Available online at www.sciencedirect.com

ScienceDirect

Journal of Theoretical Biology 252 (2008) 288–294

Journal of Theoretical Biology

www.elsevier.com/locate/yjtbi

Relaxing the zero-sum assumption in neutral biodiversity theory

Bart Haegeman[a,b], Rampal S. Etienne[c,*]

[a]*INRA, UR50, Laboratory of Environmental Biotechnology, Avenue des Etangs, F-11100 Narbonne, France*
[b]*INRA-INRIA Research Team MERE, UMR Systems Analysis and Biometrics, 2 place Pierre Viala, F-34060 Montpellier, France*
[c]*Community and Conservation Ecology Group, Centre for Ecological and Evolutionary Studies, University of Groningen, P.O. Box 14, 9750 AA Haren, The Netherlands*

Received 23 September 2007; received in revised form 13 December 2007; accepted 22 January 2008
Available online 2 February 2008

Abstract

The zero-sum assumption is one of the ingredients of the standard neutral model of biodiversity by Hubbell. It states that the community is saturated all the time, which in this model means that the total number of individuals in the community is constant over time, and therefore introduces a coupling between species abundances. It was shown recently that a neutral model with independent species, and thus without any coupling between species abundances, has the same sampling formula (given a fixed number of individuals in the sample) as the standard model [Etienne, R.S., Alonso, D., McKane, A.J., 2007. The zero-sum assumption in neutral biodiversity theory. J. Theor. Biol. 248, 522–536]. The equilibria of both models are therefore equivalent from a practical point of view. Here we show that this equivalence can be extended to a class of neutral models with density-dependence on the community-level. This result can be interpreted as robustness of the model, i.e. insensitivity of the model to the precise interaction of the species in a neutral community. It can also be interpreted as a lack of resolution, as different mechanisms of interactions between neutral species cannot be distinguished using only a single snapshot of species abundance data.

Keywords: Density-dependence; Neutral model; Detailed balance; Abundance vector

1. Introduction

Neutral biodiversity theory has received much attention since it was revived by Hubbell (2001) and major theoretical advances have been made (reviewed, for instance, in Etienne and Alonso, 2007). Most authors have focused on its most essential assumption, the neutrality assumption, which states that all individuals, regardless of species, behave identically under identical circumstances. Although this assumption seems unrealistic for most ecological communities, neutral theory has been recognized as providing useful null models for comparison with alternative models incorporating species differences and has sharpened our tests to detect deviations from neutrality (Alonso et al., 2006; McGill et al., 2006). We need to be aware, however, that rejection of neutral models failing such tests may not be (completely) due to the neutrality assumption; it may be due to other assumptions in the model, which is, after all, only a single, particular implementation of neutral theory (Etienne, 2007).

In the standard neutral model of biodiversity (Hubbell, 2001), such an assumption is the zero-sum assumption. It states that the individuals in an ecological community play a zero-sum game: each death or emigration event is immediately followed by a birth or immigration event, such that the community remains saturated and therefore introduces a coupling between species abundances. If the resources (e.g. space, light) are constant over time, the zero-sum assumption implies that the total number of individuals in the community is also constant over time. In the standard neutral model it is indeed assumed that community size is fixed. In this paper we will refer to this model as the fixed-community-size (fcs) model.

It was shown recently that a neutral model with independent species—hereafter called the independent-species model (ind)—and thus without any coupling

---

[*] Corresponding author. Tel.: +31 503632230.
E-mail addresses: bart.haegeman@inria.fr (B. Haegeman), r.s.etienne@rug.nl (R.S. Etienne).

0022-5193/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2008.01.023

between species abundances has the same sampling formula (given a fixed number of individuals in the sample) as the fixed-community-size model (Etienne et al., 2007). The equilibrium of both models are therefore equivalent from a practical point of view. In this paper we will show that this equivalence can be extended to a class of neutral models with density-dependence on the community-level. This class includes the fixed-community-size model and the independent-species model as special cases.

We start by giving an alternative proof for the sampling formula of the independent-species model in equilibrium, because it provides a nice and simple introduction to our extension to community-level density-dependent (cdd) neutral models. This alternative proof is based on the fact that the equilibrium state of the independent-species model can be written as a combination of the equilibrium states of fixed-community-size models for different community sizes. Next, we introduce the novel neutral models with community-level density-dependence, and show that their equilibrium state satisfies the same property. This means that the density-dependent neutral models all have the same sampling formula as the fixed-community-size model and the independent-species model.

2. Independent species

We formulate the independent species model as in Etienne et al. (2007), so it only differs from the fixed-community-size model in that there is no coupling between the population sizes (abundances) of the different species. To briefly summarize, the local community receives immigrants from the metacommunity. This metacommunity consists of S_M species, and is assumed to be so large that only relative abundances have to be taken into account. We denote the relative abundance of species k in the metacommunity by p_k (see Table 1 for an overview of our notation). Because the local community is smaller than the metacommunity and because it is dispersal-limited, it contains usually (much) fewer species than the metacommunity, but potentially it can contain all S_M species, so we will keep track of the abundances of all S_M species, even if they are (temporarily) zero. We denote the abundance of species k in the local community by N_k and define the local community abundance vector, $\vec{N} = (N_1, \ldots, N_{S_M})$. We denote the total abundance of the local community by J_L, that is,

$$J_L = \sum_k N_k.$$

The local community dynamics are determined by the repetitive occurrence of birth, death and immigration events. Due to the neutrality assumption, the birth rate per individual (denoted by β), the death rate per individual (denoted by δ), and the immigration rate from meta- to local community (denoted by λ) do not depend on the identity of the involved species. For species k with abundance N_k, the rate of abundance increase g_{N_k} and the rate of abundance decrease r_{N_k} are given by

Table 1
Explanation of mathematical symbols

Symbol	Explanation
S_M	Number of species in metacommunity
$\vec{p} = (p_1, \ldots, p_{S_M})$	Relative abundance vector of metacommunity
$\vec{N} = (N_1, \ldots, N_{S_M})$	Abundance vector of local community
$J_L = \sum_k N_k$	Local community size
$\vec{n} = (n_1, \ldots, n_{S_M})$	Abundance vector of sample
$J = \sum_k n_k$	Sample size
$P_{fcs}, P_{ind}, P_{cdd}$	Equilibrium probability distribution for the fixed-community-size model, the independent-species model, and the community-level density-dependent model, respectively. Thus, $P_{cdd}(\vec{N} \mid \ldots)$ denotes the probability of abundance vector $\vec{N}$, $P_{cdd}(J_L \mid \ldots)$ denotes the probability of local community size J_L, etc.
$P_{fcs,hyp}, P_{ind,hyp}, P_{cdd,hyp}$	Equilibrium probability distribution of the sample composition (i.e. the abundance vector $\vec{n}$) under sampling without replacement, described by the hypergeometric distribution
β, δ, λ	Birth, death, and immigration rate; constant for the independent-species model, and a function of the local community size J_L for the density-dependent model
g_{N_k}, r_{N_k}	Rate of abundance increase (by birth and immigration) and decrease (by death) of species k with abundance N_k; constant for the independent-species model, and a function of the local community size J_L for the density-dependent model
I	Fundamental dispersal number, or ratio of immigration rate λ and birth rate β for the independent-species and density-dependent model
R	Ratio of birth rate β and death rate δ for the independent-species and density-dependent model

$$g_{N_k} = \beta N_k + \lambda p_k,$$

$$r_{N_k} = \delta N_k.$$

We can then write down the corresponding master equation (Van Kampen, 1992). This is a differential equation for the probability $P(N_k, t)$ that at time t the abundance of species k in the local community is N_k:

$$\frac{\mathrm{d}}{\mathrm{d}t} P(N_k, t) = g_{N_k - 1} P(N_k - 1, t) + r_{N_k + 1} P(N_k + 1, t) - g_{N_k} P(N_k, t) - r_{N_k} P(N_k, t). \tag{1}$$

All species $k = 1, \ldots, S_M$ satisfy such an equation, without any coupling between them. Because the species are independent, the probability $P(\vec{N}, t)$ that at time t the abundance vector is $\vec{N}$ can be written simply as a product of the probabilities $P(N_k, t)$ for all the species $k = 1, \ldots, S_M$:

$$P(\vec{N}, t) = \prod_{k=1}^{S_M} P(N_k, t). \tag{2}$$

Eqs. (1) and (2) specify mathematically the independent-species model.

In this paper we are exclusively interested in the equilibrium probability distribution for the abundance vector $\vec{N}$, $P(\vec{N})$. For the independent-species model this boils down, according to (2), to finding the equilibrium probability distribution of the abundance N_k of each species, $P(N_k)$. This can be found by requiring that the right-hand side of (1) vanishes for all $N_k = 0, 1, 2, \ldots$. The solution, which we denote by P_{ind}, is given by (Kendall, 1948)

$$P_{ind}(N_k \,|\, p_k, I, R) = (1-R)^{Ip_k} \frac{R^{N_k}(Ip_k)_{N_k}}{N_k!}, \tag{3}$$

with $R = \beta/\delta$ and $I = \lambda/\beta$, and where we used the Pochhammer notation

$$(a)_n = \prod_{k=0}^{n-1}(a+k).$$

To show this, we first note that in the state P_{ind} and for each species k, the frequency of transitions from $N_k + 1$ to N_k equals the frequency of transitions from N_k to $N_k + 1$:

$$\begin{aligned} r_{N_k+1}&P_{ind}(N_k+1 \,|\, p_k, I, R) \\ &= \delta(N_k+1)\frac{R(Ip_k+N_k)}{N_k+1}P_{ind}(N_k \,|\, p_k, I, R) \\ &= (\lambda p_k + \beta N_k)P_{ind}(N_k \,|\, p_k, I, R) \\ &= g_{N_k}P_{ind}(N_k \,|\, p_k, I, R). \end{aligned} \tag{4}$$

Therefore, the right-hand side of (1) becomes

$$\begin{aligned} g_{N_k-1}&P_{ind}(N_k-1 \,|\, p_k, I, R) + r_{N_k+1}P_{ind}(N_k+1 \,|\, p_k, I, R) \\ &- g_{N_k}P_{ind}(N_k \,|\, p_k, I, R) - r_{N_k}P_{ind}(N_k \,|\, p_k, I, R) \\ &\quad = (r_{N_k+1}P_{ind}(N_k+1 \,|\, p_k, I, R) - g_{N_k}P_{ind}(N_k \,|\, p_k, I, R)) \\ &\qquad - (r_{N_k}P_{ind}(N_k \,|\, p_k, I, R) - g_{N_k-1} \\ &\qquad \times P_{ind}(N_k-1 \,|\, p_k, I, R)) \\ &\quad = 0 - 0 = 0, \end{aligned}$$

so the right-hand side of (1) indeed vanishes when we substitute (3). The fact that the terms cancel out two by two is a general property of this kind of master equation, and is called detailed balance, as every elementary process is balanced by its inverse process (Van Kampen, 1992, Chapter 5) due to (4). Note that the prefactor in (3) is such that the probabilities $P_{ind}(N_k \,|\, p_k, I, R)$ sum to one. A technical detail the importance of which will become clear in the next section: in order that this sum converges, we have to assume that $R < 1$ or $\beta < \delta$.

The equilibrium probability distribution P_{ind} of the abundance vector $\vec{N}$ for the independent-species model follows from (2) and (3):

$$\begin{aligned} P_{ind}(\vec{N} \,|\, \vec{p}, I, R) &= \prod_k P_{ind}(N_k \,|\, p_k, I, R) \\ &= (1-R)^I R^{\sum_k N_k} \prod_k \frac{(Ip_k)_{N_k}}{N_k!}. \end{aligned} \tag{5}$$

In contrast with Hubbell's neutral model with fixed-community-size, the size of the local community

$$J_L = \sum_k N_k$$

varies over time in the independent-species neutral model. Even in the equilibrium state the community size fluctuates. Its probability distribution, denoted by $P_{ind}(J_L \,|\, \vec{p}, I, R)$, can be computed as follows:

$$\begin{aligned} P_{ind}(J_L \,|\, \vec{p}, I, R) &= \sum_{\vec{N}:\sum_k N_k = J_L} P_{ind}(\vec{N} \,|\, \vec{p}, I, R) \\ &= (1-R)^I R^{J_L} \frac{(I)_{J_L}}{J_L!}, \end{aligned} \tag{6}$$

where the sum is over all abundance vectors that correspond to community size J_L. The species-independent distribution $P_{ind}(\vec{N} \,|\, \vec{p}, I, R)$, given by (5), is closely related to the equilibrium distribution for the fixed-community-size model, which we denote by $P_{fcs}(\vec{N} \,|\, \vec{p}, I, J_L)$. The latter is given by (Etienne and Alonso, 2005; Etienne et al., 2007; Rannala, 1996)

$$P_{fcs}(\vec{N} \,|\, \vec{p}, I, J_L) = \frac{J_L!}{(I)_{J_L}} \prod_k \frac{(Ip_k)_{N_k}}{N_k!}. \tag{7}$$

Combining (5)–(7), we find

$$\begin{aligned} P_{ind}(\vec{N} \,|\, \vec{p}, I, R) &= (1-R)^I R^{J_L} \frac{(I)_{J_L}}{J_L!} P_{fcs}(\vec{N} \,|\, \vec{p}, I, J_L) \\ &= P_{ind}(J_L \,|\, \vec{p}, I, R) P_{fcs}(\vec{N} \,|\, \vec{p}, I, J_L). \end{aligned} \tag{8}$$

The equilibrium distribution for the independent-species model can thus be written as a combination of the equilibrium distributions $P_{fcs}(\vec{N} \,|\, \vec{p}, I, J_L)$ for fixed-community-size models having different community sizes J_L with weights $P_{ind}(J_L \,|\, \vec{p}, I, R)$ that are simply the probabilities that the independent-species model has community size J_L. Stated differently, taking the independent-species equilibrium distribution and conditioning on a given community size J_L, yields the fixed-community-size equilibrium distribution for that community size J_L,

$$\begin{aligned} P_{ind}(\vec{N} \,|\, \vec{p}, I, R, J_L) &= \frac{P_{ind}(\vec{N} \,|\, \vec{p}, I, R)}{P_{ind}(J_L \,|\, \vec{p}, I, R)} \\ &= P_{fcs}(\vec{N} \,|\, \vec{p}, I, J_L), \end{aligned} \tag{9}$$

where we used (8).

As noted by Etienne et al. (2007), the connection between both models can be expressed on the level of sampling properties. To show this, we consider a sample of size $J \leqslant J_L$ from both equilibrium distributions (sampling without replacement), and denote the sample abundance vector by $\vec{n} = (n_1, \ldots, n_{S_M})$ with

$$J = \sum_k n_k.$$

The probability to obtain the sample abundance vector $\vec{n}$ from the local community abundance vector $\vec{N}$ is given by the multivariate hypergeometric distribution (as this is the appropriate distribution for sampling without replacement)

$$P_{hyp}(\vec{n}\,|\,\vec{N}, J_L, J) = \frac{\prod_{k=1}^{S_M}\binom{N_k}{n_k}}{\binom{J_L}{J}}.$$

The sample abundance distribution for the fixed-community-size model, denoted by $P_{fcs,hyp}$, is given by

$$\begin{aligned} &P_{fcs,hyp}(\vec{n}\,|\,\vec{p}, I, J_L, J)\\ &\quad = \sum_{\vec{N}} P_{hyp}(\vec{n}\,|\,\vec{N}, J_L, J) P_{fcs}(\vec{N}\,|\,\vec{p}, I, J_L), \end{aligned} \tag{10}$$

and the sample abundance distribution for the independent-species model, denoted by $P_{ind,hyp}$, is given by

$$\begin{aligned} &P_{ind,hyp}(\vec{n}\,|\,\vec{p}, I, R, J)\\ &\quad = \sum_{J_L} P_{ind}(J_L\,|\,\vec{p}, I, R) \sum_{\vec{N}} P_{hyp}(\vec{n}\,|\,\vec{N}, J_L, J)\\ &\qquad \times P_{ind}(\vec{N}\,|\,\vec{p}, I, R, J_L), \end{aligned} \tag{11}$$

where the sums over $\vec{N}$ are taken over all abundance vectors with community size J_L and compatible with the sample abundance vector $\vec{n}$.

The equilibrium distribution of both models obey identical sampling formulas, that is,

$$P_{ind,hyp}(\vec{n}\,|\,\vec{p}, I, R, J) = P_{fcs,hyp}(\vec{n}\,|\,\vec{p}, I, J_L, J). \tag{12}$$

Eq. (12) can be derived as follows. First we note the sampling invariance of the fixed-community-size equilibrium distribution (Etienne and Alonso, 2005)

$$P_{fcs,hyp}(\vec{n}\,|\,\vec{p}, I, J_L, J) = P_{fcs}(\vec{n}\,|\,\vec{p}, I, J). \tag{13}$$

Then the derivation is straightforward:

$$\begin{aligned} &P_{ind,hyp}(\vec{n}\,|\,\vec{p}, I, R, J)\\ &\quad = \sum_{J_L} P_{ind}(J_L\,|\,\vec{p}, I, R) \sum_{\vec{N}} P_{hyp}(\vec{n}\,|\,\vec{N}, J_L, J)\\ &\qquad \times P_{ind}(\vec{N}\,|\,\vec{p}, I, R, J_L)\\ &\quad = \sum_{J_L} P_{ind}(J_L\,|\,\vec{p}, I, R) \sum_{\vec{N}} P_{hyp}(\vec{n}\,|\,\vec{N}, J_L, J)\\ &\qquad \times P_{fcs}(\vec{N}\,|\,\vec{p}, I, J_L)\\ &\quad = \sum_{J_L} P_{ind}(J_L\,|\,\vec{p}, I, R) P_{fcs,hyp}(\vec{n}\,|\,\vec{p}, I, J_L, J)\\ &\quad = \sum_{J_L} P_{ind}(J_L\,|\,\vec{p}, I, R) P_{fcs}(\vec{n}\,|\,\vec{p}, I, J)\\ &\quad = P_{fcs}(\vec{n}\,|\,\vec{p}, I, J)\\ &\quad = P_{fcs,hyp}(\vec{n}\,|\,\vec{p}, I, J_L, J), \end{aligned}$$

where we used (11) in the second line, (9) in the third line, (10) in the fourth line, the law of total probability in the sixth line, and (13) in the fifth and seventh line. Eq. (12) was proved by Etienne et al. (2007) in the light of their general sampling theory, but we have presented a different derivation here, to introduce our extension to community-level density-dependent neutral models. Also, although perhaps trivial, the fact that we have detailed balance, has not been stressed in the literature.

3. Community-level density-dependence

The independent-species equilibrium distribution has the property that by conditioning on community size J_L, one retrieves the fixed-community-size equilibrium distribution, see (9). Other neutral models can be constructed with the same property. We introduce here a class of such models with density-dependence relative to the community density (instead of relative to the species density, see Discussion). We will refer to this type of model as the community-level density-dependent model.

In the independent-species model, the birth rate β, death rate δ and immigration rate λ are constants. In the density-dependent model, we take rates that depend on the community size

$$J_L = \sum_k N_k.$$

We denote them by $\beta(J_L)$, $\delta(J_L)$, and $\lambda(J_L)$, respectively. For species k with abundance N_k, the rate of abundance increase g_{N_k} and the rate of abundance decrease r_{N_k} also depend on J_L, and are given by

$$g_{N_k}(J_L) = \beta(J_L)N_k + \lambda(J_L)p_k,$$

$$r_{N_k}(J_L) = \delta(J_L)N_k.$$

Moreover, we have to impose a condition on the birth and immigration rates in order to keep the formulas tractable: the fundamental dispersal number I (Etienne and Alonso, 2005) must be independent of community size, that is,

$$I = \frac{\lambda(J_L)}{\beta(J_L)} \quad \text{is independent of } J_L. \tag{14}$$

This assumption will be discussed below.

Our analysis of the density-dependent model is analogous to the independent-species case in the previous section. We first formulate the master equation. Due to the coupling between species, we can no longer write down an equation like (1) in terms of the abundance of a single species. Also property (2) is no longer satisfied. Instead, we have to consider the entire abundance vector $\vec{N}$. Denoting by $P(\vec{N}, t)$ the probability that at time t the abundance vector is $\vec{N}$, the master equation reads

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} P(\vec{N}, t) = &\sum_k g_{N_k-1}(J_L-1) P(\vec{N}-\vec{e}_k, t)\\ &+ \sum_k r_{N_k+1}(J_L+1) P(\vec{N}+\vec{e}_k, t)\\ &- \sum_k g_{N_k}(J_L) P(\vec{N}, t) - \sum_k r_{N_k}(J_L) P(\vec{N}, t), \end{aligned} \tag{15}$$

where we used basis vectors $\vec{e}_k$ with components

$$(\vec{e}_k)_l = \begin{cases} 1 & \text{if } k = l, \\ 0 & \text{otherwise.} \end{cases}$$

Eq. (15) specifies mathematically the community-level density-dependent model.

To find the equilibrium distribution of this model, we require that the right-hand side of (15) vanishes for all vectors $\vec{N}$. Below we show that the solution to setting (15) equal to zero, which we denote by $P_{cdd}(\vec{N}\,|\,\vec{p},\lambda,\beta,\delta)$, is given by

$$P_{cdd}(\vec{N}\,|\,\vec{p},\lambda,\beta,\delta) = P_{cdd}(J_L\,|\,\vec{p},\lambda,\beta,\delta)P_{fcs}(\vec{N}\,|\,\vec{p},I,J_L), \tag{16}$$

where the probability distribution of the local community size J_L satisfies

$$P_{cdd}(J_L\,|\,\vec{p},\lambda,\beta,\delta) = P_{cdd}(0\,|\,\vec{p},\lambda,\beta,\delta)\frac{(I)_{J_L}}{J_L!}\prod_{k=1}^{J_L}\frac{\beta(k-1)}{\delta(k)}. \tag{17}$$

The coefficients $P_{cdd}(0\,|\,\vec{p},\lambda,\beta,\delta)$ follow from the requirement that the probabilities $P_{cdd}(J_L\,|\,\vec{p},\lambda,\beta,\delta)$ must sum to one:

$$P_{cdd}(0\,|\,\vec{p},\lambda,\beta,\delta) = \left(1 + \sum_{J_L=1}^{\infty}\frac{(I)_{J_L}}{J_L!}\prod_{k=1}^{J_L}\frac{\beta(k-1)}{\delta(k)}\right)^{-1}.$$

To prove (16), we show, analogously to the independent-species case considered above, that in the state P_{cdd} and for each species k, the frequency of transitions from population size N_k+1 (and total community size J_L+1) to N_k (and community size J_L, all else being equal) equals the frequency of transitions from N_k to N_k+1:

$$\begin{aligned}
&r_{N_k+1}(J_L+1)P_{cdd}(\vec{N}+\vec{e}_k\,|\,\vec{p},\lambda,\beta,\delta)\\
&\quad = (\delta(J_L+1)(N_k+1))P_{cdd}(J_L+1\,|\,\vec{p},\lambda,\beta,\delta)\\
&\qquad \times P_{fcs}(\vec{N}+\vec{e}_k\,|\,\vec{p},I,J_L+1)\\
&\quad = (\delta(J_L+1)(N_k+1))\frac{I+J_L}{J_L+1}\frac{\beta(J_L)}{\delta(J_L+1)}P_{cdd}(J_L\,|\,\vec{p},\lambda,\beta,\delta)\\
&\qquad \times \frac{J_L+1}{I+J_L}\frac{N_k+Ip_k}{N_k+1}P_{fcs}(\vec{N}\,|\,\vec{p},I,J_L)\\
&\quad = (\beta(J_L)(N_k+Ip_k))P_{cdd}(J_L\,|\,\vec{p},\lambda,\beta,\delta)P_{fcs}(\vec{N}\,|\,\vec{p},I,J_L)\\
&\quad = (\beta(J_L)N_k+\lambda(J_L)p_k)P_{cdd}(\vec{N}\,|\,\vec{p},\lambda,\beta,\delta)\\
&\quad = g_{N_k}(J_L)P_{cdd}(\vec{N}\,|\,\vec{p},\lambda,\beta,\delta),
\end{aligned} \tag{18}$$

where we used (16) in the second line, (17) in the third line, (7) in the fourth line, and (14) and (16) in the sixth line. Here we see that the assumption of I being independent of J_L, as expressed in (14), is crucial. For the right-hand side of the master equation (15), we then obtain

$$\begin{aligned}
&\sum_k g_{N_k-1}(J_L-1)P_{cdd}(\vec{N}-\vec{e}_k\,|\,\vec{p},\lambda,\beta,\delta)\\
&\quad + \sum_k r_{N_k+1}(J_L+1)P_{cdd}(\vec{N}+\vec{e}_k\,|\,\vec{p},\lambda,\beta,\delta)\\
&\quad - \sum_k g_{N_k}(J_L)P_{cdd}(\vec{N}\,|\,\vec{p},\lambda,\beta,\delta)\\
&\quad - \sum_k r_{N_k}(J_L)P_{cdd}(\vec{N}\,|\,\vec{p},\lambda,\beta,\delta)\\
&\qquad = \sum_k \Big(r_{N_k+1}(J_L+1)P_{cdd}(\vec{N}+\vec{e}_k\,|\,\vec{p},\lambda,\beta,\delta)\\
&\qquad\quad - g_{N_k}(J_L)P_{cdd}(\vec{N}\,|\,\vec{p},\lambda,\beta,\delta)\Big)\\
&\qquad\quad - \sum_k \Big(r_{N_k}(J_L)P_{cdd}(\vec{N}\,|\,\vec{p},\lambda,\beta,\delta)\\
&\qquad\quad - g_{N_k-1}(J_L-1)P_{cdd}(\vec{N}-\vec{e}_k\,|\,\vec{p},\lambda,\beta,\delta)\Big)\\
&\qquad = \sum_k 0 - \sum_k 0 = 0,
\end{aligned}$$

where we used (16) and (18). All terms cancel out two by two, which is again an instance of detailed balance. Thus, the right-hand side of (15) indeed becomes zero for the solution (16).

We noted that the equilibrium distribution (5) for the independent-species model only exists only if birth and death rates satisfy $\beta < \delta$. Analogously, we have to impose some conditions on the functions β and δ, in order to guarantee existence and uniqueness of the equilibrium distribution (16) and (17). These conditions basically prevent the community size to grow without bound, by imposing that the birth rate $\beta(J_L)$ does not exceed the death rate $\delta(J_L)$ for large community size J_L. A set of such conditions is

$$\beta(J_L) > 0 \quad \text{for all } J_L \leqslant M,$$

$$\delta(J_L) > 0 \quad \text{for all } J_L,$$

$$\delta(J_L) \geqslant \beta(J_L)(1+\varepsilon) \quad \text{for all } J_L > M, \tag{19}$$

for an integer M and a positive number $\varepsilon > 0$. We refer to Van Kampen (1992, Chapter 5) for a more elaborate discussion.

By way of illustration, we show that a common model for logistic growth satisfies these conditions. Consider an immigration rate λ independent of community size. In order to satisfy condition (14), the birth rate β must also be community-size independent. We assign density-dependence to the death rate δ:

$$\delta(J_L) = d + (\beta - d)\frac{J_L}{K},$$

with integer K and $0 < d < \beta$. To verify that conditions (19) are fulfilled, one can take $M = K$ and $\varepsilon = (1/K)(1 - d/\beta) > 0$, so that indeed

$$\begin{aligned}\delta(J_L) &= d + (\beta - d)\frac{J_L}{K} \\ &= d + (\beta - d)\left(1 + \frac{J_L - K}{K}\right) \\ &= \beta\left(1 + \frac{J_L - K}{K}\left(1 - \frac{d}{\beta}\right)\right) \geqslant \beta(1+\varepsilon),\end{aligned}$$

for $J_L > K$. For $J_L = K$, the birth rate β equals the death rate δ, and thus K can be interpreted as a carrying capacity, in the sense that at sizes larger than K the death rate exceeds the birth rate. This does not imply, however, that the community size cannot become larger than K. Immigration and stochastic fluctuations may lead to community sizes that exceed the carrying capacity K. This choice of birth and death rates is just an example; density-dependence in the birth rate is also possible as long as $\lambda(J_L)/\beta(J_L)$ is independent of J_L. We discuss this further below.

Obviously, the independent-species model is included in the class of density-dependent models. One simply has to take rates β and δ independent of community size. Conditions (19) then yield $\delta > \beta > 0$, as we found in the previous section. Interestingly, also the model with fixed-community-size J_L can be reconstructed in the class of community-level density-dependent models. To show this, we take rates β and δ as follows:

$$\beta(J_L - 1) = b \quad \text{and} \quad \beta(k) = 0 \ \text{for}\ k \neq J_L - 1,$$

$$\delta(J_L) = d \quad \text{and} \quad \delta(k) = 0 \quad \text{for}\ k \neq J_L,$$

with $b \gg d > 0$, and the initial condition that the community size is J_L. If the community size equals J_L, the only event that can happen is the death of an individual (because $\beta(J_L) = 0$ and, because of (14), $\lambda(J_L) = 0$). The corresponding decrease in community size is then quickly (because $b \gg d$) followed by a new event that increase the community size again (because $\delta(J_{L-1}) = 0$). This closely resembles Hubbell's neutral model. In fact, the larger the separation between time scales $1/d \gg 1/b$, the better this density-dependent model approximates the fixed-community-size model. The community-size equilibrium distribution (17) is given by

$$P_{cdd}(k \mid \vec{p}, \lambda, \beta, \delta) = \begin{cases} \dfrac{d}{b+d} & \text{if } k = J_L - 1, \\ \dfrac{b}{b+d} & \text{if } k = J_L, \\ 0 & \text{otherwise,} \end{cases}$$

which is sharply peaked at J_L (because $b \gg d$). Note, however, that without the initial condition this model has several equilibria, as a community with size $k \notin \{J_L - 1, J_L\}$ will never change (because $\beta(k) = \delta(k) = 0$). This is consistent with the fact that conditions (19) are not satisfied for this model. To make the model satisfy these conditions, we can alternatively take rates β and δ as follows:

$$\beta(k) = a_3 \quad \text{and} \quad \delta(k) = a_2 \quad \text{for all}\ k < J_L,$$

$$\beta(k) = a_1 \quad \text{and} \quad \delta(k) = a_2 \quad \text{for}\ k = J_L,$$

$$\beta(k) = a_2 \quad \text{and} \quad \delta(k) = a_3 \quad \text{for all}\ k > J_L,$$

with rates a_1, a_2 and a_3 satisfying $0 < a_1 \ll a_2 \ll a_3$. Whenever the community has a size different from J_L it moves with very high probability in the direction of size J_L. Again, the larger the separation between time scales $1/a_1 \gg 1/a_2 \gg 1/a_3$, the better this density-dependent model approximates the fixed-community-size model. One can take $M = J_L$ and $\varepsilon = a_3/a_2 - 1 > 0$ to see that conditions (19) are satisfied. The model has therefore a unique equilibrium distribution P_{cdd}, for which the community-size distribution (17) is again sharply peaked at J_L.

Finally, the parallel between (8) and (16) suggests that the fixed-community-size and density-dependent equilibrium are strongly linked. Indeed, the density-dependent equilibrium conditioned on community size J_L gives the fixed-community-size equilibrium distribution

$$P_{cdd}(\vec{N} \mid \vec{p}, \lambda, \beta, \delta, J_L) = P_{fcs}(\vec{N} \mid \vec{p}, I, J_L),$$

a property analogous to (9). Together with the sampling invariance of the fixed-community-size equilibrium (13), it is easy to see that the class of density-dependent models all have the same sampling formula:

$$P_{cdd,hyp}(\vec{n} \mid \vec{p}, \lambda, \beta, \delta, J) = P_{fcs,hyp}(\vec{n} \mid \vec{p}, I, J_L, J),$$

which is the equivalent of (12).

4. Discussion

In Hubbell's neutral model the zero-sum assumption couples different species together. Here we have shown that the structure of the equilibrium distribution seems to be rather insensitive to this assumption. Indeed, we have introduced a class of models where species are coupled differently, including the model with independent species, and computed explicitly their equilibrium distribution. By conditioning on J_L we recover the fixed-community-size equilibrium. Therefore, we also obtain identical sampling formulas for all these models. Strictly, we have only demonstrated equivalence of the sampling formulas for the local community model, where the local community is connected to a metacommunity with given relative abundances p_k. However, Etienne et al. (2007) have shown that the metacommunity model can also be regarded as a birth–death–immigration model where speciation is modelled as immigration from an infinite species pool with all relative abundances vanishingly small. Therefore our results also apply to the metacommunity, and thus increase the generality of the Ewens sampling formula that describes this metacommunity (Ewens, 1972; Hubbell, 2001).

The class of neutral models introduced here have birth, growth and immigration rates that depend on the community size J_L. We called this density-dependence relative to the community density, in contrast with density-dependence relative to the species density, as discussed, e.g. in Volkov et al. (2005). In the latter case, the birth, growth and immigration rates of species k depend on the abundance N_k of species k. Note that only density-dependence relative to the community density is compatible with a strict interpretation of neutrality, defined by the condition that all individuals behave the same, irrespective of the species they belong to. Density-dependence relative to the species density is compatible with a broader notion of neutrality, defined by the condition that all individuals behave the same under the same intraspecific circumstances, which has been termed symmetry (Etienne and Olff, 2005; Etienne, 2007; Alonso et al., 2008).

The community-level density-dependent models discussed here can be interpreted as perturbations of Hubbell's neutral model. We showed that the equilibrium state of these models can be computed explicitly, and that these equilibria satisfy the same sampling formula as the fixed-community-size model. We expect that these statements no longer hold when the model is further perturbed, for example by dropping assumption (14). Nevertheless, our results suggest that Hubbell's neutral model is just one representative of a much larger class of models that all yield similar ecological predictions. This could explain the success of neutral community theory in reproducing reasonable patterns, despite its unreasonable assumptions, as also argued by Pueyo et al. (2007).

Our results hinge on the validity of assumption (14), and this assumption therefore needs scrutiny. It states that immigration and local reproduction must depend on community size in the same way. This is not at all unreasonable if the density-dependence occurs in the establishment phase rather than the dispersal phase of recruitment, because in the establishment phase it no longer matters whether the propagule originated from a local individual or an immigrant; they both "feel" the presence of the community equally. In fact, one may argue that this is even required by the neutrality assumption, and therefore not an additional assumption. Put in mathematical terms: if $\beta(J_L) = \beta_1 s(J_L)$ and $\lambda(J_L) = \lambda_1 s(J_L)$ where β_1 and λ_1 are density-independent birth and immigration rates and $s(J_L)$ is a density-dependent survival probability, then $I = \lambda(J_L)/\beta(J_L) = \lambda_1/\beta_1$ is density-independent, as required by assumption (14).

Our results shine some new light on the debate whether different models of density-dependence can be distinguished using species abundance snapshot data. Although the debate concerning different models of *species-level* density-dependence (Chave et al., 2006; Volkov et al., 2006) remains unsettled as long as sampling formulas for such models are lacking, we find that a single snapshot of species abundance data cannot distinguish between different models of *community-level* density-dependence, that is, different mechanisms of interactions between neutral species. This may be interpreted as a lack of resolution of such data. However, incorporating more information (McGill et al., 2007), e.g. temporal and spatial (Etienne, 2007) variation in species abundance distributions or phylogenetic structure, may again enable distinguishing between these models.

Acknowledgements

We thank two anonymous reviewers and David Alonso for their constructive comments. R.S.E. acknowledges the financial support of the Netherlands Organisation for Scientific Research (NWO).

References

Alonso, D., Etienne, R.S., McKane, A.J., 2006. The merits of neutral theory. Trends Ecol. Evol. 21, 451–457.

Alonso, D., Ostling, A., Etienne, R.S., 2008. The implicit assumption of symmetry and the species abundance distribution. Ecol. Lett. 11, 93–105.

Chave, J., Alonso, D., Etienne, R.S., 2006. Comparing models of species abundance. Nature 414, E1–E2.

Etienne, R.S., 2007. A neutral sampling formula for multiple samples and an "exact" test of neutrality. Ecol. Lett. 10, 608–618.

Etienne, R.S., Alonso, D., 2005. A dispersal-limited sampling theory for species and alleles. Ecol. Lett. 8, 1147–1156 (Erratum in Ecol. Lett. 9, 500).

Etienne, R.S., Alonso, D., 2007. Neutral community theory: how stochasticity and dispersal-limitation can explain species coexistence. J. Stat. Phys. 128, 485–510.

Etienne, R.S., Olff, H., 2005. Confronting different models of community structure to species-abundance data: a Bayesian model comparison. Ecol. Lett. 8, 493–504.

Etienne, R.S., Alonso, D., McKane, A.J., 2007. The zero-sum assumption in neutral biodiversity theory. J. Theor. Biol. 248, 522–536.

Ewens, W.J., 1972. The sampling theory of selectively neutral alleles. Theor. Popul. Biol. 3, 87–112.

Hubbell, S.P., 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, Princeton, NJ.

Kendall, D.G., 1948. On some modes of population growth giving rise to R.A. Fisher's logarithmic series distribution. Biometrika 35, 6–15.

McGill, B.J., Maurer, B.A., Weisser, M.D., 2006. Empirical evaluation of neutral theory. Ecology 87, 1411–1423.

McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K., Dornelas, M., Enquist, B.J., Green, J., He, F., Hurlbert, A., Magurran, A.E., Marquet, P.A., Maurer, B.A., Ostling, A., Soykan, C.U., Ugland, K., White, K., 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecol. Lett. 10, 995–1015.

Pueyo, S., He, F., Zillio, T., 2007. The maximum entropy formalism and the idiosyncratic theory of biodiversity. Ecol. Lett. 10, 1017–1028.

Rannala, B., 1996. The sampling theory of neutral alleles in an island population of fluctuating size. Theor. Popul. Biol. 50, 91–104.

Van Kampen, N.G., 1992. Stochastic Processes in Physics and Chemistry. Elsevier, Amsterdam, The Netherlands.

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., He, F., Hubbell, S.P., Maritan, A., 2006. Comparing models of species abundance. Nature 414, E1–E2.