



WEB TOOL

REVISED

mfSBA: Multifractal analysis of spatial patterns in ecological communities [v2; ref status: indexed, <http://f1000r.es/347>]

Leonardo A. Saravia

Instituto de Ciencias Básicas, Universidad Nacional de General Sarmiento, Buenos Aires, Argentina

v2 First published: 16 Jan 2014, 3:14 (doi: [10.12688/f1000research.3-14.v1](https://doi.org/10.12688/f1000research.3-14.v1))
 Latest published: 07 Apr 2014, 3:14 (doi: [10.12688/f1000research.3-14.v2](https://doi.org/10.12688/f1000research.3-14.v2))

Abstract

Multifractals have been applied to characterize complex communities in a spatial context. They were developed for nonlinear systems and are particularly suited to capture multiplicative processes observed in ecological systems. Multifractals characterize variability in a scale-independent way within an experimental range. I have developed an open-source software package to estimate multifractals using a box-counting algorithm (available from <https://github.com/lsaravia/mfsba> and permanently available at doi: [10.5281/zenodo.8481](https://doi.org/10.5281/zenodo.8481)). The software is specially designed for two dimensional (2D) images such as the ones obtained from remote sensing, but other 2D data types can also be analyzed. Additionally I developed a new metric to analyze multispecies spatial patterns with multifractals: spatial rank surface, which is included in the software.

Open Peer Review

Referee Status:

	Invited Referees	
	1	2
version 2 published 07 Apr 2014	report	report
version 1 published 16 Jan 2014	report	report

- Yuxin Zhang**, Chinese Academy of Sciences China
- Edmund Hart**, National Ecological Observatory Network USA

Discuss this article

Comments (2)

Corresponding author: Leonardo A. Saravia (lsaravia@ungs.edu.ar)

How to cite this article: Saravia LA. **mfSBA: Multifractal analysis of spatial patterns in ecological communities [v2; ref status: indexed, <http://f1000r.es/347>]** *F1000Research* 2014, 3:14 (doi: [10.12688/f1000research.3-14.v2](https://doi.org/10.12688/f1000research.3-14.v2))

Copyright: © 2014 Saravia LA. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](https://creativecommons.org/licenses/by/4.0/) (CC0 1.0 Public domain dedication).

Grant information: The author(s) declared that no grants were involved in supporting this work.

Competing interests: No competing interests were disclosed.

First published: 16 Jan 2014, 3:14 (doi: [10.12688/f1000research.3-14.v1](https://doi.org/10.12688/f1000research.3-14.v1))

First indexed: 26 Jun 2014, 3:14 (doi: [10.12688/f1000research.3-14.v2](https://doi.org/10.12688/f1000research.3-14.v2))

REVISED Amendments from Version 1

I would like to thank the reviewers for their comments. In response to the referees I made available a pre-compiled version of the executables for Windows and a corrected version of the R scripts tested under that operative system (available at <https://github.com/lsaravia/mfsba/releases> and doi: [10.5281/zenodo.8481](https://doi.org/10.5281/zenodo.8481)). I added the description of the program to perform a randomization confidence envelope that is included in mfSBA. The following is a summary of the changes made to the manuscript:

- I have added some examples and clarifications to the introduction.
- I have explained that the SRS/MFA methodology is novel; I do not intend to make a comparison with other methods here, but present a description of the capabilities of the software. I also included an interpretation of multifractal spectra D_q based on Hill's generalized diversity.
- The limitations of mfSBA have been more clearly stated.
- I have included an explanation on why D_q 's are different for Figure 2, which also helps the interpretation of SRS.
- I have included a supplementary figure comparing a multispecies spatial distribution generated by a neutral model.

See referee reports

Introduction

Multifractals and fractals are related techniques mainly used in physics to characterize the scaling behavior of a system; they differ in that fractals look at the geometry of presence/absence patterns, while multifractals look at the arrangement of quantities such as population densities or biomass¹. Scaling laws are an emergent general feature of ecological systems, and there is no *a priori* reason that power laws apply to ecological communities. If they do apply, they reflect constraints in their organization that can provide clues about the underlying mechanisms^{2,3}. For example: in semi-arid vegetation patterns, power-laws are produced by facilitative interactions between plants against water scarcity^{4,5}. In inter-tidal mudflat ecosystems the loss of power-law patterns is indicative of a degradation of the system⁶. Changes in power-law vegetation patterns can signal the transition from a facilitation-dominated regime to a competition-dominated one⁷. The previous examples deal with patch-size distributions but in several cases the definition of patches is not simple. Such cases are handled naturally by multifractals because they use densities or biomass directly.

Multifractals require that the object under study should be statistically self-similar, which means that a power-law could be fitted to data in a range of scales. But that does not mean that the power-law must be the best possible model. We can analyze the data without claiming that it is an exact multifractal⁸. One of the advantages of multifractals is that they require fewer conditions on data than more classical statistics such as autocorrelation and variograms. These usually require isotropy and stationarity⁹ but multifractals can be used with anisotropic data¹⁰ and are inherently non-stationary^{11,12}. Anisotropy and non-stationarity are often seen in spatial ecological distributions¹³.

Multifractals are associated with systems governed by random multiplicative processes¹⁴. In ecological systems, these processes can be given as the interaction of survival probabilities and

compound growth¹⁵. Moreover, the presence of multiplicative process is argued to produce the log-normal-like shape of species-abundance distributions¹⁶. Also, random processes with spatial correlations can generate multifractals¹⁴; these kind of processes are part of neutral community models^{17,18} and are observed in natural communities¹⁹. Thus there are *a priori* reasons to think that multifractals can be applied to spatial ecological data. Indeed, they have been applied to vegetal communities²⁰, tropical forest²¹, microphytobentos and periphyton biomass patterns^{1,22}, and to the characterization of species-area relationships²³⁻²⁵.

Rank-abundance distributions are a representation of species-abundance distributions (SAD) that are a classical description of communities²⁶. These have been used to compare different communities and to compare models and data, but different mechanisms can produce nearly identical SADs²⁷. SADs are often presented using rank-abundance diagrams (RAD) where the log-abundance is plotted on the y-axis vs. rank on the x-axis²⁶. RADs are equivalent to cumulative distributions²⁸ and thus are a robust way to visualize the SAD without losing information²⁹. If the rank of each species is incorporated in its spatial distribution, it forms a surface: the species-rank surface (SRS). This SRS can be analyzed and compared using multifractals.

Here I present an open source software package that can be used for quantitative multifractal analysis (MFA) of densities, biomass or other continuous variables distributed in space. In addition the software can analyze SRS using MFA, though this kind of analysis is completely novel. A detailed description of the advantages of using SRS and MFA is outside the scope of this work and will be presented elsewhere. This software represents a step to make easier the use of multifractals for spatial pattern analysis in ecology. One of its advantages is that it can be integrated with another widely used software for quantitative analysis: the R statistical software³⁰.

Multifractal analysis

Several good introductions to multifractal methods applied to ecology are available^{20,31}; thus I will only give a brief overview. Multifractals analyze the scaling properties of quantities distributed in a space that we assume to be two dimensional (a plane), but MFA can be used with one dimensional (time series) or three dimensional data³². A classical way to characterize multifractals is using the generalized dimensions D_q ³³, also called Renyi dimensions³⁴. D_q has been used to portray the probabilistic structure of attractors derived from dynamical systems¹⁰.

Another way to characterize multifractals is using the so called spectrum of singularities. This spectrum describes multifractals as interwoven sets each one with a singularity exponent α and a fractal dimension $f(\alpha)$ ³⁵. The two multifractal representations are equivalent, they display the same information in a different format. But with the spectrum of singularities, two quantities are estimated (α & $f(\alpha)$) from data and are obtained with error. Instead, with generalized dimension only one quantity is estimated D_q , thus this method is preferred for statistical comparisons.

The multifractal spectra D_q is related to the Hill's generalized diversity index N_q , where q is an arbitrary real number³⁶. There are

special cases of q that are of common use in ecology: for $q = 0$ N_q corresponds to the number of species, $q = 1$ to the Shannon index, and $q = 2$ to the reciprocal of the Simpson's diversity index. Thus N_q defines different diversity measures as q changes, which vary in their sensitivity to rare species. In a similar way D_q focuses on regions of the plane with higher densities if q is greater and in regions with lower densities if q is lower or negative.

There are two important differences, the first is that N_q is calculated at one predetermined scale of measurement, and D_q is related to how N_q changes with scale³⁷. The second difference is that D_q can be calculated on any quantity distributed in space, not only the number of species. The multifractal formalism was originally developed for additive quantities³⁸, and was later extended to non-additive quantities³⁹. This means that to apply MFA, the quantity distributed over space must increase in a mathematical sense. For example, if you have ten species in an area A1 and ten species in an area A2 the number of species in the sum A1+A2 will be greater than or equal to ten. If the number of species were additive the sum A1+A2 has to be 20, but that is not generally true. A limitation of mfsBA in its present version is that it only estimates multifractal spectra for additive quantities.

Estimation

To estimate multifractal spectra I used the method of moments based on box-counting³⁸. I estimate generalized dimensions and the spectrum of singularities at the same time using the canonical method³⁵. Here I describe only the D_q estimation; the steps for α and $f(\alpha)$ estimation are identical (only the formulae to calculate the quantities are different and can be found in the appendix of Saravia *et al.* (2012)¹).

The spatial distribution that we are analyzing is covered with a grid, which is divided into $N(\epsilon)$ squares of side ϵ . The contents of each square is called $\mu_i(\epsilon)$. Then the so called partition function is computed as:

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

Where q is called moment order. The operation is performed for different values of ϵ and q , within a predetermined range. The generalized dimension is calculated as:

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

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

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

In practical cases, as the limit can not be assessed, the dimensions are estimated as the slope of $\log(Z_q)$ versus $\log(\epsilon)$ (Figure 1) divided by $q - 1$ as shown in equation (1). This is done for different q , provided that it is a real number which yields a graph of D_q in terms of q , called the spectrum of generalized dimensions (Figure 2).

To be an approximate multifractal, the relationship $\log(Z_q)$ versus $\log(\epsilon)$ should be well described by a linear relationship, although a linear relationship with superimposed oscillations is also acceptable³¹. A range of q and ϵ is fixed and then D_q is estimated using linear regressions. The coefficient of determination (R^2) can be used as a descriptive measure of goodness of fit²³.

Use of mfsBA software

The software was built and tested under Ubuntu 12.04 LTS Linux environment, using the GNU C++ compiler (v4.6.3). It requires the libtiff library for reading tiff images (<http://www.libtiff.org/>). It can be compiled under Windows environments using the GNU compiler and utilities for that operative system (<http://www.mingw.org/>).

You can download or clone mfsBA from <https://github.com/laravia/mfsba> (using `git clone` or as a zip file) and build it using the make utility.

```
make -f mfsBA.mak
```

You can run it from the command line using the following command structure:

```
mfsBA inputFile qFile minBox maxBox numBox-
Sizes option
```

the parameters are:

- **inputFile**: this file can have only two formats: 1) one-layer tiff 2) "sed" file format. Sed is an ascii format I invented to use with my own stochastic cellular automata models to represent a square grid of values. It has a header of two lines: the first line describes the two dimensions X Y of the data, and the second line describes the type of data. For this program, the type must be BI, which means that the values stored in the grid are real numbers with double precision. See the example file with a ".sed" extension.
- **qFile**: this is a sed file with a vector of values representing the q 's used to calculate the multifractal spectrum.
- **minBox,maxBox,numBoxSizes**: Minimum box size, maximum box size and maximum number of box sizes. The program uses box sizes in powers of two: if **maxBox** is greater than half of the image size, it is set to that value. If the number of boxes between **minBox** and **maxBox** is greater than **numBoxSizes**, the latter number of boxes will be used, discarding the biggest ones.
- **option**: is an upper case character with four possibilities: N,S,D,A.
 - N: analyze the input file as is.
 - S: normalize the input file then analyze it. Normalization is done summing all the pixels values and dividing each pixel by that total. After that the sum of all values is one.

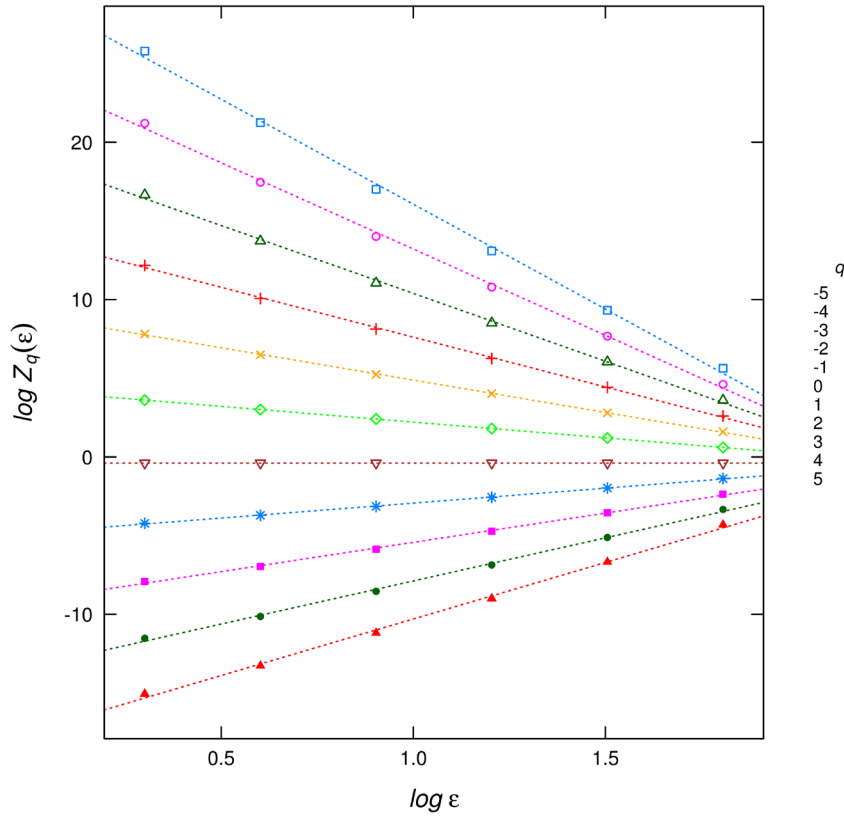


Figure 1. Plot of the linear regressions for different q used to estimate the D_q multifractal spectrum.

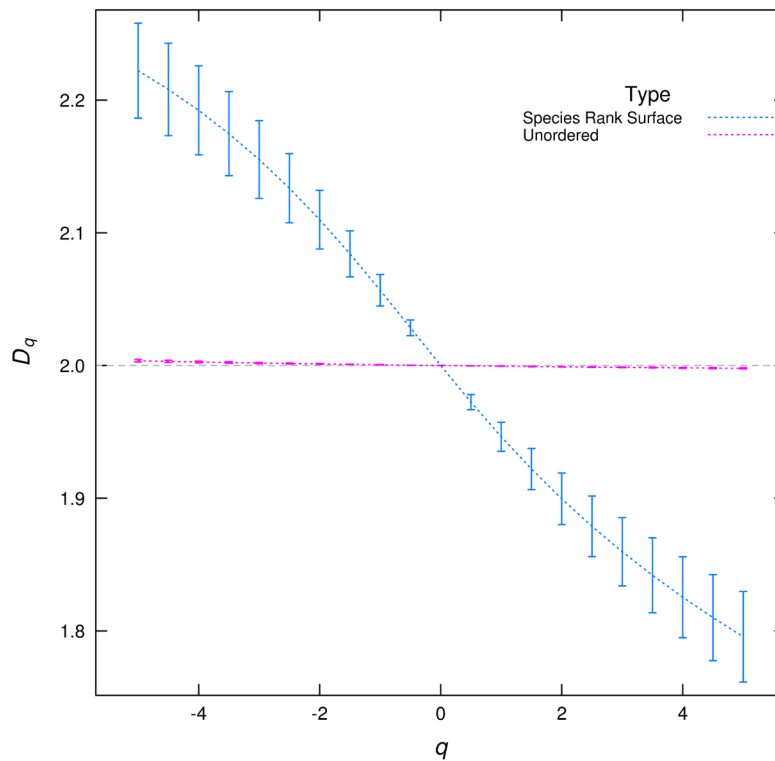


Figure 2. D_q multifractal spectrum calculated from species spatial distributions. If the multispecies distribution is analyzed unordered (with species numbers assigned by the simulation software) the D_q is almost flat corresponding to a uniform plus random noise distribution. But when the species rank surface (SRS) is used the D_q spectrum has a wide range of values, corresponding with a highly heterogeneous distribution, formed of valleys for the most abundant species and peaks for rare species. The error bars are the standard deviation obtained from the linear regressions used to estimate D_q .

- D: add 1 to all the pixels then normalize as in S.
- A: normalize as in S and save the normalized image as a sed file.

Examples of input files are included with the source code, thus after compiling you could run the following command assuming a linux system:

```
./mfSBA b4-991008bio.sed q21.sed 2 256 20 S
```

Output

The program generates four output files, attaching a prefix to the original input file name:

- *t.inputFile*: this file has a header line with field names and $q + 2$ columns. The first two columns are the box sizes and log box sizes used in the estimation. After that, each column corresponds to $\log(Z_q(\epsilon))$ of equation (2) with the q 's specified in the **qFile**. This file could be used to visually check the linearity assumptions to calculate D_q as in Figure 1.
- *a.inputFile* & *f.inputFile*: are similar to the previous file, but used to calculate α and $f(\alpha)$. The formulae are described in the appendix of Saravia *et al.* (2012)¹.
- *s.inputFile*: this file has a header line with field names and 10 columns. The first column is q . The second column is called Tau and is the result of the regression to calculate the limit in equation 2. Thus to obtain D_q we have to divide it by $q - 1$, except in the case of $q = 1$ that we take the value of the next column to get D_q . The third column is the value of α and the fourth column $f(\alpha)$. After that, columns are the corresponding coefficients of determination R^2 and standard deviations.

Species rank surface

I propose to extend the analysis of SAD attaching the rank of each species to its spatial distribution. In this way, the multivariate spatial distribution of all species can be summarized into a univariate distribution. I called this spatial distribution the species-rank surface (SRS), and it can be analyzed and compared using MFA. To construct the SRS, I first calculate the rank-ordering of the species by their abundance from biggest to smallest, starting from one. Then the rank is assigned to the spatial position of the individuals of each species, forming a surface. This landscape has valleys formed by the most abundant species and peaks determined by the rarest species, and the standard MFA can be applied. SRS is additive, in the sense explained earlier, because ranks are not recalculated when the scale changes. The program used to calculate this is called multiSpeciesSBA, and is included with the mfSBA source code. You can compile it using the following command:

```
make -f multiSpeciesSBA.mak
```

Then all the input files and parameters are identical to mfSBA except that the program expects an **inputFile** containing a multispecies

distribution. So the **inputFile** should be composed of integer numbers each one representing one species. An example of a sed file with a multispecies spatial distribution is given in t64-0100.sed, this file was obtained using a spatially explicit neutral model with 64 species (available at <https://github.com/lSaravia/neutral>). You can use the following command to perform the MFA:

```
./multiSpeciesSBA t64-0100.sed q21.sed 2 128 20 N
```

A confidence envelope for spatial randomness

To determine if D_q spectra are different from ones produced by a random spatial distribution of the quantity analyzed, I developed a randomization test. I shuffled all the positions of the original distribution and recalculated D_q . The procedure is repeated many times (e.g. 1000) and the highest and lowest tails determine a confidence envelope⁴⁰. If the actual values of D_q falls outside the envelope the spatial pattern is not random for that particular q . This program can be compiled using the following command:

```
make -f mfSBArnz.mak
```

The command to run the analysis has a similar structure to the previous:

```
mfSBArnz inputFile outFile qFile minBox maxBox numBoxSizes option numSimul P
```

All the parameters previously mentioned have the same meaning, I will explain only the new ones:

- **outFile**: this is the name of a file with all the D_q calculations specified with the numSimul parameter. It has the same format as the s.inputFile with an additional first column labeling the original D_q or a randomization. This file could be used to calculate the confidence envelope with a different P without doing the randomizations again. The program generates another file with an added ".rnz" extension, which contains the summarized results: the first row contains the number of randomizations and the size of the tail, the second row the p-value requested, the third row the field names, and from the fourth row there are five columns: the original values of q and D_q , two columns representing the randomization envelope at the requested p-level (DqMax & DqMin) and a column with 1/-1 indicating if the original D_q falls outside the envelope or 0 if it was inside the envelope.
- **numSimul**: the number of randomizations, a bigger number will give a more accurate confidence envelope at the requested p-level⁴¹.
- **P**: a two tailed p-value for the confidence envelope, that is the chance than D_q falls outside the envelope if the spatial pattern is random (Type I error).

An example of this analysis can be performed with the following command:

```
./mfSBArnz t64-0100.sed t64-0100 q21.sed 2
256 20 S 1000 0.05
```

If the software is used under Windows the name of the executable program should be changed, deleting “./” and adding the “.exe” extension, for this last example it should be mfSBArnz.exe.

R integration

Included with the source is a set of functions as an example to integrate the mfSBA software with the R language (<http://www.r-project.org/>). You should have compiled or downloaded the executables in the same folder as the R scripts. Then you can load the functions inside R with:

```
source('Fun_MFA.r')
```

and then run the same given examples:

```
dq1<- calcDq_mfSBA("b4-991008bio.sed", "q21.sed
2 256 20 S")
```

An interesting example is to compare the D_q from the example multispecies spatial distribution untransformed

```
dq1<- calcDq_mfSBA("t64-0100.sed", "q21.sed 2
512 20 S", T)
```

```
dq1$Site <- "Untransformed"
```

with the D_q from SRS

```
dq<- calcDq_multiSBA("t64-0100.sed", "q21.sed 2
512 20 S", T)
```

```
dq$Site <- "Species Rank Surface"
```

```
dq <- rbind(dq, dq1)
```

and plot D_q with

```
plot_DqCI(dq)
```

In this plot (Figure 2), we can see that the two D_q spectra are different. The D_q calculated from the unordered distribution is nearly flat, this corresponds to an almost constant spatial distribution with uncorrelated random noise. The structure of SRS is lost when species are assigned with a different order, valleys are formed by the most abundant species and peaks of rare species are destroyed. This is similar to comparing a RAD with a figure made with the ranks disordered on the x axis. Although the species distribution is the same, the two surfaces are different (Supplementary Figure 1), which is reflected in their D_q spectra.

The plot of the *t.inputFile* (Figure 1) gives a visual check of the regressions to obtain D_q :

```
plotDqFit("t.t64-0100.sed", "q21.sed")
```

additionally the R^2 values could be easily checked:

```
hist(dq1$R.Dq)
```

All the examples and more graphics are included in the file testMFA.r, you should change this file to reflect the folder where you downloaded the software.

Conclusion

The multifractal spectrum can be used to describe and compare spatial patterns of biomass, density, height, point patterns, or any continuous variable. The condition is that the distribution of the variable in space must be additive. The mfSBA software is especially useful for remote sensing data because it can be used with tiff images. Multifractal patterns could be produced by the existence of multiplicative interaction between species and by spatially correlated random processes such as dispersal and growth¹⁴. Plant and animal species are generally aggregated in space thus is very likely that multifractal analysis can be used in a wide range of cases.

The analysis of SRS using D_q adds a new dimension to the comparison of species spatial distributions, because it can be used to compare spatial distributions of all species at the same time and also the abundances are accounted. An exploration of the results of different spatial patterns should be needed as a continuation of the present work.

The analysis of SRS using D_q adds a new dimension to the comparison of species spatial distributions, because it can be used to compare spatial distributions of all species at the same time and as the species ranks are used, this is a spatial version of SAD. The method presented here analyzes SRS as a surface, which is different to SAR and to multifractals calculated with the number of species as in Borda-de-Agua (2002)²³. An exploration of SRS analysis resulting from different spatial patterns combined with different SADs will be needed as a continuation of the present work. Also the software should be extended to calculate MFA for non-additive distributions such as the number of species.

The software presented here is oriented to obtain multifractal spectra for comparisons, rather than to obtain the true value. While the estimation methods used in mfSBA could be improved^{31,42}, it has been used without trouble with the kind of data obtained in ecological studies^{1,43}.

Multifractals can be successfully used to analyze several aspects of community spatial structure. With the advent of the big data era in ecology⁴⁴ and the use of new technology to acquire spatial data⁴⁵, new methods to analyze complex data sets are needed and multifractals could be an interesting addition to the ecologist's toolbox.

Software availability

Zenodo: Multifractal estimation using a standard box-counting algorithm: version 2, doi: [10.5281/zenodo.848146](https://doi.org/10.5281/zenodo.848146)

GitHub: Multifractal estimation using a standard box counting algorithm, <https://github.com/saravia/mfsba>

Competing interests

No competing interests were disclosed.

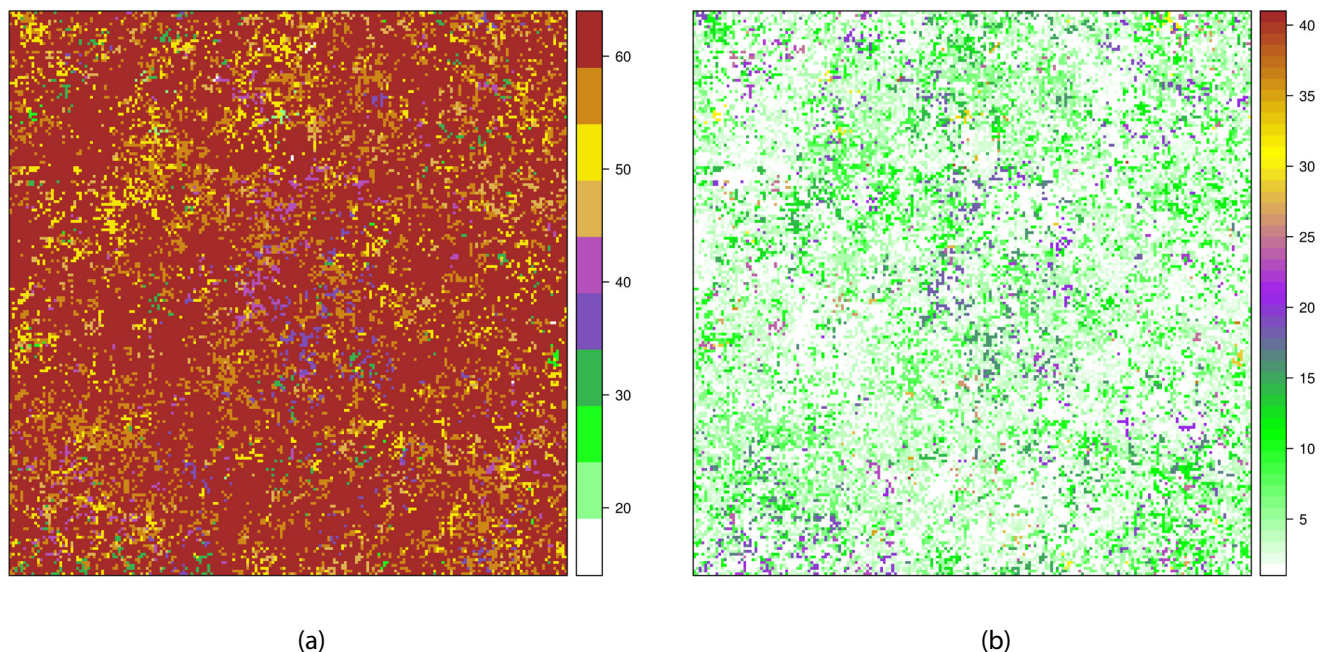
Grant information

The author(s) declared that no grants were involved in supporting this work.

Acknowledgments

I am grateful to Fernando R. Momo for our great conversations about ecological theory and David J. Harris for his language revision.

Supplementary material



Supplementary Figure 1. Comparison of a multispecies spatial distribution generated by a neutral model with 64 possible species. Both figures represent the same species distribution but with different numeric values assigned to species. **(a)** The values are assigned by the simulation program. The most abundant species have higher values and the range 1–64 represents all possible species. **(b)** Spatial Rank Surface: each species is assigned to its rank as explained in the main text: the most abundant species has the lowest value and represent valleys, rare species represent peaks. The range is lower because some species did not appear.

References

- Saravia LA, Giorgi A, Momo F: **Multifractal growth in periphyton communities.** *Oikos*. 2012; **121**(11): 1810–1820.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Brown JH, Gupta VK, Li BL, *et al.*: **The fractal nature of nature: power laws, ecological complexity and biodiversity.** *Philos Trans R Soc Lond B Biol Sci*. **357**(1421): 2002; 619–626.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Solé RV, Bascompte J: **Self-organization in complex ecosystems.** Princeton University Press, 2006.
[Reference Source](#)
- Scanlon TM, Caylor KK, Levin SA, *et al.*: **Positive feedbacks promote power-law clustering of Kalahari vegetation.** *Nature*. 2007; **449**(7159): 209–212.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kéfi S, Rietkerk M, Alados CL, *et al.*: **Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems.** *Nature*. 2007; **449**(7159): 213–217.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Weerman EJ, Van Belzen J, Rietkerk M, *et al.*: **Changes in diatom patch-size distribution and degradation in a spatially self-organized intertidal mudflat ecosystem.** *Ecology*. 2012; **93**(3): 608–618.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Manor A, Shnerb NM: **Facilitation, competition, and vegetation patchiness:**

- From scale free distribution to patterns.** *J Theor Biol.* 2008; **253**(4): 838–842.
[PubMed Abstract](#) | [Publisher Full Text](#)
8. Lévy Véhel J: **Introduction to the multifractal analysis of images.** In: Fisher Y editor. *Fractal Image Encoding and Analysis.* Springer Verlag. 1998; 299–341.
 9. Fortin MJ, Dale MRT, ver Hoef J: **Spatial Analysis in Ecology.** *Encyclopedia of Environmetrics.* John Wiley & Sons, Ltd. 2006.
[Publisher Full Text](#)
 10. Harte D: **Multifractals: Theory and Applications.** Chapman & Hall/CRC, 2001.
[Reference Source](#)
 11. Laurie H, Perrier E: **A multifractal model for spatial variation in species richness.** *Ecological Complexity.* 2010; **7**(1): 32–35.
[Publisher Full Text](#)
 12. Bez N, Bertrand S: **The duality of fractals: roughness and self-similarity.** *Theoretical Ecology.* 2011; **4**(3): 371–383.
[Publisher Full Text](#)
 13. Plotkin JB, Chave J, Ashton PS: **Cluster analysis of spatial patterns in Malaysian tree species.** *Am Nat.* 2002; **160**(5): 629–644.
[PubMed Abstract](#) | [Publisher Full Text](#)
 14. Stanley HE, Meakin P: **Multifractal phenomena in physics and chemistry.** *Nature.* 1988; **335**: 405–409.
[Publisher Full Text](#)
 15. McGill BJ: **Strong and weak tests of macroecological theory.** *Oikos.* 2003; **102**(3): 679–685.
[Publisher Full Text](#)
 16. May RM, Crawley MJ, Sugihara G: **Communities: patterns.** In: May RM, McLean AR, editors. *Theoretical ecology: principles and applications.* New York: Oxford University Press. 2007; 111–131.
[Reference Source](#)
 17. Houchmandzadeh B, Vallade M: **Clustering in neutral ecology.** *Phys Rev E Stat Nonlin Soft Matter Phys.* 2003; **68**(6 Pt 1): 1–7.
[PubMed Abstract](#) | [Publisher Full Text](#)
 18. Yakimov BN, Iudin DI, Soltsev LA, *et al.*: **Multifractal analysis of neutral community spatial structure.** *J Theor Biol.* 2014; **343**: 44–53.
[PubMed Abstract](#) | [Publisher Full Text](#)
 19. Condit R, Ashton PS, Baker P, *et al.*: **Spatial patterns in the distribution of tropical tree species.** *Science.* 2000; **288**(5470): 1414–1418.
[PubMed Abstract](#) | [Publisher Full Text](#)
 20. Scheuring I, Riedi RH: **Application of multifractals to the analysis of vegetation pattern.** *Journal of Vegetation Science.* 1994; **5**(4): 489–496.
[Publisher Full Text](#)
 21. Manrubia SC, Solé RV: **Self-organized criticality in rainforest dynamics.** *Chaos, Solitons & Fractals.* 1996; **7**(4): 523–541.
[Publisher Full Text](#)
 22. Seuront L, Spilmont N: **Self-organized criticality in intertidal microphytobenthos patch patterns.** *Physica A.* 2002; **313**(3–4): 513–539.
[Publisher Full Text](#)
 23. Borda-de-Água L, Hubbell SP, McAllister M: **Species-Area Curves, Diversity Indices, and Species Abundance Distributions: A Multifractal Analysis.** *Am Nat.* 2002; **159**(2): 138–155.
[PubMed Abstract](#) | [Publisher Full Text](#)
 24. Yakimov BN, Bossuyt B, Iudin DI, *et al.*: **Multifractal diversity-area relationship at small scales in dune slack plant communities.** *Oikos.* 2008; **117**(1): 33–39.
[Publisher Full Text](#)
 25. Laurie H, Perrier E: **Beyond species area curves: application of a scale-free measure for spatial variability of species richness.** *Oikos.* 2011; **120**(7): 966–978.
[Publisher Full Text](#)
 26. McGill BJ, Etienne RS, Gray JS, *et al.*: **Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework.** *Ecol Lett.* 2007; **10**(10): 995–1015.
[PubMed Abstract](#) | [Publisher Full Text](#)
 27. Chave J, Muller-Landau HC, Levin SA: **Comparing Classical Community Models: Theoretical Consequences for Patterns of Diversity.** *Am Nat.* 2002; **159**(1): 1–23.
[PubMed Abstract](#) | [Publisher Full Text](#)
 28. Newman MEJ: **Power laws, Pareto distributions and Zipf's law.** *Contemporary Physics.* 2005; **46**(5): 323–351.
[Publisher Full Text](#)
 29. Li L, Alderson D, Doyle JC, *et al.*: **Towards a Theory of Scale-Free Graphs: Definition, Properties, and Implications.** *Internet Mathematics.* 2005; **2**(4): 431–523.
[Publisher Full Text](#)
 30. R Core Team: **R: A Language and Environment for Statistical Computing.** Vienna, Austria: R Foundation for Statistical Computing. 2013.
[Reference Source](#)
 31. Borda-de-Água L, Hubbell SP, He F: **Scaling biodiversity under neutrality.** *Scaling Biodiversity.* 2007; **347**–375.
[Reference Source](#)
 32. Lopes R, Betrouni N: **Fractal and multifractal analysis: A review.** *Med Image Anal.* 2009; **13**(4): 634–649.
[PubMed Abstract](#) | [Publisher Full Text](#)
 33. Grassberger P: **Generalized dimensions of strange attractors.** *Phys Lett A.* 1983; **97**(6): 227–230.
[Publisher Full Text](#)
 34. Renyi A: **Probability Theory.** Amsterdam: North-Holland. 1970.
[Reference Source](#)
 35. Chhabra AB, Jensen RV: **Direct determination of the f(alfa) singularity spectrum.** *Phys Rev Lett.* 1989; **62**(12): 1327–1330.
[Publisher Full Text](#)
 36. Hill MO: **Diversity and Evenness: A Unifying Notation and Its Consequences.** *Ecology.* 1973; **54**(2): 427–432.
[Publisher Full Text](#)
 37. Ricotta C: **From theoretical ecology to statistical physics and back: self-similar landscape metrics as a synthesis of ecological diversity and geometrical complexity.** *Ecol Model.* 2000; **125**(2–3): 245–253.
[Publisher Full Text](#)
 38. Halsey TC, Jensen MH, Kadanoff LP, *et al.*: **Fractal measures and their singularities: The characterization of strange sets.** *Phys Rev A.* 1986; **33**(2): 1141–1151.
[PubMed Abstract](#) | [Publisher Full Text](#)
 39. Lévy Véhel J, Vojak R: **Multifractal Analysis of Choquet Capacities.** *Adv Appl Math.* 1998; **20**(1): 1–43.
[Publisher Full Text](#)
 40. Crowley PH: **Resampling methods for computation-intensive data analysis in ecology and evolution.** *Ann Rev Ecol Syst.* 1992; **23**: 405–447.
[Publisher Full Text](#)
 41. Wiegand T, Moloney KA: **Rings, circles, and null-models for point pattern analysis in ecology.** *Oikos.* 2004; **104**(2): 209–229.
[Publisher Full Text](#)
 42. De Bartolo SG, Primavera L, Gaudio R, *et al.*: **Fixed-mass multifractal analysis of river networks and braided channels.** *Phys Rev E Stat Nonlin Soft Matter Phys.* 2006; **74**(2 pt 2): 26101.
[PubMed Abstract](#) | [Publisher Full Text](#)
 43. Saravia LA, Giorgi A, Momo F: **Multifractal spatial patterns and diversity in an ecological succession.** *PLoS ONE.* 2012; **7**(3): e34096.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 44. Hampton SE, Strasser CA, Tewksbury JJ, *et al.*: **Big data and the future of ecology.** *Front Ecol Environ.* 2013; **11**(3): 156–162.
[Publisher Full Text](#)
 45. Michener WK, Jones MB: **Ecoinformatics: supporting ecology as a data-intensive science.** *Trends Ecol Evol.* 2012; **27**(2): 85–93.
[PubMed Abstract](#) | [Publisher Full Text](#)
 46. Saravia LA: **Multifractal estimation using a standard box-counting algorithm.** *ZENODO.* 2014.
[Data Source](#)

Open Peer Review

Current Referee Status:



Version 2

Referee Report 26 June 2014

doi:10.5256/f1000research.4039.r4385



Edmund Hart

National Ecological Observatory Network, Boulder, CO, USA

The author has addressed a majority of my concerns. I felt the explanations and cases used were much clearer with this set of revisions than in the past, and I can foresee how I could use the software. I also believe that providing windows binaries will also significantly help users. My only trouble spot was building the package from source and it's dependency on libtiff. However I think that is an issue that can be resolved on the issues page of the repo. My one thought for what might improve the paper for the more novice users would be to include that issues could be posted on the github repo, not just cloning the source. Other than that I think the paper is substantially improved and is strong work.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Referee Report 22 May 2014

doi:10.5256/f1000research.4039.r4386



Yuxin Zhang

State Key Laboratory of Urban and Regional Ecology, Chinese Academy of Sciences, Beijing, China

I tested the revised R script, it works well under Window OS.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Version 1

Referee Report 14 February 2014

doi:10.5256/f1000research.3496.r3330

**Edmund Hart**

National Ecological Observatory Network, Boulder, CO, USA

I think this software package will make a needed contribution to ecology. It meets are criterion for technical merit. However I believe the paper can still be substantially improved. There are two aspects which will benefit from revision.

Main text*Introduction:*

I think the introduction would benefit greatly from clear examples of the benefits and uses of mfSBA. It seems that the author should address both the benefits of multifractal analysis as well as where the SRS improves upon comparisons made by SAD's. It seems that the author has already made some publications that could address this, and it would be useful to have a brief explanation.

In general I found the Introduction lacking specificity. For instance the author states in the first paragraph: "If they do apply, they reflect constraints in their organization that can provide clues about the underlying mechanism" What constraints? What underlying mechanism is the author referring to? How does it provide clues about this these mechanisms?

I finished reading the introduction and was unclear both about what multifractal were, and why I would want to use them.

Methods and conclusion

I think the most important part here is a clear explanation of the quantities of interest that I can get from mfSBA, and how to interpret them. While parameters like slope and intercept seem obvious to most people now, I had no idea how to interpret q or D_q , or how I would use them.

Furthermore the author focuses on the ability of this new method to make comparisons, stating: "The analysis of SRS using Dq adds a new dimension to the comparison of species spatial distributions, because it can be used to compare spatial distributions of all species at the same time and also the abundances are accounted."

However, the comparison the author makes is very unclear. It would be nice to see a comparison where simulation conditions were clearly made, and perhaps even performed in R so users could tweak them.

Figure 2 provides output for a comparison, but it's unclear how to interpret that comparison. In summary the main text would benefit from a clearer presentation of use cases, what parameters are being estimated, and how to understand those parameters in different scenarios, both simulated and real data.

Software packaging.

I think the average ecologist will find this package still rather intimidating. It requires the installation of secondary libraries as well as a C compiler. I wouldn't even know how to begin to install this on a windows machine. I would suggest the author create an R package with precompiled binary libraries for windows and mac. This combined with an improved manuscript will provide ecologists with a powerful tool.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Referee Report 24 January 2014

doi:10.5256/f1000research.3496.r3170



Yuxin Zhang

State Key Laboratory of Urban and Regional Ecology, Chinese Academy of Sciences, Beijing, China

This is a useful Web Tool for the researcher, particular the ecologist in the community ecology, who wants to use the multifractal method describing the SAD and point pattern of species distribution. The paper is well written. However, the script (followed with the instruction in the “R integration” section) don't work well in Windows with R 3.0.2 (64bit). The author needs to revise the R scripts or provide detailed instruction on it, before the work is acceptable. Reports on the performance of the software in GNU for the Windows environment are also needed.

There are two minor comments:

1. The q is usually an integer in community ecology, because this will directly correspond to the Hill numbers.
2. The range of q in this paper is from -5 to 5. Can the users of this tool set the range of q for their own preference?

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Reader Comment 15 Feb 2014

Leonardo Saravia, Universidad Nacional de General Sarmiento, Argentina

I appreciate Dr. Zhang's comments on the manuscript. As the software was developed and tested under Linux it should not be expected to work flawlessly under windows. I agree that it would be a great benefit to release at least a windows working version and I will try to update the software as time and resources are available. But I think that the software should be evaluated as is, and not judged because it's available or not under different operative systems.

It will be useful to open issues on github with the required enhancements to facilitate the development of the software: <https://github.com/lSaravia/mfsba/issues>.

The q 's can be real numbers or integers in any range and are specified in the **qFile** as explained on the main text.

Competing Interests: No competing interests were disclosed.

Discuss this Article

Version 2

Author Response 28 Apr 2014

Leonardo Saravia, Universidad Nacional de General Sarmiento, Argentina

I appreciate Dr. Yakimov's comments on the manuscript. As a response to them and to the reviewer's comments, I have made several changes in the version 2 of the article:

1. I have added windows executables available at <https://github.com/lsaravia/mfsba/releases> and Zenodo. R for windows can be downloaded from <http://cran.r-project.org/bin/windows/base/>.
2. I have added this clarification to the version 2 of the paper.
3. I have stated that the coefficient of determination (R^2) can be used as a descriptive measure of goodness of fit, not as a criterion for multifractality. Besides that I am not assuming that the objects analyzed are exact multifractals. I think that the Akaike criterion could be a good addition and anyone could apply it because the data used for regressions are in the file called *.inputFile*. So you can take the data on this file and calculate AIC for linear and quadratic regressions and compare them. I should add to the model set to compare by AIC, the functional form of power law with oscillations specified by Borda-de-Agua¹.
4. As SRS is a surface, this example shows the difference between two surfaces. Here I intended to describe the software, and the SRS will be described in another article. Anyway I can answer that if you have co-occurring species you have to sum all ranks because that is what the algorithm does in each box. Other options could be investigated like taking the maximum or minimum or other function but that is not implemented.

About the neutral model, I used it as an example of a not random distribution of species, I added a more complete description in its github repository <https://github.com/lsaravia/Neutral>

References

1. Borda-de-Água L, Hubbell SP, He F, (2007). Scaling biodiversity under neutrality. Scaling biodiversity. Cambridge: Cambridge University Press. pp. 347–375.

Competing Interests: I am the author of the article.

Version 1

Reader Comment 04 Feb 2014

Basil Yakimov, Nizhny Novgorod State University, Russian Federation

Several comments on the paper and the software:

1) I highly appreciate the Author's intention to provide ecologists with a tool for empirical multifractal analysis. However I think that the form of C++ program created for the use in Linux environment is unsuitable. Numerous Windows-users cannot apply it or even to test. The same is true for R-interface. It works only in Linux environment.

2) After formula (3) Author states that «dimensions are estimated as the slope of $\log(Z_q)$ versus $\log(e)$ », whereas it applies to mass exponent $\tau(q)$ estimation. It should be stated that mass exponents are divided by $(q-1)$ to get D_q .

3) Author proposes application of R^2 as a criterion of multifractality. That is a very bad choice. In my experience R^2 depends strongly on q . It is in no way a measure of linearity. Instead I recommend application of Akaike information criterion for linear and quadratic fits in double log space.

4) I think illustration of program performance with species rank surface is not a good idea. Species rank surface and its analysis is something unknown to ecologists. As I understood it is proposed here for the first time. Section with its description leaves too many questions. How to construct the surface for the case of co-occurring species (i.e. not when each individual has its own cell of space)? What kind of neutral model is used for illustration? What is the interpretation of the final spectra?

Competing Interests: No competing interests
