Ecology Letters, (2004) **7**: 60–68 doi: 10.1046/j.1461-0248.2003.00549.x

REPORT

Power-law species–area relationships and self-similar species distributions within finite areas

Arnošt L. Šizling[1*] and David Storch[2,3]
[1]Department of Philosophy and History of Science, Faculty of Sciences, Charles University, Viničná 7, 128 44-CZ Praha 1, Czech Republic
[2]Center for Theoretical Study, Charles University, Jilská 1, 110 00-CZ Praha 1, Czech Republic
[3]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA
**Correspondence: E-mail: arnost.l.sizling@seznam.cz*

Abstract

The species–area relationship (SAR) is often expressed as a power law, which indicates scale invariance. It has been claimed that the scale invariance – or self-similarity at the community level – is not compatible with the self-similarity at the level of spatial distribution of individual species, because the power law would only emerge if distributions for all species had identical fractal dimensions (*FD*). Here we show that even if species differ in their *FD*, the resulting SAR is approximately linear on a log–log scale because observed spatial distributions are inevitably spatially restricted – a phenomenon we term the 'finite-area effect'. Using distribution atlases, we demonstrate that the apparent power law of SARs for central European birds is attributable to this finite-area effect affecting species that indeed reveal self-similar distributions. We discuss implications of this mechanism producing the SAR.

Keywords

Biogeography, birds, fractals, macroecology, scale invariance, scaling, species-richness patterns.

Ecology Letters (2004) 7: 60–68

INTRODUCTION

One of the most general ecological patterns is the increase in the number of species with the area sampled. There is no consensus concerning the importance of individual mechanisms contributing to the pattern or the exact shape of this species–area relationship (hereafter SAR). The shape of the SAR can be approximated by many functions, including the exponential (Gleason 1922), the logistic (He & Legendre 1996, 2002) and even more complex equations (for review see Tjørve 2003). However, it seems that a power law represents considerably good approximation of the SAR (Arrhenius 1921; Rosenzweig 1995; Storch *et al.* 2003b), at least within particular spatial scales.

The power law implies self-similarity or scale-invariance (Gisiger 2001). Harte *et al.* (1999) provided a model that explicitly related the power-law SAR to the self-similarity at the community level. This model is based on the assumption that if a species is present within an area, the probability of its occurrence within a constant portion of the original area is also constant, regardless of the absolute size of that area. However, using slightly different formalism, Lennon *et al.* (2002) claimed that the self-similarity at the level of the distribution of individual species would lead to the power-law SAR only if all species had equal fractal dimensions (*FD*). As the *FD* is closely related to species occupancy, and occupancy varies extensively among species, this condition cannot be fulfilled in most real situations, and, as also noted by Harte *et al.* (2001), the self-similarity at the species level is incompatible with the community-level self-similarity. On the contrary, there is some evidence that species distributions are indeed self-similar (Kunin 1998; Witte & Torfs 2003). There is therefore a controversy between the apparent power law of the SAR and the observed structure of species spatial distribution, which should be resolved.

Here we show that within any real landscape, even self-similar spatial distributions of species which differ in their *FD* result in SARs that are close to a power law. The reason is that the location of any study plot within a finite region is constrained by the boundary of the region and sufficiently large plots therefore inevitably contain the species in focus. The relationship between the probability of species occurrence and area thus cannot increase over all scales, which changes the shape of resulting SARs. We call this the 'finite-area effect'. We assess the reliability of the finite-area effect using numerical simulations and bird distribution data in central Europe.

THEORY

The SAR derived from the species probability of occupancy

The SAR cannot be characterized by one particular curve that assigns one species richness value to each area, as different study plots of equal area differ substantially in species number (Storch *et al.* 2003b). It is therefore necessary to express the relationship between area and the most likely species number for all plots of this area. In the following text we express the SAR as the relationship between area and mean of all species numbers (hereafter $\bar{S}$). Given the probability of occurrence for each species within the study plot, mean number of species can be calculated as

$$\bar{S}_{(A)} \equiv \sum_{i=1}^{S_{\text{tot}}} p_{\text{occ } i(A)} \tag{1}$$

where S_{tot} is the total number of species considered and p_{occ} is the probability of occurrence of species i within plots whose area is A. This is in accord with most models of the SAR based on random species distributions (Coleman 1981; Williams 1995; Muriel & Mangel 1999; He & Legendre 2002).

Self-similarity of species spatial distribution

The spatial distribution of a species is self-similar if there is a structure that occurs repeatedly on different spatial scales (Hastings & Sugihara 1993). Then, if we plot a grid of quadrates whose side length (hereafter L) corresponds to the size of the repeated structure over the distribution, the relationship between number of occupied quadrates n and L is given by formula

$$n = n_0 L^{-FD} \tag{2}$$

where n_0 is constant, and FD is Hausdorff or fractal dimension (Hastings & Sugihara 1993). Generalizing this relationship for all possible L (i.e. not only those corresponding to the size of the repeated structure), and approximating the probability of occupancy by a proportion of non-overlapping occupied quadrates, i.e. $p_{\text{occ}} \cong n/(A_{\text{tot}}/L^2)$, where A_{tot} is the total area of the grid, we obtain the formula

$$p_{\text{occ}} \cong \pi L^{2-FD} \tag{3}$$

where π is the probability of occupancy for $L = 1$ ($\pi \equiv n_0/A_{\text{tot}}$) (Lennon *et al.* 2002). If we replace L^2 with area A, we get the approximate relationship between area and probability of species occurrence for that area

$$p_{\text{occ}} \cong \pi A^{1-\text{FD}/2}. \tag{4}$$

This formula can be treated as a species–area relationship for just one species, where species number is replaced by the probability of species occurrence and the term $1 - FD/2$ corresponds to z in classical Arrhenius (1921) equation, i.e.

$$z_i \equiv 1 - \frac{FD}{2}. \tag{5}$$

Note that the relationship in eqn 4 is based on two simplifying assumptions, the first one is that eqn 2 is valid for all lengths, L, the second one comprises the estimation of probability of species occurrence by the division of the number of occupied quadrates by the total number of non-overlapping quadrates of that particular size.

The finite-area effect

According to eqn 4, the probability of occupancy would increase to infinity over infinite areas. However, probability is bounded within the interval [0,1], and thus eqn 4 can be valid only within a particular range of areas $\leq A_{\text{tot}}$. Within any real (i.e. finite) census area we can envisage area of saturation, A_{sat}, as a minimum area of the study plot that always contains the species, regardless of its location (see Fig. 1a), i.e. A_{sat} is bigger than any distributional 'gap' of respective shape. Beyond A_{sat}, the p_{occ} is inevitably equal to 1. We call this the finite-area effect, as the existence of A_{sat} is the direct consequence of the fact that we can study species distributions and diversity patterns only within finite census areas in which the location of sample plots is constrained by the boundary of the total study area (e.g. whole continent).

The finite-area effect has important implications for the relationship between p_{occ} and A in self-similarly distributed species, which can be expressed as a continuous function using formula

$$\begin{aligned} &p_{\text{occ } i} = \pi_i A^{z_i} \quad \text{for all } 0 \leq A \leq A_{\text{sat } i}, \quad \text{and} \\ &p_{\text{occ } i} = 1 \quad \text{for all } A_{\text{sat } i} \leq A \leq A_{\text{tot}}. \end{aligned} \tag{6}$$

If we order species according to their value of A_{sat}, we can combine eqn 1 with eqn 6 to express the SAR as

$$\bar{S}_{(A)} = \sum_{i=S_{\text{sat}(A)}+1}^{S_{\text{tot}}} \pi_i A^{z_i} + S_{\text{sat}(A)} \tag{7}$$

where $S_{\text{sat}(A)}$ is the number of species whose relationship between p_{occ} and A has reached saturation (i.e. species with A_{sat} lower than A). The exact pattern of the increase of species number with area will therefore be dependent on the values of z_i and $A_{\text{sat } i}$ for individual species (see Fig. 1b). Let us call the proposition that the SAR can be derived from these properties of species distributions using eqn 7 the *finite-area model*.

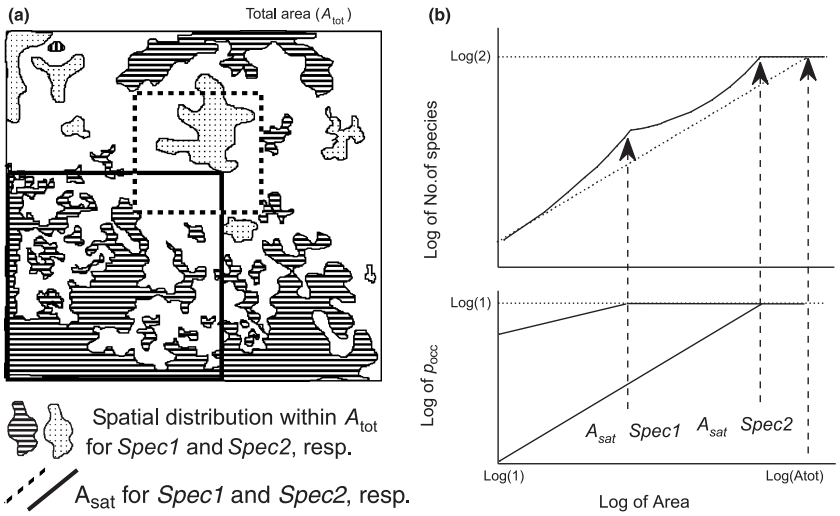

Figure 1 (a) The geometric representation of A_{sat} within a rectangular area A_{tot}. A_{sat} represents the minimum area of a plot where the respective species is necessarily present regardless of the location of the plot ($p_{occ} = 1$ for all areas $\geq A_{sat}$); different species (here *Spec*1 and *Spec*2) differ in their A_{sat}. (b) The effect of A_{sat} for the emergence of SAR. The linear relationship between A and p_{occ} in a log–log space, which characterizes the self-similarly distributed species, is valid only up to A_{sat}, because then $p_{occ} = 1$ (below). This affects the shape of the SAR (above) resulting from summing p_{occ} for both species (solid line). The slope of the dotted line is the apparent slope of the SAR in a log–log space calculated using eqn 8. The A_{sat} pushes the slope of the SAR in the log–log scale down.

Shape of the SAR

According to the finite-area model, the SAR is affected both by the self-similarity expressed as the linear increase of probability of species occurrence with area in a log–log space, and by the finite-area effect. It is therefore necessary to explore how these effects combine to produce the resulting SAR.

There are two components of the shape of the SAR. The first component comprises the intervals between subsequent A_{sat} for individual species. Here $S_{sat(A)}$ is constant and only z_i of individual species with A_{sat} higher than the values within that interval contribute to the increase of species number within that interval, because all species with lower A_{sat} have $p_{occ} = 1$ (see Fig. 1b) within the entire interval. For any interval between $A_{sat\,j}$ and $A_{sat\,j+1}$, the SAR can therefore be expressed using eqn 7, where $S_{sat(A)}$ remains constant (i.e. $S_{sat(A)} = j$). The shape of the respective curve can be studied analytically (Appendix 1), and these analyses show, in accord with Lennon *et al.* (2002), that the SAR within these intervals is always convex (upward accelerating) in a log–log space. The second component represents the effect of A_{sat}, which opposes this tendency and leads to the decrease of the resulting slope of the whole SAR (see Fig. 1b and Appendix 1).

There are, therefore, two opposite tendencies affecting the SAR: the tendency of upward accelerating increase of the species number caused by interspecific differences in *FD* and of downward decrease because of the finite-area effect. The strength of these opposing forces depends on the frequency distribution of A_{sat} and z_i, and on the relationship between these variables. It would be very complicated to prove analytically as to which distributions would lead to particular SARs, so we performed numerical simulations assuming various frequency distributions for both variables. We investigated three types of distribution for A_{sat} (symmetric distribution, and distributions with negative and positive skewness, Fig. 2a), and two types of frequency distribution for z_i (uniform and independent of A_{sat}, and positively dependent on A_{sat} such that $z_i = A_{sat\,i}/A_{tot}$; Fig. 2b). All six combinations of distributions of A_{sat} and z_i were explored. We did not consider the situation of the negative dependency between the variables, because this is unrealistic – note that both variables depend on occupancy (number of occupied grid cells) such that species that occupy many cells necessarily have low A_{sat}, high *FD* and consequently low z. There is actually clear statistical (although not straightforward) relationship between these variables, but its exploration is beyond the scope of this paper.

We performed 1000 simulations for 500 species. The resulting SARs (Fig. 2c) are not exactly linear in a log–log space, but in most cases there is no apparent curvilinearity. Considering the high variance of observed species numbers within plots of equal area (Storch *et al.* 2003b), the deviations of SARs from the power law would be

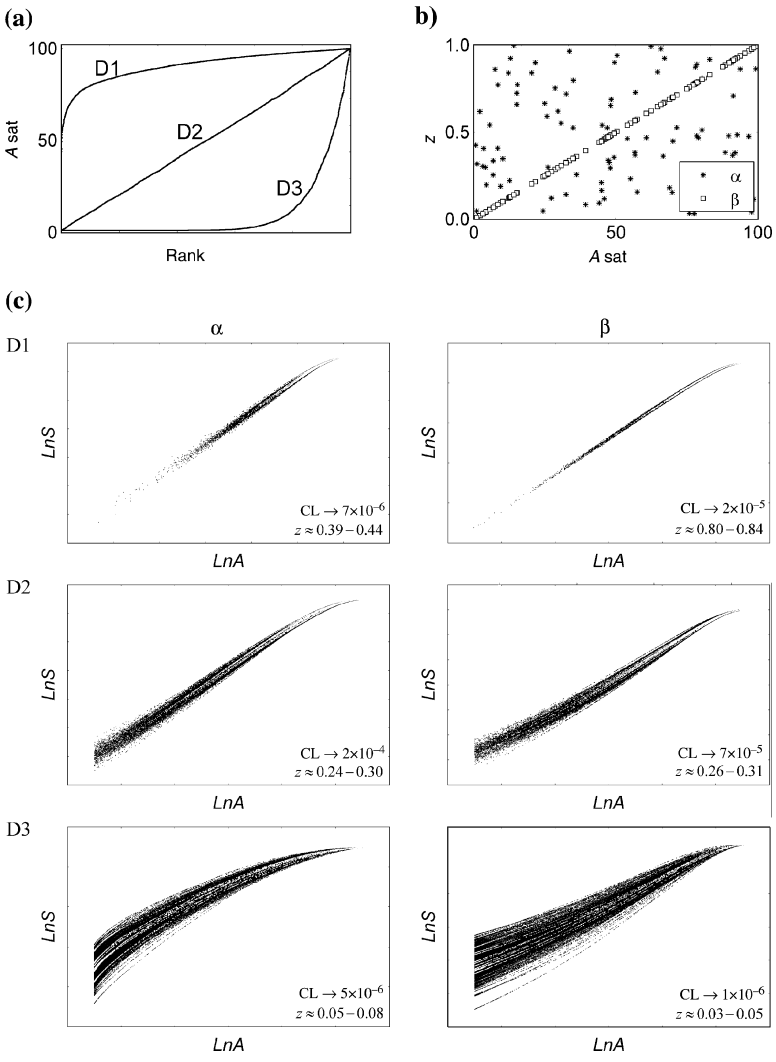

Figure 2 The settings and results of the simulations of the finite-area model. (a) The three types of the distribution of A_{sat} used, expressed using the relationship between species rank and A_{sat}, (b) the two types of the relationship between A_{sat} and z (α – not related, β – linearly dependent), and (c) the results of the simulations for 500 species, for all combinations of these settings. Each dotted line represents one of 1000 simulation runs, except the first 10 simulations that are represented by the white lines to visualize the shape of the SAR for one simulation run. The outlines, given by the most upper and the lowest simulation, represent a 99% confidence interval with 95% likelihood (Jílek 1988). The shapes of simulated SARs are apparently not linear in a log–log scale, but are quite close to the linearity. The numbers in the right lower corners refer to the value of the curvilinearity (CL), (see text) and to the range of the slopes of simulated SARs, z.

undetectable. The slopes of simulated SARs vary mostly accordingly to the changes of distribution of A_{sat} (Fig. 2a), such that the documented slopes of mainland SARs ($z \approx 0.08 - 0.25$; see Connor & McCoy 1979) correspond well to the distribution which is somewhere between distribution D2 (uniform) and D3 (characterized by prevalence of lower values of A_{sat}).

We have also explored the relationship between the total species number and the curvilinearity of emerging SARs, by performing 100 simulations of different species numbers (50, 100, 150, 200,…, 500) and measuring curvilinearity for each case. The curvilinearity has been calculated as a mean of sum of squares of distances from the regression line for all points of A_{sat} in a log–log space. It did not change with number of species, but its variance did, indicating that the SARs converge to some particular shape with increasing species number.

Implications for the slope of the SAR

If the shape of the SAR can be expressed approximately as a power law, then its approximate slope in a log–log space can be calculated from the extreme points of the relationship, i.e. the maximum ($A = A_{tot}; \bar{S} = S_{tot}$) and minimum ($A = 1; \bar{S} = \sum \pi_i$). The approximate slope is then

$$Z = \frac{\ln\left(\frac{S_{tot}}{\sum \pi_i}\right)}{\ln(A_{tot})}. \quad (8)$$

This equation has two implications. First, the slope depends on π_i, i.e. on the intercept of the regression line in a log–log space, which can be estimated by the relative species occupancy in the smallest area sampled (grid cells). Obviously, if all species occurred everywhere, the slope of the SAR would be zero, whereas if every species occupied just

one grid cell, the slope would converge to 1. Consequently, assuming the spatial self-similarity, and knowing the total number of species S_{tot} within A_{tot}, it could be possible to predict the approximate slope of the SAR using only the data of the number of occupied grid cells for each species. And as relative species occupancies are related to the species-abundance distribution (Nee *et al.* 1991; Storch & Šizling 2002), it might be ultimately possible to derive the shape and slope of the SAR just from the knowledge of species total abundances (cf. Preston 1960; Sugihara 1980; Harte *et al.* 2001; He & Legendre 2002).

Second, although the mainland SAR has often been attributed to an increase in the number of habitats with area (Rosenzweig 1995), the slope of the SAR cannot be attributed only to the habitat effect, because the amount of suitable habitats represents the upper limit for species occupancies, and therefore the lower limit for the slope of the SAR. The SAR for species will always be steeper than the SAR for habitats, because species occupy only a portion of suitable habitat (Storch *et al.* 2003b).

Testing the model using empirical data

We empirically tested our theory by determining whether the relationship between area and species probability of occupancy follows eqn 6, whether species reveal self-similar distributions, and whether observed SARs can be predicted by the finite-area model. We have used the data on the distribution of birds in central Europe on two scales of resolution (Fig. 3), that of basic grid cell size of 10′ in longitude and 6′ in latitude, that is, *c.* 11.1 × 12 km (the Czech Republic, hereafter CR; Šťastný *et al.* 1996) and that of basic grid cell size of 50 km × 50 km (central Europe, hereafter CE; Hagemeijer & Blair 1997). Both data sets consist of 16 × 16 grid cells, containing the information about probable or confirmed breeding of all bird species within each cell (see Storch & Šizling 2002).

First, we tested whether eqn 6 represents a reliable description of the observed species distributions, i.e. whether the relationship between the probability of species occupancy and area is linear in the log–log space within respective intervals, and whether the deviation from this relationship does not affect the SAR calculated using the finite-area model (eqn 7). The probabilities $p_{occ\ i}$ have been calculated as a proportion of all possible plots of given area within the grid, which contained at least one record of the respective species (Storch *et al.* 2003b). For each species we calculated z_i and $A_{sat\ i}$ by extracting the slope and the intercept of the regression line for the relationship between $\ln(A)$ and $\ln(p_{occ})$ for $A < A_{sat}$ (i.e. not considering the values of $p_{occ} = 1$). Then we defined residuals $\varepsilon_{i(A)}$ for each species and area as $\varepsilon \equiv p_{occ\ observed} - p_{occ\ predicted}$, where $p_{occ\ predicted}$ were calculated using eqn 6 and the extracted parameters. If the relationship between $\ln(p_{occ})$ and $\ln(A)$ for a species is actually linear within the interval, residuals $\varepsilon_{i(A)}$ should be close to zero for all areas.

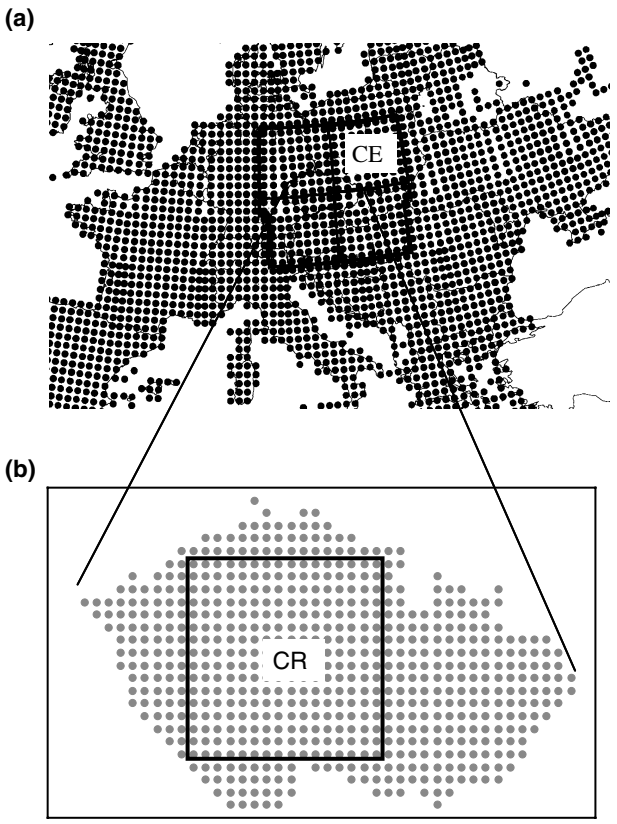

Figure 3 The studied census areas: grids of 16 × 16 cells, located within central Europe (CE) (a) and the Czech Republic (CR) (b). For details see Storch & Šizling (2002).

The analyses show that all $\varepsilon_{i(A)}$ are close to zero for areas larger than 6 × 6 grid cells (Fig. 4a), but they differ significantly from zero in smaller areas. This deviation in smaller areas is, however, quite small – note that the mean of $\varepsilon_{i(A)}$ for all species within an area is actually equal to the difference between the observed and predicted species number divided by the total species number – and thus the value of 0.04 means that for a sample of 100 species the prediction will differ from the observation by only four species. Spatial distribution of individual species is therefore reasonably well represented by eqn 6, and the resulting SAR can be well predicted by the finite-area model.

The deviation from the finite-area model in small areas could be either the result of the violation of the assumption of self-similarity for these areas, or because of the approximate nature of eqn 6, which could represent an inaccurate approximation of the relationship between A and p_{occ} for smaller areas in self-similarly distributed species. We tested this second possibility by calculating residuals $\varepsilon_{i(A)}$ for

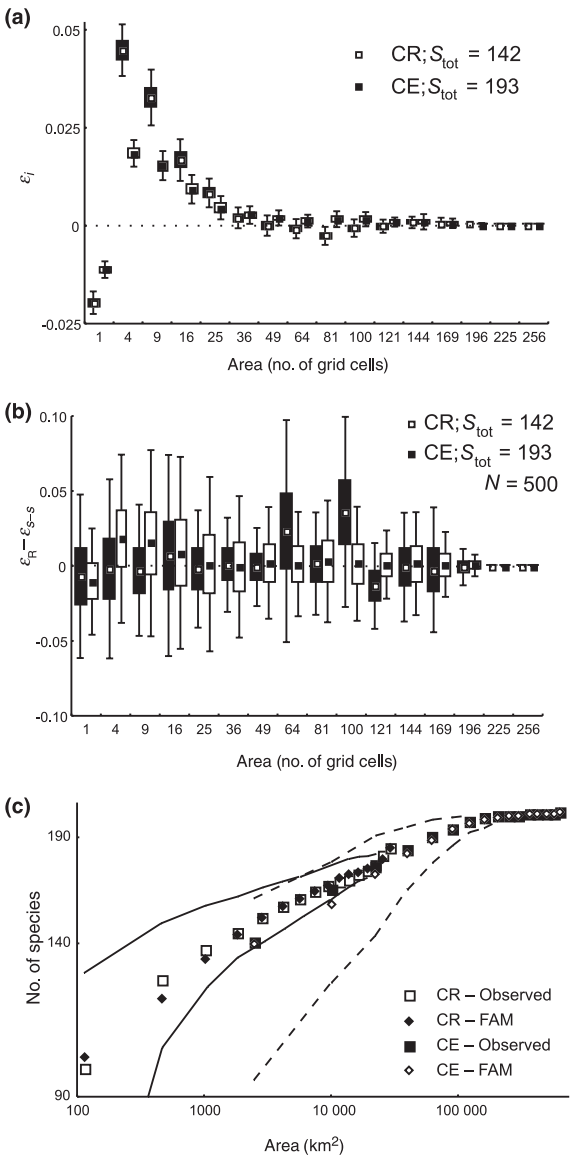

Figure 4 Results of the test of our theory using the bird distribution data. (a) The residuals of observed p_{occ} from p_{occ} predicted by the finite area model ($\varepsilon \equiv p_{occ\ observed} - p_{occ\ predicted}$) for each area wihin the Czech Republic (CR) and central Europe (CE). Boxes and whiskers represent 50 and 95% confidence intervals of means ($\pm$0.67SE and $\pm$1.96SE), respectively, which are important for the estimation of the difference between observed and predicted species numbers (see text). (b) The deviation between ε_R for the comparison between the linear model and observations, and ε_{S-S} for the comparison between the linear model and the simulated self-similar distributions (for details see text). Boxes and whiskers represent 50 and 95% confidence intervals, respectively. The deviations between ε_R and ε_{S-S} apparently do not differ systematically from zero. (c) The comparison between observed SARs and the SARs predicted using the finite-area model (FAM). Full and dashed lines represent 95% confidence intervals of the observed species numbers for CR and CE, respectively.

500 simulated assemblages of self-similarly distributed species (see Appendix 2) with *FD*s equal to the observed *FD*s, and by comparing these residuals with observed $\varepsilon_{i(A)}$. The difference between the residuals from the two data sets is close to zero (Fig. 4b). As the simulated species distributions were exactly self-similar with known *FD*, the deviation of the finite-area model is therefore not attributable to the violation of the assumption of spatial self-similarity of species distribution, but is because of the fact that eqn 6 is just an approximate expression of self-similarity. Therefore, it is not possible to reject the hypothesis of self-similar spatial distribution of species, and the SAR is in our case attributable to the collective effect of self-similarly distributed species (Fig. 4c). Moreover, the approximation of the self-similarity with the finite-area model (eqn 6) is considerably accurate and as it is possible to deal with it analytically, it represents very useful tool for studying the relationship between SARs and species distributions.

DISCUSSION

Our results show that there is a relationship between the spatial self-similarity and the SAR, postulated by Harte *et al.* (1999). However, our theory differs from the model of Harte *et al.* (1999) in several points. First, and most important, we consider finite grids composed of finite numbers of cells, some of which are occupied by particular species. There is no room for discussion on whether the self-similarity holds down the level of the distribution of individuals, and whether it implies some distribution of species abundances. As Hubbell (2001) and others have pointed out, within the smallest scales the SAR certainly does not have a form of the power law, and the self-similarity does not hold down the level of individuals. We just show that when dealing with sufficiently large sampling plots (grid cells), the assumption of self-similarity is valid.

The second important difference is that our theory does not rely on a particular way of measuring the self-similarity. The model of Harte *et al.* (1999) assumes that the area is represented by 'golden rectangle', and self-similarity is represented by the constant probability of occurrence of a species within exactly half the original area. As Maddux (2004) pointed out, this assumption leads to the violation of transition invariance: if the probability is defined for, say, right or left half of the rectangle, the half which is located in the middle of the original area has different value of the probability, and thus the self-similarity concerns only particular plots within the original areas. This is not the case for our theory, because this operates with the probability of species occurrence within all possible plots of a particular area.

Third, we assume that species have different occupancies, and thus different *FD*s. Although Harte *et al.* (1999) have not made any explicit statement concerning the self-similarity at

the level of the distribution of individual species, Lennon *et al.* (2001) realized, in accord with Harte *et al.* (2001), that the power law would emerge only if all species had equal *FD*s. Our theory, on the contrary, shows that although within the intervals between consecutive A_{sat} for individual species the SAR is upward accelerating in a log–log space if the species differ in their *FD*, contributions of the finite-area effect result in SARs which are close to the power law. Considering the variance of observed species numbers within samples of equal areas, it is not surprising that observed SARs can be often well described as power laws (Storch *et al.* 2003b) even if the species spatial distribution is actually self-similar.

In our data set, the species reveal self-similar spatial distributions, and the resulting SAR can be predicted from the properties of these distributions. But why should species distributions be self-similar? One possibility is that this feature is imposed on species by the environment, i.e. that natural landscapes have self-similar properties. There is some evidence supporting this argument. Storch *et al.* (2002) showed that the spatial variability of biologically relevant parameters reveal spectral properties indicating self-similarity (so called $1/f$ spectra; see Halley 1996). In addition, if the species distribution is strongly affected by altitude, i.e. if species are confined to only particular elevations, their distribution could be self-similar, because altitude is related to the ruggedness of the earth surface, which reveals fractal properties (Mandelbrott 1977) – after all, many altitudinal changes are related to the system of water drainage, which is naturally fractal (Peckham 1995; Veitzer & Gupta 2000). We have good evidence that altitude is indeed the most important factor affecting the distribution of birds within the Czech Republic (Storch *et al.* 2003a).

However, for the same data set used here, Storch *et al.* (2003b) showed that the shape and slope of the SAR are not attributable only to the effect of habitat, and moreover, that the SARs predicted from the distribution of habitats differ substantially from the observed SARs, which are actually much closer to the power law. The power law in this case emerges because of the spatial aggregation of species which is not fully attributable to habitats. It is thus necessary to look for spatial population processes that generate self-similarity, i.e. to consider the dynamic nature of species assemblages (Adler & Lauenroth 2003). There are plenty of spatial population models that can be parameterized to produce particular spatial distributions including the self-similar distributions (e.g. Hubbell 2001), but this generality is their weakness rather than strength. For now, we do not know any particular reason why the spatial population dynamics should preferentially lead to self-similar spatial distribution.

There is also a possibility that the apparent self-similarity of species distribution is simply the result of the fact that the distribution is affected simultaneously by many factors acting on different scales of resolution: whereas in some parts of a species range the distribution is affected mostly by the habitat availability, in other places it is driven by spatial population processes, combined with spatially restricted interspecific interactions and so on. Then the scale invariance would represent a 'neutral' distribution, in a sense more random than that provided by models of random distribution based on equal density of probability of occupancy across all places (i.e. Poisson distribution). This argument is similar to that of why there are the $1/f$ spectra of environmental and population variation in time (i.e. that it is because of independent effects of factors acting within different scales; see Halley 1996), or similar argumentation about the spectral properties of physical surfaces (Sayles & Thomas 1978). This idea, however, would deserve further theoretical consideration.

ACKNOWLEDGEMENT

We thank Andrew Allen, John Harte, Ethan White, Kevin Gaston and two anonymous referees for helpful comments. The study was supported by the Grant Agency of the Czech Republic (GACR 206/03/D124). D.S. was supported by the International Program of Santa Fe Institute.

REFERENCES

Adler, P.B. & Lauenroth, W.K. (2003). The power of time: spatiotemporal scaling of species diversity. *Ecol. Lett.*, 6, 749–756.

Arrhenius, O. (1921). Species and area. *J. Ecol.*, 9, 95–99.

Coleman, D.B. (1981). On random placement and species–area relations. *Math. Biosci.*, 54, 191–215.

Connor, E.F. & McCoy, E.D. (1979). The statistics and biology of the species–area relationship. *Am. Nat.*, 113, 791–833.

Gisiger, T. (2001). Scale invariance in biology: coincidence or footprint of a universal mechanism? *Biol. Rev.*, 76, 161–209.

Gleason, H.A. (1922). On the relation between species and area. *Ecology*, 3, 158–162.

Hagemeijer, W.J.M. & Blair, M.J. (1997). *The EBCC Atlas of European Breeding Birds*. T. & A.D. Poyser, London.

Halley, J.M. (1996). Ecology, evolution and 1/f noise. *Trends Ecol. Evol.*, 11, 33–37.

Harte, J., Kinzig, A. & Green, J. (1999). Self-similarity in the distribution and abundance of species. *Science*, 284, 334–336.

Harte, J., Blackburn, T. & Ostling, A. (2001). Self-similarity and the relationship between abundance and range size. *Am. Nat.*, 157, 374–386.

Hastings, H.M. & Sugihara, G. (1993). *Fractals, a User's Guide for the Natural Sciences*. Oxford University Press, Oxford.

He, F.L. & Legendre, P. (1996). On species–area relations. *Am. Nat.*, 148, 719–737.

He, F.L. & Legendre, P. (2002). Species diversity patterns derived from species–area models. *Ecology*, 85, 1185–1198.

Hubbell, S.P. (2001). *A Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, NJ.

Jílek, M. (1988). *Statistical and Tolerance Limits*. SNTL, Praha, pp. 52–67 (in Czech).

Kunin, W.E. (1998). Extrapolating species abundances across spatial scales. *Science*, 281, 1513–1515.
Lennon, J.J., Koleff, P., Greenwood, J.J.D. & Gaston, K.J. (2001). The geographical structure of British bird distributions: diversity, spatial turnover and scale. *J. Anim. Ecol.*, 70, 966–979.
Lennon, J.J., Kunin, W.E. & Hartley, S. (2002). Fractal species distributions do not produce power-law species area distribution. *Oikos*, 97, 378–386.
Maddux, R.D. (2004). *Am. Nat.*, in press.
Mandelbrot, B.B. (1977). *Fractals, Form, Chance and Dimension*. Freeman, San Francisco, CA.
Muriel, N-N. & Mangel, M. (1999). Species–area curves based on geographic range and occupancy. *J. Theor. Biol.*, 196, 327–342.
Nee, S., Gregory, R.D. & May, R.M. (1991). Core and satellite species: theory and artefacts. *Oikos*, 62, 83–87.
Peckham, S. (1995). New results for self similar trees with applications to river networks. *Water Resour. Res.*, 31, 1023–1029.
Preston, F.W. (1960). Time and space variation of species. *Ecology*, 41, 611–627.
Rosenzweig, M.L. (1995). *Species Diversity in Space and Time*. Cambridge University Press, Cambridge.
Sayles, R.S. & Thomas, T.R. (1978). Surface topography as a non-stationary random process. *Nature*, 271, 431–434.
Šťastný, K., Bejček, V. & Hudec, K. (1996). *Atlas of Breeding Bird Distribution in the Czech Republic 1985–1989*. Nakladatelství a vydavatelství H & H, Jihlava, (in Czech).
Storch D., Gaston K. & Cepák J. (2002). Pink landscapes: *1/f* spectra of spatial environmental variability and bird community composition. *Proc. R Soc. Lond. B*, 269, 1791–1796.
Storch, D., Konvicka, M., Benes, J., Martinková, J. & Gaston, K.J. (2003a). Distributions patterns in butterflies and birds of the Czech Republic: separating effects of habitat and geographical position. *J. Biogeogr.*, 30, 1195–1205.
Storch, D. & Šizling, A.L. (2002). Patterns in commoness and rarity in central European birds: reliability of the core-satellite hypothesis. *Ecography*, 25, 405–416.
Storch, D., Šizling, A.L. & Gaston, K.J. (2003b). Geometry of the species–area relationship in central European birds: testing the mechanism. *J. Anim. Ecol.*, 72, 509–519.
Sugihara, G. (1980). Minimal community structure: an explanation of species abundance patterns. *Am. Nat.*, 116, 770–787.
Tjørve, E. (2003). Shapes and functions of species–area curves: a review of possible models. *J. Biogeogr.*, 30, 827–835.
Veitzer, S. & Gupta, V. (2000). Random self-similar river networks and derivations of generalized Horton laws in terms of statistical simple scaling. *Water Resour. Res.*, 36, 1033–1048.
Williams, M.R. (1995). An extreme-value function model of the species incidence and species–area relationship. *Ecology*, 76, 2607–2616.
Witte, J.-P.M. & Torfs, P.J.J.F. (2003). Scale dependency and fractal dimension of rarity. *Ecography*, 26, 60–68.

Editor, Pablo Marquet
Manuscript received 8 September 2003
First decision made 20 October 2003
Manuscript accepted 3 November 2003

APPENDIX 1: THE SHAPE OF THE SAR BETWEEN SUBSEQUENT A_{sat}

To study the shape of the SAR, it is useful to express the first and the second derivatives of the SAR in a log–log space (hereafter lnSAR). If the second derivative is positive, then the lnSAR is convex (upward accelerating), if it is negative, the lnSAR is concave, and if it is equal to zero, the lnSAR is linear within the studied intervals, i.e. the SAR can be expressed as a power law. Using the finite-area model (eqn 7) the first derivative of the lnSAR, i.e. of the function $\ln(\bar{S}(\ln(A)))$ (hereafter $\ln \bar{S}(\ln A)$, between $A_{\text{sat } j}$ and $A_{\text{sat } j+1}$, can be expressed as

$$\frac{\mathrm{d}}{\mathrm{d}\ln A}\ln \bar{S}_{(A)} = \frac{\sum_{i=S_{\text{sat}(A)}+1}^{S_{\text{tot}}} \pi_i z_i A^{z_i}}{\sum_{i=S_{\text{sat}(A)}+1}^{S_{\text{tot}}} \pi_i A^{z_i} + S_{\text{sat}(A)}}. \quad \text{(A1)}$$

The second derivative is then

$$\frac{\mathrm{d}}{\mathrm{d}\ln A}\left(\frac{\mathrm{d}}{\mathrm{d}\ln A}\ln \bar{S}_{(A)}\right) = \frac{\left(\sum_{i=S_{\text{sat}(A)}+1}^{S_{\text{tot}}} \pi_i z_i^2 A^{z_i}\right)\left(\sum_{i=S_{\text{sat}(A)}+1}^{S_{\text{tot}}} \pi_i A^{z_i} + S_{\text{sat}(A)}\right) - \left(\sum_{i=S_{\text{sat}(A)}+1}^{S_{\text{tot}}} \pi_i z_i A^{z_i}\right)^2}{\left(\sum_{i=S_{\text{sat}(A)}+1}^{S_{\text{tot}}} \pi_i A^{z_i} + S_{\text{sat}(A)}\right)^2}, \quad \text{(A2)}$$

which after multiplying gives

$$\frac{\mathrm{d}^2}{\mathrm{d}\ln A^2}\ln \bar{S} = \frac{\sum_{(\forall i,j)} \pi_i \pi_j \left(z_i - z_j\right)^2 A^{\left(z_i + z_j\right)} + S_{\text{sat}(A)} \sum_{i=S_{\text{sat}(A)}+1}^{S_{\text{tot}}} \pi_i z_i^2 A^{z_i}}{\left(\sum_{j=S_{\text{sat}(A)}+1}^{S_{\text{tot}}} \pi_i A^{z_i} + S_{\text{sat}(A)}\right)^2}. \quad \text{(A3)}$$

As $S_{\text{sat}(A)}$ differs from zero for all $j > 0$, and thus the second additive term in the numerator is positive, the ln SAR is always convex (upward accelerating) within individual intervals between $A_{\text{sat } j}$ and $A_{\text{sat } j+1}$, except the case of $A < A_{\text{sat } 1}$ (i.e. $S_{\text{sat}(A)} = 0$) and $z_i = z_j$ for all combinations i, j, when it is linear in the log–log space. As z_i drops down in the point $A_{\text{sat } i}$ the slope of the ln SAR also drops down here, according to eqn A1 (Fig. 1b). This is the moment that pushes the slope of the whole lnSAR down, although different z_i push the slope always up.

APPENDIX 2: CONSTRUCTION OF THE SELF-SIMILAR SPATIAL DISTRIBUTIONS

To test the hypothesis that the spatial distribution of birds within the CR and CE can be considered as self-similar we simulated 500 species assemblages consisting from the same number of species as observed (i.e. 142 for the CR and 193 for CE), each of them having the *FD* equal to the *FD* of respective species. The self-similar distributions of these simulated species were placed randomly over the grid of 16 × 16 cells.

The construction of a self-similar distribution started with a square whose side was two times longer than the side of the whole grid. In each subsequent step, the square was scaled down by the linear factor k, and four squares of the resulting size were placed randomly within the original square without an overlap (see Fig. A1). This procedure was repeated until the squares were smaller than the basic grid cell. The *FD*s of these constructed distributions are equal to $\ln(4)/\ln(k)$ (Mandelbrot 1977; Hastings & Sugihara 1993), and thus it is easy to set the *FD* by adjusting k. The *FD*s used for the construction were estimated using the coefficient z in the finite-area model (see eqns 4 and 5) obtained by extracting the slope of the regression line for the relationship between $\log A$ and $\log p_{occ}$ for $A < A_{sat}$. This estimate (hereafter p*BD*) is better than the classical box counting (Hastings & Sugihara 1993), which was proven by 600 simulations of self-similar distributions with *FD*s of 0.1, and 0.2, and 0.3,…, 2.0 (30 simulations for each *FD*). We calculated both classical *BD* and p*BD* for every simulation, and compared these measurements with *FD*s used for the construction of these self-similar distributions. The difference was 0.26 ± 0.19 for *BD* and 0.09 ± 0.19 for p*BD*.

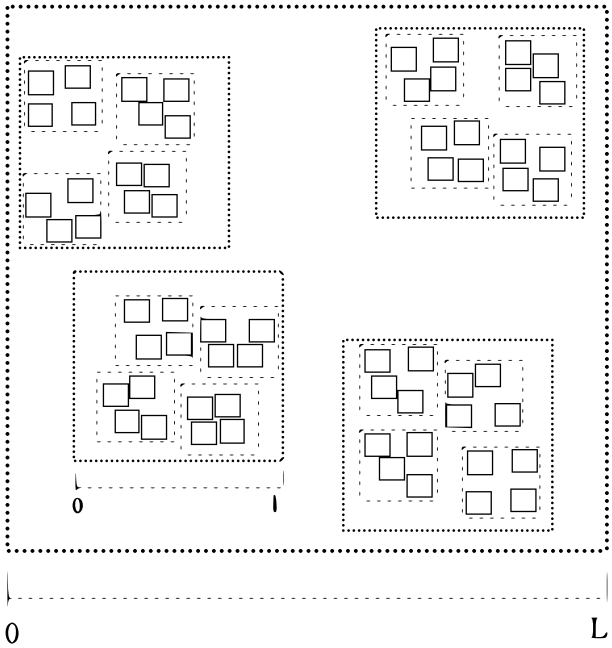

Figure A1 The first three stages of the construction of a random self-similar distribution, which was used for the testing the hypothesis that the spatial distribution of birds within the Czech Republic (CR) and central Europe (CE) can be considered as self-similar. Elementary steps of the construction comprise lessening of each square by the factor k (i.e. $L = kl$) and the random placement of the four obtained copies within the original square without overlaps. The grid of 16 × 16 cells whose total area was one-fourth of L^2 was then placed randomly within the area, and p_{occ} for each area was calculated using all possible quadrats of that area within the grid.