Journal of Theoretical Biology 265 (2010) 517–523

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

A spatially explicit model for tropical tree diversity patterns

Sz. Horvát [a,b], A. Derzsi [a,c], Z. Néda [a,d,*], A. Balog [e,f]

[a] Babeş-Bolyai University, Department of Physics, Kogălniceanu street 1, RO-400084 Cluj-Napoca, Romania
[b] University of Bergen, Department of Physics and Technology, Allégaten 55, N-5007 Bergen, Norway
[c] Research Institute for Solid State Physics and Optics, Konkoly-Thege street 29-33, H-1121 Budapest, Hungary
[d] Hungarian University Federation of Cluj, Tipografiei street 12/I, RO-400101 Cluj-Napoca, Romania
[e] Institute of Ecology, Friedrich-Schiller-University, Dornburger street 159, D-07749 Jena, Germany
[f] Yale University, Greeley Memorial Laboratory, 370 Prospect Street New Haven, CT, USA

ARTICLE INFO

Article history:
Received 15 January 2010
Received in revised form
14 May 2010
Accepted 26 May 2010
Available online 2 June 2010

Keywords:
Neutral model
Spatial distribution of species
Population dynamics
Lattice models
Monte Carlo simulations

ABSTRACT

A simple two-parameter model resembling the classical voter model is introduced to describe macroecological properties of tropical tree communities. The parameters of the model characterize the speciation- and global-dispersion rates. Monte Carlo type computer simulations are performed on the model, investigating species abundances and the spatial distribution of individuals and species. Simulation results are critically compared with the experimental data obtained from a tree census on a 50 hectare area of the Barro Colorado Island (BCI), Panama. Fitting to only two observable quantities from the BCI data (total species number and the slope of the log–log species-area curve at the maximal area), it is possible to reproduce the full species-area curve, the relative species abundance distribution, and a more realistic spatial distribution of species.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Macroecology (Brown, 1995; Pielou, 1977; Bell, 2001) studies the relationships between organisms and their environment at large spatial scales in order to characterize and explain universal statistical patterns of abundance, distribution, and diversity (May, 1975; Harte, 2000; Harte et al., 1999). A variety of models and methods from statistical physics are appropriate for the study of ecosystem dynamics, in which the main "entities" are either many individual organisms within populations, or many species within local, regional or continental communities (Dewdney, 2000; Chave, 2004; Enquist et al., 2002; Norris, 2003; Chave et al., 2002).

The remarkable regularities in patterns of data on how species originate, persist, assemble in groups, and eventually go extinct, suggest the existence of general mechanisms from which the biodiversity and the structure of ecological communities originate (Brown, 1995; Hubbell, 2001; Tokeshi, 1999). The difficulties in explaining the biological diversity of such systems originate from the very different spatial and temporal scales, starting with the evolution and biogeographic distribution of species, and ending with individual births and deaths in local communities. Present approaches for describing the dynamics of population genetics and ecology consider either the importance of the genetic fitness or the influence of random events governing birth, death and migration phenomena. There are thus two main, but conflicting perspectives on the nature of ecological communities: the niche-assembly and the dispersal-assembly perspective (Brown, 1995; Hubbell, 2001). Here we intend to use the latter approach, constructing a simple spatially explicit stochastic model for describing the relative species abundance and the spatial diversity patterns of trees in a tropical forest (Condit et al., 2002; Volkov et al., 2004a, 2007). The present work is motivated by a recent study (Zillio et al., 2005), where a one-parameter, spatially explicit model is applied for describing tropical tree diversity patterns.

2. Neutral models

Recently it has been argued in the scientific community that systems of ecologically similar species whose interaction with each other is primarily competitive can be successfully approached by using neutral models. One of the first and most widely used neutral models is that of Hubbell (2001) and Volkov et al. (2003), who assumed in his theory that within groups of ecologically similar species, individuals fill the landscape up to a point of saturation and the dynamics of all these species is governed by the same birth-, death-, migration- and mutation-rates.

* Corresponding author at: Babeş-Bolyai University, Department of Physics, Kogălniceanu street 1, RO-400084 Cluj-Napoca, Romania.
E-mail address: zneda@phys.ubbcluj.ro (Z. Néda).

0022-5193/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2010.05.032

This simple neutrality principle allows us to study a wide variety of systems by means of stochastic computer simulations, and to use the methods of statistical mechanics to get analytical results (Alonso et al., 2006; Azaele et al., 2006; Etienne and Alonso, 2007; Pigolotti et al., 2004; Volkov et al., 2004b).

Tropical forests (Condit et al., 2002; Volkov et al., 2004a, 2007; Condit, 1998) contain many more tree species than temperate or boreal ones (two orders of magnitude more), and it is believed that such systems are much closer to the neutrality assumption. Since the number of mature individual trees in equal-area samples of each forest type is almost the same, the increase in the abundance of a particular species needs to be compensated by a decrease in another species's population.

In order to develop an appropriate model, one needs to take into account the whole set of factors influencing community composition: births, deaths, immigration, and (on longer time scales) speciation. These processes interact in a complex manner and produce the commonly observed empirical patterns of diversity and distribution of abundances. The model should prescribe, for each generation, a method for choosing the species identity of the individuals that die and those that will occupy (by birth or immigration) the vacancies created. A simple method which describes fairly well the above-mentioned processes is the one developed by Hubbell and it is called "zero-sum ecological drift" (Hubbell, 2001; Volkov et al., 2003). This method implies that the replacing species are drawn at random from the existing community of species. The "ecological drift" is connected to the random replacement process and assumes that all individuals, regardless of species, have equal probabilities of giving birth, dying, immigrating or acquiring a mutation to generate a speciation event. The ecological drift does not imply an equal chance for each species to fill a given vacancy. Obviously, the greater probability of being drawn into a vacancy belongs to the more abundant species. The basic assumption is that individuals are equal, but species, as collective entities, are not. In Hubbell's model (Hubbell, 2001; Volkov et al., 2003), the ecological drift, without any additional mechanisms, is enough to produce the patterns of species abundance and diversity observed in nature. In particular, the community of tropical trees is very well described by these neutral models and neutrality will be also the basis of the approach considered in this work. Contrary to earlier modelling efforts (Hubbell, 2001; Volkov et al., 2003, 2005), here we continue the novel idea introduced by Zillio et al. (2005), and a spatially explicit model resembling the classical voter model will be studied. In our approach, however, a slightly more complex, two-parameter version is considered which is not suitable for an analytical study. We argue why it is important to extend model considered in Zillio et al. (2005), and we determine the best parameters for optimally reproducing the measured macro-ecological patterns.

3. A simple spatially explicit model

Most of the previous models used for understanding relevant ecological measures are mean-field like approaches in the sense that the spatiality of the individuals is lost, and the only relevant quantity that characterizes the system is the number of individuals belonging to each species (Dewdney, 2000; Chave, 2004; Enquist et al., 2002; Norris, 2003; Chave et al., 2002; Condit et al., 2002; Volkov et al., 2003, 2004a, 2004b, 2005, 2007; Zillio et al., 2005; Alonso et al., 2006; Azaele et al., 2006; Etienne and Alonso, 2007; Pigolotti et al., 2004; Condit, 1998).

Recently Zillio et al. (2005) took a step beyond the mean-field approach by considering a simple, but still realistic spatially explicit model. Their model consists of individuals placed on a lattice, and apart from speciation, only first neighbour interactions are taken into account. The community of tropical trees, where the coordinates of individuals are fixed, is especially suited for such an approach. This one-parameter model is successful in reproducing realistic relative species abundance (RSA) curves, but as it will be seen in Section 6 (Fig. 6a), it produces unrealistically compact homogeneous islands of tree species.

The model considered in the present study is discrete both in time and space and it is also inspired by the classical voter model (Holley and Liggett, 1975). Individuals are placed on a predefined lattice. For the sake of simplicity, a simple square lattice was considered. Each individual belongs to a given species, and the state variable characterizing each lattice site encodes these species. It is assumed that the lattice with size $L \times L$ is always completely filled up, corresponding to a constant ecological saturation of the territory. Thus the total number of individuals in the system is always $N = L \times L$. The number of possible states (species) is W, ($W \gg 1$). The dynamics of the system is governed by two adjustable parameters: p and q ($0 < p, q < 1$).

In the beginning a randomly chosen Potts state (species) is assigned to each lattice site (individual). Starting from this random initial condition, at each discrete time step a lattice site is randomly chosen and its state is changed. This process models the disappearance (death) of a tree, and the immediate occupation of the freed-up space by a newly born sapling. The p and q parameters govern the selection of the species identity of the new individual that occupies this free position as follows:

- With probability $1-q-p$, the state of the chosen site is changed to one of its eight Moore-neighbours. This is the most probable outcome, and models the quite common case of the new individual being an offspring of a neighbouring tree.
- With probability q, a randomly selected species is assigned to the chosen site. This species is selected with a uniform probability from the ensemble of possible species. If W is big enough and q much smaller than one, this process is suitable for modelling speciation or immigration into the considered territory.
- With probability p, the species identity of the new tree is chosen to be the same as the species of a randomly selected individual from among the remaining $N-1$ ones. This rule models the diffusion of the seeds in the considered area, allowing seeds originating from faraway individuals to reach the location.

It will be seen later that in order to get realistic results, the parameters p and q need to be selected as $0 < q \ll p < 1$.

By settings $p=0$ and choosing $W \gg N$, the analytically tractable model described in Zillio et al. (2005) is regained. However, when $p > 0$, the analytical treatment of the model becomes difficult due to the combined short- and long-range interactions. The model was studied thus by Monte Carlo (MC) simulations. One MC step is defined as N updates of lattice sites. In order to minimize the effect of boundaries, periodic boundary conditions were used.

4. Experimental data: The Barro Colorado Island dataset

The experimental results used for testing the model are from a detailed 50 hectare tropical forest tree census in Barro Colorado Island (BCI), realized by the Smithsonian Tropical Research Institute, Center for Tropical Forest Science (CTFS) (Condit, 1998). BCI is located in the Atlantic watershed of the Gatun Lake (Panama) and was declared a biological reserve in 1923. It has been administrated by the Smithsonian Tropical Research Institute since 1946. From the viewpoint of ecological studies, this island is ideal because it is covered with a rainforest that is still

undisturbed by humans. The flora and fauna of BCI have been studied extensively and inventories have reported 1369 plant species, 93 mammal species (including bats), 366 avian species (including migratory), and 90 species of amphibians and reptiles. The tropical tree census was performed only on a small part of the island, precisely on a 1000×500 m (50 hectare) area. The first census was completed in 1982, revealing a total of approximately 240,000 stems of 303 species of trees and shrubs. The importance of this CTFS programme consists in the fact that in each census all free-standing woody stems at least 1 cm diameter at breast height are identified, tagged, and mapped, and hence accurate statistics can be made. Data are publicly available for the years 1982, 1985, 1990 and 1995 (Hubbell et al., 2005).

The BCI database contains more than 320,000 individuals in total, belonging to 316 different tree species (Condit, 1998). The density of individuals for the 1995 census is illustrated in Fig. 1 with a greyscale code (darker shades correspond to higher densities). In the western part of the studied area lies a relatively large swampland (lighter region in the left side of Fig. 2, i.e. less trees). This area is ecologically quite different from the other parts (see for example the spatial distribution of the most abundant species, *Hybanthus prunifolius*, in Fig. 2). Therefore, in determining the relevant macroecological measures, only trees located in the eastern half of the region (500×500 m, 25 ha) were considered.

5. Relevant macroecological measures. Results for the Barro Colorado Island dataset

In order to describe the statistics of species sizes in ecological communities, the main measure which is usually considered is the relative species abundance (RSA) distribution. Provided that our data are spatially accurate, and we know the location of each individual, we can also investigate the species-area scaling and the spatial auto-correlation function for the individuals of a given species. These measures are briefly discussed here.

5.1. Relative species abundance distribution (RSA)

In ecology, the *abundance* of a species is understood as the relative number of individuals belonging to that species in an ecosystem. The *relative species abundance distribution* is introduced for characterizing the frequency of species with a given abundance (Volkov et al., 2007; Preston, 1962). Three different types of plots are generally used for representing species-abundance distributions. Historically the first, and thus also the most widespread, representation of species abundances is due to Preston (1962), who sorted the species of a sample into abundance intervals of consecutively doubling lengths ([1,2), [2,4), [4,8), etc.), and plotted the number of species found within these "octaves". Preston's method of plotting is motivated by the fact that abundances can vary over several orders of magnitude and there are far fewer abundant species than rare ones. Using fixed-length abundance intervals would result in large statistical fluctuations at the tail of the curve. In Fig. 3a the shape of the RSA Preston plot for the Barro Colorado tropical tree census for year 1995 is illustrated.

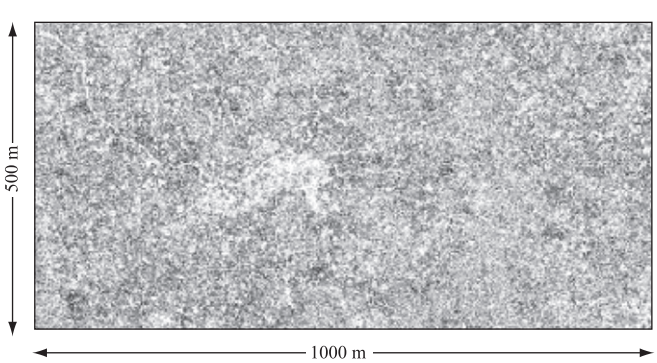

Fig. 1. Density of trees in the BCI census (census year: 1995).

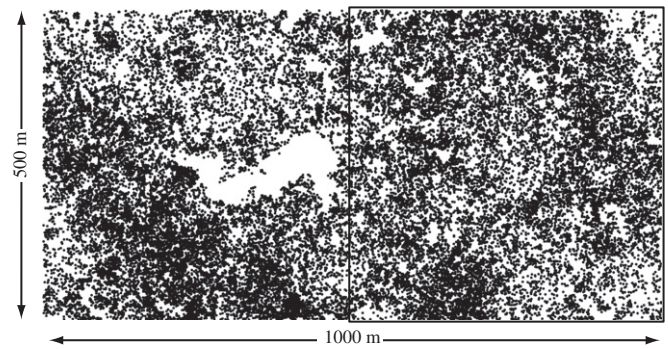

Fig. 2. Spatial distribution of the most abundant species in the BCI census, *Hybanthus prunifolius* (census year: 1995). The square on the right shows the region in which the relevant macroecological measures were computed.

The second and mathematically most rigorous way of representing the species-abundance distribution is plotting the associated probability density function $\rho(s)$, i.e. the probability for finding a species with a given s abundance. For most of the neutral-like communities this distribution function has a tilted J shape on a log–log scale. Fig. 3b shows this probability density for the Barro Colorado tropical tree census (year 1995). It is worth mentioning that $\rho(s)$ can be derived from the Preston plot by dividing the number of species in each interval with the length of the interval, and plotting on log–log scale this quantity versus the mean abundance in the given interval.

A third way of representing the species abundance distribution is arranging the species in decreasing order by their abundances and plotting the species's rank versus its abundance (rank–abundance plot) on a log–linear scale (Norris, 2003; Chave et al., 2002). This representation is inspired by several abundance studies in sociology and economics, leading to the very general Pareto–Zipf distribution. For the Barro Colorado tropical tree census the rank–abundance curve for the same, 1995 census year, is plotted in Fig. 3c. It is also easy to realize that the rank–abundance plot is related to the cumulative distribution function corresponding to the $\rho(s)$.

In Figs. 3a–c, for reference, the data are compared with the prediction of the Fisher log series distribution described in Fisher et al. (1943), with which the probability density plot and the rank–abundance plot show good agreement. Data for other census years give similar results.

5.2. Species–area relationship (SAR)

Generally, the number of detected species does not scale linearly by increasing the size of the sampled territory. In order to characterize this dependence, the species–area relationship is studied. This is done by considering larger and larger territories, and counting the number of species present within these areas (Harte, 2000; Harte et al., 1999; Zillio et al., 2005). For a better statistics, an average species number is calculated on several territories with similar areas.

The dependence of the average species number as a function of the sampling area defines the SAR curve. In Fig. 4 the

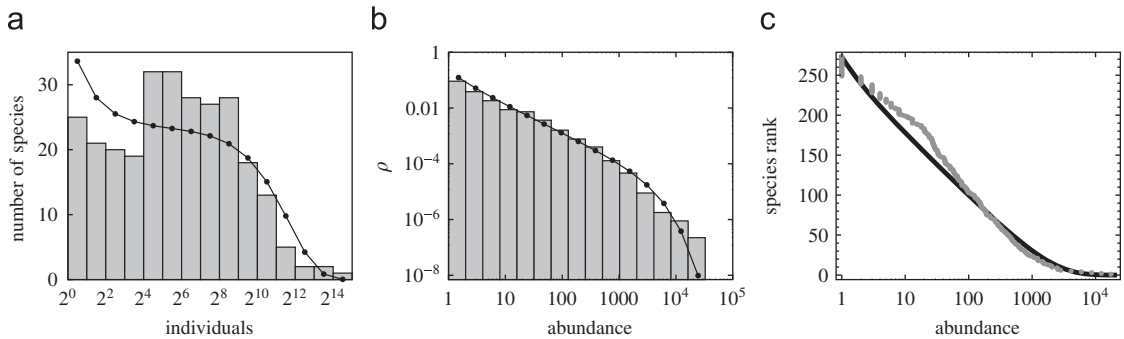

Fig. 3. RSA for Barro Colorado Island tropical tree census (census year: 1995). Three different ways of presenting the RSA: (a) the Preston plot, (b) the probability density function, and (c) the rank–abundance plot. Measurement data is in grey. The predictions of the Fisher log series for $\alpha = 33.64$ are shown in black.

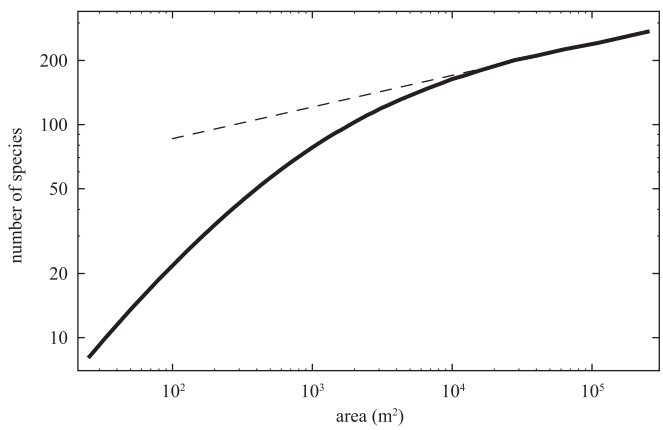

Fig. 4. Species-area curve for the BCI data (census year: 1995). In the limit of large areas, the SAR is compared to a power-law trend with exponent 0.15 (dashed line). Both scales are logarithmic.

characteristic shape of this curve is shown for the 1995 census year BCI data.

For the BCI dataset the species–area scaling is not obvious at all (Plotkin et al., 2000). We believe that the reason for this is that the data were collected from relatively small area. As it is visible in Fig. 4, a trend resembling a power-law appears only in the limit of larger areas.

5.3. The spatial auto-correlation function

If the individuals are restricted to a given spatial position, like trees for example, one can characterize the spatial distribution of the individuals from a given species using the spatial auto-correlation function. The way we construct the $C(r)$ auto-correlation function is the following: first a check-board type uniform mesh is considered, and the territory is divided into small, square-like domains (labelled by coordinates i,j) and the number of individuals from the considered species is determined in each domain ($N_{i,j}$). The auto-correlation function for a relative coordinate p,q ($C_{p,q}$) is calculated as

$$C_{p,q} = \langle (N_{i,j} - \langle N \rangle)(N_{i+p,j+q} - \langle N \rangle) \rangle_{i,j} \tag{1}$$

Here $\langle N \rangle$ denotes the average number of individuals of the considered species in the constructed domains: $\langle N \rangle = \langle N_{i,j} \rangle_{i,j}$. Since there is no reason to assume that the distribution of the individuals is non-isotropic, we can calculate the average of the $C_{p,q}$ values for all p,q values that are inside a ring with radius r and width Δr ($r \le \sqrt{p^2+q^2} \le r+\Delta r$), considering a reasonably small Δr value:

$$C(r) = \langle C_{p,q} \{ r \le \sqrt{p^2+q^2} \le r+\Delta r \} \rangle_{p,q} \tag{2}$$

The $C(r)$ auto-correlation function, characterizing the spatial distribution of the 10 most abundant species in the 1995 BCI dataset was calculated. For all of them, a power-law type decrease has been observed. As an example, in Fig. 5 the spatial auto-correlation function of the three most abundant species is plotted on a log–log scale. The linear trend suggests a power-law like decrease for $C(r)$.

6. MC simulation results

Given enough simulation time, the model converges to a statistically stable stationary state. For $0 < q \ll 1$, several species will co-exist in the final stationary state. Simulations performed on several systems proved that in order for the final, statistically stable state to be realistic, both $q \neq 0$ and $p \neq 0$ are necessary. If $q=0$, eventually one species will replace all others. Setting $p=0$ (which reduces the model to the one studied by Zillio et al., 2005) results in unrealistically compact islands of species (Fig. 6a), in obvious contradiction with the fractal-like intermixing observed in the BCI dataset (Fig. 6c). By choosing $p > 0$, it is possible to get a more realistic spatial distribution for the abundant species (Fig. 6b).

One important aim of the present study was to find those model parameters that reproduce the experimental data best. In order to compare the simulation results with the measurements, a unit of area needed to be defined on the lattice. The unit of area was chosen, so that the density of trees would be the same in the model as in the BCI data. The 1995 census was used for all comparisons.

The parameters p and q were optimized to reproduce two quantities from the BCI dataset: (1) the number of species present on a 25 ha area; (2) the slope of the log–log SAR curve in the neighbourhood of the 25 ha area. By optimizing for these two quantities only, the RSA, as well as the SAR for the whole range of available area values could be reproduced.

An extensive search of the $p-q$ parameter space by MC simulations suggested $p=0.3$ and $q=1.3 \times 10^{-4}$ as optimal parameter values. In Fig. 7, the time evolution of the total species number (W_s) in the system is illustrated (thick continuous line). After 5000 MC steps the W_s species number reaches a stable limit and fluctuates around $W_s \approx 300$. In the same figure, the time evolution of the population for two selected species is also shown, one which appears at a later time moment and prevails, and one that appears quicker but gets extinct during the simulation (thin dashed and continuous lines, respectively).

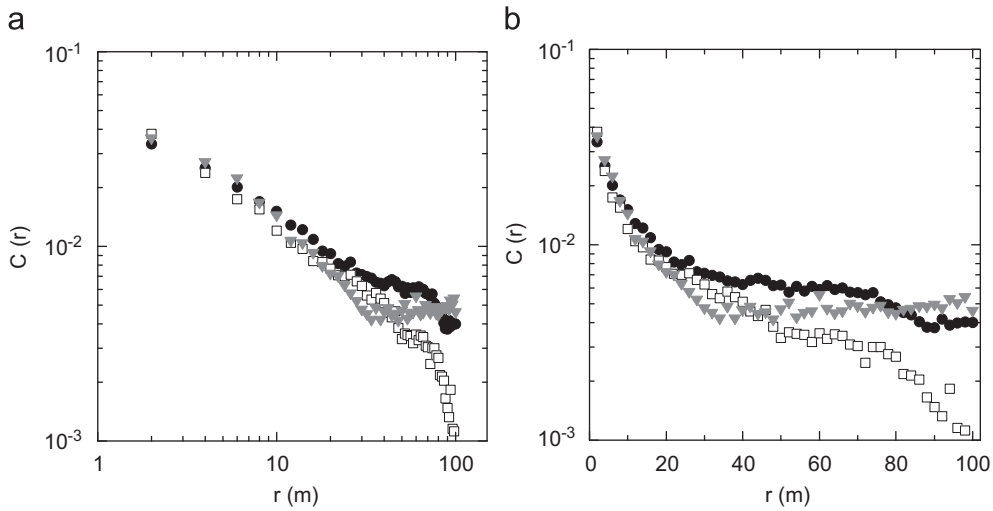

Fig. 5. The $C(r)$ spatial auto-correlation function for the three most abundant species of the BCI census: *Hybanthus prunifolius* (white squares), *Faramea occidentalis* (black circles), and *Trichilia tuberculata* (grey triangles). Census year: 1995. The same function is plotted on a log–log scale on panel (a) and a linear-log scale on panel (b).

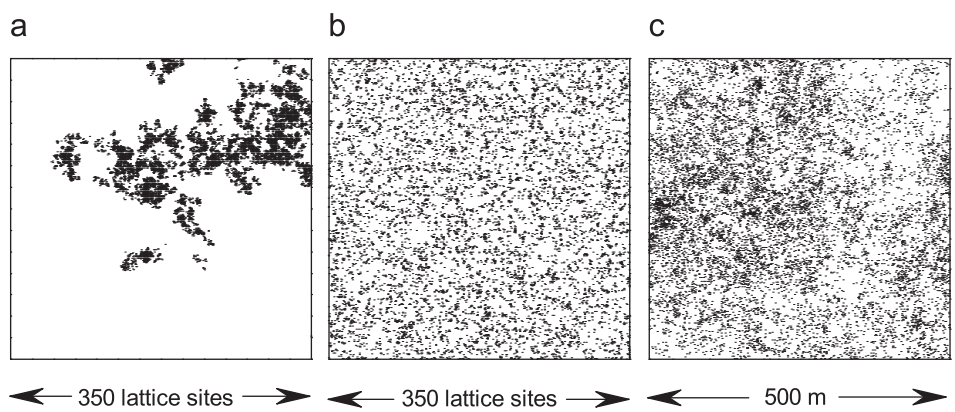

Fig. 6. Spatial distribution of the abundant species. In panels (a) and (b) simulation results are presented. Simulations were done on a 500×500 lattice with parameters $q=1.3 \times 10^{-4}$, $p=0$ (a) and $p=0.3$ (b). In panel (c) the spatial distribution of the abundant *Faramea occidentalis* species in the BCI census is shown (census year: 1995).

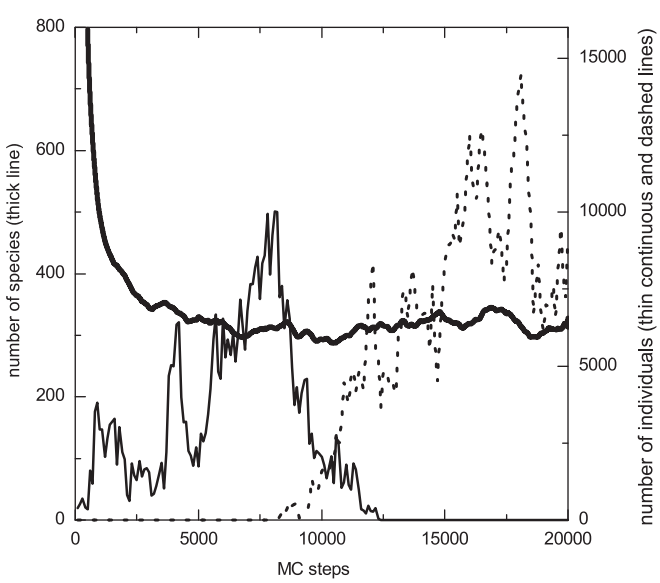

Fig. 7. Time evolution of the total number of species W_s (thick continuous line) and the population of two selected species (continuous and dashed thin line) on a 350×350 square lattice, $q=1.3 \times 10^{-4}$, $p=0.3$.

The total number of species fluctuates in time, so after the optimal p and q parameters have been found, a new set of simulations were run for the purpose of comparing the SAR along the full range of available area values: after reaching equilibrium, these simulations were run for a few more MC steps until the fluctuating species number on a 25 ha sub-lattice reached the same value that is observed in the 1995 census data, which is 273 species. The results are plotted in Fig. 8. Even though the simulation parameters were optimized to reproduce the measured species number and the slope of the log–log SAR at a 25 ha area only, the SAR is reproduced remarkably well over the complete range of areas.

The shape of the RSA curve is also in good agreement with the BCI results. The RSAs resulting from the measurement and the simulation are compared in Fig. 9. Both the Preston plot (Fig. 9a) and the probability density function (Fig. 9b) resembles the experimentally observed ones (Fig. 3).

Unfortunately, this spatially explicit model does not reproduce the observed power-law type decay (Fig. 5) of the spatial auto-correlation function. The auto-correlation functions for the abundant species that result from the model show an exponential decay for all reasonable p and q parameter values. The auto-correlation function of the most abundant species obtained for the above discussed optimal p and q values is shown in Fig. 10.

To test the inter-dependence of the two main spatial measures we used, the SAR and the spatial auto-correlation function, we calculated the SAR function for the zero-correlation case, i.e. when a set of trees is scattered completely randomly in an area. It can be shown that the zero-correlation species area relation is very well approximated by the formula $S(a)=\sum_k(1-e^{-(a/A)n_k})$, where A is the total area of the considered territory, and n_k is the number of individuals belonging to species k. Summation is over all species in the territory.

First, the SAR was calculated for the case of all the trees from the 1995 BCI dataset scattered randomly. This resulted in a SAR that is different from the actual measured one. Then we used an optimization algorithm based on simulated annealing to produce a set of 112,543 individuals belonging to 273 species (the same numbers that are found in the 1995 BCI census), so that the zero-correlation SAR would be as close as possible to the one measured in the BCI data. This resulted in a curve that is nearly indistinguishable from the measured one (Fig. 11). The RSA of this artificially created set of trees matches the experimentally measured RSA well. From this it can be concluded that the SAR curve and the spatial auto-correlation function do not strongly depend on each other. A simple, zero-correlation model can reproduce both the experimentally measured SAR and RSA. This is a useful lesson, and therefore, it is desirable to test both spatial measures (both the SAR and the auto-correlation) when comparing the predictions of a theoretical model with experimental findings.

7. Discussion and conclusions

A simple spatially explicit model, which was first considered in Zillio et al. (2005), was extended here for explaining the statistical properties of individuals and species in a tropical tree community. The model is defined on a square lattice and it is inspired by the classical voter model (Holley and Liggett, 1975). It describes in a realistic manner the spatiality of individuals, birth and death processes, diffusion of seeds and speciation events. Within the model the dynamics of species is governed by two adjustable parameters, p and q. The parameter q defines the probability of speciation, or immigration events. The parameter p describes the probability for the global diffusion of seeds within the considered area.

The system was studied by large scale Monte Carlo simulations, and the p and q parameters were adjusted, so that all experimentally studied macroecological measures would be reproduced optimally. For reasonable values of p and q, the model is successful in reproducing the shape of the species–area and RSA curves, as measured in the BCI dataset. The species–area function is reproduced remarkably well, over the full range of areas. The model also generates a visually good spatial distribution of individuals within a species. The optimal values of the model parameters were found to be $p=0.3$ and $q=1.3\times10^{-4}$.

Despite the fact that the model is simple and ignores many relevant biological processes, it is successful in reproducing several measures of spatial distribution of tree species at

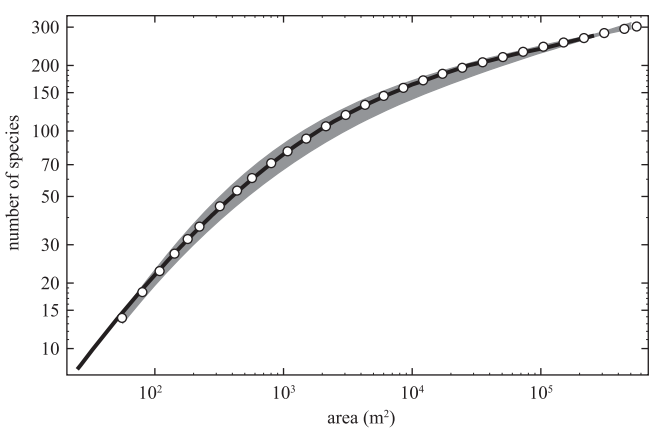

Fig. 8. Species-area curves: comparison of simulation (grey shaded area) and experimental (black line) results. The shaded area was computed from a set of 50 simulation runs, and represents the mean value ±2 standard deviations. The white circles represent one simulation result that fits the experimental data best. Simulation parameters: $L=500$, $p=0.3$ and $q=1.3\times10^{-4}$. Experimental data: BCI 1995.

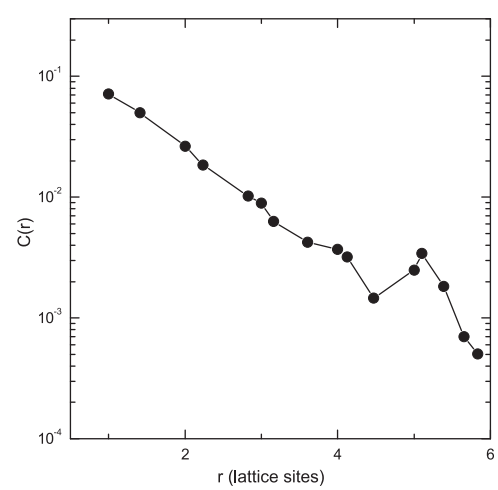

Fig. 10. The auto-correlation function of the most abundant species obtained after 50,000 MC steps. Note that the scale for $C(r)$ is logarithmic, while the scale for the radius is linear. Simulation parameters: $L=500$, $p=0.3$ and $q=1.3\times10^{-4}$.

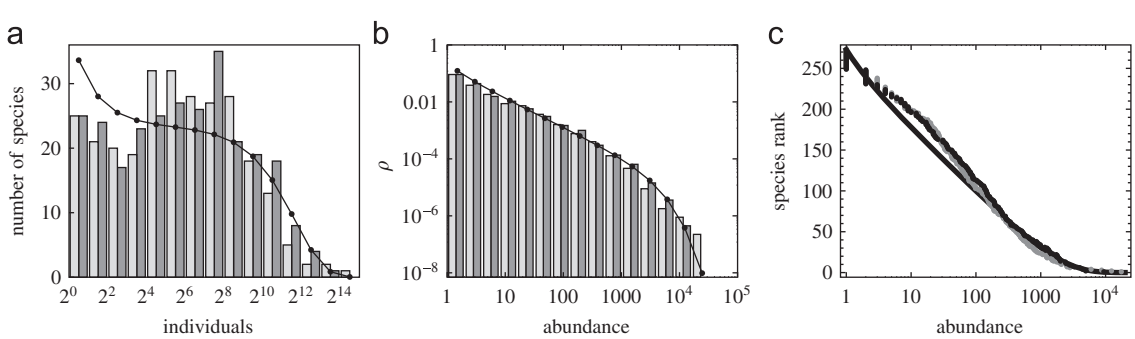

Fig. 9. Simulation results (dark grey) for the RSA curve after 50,000 MC steps, compared with the measurement data (light grey), and the predictions of the Fisher log series for $\alpha=33.64$ (solid black lines). Panel (a) shows the Preston plot, (b) the probability density function, and (c) the rank–abundance plot. Simulation parameters: $L=500$, $p=0.3$ and $q=1.3\times10^{-4}$.

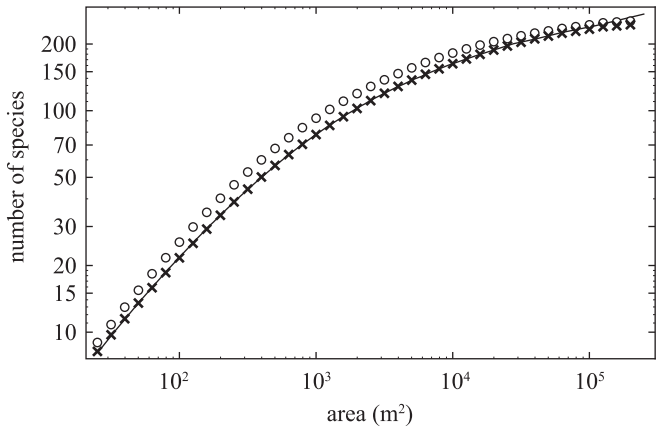

Fig. 11. The species-area relation corresponding the hypothetical case of the trees being scattered completely randomly (zero-correlation case). The circles show the SAR that is obtained by scattering the trees of the 1995 BCI census randomly, and the crosses show the same for a list of trees generated artificially to reproduce the experimentally measured SAR as well as possible. The solid line shows the SAR measured in the BCI data. Note that the crosses overlap the solid line for all but the largest areas.

the same time. Thus we believe it to capture the nature of the essential processes shaping these distributions, and to be a robust foundation for further complex models aiming to reproduce an even wider set of ecologically relevant measures.

It can be used with success to estimate the total number of species in an area and its fluctuations, from small scale samples. It can also predict qualitative effects of human impact that changes factors influencing the parameters p and q. As an example, factors that increase the value of p (which characterizes the rate of the global dispersal) lead to a lower total species number. Factors decreasing p may lead to the formation of compact islands.

Because of human influence, today species are disappearing with a higher rate than ever before, despite the numerous programs that have been started to stop this decline. In order to maintain forest biodiversity, it is important to identify the main processes that have an impact on it (Plotkin et al., 2000). Studies based on simple mathematical models, that take into account only a few primary mechanisms, are useful in such sense.

Beside all these successes, the model fails to reproduce the experimentally observed power-law decay of the spatial auto-correlation function of the individuals within a species as a function of separation distance. As one would naturally expect for lattice models where the local interactions dominate ($p,q<1$), simulations lead to an exponential decay. This is an indication that in reality long-range interactions might be more important than it was considered in this simple approach. The logical next step in improving the model is including distance dependent dispersal. Such models would necessarily need to operate with more parameters, which diminishes the elegance of the model. Another way to make the model more precise would be to make use of not only dispersal-assembly but also niche-assembly theories (Gravel et al., 2006; Potts et al., 2004).

As final conclusion, the present simple model is successful in reproducing the majority of the experimentally available macro-ecological measures.

A complete model is expected to be able to reproduce all measurable macroecological quantities. Focusing on global measures only, such as the SAR and RSA, is not sufficient. This study is such an attempt, and present a robust foundation for potential future modelling efforts.

Acknowledgements

The authors wish to thank Dr. Joshua Plotkin for useful comments. The research was financed by MTA-KMEI. Financial support of Sz. H. was provided from programs co-financed by *The sectoral operational programme human resources development*, Contract POSDRU 6/1.5/S/3—doctoral studies: through science towards society and an MTA-KMEI research Grant.

References

Alonso, D., Etienne, R.S., McKane, A.J., 2006. The merits of neutral theory. Trends Ecol. Evol. 21, 451–457.
Azaele, S., Volkov, I., Banavar, J.R., Maritan, A., 2006. Linear one-step processes with artificial boundaries. J. Stat. Phys. 125, 491–511.
Bell, G., 2001. Neutral macroecology. Science 293, 2413–2418.
Brown, J.H., 1995. Macroecology. University of Chicago Press.
Chave, J., 2004. Neutral theory and community ecology. Ecol. Lett. 7, 241–253.
Chave, J., Muller-Landau, H.C., Levin, S.A., 2002. Comparing classical community models: theoretical consequences for patterns of diversity. Am. Nat. 159, 1–23.
Condit, R., 1998. Tropical Forest Census Plots: Methods and Results from Barro Colorado. Springer.
Condit, R., Pitman, N., Leigh, E.G., Chave, J., Terborgh, J., Foster, R.B., Nunez, P., Aguilar, S., Valencia, R., Villa, G., Muller-Landau, H.C., Losos, E., Hubbell, S.P., 2002. Beta-diversity in tropical forest trees. Science 295, 666–669.
Dewdney, A., 2000. A dynamical model of communities and a new species-abundance distribution. Biol. Bull. 198, 152–165.
Enquist, B.J., Sanderson, J., Weiser, M.D., Bell, G., 2002. Modeling macroscopic patterns in ecology. Science 295, 1835–1837.
Etienne, R.S., Alonso, D., 2007. Neutral community theory: how stochasticity and dispersal-limitation can explain species coexistence. J. Stat. Phys. 128, 485–510.
Fisher, R.A., Corbet, A.S., Williams, C.B., 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. J. Anim. Ecol. 12 (1), 42–58.
Gravel, D., Canham, C.D., Beaudet, M., Messier, C., 2006. Reconciling niche and neutrality: the continuum hypothesis. Ecol. Lett. 9 (4), 399–409.
Harte, J., 2000. Scaling and self-similarity in species distribution: implications for extinction, species richness, abundance and range. In: Brown, J.H., West, G.B. (Eds.), Scaling in Ecology. Oxford University Press.
Harte, J., Kinzig, A., Green, J., 1999. Self-similarity in the distribution and abundance of species. Science 284, 334–336.
Holley, R.A., Liggett, T.M., 1975. Ergodic theorems for weakly interacting infinite systems and the voter model. Ann. Probab. 3, 643–663.
Hubbell, S.P., 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press.
Hubbell, S.P., Condit, R., Foster, R.B., 2005. Barro colorado forest census plot data.
May, R.M., 1975. Patterns of species abundance and diversity. In: Cody, M.L., Diamond, J.M. (Eds.), Ecology and Evolution of Communities. Harvard University Press.
Norris, S., 2003. Neutral theory: a new, unified model for ecology. BioScience 53, 124–129.
Pielou, E.C., 1977. Mathematical Ecology. John Wiley & Sons.
Pigolotti, S., Flammini, A., Maritan, A., 2004. Stochastic model for the species abundance problem in an ecological community. Phys. Rev. E 70, 011916.
Plotkin, J.B., Potts, M.D., Yu, D.W., Bunyavejchewin, S., Condit, R., Foster, R., Hubbell, S., LaFrankie, J., Manokaran, N., Seng, L.H., Sukumar, R., Nowak, M.A., Ashton, P.S., 2000. Predicting species diversity in tropical forests. Proc. Nat. Acad. Sci. 97 (20), 10850–10854.
Potts, M., Davies, S., Bossert, W., Tan, S., Supardi, M.N., 2004. Habitat heterogeneity and niche structure of trees in two tropical rain forests. Oecologia 139 (3), 446–453.
Preston, F.W., 1962. The canonical distribution of commonness and rarity: part i. Ecology 43, 185–215.
Tokeshi, M., 1999. Species Coexistence. Blackwell.
Volkov, I., Banavar, J.R., Hubbell, S.P., Maritan, A., 2003. Neutral theory and relative species abundance in ecology. Nature 424, 1035–1037.
Volkov, I., Banavar, J.R., Maritan, A., Hubbell, S.P., 2004a. Neutral theory (communication arising): the stability of forest biodiversity. Nature 427, 696.
Volkov, I., Banavar, J.R., Maritan, A., 2004b. Organization of ecosystems in the vicinity of a novel phase transition. Phys. Rev. Lett. 92, 218703.
Volkov, I., Banavar, J.R., He, F., Hubbell, S.P., Maritan, A., 2005. Density dependence explains tree species abundance and diversity in tropical forests. Nature 438, 658–661.
Volkov, I., Banavar, J.R., Hubbell, S.P., Maritan, A., 2007. Patterns of relative species abundance in rainforests and coral reefs. Nature 450, 45–49.
Zillio, T., Volkov, I., Banavar, J.R., Hubbell, S.P., Maritan, A., 2005. Spatial scaling in model plant communities. Phys. Rev. Lett. 95, 098101–098104.