Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Theoretical Biology 232 (2005) 119–126

Journal of Theoretical Biology

www.elsevier.com/locate/yjtbi

Range size in mid-domain models of species diversity

Héctor T. Arita*

Instituto de Ecología, Universidad Nacional Autónoma de México, Apartado Postal 70-275, CP 04510 México DF, México

Received 30 January 2004; received in revised form 29 June 2004; accepted 4 August 2004
Available online 23 September 2004

Abstract

Geographical patterns of species diversity have been examined using mid-domain null models, in which the ranges of individual species are simulated by randomly arranging them on a bounded one- or two-dimensional continent. These models have shown that structured patterns in the geographical distribution of biodiversity can arise even under a fully stochastic procedure. In particular, mid-domain models have demonstrated that the random generation of ranges of different sizes and locations can produce a gradient of species diversity similar to the one found in real assemblages, with a peak at the middle of a continent. A less explored feature of mid-domain models is the pattern of range-size frequency distribution. Numerical simulations have provided some insights about the geographic pattern of average range size, but no exploration of the shape of range-size frequency distributions has been carried out. Here I present analytical and numerical models that generate explicit predictions for patterns of range size under the assumptions of mid-domain models of species diversity. Some generalizations include: (1) Mid-domain models predict no geographic gradient of average range size; the mean range size of species occurring at any point on a continent is constant (0.5 of the extent of the continent in the one-dimensional model, 0.25 of the area of the continent in the two-dimensional case); (2) Variance in range size is lowest at the middle of a continent and highest near the corners of a square-shaped continent; (3) The range-size frequency distribution is highly right-skewed at any point of a continent, but the skewness is highest near the corners. Despite their alleged weaknesses, mid-domain models are adequate null models against which real-world patterns can be contrasted.

Keywords: Biogeography; Mid-domain effect; Null models; Range size; Species diversity

1. Introduction

One of the oldest and most thoroughly studied patterns of biogeography is the tendency for species richness (the number of species occurring at a given site) to increase from the poles towards the tropics (Hawkins, 2001; Willig et al., 2003). More than 30 evolutionary and ecological hypotheses have been posed to explain this gradient of species diversity (Hawkins et al., 2003), and there is no single best explaining cause for it, as the details of the gradient vary with geographic location, scale, and geological history of the area of study. A related pattern, the latitudinal gradient of the size of species ranges, has been the focus of interest since Stevens (1989) proposed the so-called Rapoport's rule as a possible explanation for the higher species richness in the tropics. The rule postulates that, on average, tropical species have smaller ranges than temperate taxa, due to the tolerance of temperate species to a broader gamut of environmental conditions. Rapoport's "rule" has been shown to apply for some groups in some continents, but exceptions are numerous, and its generality has been seriously questioned (Gaston et al., 1998; Kerr, 1999; Gaston, 2003).

A new perspective to the study of latitudinal gradients of species richness came from the mid-domain models developed during the 1990s (Colwell and Hurtt 1994; Willig and Lyons, 1998; Lees et al., 1999; Colwell and Lees, 2000; Laurie and Silander, 2002; Grytnes, 2003; Colwell et al., 2004; Pimm and Brown, 2004). These models, which use randomization algorithms to simulate

---

*Tel.: +52-55-56-22-89-96; fax: +52-55-56-22-19-76.
E-mail address: arita@ecologia.unam.mx (H.T. Arita).

0022-5193/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2004.08.004

the spatial arrangement of species within a bounded domain that represents a continent, predict patterns for the latitudinal gradient of species richness that in some cases seems to coincide with reality (Lees et al. 1999; Colwell and Lees, 2000; McCain, 2003). In particular, due to geometric constraints imposed by hard boundaries to the distribution of species, models predict a parabolic curve of species richness, with a peak of 0.5 times the total number of species at the center of a latitudinal gradient, similar to that found in real continents. In two-dimensional models, a similar peak in species richness, with a maximum value of 0.25 times the total number of species, appears in the center of a bounded domain defined by a latitude and a longitude (Bokma et al., 2001). The name "mid-domain models" comes from these predicted patterns of highest richness at the middle of the gradient.

Although the focus of mid-domain models has been on the gradient of species richness, the models also predict particular patterns for the latitudinal gradient of range sizes (Lyons and Willig, 1997; Colwell and Lees, 2000; Koleff and Gaston, 2001). In particular, numerical simulations generate no gradient at all if the average range of species intersecting a given latitude is used as the metric, following Stevens (1989), or a reversed Rapoport pattern if the metric used is the average range of species whose range midpoint coincides with the focal latitude, following the method of Rohde et al. (1993). The focus has been on the pattern of average range size as a function of latitude, but the pattern of variation at each latitude has not been examined. Also, the discussion has centered mostly on one-dimensional domains.

The study of range-size frequency distributions has important theoretical and practical implications, although the details of such distributions are still very poorly understood (Gaston, 2003). At the global and continental scales, several mathematical models have been tested to describe the patterns of range-size frequency distribution, but no general model has been developed (Williamson and Gaston, 1999). At the regional and local scales, the patterns of range-size variation among species are linked to a wide variety of macroecological parameters, such as the slope z of species–area relationships (Rosenzweig, 1995; Ney-Nifle and Mangel, 1999; Lyons and Willig, 2002), beta diversity (Rodríguez and Arita, 2004), and in general to the patterns in the scaling of species diversity (Arita and Rodríguez, 2002). Additionally, in conservation planning, local and regional range-size frequency distributions are key tools in the identification of rare species, those taxa with the most restricted distributions among a particular set (Gaston, 2003). In all those cases, the value of the average range is of little use, and a full description of the frequency distribution of range sizes is needed.

Surprisingly, very little theoretical or empirical work has been done to explore the relationship between continental, regional, and local patterns of range-size variation. In particular, no model has explored the implications of mid-domain models on the detailed shape of range-size frequency distributions, and the emphasis up to now has been on numerical simulations focusing only on average range, and limited to the one-dimensional case. Here I develop analytical models allowing the prediction of the exact frequency distribution of range sizes for species occurring at different latitudinal positions under the assumptions of a fully stochastic one-dimensional mid-domain model. Additionally, I explore theoretical implications of mid-domain models for the two-dimensional case, for which no description of the predicted range-size frequency distribution had been previously reported.

2. The one-dimensional case

The model is based on the one-dimensional fully stochastic mid-domain model developed by Colwell and Hurtt (1994), Willig and Lyons (1998), and Colwell and Lees (2000). In this model, a continent is simulated by a domain on the interval (0, 1), and the ranges of a set of species are modeled by lines arranged within the domain (Fig. 1). The continent is bounded by "hard" limits, so species cannot occur outside the (0, 1) domain. Each range is defined by the limits L_1 and L_2 ($0 \leqslant L_1, L_2$; $L_2 > L_1$), so the size of the range (the length of the corresponding line) is $R = L_2 - L_1$, and the midpoint is $M = (L_1 + L_2)/2$. For a given point p along the domain, which is analogous to a latitude, the species richness is the number of lines intersecting that point.

The fully stochastic mid-domain model consists in generating species ranges that are randomly located along the domain. To do so, Colwell and Hurtt (1994), in their model 2, proposed a method in which points (M, R) are randomly selected from the universe of all possible pairs of values, which are arranged on an isosceles triangle on an M vs. R plot (Fig. 2). Alternatively, the range of each species can be defined by the position of its extreme points (L_1 and L_2; Willig

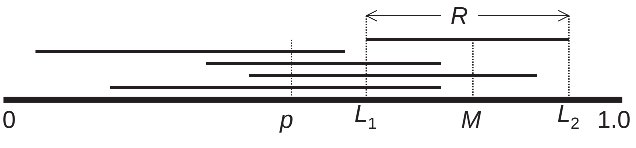

Fig. 1. The geographic distribution of species as represented in one-dimensional mid-domain models. A bounded continent is modeled by a domain on the interval (0, 1). The ranges of species are represented with lines, five of which are pictured here, lying within the domain. Species richness at a given point p is the count of ranges that intersect that point. The size and position of each range are determined by generating two random points (L_1, L_2) along the domain. The midpoint of each range (M) is the average value of L_1 and L_2, and the range size (R) is the distance between those points.

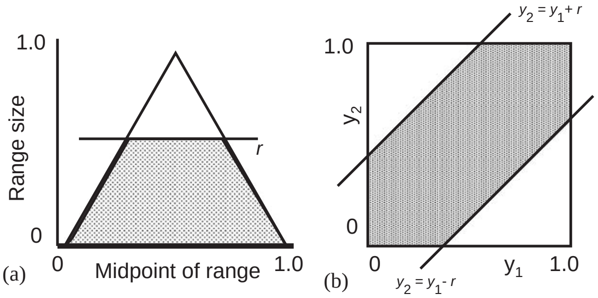

Fig. 2. Two ways of representing permissible values for species ranges arranged on a one-dimensional domain. Possible pairs of (location of midpoint M, range size R) values arrange on a triangle defined by the points (0, 0), (1, 0), and (0.5, 1) (a). Points for which $R<r$, where r is an arbitrary value, lie below a horizontal line (shaded area). Points (y_1, y_2) constituting the extremes of species ranges form a unit square on a (y_1, y_2) plot (b). Points for which the range size $R=|Y_2-Y_1|$ is equal to a constant r generate pairs of straight lines with slope 1 and Y_2intercept equal to $\pm r$. The probability $P(R\leqslant r)$ is equal to the shaded area bounded by the two parallel lines.

and Lyons, 1998; Colwell and Lees, 2000). To generate a species range, two points (Y_1, Y_2) are randomly placed along the domain, such that $L_1 = MIN(Y_1, Y_2)$ and $L_2 = MAX(Y_1, Y_2)$. Thus, for each Y_i the density function is $f(y) = 1(0\leqslant y\leqslant 1)$, and, because Y_1 and Y_2 are independent, their joint density is $f(y_{1,}y_2) = f(y_1)\, f(y_2) = 1$. The range size R is the latitudinal extent and is a function of Y_1 and Y_2, $R(Y_1, Y_2) = MAX(Y_1, Y_2) - MIN(Y_1, Y_2) = |Y_1 - Y_2|$, so $0\leqslant R\leqslant 1$. The universe of possible values for Y_1, Y_2 can be represented by a unit square (Fig. 2).

2.1. Continental frequency distribution of range sizes

To deduce the frequency distribution of range sizes, we need to find $F_R(r) = P(R\leqslant r)$, where r is an arbitrary limit ($0\leqslant r\leqslant 1$). $F_R(r)$ can be found using either the M vs. R plot of Colwell and Hurtt (1994) or a model based on the extreme points of species ranges (a "two-hit" model). Here I develop the latter option, because the procedure will be required in other sections of the paper. However, a similar demonstration, with identical results, is possible using the triangular model. In Fig. 2a, it is easy to see that the region over which $R\leqslant r$ is the set of points of the triangle below an arbitrary line r, so $F_R(r)$ is the shaded area proportional to the area of the triangle of permissible M, R points.

Using the two-hit model, to find $F_R(r)$ we need to define the points (y_1, y_2) for which $|y_1 - y_2|\leqslant r$. These points correspond to the shaded area in Fig. 2b. Arbitrarily defining y_2 as the dependent variable, pairs of values yielding a constant r value arrange along straight lines with slope$\,=1$ and intercept $y_2 = r$ (for $y_2\geqslant y_1$) or $y_2 = -r$ (for $y_2\leqslant y_1$). Because the bivariate density function $f(y_1, y_2) = 1$ is uniform over the square $0\leqslant y_1\leqslant 1$, $0\leqslant y_2\leqslant 1$, the distribution function $F_R(r) =$ $P(R<r)$ is the volume of a solid with height equal to 1 and cross section equal to the shaded area shown in Fig. 2b. Therefore,

$$F_R(r) = P(R<r) = \int\!\!\int_{|y_1-y_2|<r} f(y_1, y_2)\mathrm{d}y_1\,\mathrm{d}y_2$$

$$= \int\!\!\int_{|y_1,y_2|<r} (1)\mathrm{d}y_1\,\mathrm{d}y_2. \qquad (1)$$

Although the distribution function can be obtained by integration, it is easier to find the result using simple geometry. Since the side of the triangles not shaded in Fig. 2b is $1-r$, it is easy to show that the distribution function is $F_R(r) = 1-(1-r)^2 = 2r-r^2$. Thus, the density function is:

$$f_R(r) = \frac{\mathrm{d}}{\mathrm{d}r}(2r - r^2) = 2 - 2r, \qquad (2)$$

the expected value for R is

$$E(R) = \int_0^1 r(2 - 2r)\mathrm{d}r = 1/3. \qquad (3)$$

and the variance is

$$V(R) = E(R^2) - [E(R)]^2 = \int_0^1 r^2(2 - 2r)\mathrm{d}r - (1/3)^2$$

$$= 1/18 \qquad (4)$$

Thus, the frequency distribution of the size of ranges is a linear decreasing function of range size, with a mean value of 0.33 and variance 0.055. According to the fully stochastic model, then, species with small ranges should be more common than widespread species on a continental scale. Colwell and Hurtt (1994) reported this distribution based on numerical simulations. Before that, MacArthur (1957) reported Eq. (2) but, in the context of his analysis of the relative abundance of species, went no further in exploring the implications of bounded limits on the frequency distribution of abundances (Colwell and Lees, 2000).

2.2. Frequency distribution of range size at a given latitude

Next, we are interested in finding the frequency distribution of ranges for species occurring at a given point p ("latitude") located along the domain. I describe in detail the procedure to find the distribution for the case where $p\leqslant 0.5$, and present the general formula for any p. The probability that a range defined by two random variables Y_1, Y_2 on the interval (0, 1) intersects a point p is $P=2p-2p^2$ (Willig and Lyons, 1998; Colwell and Lees, 2000). The rationale is that in order for the range of a species to intersect p, its limits $L_1 = MIN(y_1, y_2)$ and $L_2 = MAX(y_1, y_2)$ must be in opposite sides of p. Therefore, the probability of intersection is one minus the probability $P(y_1, y_2<p) = p^2$ minus the probability

$P(y_1,\ y_2 > p) = (1-p)^2$. This approach is illustrated in Fig. 3b, where points forming ranges intersecting point p lie on the two rectangles of area $p(1-p)$, which add up to $2p-2p^2$. Values for which the range $R_p < r$ for a given p (shaded areas in Fig. 3) are defined by these rectangles and by the two lines shown in Fig. 2b.

When obtaining the distribution function of R_p, because of geometric constraints, two different cases ($p \leqslant 0.5$ and $p \geqslant 0.5$), with three variants each, are possible. For $p \leqslant 0.5$, the three possible outcomes yield the following distribution functions:

$$F_{R_p}(r) = \begin{cases} \frac{r^2}{2(p-p^2)} & \text{for } r \leqslant p, \\ \frac{r-p/2}{1-p} & \text{for } p \leqslant r \leqslant (1-p), \\ 1 - \frac{1}{2(p-p^2)} + \frac{r}{p-p^2} - \frac{r^2}{2(p-p^2)} & \text{for } r \geqslant (1-p), \end{cases} \tag{5}$$

which correspond to the following density functions:

$$f_{R_P} = \begin{cases} \frac{r}{p-p^2} & \text{for } r \leqslant p, \\ \frac{1}{1-p} & \text{for } p \leqslant r \leqslant (1-p), \\ \frac{1-r}{p-p^2} & \text{for } r \geqslant (1-p), \end{cases} \tag{6}$$

Therefore, the expected value and variance for R_p are:

$$\begin{aligned} E(R_p) &= \int_0^1 r f_{R_p}(r)\,\mathrm{d}r \\ &= \int_0^p \left[\frac{r^2}{p-p^2}\right] \mathrm{d}r + \int_p^{1-p} \left[\frac{r}{1-p}\right] \mathrm{d}r \\ &\quad + \int_{1-p}^1 \left[\frac{r(1-r)}{p-p^2}\right] \mathrm{d}r = \frac{1}{2}, \\ V(R_p) &= E(R_p^2) - [E(R_p)]^2 \\ &= \int_0^1 r^2 f_{R_p}(r)\,\mathrm{d}r - (1/2)^2 = \frac{1-2(p-p^2)}{12}. \end{aligned} \tag{7}$$

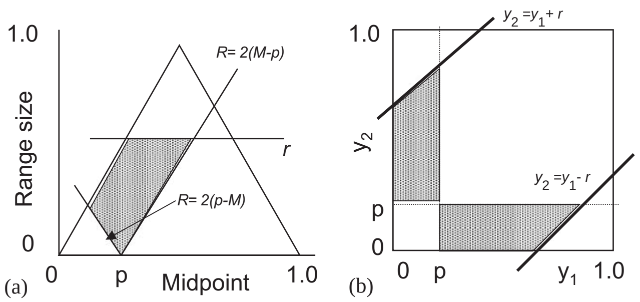

Fig. 3. Two graphic representations of the probability of a given range size in mid-domain models. Probability that $R \leqslant r$ for ranges intersecting the point p along the domain on the interval (0, 1) is equal to the shaded area on the triangle formed by permissible (M, R) points, divided by the area of the triangle (a). The same probability shown for the "two-hit" model: shaded areas on the unit square defined by a sample (Y_2, Y_1) of size $n = 2$ from the uniform distribution (b). Both figures show the case where $p \leqslant r \leqslant (1-p)$.

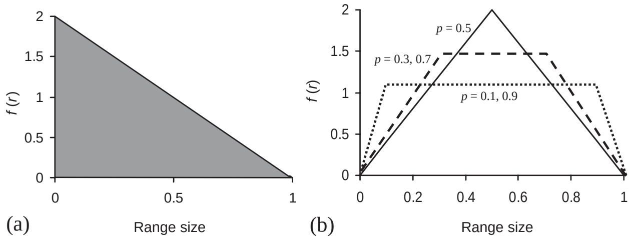

Fig. 4. Density function of range size in a one-dimensional mid-domain model. Function for all species in a continent under the fully stochastic mid-domain model (a). The same function for species whose range intersect a given point $p = 0.1$ or 0.9, 0.3 or 0.7, and 0.5 along the domain on the interval (0, 1) (b).

With a similar approach, the case where $p \geqslant 0.5$ yields:

$$\begin{aligned} E(R_p) &= \int_0^1 r f_{R_p}(r)\,\mathrm{d}r = \int_0^{1-p} \left[\frac{r^2}{p-p^2}\right] \mathrm{d}r \\ &\quad + \int_{1-p}^{p} \left[\frac{r(1-p)}{p-p^2}\right] \mathrm{d}r + \int_p^1 \left[\frac{r-r^2}{p-p^2}\right] \mathrm{d}r = \frac{1}{2}, \\ V(R_p) &= \frac{1-2(p-p^2)}{12}. \end{aligned} \tag{8}$$

Identical results can be obtained from the triangular model (Fig. 3a). Points corresponding to ranges intersecting point p are bounded by lines crossing the point (p, 0) and having slopes ± 2, and by the sides of the triangle of permissible M, R points (Laurie and Silander, 2002). It can be shown that the possible combinations of lines yield exactly the two cases with three variants each found for the two-hit model, and therefore derive also into Eqs. (5)–(8).

The density functions for several values of p are shown in Fig. 4b. Note that functions for points p and $1-p$ have identical values, and that the expected value for range size is constant regardless of the value of p. This finding coincides with the numerical results of Colwell and Hurtt (1994) and Willig and Lyons (1998) that the fully stochastic model predicts no gradient of average range size along the domain. However, previous studies had not examined the pattern of variation from the mean. Here I show that the fully stochastic mid-domain model predicts the existence of a gradient in the variance of the range-size frequency distribution, this being lowest (1/24) at the middle of the gradient and highest (1/12) at both ends of the latitudinal gradient.

3. The two-dimensional model

3.1. Area of range at the continental scale

The model discussed in section 2 can be readily extended to a two-dimensional case simply by imagining a square-shaped continent forming a two-dimensional domain on the interval (0, 1). The range of a species in this continent will be a rectangle defined by four points (Y_1, Y_2, X_1, X_2), two in each dimension of the domain, which are independent random variables drawn from the uniform distribution on the interval (0, 1). As in the one-dimensional model, the range of a species has a given extent in each dimension (R_y,R_x), defined by the position of the Y_1, Y_2, X_1, X_2 as $R_y = MAX(Y_1, Y_2) - MIN(Y_1, Y_2)$ $R_x = MAX(X_1, X_2) - MIN(X_1, X_2)$. We know that the variables are independent, and from Eq. (2), that the density function for each range along one of the dimensions is $f_{R_i}(r_i) = 2(1 - r_i)$ $(0 \leqslant r_i \leqslant 1)$. Therefore, the joint density function is given by:

$$f_{R_xR_y}(r_x, r_y) = f(r_x)f(r_y) = 4(1 - r_x)(1 - r_y). \qquad (9)$$

We define the area of the range as $A = R_xR_y$, and need to find $F_A(a) = P(A < a)$, where a is an arbitrary limit. The region over which $f(r_x,r_y)$ is non-zero is a square of side 1. The line $r_xr_y = a$, for $0 \leqslant a \leqslant 1$, is a hyperbolic curve under which any point (r_x,r_y) will satisfy $r_xr_y \leqslant a$ (Fig. 5a). Therefore, for $0 \leqslant a \leqslant 1$, $F_A(a) = P(r_xr_y < a)$ will be the volume defined by the integral below the line $r_xr_y = a$ and the joint density function. Hence, the distribution function is given by:

$$F_A(a) = \int_a^1 \int_0^{a/r_x} 4(1 - r_x)(1 - r_y)\,\mathrm{d}r_y\,\mathrm{d}r_x + \int_0^a \int_0^1 4(1 - r_x)(1 - r_y)\,\mathrm{d}r_y\,\mathrm{d}r_x, \qquad (10)$$

where the first double integral is the volume below the curve $r_xr_y = a$ on the interval (a, 1), and the second double integral is the volume above the rectangle to the left of the curve. The solution is:

$$F_A(a) = 5a^2 - 4a - 2(2a + a^2)\ln a, \qquad (11)$$

so the density function is:

$$f_A(a) = \frac{\mathrm{d}}{\mathrm{d}a}F_A(a) = 8(a - 1) - 4(a + 1)\ln a, \qquad (12)$$

and the expected value and variance of the distribution are:

$$E(A) = \int_0^1 af_A(a)\,\mathrm{d}a = \int_0^1 a[8(a - 1) - 4(a + 1)\ln a]\,\mathrm{d}a = \frac{1}{9},$$

$$V(A) = \int_0^1 a^2f_A(a)\,\mathrm{d}a - \left(\frac{1}{9}\right)^2 = \frac{1}{36} - \frac{1}{81} = 0.0154. \qquad (13)$$

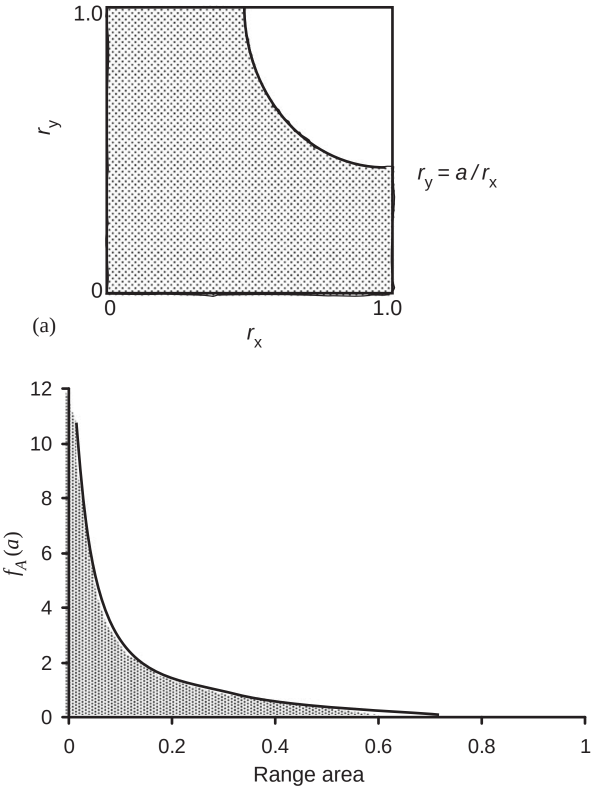

Fig. 5. Range size in a two-dimensional fully stochastic mid-domain model. On a (R_x, R_y) plot, where R_i is the range size in axes i, points for which range area $A = R_x$, R_y is equal to a constant a generate a hyperbolic line. The probability $P(A < a)$ is defined by the shaded area shown (a). Corresponding density function for range area (b).

3.2. Area of range at a given point

The species richness at any point $p_{xy} = (p_x, p_y)$, where p_x and p_y are coordinates on the two-dimensional square continent can be predicted under the fully stochastic mid-domain model by a simple extension of the binomial procedure described in Section 2.2. Since the probability that the range of a species intersects a point p_i on any of the two one-dimensional gradients is $P = 2(p_i - p_i^2)$, and assuming independence of the distributions on the two dimensions, then the probability that the range of a species intersects point p_i is $P = 4(p_x - p_x^2)(p_y - p_y^2)$ (Bokma et al., 2001). The equation predicts that 1/4 of species will occur at the center of the continent, and that species richness should decline towards the ends of the continent following a paraboloid curve.

The expected value and variance for the area of range of those species could in principle be deduced for any point p_{xy} from Eq. (6) and their equivalents when p_x or p_y are > 0.5. However, given the high number of possible combinations of p_i and r values that should have to be examined in each case, the procedure would be rather cumbersome and impractical. Instead, I performed numerical simulations in which the ranges of five million species were defined by generating four random numbers on the interval (0, 1) to determine the points Y_1, $Y_{2,}$, X_1, X_2. Then, for points located within the square-shaped continent, I calculated the corresponding species richness, mean area of range, and variance of range. The following generalizations derive from the simulation (Fig. 6).

Species richness follows the paraboloid equation reported above. In the middle of the continent, 25% of species intersected the point $p_x = p_y = 0.5$. In contrast, only 3.24% occurred in the point $p_x = p_y = 0.1$. Average range is constant regardless of the location of $p_{x,y}$. In all cases, a mean area of range of 0.250 was obtained from the simulations, meaning that on average, species occurring at any given point in the domain occur in 1/4 of the continent. Note that this average range area is equal to the squared value of the expected mean linear range size of the one-dimensional case (Eq. (7)). In contrast, variance in area of range was lowest at the middle of the continent (0.023) and highest near the corners of the continent; for example, variance was 0.039 in the point (0.1, 0.1). Variance values followed a symmetrical pattern around the middle of the continent; for example, points (0.2, 0.6), (0.6, 0.2), (0.2, 0.4), (0.4, 0.2), (0.4, 0.8), (0.6, 0.8), (0.8, 0.4), and (0.8, 0.6) had the same variance (0.0275 in this case). All frequency distributions were highly skewed towards large range sizes. However, skewness was highest close to the domain limits and lowest at the midpoint of the domain (Fig. 6).

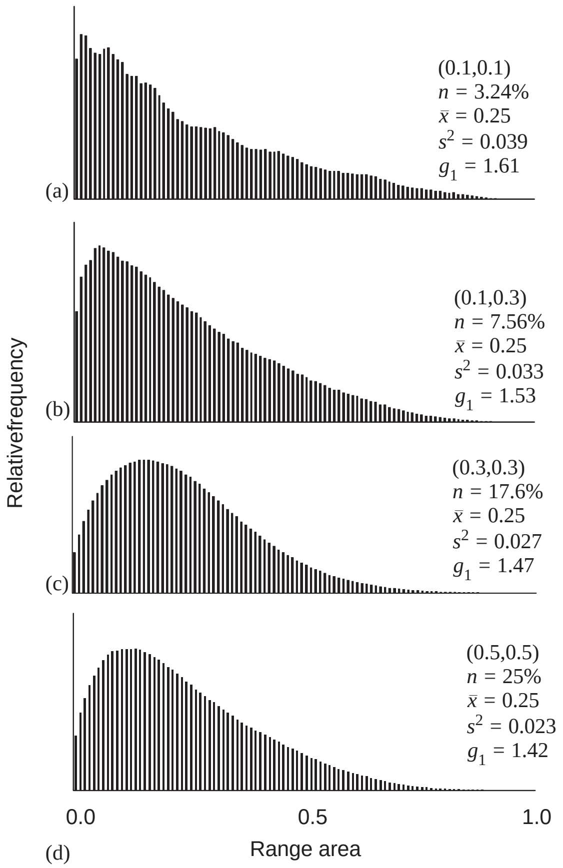

Fig. 6. Range-area frequency distributions at four points on a two-dimensional domain derived by randomly locating the range of five million species. Case where $(p_x, p_y) = (0.1, 0.1)$; 3.24% of species, mean range area $\bar{A} = 0.25$, variance $S_A^2 = 0.039$, skewness $g_1 = 1.61$ (a). Case where $(p_x, p_y) = (0.1, 0.3)$; 7.56% of species, mean range area $\bar{A} = 0.25$, variance $S_A^2 = 0.033$, skewness $g_1 = 1.53$ (b). $(p_x, p_y) = (0.3, 0.3)$; 17.6% of species, mean range area $\bar{A} = 0.25$, variance $S_A^2 = 0.027$, skewness $g_1 = 1.47$ (c). $(p_x, p_y) = (0.5, 0.5)$; 25% of species, mean range area $\bar{A} = 0.25$, variance $S_A^2 = 0.023$, skewness $g_1 = 1.42$ (d).

4. Discussion

The analytical models presented herein confirm previous observations based on numerical simulations and provide new predictions regarding the spatial distribution of species and its consequence on species richness. Models presented here prove that the fully stochastic procedure of arranging species on a bounded continent should generate a pattern in which species richness peaks at the middle of one-dimensional domains, declining towards the borders following a parabolic curve (Willig and Lyons, 1998; Colwell and Lees, 2000). A similar pattern is demonstrated here for the two-dimensional case. Also confirmed mathematically is the observation that under the assumptions of the fully stochastic model no gradient in average range size should be observable along a one-dimensional domain. My numerical simulations also show that the same pattern can be generalized to the two-dimensional case.

A pattern that had not been examined in previous studies is the spatial gradient in the variance of range sizes and the changes in the shape of the frequency distribution of range sizes. Because Rapoport's rule only predicts a gradient of average range size, very little attention has been given to variation around that average. Models presented here caution future studies contrasting empirical data with the predictions of mid-domain models to include an analysis of the shape of range-size histograms at different latitudes, and not only of the corresponding averages. Results also imply that the composition of species assemblages in terms of their rarity (or how small is their range) should show a latitudinal gradient under the assumptions of mid-domain models. In particular, near the extremes of the domain, the frequency distribution of range sizes becomes almost uniform, with the same percentage of restricted (rare) vs. widespread species. In contrast, in the middle of the domain, the model predicts that most species will have ranges of intermediate size, with relatively few widespread and restricted species.

The two-dimensional model predicts a curvilinear frequency distribution of range size that contrasts with the linear decreasing pattern for the unidimensional case. This curve is very similar to that reported in the literature for continental assemblages of vertebrates, which show a unimodal, highly right-skewed species-range size distribution. The pattern, which has been dubbed "a hollow curve", and described most frequently using a log-normal curve, is one of the most pervasive features of the distribution of animal species in continents (Anderson, 1977; Williamson and Gaston, 1999; Gaston, 2003). The model presented here provides a new description for species-range size distributions, and constitutes an adequate null model for empirical comparisons.

Models that are analogous to those presented here have been developed to simulate the abundances of species in ecological communities (the second "broken stick" model, MacArthur, 1957), the arrangement of the ranges of species on a continent (Pielou, 1977), the patterns of use of ecological resources arranged along a gradient (De Vita, 1979; Sugihara, 1986), and the flowering patterns of plant species along a temporal scale (Cole, 1981). However, all of these studies focused on patterns of overlap to estimate competition between species. Colwell and Hurtt (1994) were the first to analyse the number of overlaps on a particular point on the gradient, thus discovering the mid-domain effect of species diversity. Subsequently, mid-domain models have focused mostly on patterns of species diversity. Models presented here emphasize the spatial variation in range size, a feature poorly examined in previous studies.

Mid-domain models constitute a major advance in our understanding of null models of species richness. Before them, it was assumed that a random placement of several species on a continent would yield a uniform distribution of diversity. The first mid-domain models showed otherwise, demonstrating that particular patterns of species richness and average range size would appear even when the extent and position of the ranges of species were determined at random (Colwell and Lees, 2000). Therefore, adequate null models should take into account those "background" patterns. More recently, some proponents of mid-domain model have suggested that the models themselves explain much of the diversity patterns seen in nature (Jetz and Rahbeck, 2001). In contrast, other authors have dismissed mid-domain models, considering them unrealistic and flawed (Hawkins and Diniz-Filho, 2002; Zapata et al., 2003). As other models of ecological systems, mid-domain models are indeed "unrealistic". However, as other null models, they do not pretend to reproduce with detail the processes and patterns that can be seen in nature. Instead, by simplifying a very complex system, mid-domain models extract the essential components of natural patterns. In that sense, they constitute adequate and perfectly valid null models for ecological studies of species richness (Colwell et al., 2004).

Regardless of the position that one takes, models presented here show particular predictions that should be considered in future analyses of mid-domain models. In particular, the pattern in which the average range size remains invariant despite dramatic changes in the variance and skewness of the frequency distribution constitute an explicit benchmark against which empirical data could be compared to test the hypothesis that random arrangement of species explain much of the diversity patterns in nature. Skeptics and proponents alike should find predictions presented here useful for testing their own ideas to increase our understanding of geographic patterns of diversity.

Acknowledgements

I thank Andrés Christen, Gerardo Rodríguez, Pilar Rodríguez, Jorge Soberón, Ella Vázquez, and two anonymous reviewers for helpful ideas and suggestions, and Rocío Graniel for help with the literature search. I also thank the National Autonomous University of Mexico (UNAM) for providing the necessary spatial and temporal opportunities for developing the models presented here. This work was financially supported by DGAPA-UNAM.

References

Anderson, S., 1977. Geographic ranges of North American terrestrial mammals. Am. Mus. Novitates 2629, 1–15.

Arita, H.T., Rodríguez, P., 2002. Geographic range, turnover rate, and the scaling of species diversity. Ecography 25, 541–553.

Bokma, F., Bokma, J., Mönkkönen, M., 2001. Random processes and geographic species richness patterns: why so few species in the north? Ecography 24, 43–49.

Cole, B., 1981. Overlap regularity and flowering phenology. Am. Nat. 117, 993–997.

Colwell, R.K., Hurtt, G.C., 1994. Nonbiological gradients in species richness and a spurious Rapoport effect. Am. Nat. 144, 570–595.

Colwell, R.K., Lees, D.C., 2000. The mid-domain effect: geometric constraints on the geography of species richness. Trends Ecol. Evol. 15, 70–76.

Colwell, R.K., Rahbek, C., Gotelli, N.J., 2004. The mid-domain effect and species richness patterns: What have we learned so far? Am. Nat. 163, E1–E23.

De Vita, J., 1979. Niche separation and the broken-stick model. Am. Nat. 114, 171–178.

Gaston, K.J., 2003. The Structure and Dynamics of Geographic Ranges. Oxford University Press, Oxford.

Gaston, K.J., Blackburn, T.M., Spicer, J.I., 1998. Rapoport's rule: Time for an epitaph? Trends Ecol. Evol. 13, 70–74.

Grytnes, J.A., 2003. Ecological interpretations of the mid-domain effect. Ecol. Lett. 6, 883–888.

Hawkins, B.A., 2001. Ecology's oldest pattern? Trends Ecol. Evol. 16, 470.

Hawkins, B.A., Diniz-Filho, J.A.F., 2002. The mid-domain effect cannot explain the diversity gradient of Nearctic birds. Global. Ecol. Biogeog. 11, 419–426.

Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guegan, J.F., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M., Porter, E.E., Turner, J.R.G., 2003. Energy, water, and broad-scale geographic patterns of species richness. Ecology 84, 3105–3117.

Jetz, W., Rahbeck, C., 2001. Geometric constraints explain much of the species richness patterns in African birds. Proc. Natl. Acad. Sci. USA 98, 5661–5666.

Kerr, J.T., 1999. Weak links: 'Rapoport's rule' and large-scale species richness patterns. Global. Ecol. Biogeog. 8, 47–54.

Koleff, P., Gaston, K.J., 2001. Latitudinal gradients in diversity: real patterns and random models. Ecography 24, 341–351.

Laurie, H., Silander Jr., J.A., 2002. Geometric constraints and spatial patterns of species richness: critique of range-based null models. Diver. Distrib. 8, 351–364.

Lees, D.C., Kremen, C., Andriamampianina, L., 1999. A null model for species richness gradients: bounded range overlap of butterflies and other rainforest endemics in Madagascar. Biol. J. Linn. Soc. 67, 529–584.

Lyons, K.S., Willig, M.R., 1997. Latitudinal patterns of range size: methodological concerns and empirical evaluations for New World bats and marsupials. Oikos 79, 568–580.

Lyons, K.S., Willig, M.R., 2002. Species richness, latitude and scale-sensitivity. Ecology 8, 47–58.

MacArthur, R.H., 1957. On the relative abundance of bird species. Proc. Natl. Acad. Sci. US 43, 293–295.

McCain, C.M., 2003. North American desert rodents: a test of the mid-domain effect in species richness. J. Mamm. 84, 967–980.

Ney-Nifle, M., Mangel, M., 1999. Species-area curves based on geographic range and occupancy. J. Theor. Biol. 196, 327–342.

Pielou, E.C., 1977. The statistics of biogeographic range maps: sheaves of one-dimensional ranges. Bull. Int. Stat. Inst. 47, 11–122.

Pimm, S.L., Brown, J.H., 2004. Domains of diversity. Science 304, 831–833.

Rodríguez, P., Arita, H.T., 2004. Beta diversity and latitude in North American mammals: testing the hypothesis of covariation. Ecography 27, 547–556.

Rohde, K., Heap, M., Heap, D., 1993. Rapoport's rule does not apply to marine teleost fish and cannot explain latitudinal gradients in species richness. Am. Nat. 142, 1–16.

Rosenzweig, M.R., 1995. Species Diversity in Space and Time. Cambridge University Press, Cambridge.

Stevens, G.C., 1989. The latitudinal gradient in geographical range: How so many species coexist in the tropics. Am. Nat. 133, 240–256.

Sugihara, G., 1986. Shuffled sticks: on calculating nonrandom niche overlaps. Am. Nat. 127, 554–560.

Williamson, M., Gaston, K.J., 1999. A simple transformation for sets of range sizes. Ecography 22, 674–680.

Willig, M.R., Lyons, S.K., 1998. An analytical model of latitudinal gradients of species richness with an empirical test for marsupials and bats in the New World. Oikos 81, 93–98.

Willig, M.R., Kaufman, D.M., Stevens, R.D., 2003. Latitudinal gradients of biodiversity: Pattern, process, scales, and synthesis. Ann. Rev. Ecol. Evol. System 34, 273–309.

Zapata, F.A., Gaston, K.J., Chown, S.L., 2003. Mid-domain models of species richness gradients: assumptions, methods and evidence. J. Anim. Ecol. 72, 677–690.