Miroslav Holeček, Olga Červená

A CONTRIBUTION TO THE PROBLEM OF SPECIES-AREA RELATIONSHIP: ONE SPECIES ON AN INFINITE PLANE

Key words: Self-similarity, scale dependence, species-area relationship, probability.

1 Introduction

The relationship between the number of species and the area sampled is one of the oldest and best-documented patterns in community ecology. A species-area relationship (SAR) describes the dependence of the number of species found in a censused patch of habitat on the area of that patch. It is conventional to use power law of an SAR,

$$S(A) = cA^z, \tag{1}$$

where S is the number of species found in an area A, and c and z are constants [1]. The exponent z is close to 0.25 for several theoretical models and much field data [2], but there are also data where z is less than 0.25 (e.g. with smaller, plant-sizes quadrants [3]), and data where the slope is greater than 0.25 (e.g. on islands, and at large scales generally). In general, z may take on different values at very different spatial scales. The equation (1) is interesting and useful if z is constant across a large range of areas; an enormous amount of experimental evidence [2] suggests this power-low form is indeed often valid over a range of at least one or two orders of magnitude of area. If the fraction of species in area A that are also found in one-half of that area is independent of A, distribution of species is *self-similar*. It can be easily shown that the self-similarity gives the power law (1) with a constant z [4].

Any concrete distribution of species over an area is influenced by an enormous amount of effects: How many individuals belong to individual species, what is character of landscape and niches they live in, how individuals and species interact each other because of their entanglement into an extremely complicated net of mutual relations (such as food chains), how they move, migrate and so on. From this point of view, such an extremely simple law as (1) with a constant z looks like a miracle. A striking fact is an (at least approximate) *self-similarity*, which calls for an explanation based on some simple, fundamental principles. The problem, however, is that we have no some basic "ecological" laws (such as those in physics) and thus we do not know *what* ideas (physical laws?) should be important for such considerations.

In this paper we use only geometrical ideas to formulate an extremely simple model of a species distribution. A big advantage of this model is its formulation in an infinite plane which guarantees *scale invariance*and enables us to find results in a pure self-similar form. We neglect all complications and imagine individuals of any species as points randomly distributed over an infinite plane. In fact, the "infinite plane" means that the area where a species is studied is much larger than a characteristic region of a typical "cluster" formed by individuals of this species.

2 Randomly distributed points on a plane

The body of the problem of distribution of species lies in the following question [4]: *If we know that an individual of a species is in a rectangle what is a probability p that we can find an individual of the same species in a given half (e.g. right) of the rectangle.* If this probability does not depend on the size of rectangle, we obtain the power law (1) and the relation between p and z is [4]

$$p = 2^{-z}. \tag{2}$$

In fact, this dependence describes an area relationship for one species. The question whether it can model a species-relationship will not be solved here.

Let us imagine individuals of one species as points randomly distributed over an (infinite) plane in which a rectangle R with sides L and $\sqrt{2}L$ is located. The above problem can be formulated for this situation as well: If we know that there is a point in R what is a probability p that there is a point in its right half R_{R} (a rectangle $L \times L/\sqrt{2}$).

However, the infinity of plane makes a problem: The probability of a concrete distribution of points on infinite plane is ill-defined (there is no probability measure to define it). On the other hand, the infinity of plane guarantees a manifest *self-similarity.* Indeed, there is no length parameter connected with an infinite plane and thus the probability p (whatever its mathematical meaning may be) cannot depend on L. In other words, if we use a rectangle with sides L', $\sqrt{2}L'$ the probability of finding a point in its left half must be the same as when using the previous one.

Therefore we must formulate the previous problem much more carefully. Let us imagine a *concrete* cluster of N points in an infinite plane and a ("blind") observer having only one more information about this cluster (excepting the fact that it includes N points): namely, that *there is at least one point somewhere within the rectangle.* The observer may denote this (unknown) point a and think as follows: Surely there is a (unique) square with sides parallel to the sides of the rectangle having a at its center so that

1. all points of the cluster are within this square,
2. there is at least one point laying on the border of the square.

The observer denotes this (unknown) point b and the side of the square by an (unknown) positive real number l (see Fig. 1). The crucial idea of our consideration consists in the fact that the observer can determine the probability of finding a point in right half of the rectangle with respect to the parameter l. In other words, though he does not know this parameter he can calculate the probability $P = f(N, l, L)$ as a function of this parameter.

Let us calculate the probability P. Denoting by c an arbitrary point different from a and b (see Fig. 1) the probability P is defined as follows

$$P = P(a \in R_{\mathrm{R}}) + P[(a \notin R_{\mathrm{R}}) \wedge (b \in R_{\mathrm{R}})] + P[(a \notin R_{\mathrm{R}}) \wedge (b \notin R_{\mathrm{R}}) \wedge (c \in R_{\mathrm{R}})], \tag{3}$$

Now we introduce co-ordinate x, y of point a in the left half of rectangle R_{L}, furthermore we mark by $B(x, y)$ the intersection of our square with the right half of rectangle R_{R} and the intersection of the border of square with the right half of rectangle by

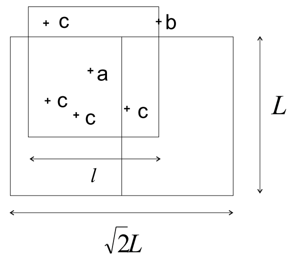

Figure 1: A square area with side l includes $N=6$ individuals, one individual a is in the center, the individual b is on the border and the others $N-2$ individuals c are inside the square.

$h(x,y)$ (see Fig. 2). The probability that at least one of N individuals (which are in square with the area l^2) falls into the patch B is $1-\left(\frac{l^2-B}{l^2}\right)^N$.

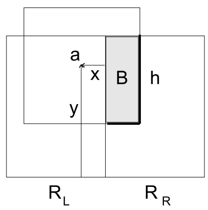

Figure 2: B - the intersection of square with area l^2 with the right half of rectangle R_R, h - the intersection of the border of the square with the right half of rectangle R_R, x, y co-ordinates of point a in the left half of rectangle R_L.

The equation (3) can be written in the form:

$$\begin{aligned} P &= \frac{1}{2} + \frac{1}{2}\int_0^L\int_0^{L/\sqrt{2}} \frac{h(x,y)}{4l}\frac{\sqrt{2}dxdy}{L^2} + \\ &+ \frac{1}{2}\left(1-\int_0^L\int_0^{L/\sqrt{2}} \frac{h(x,y)}{4l}\frac{\sqrt{2}dxdy}{L^2}\right)\int_0^L\int_0^{L/\sqrt{2}}\left(1-\left(1-\frac{B(x,y)}{l^2}\right)^{N-2}\right)\frac{\sqrt{2}dxdy}{L^2}, \end{aligned} \tag{4}$$

where $B(x,y)$ and $h(x,y)$ depend on L,l and a concrete position of the point a, i.e. on the coordinates x, y:

$(L \geq l)$

$$
\begin{array}{ll}
B(x,y) = (l/2-x)(y+l/2), & h(x,y) = l-x+y \\
& \quad if \quad x \in (0,l/2),\, y \in (0,l/2) \\
B(x,y) = l(l/2-x), & h(x,y) = 2(l-x) \\
& \quad if \quad x \in (0,l/2),\, y \in (l/2, L-l/2) \\
B(x,y) = (l/2-x)(L-y+l/2), & h(x,y) = l+L-x-y \\
& \quad if \quad x \in (0,l/2),\, y \in (L-l/2, L) \\
B(x,y) = 0, & h(x,y) = 0 \\
& \quad if \quad x \in (l/2, L/\sqrt{2}),\, y \in (0,L)
\end{array}
\tag{5}
$$

$(L < l < \sqrt{2}L)$

$$
\begin{array}{ll}
B(x,y) = (l/2-x)(y+l/2), & h(x,y) = l-x+y \\
& \quad if \quad x \in (0,l/2),\, y \in (0,L-l/2) \\
B(x,y) = L(l/2-x), & h(x,y) = L \\
& \quad if \quad x \in (0,l/2),\, y \in (L-l/2, l/2) \\
B(x,y) = (l/2-x)(L-y+l/2), & h(x,y) = l+L-x-y \\
& \quad if \quad x \in (0,l/2),\, y \in (l/2, L) \\
B(x,y) = 0, & h(x,y) = 0 \\
& \quad if \quad x \in (l/2, L/\sqrt{2}),\, y \in (0,L)
\end{array}
\tag{6}
$$

$(\sqrt{2}L < l < 2L)$

$$
\begin{array}{ll}
B(x,y) = (y+l/2)L/\sqrt{2}, & h(x,y) = L/\sqrt{2} \\
& \quad if \quad x \in (0, l/2-L/\sqrt{2}),\, y \in (0,L-l/2) \\
B(x,y) = L^2/\sqrt{2}, & h(x,y) = 0 \\
& \quad if \quad x \in (0, l/2-L/\sqrt{2}),\, y \in (L-l/2, l/2) \\
B(x,y) = (L-y+l/2)L/\sqrt{2}, & h(x,y) = L/\sqrt{2} \\
& \quad if \quad x \in (0, l/2-L/\sqrt{2}),\, y \in (l/2, L) \\
B(x,y) = (l/2-x)(y+l/2), & h(x,y) = l-x-y \\
& \quad if \quad x \in (l/2-L/\sqrt{2}, L/\sqrt{2}),\, y \in (0,L-l/2) \\
B(x,y) = L(l/2-x), & h(x,y) = L \\
& \quad if \quad x \in (l/2-L/\sqrt{2}, L/\sqrt{2}),\, y \in (L-l/2, l/2) \\
B(x,y) = (l/2-x)(L-y+l/2), & h(x,y) = l+L-x-y \\
& \quad if \quad x \in (l/2-L/\sqrt{2}, L/\sqrt{2}),\, y \in (l/2, L)
\end{array}
\tag{7}
$$

$(2L < l < 2\sqrt{2}L)$

$$
\begin{array}{ll}
B(x,y) = L^2/\sqrt{2}, & h(x,y) = 0 \\
& \quad if \quad x \in (0, l/2-L/\sqrt{2}),\, y \in (0,L) \\
B(x,y) = L(l/2-x), & h(x,y) = L \\
& \quad if \quad x \in (l/2-L/\sqrt{2}, L/\sqrt{2}),\, y \in (0,L)
\end{array}
\tag{8}
$$

$(4L < l)$

$$B(x,y) = L^2/\sqrt{2}, \quad h(x,y) = 0 \quad if \quad x \in (0, L/\sqrt{2}),\ y \in (0, L) \tag{9}$$

For simplification we use notation as follows:

$$r = \frac{L}{l} \quad \alpha = \frac{x}{l}, \quad \beta = \frac{y}{l}, \quad \sigma(\alpha,\beta) = \frac{B(x,y)}{l^2}, \quad \lambda(\alpha,\beta) = \frac{h(x,y)}{l} \tag{10}$$

Denoting

$$I_\lambda = \int_0^r \int_0^{r/\sqrt{2}} \lambda(\alpha,\beta) d\alpha d\beta, \quad I_\sigma = \int_0^r \int_0^{r/\sqrt{2}} \left(1 - (1-\sigma(\alpha,\beta))^{N-2}\right) d\alpha d\beta \tag{11}$$

we can write (4) in the form:

$$P = \frac{1}{2} + \frac{1}{4\sqrt{2}r^2} I_\lambda + \frac{1}{\sqrt{2}r^2}\left(1 - \frac{1}{2\sqrt{2}r^2} I_\lambda\right) I_\sigma. \tag{12}$$

After a straightforward integration, we obtain:
if $r \geq 1$

$$I_\lambda = \frac{3r-1}{4}, \tag{13}$$

$$I_\sigma = \frac{1}{2} - \frac{2}{(N-1)} \sum_{i=1}^{N-1} \frac{3^i - 2^i}{i\ 4^i} + \frac{(r-1)\left(N - 3 + 2^{2-N}\right)}{2\ (N-1)}, \tag{14}$$

if $1/\sqrt{2} < r < 1$

$$I_\lambda = \frac{3r-1}{4}, \tag{15}$$

$$I_\sigma = \frac{r}{2} - \frac{2}{(N-1)} \sum_{i=1}^{N-1} \frac{3^i - 2^i(2-r)^i}{i\ 4^i} + \frac{(1-r)\left(r - 2 + 2\left(1-\frac{r}{2}\right)^N\right)}{r\ (r-2)\ (N-1)}, \tag{16}$$

if $1/2 < r < 1/\sqrt{2}$

$$I_\lambda = -r^3 + \frac{r^2}{2}\left(4\sqrt{2}+1\right) - \frac{3r}{4}\left(\sqrt{2}+1\right) + \frac{1}{4}, \tag{17}$$

$$\begin{aligned} I_\sigma = {} & \frac{r^2}{\sqrt{2}} + \frac{\sqrt{2}(1-\sqrt{2}r)}{r(N-1)}\left[\left(1 - \frac{r^2}{\sqrt{2}}\right)^{N-1} - \left(1 - \frac{r}{2\sqrt{2}}\right)^{N-1}\right] \\ & - \frac{1 - r(\sqrt{2}+1) + \sqrt{2}r^2}{2}\left(1 - \frac{r^2}{\sqrt{2}}\right)^{N-2} - \frac{2}{N-1}\sum_{k=1}^{n-1}\binom{N-1}{k} \\ & \frac{1}{4^k\ k}\left[\left(2r\left(\sqrt{2}r-1\right)\right)^k - \left(\sqrt{2}r-1\right)^k + \left(-\sqrt{2}r\right)^k - \left(-2\sqrt{2}r^2\right)^k\right] \\ & + \frac{r-1}{r\ (N-1)}\left[\left(1 - \frac{r}{2} + \frac{r^2}{\sqrt{2}}\right)^{N-1} - \left(1 - \frac{r^2}{\sqrt{2}}\right)^{N-1}\right] \end{aligned} \tag{18}$$

if $1/2\sqrt{2} < r < 1/2$

$$I_\lambda = r^2\left(\sqrt{2}r - \frac{1}{2}\right), \tag{19}$$

$$I\sigma = r^2\left(\sqrt{2} - \frac{1}{\sqrt{2}}\right) - r\left(\frac{1}{2} - \frac{r}{\sqrt{2}}\right)\left(1 - \frac{r^2}{\sqrt{2}}\right)^{N-2} \tag{20}$$

$$-\frac{1}{N-1}\left[\left(1 - \frac{r}{2} + \frac{r^2}{\sqrt{2}}\right)^{N-1} - \left(1 - \frac{r^2}{\sqrt{2}}\right)^{N-1}\right] \tag{21}$$

and if $r < 1/2\sqrt{2}$

$$I_\lambda = 0, \tag{22}$$

$$I_\sigma = \frac{r^2}{\sqrt{2}}\left(1 - \left(1 - \frac{r^2}{\sqrt{2}}\right)^{N-2}\right). \tag{23}$$

Substitution integrals I_λ, I_σ into (12), we get following values for separate r:

$$P_{(r\geq 1)} = \frac{1}{2} + \frac{3r-1}{2^4\sqrt{2}r^2} + \frac{1}{\sqrt{2}r^2}\left(1 - \frac{3r-1}{2^3\sqrt{2}r^2}\right)s(N,r), \tag{24}$$

where

$$s(N,r) = \frac{1}{2} - \frac{2}{(N-1)}\sum_{i=1}^{N-1}\frac{3^i - 2^i}{i\ 4^i} + \frac{(r-1)\left(N-3+2^{2-N}\right)}{2(N-1)}, \tag{25}$$

$$P_{(1/\sqrt{2}<r<1)} = \frac{1}{2} + \frac{3r-1}{2^4\sqrt{2}r^2} + \frac{1}{\sqrt{2}r^2}\left(1 - \frac{3r-1}{2^3\sqrt{2}r^2}\right)t(N,r), \tag{26}$$

where

$$t(N,r) = \frac{r}{2} - \frac{2}{(N-1)}\sum_{i=1}^{N-1}\frac{3^i - 2^i(2-r)^i}{i\ 4^i} - \frac{(1-r)\left(r-2+2\left(1-\frac{r}{2}\right)^N\right)}{r\ (r-2)\ (N-1)}, \tag{27}$$

$$P_{(1/2<r<1/\sqrt{2})} = \frac{1}{2} + \frac{u(N,r)}{16\ r^2} + \frac{1}{\sqrt{2}r^2}\left(1 - \frac{u(N,r)}{8\ r^2}\right)v(N,r), \tag{28}$$

where

$$u(N,r) = -2\sqrt{2}r^3 + r^2\left(8+\sqrt{2}\right) - \frac{1}{\sqrt{2}}\left(3r\left(\sqrt{2}+1\right) - 1\right) \tag{29}$$

and

$$v(N,r) = \frac{r^2}{\sqrt{2}} + \frac{\sqrt{2}(1-\sqrt{2}r)}{r(N-1)}\left[\left(1 - \frac{r^2}{\sqrt{2}}\right)^{N-1} - \left(1 - \frac{r}{2\sqrt{2}}\right)^{N-1}\right] \tag{30}$$

$$-\frac{1 - r(\sqrt{2}+1) + \sqrt{2}r^2}{2}\left(1 - \frac{r^2}{\sqrt{2}}\right)^{N-2} - \frac{2}{N-1}\sum_{k=1}^{n-1}\binom{N-1}{k}$$

$$\frac{1}{4^k\ k}\left[\left(2r\left(\sqrt{2}r-1\right)\right)^k-\left(\sqrt{2}r-1\right)^k+\left(-\sqrt{2}r\right)^k-\left(-2\sqrt{2}r^2\right)^k\right]$$
$$+\frac{r-1}{r\ (N-1)}\left[\left(1-\frac{r}{2}+\frac{r^2}{\sqrt{2}}\right)^{N-1}-\left(1-\frac{r^2}{\sqrt{2}}\right)^{N-1}\right]$$

$$P_{(1/2\sqrt{2}<r<1/2)}=\frac{1}{2}+\frac{\sqrt{2}r-\frac{1}{2}}{4\sqrt{2}}+\frac{1}{\sqrt{2}r^2}\left(1-\frac{\sqrt{2}r-\frac{1}{2}}{2\sqrt{2}}\right)w(N,r), \tag{31}$$

where

$$w(N,r) \quad = \quad r^2\left(\sqrt{2}-\frac{1}{\sqrt{2}}\right)-r\left(\frac{1}{2}-\frac{r}{\sqrt{2}}\right)\left(1-\frac{r^2}{\sqrt{2}}\right)^{N-2} \tag{32}$$
$$-\frac{1}{N-1}\left[\left(1-\frac{r}{2}+\frac{r^2}{\sqrt{2}}\right)^{N-1}-\left(1-\frac{r^2}{\sqrt{2}}\right)^{N-1}\right] \tag{33}$$

$$P_{(0<r<1/2\sqrt{2})}=1-\frac{1}{2}\left(1-\frac{r^2}{\sqrt{2}}\right)^{N-2} \tag{34}$$

The result is pictured by a three-dimensional graph at Fig. 3 where $N \in [3, 100]$. Let us

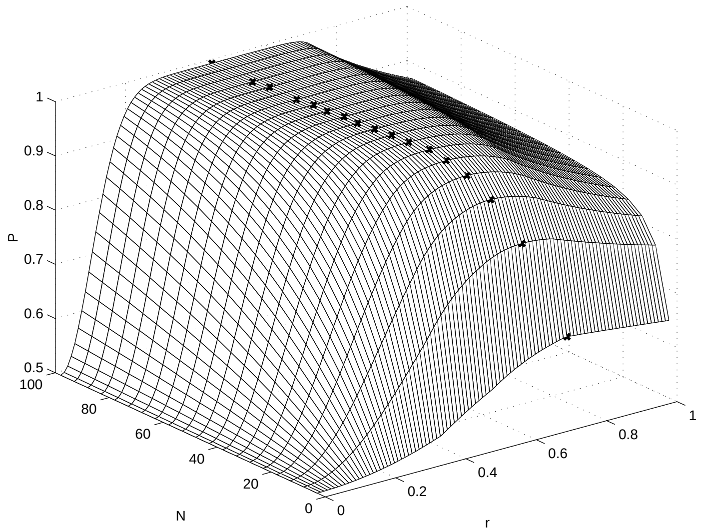

Figure 3: Probability of occurrence of at least one of N individuals in the right half of rectangle area in dependence on number of individuals N and on the rate $r = L/l$. The asterisk denotes its maximal value for a concrete N.

notice that, for each N, the probability reaches a maximal value at a concrete $r = L/l$, see Figs. 3, 4. For our "blind" observer knowing only N this maximal probability is a significant number characterizing in a way the situation (without any additional

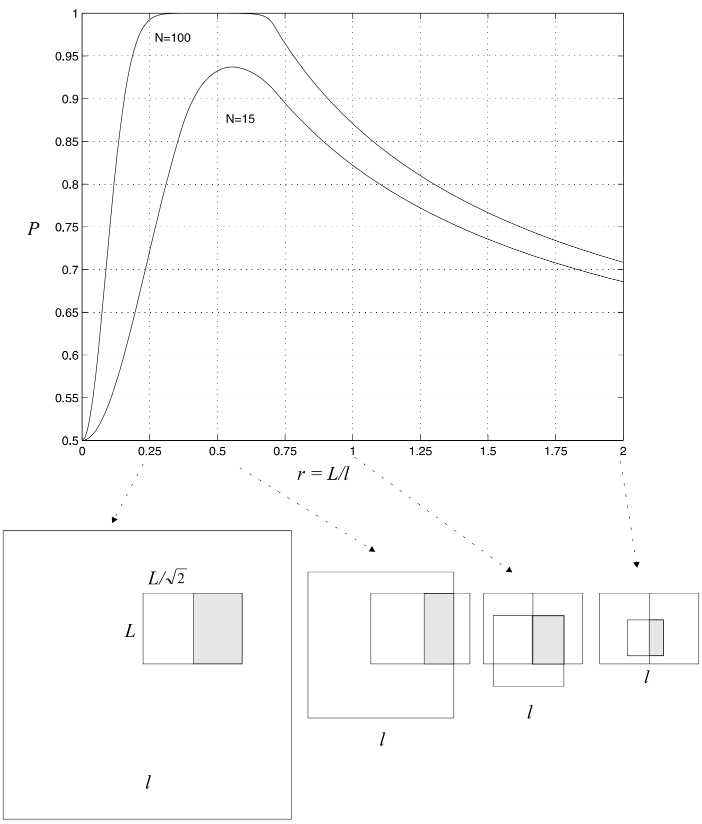

Figure 4: Probability of occurrence of at least one of $N = 15; 100$ individuals in right half of the rectangle.

knowledge). Moreover, this number does not depend on L and thus it describes the situation in a clear *self-similar* way. In the diagram at Fig. 5 its dependence on N is shown. Using the formula (2) we get also the dependence of the parameter z on N (see the second diagram at Fig. 5). A meaning of this dependence with respect to a species-area relationship is not clear. Having many species distributed over the plane each of them have different number of individuals and the maximal probability for each of them is different.

3 Conclusion

In this paper, we formulate a problem of finding a point at right half of a rectangle located on an infinite plane when we know that at least one point from a group of N randomly distributed points on the plane lies within the rectangle. Though the problem is ill-defined we find a consistent formulation using the size l of a minimal square including all N points. This length is eliminated at the end by studying only

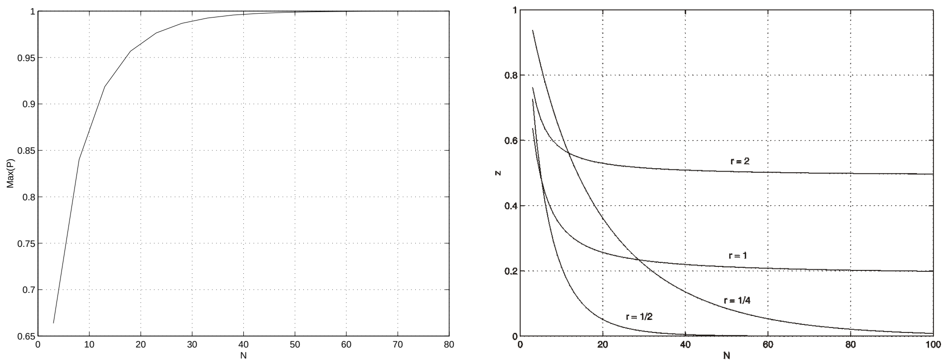

Figure 5: Maximal probability of occurrence of at least one of N individuals in right half and z in dependence on N for $r = 0.25;\ 0.5;\ 1;\ 2$.

the maximal probability for each N.

This problem has a close relation to the typical problem in community ecology, namely the relationship between the number of species and the area sampled. The crucial question concerning this pattern is its evident *self-similarity* being valid at a broad extent of scales. Therefore we have found the results concerning randomly distributed points in a manifest self-similar way. Though we have not derived the correct exponent of the power law (1) we outlined a way which might lead to such a result. It needs, however, a clear interpretation of *what* the maximal probability mean and *how* to manipulate correctly with number of individuals of various species.

Acknowledgement

The authors are indebt to Dr. David Storch from the Center of Theoretical Study in Prague for explaining them the problem of species-area relationship in ecology. The work has been supported by the Czech Ministry of Education, Project No. MSM230000009.

Resumé: V práci je navržen způsob, kterým je možno přistoupit k objasňování škálového vztahu mezi počtem druhů a velikostí plochy, na níž se vyskytují. Metoda je založena na náhodném rozložení bodů na ploše.

Reference

[1] Harte J.: Scaling and Self-Similarity in Species Distributions: Implications for Extinction, Species Richness, abundance, and Range *Scaling in Biology*: 325-342, *Oxford Univ. Press*, Oxford 2000

[2] Rosenzweig M.L.: Species diversity in space and time, *Cambridge Univ. Press*, Cambridge 1995

[3] Crawley M.J., Harral J.E.: Scale Dependence in Plant Biodiversity, *Science*, **291**: 864-868, 2001

[4] Harte J., Kinzig A., Green J.: Self-Similarity in the Distribution and Abundance of Species, *Science*, **284**: 334-336, 1999

Dr. RNDr. Miroslav Holeček, RNDr. Olga Červená, Institute of Interdisciplinary Studies, University of West Bohemia, Husova 11, 306 14 Pilsen, Czech Republic Tel.: +420/19/723 50 85; fax: +420/19/723 64 43; e-mail:holecek@ums.zcu.cz