What drives community dynamics?
Crispin M. Mutshinda1,*, Robert B. O’Hara1 and Ian P. Woiwod2

1Department of Mathematics and Statistics, University of Helsinki, PO Box 68 (Gustaf Hällströmin katu 2b), 00014 Helsinki, Finland
2Plant and Invertebrate Ecology Department, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK

Received 30 March 2009
Accepted 21 April 2009

The search for general mechanisms of community assembly is a major focus of community ecology. The common practice so far has been to examine alternative assembly theories using dichotomist approaches of the form neutrality versus niche, or compensatory dynamics versus environmental forcing. In reality, all these mechanisms will be operating, albeit with different strengths. While there have been different approaches to community structure and dynamics, including neutrality and niche differentiation, less work has gone into separating out the temporal variation in species abundances into relative contributions from different components. Here we use a refined statistical machinery to decompose temporal fluctuations in species abundances into contributions from environmental stochasticity and inter-/intraspecific interactions, to see which ones dominate. We apply the methodology to community data from a range of taxa. Our results show that communities are largely driven by environmental fluctuations, and that member populations are, to different extents, regulated through intraspecific interactions, the effects of interspecific interactions remaining broadly minor. By decomposing the temporal variation in this way, we have been able to show directly what has been previously inferred indirectly: compensatory dynamics are in fact largely outweighed by environmental forcing, and the latter tends to synchronize the population dynamics.

Keywords: community drivers; compensatory dynamics; density dependence; environmental stochasticity; neutrality; niche

1. INTRODUCTION
All natural populations fluctuate in abundance due to environmental stochasticity, demographic stochasticity and density-dependent regulation (Lande et al. 2003; Wilson & Lundberg 2006). Some form of density dependence seems inevitable (Cooper 2003), but deciding how strongly density dependence affects a population’s dynamics can ultimately only be done with empirical studies. Species also exist as part of communities, so between-species interactions may be important as well (Bower 1962; Murdock 1994; Shorrocks & Sevenster 1995; Miyashita 2001; Viljugrein et al. 2005).

A major focus of community ecology is to identify the assembly rules that shape species assemblages, as well as any exogenous factors that the rules act with to create and maintain a community. The common practice so far has been to examine alternative assembly theories using dichotomist approaches of the form neutrality versus niche (e.g. Fargione et al. 2003) or compensatory dynamics versus environmental forcing (Houllahan et al. 2007). But in reality all these forces will be operating (e.g. Chave et al. 2002), albeit with different strengths.

While there have been different approaches to community structure and dynamics, encompassing neutral mechanisms (Caswell 1976; Bell 2000; Hubbell 2001) and niche differentiation (Gause 1934; Hutchinson 1957; Armstrong & McGehee 1980), less work has gone into separating out the temporal variation in species abundances into contributions from different factors.

Neutral models assume that all individuals of interacting species at a single trophic level are ecologically equivalent. In reality, ecological communities are shaped by several factors: individuals interact with conspecifics and with individuals of other species, and species respond to abiotic influences (e.g. Enquist et al. 2002). These factors act in concert with demographic stochasticity to generate population fluctuations in both time and space (Ylikarjula 2000), fluctuations that neutral theory is unable to replicate (Mutshinda et al. 2008). Consequently, a central problem for understanding community dynamics is to determine how strong (relatively) these different factors are. Here we decompose temporal fluctuations in species abundances into contributions from different factors—namely environmental stochasticity and intra-/interspecific interactions—to find which ones dominate the dynamics.

2. MATERIAL AND METHODS
(a) Description of data
We used community time series from a range of taxa: moths, fishes, macrocrustaceans, birds and rodents. For all of the datasets, the sampling methods were consistent over time and the data are recorded as counts of individuals. The moth data consist of yearly totals of light-trapping counts of 12 common noctuid moths from the Rothamsted Insect Survey in the UK (Woiwod & Harrington 1994). The data span a 40-year period (1964 to 2003) and come from the Geescroft I station, in a small piece of woodland on Rothamsted Farm in Hertfordshire, UK (Woiwod & Gould 2008). The fish and

* Author for correspondence (crispin.mutshindamwanza@helsinki.fi).

crustacean data both come from Hinkley Point in the Bristol Channel, UK and were collected between 1981 and 2001 (the methods are described further in Henderson & Holmes 1990). The data contain the nine and seven most common fish and crustacean species, respectively. The bird data consist of yearly abundances of five species at Hubbard Brook Experimental Forest, an area of land in central New Hampshire, USA that functions as an outdoor laboratory for ecological studies. The data cover a 16-year period, 1969–1984 (Holmes et al. 1986). The rodent data come from trapping webs throughout the Sevilla National Wildlife Refuge in central New Mexico, USA over the period 1989–2003. At each site, three sampling webs were sampled for three consecutive nights (Frigngens 2006). This replicated feature allows us to estimate the sampling variation.

Species that were not regularly observed (i.e. that only had abundances of zero or one individual) were left out of all of the datasets. It is also assumed that each yearly sample was collected using the same effort and method, so that the time series reflect real temporal change in abundance at one locality. Species’ scientific names, common names, taxa and sampling periods, as well as mean population densities and standard deviations, are given in table S1 in the electronic supplementary material.

(b) Model specification

The underlying model assumed for the abundance of species $i$ at time $t$, $N_{i,t}$, is a discrete-time stochastic Gompertz model (e.g. Saitoh et al. 1997; Ives et al. 2003; Jacobson et al. 2004; Dennis et al. 2006), including the effects of interactions with other species. That is,

$$N_{i,t} \sim N_{i,t-1} \exp \left\{ r_i \left[ 1 - \sum_{j=1}^{S} a_{ij} \log N_{j,t-1}/k_i + e_{ij} \right] \right\},$$

$t = 2, 3, \ldots, T,$

(2.1)

where $r_i$ and $k_i$ are respectively the intrinsic growth rate and the natural logarithm of the carrying capacity of species $i$; $a_{ij}$ is the coefficient of interaction between species $i$ and $j$, replicating the per capita effect of species $j$ on the growth of species $i$ from time $(t-1)$ to time $t$ with all intraspecific coefficients, $a_{ii}$, set to 1. Environmental noise enters the dynamics through the zero-mean Gaussian random shocks $e_{ij}$. If we denote by $n_{ij}$ the natural logarithm of $N_{i,t}$, then, on the natural logarithmic scale, equation (2.1) reads as follows:

$$n_{i,t} \mid n_{i,t-1} = n_{i,t-1} + r_i \left[ 1 - \sum_{j=1}^{S} a_{ij} n_{j,t-1}/k_i \right] + e_{ij}.$$

$t = 2, 3, \ldots, T.$

(2.2)

The community model can be compactly written in matrix form as

$$n_i \mid n_{i-1} = n_{i-1} + R(I_S - A_{i-1}) + e_i,$$

for $t = 2, 3, \ldots, T,$

(2.3)

where $n_i$ is the vector of log-abundances of all the $S$ species at time $t$, $R$ is an $S \times S$ diagonal matrix with intrinsic growth rates $r_i$ on the diagonal, $I_S$ is the $S$-dimensional vector with all elements equal to 1 and $A_{ij} = a_{ij}/k_i$. Thus, elements on the leading diagonal of $A$ are inverse carrying capacities, and off-diagonal terms are the interaction coefficients between species, scaled by the carrying capacity of the focal species (i.e. species $i$ for $A_{i-1}$). $e_i = [e_{i1} \ldots e_{iS}]$ is the vector of zero-mean environmental disturbances with one element by population, which are assumed to be serially independent and multivariate normally distributed, with a covariance matrix denoted by $C$ (e.g. Ives et al. 2003; Ripa & Ives 2003). The off-diagonal terms of the ‘environmental covariance matrix’ $C$ represent the common responses of different species to the environment. Subsequently, $n_{ij} = c_{ij}/(c_{ij}s_{ij})^{1/2}$ quantifies the correlation between the responses of species $i$ and $j$ to environmental fluctuations.

We would expect many off-diagonal terms of $A$ to be at or close to zero (i.e. there would be no interaction between the species), owing to the documented predominance of weak interspecific interactions (e.g. Kokkoris et al. 1999; Turnbull et al. 2005). Therefore we used stochastic search variable selection (SSVS; George & McCulloch 1993) to identify the set of non-zero interactions.

SSVS is based on letting regression coefficients to be in two states: either they are close to zero, and hence have no significant effect on the model, or they are estimated freely. This is done by modelling a binary indicator $g_{ij}$, for $i \neq j$, and assuming $g_{ij} \sim \text{Bernoulli}(p)$, with $g_{ij} = 1$ when species $j$ is included in the dynamics of species $i$ and $g_{ij} = 0$ otherwise. This controls the (prior) variance of the interaction coefficient $a_{ij}$: $a_{ij}$ is assumed to be normally distributed with mean 0 and variance $\tau_{ij}$, and we let $\tau_{ij} = (1 - g_{ij})x_1 + g_{ij}x_2$. The non-negative numbers $x_1$ and $x_2$ are selected so that the prior distribution of $a_{ij}$ is wide under inclusion of species $j$ in the dynamics of species $i$ and forced to be close to zero if species $j$ is excluded (i.e. $x_2$ should be large and $x_1$ small). If inclusion of species $j$ in the dynamics of species $i$ is not supported by the data, the prior with variance $\tau_{ij}$ will tend to be selected more often the prior with variance $x_2$, and vice versa. For the purposes of our analysis, we took $x_1 = 0.01$, $x_2 = 10$, and set the prior probability of a non-zero interspecies interaction (i.e. $p$) to 0.2. We used Bayes factors (BFs) to evaluate the amount of evidence in the data in favour of any particular interspecific interaction. Since $\text{PosteriorOdds} = \text{PriorOdds} \times \text{BF}$, the BF is the factor by which the data change the prior odds of including versus not including a particular interspecific interaction into posterior odds. A BF that is larger than 1 suggests that the event is more probable than previously thought, and vice versa.

It follows from equation (2.3) that the variance–covariance matrix, $V$, of temporal fluctuations in species abundances can be decomposed (e.g. Ripa & Ives 2003) as

$$V = BB^T + C,$$

(2.4)

where $B = (RA)$; that is, $b_{ij} = r_i/k_i$ and $b_{i,j \neq i} = r_{i,j} / k_i$.

Consequently, the temporal variances of individual species abundances (diagonal elements of $V$) can be decomposed into the contributions $\nu_{ij}^2 (r_i/k_i)^2(1/k_i)$ and $\nu_{ij}^2 (r_i/k_i)^2s_{ij}^2$ and $c_{ij}$ from intraspecific interactions, interspecific interactions and environmental forcing, respectively.

Population abundance time series are often fraught with sampling errors (e.g. Clark & Bjornstad 2004). In a state-space framework, $n_{i,t}$ is related to its observed counterpart, $x_{i,t}$, through a sampling model. Here we assume that

$$x_{i,t} \mid n_{i,t}, s_{i,t}^2 \sim N(n_{i,t}, s_{i,t}^2),$$

(2.5)

where $s_{i,t}$ is the standard error of $x_{i,t}$ as an estimate of $n_{i,t}$.

(c) Model fitting

The model was fitted to the data with a Bayesian approach (Gelman et al. 2003; McCarthy 2007) using Markov chain
Monte Carlo (MCMC) methods (Gillks et al. 1996) through the Bayesian freeware OpenBUGS (Thomas et al. 2006). We placed on $C$ an inverse Wishart prior $C^{-1} \sim \text{Wishart}(\omega, S)$ with scale matrix $\omega$ and the smallest possible number of degrees of freedom (i.e. $\text{rank}(C) = S$) as an attempt to impose non-informativeness. The elements of $W$ were set as $w_{i,i} = 1$ and $w_{i,j} = 0$. All intrinsic growth rates $r_i$ were assigned independent $N(0, 1) I(0, +\infty)$ priors, where $I(\cdot)$ denotes the indicator function. The log-carrying capacities, $k_i$, were assigned Uniform$(l_i, u_i)$ priors, with suitably selected hyper-parameters $l_i$ and $u_i$ to constrain $k_i$ to the range of the data for identifiability.

We used the replicated feature of the rodent dataset to derive estimates of sampling errors $s_i = s_j$ as 0.223 for Dipodomys ordii, 0.141 for Dipodomys spectabilis and 0.36 for Perognathus flavus. We used the model with and without sampling error to fit the rodent data, but the results were similar. Given the larger sample sizes and longer observational periods for the other datasets used here, we expected the amount of sampling variation to be similar or less. Hence, we fitted the ‘process-error-only’ version of the model to all these datasets.

For each dataset, we ran 60 000 iterations of three MCMC with a burn-in period of 15 000 iterations, and thinned the chains to every 25th observation. The results were broadly robust to changes in the range of hyper-parameters. We checked the model adequacy by comparing posterior predictive distributions of future data to the data that have actually occurred (Rubin 1984). For all datasets analysed, the data were consistent under the posterior predictive distributions, as exemplified by fig. S1 in the electronic supplementary material. Much of the remaining variation was due to intraspecific interactions, explaining between 11 and 54 per cent of variation in individual moth species, between 2 and 10 per cent in fish species, between 6 and 15 per cent in macrocrustaceans, between 1 and 17 per cent in birds, and between 11 and 37 per cent in rodents. Interspecific interactions were found to be broadly weak: they explained less than 10 per cent of the temporal variation, and the BF s for the individual interaction coefficients to be non-zero were generally between 0.002 and 1.25.

The environmental correlations were, on average, either zero or positive for most of the datasets examined (figures 4–6), with posterior means lying between 0.00 and 0.70. However, some negative correlations were detected in the bird and the rodent communities (see figs S4–S5 in the electronic supplementary material).

Incorporation of sampling variation into the model for the rodent data did not change inference, as sampling variation only accounted for less than 9 per cent of the total variance (fig. S3 in the electronic supplementary material).

3. RESULTS

In all of the datasets, the dynamics were largely dominated by environmental stochasticity, which accounted for between 40 and 95 per cent of the temporal variances in individual species’ abundances (figures 1–3; figs S2–S3 in the electronic supplementary material). Much of the remaining variation was due to intraspecific interactions, explaining between 11 and 54 per cent of variation in individual moth species, between 2 and 10 per cent in fish species, between 6 and 15 per cent in macrocrustaceans, between 1 and 17 per cent in birds, and between 11 and 37 per cent in rodents. Interspecific interactions were found to be broadly weak: they explained less than 10 per cent of the temporal variation, and the BF for the individual interaction coefficients to be non-zero were generally between 0.002 and 1.25.

The environmental correlations were, on average, either zero or positive for most of the datasets examined (figures 4–6), with posterior means lying between 0.00 and 0.70. However, some negative correlations were detected in the bird and the rodent communities (see figs S4–S5 in the electronic supplementary material).

Figure 1. Error bars (posterior mean±1 s.d.) for the proportions of temporal variation attributable to environmental forcing (black diamonds) and to intraspecific interactions (grey boxes) in the dynamics of individual moth species.

Figure 2. Error bars (posterior mean±1 s.d.) for the proportions of temporal variation attributable to environmental forcing (black diamonds) and to intraspecific interactions (grey boxes) in the dynamics of individual fish species.

Figure 3. Error bars (posterior mean±1 s.d.) for the proportions of temporal variation attributable to environmental forcing (black diamonds) and to intraspecific interactions (grey boxes) in the dynamics of individual crustacean species.
Figure 4. Posterior densities of environmental correlations for the moth data; darker shading corresponds to a correlation with a posterior mean further from 0.

Figure 5. Posterior densities of environmental correlations for the fish data; darker shading corresponds to a correlation with a posterior mean further from 0.
4. DISCUSSION

Our results show directly that environmental stochasticity is a major driver of biodiversity patterns, as previously only inferred from indirect analyses (e.g. Steele 1985; Lawton 1988; Dornelas et al. 2006; Angela et al. 2007; Houlahan et al. 2007).

Our results also support the suggestions of Angela et al. (2007), Houlahan et al. (2007) and Loreau & de Mazancourt (2008), among others, that environmental forcing broadly dominates compensatory dynamics in driving temporal fluctuations in species abundance, and that the pervasive positive correlations in species abundances are presumably environment-induced.

Ripa & Ives (2003) have looked at the dynamics of communities with correlated environments. When the interactions between species are negligible (i.e. \( a_{ij} \approx 0 \), as we found), it follows from equation (2.4) that \( v_{ij} = c_{ij} / (1 - b_{ij} b_{ij}) \), \( i,j = 1, \ldots, S \) and the correlation, \( \phi_{i,j} \), between the densities of species \( i \) and \( j \) is related to the environmental correlation, \( \rho_{i,j} \) through

\[
\phi_{i,j} = \theta_{i,j} \rho_{i,j} \tag{4.1}
\]

(eqn (5) in Ripa & Ives 2003), where

\[
\theta_{i,j} = \left( \frac{1 - b_{ij}^2}{1 - b_{ij}^2 + (b_{ij} - b_{ij})^2} \right)^{1/2},
\]

and \( b_{ij} \) is defined as in equation (2.4).

When species have similar dynamics (i.e. when \( b_{ij} \approx b_{ij} \)), the correlation between population densities is the same as the correlation in environmental noise, and is lower when the dynamics of the species are different. We found that the dynamics are fairly different, with the term \( \theta_{i,j} \) in equation (4.1) ranging between 0.70 and 1. Hence, even when there are no direct interactions between species, the correlation between population densities is not a good estimate of the environmental correlation. Houlahan et al. (2007) found that correlations between population densities tended to be positive. If our results can be extrapolated to their study (i.e. we can assume that species’ interactions are negligible, and variation between species’ dynamics is large), we can conclude that the correlations they report broadly underestimate the correlations in species’ responses to the environment.

Here we ascribed any unexplained variance as environmental. Clearly, a better approach would be to model the effect of the key environmental drivers (e.g. Ives 1995; Saether et al. 2000). However, this would require a larger effort, as the drivers are almost certainly different in the different communities, and indeed for different species in each community. Thus, we have concentrated on finding out how general our results are across different communities.

If the residual variance is not environmental variation, what could it be? For most of the datasets, it will include some sampling variation. However, we expect this to be relatively small for two reasons. Our direct evidence is that the results from the rodent dataset, where we could estimate the sampling variability, are consistent with the other results: the environmental variation still explains 57 to 87 per cent of the variation versus 0.1 to 9 per cent for sampling error (fig. S3 in the electronic supplementary material). Indirectly, we would expect the amount of sampling variation to be similar or less for other datasets we used, as their sample sizes are larger and the data span longer observation periods. However, even if the sampling variation was significantly larger, our conclusions would still hold.

Proc. R. Soc. B (2009)
The environmental variance may also be inflated by some interspecific interactions. Not all species in the community were included in the analysis, so there may be strong interactions with unseen species. Indeed, it would seem odd if there were no such interactions. In particular, species at different trophic levels (e.g., predators, parasitoids or host plants for the moths) would be expected to have an effect. Similarly, we might expect competition with other species that share a common host resource. The variation in the dynamics induced by these missing interactions will appear in our model as environmental fluctuations. Of course, it could also be argued that they genuinely are environmental effects, as they are part of the biotic environment of the community.

Our finding that environmental variation rather than interspecific competition dominates the dynamics in a diverse range of communities has important implications for the way we see the communities. The communities are not very stable, with the abundances able to bounce around as the environment buffets them. There is still some density dependence, which keeps the populations in check within some bounds, but these are generous.

We found little evidence for interspecific interactions affecting the temporal variation. This contrasts with studies using manipulative experiments (i.e., exclusions; e.g., Brown & Davidson 1997). This may partly be because manipulative experimental studies tend to overestimate the strength of interspecific interaction (but see Turnbull et al. 2005). It may also be because manipulative studies typically ask slightly different questions. They are usually carried out by excluding species and seeing the effect this has on the community (e.g., Heske et al. 1994; Brown & Davidson 1997). However, this is a test of the effect of interspecific competition on demarcating species’ niches (e.g., their geographical range limits). Our study examines species within their realized niches, where the strength of interspecific competition typically fades out, and its role in driving year-to-year fluctuations in species abundances remains minor. If we are to understand what drives the dynamics in natural systems, we need to use observations of natural systems.

Here we have been able to decompose the temporal variation in species abundances into contributions from environmental fluctuations and intra- and interspecific competition, making the evaluation of the relative strengths of each component possible. This requires both sophisticated statistical machinery and, of equal importance, high quality community data collected over long periods of time.

We wish to thank the associate editor, two anonymous reviewers and Rampal Etienne for their constructive comments on the manuscript. We are grateful to Pisces Conservation Ltd for making the Himley fish and crustacean data available, and to the Sevilleta Long Term Research Project for the rodent data. The bird data were obtained by scientists of the Hubbard Brook Ecosystem Study. Rothamsted Research is grant-aided by the UK Biotechnology and Biological Sciences Research Council. The Academy of Finland (project no. 205371) provided funding for this research to C.M.M. and R.B.O.

REFERENCES


Friggens, M. 2006 Sevilleta small mammal population data SEV008. See http://sevilleta.unm.edu/data/.


