Ecology, 86(12), 2005, pp. 3272–3278

CONTINUOUS-TIME MARKOV MODELS FOR SPECIES INTERACTIONS

MATTHEW SPENCER[1,3] AND EDWARD SUSKO[2]

[1]*Department of Mathematics and Statistics and Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia B3H 3J5 Canada*
[2]*Department of Mathematics and Statistics, Dalhousie University, Halifax, Nova Scotia, B3H 3J5 Canada*

Abstract. Discrete-time Markov chains are widely used to study communities of competing sessile species. Their parameters are transition probabilities between states (species found at points in space), estimated from repeated observations. The proportion of nonzero entries in the transition matrix has been suggested as a measure of the complexity of interspecific interactions. This is not accurate if more than one transition can occur per time interval. In such cases, continuous-time Markov chains may be better, and discrete-time models may overestimate the complexity of species interactions. We reanalyze data from a marine community. A continuous-time model with homogeneous rates is not significantly worse than the maximum-likelihood discrete-time model. Compared to the continuous-time model, the discrete-time model overestimates the complexity of interspecific interactions. We also discuss the entropy of a continuous-time Markov chain, another measure of complexity.

Key words: competition; continuous-time Markov chains; entropy; interspecific interactions; marine community dynamics; maximum likelihood; parametric bootstrap; temporal variability.

INTRODUCTION

Discrete-time Markov chain models of community dynamics are increasingly common for competing sessile species (e.g., Tanner et al. 1994, Wootton 2001, Hill et al. 2004). These models assume that the set of possible states for a given point in space is finite (typically the set of species observed in the community, plus empty space). A transition matrix **P** contains the conditional probabilities p_{ij} that a point in state j at time k will be in state i at time $k + 1$ (in other fields, it is often p_{ji} that has this meaning). Given a column vector $\mathbf{x}(k)$ of state probabilities at time k,

$$\mathbf{x}(k + 1) = \mathbf{P}\mathbf{x}(k) \tag{1}$$

(Hill et al. 2004). Most communities modeled by Eq. 1 asymptotically approach a stationary probability distribution $\boldsymbol{\pi} = \mathbf{P}\boldsymbol{\pi}$.

An appealing feature of these models is that the parameters p_{ij} have simple interpretations as probabilities of colonization, disturbance, replacement, and persistence (Hill et al. 2004). In principle, the **P** matrix can give us a great deal of information about interactions among species. For example, a community with hierarchical interactions is one in which if species A outcompetes species B, and B outcompetes C, then A also outcompetes C for all sets of species {A,B,C}. In a nonhierarchical community, there may be competitive networks such that A outcompetes B, B outcompetes C, C outcompetes A (Buss and Jackson 1979). A community with strictly hierarchical interactions will have no cycles in the directed graph corresponding the **P** matrix, other than cycles that pass through the empty space state. Along the same lines, the proportion of nonzero elements in the **P** matrix (the proportion of transitions that occur) has been suggested as a measure of the complexity of species interactions (Tanner et al. 1994), analogous to the connectance of a food web matrix (e.g., Warren 1994). Using this measure, a more complex community is one in which a greater proportion of possible replacements of one species by another are observed. This has an intuitive interpretation. If there are more possible transitions out of a given state, we will be less certain of the state one step in the future, other things being equal.

Unfortunately, these simple interpretations break down if more than one transition can occur between time k and time $k + 1$ (Wootton 2001). In such cases, it may be better to model community dynamics as a continuous-time Markov chain. Here, we illustrate the difference between continuous- and discrete-time Markov chain models. Using a hypothetical example, we

---

Manuscript received 10 January 2005; revised 6 June 2005; accepted 13 June 2005. Corresponding Editor: L. Stone.

[3] E-mail: matts@mathstat.dal.ca

show that if the underlying dynamics occur in continuous time, we may not be able to infer which interactions are possible using a discrete-time model. We can make these inferences from a continuous-time model if the transition rates are constant over time (the model is time-homogeneous). Furthermore, observing that p_{ij} is larger than p_{kl} does not necessarily allow us to infer that the rate of transitions from j to i is higher than the rate of transitions from l to k, especially if there is a high probability of more than one event per unit time. To illustrate this, we reanalyze data from one of the sites studied by Tanner et al. (1994). We fit discrete-time and homogeneous continuous-time models, and show using a likelihood ratio test that we cannot reject the hypothesis of time homogeneity, although in reality there is likely to be some variation in rates over time. The homogeneous continuous-time model predicts a lower complexity of species interactions than the discrete-time model. We also discuss a measure of entropy for continuous-time models, analogous to the relative entropy used by Hill et al. (2004) for discrete-time models.

CONTINUOUS-TIME MARKOV CHAINS

The discrete-time system of Eq. 1 has the continuous-time analogue

$$\mathbf{x}'(k) = \mathbf{Q}\mathbf{x}(k) \qquad (2)$$

where $\mathbf{x}'(k)$ is vector of first derivatives with respect to time and $\mathbf{Q}$ is the matrix of instantaneous rates. The off-diagonal elements of $\mathbf{Q}$ can take any nonnegative real values. The rate of leaving state j is $-q_{jj} = \Sigma_{i \neq j} q_{ij}$. We assume throughout that the Markov chain is irreducible (every state can be reached from any other state), which implies, among other things, that $q_{jj} \neq 0$. This ensures that the stationary distribution is unique. Irreducibility is a reasonable assumption if disturbance can return the community to empty space from any state.

There is one discrete-time transition matrix $\mathbf{P}(t)$ with a time step of t corresponding to every instantaneous rate matrix $\mathbf{Q}$. We can find this transition matrix using the matrix exponential

$$\mathbf{P}(t) = e^{\mathbf{Q}t} = \sum_{m=0}^{\infty} \frac{(\mathbf{Q}t)^m}{m!} \qquad (3)$$

(e.g., Norris 1997:62–63). This means that if we are interested only in making predictions at evenly-spaced points in time, a $\mathbf{P}$ matrix corresponding to the $\mathbf{Q}$ matrix of a continuous-time Markov chain is a perfectly good representation. The stationary distributions of the discrete- and continuous-time models are identical. The possibility of equivalent discrete- and continuous-time Markov models is well known (e.g., Logofet and Lesnaya 2000), but continuous-time models are less widely used in ecology. The matrix exponential is not straightforward to compute (Moler and Van Loan 1978), but is implemented in software such as Matlab (Mathworks, Natick, Massachusetts, USA; we used releases 13 and 14 for all calculations). Throughout, we will refer to the one-time-unit matrix $\mathbf{P}(1)$ simply as $\mathbf{P}$.

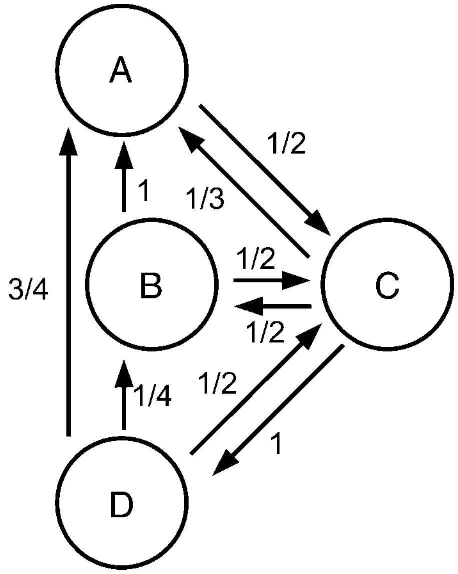

FIG. 1. Hypothetical example of a continuous-time Markov chain for three species A, B, and C, and empty space S. Transitions that may occur are shown as arrows, labeled with their instantaneous rates.

INFERRING SPECIES INTERACTIONS

Consider the hypothetical three-species plus empty space system shown in Fig. 1, where A, B, and C are species and S is empty space. In this system, A is a superior competitor to both B and C, and B is a superior competitor to C. All three species are affected by disturbance (the transition to S) at an equal rate. C is the fastest and A the slowest colonizer of empty space. The instantaneous rate matrix is

$$\mathbf{Q} = \begin{array}{c|cccc} & \mathrm{A} & \mathrm{B} & \mathrm{C} & \mathrm{S} \\ \hline \mathrm{A} & -1/2 & 1 & 3/4 & 1/3 \\ \mathrm{B} & 0 & -3/2 & 1/4 & 1/2 \\ \mathrm{C} & 0 & 0 & -3/2 & 1 \\ \mathrm{S} & 1/2 & 1/2 & 1/2 & -11/6. \end{array} \qquad (4)$$

There are three zero elements (q_{BA}, q_{CA}, and q_{CB}). Using Eq. 3, the corresponding discrete-time transition matrix for a time interval of one unit is

$$\mathbf{P} = \begin{array}{c|cccc} & \mathrm{A} & \mathrm{B} & \mathrm{C} & \mathrm{S} \\ \hline \mathrm{A} & 0.6828 & 0.4596 & 0.4039 & 0.3372 \\ \mathrm{B} & 0.0452 & 0.2683 & 0.1010 & 0.1424 \\ \mathrm{C} & 0.0785 & 0.0785 & 0.3017 & 0.2299 \\ \mathrm{S} & 0.1935 & 0.1935 & 0.1935 & 0.2905. \end{array} \qquad (5)$$

There are no zero entries in the $\mathbf{P}$ matrix, because the Markov chain is irreducible. From examining the

P matrix alone, we would conclude that the competitive interactions in this community are nonhierarchical, because we would sometimes observe transitions such as A to C. The underlying dynamics are hierarchical, because we can only reach C from A in two transitions. **Q** gives the direct interactions between species, while **P** is the net result of direct and indirect interactions over a finite time interval.

The elements corresponding to the three zeros in **Q** are the three smallest elements in **P**. We might therefore hope that the rank order of elements in **P** tells us about the relative rates of transitions in **Q**. For example, Tanner et al. (1994) suggest that ''the interaction strength is given directly by the transition probabilities.'' This is not necessarily so if the rates of change are high. For example, q_{AS} (the rate at which A colonizes empty space), is 1/3 and q_{CS} (the rate at which C colonizes empty space) is 1. The corresponding transition probabilities per unit time are $p_{\mathrm{AS}} = 0.3372$, $p_{\mathrm{CS}} = 0.2299$. Although C colonizes empty space more rapidly than A, the rate $-q_{\mathrm{CC}}$ at which C is replaced by other species is large. A point colonized by a C individual is often replaced by an individual of another species before the next observation is taken. The rate of leaving state A is much smaller, so an A individual colonizing empty space is likely to persist long enough to be sampled.

EXPECTED NUMBER OF EVENTS AND PROBABILITY OF MORE THAN ONE EVENT PER UNIT TIME

The example above showed that we do not expect the rank order of elements in **P** to tell us much about the relative rates of transitions in **Q** when the rates of change are high. The mean number r of events per unit time at steady state is -1 times the sum of the diagonal elements, weighted by the stationary probability of each state:

$$r = -\sum_j \pi_j q_{jj} \tag{6}$$

where π_j is the stationary probability of state j. The waiting time Δ_j to the next event when we are in state j is exponentially distributed with parameter $-q_{jj}$, so

$$P(>1 \text{ event per unit time}) = \sum_j \sum_{i \neq j} \pi_j \frac{q_{ij}}{-q_{jj}} P(\Delta_j + \Delta_i \leq 1). \tag{7}$$

Here, $q_{ij}/(-q_{jj})$ is the probability that a transition from state j takes us to state i, and $P(\Delta_j + \Delta_i \leq 1)$ is the probability that the sum of waiting times to leave the first state j and the second state i is no more than one unit. The summation is over states $i \neq j$ because we have to enter a different state on leaving the first state j. Δ_j and Δ_i are independent, so

$$\begin{aligned} P(\Delta_j + \Delta_i \leq 1) &= \int_0^1 f_{\Delta_j}(x) P(\Delta_i \leq 1 - x)\, dx \\ &= \begin{cases} \dfrac{q_{ii}(1 - e^{q_{jj}}) - q_{jj}(1 - e^{q_{ii}})}{q_{ii} - q_{jj}} & q_{jj} \neq q_{ii} \\ e^{q_{jj}}(q_{jj} - 1) + 1 & q_{jj} = q_{ii} \end{cases} \end{aligned} \tag{8}$$

where $f_{\Delta_j}(x)$ is the probability density of Δ_j at x. For the hypothetical case given by Eq. 4, the mean number of events per unit time is 1.0238 and the probability of more than one event per unit time at stationarity is 0.2957.

ENTROPY OF CONTINUOUS-TIME MARKOV CHAINS

The proportion of zeros in a matrix gives the same weight to rare and common transitions. Hill et al. (2004) suggest using the entropy of a discrete-time transition matrix as an index of successional complexity that takes account of differences in transition probabilities. As with the proportion of zeros, the entropy of the **P** matrix does not necessarily reflect the interactions that actually occur. One analogous entropy for a continuous-time rate matrix is the entropy of the jump matrix **S**. The jump matrix is a transition matrix whose entries s_{ij} are the conditional probabilities of the new state at a point being i when a change of state occurs, given that the current state at the location is j:

$$s_{ij} = \begin{cases} -q_{ij}/q_{jj} & j \neq i \text{ and } q_{jj} \neq 0 \\ 0 & j \neq i \text{ and } q_{jj} = 0 \end{cases}$$

$$s_{jj} = \begin{cases} 0 & q_{jj} \neq 0 \\ 1 & q_{jj} = 0 \end{cases} \tag{9}$$

(Norris 1997:87). Then the entropy $H(\mathbf{S})$ of the jump matrix is

$$H(\mathbf{S}) = -\sum_j \pi_j^{(\mathbf{S})} \sum_i s_{ij} \log s_{ij} \tag{10}$$

(Hill et al. 2004), where $\pi_j^{(\mathbf{S})}$ is the stationary probability of state j in the jump matrix. This entropy measures how uncertain we are about the next new state of a point. By convention, $s_{ij} \log s_{ij} = 0$ if $s_{ij} = 0$, but this will not happen for an irreducible model. Because the diagonals of the jump matrix are zero for nonabsorbing states, the maximum possible entropy is $H_{\max}(\mathbf{S}) = \log(s - 1)$, where s is the number of states. The relative entropy $H_r(\mathbf{S})$ for the jump matrix is then

$$H_r(\mathbf{S}) = \frac{H(\mathbf{S})}{H_{\max}(\mathbf{S})}. \tag{11}$$

For the hypothetical matrix in Eq. 4, $H(\mathbf{S}) = 0.6823$, $H_{\max}(\mathbf{S}) = 1.0986$, and $H_r(\mathbf{S}) = 0.6210$.

The jump matrix entropy takes no account of the waiting times between changes of state. Kesidis and

Walrand (1993) derive a relative entropy between two rate matrices that does account for the waiting times. However, this measure can be infinite when one of the two matrices has a zero rate where the other does not (Baldi and Piccioni 1999). Absolute measures of entropy that include waiting times can also be infinite, because increasing all the rates in the same proportion always increases the uncertainty about future states. Since an entropy that includes waiting times confounds measures of uncertainty about transitions with measures of uncertainty about waiting times, we will use only the jump chain entropy here.

FITTING HOMOGENEOUS CONTINUOUS-TIME MODELS TO REAL DATA

We estimated the time-homogeneous **Q** matrix corresponding to a real transition matrix. Tanner et al. (1994, Exposed Crest site, their Table 2) estimated transition matrices among eight species groups and empty space on a coral reef, from observations at approximately two-year intervals. It seems possible that more than one event might occur per time interval at a given point. For example, an organism might die at any time of year, and subsequently be replaced during the breeding season of another species. Although settlement may only occur at some times of the year, the breeding seasons of opportunist species that colonize empty space tend to be long (Connell 1975:477). Furthermore, the observation interval is long enough to contain two breeding seasons.

Given n_{ij}, the number of times each transition was observed, we first need to estimate **P**. The original data are observations at a number of sites of the states i at approximately equally spaced time points. At a single site, we observe a sequence of states $i_0, i_1, \ldots, i_T$ at times $0, 1, \ldots, T$. Assuming a stationary Markov chain, the probability of this sequence of observations given a set of transition probabilities $P(i_{k+1}|i_k)$ and stationary probabilities π_i is

$$\begin{aligned} &P(i_0, i_1, \ldots, i_T) \\ &\quad = P(i_T \mid i_{T-1})P(i_{T-1} \mid i_{T-2}) \cdots P(i_1 \mid i_0)\pi_{i_0} \qquad (12) \end{aligned}$$

where π_{i_0} is the stationary probability of the state observed at time 0. Summing over a number of independent and identically distributed sites and all time points and taking logs, we get the full log likelihood

$$l_{\text{full}} = c + \sum_{i,j} n_{ij} \log p_{ij} + \sum_{j} n_j \log \pi_j \qquad (13)$$

where n_{ij} is the observed number of transitions from state j to state i, p_{ij} is the predicted probability of a transition from j to i, and n_j is the observed number of sites having initial state j. We ignore c, which is the log of a multinomial coefficient and is constant for given data. We cannot obtain the initial states from aggregated data giving the number of observations of each transition, as presented in Tanner et al. (1994). We therefore ignore the contribution from initial states and maximize

$$l = \sum_{i,j} n_{ij} \log p_{ij} \qquad (14)$$

This will usually be a good approximation to maximizing the full log likelihood, because the n_j sum to the number of sites n_s, while the n_{ij} sum to $n_s T$, where T is the number of sample intervals. Furthermore, the π_j are only indirectly related to the transition probabilities, and there may be many transition matrices consistent with a given set of π_j. Eq. 14 can also be considered as a likelihood conditional upon the initial states. If the process is not assumed to be at equilibrium at time 0, the initial states are ancillary statistics, and inference should be conditional upon them (Lehmann 1986: section 10.2).

Using Eq. 14, we obtain the usual maximum-likelihood estimate $\hat{\mathbf{P}}$ of the transition matrix **P**

$$\hat{p}_{ij} = \frac{n_{ij}}{n_{\cdot j}} \qquad (15)$$

where $n_{\cdot j}$ is the total number of transitions observed out of state j (Caswell 2001:135).

Eq. 3 suggests $\log \hat{\mathbf{P}}$ as an estimate of the instantaneous rate matrix, where log is the matrix logarithm, defined such that $e^{\log \hat{\mathbf{P}}} = \hat{\mathbf{P}}$. We used the matrix log algorithm implemented in Matlab releases 13 and 14 (Mathworks), described in Davies and Higham (2003). For the Exposed Crest coral reef site in Tanner et al. (1994, their Table 2), 13 out of 72 off-diagonal elements in $\log \hat{\mathbf{P}}$ were negative. **Q** matrices with negative off-diagonal elements are not biologically meaningful, but could result from either sampling error or temporal variability.

Given observations at fixed time intervals, a homogeneous continuous-time Markov chain is a special case of a discrete-time Markov chain, but there are discrete-time models that do not correspond to any valid homogeneous continuous-time model. Thus the class of homogeneous continuous-time models is nested within the class of discrete-time models, and we can use a likelihood ratio test of homogeneity. There is almost certainly some temporal variability in the community studied by Tanner et al., since events such as storm damage and settlement may be more frequent at some times of year than others. Furthermore, the Tanner et al. (1994) data average over years with and without major cyclones. Nevertheless, if we cannot reject time homogeneity, the estimated **Q** matrix may be a good source of information about interspecific interactions.

We used maximum likelihood to find the best-fitting **Q** matrix with nonnegative off-diagonal elements. We used $\log \hat{\mathbf{P}}$ as a starting point, setting any negative off-diagonals to 2×10^{-16}. We then used nonlinear optimization of all the q_{ij} to minimize the negative log-likelihood, constraining all parameters to be nonneg-

TABLE 1. The estimated instantaneous rate matrix $\hat{\mathbf{Q}}^{+}$ for the data in Tanner et al. (1994, their Table 2).

−0.482	†	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	†	−0.555	0.0148	0.0784	0.139	0.15	0.059	0.0634
0.0245	†	0.00704	−0.767	0.0727	0.235	0.0342	0.0833	0.0369
0.000669	0.0131	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	0.00501	†	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

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

† Elements that were nonzero in the discrete-time transition matrix $\hat{\mathbf{P}}$ but not in $\hat{\mathbf{Q}}^{+}$.

ative. Once the optimization was finished, we set any parameters that were less than 4×10^{-16} to zero (with undetectable effect on the log likelihood). Using the resulting estimate $\hat{\mathbf{Q}}^{+}$, we calculated $-2\Delta l$ (twice the difference in log likelihoods) for the comparison between $e^{\hat{\mathbf{Q}}^{+}}$ and the maximum-likelihood estimate $\hat{\mathbf{P}}$.

The distribution of $-2\Delta l$ is complicated because some parameters are on the boundaries of their ranges (Self and Liang 1987), so we estimated it using a parametric bootstrap (Efron and Tibshirani 1993: section 6.5). For each bootstrap replicate, we simulated data by stochastic iteration of Eq. 1, assuming $\hat{\mathbf{Q}}^{+}$ was the true model and fixing the total number of transitions at the observed value. We took an initial multinomial sample $\mathbf{n}(0)$ of size n_s from the stationary distribution of $\hat{\mathbf{Q}}^{+}$, where n_s is 2122 (J. Tanner, *personal communication*). We then took a multinomial sample of $n_j(0)$ transitions out of each state j from the jth column of $e^{\hat{\mathbf{Q}}^{+}}$, where $n_j(0)$ is the jth entry in $\mathbf{n}(0)$. Summing the destination states over all j gives the new state vector $\mathbf{n}(1)$. We repeated this process for T steps. A total of 22 352 transitions were observed (Tanner et al. 1994; the number of observations of free space should be 14 255, not 1425 as printed; J. Tanner, *personal communication*). There were 12 sample times ($T = 11$), with some missing data due to unidentifiable or very rare species. We deleted transitions at random, assuming each was equally likely to be missing, to obtain exactly the observed number of transitions. For the kth bootstrap replicate, we then estimated the bootstrap value of the test statistic $(-2\Delta l)^{*}_{k}$ exactly as for the observed data. Finally, we estimated the probability of a $-2\Delta l$ at least as large as observed if $\hat{\mathbf{Q}}^{+}$ was the true model as $p = \#\{(-2\Delta l)^{*}_{k} \geq -2\Delta l\}/B$, where B is the number of bootstrap replicates (1000). A large p value indicates that we cannot reject the null hypothesis of time homogeneity.

$\hat{\mathbf{Q}}^{+}$ had partial log likelihood -1.7350×10^{4}, and the maximum-likelihood model $\hat{\mathbf{P}}$ had partial log likelihood -1.7340×10^{4} ($-2\Delta l = 18.6$, $p = 0.096$, maximum $(-2\Delta l)^{*}_{k} = 29.0$, median $= 11.8$). We therefore cannot reject the hypothesis of rate homogeneity in continuous time for the Tanner et al. (1994) data. Sampling error alone could explain the negative entries in $\log \hat{\mathbf{P}}$. Furthermore, $\log \hat{\mathbf{P}}$ was very similar to $\hat{\mathbf{Q}}^{+}$ (least-squares regression of off-diagonals with $\hat{\mathbf{Q}}^{+}$ as predictor: intercept, −0.001; SE, 0.0007; slope, 1.016; SE, 0.003).

There was a strong Spearman correlation (0.95) between the off-diagonal elements of $\hat{\mathbf{P}}$ and $\hat{\mathbf{Q}}^{+}$. Because the mean number of events per unit time (0.33, from Eq. 6) and the probability of more than one event per unit time (0.06, from Eq. 7) are fairly low, the relative magnitudes of the transition probabilities are a reliable guide to the relative magnitudes of the instantaneous transition rates. Despite this strong correlation, the number of zero off-diagonal entries in $\hat{\mathbf{P}}$ is very different from the number of zero off-diagonal entries in $\hat{\mathbf{Q}}^{+}$. $\hat{\mathbf{P}}$ had seven off-diagonal zeros, while $\hat{\mathbf{Q}}^{+}$ had 13 (Table 1: the same 13 that were negative in $\log \hat{\mathbf{P}}$). This implies that even though the probability of more than one event per unit time is quite small, $\hat{\mathbf{P}}$ may substantially overestimate the complexity of species interactions.

The jump matrix entropy was 1.24, and the relative entropy was 0.59, for the jump matrix corresponding to $\hat{\mathbf{Q}}^{+}$. The relative entropy uses the maximum possible jump matrix entropy given the number of states as a frame of reference, and has a maximum possible value of 1. Thus, although 90% of the entries in $\hat{\mathbf{P}}$ were nonzero, we are not particularly uncertain about the next new state at a given point. Because the number of events per unit time was small, the relative jump matrix entropy for the continuous-time model was close to that for the discrete-time model (0.58: the definition of the jump matrix for the discrete-time model is similar to Eq. 9, except that off-diagonal elements are normalized by $1 - p_{jj}$, the probability of a different state after one time unit). In general, we will not know whether this is the case unless we fit both discrete- and continuous-time models.

To show that we can accurately recover community structure even with fewer data, we treated the estimated $\hat{\mathbf{Q}}^{+}$ matrix as the true matrix, and simulated sampling from this with 10, 15, or 20 sample times (two-thirds, three-thirds, or four-thirds of the actual number of observations). The mean jump matrix entropy and mean proportion of nonzero off-diagonal elements were not significantly different from the true values (approximate P values from distributions of absolute deviations from the mean of the simulated distributions >0.39 in all cases, with 1000 replicates). Standard deviations of the simulated distributions were <0.022 for jump matrix entropy and <0.035 for proportion of nonzero off-diagonal entries, even with two-thirds of the actual number of observations.

Discussion

For the system we analyzed, Tanner et al. (1994) commented that ''the most remarkable aspect of the [discrete-time] matrices is the marked paucity of zeros, i.e., almost every conceivable transition took place, albeit at markedly different rates.'' Our analysis suggests that some of these transitions only took place indirectly. The difference between discrete- and continuous-time matrices is somewhat analogous to the difference between competitive interactions considered in isolation and the net effects of interactions within a community (Stone and Roberts 1991).

The proportion of nonzero elements in the instantaneous rate matrix is more appropriate than the proportion of nonzero elements in the discrete-time transition matrix as a measure of the complexity of direct competitive interactions between species. However, this measure is highly sensitive to sampling variability (because rare events will often be missed completely in small samples) and to observation errors (e.g., occasional misidentified species). The entropy of the jump chain is another possible measure of complexity, and is less sensitive to sampling errors. Both measures are summary statistics, and cannot tell us about the detailed topology of interactions. If the true process occurs in continuous time but the expected number of events per unit time is small enough, the transition probabilities from a discrete-time model are approximately the probabilities of direct interactions between species. However, we would still have to fit a continuous-time model in order to estimate the expected number of events. We therefore think that, if there are biological grounds for believing that interactions happen in continuous time, there is little reason to make inferences about interactions from a discrete-time model.

Temporal variability complicates the estimation of instantaneous transition rates from observations at discrete times. In many communities, some events (such as breeding) may happen during a short time interval, while other events (such as death and overgrowth) may occur during most of the year. Furthermore, the sampling interval for the data we analyzed (Tanner et al. 1994) was longer than a year, and there is good evidence for temporal variability in transition probabilities at annual or longer time scales (e.g., Usher 1979, Hill et al. 2002). There are some kinds of historical events that cannot be accurately represented by any first-order Markov model (Tanner et al. 1996). Thus, neither discrete-time models nor homogeneous continuous-time models are entirely realistic. Nevertheless, in the case we analyzed, a homogeneous continuous-time model was a reasonable fit to the data. This suggests that although temporal variability is ubiquitous in nature, ignoring it will not necessarily lead to very inaccurate results.

Conclusions

If competitive interactions between species occur in continuous time, then estimates of transition rates from a continuous-time model have a simple biological interpretation, while estimates from a corresponding discrete-time model do not. In such cases, the graph of nonzero interactions in the discrete-time model may give a misleading picture of the direct competitive interactions between species. The proportion of nonzero instantaneous rates observed is a measure of the complexity of interspecific interactions, although this statistic takes no account of the sizes of rates and is sensitive to rare events. Entropy measures such as the relative entropy of the jump chain do not have these problems, but neither measure can tell us about detailed interaction structure.

Acknowledgments

This work was funded by the Genome Atlantic/Genome Canada Prokaryotic Genome Evolution and Diversity Project. E. Susko is supported by the Natural Sciences and Engineering Council of Canada. We are grateful to Moshe Kiflawi, Andrew Roger, Lewi Stone, Phil Warren, and two anonymous reviewers for their comments on the ideas expressed here. Jason Tanner kindly answered our questions on the data analyzed here.

Literature Cited

Baldi, P., and M. Piccioni. 1999. A representation formula for the large deviation rate function for the empirical law of a continuous time Markov chain. Statistics and Probability Letters **41**:107–115.

Buss, L. W., and J. B. C. Jackson. 1979. Competitive networks: nontransitive competitive relationships in cryptic coral reef environments. American Naturalist **113**:223–234.

Caswell, H. 2001. Matrix population models: construction, analysis, and interpretation. Second edition. Sinauer, Sunderland, Massachusetts, USA.

Connell, J. H. 1975. Some mechanisms producing structure in natural communities: a model and evidence from field experiments. Pages 460–490 *in* M. Cody and J. Diamond, editors. Ecology and evolution of communities. Belknap, Harvard, Cambridge, Massachusetts, USA.

Davies, P. I., and N. J. Higham. 2003. A Schur-Parlett algorithm for computing matrix functions. SIAM Journal on Matrix Analysis and Applications **25**:464–485.

Efron, B., and R. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, New York, New York, USA.

Hill, M., J. Witman, and H. Caswell. 2002. Spatio-temporal variation in Markov chain models of subtidal community succession. Ecology Letters **5**:665–675.

Hill, M., J. Witman, and H. Caswell. 2004. Markov chain analysis of succession in a rocky subtidal community. American Naturalist **164**:E46–E61.

Kesidis, G., and J. Walrand. 1993. Relative entropy between Markov transition matrices. IEEE Transactions on Information Theory **39**:1056–1057.

Lehmann, E. 1986. Testing statistical hypotheses. Second edition. John Wiley and Sons, New York, New York, USA.

Logofet, D., and E. Lesnaya. 2000. The mathematics of Markov models: what Markov chains can really predict in forest successions. Ecological Modelling **126**:285–298.

Moler, C., and C. Van Loan. 1978. Nineteen dubious ways to compute the exponential of a matrix. SIAM Review **20**: 801–836.

Norris, J. 1997. Markov chains. Cambridge University Press, Cambridge, UK.

Self, S., and K.-Y. Liang. 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. Journal of the American Statistical Association **82**:605–610.

Stone, L., and A. Roberts. 1991. Conditions for a species to gain advantage from the presence of competitors. Ecology **72**:1964–1972.

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

Tanner, J., T. Hughes, and J. Connell. 1996. The role of history in community dynamics: a modelling approach. Ecology **77**:108–117.

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

Warren, P. H. 1994. Making connections in food webs. Trends in Ecology and Evolution **9**:136–141.

Wootton, J. 2001. Causes of species diversity differences: a comparative analysis of Markov models. Ecology Letters **4**:46–56.

SUPPLEMENT

Matlab code for testing the hypothesis of time homogeneity by parametric bootstrap is available in ESA's Electronic Data Archive: *Ecological Archives* E086-179-S1.