

Nordic Society Oikos

On the Implications of Species-Area Relationships for Endemism, Spatial Turnover, and Food Web Patterns
Author(s): J. Harte and A. P. Kinzig
Reviewed work(s):
Source: *Oikos*, Vol. 80, No. 3 (Dec., 1997), pp. 417-427
Published by: Blackwell Publishing on behalf of Nordic Society Oikos
Stable URL: http://www.jstor.org/stable/3546614
Accessed: 30/06/2012 17:41

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Blackwell Publishing and *Nordic Society Oikos* are collaborating with JSTOR to digitize, preserve and extend access to *Oikos.*

http://www.jstor.org

OIKOS 80: 417–427. Copenhagen 1997

On the implications of species-area relationships for endemism, spatial turnover, and food web patterns

J. Harte and A. P. Kinzig

Harte, J. and Kinzig, A. P. 1997. On the implications of species-area relationships for endemism, spatial turnover, and food web patterns. – Oikos 80: 417–427.

Consequences of species-area relationships (SARs) of the form $S = cA^z$ are derived. One consequence is an endemics-area relationship (EAR); it is of the same power-law form as the SAR but with an exponent z' that is a function only of z and that always exceeds unity. An explicit formula is derived for the dependence of species turnover in space on census plot size, interplot distance, and the SAR exponent z; this formula can be used to determine z over spatial scales that are too large to permit direct estimate of z by censusing of nested patches. The areal dependence of link-species patterns observed in food webs is also examined; SARs are shown to imply the approximate, but not exact, area-independence of link-species relationships of the form $L = kS^y$, where L is the number of trophic links and $y \leq 2$. A relationship between the average range of species in a habitat patch and the exponent z is also derived, leading to the result that average range is a decreasing function of both patch area and z.

J. Harte, Energy and Resources Group & Dept of Environmental Science, Policy and Management, Univ. of California, Berkeley, CA 94720, USA (jharte@violet.berkeley.edu). – A. P. Kinzig, Dept of Ecology and Evolutionary Biology & Center for Energy and Environmental Studies, Princeton Univ., Princeton, NJ 08544, USA.

Over 30 years ago, pioneering work by Preston (1962) demonstrated that a species-area relationship (SAR) of the form $S \propto A^z$, with $z \approx 0.25$, could be derived from plausible assumptions about the distribution of individuals among species. Even earlier, SARs of that approximate form had been shown to describe available data on species distributions (Arrhenius 1921), and since then there has been successful application of SARs to species distributions both on islands and on contiguous patches within larger habitats (MacArthur and Wilson 1963, Diamond 1975, Rosenzweig 1995). Conservation biologists have applied SARs to the design of preserves and to estimation of the number of extinctions expected under habitat loss (e.g., Diamond 1975, Pimm et al. 1995). And theoretical analyses have sharpened our understanding of the conditions under which SARs are likely to be observed (Caswell and Cohen 1993).

Despite these advances, a number of basic questions about SARs remain unresolved. The SAR offers information about the average distribution of species within a biome, but some species will be distributed extensively throughout a biome while others may have very limited distributions. Does the SAR provide information about the number of species with limited versus broad distributions, and how those numbers depend on the area censused? The answer to this question may have major implications for conservation biology because species extinction due to habitat degradation will depend more on the number of species confined to the degraded habitat than on the total number of species found there (Pimm et al. 1995).

A related question concerns species turnover in space or, equivalently, the degree of overlap in the species composition of habitat patches at various distances

Accepted 2 May 1997

Copyright © OIKOS 1997
ISSN 0030-1299
Printed in Ireland – all rights reserved

from one another. How does species turnover in space depend on the SAR exponent z, the distance between patches, and on patch size?

The SAR may also have implications for the existence of regularities in food webs. While a lake has a distinct boundary, a patch of forest within Amazonia or a patch of desert within the southwestern United States does not. Are hypothesized food web patterns relating the number of trophic interactions in an ecosystem to the number of species valid independently of how the boundary is drawn around the study area?

A final question concerns the relationship between species ranges and the SAR exponent z. How do average species ranges (within a taxonomic category) vary with the SAR exponent, and does a latitudinal trend in species range imply a latitudinal trend in the SAR exponent?

To address these and other related questions, we develop here a theoretical framework within which many of the implications of an SAR can be mathematically derived. In contrast to Preston (1962), Caswell and Cohen (1993), and other researchers, we do not derive an SAR from more basic assumptions, but rather we assume an SAR and then explore its consequences for the issues raised above.

Definitions and assumptions

Consider a large area comprising a distinct biome – for example, lowland tropical forest within Amazonia, the South Western Botanical Province of Australia, the Sonoran desert, or the Arctic tundra. We take the area of such an entire biome to be A_0; the area A_0 contains S_0 number of species. We assume that most of the species found in our distinct biome are found only there: that is, most of the S_0 species are found only in A_0. Initially, we treat only the case in which a species-area relationship holds across all scales within the biome down to some minimum area, A_m, so that $S = cA^z$ for all $A_m \leq A \leq A_0$; then we extend the analysis to show how the scale-dependence of z can be extracted from spatial turnover data.

Even within a biome for which the species-area relationship holds generally, violations may occur for irregularly or unusually shaped patches – for example, patches that are very narrow but very long (Condit et al. 1996). Therefore, we will consider only "nicely shaped" areas of increasing bisection within A_0. The first bisection will cut A_0 in half; these halves will have area A_1, with $A_1 = A_0/2$. Similarly, the next bisection will divide A_0 into quadrants; the areas of the quadrants are labeled A_2, with $A_2 = A_0/4$. We can continue in this fashion, halving subpatches to produce patches of area $A_i = A_0/2^i$ until the minimum area A_m is reached. (Thus, for instance, the octants have area A_3.) If the original biome of area A_0 is roughly square (or a $2L \times 1L$ rectangle, where L is some length) then the original patch and all of its subdivisions will be squares or 2×1 rectangles.

We call the patch with area A_i the P_i patch; there are S_i species within the P_i patch. We assume that favorable or unfavorable patches are randomly distributed within the biome; there are no large-scale gradients favoring significantly more species, say, in the east half over the west half, or the southeast corner over the northeast corner. This assumption is consistent with assuming that the species-area relationship holds across the biome; for instance, the eight values of S_3 – corresponding to the number of species found in each of the eight octants of A_0 – are taken to be roughly equal for all octants (with $S_3 = cA_3{}^z$), with no octant harboring significantly more or fewer species than the other octants. This assumption does not preclude habitat heterogeneity, though that heterogeneity cannot result in significant differences in the number of species found in patches of identical size. For biomes for which this assumption does not hold – or for scales within biomes at which this assumption is violated (most likely at very small areas) – our derivation below is not applicable (just as an SAR with constant z is no longer applicable under those conditions).

If the species-area relationship holds in each of the P_i patches, we can then write

$$S_i = cA_i{}^z \tag{1}$$

or, substituting our formula for A_i given above ($A_i = A_0/2^i$)

$$S_i = c(A_0/2^i)^z = 2^{-iz}cA_0{}^z = 2^{-iz}S_0 = a^iS_0 \tag{2}$$

where

$$a = 2^{-z}. \tag{3}$$

Because the exponent z can differ from one broad taxonomic group to another, in what follows we usually intend S_i to refer to the species within a specific taxonomic group.

The species-area relationship contains information about average species distributions within a biome. In particular, the SAR implies that any species in P_0 has, on average, a probability $1/2^z$ $(=a)$ of being found if we survey only half of P_0. That is, out of the S_0 species found in P_0, $S_0/2^z$ species will be found in P_1. More generally, if we know that a species is found within an area A_i and we know nothing more about the specifics of its distribution within that area, then the odds that it is found in half that area (A_{i+1}) is just $1/2^z$ $(=a)$.

But we can use the SAR to go beyond this information on the *average* probability distributions of species within the biome. Some of our S_0 species, for instance,

will be confined to one half of P_0, while others will be found in both halves. Thus the S_0 species can be divided into two distinct groups; those with broad distributions (greater than the average distribution; found in both halves of P_0) and those with limited distributions (less than the average distribution; found only in one half or the other of P_0).

We call those species with limited distributions the *endemic* species: in particular, we label those species found only in one half of P_0 (i.e., found only in one P_1 patch and not the other) the E_1 species. Continuing in this fashion, species found only in one of the four quadrants (the P_2 patches) are labeled E_2 species, species found in only one of the eight octants are labeled E_3 species, etc. Note that membership in the E_{i+1} group implies membership in the E_i group; for instance, if a species is in only one of four quadrants, it can be in only one of two halves.

Our definition of endemic – species found *only* in a sub-patch of a larger distinct biome – is entirely reasonable, but differs from some other uses of the word endemic. For instance, endemic is often applied as a label to species found only with very limited global distributions, such as the Galapagos Islands or Madagascar. The theories and formulas we develop below are *not* applicable to understanding the global distribution of "hot spots" of endemism. Our theories and formulas, however, do address the variations in species distributions implied by a species-area relationship, and identify the numbers of species confined to sub-patches within a larger biome due to the constraints of the species-area relationship; it is those species with limited distributions within the larger biome that we call endemic species.

Derivation of an SAR for endemic species

Consider again the area of A_0 with S_0 species, where $S_0 = cA_0^z$. If the area A_0 is divided in half, each of the halves of area A_1 will have $S_1 = c(A_0/2)^z = aS_0$ species. Some number of species E_1 will be found only in one of the two A_1 areas; another group of species, labeled U_0 (for "ubiquitous" species) will be found in both A_1 areas. Conservation of species numbers dictates that

$$2E_1 + U_0 = S_0. \tag{4.0}$$

But E_1 must be given by $S_0 - S_1 = (1-a)S_0$. To see that this is true, consider a patch of area A_0 with $a = 1/2^z = 0.84$ ($z = 0.25$) and $S_0 = 1000$ species. Dividing the area in half gives $S_1 = aS_0 = 840$ species in each half of P_0. If, however, there are 840 species in the left half of P_0, and 1000 species in the whole of P_0, there must be 160 species found in the right half of P_0 that are not found in the left half of P_0. Thus $E_1 = 160$, or

$$E_1 = S_0 - S_1 = (1-a)S_0 \tag{5.1}$$

and

$$U_0 = S_0 - 2E_1 = (2a-1)S_0. \tag{6.0}$$

Again, in the absence of large-scale gradients in species numbers, these relationships will hold true for the right and left halves of P_0, or the upper and lower halves of P_0.

The E_1 species can be further subdivided. Out of the E_1 species found only in the west half of A_0, some will be found only in the NW quadrant (the E_2 species) and some will be found in both the NW and SW quadrants (the U_1 species). As discussed above, we assume the E_1 species will on average be distributed within P_1 according to the probabilities dictated by the species-area relationship; that is, if only the upper half of P_1 is surveyed, each of the E_1 species will have a probability $1/2^z$ ($=a$) of being found in the surveyed area. In other words, aE_1 of the E_1 species will be found in area A_2.

Continuing with the example above, out of the 160 species found only in the West half of P_0 (the E_1 species), 134 species (a $E_1 = 0.84 \cdot 160$) will be found in the NW quadrant, which means that 26 species of the 160 species must be confined to the SW quadrant. But these species confined to one quadrant are the E_2 species; thus

$$E_2 = E_1 - aE_1 = (1-a)E_1 = (1-a)^2 S_0. \tag{5.2}$$

Furthermore, some number U_1 of the E_1 species must be found in both the upper- and lower-left quadrants, but again, if E_1 is subdivided into those species found only in one quadrant within a half (E_2 species) and those species found in both quadrants within a half (U_1 species), we have, by conservation of species number

$$E_1 = 2E_2 + U_1 \tag{4.1}$$

or

$$\begin{aligned} U_1 &= E_1 - 2E_2 = (1-a)S_0 - 2(1-a)^2 S_0 \\ &= (1-a)(2a-1)S_0. \end{aligned} \tag{6.1}$$

(Note that the U_1 species are not the only species found in both the NW and SW quadrants; some of the U_0 species – those found in both halves of P_0 – will also be found in both quadrants.)

We can continue writing equations for the E_i and U_i species in this fashion; doing so gives

$$E_i = (1-a)^i S_0 \tag{5.i}$$

and

$$U_i = (2a - 1)(1 - a)^i S_0. \quad (6.i)$$

Note that, since E_i is always subdivided into E_{i+1} and U_i groups, we can write

$$E_i = U_i + 2E_{i+1} \quad (4.i)$$

consistent with (4.1) above. Finally, our assumption that most or all of the S_0 species found in P_0 are found only in P_0 means that the S_0 species are endemic to P_0, or $E_0 = S_0$, consistent with $i = 0$ in eq. (4.i). See Fig. 1 for a pictorial representation of the various groups of endemic and ubiquitous species.

Eqs (5.i) imply that endemic species follow an endemics-area relationship (EAR) of the form

$$E_i = c' A_i^{z'} \quad (7)$$

where $c' = cA_0^{(z-z')}$ $(= S_0/A_0^{z'})$ and $1/2^{z'} = 1 - 1/2^z$, or

$$z' = -\ln(1 - 1/2^z)/\ln(2). \quad (8)$$

If $z < 1$ (as it must be), then eq. (8) implies that $z' > 1$. In other words, the total number of endemics is a rapidly increasing function of area; small patches have very few numbers of endemic species (i.e., species confined to only that patch within the larger biome), but these numbers increase quickly as patch size increases. For values of $z = 0.10$, 0.25, and 0.40, $z' = 3.90$, 2.65, and 2.05, respectively.

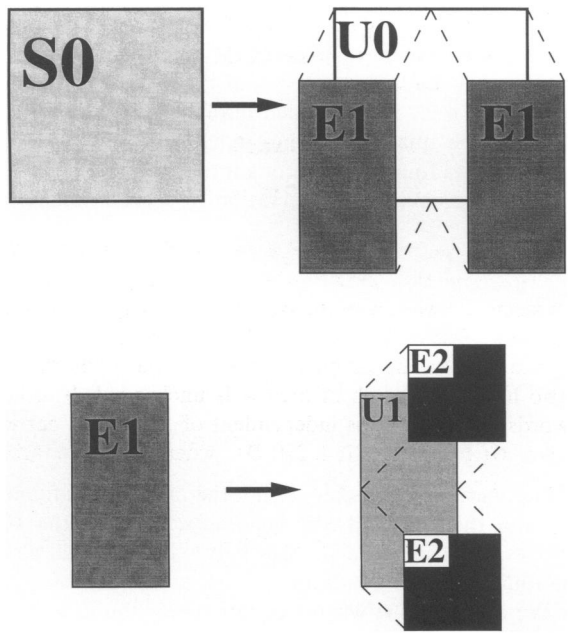

Fig. 1. The distributions of various types of species within a biome. The E_i species are endemic to a patch of area A_i, meaning they are found only in that area and nowhere else in the larger biome. Ubiquitous species, as defined in the text, are denoted by U_i. Note that the U_1 species are a subset of the E_1 species, but that U_0 and U_1 species are two distinct groups.

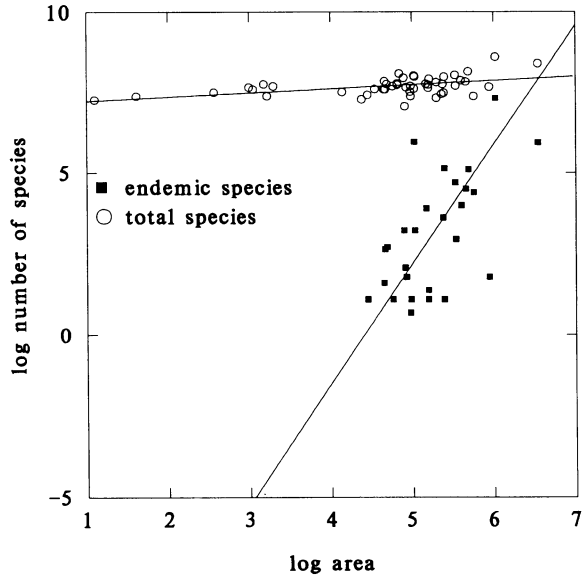

Fig. 2. Ln-ln plots of total plant species and endemic plant species versus area for each of the 48 contiguous states of the United States. The endemism data exclude states with one or no endemic species because of shortcomings of $\ln(S)$ vs $\ln(A)$ plots at small values of S (Williams 1995) and both data sets exclude exotic species. For the endemic-species data, the straight-line fit has slope $z' = 3.7$, $R^2 = 0.38$, $p < 0.001$; for the total-species data, the straight-line fit has slope $= z = 0.13$, $R^2 = 0.22$, $p < 0.01$. Best fits to slopes are determined from the geometric means of values obtained by $\ln(S)$ vs $\ln(A)$ and $\ln(A)$ vs $\ln(S)$ regressions. Data developed by The Nature Conservancy with the cooperation of John Kartesz of the North Carolina Botanical Garden (Larry Morse, pers. comm.) and from Gentry (1986).

Our eq. (7) means that the number of species confined to certain patches within a biome under the constraints of the species-area relationship can be calculated directly, without resorting to the iterative process described above. If we return to our example with $S_0 = E_0 = 1000$ and $z = 0.25$, we find, from eq. (8), that $z' = 2.65$. Furthermore, $E_{i+1}/E_i = 1/2^{z'}$, so that there are $E_1 = 1000/2^{2.65} = 159$ species confined to the left (or right) half of the biome; there are $E_2 = 159/2^{2.65} = 25$ species found only in each of the four quadrants of our biome, and $E_3 = 4$ species found only in each of the eight octants of our biome.

Data for distributions of endemic species within contiguous patches that span enough areal variations to provide a rigorous test of our predicted endemics-area relationship (eqs 7 and 8) are scant (Anderson 1995). We were able to obtain data on the total number of plant species and the total number of endemic plant species in each of the 48 contiguous states of the United States. The data for the total number of plant species are consistent with an SAR exponent of $z = 0.13$ for plant species across the continental U.S. (See Fig. 2.) Eq. (8)

implies that, for an SAR with exponent $z = 0.13$, there should be a corresponding EAR with exponent $z' = 3.5$. Fitting the data for endemic plants to our EAR gives $z' = 3.7$. This is a remarkable agreement given that our theory contains assumptions that are at best applicable to a more homogeneous biome than the 48 states. Clearly, more systematic comparisons of theory with data are required to truly test the prediction for the distribution of endemics derived here.

Finally, in derivations below, we will find it convenient to express the S_i in terms of the U_j. Consider $S_0 = E_0$. Successive application of eqs (4.i) gives

$$S_0 = U_0 + 2U_1 + 4U_2 + 8U_3 + 16U_4 \cdots \tag{9.0}$$

S_1, on the other hand, gives all of the species in one half of P_0, and consists of all of the U_0 species (those found in both halves of P_0) and the E_1 species (those found only in one or the other half of A_0), so $S_1 = U_0 + E_1$. Again, successive application of eqs (4.i) gives

$$S_1 = U_0 + U_1 + 2U_2 + 4U_3 + 8U_4 + \cdots \tag{9.1}$$

What about the S_2 species? The S_2 species (total number of species in a quadrant) will consist of all E_2 species (those found only in one of the four quadrants), plus the U_1 species (those species found only in one half, but both quadrants of that half), plus some fraction of the U_0 species (those species found in both halves of P_0, but not all will be found in a quadrant of P_0). What is that fraction f? We can write

$$S_2 = a^2 S_0 = E_2 + U_1 + fU_0 = (1-a)^2 S_0 + (2a-1)(1-a)S_0 + f(2a-1)S_0.$$

Solving for f gives $f = a$, so that

$$S_2 = E_2 + U_1 + aU_0$$

and successive iteration of eqs (4.i) gives

$$S_2 = aU_0 + U_1 + U_2 + 2U_3 + 4U_4 + \cdots \tag{9.2}$$

Continuing in this fashion gives

$$S_3 = a^2 U_0 + aU_1 + U_2 + U_3 + 2U_4 + \cdots \tag{9.3}$$

and, in general,

$$S_i = a^{i-1}U_0 + a^{i-2}U_1 + \cdots + U_{i-1} + U_i + 2U_{i+1} + 4U_{i+2} + \cdots \tag{9.i}$$

Species turnover

Species turnover in space is defined as one minus species commonality, where species commonality is the fraction of species in common to two patches. Denoting turnover by T and commonality by χ, we have $T = 1 - \chi$. Species turnover should depend on the distribution of species within a biome (i.e., the SAR exponent z), the area of surveyed patches, and the distance between surveyed patches; in particular, turnover should decrease as the surveyed area A increases, and increase as the distance between surveyed patches increases. We show here that an approximate formula for species turnover follows from the SAR and our results about species distributions derived above.

Consider two censused patches within P_0, one of area A with $S(A)$ species ($S(A) = cA^z$) and one of area A' with $S(A')$ species. We assume, for now, that the distance D between the centers of the two patches is very large compared to either $(A)^{1/2}$ or $(A')^{1/2}$. We denote by $N(A, A', D)$ the number of species in common to the two patches. Commonality, defined as the number of species in common divided by the average number of species in the two patches, $[S(A) + S(A')]/2$, is then given by

$$\chi(A, A', D) = 2N(A, A', D) \div (S(A) + S(A')). \tag{10}$$

Consider the odds that a species picked at random from the patch of area A' is also found in the patch of area A a distance D away. That odds, which we denote by $G(A, A', D)$, are given by the number of species in common to the two patches divided by the number in A', or from eq. (10):

$$G(A, A', D) = N(A, A', D)/S(A') = [S(A) + S(A')]\chi(A, A', D)/[2S(A')]. \tag{11}$$

Now suppose we double the size of the patch of area A'. Each of the species in A' is still nearly exactly a distance D away from the center of the patch of area A (because $D \gg (A')^{1/2}$) and so if a species is selected at random from the patch of area $2A'$, the odds that it also lies in the patch of area A is unchanged. In other words, $G(A, A', D)$ is independent of A' and, in particular, $G(A, A, D) = G(A, 2A, D)$, where A is an arbitrary area. Using eq. (11),

$$G(A, 2A, D) = [S(A) + S(2A)]\chi(A, 2A, D)/[2S(2A)] \tag{12}$$

and

$$G(A, A, D) = \chi(A, A, D). \tag{13}$$

Equating these two expressions, we have

$$\chi(A, 2A, D) = \frac{2S(2A)\chi(A, A, D)}{S(A) + S(2A)}$$
$$= \frac{2 \cdot 2^z}{(1^z + 2^z)} \chi(A, A, D). \quad (14)$$

Again considering $\chi(A, A', D)$, we can continue with this same reasoning. If we double the size of the patch of area A, then we increase the number of species in that patch and thus increase the odds that a species selected at random a distance D away from that patch of area A is also found there. The species number is increased by a factor of 2^z and so the odds must increase by that same factor, or

$$G(2A, 2A, D) = 2^z G(A, 2A, D) = 2^z G(A, A, D). \quad (15)$$

Again using eq. (11), $G(2A, 2A, D) = \chi(2A, 2A, D)$, and so

$$\chi(2A, 2A, D) = 2^z \chi(A, A, D). \quad (16)$$

Eq. (16) holds true when D and z remain constant, and the censused areas double. Thus, eq. (16) generalizes readily to the result $\chi(A, A, D) = (A)^z f(z, D)$, where f is some function. For the more general case in which the patch areas are unequal, eq. (14) generalizes readily to the result $\chi(A, A', D) = 2(AA')^z f(z, D) \div [(A)^z + (A')^z]$. The form of the function f is constrained, however, because χ must be a dimensionless fraction and D^2 is the only other parameter with dimensions of area. Therefore we must have

$$\chi(A, A, D) = (A/D^2)^z g(z). \quad (17)$$

To determine the function g we note that if the two patches share a common edge, so that $D^2 = A$, then the number of species in common to the two patches, $N(A, A, A^{1/2})$, is exactly given by the difference between twice the number of species in a patch of area A and the number in the patch of area $2A$, or $N(A, A, A^{1/2}) = 2c(A)^z - c(2A)^z$. $\chi(A, A, A^{1/2})$ is given by dividing this expression for N by the average number of species in a patch, cA^z, or $\chi(A, A, A^{1/2}) = [2c(A)^z - c(2A)^z]/c(A)^z = 2 - 2^z$. But since $\chi(A, A, A^{1/2})$ also equals $(A/A)^z g(z) = g(z)$, we have

$$g(z) = 2 - 2^z \quad (18)$$

for the case in which the two patches share an edge.

Consider, next, the case in which the two patches touch at one corner, so that they occupy the diagonally opposite quadrants (hereafter referred to as "opposite quadrants") of a square of area $4A$; in this case, $D = (2A)^{1/2}$. As before, we can derive an exact formula for $N(A, A, (2A)^{1/2})$. We derive such a formula for an arbitrary square, P_i, with S_i species. First, we note that the species in S_i found in two opposing quadrants (e.g., upper-right and lower-left quadrants) of patch P_i must be drawn from the U_j species, with $j \leq i$. For instance, if we consider the P_0 patch, a U_0 species (found in both halves of P_0) could be found in two diagonally opposing quadrants, but a U_1 species (found only in one half, but both quadrants of that half) cannot be found in both diagonally opposing quadrants.

Consider first the U_i species; these species are found in both the left and right halves of P_i. How many of the U_i species will be found in our opposing quadrants? Some of the U_i will be found only in the upper or lower half of P_i, and thus cannot be found in diagonally opposing quadrants. To find those upper or lower species, we need to make a different division, bisecting P_1 into upper and lower halves instead of left and right halves. The E_i species are those species found only in the upper or lower half; the U_{i+1} species are those confined to the upper (or lower) half, but found in both the left and right sections of the upper (or lower) half. Note that the U_{i+1} species from the upper/lower division are, by definition, part of the U_i species using a left/right division. Thus, the number of species in U_i that will be found in opposing upper left and lower right quadrants is just $U_i - 2U_{i+1}$. In words, take those species found in both left and right halves (U_i) and subtract the subset of those species found only in the upper or lower halves ($2U_{i+1}$). Using eq. (6), this becomes $(2a-1)^2(1-a)^i S_0$, or $(2a-1)$. Therefore, if a species is in U_i, it has a probability $(2a-1)$ of being found in diagonally opposing quadrants of P_i.

What about the U_j species, with $j < i$? What is the probability that they will be found in opposing quadrants of P_i? For these U_j species, we expect the probabilities to be lower than $(2a-1)$, since we have no a priori information about the locations within P_i of the species in the remaining U_j. That is, while we knew that all of the U_i species were already in both halves of P_i, we have no such information about each of the remaining U_j species. Some may be in both halves of P_i, some may not. A reasonable approach, however, would be to assume that the $U_{j<i}$ species and the S_i species are governed by the same distribution statistics; that is, the fraction f of the U_j species found in opposing quadrants is the same as the fraction f of the S_i species found in opposing quadrants. Eqs (9.i) give

$$S_i = a^{i-1}U_0 + a^{i-2}U_1 + \cdots + U_{i-1} + U_i + 2U_{i+1} + \cdots \quad (9.i)$$

Remembering that the U_i species have a probability $(2a-1)$ of being in opposing quadrants, and the $U_{j>i}$ species have a probability 0 of being in opposing quadrants gives

$$fS_i = f(a^{i-1}U_0 + a^{i-2}U_1 + \cdots + U_{i-1}) + (2a-1)U_i \quad (19)$$

Solving for f gives $f=(2a-1)^2$. So, the fraction of species in a patch that are found in two opposing quadrants is just $(2a-1)^2$. The number of species found in both of the two opposing quadrants of patch P_i is then $(2a-1)^2S_i$.

Using this result, we can return to the two censused patches that touch at one corner. The area of each of the censused patches is A and these patches are the diagonally opposing quadrants of a patch P_i of area $4A$; the distance between the centers of the censused patches is $(2A)^{1/2}$. Thus $N(A, A, (2A)^{1/2}) = (2a-1)^2S_i = (2^{1-z}-1)^2c(4A)^z$. Dividing by the average number of species in either patch, $c(A)^z$, we get $\chi(A, A, (2A)^{1/2}) = 4^z(2^{1-z}-1)^2$. But since $\chi(A, A, (2A)^{1/2}) = g(z)(A/2A)^z$, we have

$$g(z) = (2-2^z)^2 2^z. \tag{20}$$

The two expressions for $g(z)$ given by eqs (20) and (18) are not identical but they differ by only a very small amount if z is small. For example, for $z = 0.1$, 0.2, and 0.3, the ratios of the two expressions are 0.995, 0.978, and 0.947 respectively. They differ because the assumption that led to eq. (14), namely that the patches are separated by a distance large compared to the square root of their area, is not satisfied when the patches are touching.

If the patches are separated, then we can derive approximate, but not exact, expressions for χ and $g(z)$. In particular, consider the case in which the patches are separated by a distance D and are situated within a larger patch as shown in Fig. 3. The number of species in common to the two large opposing quadrants (within which the smaller censused patches are imbedded), each of area $D^2/2$, was derived above to be exactly $(2a-1)^2S(2D^2)$. But, if a species is in a quadrant of area $D^2/2$, then by the SAR, the odds that it is in the patch within that quadrant of area A are $[A/(D^2/2)]^z$. If we assume that the species distributions within the two opposing quadrants containing our censused patches are independent of each other, then the simultaneous odds of a species in one quadrant being in the censused patch in that quadrant and of a species in the other quadrant being in the censused patch in that quadrant are just the square of this expression, or $[A/(D^2/2)]^{2z}$. These odds can be applied to that group of species found in both quadrants of an opposing pair – the number of these species is again given by $(2a-1)^2S(2D^2)$. Thus

$$N(A, A, D) \approx (A/D^2/2)^{2z}(2a-1)^2S(2D^2). \tag{21}$$

Recalling that $\chi(A, A, D) = N(A, A, D)/S(A)$ $S(A) = c(A)^z$, and $a = 2^{-z}$, we have

$$\chi(A, A, D) \approx (A/D^2/2)^{2z}(2^{1-z}-1)^2c(2D^2)^z/cA^z$$
$$= 2^z(2-2^z)^2(A/D^2)^z. \tag{22}$$

Note that this derived expression for $g(z)$ for the case of widely separated patches, $2^z(2-2^z)^2$, is identical to that for patches touching at one corner (eq. 20). Eq. (22) is only an approximate expression for χ because the assumption of independent species distributions made above is not strictly true – species found in the lower left quadrant, for instance, are more likely to be found near the lower left corner of the upper right quadrant than near the upper right corner of that quadrant. By placing the patches of area A in the middle of the quadrants, as shown in Fig. 3, we attempt to minimize this error.

If we had chosen the two patches to lie not within opposing quadrants of the larger square but rather within two quadrants touching along a common edge, then we would, by the same argument as above, reach the conclusion that $\chi(A, A, D) \approx (2-2^z)(A/D^2)^z$, in agreement with the expression for $g(z)$ given in eq. (18).

Because the two expressions for $g(z)$ differ only slightly, we choose the simpler one (eqs 17 and 18) and hence express species turnover in space by

$$T(A, A, D) = 1 - \chi(A, A, D) \approx 1 - (2-2^z)(A/D^2)^z. \tag{23}$$

Plots of turnover as a function of D for 3 values of A (corresponding to censused patches that comprise $1/2^6$, $1/2^{10}$, and $1/2^{20}$ of the entire biome) and 2 values of the SAR exponent ($z = 0.25$ and 0.10) are shown in Fig. 4. The qualitative features of these curves can be compared to turnover data for birds collected by Cody (1993) for the rainforest (his Fig. 13.9) and mulga (his Fig. 13.11) habitats of Australia. Cody uses an empirical function for turnover that increases like (constant)D^β and finds a fitted value of 0.28 for β. In contrast, our derived function for turnover behaves like $1 - (\text{constant})D^{-2z}$ with no free parameters. His data are at least qualitatively consistent with the shapes of our derived curves. A test of our turnover formula requires site-specific determination of the SAR exponent z along with turnover data between patches of

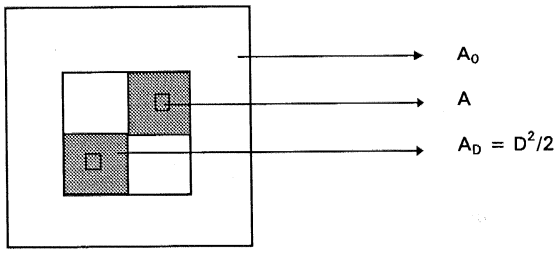

Fig. 3. Configuration of patches for the derivation of species turnover in space. The small patches of area A are the censused patches located a distance D apart. The shaded squares of area A_D are the two opposing quadrants referred to in the text and are constructed to allow determination of the number of species found in common to the two small patches.

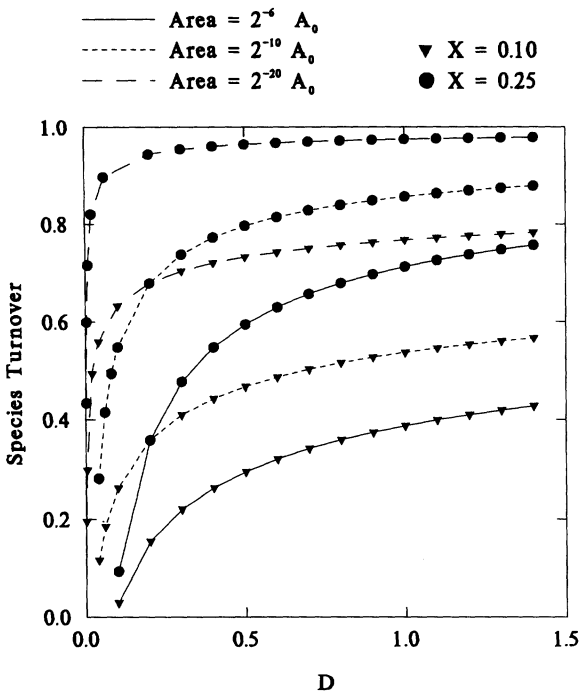

Fig. 4. Plots of species turnover in space as a function of the SAR exponent z, the area of censused plots, and the interplot distance, D. The total biome is assumed to have length and width = unity, so the maximum value of $D = \sqrt{2}$.

specified area and distance apart. In subsequent publications we will describe results of a more systematic comparison of our prediction with turnover data derived from bird distributions in Amazonia and from vegetation distributions in montane meadows.

We note that validation of eq. (23) would provide a simple method for determining the species-area exponent over large spatial scales. Direct estimation of the SAR exponent would require census data in large nested areas; such censuses are often impractical to carry out. On the other hand, use of eq. (23) permits estimation of the SAR exponent by a straightforward measure of species turnover between patches of either different area or of different distance from one another. In the likely event that the species-area exponent, z, is scale-dependent, eq. (23) combined with measurement of the dependence of turnover on interpatch distance, D, provides a means of determining z at a series of scales characterized by different ranges of D.

Species-linkage relationships and SARs

Next we look at putative relations between the number of species (S) in ecosystems and the number of predator-prey linkages (L) among those species. One such species-linkage relationship (SLR), called the constant connectance hypothesis, states $L = kS^2$, where k is the same constant in all ecosystems; a more general form for such an SLR is $L = kS^y$ for some fixed value of y and of k. The constant connectance hypothesis (Martinez 1993) is consistent with data from a wide range of ecosystems, with $k \sim 0.1$, although other forms of an SLR with $y < 2$ cannot be ruled out (Pimm et al. 1991).

The concept of scale invariance in the context of food web regularities has been used to characterize relationships that hold across ecosystems with differing numbers of species. Our interest here is in a different but related type of invariance – namely, absence of dependence of the SLR parameters (k and y) on the area within an ecosystem that is sampled. The existence of an SAR, of course, means that the degrees to which the two types of invariance hold are interconnected. As argued below, it is impossible for y and k to be strictly independent of sample area if species number depends on area. We will assume a fixed "universal" value of y and determine the magnitude of the areal dependence of k under the assumption that eqs (5i) and (6i) are valid.

To see that an SAR and an SLR generally cannot be simultaneously valid across a range of spatial scales, consider the two limiting cases: $a = (1/2)^z = 1$ and $a = 1/2$ (corresponding to $z = 0$ and $z = 1$, respectively). In the limiting case in which $a = 1$ ($z = 0$, all species are found everywhere) there is no conflict; the same species, and therefore the same linkages, are found in each P_i; if any SLR holds in any area within an ecosystem it will hold in any other. But in the limit of $a = 1/2$ ($z = 1$, each species having, in the limit, zero range), the species in any two non-overlapping patches within the ecosystem are, by the SAR, totally different from each other. Then the total number of species and linakges in P_0 are twice those in either of the P_1's. But according to the SLR there should be $(2S)^y/S^y = 2^y$ more linkages in P_0 than in P_1. For $y = 2$, for example, this implies four times more linkages in P_0 than in P_1. A similar problem arises for all pairs P_i, P_{i+1}. In summary, the SLR cannot be areal-invariant if $y > 1$, unless $a = 1$ (all species found everywhere).

To determine the magnitude of violation of areal-invariance of the SLR, we consider first, for simplicity, the case of an SLR with $y = 2$, and an SAR with an arbitrary value of z. The relationship $L = kS^2$ can be interpreted to mean that, on average, each species in the ecosystem eats a fixed fraction, k, of all the other species in the ecosystem. But, clearly, species can only eat those species that they encounter, or share space with. Species that are endemic to a small region contained within P_0 cannot eat, or be eaten by, species endemic to a non-overlapping region within P_0. So we assume that each species eats a fixed fraction, k', of the subset of species it shares space with. Obviously k' must be greater than k as long as some species do not share space (i.e., as long as $z > 0$).

Next, we decompose the linkages within the ecosystem of interest into a complete and non-overlapping set of linkages among component subsets of species that share space. For example, consider the patch P_0, with species number S_0 and L_0 linkages. The U_0 species overlap in space with one another and thus the interactions among the U_0 species should contribute a term $k'U_0U_0$ to L_0. Similarly, the U_1 species overlap in space with one another within *each* of the two halves of P_0 and thus the interactions among the U_1 species should contribute a term $2k'U_1U_1$. Similar reasoning leads to a sum of terms $2^ik'U_iU_i$ in the full expression for L_0. Next consider the cross term arising from U_0 species eating U_1 species. The number of such linkages will equal $2k'U_0U_1$, with the factor of two arising because the U_0 species can eat a fraction k' of the U_1 species in the left half of P_0 and the same fraction of the (different) U_1 species in the right half. Similarly, U_1 species eating U_0 species leads to another term, $2k'U_1U_0$, of equal magnitude.

We can proceed in this way through all of the cross terms to get

$$\begin{aligned} L_0 = k'\{&U_0(U_0 + 2U_1 + 4aU_2 + 8a^2U_3 + \cdots) \\ &+ 2U_1(U_0 + U_1 + 2U_2 + 4aU_3 + \cdots) \\ &+ 4U_2(aU_0 + U_1 + U_2 + 2U_3 + \cdots) \\ &+ 8U_3(a^2U_0 + aU_1 + U_2 + U_3 + \cdots) \\ &+ \cdots\}. \end{aligned} \tag{24.0}$$

Factors of a, a^2, etc., arise in this expression for the same reason they do in eqs (9.i) (and as explained in the derivation of those equations). We note that this expression differs from that obtained by simply taking the expression for S_0 given in eq. (9.0), squaring it, and multiplying by k'.

Similar expressions can be written for each of the L_i with careful attention to the combinatorics. For example:

$$\begin{aligned} L_1 = k'\{&U_0(U_0 + U_1 + 2aU_2 + 4a^2U_3 + \cdots) \\ &+ U_1(U_0 + U_1 + 2U_2 + 4aU_3 + \cdots) \\ &+ 2U_2(aU_0 + U_1 + U_2 + 2U_3 + \cdots) \\ &+ 4U_3(a^2U_0 + aU_1 + U_2 + U_3 + \cdots) \\ &+ \cdots\} \end{aligned} \tag{24.1}$$

and

$$\begin{aligned} L_2 = k'\{&U_0(a^2U_0 + aU_1 + aU_2 + 2a^2U_3 + \cdots) \\ &+ U_1(aU_0 + U_1 + U_2 + 2aU_3 + \cdots) \\ &+ U_2(aU_0 + U_1 + U_2 + 2U_3 + \cdots) \\ &+ U_3(2a^2U_0 + 2aU_1 + 2U_2 + 2U_3 + \cdots) \\ &+ \cdots\}. \end{aligned} \tag{24.2}$$

In each of eqs (24.i) we can rewrite the U_i that appear in that equation in terms of S_i using eqs (2) and (6). Thus each equation will be of the form:

$$L_i = k'\alpha^i S_i^{\,2}. \tag{25}$$

The degree to which the α_i depend on i measures the scale-dependence of the SLR. In other words, if $\alpha_i =$ constant for all patches P_i, then the number of linkages is given by $kS_i^{\,2}$, where $k' = k\alpha_i$, regardless of the area of the patch surveyed; if α_i does depend on i (or, equivalently, on the size of the patch surveyed), this species-linkage relationship will not be scale invariant. The α_i are determined by summing the infinite series that arise in eqs (24.i) and then summing the sums, a tedious but straightforward task. We have evaluated the expressions for the resulting summations numerically for $a = 0.84$ ($z = 1/4$) and obtained for the first three cases: $\alpha_0 = 0.875$, $\alpha_1 = 0.948$, $\alpha_2 = 0.966$. In the limit $i \to \infty$, α_i can be evaluated analytically and it approaches the limiting value of 1; in other words, an SLR of the form $L_i = k'S_i^{\,2}$ is approached as the area surveyed within a biome gets smaller (i gets larger).

We note from the numerical results above that the convergence toward this limit is extremely rapid. If $a = 1$ ($z = 0$), then all the α_i are identically equal to unity. If, however, $a \sim 0.5$ ($z \sim 1$), then the α_i are considerably smaller than the numerical values given above and, more importantly, the areal dependence is greater in the sense that convergence toward the limiting case is slower. If $y < 2$, the areal dependence is, as would be expected, even weaker than if $y = 2$.

Our results establish the nearly exact spatial scale invariance of all hypothesized link-species relationships ($y \leq 2$); even for the most scale-dependent case (the constant connectance hypothesis, with $y = 2$) and for $z = 1/4$, the link-species relationship is quite insensitive to the censused area, particularly for patches with area less than or equal to $\sim 1/4$ of the area of the entire biome.

Species ranges and SARs

We derive next a relation between average species range for those species found in P_i, which we define as $\langle r_i \rangle$, and the parameter a ($= 1/2^z$). Consider first the range of those species found in the entire biome. Eq. (9.0) provides a decomposition of S_0 into the sets U_i ($i = 0, \ldots, \infty$). The range of the species that comprise the set U_i is $\sim (A_0/2^i)^{1/2}$ because those species are found throughout an area A_i but are not found in A_{i+1} (recall that U_i is a subset of the E_i species, but not a subset of the $U_{j<i}$ species). Thus, using eq. (9.0), $\langle r \rangle$ can be written as weighted average:

$$\langle r_0 \rangle = (1/s_0) \sum_{i=0}^{\infty} 2^i U_i \sqrt{A_0/2^i}\,. \tag{26}$$

Using eqs (2) and (6i), the sum can be carried out, yielding

$$\langle r_0 \rangle = \sqrt{A_0}\,\frac{(2a-1)}{1-\sqrt{2}(1-a)}\,. \tag{27.0}$$

We note that if $a = 1$ ($z = 0$, S is independent of area, every species exists everywhere), then $\langle r_0 \rangle = \sqrt{A_0}$ as expected. The smallest possible value of a is 1/2, corresponding to $z = 1$, and thus $S = cA$. In this case, $\langle r_0 \rangle = 0$, because space is effectively fractionated into finer and finer distinct habitats for distinct species, and no species is found in more than one micropatch. For $z = 1/4$, $\langle r_0 \rangle = 0.88\sqrt{A_0}$.

More generally, exact expressions for all the $\langle r_i \rangle$ can be derived from eqs (2) and (9):

$$\langle r_i \rangle = \sqrt{A_0}(2a-1)\left\{\left(\frac{1-a}{\sqrt{2}a}\right)^i (\langle r_0 \rangle - \langle r_\infty \rangle) + \frac{\sqrt{2}}{a(\sqrt{2}+1)-1}\right\}, \tag{27.i}$$

where

$$\langle r_\infty \rangle = \frac{\sqrt{A_0}\sqrt{2}(2a-1)}{a(\sqrt{2}+1)-1}, \tag{28}$$

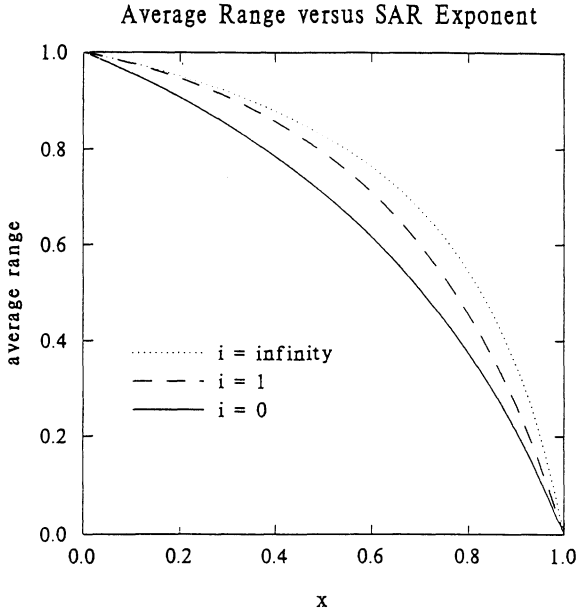

Fig. 5. Average range of all species in a biome of area $A_0/2^i$ as a function of the SAR parameter a for three values of i. The limit $i \to \infty$ corresponds to the censused patch size approaching zero. Range is expressed in units of $\sqrt{A_0}$.

and $1/2 \le a \le 1$. For $i = 0$, 1, and ∞, Fig. 5 shows the average range, in units of $\sqrt{A_0}$, as a function of the parameter a.

It follows from eq. (19) that $\langle r_i \rangle > \langle r_j \rangle$ if $i > j$. Thus, the average range of all species sampled on, say, a 100-hectare plot within some biome is *less* than the average range of all species sampled on a 1-hectare plot within that biome. The larger the exponent z, the larger this effect. For example, if $z = 1/2$ ($a = 0.71$), then $\langle r_\infty \rangle / \langle r_0 \rangle = 1.2$, whereas if $z = 1/4$ ($a = 0.84$), then $\langle r_\infty \rangle / \langle r_0 \rangle = 1.07$.

Conclusions

Under a set of assumptions that are implicit in the existence of SARs of the form $S = cA^z$, such as the absence of significant gradients in species distributions, we have shown that SARs contain considerably more predictive power than hitherto has been appreciated.

1. Defining "endemic" species in an area to be those found only in that area, we derived from the species-area relationship a new species-area relationship for endemics; it is of the form $E = c'A^{z'}$, where z' is an explicit function of z, always exceeds unity, and increases with decreasing z. For $z = 1/4$, $z' = 2.65$.

2. We have derived an approximate formula (eq. 23) for species turnover in space, defined as one minus the fraction of species in common to two separated patches. The formula predicts explicit dependence of turnover on the area A of censused patches, the distance D between patches, and the SAR exponent z. The fraction of species in common to two patches is shown to be proportional to $(A/D^2)^z$. This result, combined with measurement of the dependence of turnover on interpatch distance, D, provides a means of determining z at a series of scales characterized by different ranges of D.

3. Food web analysis has suggested that a possible power-law relationship exists between the number of species and the number of trophic links in ecosystems, but has left unresolved the question of how such relationships can appear to hold independent of ecosystem area. In particular, area dependence could result in detectable empirical violations when food web data compiled at different spatial scales are compared. Simple heuristic arguments demonstrate that such area dependence must exist and that it increases in strength as z decreases and the power relationship in the links-species relation gets more non-linear. We showed here that SARs imply the approximate, but not exact, area-independence of link-species relationships of the form links $= cS^y$, $y \le 2$; dependence on area is nearly negligible at $z = 1/4$, $y = 2$ and decreases further with decreasing z and y.

4. We derived a formula for the average range of species in a patch of a biome in terms of the SAR exponent, z, and the area of the censused patch. We showed that average range is a decreasing function of patch size and of the SAR exponent z and that the dependence of range on patch area is greatest at large z.

Initial comparisons with limited data sets describing the distribution of endemics and species turnover in space are qualitatively consistent with our predictions, but further work is needed to systematically confront our theoretical results with reliable data on plant and animal distributions.

Acknowledgements – We are grateful to Joel Cohen, Jonathan Dushoff, Ken Harte, Stephen Hubbell, Deborah Jensen, John Lawton, Neo Martinez, Helene Muller-Landau, and Stuart Pimm for enlightening discussions. We also thank Larry Morse of the Nature Conservancy for sharing with us unpublished data on plant distributions. We thank the Pew Charitable Trusts and the Mellon Foundation for financial support.

References

Anderson, S. 1995. Area and endemism. – Q. Rev. Biol. 69: 451–471.

Arrhenius, O. 1921. Species and area. – J. Ecol. 9: 95–99.

Caswell, H. and Cohen J. E. 1993. Local and regional regulation of species-area relations: a patch-occupancy model. – In: Ricklefs, R. and Schluter, D. (eds), Species diversity in ecological communities. Univ. of Chicago Press, Chicago, pp. 99–107.

Cody, M. L. 1993. Bird diversity components within and between habitats in Australia. – In: Ricklefs, R. and Schluter, D. (eds), Species diversity in ecological communities. Univ. of Chicago Press, Chicago, pp. 147–158.

Condit, R., Hubbell, S. P., LaFrankie, J. V., Sukumar, R., Manokaran, N., Foster, R. and Ashton, P. S. 1996. Species-area and species-individual relationships for tropical trees: a comparison of three 50-ha plots. – J. Ecol. 84: 549–562.

Diamond, J. 1975. The island dilemma: lessons of modern biogeographic studies for the design of nature reserves. – Biol. Conserv. 7: 129–146.

Gentry, A. 1986. Endemism in tropical versus temperate plant communities. – In: Soule, M. (ed.), Conservation biology: the science of scarcity and diversity. Sinauer, Sunderland, MA, pp. 153–181.

MacArthur, R. and Wilson, E. 1963. An equilibrium theory of insular zoogeography. – Evolution 17: 373–387.

Martinez, N. 1993. Effects of scale on food web structure. – Science 260: 242–243.

Pimm, S., Lawton, J. and Cohen, J. 1991. Food web patterns and their consequences. – Nature 350: 669–674.

– , Russell, G., Gittleman, J. and Brooks, T. 1995. The future of biodiversity. – Science 269: 347–354.

Preston, F. W. 1962. The canonical distribution of commonness and rarity: Part I. – Ecology 43: 185–215.

Rosenzweig, M. L. 1995. Species diversity in space and time. – Cambridge Univ. Press, Cambridge.

Williams, M. 1995. An extreme-value function model of the species incidence and species-area relations. – Ecology 76: 2607–2616.