Sensitivity analysis of Markov models for communities of competing sessile organisms

MATTHEW SPENCER

Department of Mathematics and Statistics and Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, B3H 3J5, Canada

Summary

1. Communities of competing sessile organisms are often modelled using Markov chains. Sensitivity analysis of the stationary distribution of these models tells us how we expect the abundance of each organism to respond to changes in interactions between species. This is important for conservation and management.

2. Markov models for such communities have usually been formulated in discrete time. Each column of the discrete-time transition matrix must sum to 1 (column stochasticity). Sensitivity analysis therefore involves defining a pattern of compensation that maintains column stochasticity as a single transition probability changes. There is little biological theory about the appropriate compensation pattern, but the usual choices involve changing only the elements of a single column of the transition matrix.

3. I argue that if the underlying dynamics occur in continuous time, each transition probability is the net outcome of direct and many indirect interactions.

4. Determining the consequences of changing a single direct interaction will often be of interest. I show how this can be achieved using a continuous-time model. The resulting discrete-time compensation pattern is quite different from those that have been considered elsewhere, with changes occurring in many columns.

5. I also show how to determine which direct interactions are being changed under any discrete-time compensation pattern.

Key-words: competition, interspecific interactions, marine communities, stationary distribution.

Introduction

Most models of interspecific interactions do not work very well for species-rich systems. Although the parameters of classical predator–prey models can be estimated from time-series data (Harrison 1995), we often misidentify the system even in simple cases (Jost & Arditi 2000). For complex communities, we are often unable to predict the qualitative effects of major perturbations. For example, we cannot give precise answers to apparently simple questions such as whether a seal cull will increase the yield of a fishery (Yodzis 1998).

One exception is the case of multispecies Markov models. These have been applied to communities of sessile organisms such as trees, corals, mussels and sponges (e.g. Usher 1979; Tanner *et al*. 1994; Wootton 2001a;

Hill *et al*. 2004). Their parameters are estimated easily from field data, and they make surprisingly good predictions of the effects of species removals (Wootton 2001b). Markov models for communities of competing sessile organisms assume typically that the state of the community at a given point in space is defined by the single species (or empty space) present at that point, that the set of possible states is fixed and finite, and that the future state at a point depends only on the present state, not on states at any previous times.

These models have almost always been formulated in discrete time. Their parameters are the conditional probabilities of observing a species *i* at a point in space at time $T + t$, given that we saw species *j* at that point at time *T*. These parameters, known as transition probabilities, cannot usually be interpreted as probabilities that species *i* can directly replace species *j*, because they include the effects of indirect transitions. Unless it is physically impossible for more than one transition to occur in a unit of time, observing *j* at time *T* and *i* at

Correspondence: Matthew Spencer, School of Biological Sciences, University of Liverpool, Crown Street, Liverpool, L69 72B, UK. E-mail: m.spencer@liverpool.ac.uk

time $T + 1$ does not rule out the possibility that k first replaced j, and was in turn replaced by i (Wootton 2001a). This makes it difficult to predict how the system will respond to changes in the direct interactions between species. For example, suppose we increase the rate of disturbances that remove living organisms, leaving empty space. The probabilities of transitions to empty space from all species will probably increase, but probabilities of transitions from species i to species j via empty space are also likely to increase. Because the discrete-time model does not separate direct and indirect transitions, it does not tell us how to deal with this situation. Questions of this kind, known as sensitivity analyses, are important to management and conservation, but we do not know how to answer them using discrete-time Markov models. If one species may replace another at any time, formulating the model in continuous time can solve this problem. The parameters of a continuous-time Markov model are the instantaneous rates of direct replacement of one species by another. Unlike the discrete-time transition probabilities, these rates can be interpreted as the intensities of direct interactions between species. Using continuous-time models, we can predict how an increase in the rate of one kind of transition will affect the transition probabilities observed over a given time interval, and thus determine the consequences of changing only one direct interaction rate. Continuous-time Markov chains are rarely used in community ecology (two exceptions are Cohen & Singer 1979; Singer & Cohen 1980). In this paper, I first describe the usual discrete-time models and list some suggestions that have been made for doing sensitivity analyses of these models. Next, I introduce continuous-time models, describe some aspects of parameter estimation and develop sensitivity analyses in continuous time, showing how we can uniquely answer questions about the effects of a change in one direct interaction rate. I show how changing a single direct interaction rate affects the discrete-time transition matrix, and how we can calculate the pattern of changes in direct interaction rates corresponding to a given change in the discrete-time transition matrix. Different kinds of sensitivity analysis involving a change in some focal discrete-time transition probability are asking different biological questions. I show that they can have qualitatively different answers. This is important, because such methods have been used in a number of major long-term ecological studies (Tanner *et al.* 1994; Wootton 2001b; Hill *et al.* 2004). Finally, I consider some possibilities for the case where interaction rates vary over time.

Discrete-time Markov chains

In discrete time, a Markov chain with s possible states is defined by an $s \times s$ transition matrix $\mathbf{P}(t)$, whose entries $p_{ij}(t)$ are the conditional probabilities of observing state i at a point at time $T + t$, given that the point is in state j at time T [this convention is common in ecology,

but in many other fields the conditional probability of i given j would be $p_{ji}(t)$]. The $\mathbf{P}(t)$ matrix is therefore column-stochastic (each column sums to 1). Let $\mathbf{x}(T)$ be a column vector whose elements $x_1(T), x_2(T), \ldots, x_s(T)$ are the probabilities of each state at some time T. Then $\mathbf{x}(T)$ and $\mathbf{x}(T + t)$ are related by:

$$\mathbf{x}(T + t) = \mathbf{P}(t)\mathbf{x}(T) \qquad \text{eqn 1}$$

Here, in common with most other studies, I will focus on homogeneous, linear Markov chains. Homogeneous means that the conditional probabilities of states at time $T + t$ given the state at time T do not vary with T. Linear means that the conditional probabilities do not depend on the frequencies of states. Such models ignore important features of the real world such as density dependence and temporal variability in environmental conditions, but can always be thought of as local approximations (cf. Caswell 2001: 29–31).

Sensitivity of the stationary distribution in a discrete-time model

An important aspect of the study of a model is determining the consequences of changing the parameter values. For example, most communities modelled by eqn 1 asymptotically approach a stationary probability distribution of states, given by the dominant right eigenvector $\mathbf{w}_1$ of $\mathbf{P}(t)$, normalized to sum to 1 (Hill *et al.* 2004). I assume that the Markov chain is irreducible (any state can be reached from any other, although not necessarily in a single transition) and has a unique stationary distribution. It will often be interesting to know how this stationary distribution will change with the transition probabilities $p_{ij}(t)$. Examining the derivatives (sensitivities) of the stationary distribution with respect to the $p_{ij}(t)$ is one way to do this. Applications include finding management strategies that maximize the amount of space filled by a native species, or minimize the amount filled by an invader. Similar ideas can be applied to other community properties (Hill *et al.* 2004). Because the $\mathbf{P}(t)$ matrix is column-stochastic, an increase in $p_{ij}(t)$ must be compensated by a decrease in one or more other elements in column j. Let the compensation pattern $\mathbf{D}_\mathbf{v}[\mathbf{P}(t)]^{(ij)}$ be a matrix whose (m, n)th element $\mathbf{D}_\mathbf{v}[\mathbf{P}(t)]_{mn}^{(ij)}$ is the partial derivative of $p_{mn}(t)$ with respect to $p_{ij}(t)$. Then the sensitivity of the stationary distribution with respect to $p_{ij}(t)$ is:

$$\frac{d\mathbf{w}_1}{dp_{ij}(t)} = \sum_{m,n} \mathbf{D}_\mathbf{v}[\mathbf{P}(t)]_{mn}^{(ij)} \frac{\delta \mathbf{w}_1}{\delta p_{mn}(t)} \qquad \text{eqn 2}$$

(cf. Caswell 2001: 253, eqn 9·139). Caswell (2001: 253) suggests several possibilities for $\mathbf{D}_\mathbf{v}[\mathbf{P}(t)]^{(ij)}$, including:
• Specific compensation. $\mathbf{D}_\mathbf{v}[\mathbf{P}(t)]_{ij}^{(ij)}$ is 1. One element in column j is −1. I will consider the obvious choice, element (j, j). This means that increasing one of the probabilities of leaving a state is compensated by decreasing the probability of staying in that state. All other columns have only zeros.

• Uniform compensation. $\mathbf{D}_v[\mathbf{P}(t)]_{ij}^{(ij)}$ is 1. All other elements of column j are $-1/(s-1)$. All other columns have only zeros.

• Proportional compensation. $\mathbf{D}_v[\mathbf{P}(t)]_{ij}^{(ij)}$ is 1. All other elements (m, j) of column j are $-p_{mj}/(1-p_{ij})$. All other columns have only zeros.

It is not usually obvious which pattern to choose. One approach is to try several choices from the list above and hope they give similar results (Hill *et al.* 2004). In the following sections, I develop the alternative continuous-time formulation, and show that changing a single instantaneous rate can affect every column of $\mathbf{P}(t)$. This is because the $p_{ij}(t)$ are composites of direct and indirect transitions, rather than independent parameters. I also show how to determine the changes in direct interactions corresponding to a given discrete-time compensation pattern.

Continuous-time Markov chains

It is often reasonable to assume that transitions can occur at any time. For example, if organisms have long breeding seasons, colonization may occur during a large part of each year (Connell 1975: 477). Disturbances such as storms could occur at any time, although they might be more frequent at some times of year than others. Overgrowth of one sessile organism by another is likely to be possible at any time. Finally, if the time step is longer than a year, more than one event could happen per time step even if events are limited to particular seasons. If transitions may occur at any time, then the homogeneous continuous-time analogue of eqn 1 is the system of differential equations:

$$\mathbf{x}' = \mathbf{Q}\mathbf{x} \qquad \text{eqn 3}$$

where $\mathbf{x}'$ is a vector of first derivatives with respect to time and $\mathbf{Q}$ is a matrix of instantaneous rates. For a scalar differential equation $x' = ax$, the solution is $x(t) = e^{at}x(0)$, where the scalar exponential is defined by the series:

$$e^{at} = \sum_{m=0}^{\infty} \frac{(at)^m}{m!} \qquad \text{eqn 4}$$

Similarly, it can be shown that the solution of eqn 3 is:

$$\mathbf{x}(t) = e^{\mathbf{Q}t}\mathbf{x}(0) \qquad \text{eqn 5}$$

where the matrix exponential is:

$$e^{\mathbf{Q}t} = \sum_{m=0}^{\infty} \frac{(\mathbf{Q}t)^m}{m!} \qquad \text{eqn 6}$$

(e.g. Norris 1997: 62–63). Comparing eqns 1 and 5 suggests that $e^{\mathbf{Q}t}$ must be the transition matrix $\mathbf{P}(t)$. We can see from eqn 6 that for small t, $\mathbf{P}(t) \approx \mathbf{I} + \mathbf{Q}t$. The $p_{ij}(t)$ are conditional probabilities, which must be real and between 0 and 1. Thus if any of the off-diagonal entries q_{ij} were complex or negative, there would also be complex or negative $p_{ij}(t)$ for small t. Furthermore, any transition out of a state j must be into some other state

i, so the rate of leaving a state j is the sum of the rates from j to each other state i:

$$
\begin{aligned}
\frac{dx_j}{dt} &= q_{j1}x_1 + q_{j2}x_2 + \ldots - (q_{1j} + \ldots + q_{j-1,j} \\
&\quad + q_{j+1,j} + \ldots + q_{sj})x_j + \ldots + q_{js}x_s \\
&= q_{j1}x_1 + q_{j2}x_2 + \ldots - \sum_{m \neq j} q_{mj}x_j + \ldots + q_{js}x_s
\end{aligned} \qquad \text{eqn 7}
$$

Thus, to account for the fate of every point leaving state j, the diagonal entries of $\mathbf{Q}$ must be:

$$q_{jj} = -\sum_{m \neq j} q_{mj} \qquad \text{eqn 8}$$

We therefore define Q as the set of valid $\mathbf{Q}$ matrices, having real and non-negative off-diagonals, and diagonal entries given by eqn 8. Let $\mathbf{W}$ be a matrix whose columns are the right eigenvectors of $\mathbf{Q}$, and let $\mathbf{\Lambda}$ be a diagonal matrix of the corresponding eigenvalues. Then $\mathbf{Q}$ can almost always be diagonalized as:

$$\mathbf{Q} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^{-1} \qquad \text{eqn 9}$$

The eigenvalues of $\mathbf{P}(t)$ are then the scalar exponentials of the eigenvalues of $\mathbf{Q}$. $\mathbf{P}(t)$ and $\mathbf{Q}$ have the same eigenvectors.

Estimating the continuous-time model

In most cases, we do not have observations in continuous time. Instead, we know the number n_{ij} of observations of state j at time T for which we observed state i at time $T + t$. The maximum likelihood estimate of $p_{ij}(t)$ is then:

$$\hat{p}_{ij}(t) = \frac{n_{ij}}{n_{.j}} \qquad \text{eqn 10}$$

where $n_{.j}$ is the total number of transitions observed out of state j (Caswell 2001: 135; Bladt & Sørensen 2005). Estimating a continuous-time Markov model given the discrete-time transition matrix is known as the embeddability problem. Equation 6 suggests that we could obtain an estimate $\hat{\mathbf{Q}}$ using the matrix logarithm:

$$\hat{\mathbf{Q}} = \frac{1}{t} \log \hat{\mathbf{P}}(t) \qquad \text{eqn 11}$$

If $\hat{\mathbf{Q}} \in Q$, this is the maximum likelihood estimator (Singer & Cohen 1980; Bladt & Sørensen 2005). In practice, things are more complicated. First, there are some discrete-time transition matrices which cannot be embedded in a homogeneous continuous-time Markov chain. Secondly, there are some discrete-time transition matrices which can be embedded in more than one homogeneous continuous-time Markov chain. Singer & Spilerman (1976; their Table 1) summarize the conditions under which a transition matrix may be uniquely embeddable in a homogeneous process. In many cases, we find that the maximum likelihood matrix $\hat{\mathbf{P}}(t)$ is not embeddable in a homogeneous continuous-time Markov chain, because $1/t \log \hat{\mathbf{P}}(t)$ has negative or complex off-diagonals. Two possible explanations are that the true transition matrix $\mathbf{P}(t)$ is not generated by a homogeneous

process, or that the true transition matrix $\mathbf{P}(t)$ is embeddable, but $\hat{\mathbf{P}}(t)$ is not embeddable because of sampling variability. If the latter explanation is true, we should be able to find a $\hat{\mathbf{Q}} \in Q$ for which $e^{\hat{\mathbf{Q}}t} \approx \hat{\mathbf{P}}(t)$. A good approach is to find the maximum likelihood $\hat{\mathbf{Q}} \in Q$ (Kalbfleisch & Lawless 1985). Elsewhere, we describe one way of performing this for ecological data, together with a likelihood ratio test of the hypothesis that the true $\mathbf{P}(t)$ is embeddable in a homogeneous continuous-time Markov process (Spencer & Susko 2005). The problem of uniqueness in the presence of sampling variability is less easily solved. Possible approaches include Bayesian analyses (Geweke *et al*. 1986; Bladt & Sørensen 2005), and determining whether there are any $\mathbf{P}$ matrices reasonably close to the estimated transition probability matrix, but with quite different logs (Singer & Spilerman 1976). However, it has been suggested that multiple solutions to eqn 11 are only likely when the time step t is very long (Kalbfleisch & Lawless 1985).

Interpretation of sensitivity analyses from a continuous-time model

Once we have a continuous-time model, we can determine how the system responds to changes in each of the direct interaction rates. Unlike the discrete-time case, we can find a unique answer to the question 'what will happen if just one direct interaction rate is changed?'. By definition (eqns 7 and 8), a change in an element q_{ij} alone will result in an equal and opposite change in q_{jj}. This is the unique compensation pattern that corresponds to changing only one direct interaction rate in a continuous-time model. To see why this is so, consider the equations for three states i, j and k out of a system of s states:

$$\frac{dx_i}{dt} = q_{i1}x_1 + \ldots - (q_{1i} + \ldots + q_{si})x_i$$
$$+ \ldots + q_{ij}x_j + \ldots + q_{is}x_s$$
$$\frac{dx_j}{dt} = q_{j1}x_1 + \ldots - (q_{1j} + \ldots + q_{ij} + \ldots \quad \text{eqn 12}$$
$$+ q_{kj} + \ldots + q_{sj})x_j + \ldots + q_{js}x_s$$
$$\frac{dx_k}{dt} = q_{k1}x_1 + \ldots + q_{kj}x_j + \ldots$$
$$- (q_{1k} + \ldots + q_{sk})x_k + \ldots + q_{ks}x_s$$

Note that the terms in brackets are the rates of leaving each state, as defined in eqn 8. Now consider increasing a single direct interaction rate q_{ij} by a tiny amount δ. This is what we do when we look at the sensitivity of the system to a change in q_{ij} alone:

$$\frac{dx_i}{dt} = q_{i1}x_1 + \ldots - (q_{1i} + \ldots + q_{si})x_i + \ldots$$
$$+ (q_{ij} + \delta)x_j + \ldots + q_{is}x_s$$
$$\frac{dx_j}{dt} = q_{j1}x_1 + \ldots - (q_{1j} + \ldots + (q_{ij} + \delta) + \ldots \quad \text{eqn 13}$$
$$+ q_{kj} + \ldots + q_{sj})x_j + \ldots + q_{js}x_s$$
$$\frac{dx_k}{dt} = q_{k1}x_1 + \ldots + q_{kj}x_j + \ldots$$
$$- (q_{1k} + \ldots + q_{sk})x_k + \ldots + q_{ks}x_s$$

We have increased the rate at which we go from state j to state i and therefore, by definition, if nothing else has changed we have increased the rate at which we leave state j. Because q_{jj} is -1 times the rate of leaving state j, we must decrease q_{jj} by the same amount as the increase in q_{ij}. At first sight this appears no different from the arbitrary compensation patterns for discrete-time models discussed in the section 'Sensitivity of the stationary distribution in a discrete-time model'. It appears that we could consider any pattern of changes in rates $q_{\cdot j}$ out of state j, provided we ensure that each column of $\mathbf{Q}$ sums to zero. However, such changes address a different question: 'what will happen if we simultaneously change several rates of transition from state j to other states?'. As an example, consider what will happen if we increase q_{ij} by a tiny amount δ, and simultaneously decrease q_{kj} by the same amount:

$$\frac{dx_i}{dt} = q_{i1}x_1 + \ldots - (q_{1i} + \ldots + q_{si})x_i + \ldots$$
$$+ (q_{ij} + \delta)x_j + \ldots + q_{is}x_s$$
$$\frac{dx_j}{dt} = q_{j1}x_1 + \ldots - (q_{1j} + \ldots + (q_{ij} + \delta) + \ldots \quad \text{eqn 14}$$
$$+ (q_{kj} - \delta) + \ldots + q_{sj})x_j + \ldots + q_{js}x_s$$
$$\frac{dx_k}{dt} = q_{k1}x_1 + \ldots + (q_{kj} - \delta)x_j + \ldots$$
$$- (q_{1k} + \ldots + q_{sk})x_k + \ldots + q_{ks}x_s$$

The rate of leaving state j has not changed, because the increase in q_{ij} is balanced by the decrease in q_{kj}. However, we have changed two parameters, not one. We might sometimes be interested in sensitivity analyses that change multiple interaction rates, if each parameter was a function of some environmental variable. However, we would not expect the rates of leaving each state to remain constant.

The Markov models discussed here assume that if the rate of leaving some state j decreases because a parameter q_{ij} is reduced, the extra individuals remaining in j may be affected by other ecological processes, but the rates of these processes do not change. It might be more realistic to allow rates to be affected by abundances. For example, $q_{ij} = a_{ij}x_i$ leads to a Lotka–Volterra model with a per-capita interaction rate a_{ij}. We cannot use the techniques described here to analyse such models, but it would be interesting to see whether they make better predictions of transition probabilities in field data. If we can specify a functional form for each q_{ij}, it will lead to a clearly defined (but different) sensitivity analysis of the effects of changing each parameter.

Sensitivity of the stationary distribution in a continuous-time model

I now calculate the sensitivity of the stationary distribution to a change in only one direct interaction rate. Caswell (2001: 250) showed that the partial derivative of $\mathbf{w}_1$ with respect to element a_{ij} of some matrix $\mathbf{A}$ is:

$$\frac{\delta}{\delta a_{ij}}\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = \frac{\delta\mathbf{w}_1}{\delta a_{ij}} - \mathbf{w}_1\sum_{m=1}^{s}\frac{\delta w_m^{(1)}}{\delta a_{ij}} \quad \text{eqn 15}$$

where $w_m^{(1)}$ is the mth element of $\mathbf{w}_1$, the sum is over all s states, and:

$$\frac{\delta \mathbf{w}_1}{\delta a_{ij}} - w_j^{(1)} \sum_{m=2}^s \frac{\bar{v}_i^{(m)}}{\lambda_1 - \lambda_m} \mathbf{w}_m \qquad \text{eqn 16}$$

with $\bar{v}_i^{(m)}$ being the complex conjugate of the ith element of the mth left eigenvector of $\mathbf{Q}$. For a continuous-time Markov chain, the dominant eigenvalue λ_1 is zero. This means that the sum of $\mathbf{w}_1$ is constant, so the second term on the right-hand side of eqn 15 is always zero, and:

$$\frac{\delta}{\delta q_{ij}} \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = \frac{\delta \mathbf{w}_1}{\delta q_{ij}} = -w_j^{(1)} \sum_{m=2}^s \frac{\bar{v}_i^{(m)}}{\lambda_m} \mathbf{w}_m \qquad \text{eqn 17}$$

As discussed above, a change in an element q_{ij} will be compensated by an equal and opposite change in q_{jj}. To answer questions about the response of the system to changes in a single direct interaction rate, we require the total derivative:

$$\frac{d\mathbf{w}_1}{dq_{ij}} = \frac{\delta \mathbf{w}_1}{\delta q_{ij}} + \frac{\delta \mathbf{w}_1}{\delta q_{jj}} \frac{\delta q_{jj}}{\delta q_{ij}} = -w_j^{(1)} \sum_{m=2}^s \frac{\bar{v}_i^{(m)} - \bar{v}_j^{(m)}}{\lambda_m} \mathbf{w}_m \qquad \text{eqn 18}$$

Compensation in the corresponding discrete-time transition matrix

I now calculate the derivative of $\mathbf{P}(t)$ with respect to q_{ij}, assuming $\mathbf{Q}$ is diagonalizable. This leads to the conclusion that compensation patterns other than the single-column ones discussed in the section 'Sensitivity of the stationary distribution in a discrete-time model' are relevant to some biological questions, including the important class of questions about the effects of changing a single direct interaction.

DERIVATIVES OF P(T) WITH RESPECT TO q_{ij}

For a function $f(\mathbf{y})$ of a vector $\mathbf{y}$, the directional derivative $\mathbf{D}_\mathbf{V}[f(\mathbf{y})]$ in direction $\mathbf{V}$ is:

$$\mathbf{D}_\mathbf{V}[f(\mathbf{y})] \equiv \nabla f \mathbf{V}$$
$$= \lim_{h \to 0} \frac{f(\mathbf{y} + h\mathbf{V}) - f(\mathbf{y})}{h} \qquad \text{eqn 19}$$
$$= \sum_i v_i \frac{\delta f}{\delta y_i}$$

where v_i and y_i are the ith elements of $\mathbf{V}$ and $\mathbf{y}$, respectively. The direction vector $\mathbf{V}$ gives the relative size of the change in each element of $\mathbf{y}$. Here, we apply a similar idea to a function of a matrix.

If $\mathbf{Q}$ can be diagonalized as in eqn 9 (which will almost always be the case), the first directional derivative of $\mathbf{P}(t)$ in direction $\mathbf{V}$ (where $\mathbf{V}$ is a matrix the same size as $\mathbf{Q}$) is:

$$\mathbf{D}_\mathbf{V}(\mathbf{Q}) = \mathbf{W}[\mathbf{B} \odot \Phi(t)]\mathbf{W}^{-1} \qquad \text{eqn 20}$$

(Najfeld & Havel 1995; theorem 4·5). $\odot$ is the Hadamard (elementwise) product, $\mathbf{B} = \mathbf{W}^{-1}\mathbf{V}\mathbf{W}$ and $\Phi(t)$ has entries:

$$\phi_{ij}(t) = \begin{cases} t e^{\lambda_i t}, & \lambda_i = \lambda_j \\ (e^{\lambda_i t} - e^{\lambda_j t})/(\lambda_i - \lambda_j), & \text{otherwise} \end{cases} \qquad \text{eqn 21}$$

$\mathbf{V}$ is called a direction by analogy with the vector case (eqn 19): v_{ij} gives the relative size of the change in q_{ij}. In most applications of eqn 19, the direction $\mathbf{V}$ is normalized to have unit length. However, if we are interested in changing one direct interaction, we want a unit change in a single q_{ij} and the corresponding change in the diagonal element. Thus $\mathbf{V}$ has zeros everywhere except for $v_{ij} = 1$ and $v_{jj} = -1$. Using the chain rule, the total derivative of the stationary distribution with respect to q_{ij} is then:

$$\frac{d\mathbf{w}_1}{dq_{ij}} = \sum_{m,n} \mathbf{D}_\mathbf{V}(\mathbf{Q})_{mn}^{(ij)} \frac{\delta \mathbf{w}_1}{\delta p_{mn}(t)} \qquad \text{eqn 22}$$

where m and n are summed over all states s, $\mathbf{D}_\mathbf{V}(\mathbf{Q})_{mn}^{(ij)}$ is the (m,n)th element of the directional derivative of $\mathbf{P}(t)$ with respect to q_{ij}, and $p_{mn}(t)$ is element m, n of the $\mathbf{P}(t)$ matrix. This is identical to eqn 18. Comparing eqns 22 and 2 shows that $\mathbf{D}_\mathbf{V}(\mathbf{Q})^{(ij)}$ is the compensation pattern for $\mathbf{P}(t)$ resulting from a change in a single direct interaction rate q_{ij}. In general, there are non-zero entries in all columns of $\mathbf{D}_\mathbf{V}(\mathbf{Q})^{(ij)}$.

CHANGES IN DIRECT INTERACTION RATES CORRESPONDING TO A GIVEN CHANGE IN $\mathbf{P}(t)$

Equation 20 gives the change in $\mathbf{P}(t)$ corresponding to a given change in $\mathbf{Q}$. We can also invert this equation to determine what pattern of changes $\mathbf{V}_c$ in direct interaction rates would result from some specified pattern of changes $\mathbf{D}_\mathbf{V}(\mathbf{P})$ in the transition probability matrix. We can rearrange eqn 20 as follows:

$$\mathbf{D}_\mathbf{V}(\mathbf{P}) = \mathbf{W}[\mathbf{B} \odot \Phi(t)]\mathbf{W}^{-1}$$
$$\mathbf{W}^{-1}\mathbf{D}_\mathbf{V}(\mathbf{P})\mathbf{W} = \mathbf{B} \odot \Phi(t) \qquad \text{eqn 23}$$
$$[\mathbf{W}^{-1}\mathbf{D}_\mathbf{V}(\mathbf{P})\mathbf{W}] \odot^{-1} \Phi(t) = \mathbf{B} = \mathbf{W}^{-1}\mathbf{V}_c\mathbf{W}$$
$$\mathbf{V}_c = \mathbf{W}[[\mathbf{W}^{-1}\mathbf{D}_\mathbf{V}(\mathbf{P})\mathbf{W}] \odot^{-1} \Phi(t)]\mathbf{W}^{-1}$$

where $\odot^{-1}$ denotes elementwise division. If $\mathbf{Q}$ can be diagonalized as in eqn 9 and has distinct eigenvalues, then $\mathbf{V}_c$ exists and is unique, because all the elements of $\Phi(t)$ are positive. This means that in such cases, there is only one set of changes in direct interaction rates corresponding to a specified change in the elements of $\mathbf{P}(t)$. This is useful because it can help us to understand what biological question is being addressed by a sensitivity analysis involving a given compensation pattern.

A biological example

Table 1 shows the $\mathbf{Q}$ matrix for a coral reef community (Tanner *et al.* 1994; Exposed Crest site), estimated by the maximum likelihood method described in Spencer & Susko (2005). We were unable to reject the hypothesis of homogeneity in continuous time (parametric bootstrap test, $P = 0·072$, reported as $P = 0·096$ in Spencer & Susko (2005), but this was with an incorrect way of sampling missing data), and the transition probabilities

© 2006 The Author.
Journal compilation
© 2006 British
Ecological Society,
*Journal of Animal
Ecology*, **75**,
1024–1033

Table 1. The estimated instantaneous rate matrix **Q** (Spencer & Susko 2005) for the coral reef community at the Exposed Crest site in Tanner *et al.* (1994; their Table 2)[a]

−0·482	0	0·014	0·00929	0	0·117	0·0563	0·0601	0·0278
0·0121	−0·574	0·00342	0·0255	0·0231	0·0125	0·0269	0·0305	0·0102
0·0443	0	−0·555	0·0148	0·0784	0·139	0·15	0·059	0·0634
0·0245	0	0·00704	−0·767	0·0727	0·235	0·0342	0·0833	0·0369
0·000669	0·0131	0	0	−1·3	0	0	0·00291	0·0071
0·0154	0·0255	0·0194	0·0679	0·0295	−2·04	0·0302	0·0302	0·0256
0	0	0·00501	0	0	0·0343	−0·861	0·00681	0·00916
0·00828	0	0·00778	0·0203	0·0583	0·00557	0·0206	−1·04	0·0216
0·376	0·535	0·499	0·629	1·04	1·5	0·543	0·768	−0·202

[a]The states are ordered encrusting acroporid corals, tabular Acropora, bushy *Acropora*, staghorn *Acropora*, soft corals, algae, massive corals, pocilloporid corals, free space.
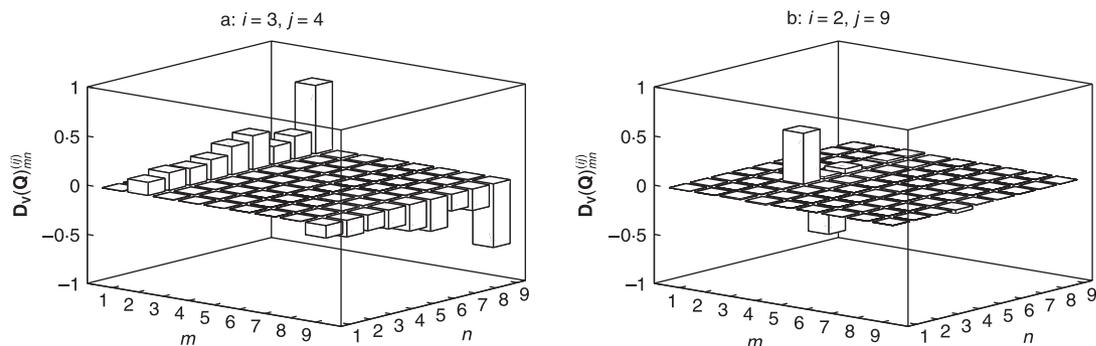
Fig. 1. Directional derivatives $\mathbf{D_V}(\mathbf{Q})^{(ij)}$ (eqn 20) of the one-time-unit transition matrix $e^\mathbf{Q}$ with respect to the elements (a) q_{34} and (b) q_{29} of the **Q** matrix shown in Table 1. Matrix elements are ordered as in Table 1.

from the best homogeneous model were very close to the discrete-time maximum likelihood estimates. This does not mean that the true situation is homogeneous, but we can hope that it is close enough to homogeneity that we may learn something useful about the system from a homogeneous model. I define one unit of time as the mean time interval between samples (26 months), so that the Tanner *et al.* (1994) **P** matrix is **P**(1). I report comparisons between sensitivity of the continuous-time matrix **Q** and the one-time-unit discrete-time matrix $e^\mathbf{Q}$. The results for comparing sensitivity of **Q** and the original matrix **P**(1) were almost indistinguishable, because $e^\mathbf{Q}$ and **P**(1) are very similar. Matlab code for these analyses is available at http://www.mathstat.dal.ca/~matts/. For some entries (*i, j*), the directional derivative $\mathbf{D_V}(\mathbf{Q})^{(ij)}$ of $e^\mathbf{Q}$ (eqn 20) shows changes that are mainly confined to column *j*, and can therefore be approximated by one of the compensation patterns Caswell (2001) suggests. However, this is not always the case. The derivative of $e^\mathbf{Q}$ with respect to element $q_{2,9}$ has large positive values in row 2 and large negative values in row 9, for all columns (Fig. 1b). State 2 is tabular *Acropora* and state 9 is free space.

Increasing the rate of transitions from free space to tabular *Acropora* increases the probability of reaching tabular *Acropora* from any other state in one unit of time. This is because the transition to free space has the largest rate from every non-empty state (row 9 of Table 1). Thus, the two-state path from any other state to tabular

Acropora has a probability that is strongly dependent on the rate from free space to tabular *Acropora*. Figure 2 shows the corresponding sensitivities of the stationary distribution. In both cases, sensitivities calculated using the continuous-time matrix (eqn 18) generally show similar qualitative patterns to sensitivities calculated using the discrete-time matrix (eqn 2) and any of the three columnwise compensation patterns discussed above. For example, under all compensation patterns, increasing the transition rate or probability from state 4 (staghorn *Acropora*) to state 3 (bushy *Acropora*) has the largest (and positive) effect on the stationary probability of bushy *Acropora* (Fig. 2a). Increasing the transition rate or probability from state 9 (free space) to state 2 (tabular *Acropora*) has the largest (and positive) effect on the stationary probability of tabular *Acropora* (Fig. 2b). Nevertheless, there are some important differences. For example, under uniform or proportional compensation in discrete time, the second-largest effect of increasing the rate or probability of transitions from staghorn to bushy *Acropora* is a decrease in the probability of free space. Under continuous-time or specific compensation in discrete time, the second-largest effect is a decrease in the probability of staghorn *Acropora*. This is because the uniform and proportional compensation patterns have relatively large decreases in the probability of transitions from staghorn *Acropora* to free space. The specific compensation pattern reduces only the probability that staghorn *Acropora* persists, and this is close to the
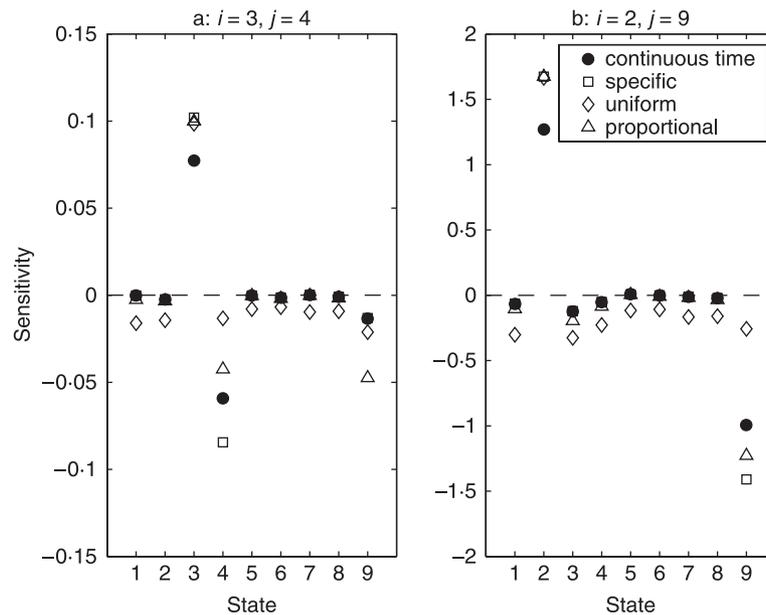
Fig. 2. Derivatives (sensitivities) of the stationary distribution for the community in Table 1 with respect to elements (a) 3,4; (b) 2,9. Filled circles are derivatives with respect to an element of the continuous-time rate matrix **Q** (eqn 18). Squares are the derivatives with respect to elements of the one time-unit transition matrix e^Q (eqn 2) and specific compensation [altering element (j, j) to compensate for the change in element (i, j)]. Diamonds are the same but with uniform compensation [altering all elements (m, j); $m \neq i$ in column j equally]. Triangles are proportional compensation [altering all elements (m, j); $m \neq i$ in column j in proportion to their magnitudes]. The horizontal axis indexes states in the stationary distribution (ordered as in Table 1). Panels (a) and (b) are drawn to different vertical scales.

compensation pattern resulting from the continuous-time model. For both parameter sensitivities that I examined ($i = 3, j = 4$ and $i = 2, j = 9$), the largest changes are an increase in the stationary probability of state i and a decrease in the stationary probability of state j. For states i and j, specific compensation in element (j, j) overestimates the sensitivity. Surprisingly, this is true even for sensitivity to $q_{3,4}$ (Fig. 2a), for which the derivative of e^Q was qualitatively similar to specific compensation. This is because the magnitudes of the non-zero entries in the specific compensation pattern (1 for elements 3,4; −1 for elements 4,4) are greater than the magnitudes of the large entries in the derivative of e^Q (0·52 for elements 3,4; −0·48 for element 4,4). For states other than i and j, uniform compensation consistently overestimates the absolute sensitivity in both cases. This may be because most of the elements in column j of e^Q are little affected by changes in q_{ij}. Proportional compensation overestimates the sensitivity of the probability of empty space (state 9) in both cases. This may be because the rate of transition to free space is the largest for all non-empty states, and so is the most altered under proportional compensation. The magnitudes of the sensitivities are generally much greater for $i = 2, j = 9$ (Fig. 2b) than for $i = 3, j = 4$ (Fig. 2a: this panel is drawn with a much smaller vertical axis range). Nevertheless, the proportional differences between the compensation pattern implied by continuous-time dynamics and all three discrete-time columnwise compensation patterns can be quite large in both cases. The proportional difference in sensitivity under the compensation pattern implied by continuous-time and sensitivity under a columnwise discrete-time compensation pattern is:

$$\left| \frac{d\mathbf{w}_1}{dq_{ij}} - \frac{d\mathbf{w}_1}{dp_{ij}(t)} \right| \Big/ \left| \frac{d\mathbf{w}_1}{dq_{ij}} \right| \qquad \text{eqn 24}$$

For example, for $i = 3, j = 4$ (Fig. 2a), the proportional difference in sensitivity of staghorn *Acropora* is 0·43 under specific compensation, 0·78 under uniform compensation and 0·28 under proportional compensation. For $i = 2, j = 9$ (Fig. 2b), the proportional difference in sensitivity of empty space is 0·42 under specific compensation, 0·74 under uniform compensation and 0·23 under proportional compensation. Tables 2 and 3 show the changes in direct interaction rates corresponding to a unit infinitesimal change in p_{29}, and specific or uniform compensation, respectively. These are calculated using eqn 23, and illustrate two important points. First, the discrete-time compensation patterns used here (and almost all others) result in changes to many direct interaction rates (all the elements of Tables 2 and 3 are non-zero, although some are small). They are therefore addressing a different question from that answered by the continuous-time sensitivity analysis. Secondly, different compensation patterns correspond to quite different biological changes. For example, to achieve specific compensation, we need to make large increases in the rates of transitions to empty space from all other states (Table 2, row 9), and large decreases in the rates of transition from all states other than 9 to state 2

Table 2. Changes in the direct interaction rates **Q** corresponding to a unit infinitesimal change in p_{29} and specific compensation in element p_{99}. Data from Tanner *et al.* (1994; their Table 2)[a]

−0·0052	−0·0078	−0·0072	−0·0093	−0·0177	−0·0303	−0·0079	−0·0122	0·0196
−0·2993	−0·4412	−0·4095	−0·5244	−0·9531	−1·5334	−0·4497	−0·6712	1·4927
−0·0121	−0·0182	−0·0168	−0·0217	−0·0417	−0·0716	−0·0186	−0·0286	0·0454
−0·0074	−0·0111	−0·0102	−0·0132	−0·0255	−0·0441	−0·0113	−0·0174	0·0272
0·0013	0·0019	0·0018	0·0023	0·0045	0·0078	0·0020	0·0030	−0·0046
0·0000	0·0000	0·0000	0·0000	0·0000	0·0001	0·0000	0·0000	0·0000
−0·0019	−0·0028	−0·0026	−0·0034	−0·0065	−0·0113	−0·0029	−0·0045	0·0069
−0·0046	−0·0069	−0·0064	−0·0082	−0·0161	−0·0280	−0·0071	−0·0109	0·0166
0·3291	0·4859	0·4509	0·5779	1·0561	1·7109	0·4955	0·7417	−1·6037

[a]The states are ordered encrusting acroporid corals, tabular *Acropora*, bushy *Acropora*, staghorn *Acropora*, soft corals, algae, massive corals, pocilloporid corals, free space.

Table 3. Changes in the direct interaction rates **Q** corresponding to a unit infinitesimal change in p_{29} and uniform compensation in other elements of column 9. Data from Tanner *et al.* (1994; their Table 2)[a]

0·0277	0·0406	0·0377	0·0481	0·0853	0·1328	0·0413	0·0608	−0·1509
−0·3010	−0·4436	−0·4117	−0·5273	−0·9587	−1·5433	−0·4522	−0·6751	1·4986
0·0228	0·0332	0·0308	0·0393	0·0682	0·1032	0·0338	0·0492	−0·1329
0·0278	0·0406	0·0377	0·0482	0·0845	0·1292	0·0414	0·0606	−0·1536
0·0596	0·0886	0·0822	0·1056	0·1976	0·3297	0·0905	0·1372	−0·2661
0·0829	0·1242	0·1151	0·1480	0·2848	0·4921	0·1269	0·1950	−0·3422
0·0428	0·0632	0·0586	0·0752	0·1376	0·2233	0·0644	0·0966	−0·2066
0·0455	0·0673	0·0624	0·0801	0·1474	0·2407	0·0686	0·1031	−0·2161
−0·0082	−0·0141	−0·0130	−0·0172	−0·0466	−0·1078	−0·0147	−0·0274	−0·0302

[a]The states are ordered encrusting acroporid corals, tabular *Acropora*, bushy *Acropora*, staghorn *Acropora*, soft corals, algae, massive corals, pocilloporid corals, free space.

(Table 2, row 2). To achieve uniform compensation, we make large decreases in the rates of transition from all states other than 9 to state 2 (Table 3, row 2), but make small decreases in the rates of transitions to empty space from all other states (Table 3, row 9). It is not clear whether these particular patterns of change are of biological interest.

Temporal variability

The example above came from a tropical habitat where seasonal variation may be relatively unimportant (although rates certainly do vary over time in all systems). This is unlikely to be true in all cases. For example, transition probabilities were significantly different between summer and winter for a mussel bed community on the north-east coast of the United States (Wootton 2001b). A homogeneous continuous-time model was not a good fit to the annual transition probabilities (M. Spencer, unpublished analysis). If we can find seasons within which transitions are homogeneous, then there is a relatively simple solution. Suppose we have a case where $\mathbf{P}(t) = \mathbf{P}_1 \mathbf{P}_2 = e^{\mathbf{Q}_1 t_1} e^{\mathbf{Q}_2 t_2}$, with $t = t_1 + t_2$, and t_1, t_2 known. For example, in the mussel bed system (Wootton 2001b), $t_1 = 8$ months for the winter transition matrix, and $t_2 = 4$ months for the summer transition matrix. If homogeneity held within seasons, we could obtain separate estimates of $\mathbf{Q}_1$ and $\mathbf{Q}_2$ from the corresponding transition probability matrices (Kalbfleisch & Lawless 1985), using the methods discussed in the section 'Estimating the continuous-time model'. The m, nth element of $\mathbf{P}(t)$ is:

$$p_{mn}(t) = \sum_k p_{mk}^{(1)} p_{kn}^{(2)} \qquad \text{eqn 25}$$

where $p_{mk}^{(1)}$ is the m, kth element of $\mathbf{P}_1$. Taking the derivative of each term in the sum, the derivatives of $\mathbf{P}(t)$ with respect to the i, jth element of $\mathbf{Q}_1$ are given by:

$$\mathbf{D}_\mathbf{V}(e^{\mathbf{Q}_1 t_1} e^{\mathbf{Q}_2 t_2})^{(ij,1)} = \mathbf{D}_\mathbf{V}(e^{\mathbf{Q}_1 t_1})^{(ij)} \mathbf{P}_2 \qquad \text{eqn 26}$$

Similarly, the derivatives of $\mathbf{P}(t)$ with respect to the i, jth element of $\mathbf{Q}_2$ are:

$$\mathbf{D}_\mathbf{V}(e^{\mathbf{Q}_1 t_1} e^{\mathbf{Q}_2 t_2})^{(ij,2)} = \mathbf{P}_1 \mathbf{D}_\mathbf{V}(e^{\mathbf{Q}_2 t_2})^{(ij)} \qquad \text{eqn 27}$$

We could substitute these derivatives into eqn 22 to calculate the sensitivity of the stationary distribution. The generalization to simultaneous changes in q_{ij} in both $\mathbf{Q}_1$ and $\mathbf{Q}_2$ is straightforward:

$$\mathbf{D}_\mathbf{V}(e^{\mathbf{Q}_1 t_1} e^{\mathbf{Q}_2 t_2})^{(ij)} = \mathbf{D}_\mathbf{V}(e^{\mathbf{Q}_1 t_1})^{(ij)} \mathbf{P}_2 + \mathbf{P}_1 \mathbf{D}_\mathbf{V}(e^{\mathbf{Q}_2 t_2})^{(ij)} \quad \text{eqn 28}$$

However, transitions may not be homogeneous even within the fixed seasons (this is the case for Wootton's data: M. Spencer, unpublished analysis). Models of temporal variability use extra parameters, and a homogeneous model within each season already has the same number of parameters as observations (counts of each transition in each season). Thus, a model that is not homogeneous within seasons will not be identifiable from data of the kind discussed here. The general conclusion that the single-column compensation patterns discussed in the section 'Sensitivity of the stationary

distribution in a discrete-time model' correspond to changes in many direct interactions will still hold, but we will not know what compensation pattern is associated with changes in only one direct interaction rate.

Discussion

Sensitivity analysis is important in many contexts, including designing management strategies and sampling schemes. To obtain useful information, we have to know that the sensitivity analysis we use is biologically relevant. However, the biological meaning of the compensation patterns used in existing sensitivity analyses of discrete-time Markov models is unclear. Continuous-time models may clarify the changes in interspecific interactions that are implied by a given compensation pattern, and will therefore help us to ask the right questions. There are many further avenues for research in continuous-time models. First, the maximum-likelihood approach allows us to estimate a best-fitting continuous-time model. However, there could be other models that fit almost as well that have quite different structures. There are several possible approaches to dealing with this uncertainty (Singer & Spilerman 1976; Geweke *et al.* 1986).

Secondly, the transition probability matrix is usually averaged over large numbers of sample points and time intervals, even though there is often significant spatial and temporal heterogeneity (Tanner *et al.* 1996). Homogeneous models are caricatures at best. Making the transition rates simple functions of space and time (Kalbfleisch & Lawless 1985) might give better models. Other options include random-effects models in which transition rates are sampled from some parametric distribution (Yang 1994), and hidden Markov models in which different rate matrices operate at different times (Tucker & Anand 2005). Thirdly, there is an obvious connection between neutral models of biodiversity (Hubbell 1997) and continuous-time Markov chains. Abundance distributions do not provide very sensitive tests of neutral models, because many different models can generate similar abundance distributions (He 2005). Comparing the estimated distribution of interaction strengths with the predicted distribution from a neutral model might tell us why neutral models sometimes fail (Wootton 2005). In conclusion, modelling communities of sessile organisms as continuous-time rather than discrete-time Markov chains seems to offer increased biological understanding and a clearer view of the consequences of changing single interspecific interactions. It may therefore be worth applying continuous-time models more widely than they have been so far.

Acknowledgements

References

Bladt, M. & Sørensen, M. (2005) Statistical inference for discretely observed Markov jump processes. *Journal of the Royal Statistical Society, Series B*, **67**, 395–410.

Caswell, H. (2001) *Matrix Population Models: Construction, Analysis, and Interpretation*, 2nd edn. Sinauer, Sunderland, MA.

Cohen, J.E. & Singer, B. (1979) Malaria in Nigeria: constrained continuous-time Markov models for discrete-time longitudinal data on human mixed-species infections. *Lectures on Mathematics in the Life Sciences*, **12**, 69–133.

Connell, J.H. (1975) Some mechanisms producing structure in natural communities: a model and evidence from field experiments. *Ecology and Evolution of Communities* (eds M.L. Cody & J.L. Diamond), pp. 460–490. Belknap, Harvard.

Geweke, J., Marshall, R.C. & Zarkin, G.A. (1986) Exact inference for continuous time Markov models. *Review of Economic Studies*, **53**, 653–669.

Harrison, G.W. (1995) Comparing predator–prey models to Luckinbill's experiment with *Didinium* and *Paramecium*. *Ecology*, **76**, 357–374.

He, F. (2005) Deriving a neutral model of species abundance from fundamental mechanisms of population dynamics. *Functional Ecology*, **19**, 187–193.

Hill, M.F., Witman, J.D. & Caswell, H. (2004) Markov chain analysis of succession in a rocky subtidal community. *American Naturalist*, **164**, E46–E61.

Hubbell, S.P. (1997) A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs*, **16**, S9–S21.

Jost, C. & Arditi, R. (2000) Identifying predator–prey processes from time-series. *Theoretical Population Biology*, **57**, 325–337.

Kalbfleisch, J.D. & Lawless, J.F. (1985) The analysis of panel data under a Markov assumption. *Journal of the American Statistical Association*, **80**, 863–871.

Najfeld, I. & Havel, T.F. (1995) Derivatives of the matrix exponential and their computation. *Advances in Applied Mathematics*, **16**, 321–375.

Norris, J.R. (1997) *Markov Chains*. Cambridge University Press, Cambridge, UK.

Singer, B. & Cohen, J.E. (1980) Estimating malaria incidence and recovery rates from panel surveys. *Mathematical Biosciences*, **49**, 273–305.

Singer, B. & Spilerman, S. (1976) The representation of social processes by Markov models. *American Journal of Sociology*, **82**, 1–54.

Spencer, M. & Susko, E. (2005) Continuous-time Markov models for species interactions. *Ecology*, **86**, 3272–3278.

Tanner, J.E., Hughes, T.P. & Connell, J.H. (1994) Species coexistence, keystone species, and succession: a sensitivity analysis. *Ecology*, **75**, 2204–2219.

Tanner, J.E., Hughes, T.P. & Connell, J.H. (1996) The role of history in community dynamics: a modelling approach. *Ecology*, **77**, 108–117.

Tucker, B.C. & Anand, M. (2005) On the use of stationary versus hidden Markov models to detect simple versus complex ecological dynamics. *Ecological Modelling*, **185**, 177–193.

Usher, M.B. (1979) Markovian approaches to ecological succession. *Journal of Animal Ecology*, **48**, 413–426.

Wootton, J.T. (2001a) Causes of species diversity differences:

a comparative analysis of Markov models. *Ecology Letters*, **4**, 46–56.

Wootton, J.T. (2001b) Prediction in complex communities: analysis of empirically derived Markov models. *Ecology*, **82**, 580–598.

Wootton, J.T. (2005) Field parameterization and experimental test of the neutral theory of biodiversity. *Nature*, **433**, 309–312.

Yang, Z. (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution*, **39**, 306–314.

Yodzis, P. (1998) Local trophodynamics and the interaction of marine mammals and fisheries in the Benguela ecosystem. *Journal of Animal Ecology*, **67**, 635–658.