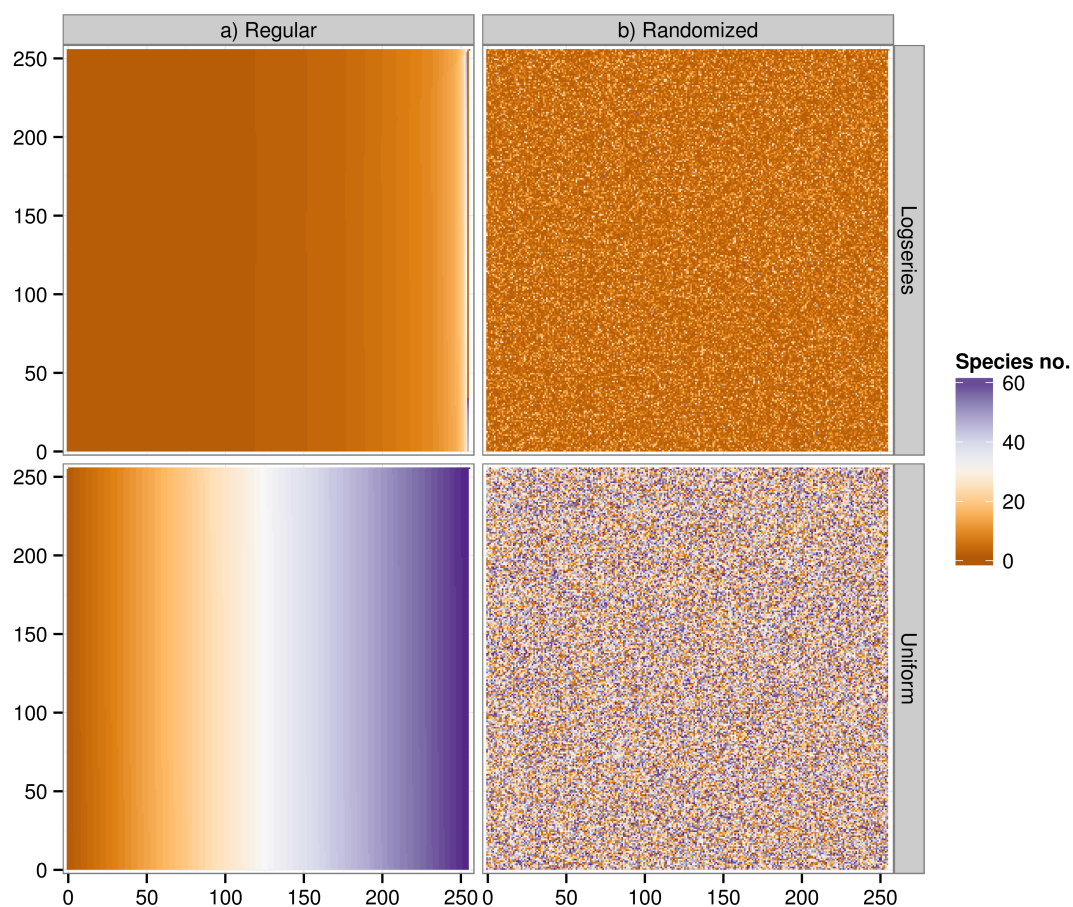


Figure 1: Spatial patterns generated with logseries and uniform species abundance distribution (all species have the same density) with 64 species and a grid with side=256. a) Regular: species are distributed in vertical bands. b) Randomized: the position of species are distributed at random in space.

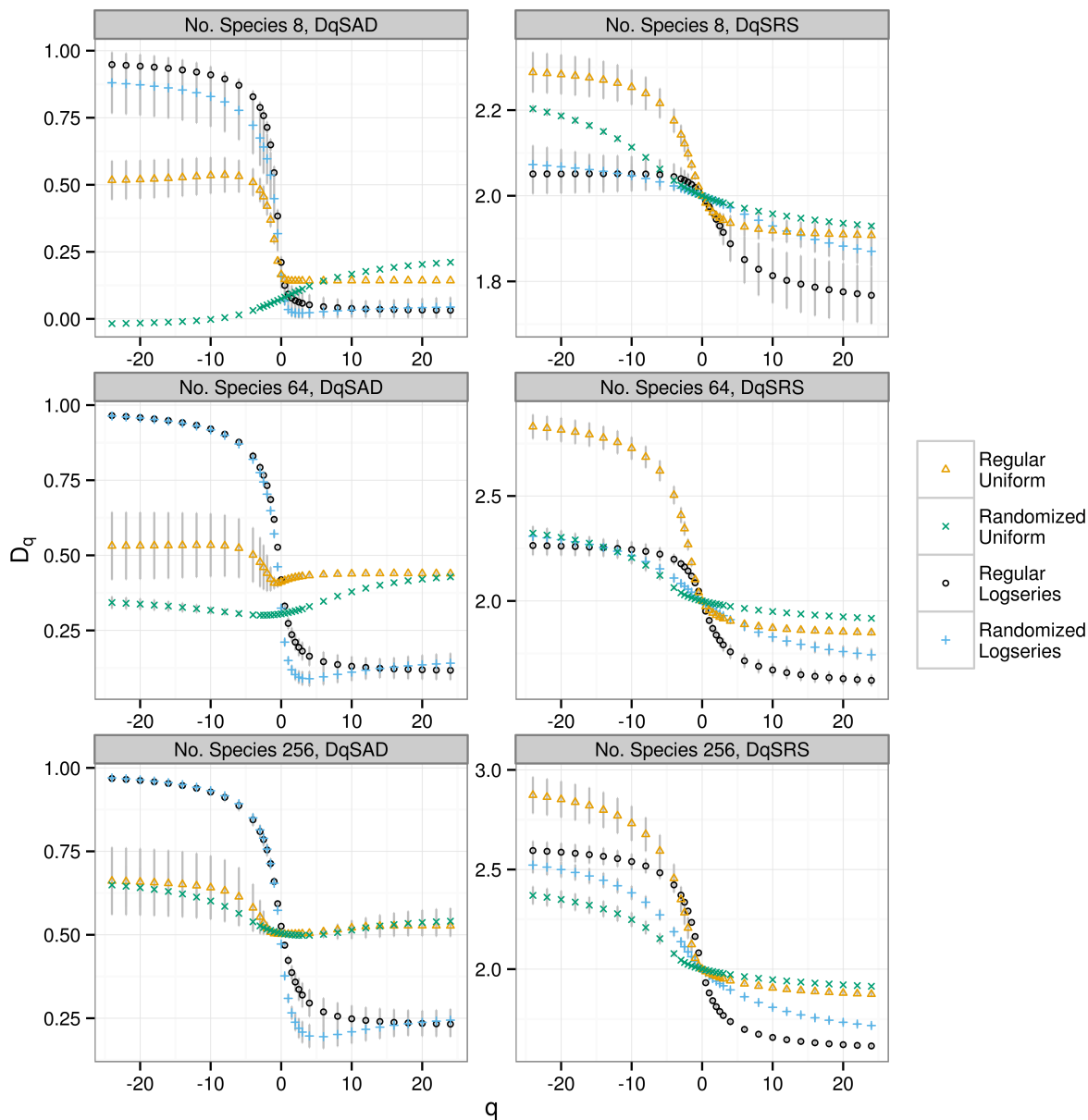


Figure 2: Generalized dimension spectra D_q of simulated species spatial patterns. The points are means of 10 simulated patterns using a spatial grid of side=256. A logseries or uniform species abundance distribution were used, with 8,64 and 256 species. Two forms of generalized dimensions were estimated: DqSRS, from species rank surface D_q^{SRS} . DqSAD, estimated from species abundance distribution D_q^{SAD} . I use two spatial patterns: Regular, the species are distributed in vertical bands, Randomized the spatial distribution of species was randomized.

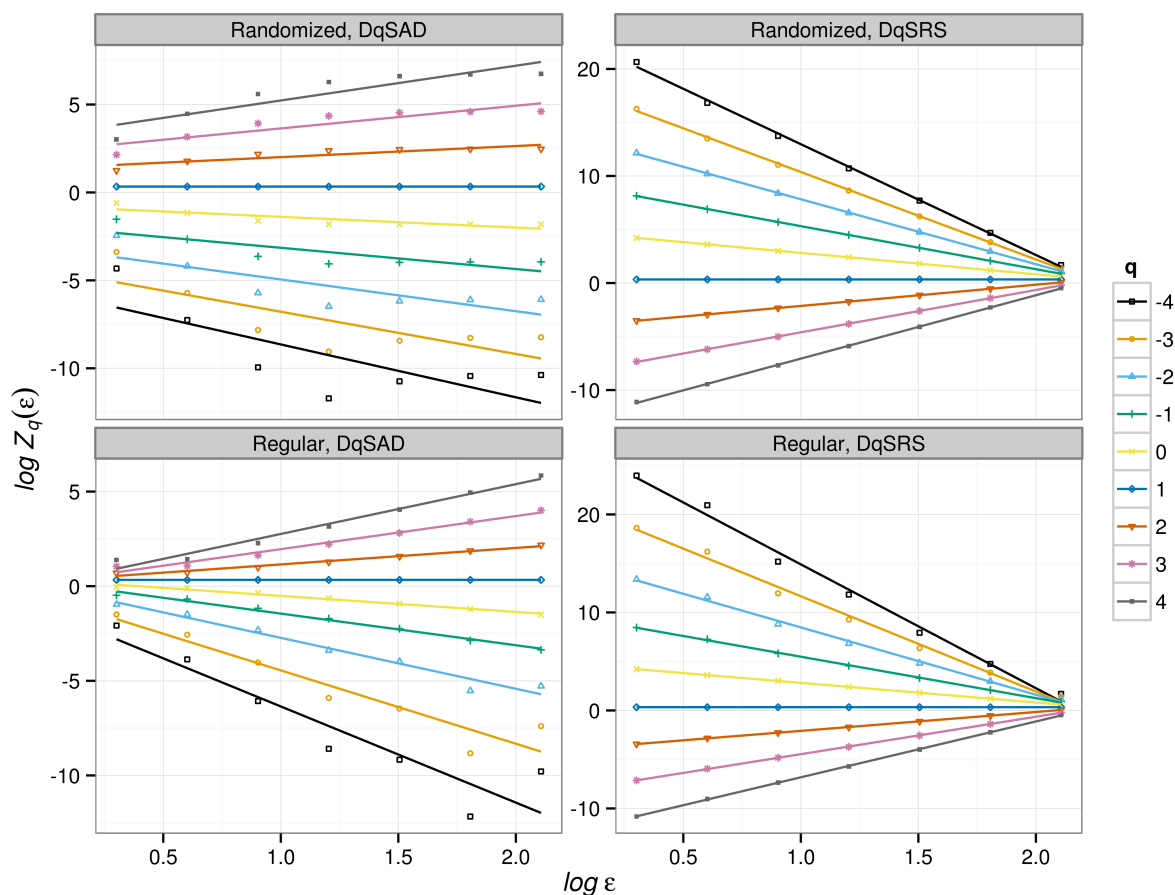


Figure 3: Linear fit of generalized dimension (D_q) estimation from simple species spatial patterns with 64 species and a uniform abundance distribution. The spatial grid has a side=256 sites and two different spatial patterns: Regular, a regular spatial pattern with species distributed in vertical bands of equal width. Randomized, the positions of species in the grid are randomized. Two kinds of generalized dimension were estimated: DqSRS corresponds the fit of D_q^{SRS} (see text) and DqSAD is the fit from the estimation of D_q^{SAD} (see text). $Z_q(\epsilon)$ corresponds to the partition function calculated for a box with side ϵ , q is the moment order.

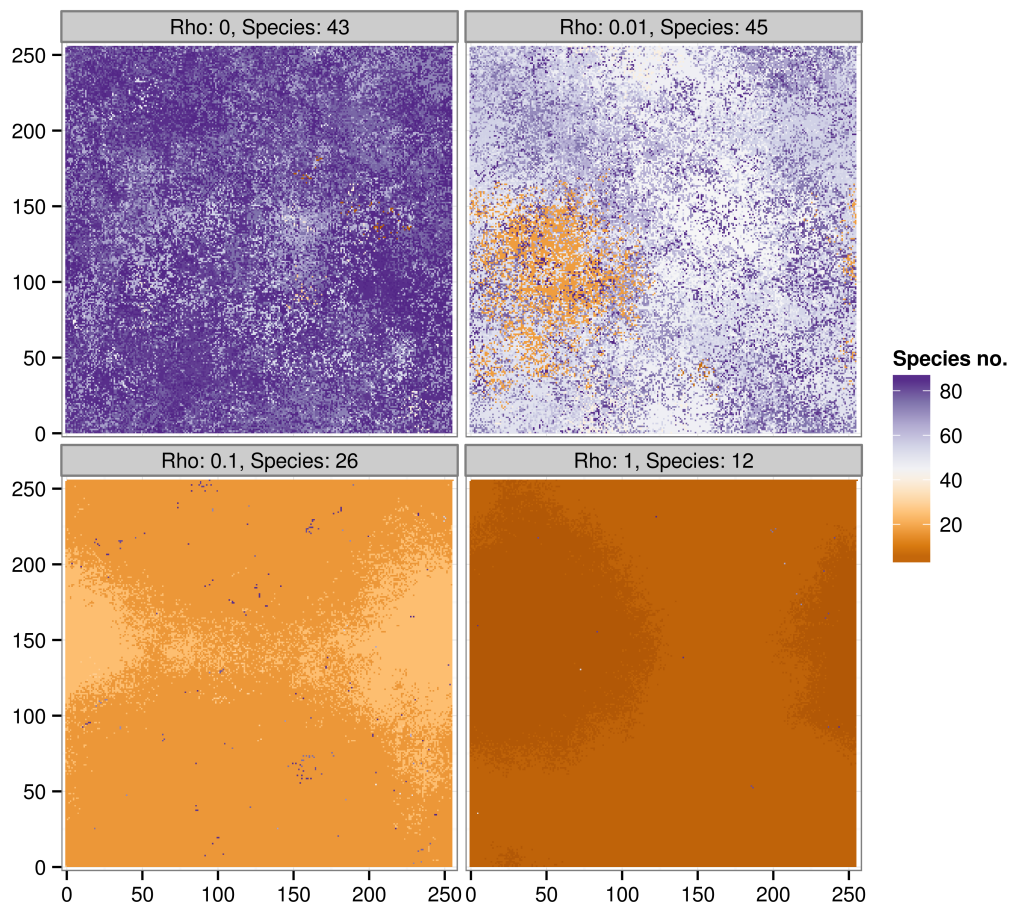


Figure 4: Spatial patterns generated with a spatial neutral/hierarchical model. Rho is the parameter that determines the degree of neutrality. When this parameter is 0 the model is completely neutral and there is no competitive replacement of species. When ρ is 1 competitive superior species always replaces inferior ones and the model is completely hierarchical. *Species* is the number of species actually present plot. The simulations use a metacommunity with a logseries abundance distribution with 86 species and a simulation grid side=256.

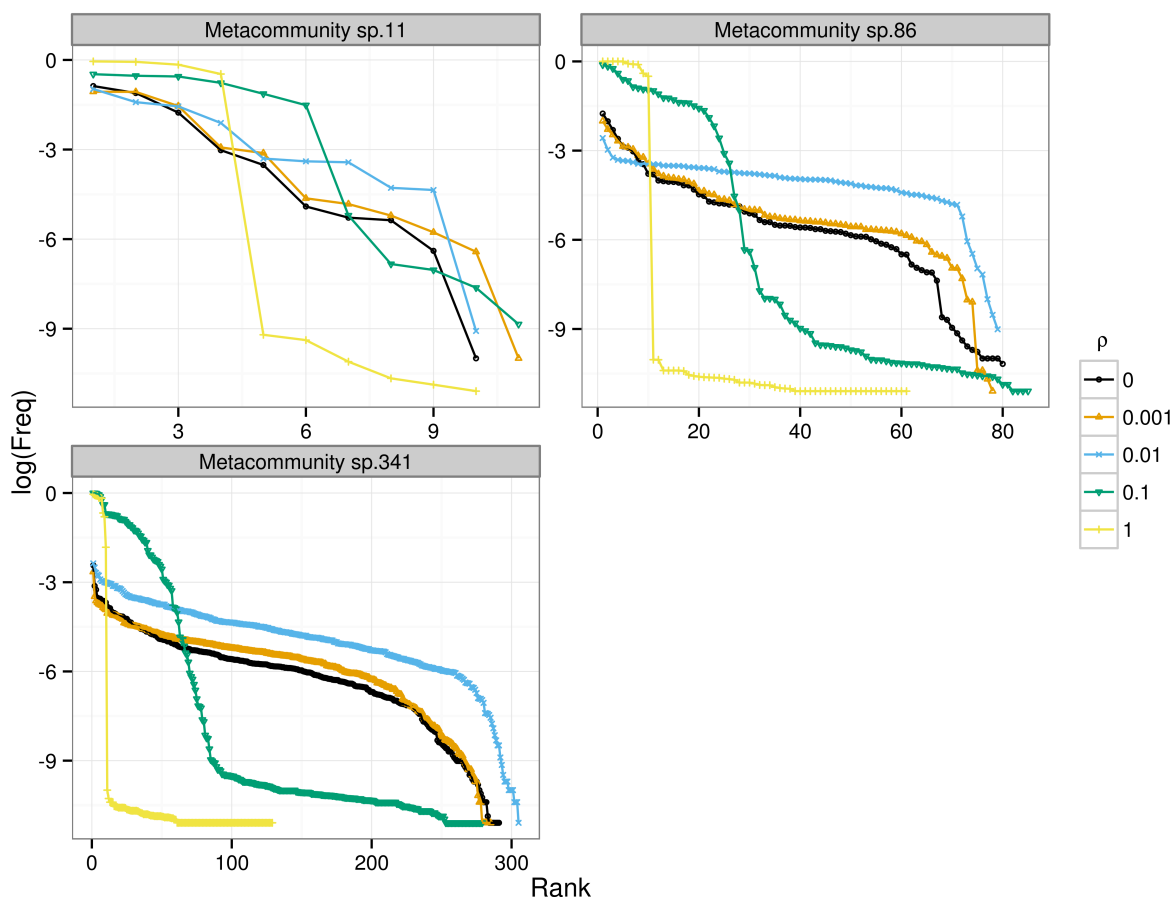


Figure 5: Rank abundance diagrams of communities patterns generated with a spatial neutral/hierarchical model with different number of species in the metacommunity (labeled in each subfigure). ρ is the parameter that determines the degree of neutrality. When this parameter is 0 the model is completely neutral and there is no competitive replacement of species. When ρ is 1 competitive superior species always replaces inferior ones and the model is completely hierarchical. The simulations use a metacommunity with a logseries abundance distribution with 11, 86 and 341 species and a simulation grid side=256, the other parameters used were MortalityRate=0.2, DispersalDistance=0.4 (2.5 grid units), ColonizationRate=0.001. The ranks were calculated with averages of species densities over 50 simulations.

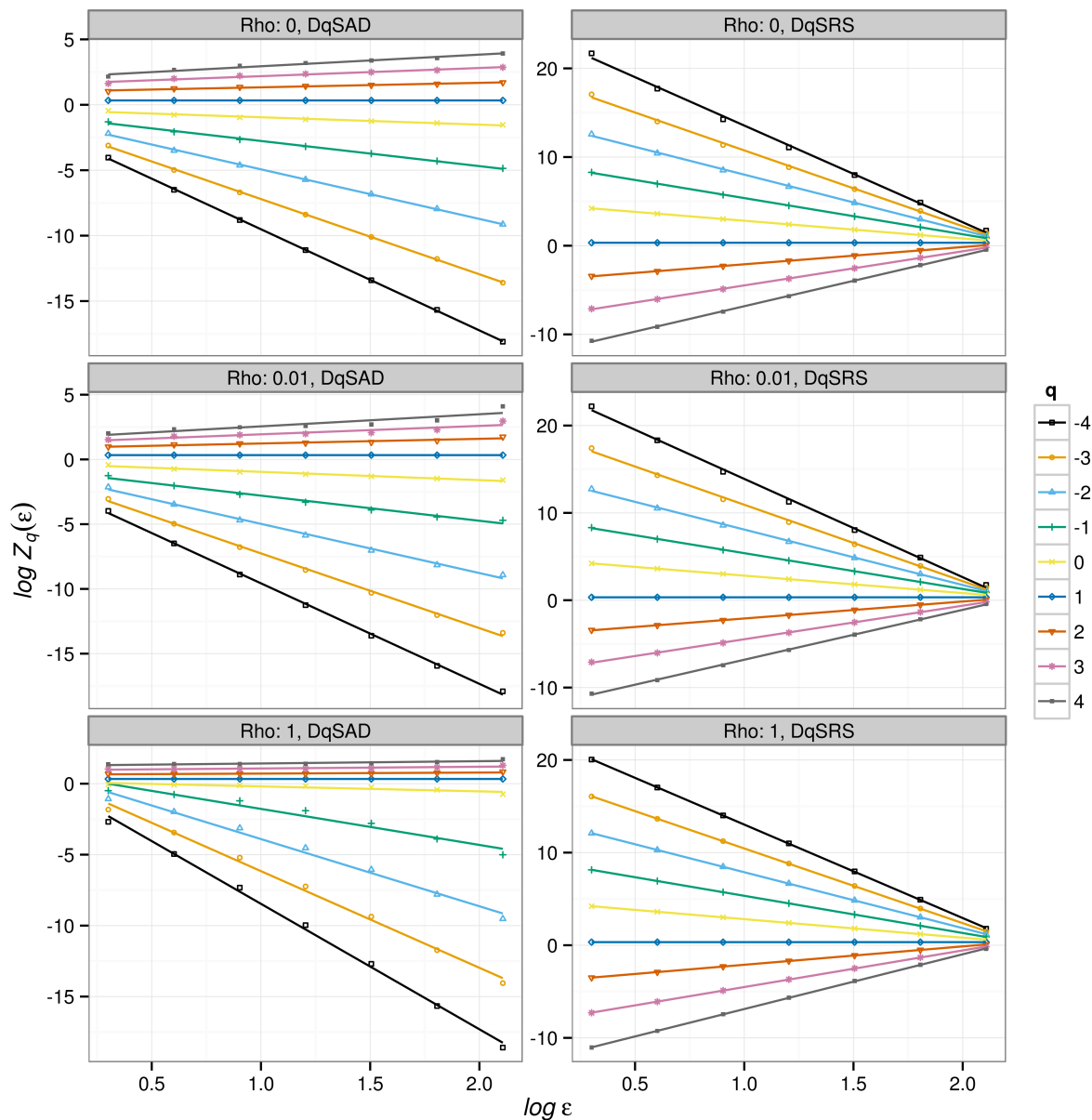


Figure 6: Linear fit of generalized dimension (D_q) estimation from spatial patterns generated with a spatial neutral/hierarchical model. Rho is the parameter that determines the degree of neutrality. When this parameter is 0 the model is completely neutral and there is no competitive replacement of species. When ρ is 1 competitive superior species always replaces inferior ones and the model is completely hierarchical. $Species$ is the number of species actually present plot. The simulations use a metacommunity with a logseries abundance distribution with 86 species and a simulation grid side=256 sites.

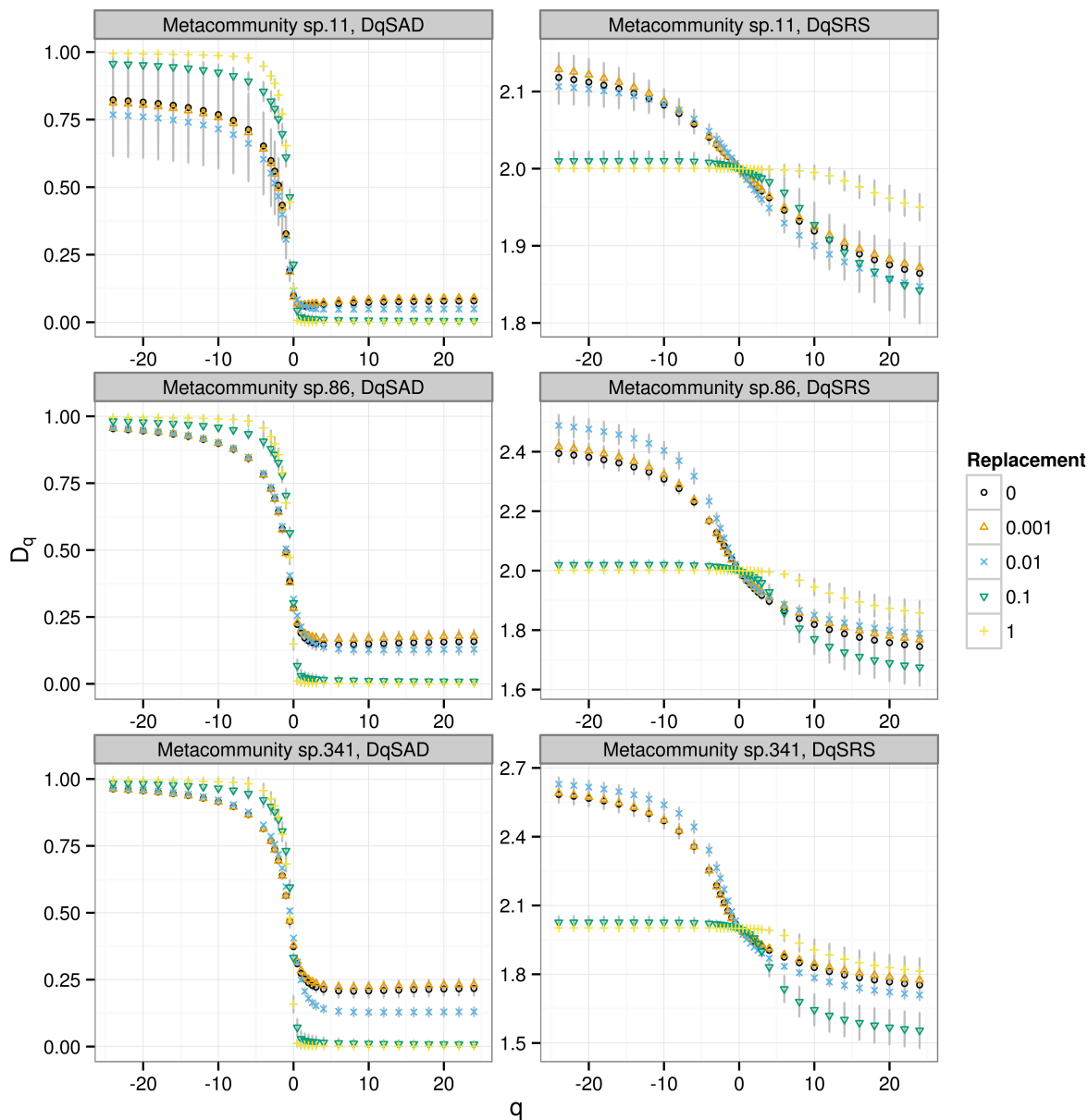


Figure 7: Generalized dimension spectra D_q of spatial patterns generated with a spatial neutral/hierarchical model. *Replacement* is the parameter ρ , that determines the degree of neutrality. When this parameter is 0 the model is completely neutral and there is no competitive replacement of species. When ρ is 1 competitive superior species always replaces inferior ones and the model is completely hierarchical. The points are means and vertical lines are standard deviation of 50 simulated patterns. Simulations use a metacommunity with a logseries abundance distribution with 11,86 and 341 species. The simulation grid side is 256, and the other parameters are given in the main text.

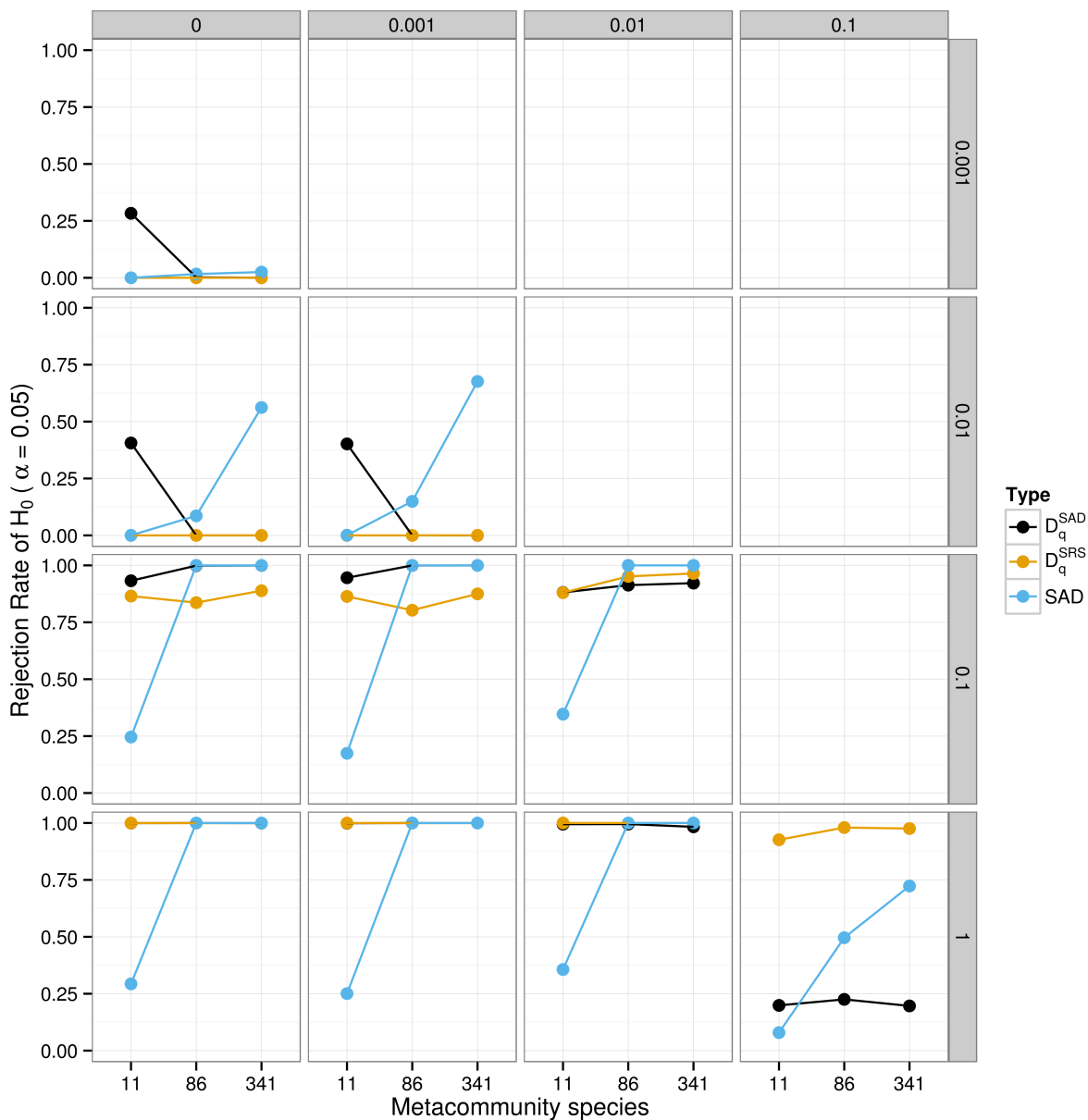


Figure 8: Power of the Anderson-Darling test for the hypothesis of differences between simulated neutral/hierarchical communities. The test uses generalized dimensions curves calculated from SAD (D_q^{SAD}), generalized dimensions calculated from the species rank surfaces (D_q^{SRS}) and the species abundance distributions (SAD). The compared communities differ only in parameter ρ (across panels) that determines the degree of neutrality/hierarchy. The number of comparisons to calculate the frequency is 2500 in all cases. Simulations use a metacommunity with a logseries abundance distribution with 11,86 and 341 species; a grid side of 256 sites, the other parameters are given in the main text.