Past states of continuous-time Markov models for ecological communities

Matthew Spencer

School of Biological Sciences, University of Liverpool, Liverpool, L69 7ZB, UK.

Phone +44 (0)151 795 4399

Abstract

Discrete-time Markov chains are often used to model communities of sessile organisms. The community is described by a set of discrete states, which may represent species or groups of species. Transitions between states are modelled using a stochastic matrix. A recent study showed how the time-reversal of such a Markov chain can be used to estimate the distribution of time since the last occurrence of some state of interest (such as free space) at a point, given the current state of the point. However, if the underlying process operates in continuous time but is observed at regular intervals, this distribution describes the time since the last possible observation of the state of interest, rather than the time since its last occurrence. We show how to obtain the distribution of time since the last occurrence of a state of interest for a continuous-time homogeneous Markov chain. The expected time since the last occurrence of free space can be interpreted as a measure of the successional rank of a state. We show how to distinguish between different ways in which a state can have high successional rank. We apply our results to a marine subtidal community.

Key words: Continuous time Markov chains, community ecology, first passage time, marine subtidal, time reversal

Preprint submitted to Elsevier Science 22 June 2007
1 Introduction

Markov chains are often used to model the dynamics of communities of sessile organisms [1–4]. It is assumed that a point in space can be in one of a finite number of states 1...s, such as different species or groups of species, and free space. In most cases, the dynamics of the system are assumed to be described by a temporally homogeneous, ergodic discrete-time Markov chain. Temporally homogeneous means that the conditional probability of each state one unit of time in the future, given the current state, does not depend on time. Ergodic means that any state is eventually reachable from any other state, the expected return time to any state is finite, and the probability of returning to any state i in n time steps is non-zero for all sufficiently large n. Let \( P(t) \) be an s-by-s matrix of transition probabilities, whose entries \( p_{ij}(t) \) are the conditional probabilities that a point in state \( i \) at time \( \tau \) will be in state \( j \) at time \( \tau + t \) (note that in much of the ecological literature, the transition probability matrix is transposed relative to this definition). We will sometimes refer simply to the transition probability matrix \( P \), when the time step \( t \) is fixed. If \( x(\tau) \) is an s-by-1 vector of state probabilities at time \( \tau \), then

\[
x^T(\tau + t) = x^T(\tau)P(t)
\]  

where \( x^T \) denotes the transpose of \( x \). \( P(t) \) is a stochastic matrix, with rows summing to 1. Under the above assumptions, there is a unique vector of stationary probabilities \( \pi \) such that \( \pi^T = \pi^TP(t) \), to which the state probabilities will converge [5, Theorem 1.8.3].

Applications of Markov chains in community ecology have mainly addressed

---

*Email address: m.spencer@liverpool.ac.uk* (Matthew Spencer).
questions such as quantifying the effects of species removals [4,6,7], comparisons of transition probabilities across communities [3,4], turnover rates and recurrence times [4], effects of hypothetical modifications of transition probabilities [3,8], and identification of keystone species [2]. Recently, [9] showed how Markov chains can be used to make inferences about past states. Specifically, they calculated the distribution of times $T_j$ since the last observation of some state $j$ (such as bare rock) at a point, given the present state of that point. In this paper, we briefly review the results presented in [9]. We show how the distribution of times since the last observation of some state at a point in a discrete-time chain, conditional on the present state, depends on the observation interval, and therefore may not tell us about the most recent occurrence of a given state in the past. We then describe analogous results for a time-homogeneous continuous-time Markov chain. If this is an appropriate model, then the resulting distribution of first passage times in the time-reversed chain can be interpreted as the distribution of times since the last occurrence of a given state at a point in space, conditional on the current state. This distribution is independent of the intervals between observations of the chain. We suggest that the expected first passage time to empty space in the time-reversed version of a Markov model for an ecological community provides a way to rank states by successional level. We show how the jump chain corresponding to a continuous-time Markov model can be used to calculate the expected number of events at a point since the last occurrence of bare rock, given the current state of the point. This allows us to distinguish between states which have high successional rank because they persist for a long time, and states which have high successional rank because they are likely to be separated from bare rock by a long series of transitions between intermediate states. All our results are based on well-known theory concerning the
transient behaviour of Markov chains, for which [10] is excellent source.

2 First passage times in time-reversed discrete-time Markov chains

To obtain the distribution of time since the last occurrence of some given state (such as bare rock) at a point in space with a known current state, [9] first obtained a time-reversed transition probability matrix $P^R$ under the assumption that state probabilities are stationary:

$$P^R = \Pi^{-1} P^T \Pi$$

[10, Definition 2.7], where $\Pi$ is a diagonal matrix of stationary probabilities. [9] then obtained the distribution of first passage times to state $j$ in the time-reversed chain. Without loss of generality, we will assume that the state of interest in the past is $j = 0$. If we write

$$P^R = \begin{bmatrix} p_{00}^R \left( P_0^R \right)^T \\ r \\ T \end{bmatrix}$$

then we can construct a chain in which state 0 is absorbing, with transition probability matrix

$$P_0^R = \begin{bmatrix} 1 \\ 0^T \\ r \\ T \end{bmatrix}$$

Then the cumulative distribution function for $T_0$ given current state $i$ is

$$F_{i0}^d(t) = P_i(T_0 \leq t) = \left[P_0^R \right]_{i0}^t t = 1, 2, \ldots$$

[10, p. 80], where the superscript $d$ in $F_{i0}^d(t)$ indicates that this is the cumulative distribution for the discrete-time chain, and the subscript $i$ in $P_i(T_0 \leq t)$
indicates that the distribution is conditional on the current state being $i$.

Equation 5 gives us the cumulative distribution for $T_0$ in the original chain because we have not changed the probability of entering state 0 from any other state, but once we have entered state 0 we never leave it. [9] applied this method to the Markov model of a rocky subtidal community in the Gulf of Maine developed by [4].

3 Interpretation of the first passage time in the time-reversed discrete-time chain

If $P(t)$ is the transition probability matrix for a discrete-time Markov chain with time step $t$, then the transition probability matrix for a chain observed at time intervals of length $nt$, where $n$ is a positive integer, is $[P(t)]^n$ [10, p. 16]. We can apply the method described in the previous section to calculate $F_{dij}(t)$ for chains with different intervals between observations. For example, in the system studied by [4], state 14 is the polychaete *Spirorbis spirorbis*, and state 15 is bare rock. Figure 1 shows the cumulative distributions $F_{14,15}^d(t)$, with time steps of one, two and five years. Even though the system is undergoing the same dynamics in each case, as the interval between observations increases, the value of $F_{14,15}^d(t)$ for any fixed, finite integer time $t$ decreases. In other words, as the interval between observations increases, so does the apparent time since the last occurrence of bare rock at a point. If there may be multiple events at a point in space between two observation times, then we should interpret $F_{dij}(t)$ as giving us the cumulative distribution of time since the last possible observation of state $j$ given the current state $i$, not necessarily the last occurrence of state $j$.
4 First passage times in time-reversed continuous-time Markov chains

We would like to know the distribution of time since the last occurrence of state $j$, given that a point is currently in state $i$. If we suppose that events can occur at any time, and that the rate of events is constant, then we may be able to find a matrix $Q$ such that

$$P(t) = e^{Qt} = \sum_{k=0}^{\infty} \frac{(Qt)^k}{k!}$$

for any non-negative real time $t$ [5, Theorem 2.1.1]. We refer to $Q$ as the generator or rate matrix for a time-homogeneous continuous-time Markov chain.

The off-diagonal elements $q_{ij}$ of $Q$ are the rates (non-negative, real, with dimensions time$^{-1}$) of transitions from state $i$ to state $j$. The diagonal elements $q_{ii}$ are defined as $-\sum_{j \neq i} q_{ij}$, so $-q_{ii}$ is the rate of leaving state $i$.

Using the homogeneous continuous-time model, we can calculate the distribution of times since the last observation of some state $j$ given current state $i$ as the interval between observations becomes arbitrarily small. The methods are similar to the discrete-time case. Using Theorem 1 in Appendix A, we can find a time-reversed generator matrix $Q^R$:

$$Q^R = \Pi^{-1}Q^T\Pi$$

Assuming without loss of generality that $j = 0$, and writing $Q^R = \begin{bmatrix} -q_0^R & (q_0^R)^T \\ r & T \end{bmatrix}$,
we construct a chain in which state 0 is absorbing,

\[ Q_0^R = \begin{bmatrix} 0 & 0^T \\ r & T \end{bmatrix} \]  

(8)

Now the cumulative distribution of first passage times to state 0 in the time-reversed chain is given by

\[ F_{i0}(t) = [e^{Q_0^R t}]_{i0} \quad t \geq 0 \]  

(9)

Theorems 2 and 3 in Appendix A give the probability density function, mean and variance for the first passage time to state 0 in the time-reversed chain.

We estimated the generator matrix (Table 1) for the system studied by [4] as described in Appendix B. Figure 2 shows the cumulative distributions of first passage times to state 15 (bare rock) from each other state, calculated from the discrete-time model with observation interval one year (solid lines) and the continuous-time model (broken lines). The distributions from the continuous-time model always lie above those for the discrete-time model. This is because if we observe the system continuously, we will see all occurrences of state \( j \) that occurred at observation times in the discrete-time chain, in addition to the occurrences between observation times (Theorem 4, Appendix A). As a result, the expected first passage time for the continuous-time chain will be less than that for the discrete-time chain (Theorem 5, Appendix A). The difference between the continuous- and discrete-time distributions may be large. For example, with current state 14 (\textit{Spirobranchus spirorbis}), the probability that the last occurrence of state 15 (bare rock) was no more than 10 years ago is 0.81 (from the continuous-time model). However, the probability that an occurrence of bare rock would have been observed in the last 10 years if we
made observations at one-year intervals is only 0.55.

5 Expected first passage time in the time-reversed chain as a measure of successional rank

Table 2 gives the mean \(E_i[T_{15}]\) and variance \(V_i[T_{15}]\) of first passage time to bare rock (state 15) in the time-reversed continuous-time chain based on data in [4], conditional on each other current state \(i\). Points currently in states for which \(E_i[T_{15}]\) is small are likely to have been bare rock more recently than points currently in states for which \(E_i[T_{15}]\) is large. Since bare rock is the initial stage of a successional process, \(E_i[T_{15}]\) tells us how far from this initial stage a point in space is likely to be, given its current state. This information is difficult to obtain in other ways when succession is stochastic and many alternative pathways are possible. However, Table 2 also shows that the variance of time since the last occurrence of bare rock is large for all states. The difference between the smallest and largest values of \(E_i[T_{15}]\) among current states is 7.93 years, but for any given current state \(V_i[T_{15}]\) is at least 70. We are quite uncertain about the past states of a point in space, given its present state.

For the continuous-time model based on data in [4], there is little relationship between the stationary probability of a state and the expected time since the last occurrence of bare rock (Figure 3a, Pearson correlation 0.21: we do not test the statistical significance of this and other correlations, because we have sampled the entire population of states). However, there is quite a strong negative relationship between the rate of leaving a state and the expected time since the last occurrence of bare rock (Figure 3b, Pearson correlation \(-0.74\)). In other words, if we observe a species such as the sea anemone *Urticina*
**crassicornis** or the sponge *Mycale lingua* that is likely to persist at a point for a long time, it is unlikely that the point was bare rock recently, while the opposite is true for a species such as *Spirorbis spirorbis* that is unlikely to persist at a point for a long time. The relationship between the stationary probability and rate of leaving a state is negative but relatively weak (Figure 3c, Pearson correlation $-0.37$). Species that are relatively poor at persisting may be able to maintain moderately high abundance by colonizing bare rock.

### 6 The number of events since the last occurrence of bare rock

The negative relationship between the rate of leaving a state and the expected time since the last occurrence of bare rock (Figure 3b) suggested that states with high successional rank tend to be highly persistent. This could mean that states with different successional rank differ mainly in the length of time it takes for them to colonize bare rock (Figure 4a). However, it is also possible for a state to have high successional rank if it is likely to be separated from bare rock by a long series of transitions among intermediate states (Figure 4b). In order to see which of these patterns occurs in a continuous-time Markov chain, we need to calculate the expected number of events (changes of state) since the last occurrence of bare rock for each current state at a point in space.

The jump matrix $S$ associated with a continuous-time Markov chain is a stochastic matrix whose entries $s_{ij}$ are the probabilities of each state $j$ being the new state when a transition occurs, given that the current state is
\[ s_{ij} = \begin{cases} 
  -q_{ij}/q_{ii} & j \neq i, q_{ii} \neq 0 \\
  0 & j \neq i, q_{ii} = 0 
\end{cases} \]
\[ s_{ii} = \begin{cases} 
  0 & q_{ii} \neq 0 \\
  1 & q_{ii} = 0 
\end{cases} \]  

\[ (10) \]

\[ [5, \text{p. 87}]. \] We can calculate S for the time-reversed chain \( Q^R \). As before, we assume without loss of generality that the state of interest in the past is \( j = 0 \), and define the absorbing jump matrix

\[ S_0 = \begin{bmatrix} 
  1 \\
  0^T \\
  r \\
  T 
\end{bmatrix} \]  

\[ (11) \]

Let \( J_0 \) be the number of jumps from state \( i \) to state 0 in the absorbing jump chain for \( Q^R \). Then we can apply Theorem 6 in Appendix A to obtain the mean \( E_i[J_0] \) and variance \( V_i[J_0] \) of the number of jumps since the last occurrence of state 0, given that the current state is \( i \).

Table 3 gives the mean \( (E_i[J_{15}]) \) and variance \( (V_i[J_{15}]) \) of the number of jumps since the last occurrence of bare rock (state 15) given the current state, for the continuous-time chain based on data in [4]. The Pearson correlation between the time since the last occurrence of bare rock \( (E_i[T_{15}]) \) and the number of jumps since the last occurrence of bare rock \( (E_i[J_{15}]) \) is 0.87 (Figure 5). In other words, states for which the expected time since the last occurrence of bare rock is longer are also expected to be separated from bare rock by a larger number of intermediate states. However, the differences among states in the expected number of jumps since the last occurrence of bare rock are small relative to the variance in number of jumps for any given state. From Table 3,
the difference between the largest and smallest values of $E_i(J_{15})$ is 1.48, while $V_i(J_{15})$ is at least 16.54 for all states.

7 Conclusions

Markov models for communities are usually formulated in discrete time. In many ecological communities, it is reasonable to assume that some events such as mortality can occur at any time, and even seasonal events such as reproduction may occur over quite long time periods. If this is the case, then a continuous-time model might be more appropriate for inferences about some kinds of properties than a discrete-time model, which predicts the state of the system only at discrete observation points. If the underlying dynamics of the system are in continuous time, a discrete-time model will miss many events that occur between observation intervals. Here, we showed that the expected time since the last occurrence of some past state (estimated from a continuous-time model) will be less than the expected time since the last observation of this state (estimated from a discrete-time model of the same system). Elsewhere, we showed that inferences about the complexity of interaction networks [11] and the consequences of changing interspecific interaction parameters [12] from discrete-time models may also be misleading if the true dynamics are in continuous time.

One important caveat is that we assumed homogeneity in time. This may not be true, especially in temperate habitats such as the one we studied, where it is likely that survival, reproduction and interspecific interactions will vary seasonally. There is also statistically significant interannual variability in transition probabilities in this system, although the biological consequences
of this variability may not be major [13]. We did not consider models with
temporal variability here because such models cannot be identified from time-
averaged data such as those presented in [4]. However, temporal variability
could potentially have large effects on inferences about past states. If the
sequences of states at points in space are available, these effects could be
checked by direct comparison of the calculated and observed distributions of
first passage times. Both time-averaged discrete-time models and homogeneous
continuous-time models are potentially useful caricatures of the dynamics of
a community, but it is important to remember their limitations.

We refer to states for which the expected time since the last occurrence of bare
rock is long as having high successional rank. For the Gulf of Maine subtidal
data set we analyzed, states with high successional rank tended to be good at
persisting at a point, and to be separated from bare rock by longer sequences
of events than states with low successional rank. However, other relationships
might occur in other communities.

Inferences about past states could be useful for estimating the age of a habitat
[9]. However, for the community we analyzed, the variance in time since the
last occurrence of bare rock given the current state was large relative to the
differences in expected time since the last occurrence of bare rock among
current states. In other words, for this community, knowing the current state
does not give us much information about the past. We did not consider the
consequences of uncertainty in transition probabilities [9], but such uncertainty
would be likely to further reduce the amount of information available about
the past.

Other communities might show quite different patterns in successional rank
from the ones we found, and might allow inferences about the past to be made
with greater certainty. Comparative studies of Markov models from a range
of communities [3,4] will be the best way to address these questions.

Acknowledgments

We are very grateful to Damian Clancy, Ian Nightingall, Pawel Paszek and
Andy Solow for discussions. This work was supported by a grant from the
University of Liverpool Research Development Fund, project number 6119.

References

   (1979) 413–426.
   1997.


Appendix A: time reversal and first passage times in continuous time

Here, we summarize the necessary results on time reversal and the distribution of first passage times for a stationary finite-state ergodic continuous-time homogeneous Markov chain at equilibrium. Our main sources are [10] and [5].

**Theorem 1** Let $q_{ij}$ be the instantaneous rate of transitions from state $i$ to state $j$ in a stationary finite-state ergodic continuous-time homogeneous Markov chain, and let $\pi_i$ be the stationary probability of state $i$. Then the instantaneous rate of transitions from $j$ to $i$ in the corresponding stationary time-reversed chain is $q_{ji}^R = (q_{ij}\pi_i)/\pi_j$ [5, Theorem 3.7.1].

**Proof** Let $X_t$ be the state of a stationary finite-state ergodic continuous-time Markov chain at time $t$. Then

$$P(X_t = i, X_{t+s} = j) = P(X_{t+s} = j|X_t = i)P(X_t = i)$$

$$= P(X_t = i|X_{t+s} = j)P(X_{t+s} = j) \tag{12}$$

for $s > 0$. 
Let the instantaneous rate of transitions from state $i$ to state $j$ be $q_{ij}$, and let the rate of leaving state $i$ be $q_i = \sum_{k \neq i} q_{ik} = -q_{ii}$. Then for a small time interval $h$,

$$P(X_{t+h} = i | X_t = i) \geq P(S_1 > h) \quad (13)$$

where $S_1$ is the holding time in state $i$ before the next event occurs. The inequality is because the RHS is the probability that no changes of state occur between $t$ and $t+h$, but we could also leave and return to state $i$ in two or more events.

For a process occurring at a constant rate $q_i$ per unit time, the expected number of events over time interval $h$ is $q_i h$. Divide $h$ into $N$ equal-length subintervals, and

$$P(\text{No events in interval } h) = (1 - \frac{q_i h}{N})^N \quad (14)$$

Then making the subintervals arbitrarily small

$$P(S_1 > h) = \lim_{N \to \infty} (1 - \frac{q_i h}{N})^N = e^{-q_i h} \quad (15)$$

so

$$P(X_{t+h} = i | X_t = i) \geq e^{-q_i h}$$

$$= 1 - q_i h + \frac{q_i^2 h^2}{2!} - \ldots \quad (16)$$

$$= 1 - q_i h + o(h)$$

where $o(h)$ is any function $\phi(h)$ such that $h^{-1}\phi(h) \to 0$ as $h \downarrow 0$.

Similarly, if $j \neq i$,

$$P(X_{t+h} = j | X_t = i) \geq P(S_1 \leq h, Y_1 = j, S_2 > h) \quad (17)$$

where $Y_1$ is the state of the chain after the first event, given that the initial state is $i$. The inequality is because the RHS ignores the possibility of ending in
by starting in \( i \) and undergoing more than one transition, and the possibility that only one event happens because \( S_1 \leq h, S_2 \leq h \) but \( S_1 + S_2 > h \). The probability that the new state is \( j \) after one event, given that the initial state is \( i \), is \( q_{ij}/q_i \), so

\[
P(X_{t+h} = j|X_t = i) \geq (1 - e^{-q_i h}) \frac{q_{ij}}{q_i} e^{-q_j h}
\]

(18)

\[
= q_{ij}h + o(h)
\]

(note that \( o(h) + o(h) + \ldots = o(h) \) for a finite number of terms). Thus for any \( j \)

\[
P(X_{t+h} = j|X_t = i) \geq \delta_{ij} + q_{ij}h + o(h) \tag{19}
\]

where \( \delta_{ij} \) is the Kronecker delta function (1 if \( i = j \), 0 otherwise).

Summing over all of the finite set of possible states,

\[
\sum_j P(X_{t+h} = j|X_t = i) \geq \sum_j \delta_{ij} + q_{ij}h + o(h)
\]

\[
1 = 1 + \left[ \sum_{j \neq i} q_{ij} - \sum_{k \neq i} q_{ik} \right] h + o(h) \tag{20}
\]

The LHS is 1 because we are summing over all possible states at \( t + h \), so letting \( h \downarrow 0 \), the inequality in Equation 19 must be an equality:

\[
P(X_{t+h} = j|X_t = i) = \delta_{ij} + q_{ij}h + o(h) \tag{21}
\]

[5, Theorem 2.8.2]. Substituting Equation 21 into Equation 12 gives

\[
(q_{ij}h + o(h)) P(X_t = i) = (q_{ji}^R h + o(h)) P(X_{t+h} = j) \tag{22}
\]

where \( q_{ji}^R \) is the instantaneous rate of transitions from \( j \) to \( i \) in the time-reversed process. The \( \delta_{ij} \) term in Equation 21 disappears because if \( i \neq j \) then \( \delta_{ij} = 0 \), and if \( i = j \), we can subtract \( \delta_{ij} P(X_t = i) \) from both sides. Dividing
by $h$, letting $h \downarrow 0$ and rearranging,

$$q_{ji}^R = \frac{q_{ij}P(X_t = i)}{P(X_{t+h} = j)}$$

(23)

If the chain is at equilibrium, we can substitute the stationary probabilities $\pi_i$ and $\pi_j$ of states $i$ and $j$ into Equation 23. Then the transition rates in the time-reversed process at equilibrium are

$$q_{ji}^R = \frac{q_{ij}\pi_i}{\pi_j}$$

(24)

**Theorem 2** Let $T_j$ be the first passage time to state $j$ in a stationary finite-state ergodic continuous-time homogeneous Markov chain. Without loss of generality we assume $j = 0$. Let $Q = \begin{bmatrix} -q_0 & Q_0^T \\ r & T \end{bmatrix}$ be the instantaneous rate matrix for the chain. Then the probability density of the first passage time to state 0 from initial state $i$ is given by

$$f_{i0}(t) = e_i e^{Tr}$$

for $t \geq 0$, where $e_i$ is a vector with 1 in position $i$ and 0 elsewhere [10, p. 213].

**Proof** Construct a chain from the original chain in which state 0 is absorbing, with generator

$$Q_0 = \begin{bmatrix} 0 & 0^T \\ r & T \end{bmatrix}$$

where $r = (q_i)$. This continuous-time Markov chain satisfies the forward Kol-
mogorov equation

\[ P'_0(t) = P_0(t)Q_0, \quad P_0(0) = I \]  \hspace{1cm} (26)

[5, Theorem 2.8.2], where \( P_0(t) \) is a matrix with elements \( p_{0,j}(t) = P(X_t = j|X_0 = i) \) in the absorbing chain generated by \( Q_0 \). Because state 0 is absorbing, \( p_{0,0}(t) \) is the cumulative distribution function \( F_0(t) \) for \( T_0 \) with initial state \( i \), and \( p'_{0,0}(t) \) is the corresponding probability density. Also, \( P_0(t) = e^{Q_0t} \), where \( e^{Q_0t} \) is the matrix exponential, \( \sum_{k=0}^{\infty} \frac{Q_0^k t^k}{k!} \) [5, Theorem 2.1.1]. Note that

\[
Q_0^k = \begin{bmatrix} 0 & 0^T \\ T^{k-1}r & T^k \end{bmatrix}
\]

Then

\[
e^{Q_0t} = I + \sum_{k=1}^{\infty} \frac{t^k}{k!} \begin{bmatrix} 0 & 0^T \\ T^{k-1}r & T^k \end{bmatrix} = \begin{bmatrix} 1 & 0^T \\ r(t) e^{T^t} \end{bmatrix}
\]

where

\[
r(t) \equiv \sum_{k=1}^{\infty} \frac{T^{k-1}t^k}{k!}r = T^{-1} \sum_{k=1}^{\infty} \frac{T^k t^k}{k!}r = T^{-1}(e^{T^t} - I)r
\]

[10, p. 209]. Thus in the modified chain, \( P(X_t = j|X_0 = i) = [e^{T^t}]_{ij} \), where \( i, j \in \mathcal{T} \), the set of transient states, and

\[
\frac{d}{dt}(P(X_t = 0|X_0 = i)) = \sum_{k \in \mathcal{T}} p_{ik}(t)q_{k0} = e_i e^{T^t}r
\]

\[\square\]

**Theorem 3** The first passage time \( T_0 \) to state 0 from initial state \( i \) in a stationary finite-state ergodic continuous-time homogeneous Markov chain has
mean $E_i[T_0] = e_i(-T)^{-1}1$ [10, p. 212] and variance $V_i[T_0] = 2e_i(-T)^{-2}1)$ – $(e_i(-T)^{-1}1)^2$ [14], where $T$ and $e_i$ are defined as in Theorem 2 and the subscripts $i$ in $E_i[T_0]$ and $V_i[T_0]$ indicate that they are conditional on initial state $i$.

**Proof** The Laplace-Stieltjes transform of $f_{i0}(t)$ is

$$
\phi_i(s) = \int_0^\infty e^{-st}df_{i0}(t)
$$

Let $\phi(s) = (\phi_i(s))$. Then

$$
\phi(s) = (sI - T)^{-1}r
$$

[10, Theorem 4.26]. Differentiating with respect to $s$,

$$
\phi'(s) = -(sI - T)^{-2}r
$$

For a random variable $X$, $\phi_i(-s)$ is equivalent to the moment generating function $M(s)$, and $E[X] = M'(0)$ [15, p. 146], so $\phi'_i(0) = -E[X]$. Now $r = -T1$ because the row sums of $T$ are zero, so setting $s = 0$ and using $e_i$ to select the $i$th element in Equation 33 gives $E_i[T_0]$, the expectation of $T_0$ conditional on initial state $i$:

$$
E_i[T_0] = e_i(-T)^{-2}r = e_i(-T)^{-2}(T1)1 = e_i(-T)^{-1}1
$$

Differentiating Equation 33 again and evaluating at $s = 0$ gives the second moment,

$$
E_i[T_0^2] = 2e_i(-T)^{-2}1
$$

Then $V_i[T_0] = E_i[T_0^2] - (E_i[T_0])^2 = 2e_i(-T)^{-2}1) - (e_i(-T)^{-1}1)^2$. □
Theorem 4 Let $F^d_{i0}(t)$ be the cumulative probability that a stationary finite-state ergodic continuous-time homogeneous Markov chain generated by $Q$ and with initial state $i$, observed at discrete times $1, 2, \ldots, t$, has been observed in state 0. Let $F_{i0}(t)$ be the cumulative probability that the same chain has visited state 0, whether observed or not. Then $F_{i0}(t) > F^d_{i0}(t)$ for all $0 < t < \infty$.

Proof Let $X_k$ be the state of the chain at time $k$.

$$F^d_{i0}(t) = 1 - \prod_{k=1}^{|t|} (1 - P(X_k = 0))$$

$$= 1 - \exp \sum_{k=1}^{|t|} \log(1 - P(X_k = 0))$$

and

$$F_{i0}(t) = 1 - \exp \int_0^t \log(1 - P(X_k = 0)) dk$$

Because the chain is ergodic, $P(X_k = 0) > 0$ for all $k > 0$. Thus

$$\int_0^t \log(1 - P(X_k = 0)) dk < \sum_{k=1}^{|t|} \log(1 - P(X_k = 0))$$

for $0 < t < \infty$. \square

Theorem 5 For a stationary finite-state ergodic continuous-time homogeneous Markov chain, the expected first passage time from state $i \neq 0$ to state 0 is less than the expected first time at which state 0 is observed out of the discrete observation times $1, 2, \ldots, \infty$.

Proof For a random variable $X$ with density $f(x) = 0$ for $x < 0$ and distribution function $F$,

$$E[X] = \int_0^\infty (1 - F(x)) dx$$

[16, Lemma 4.3.4]. Let $E_i[T_0]$ be the expected first passage time from state $i$ to state 0, and $E_i[T^d_0]$ be the expected time to the first observation of state 0.
for a chain observed only at discrete times 1, 2, \ldots, \infty. Then
\[ E_i[T_0^d] - E_i[T_0] = \int_0^\infty (1 - F_{i0}^d(t))dt - \int_0^\infty (1 - F_{i0}(t))dt \]
\[ = \int_0^\infty (F_{i0}(t) - F_{i0}^d(t))dt \quad (40) \]

By Theorem 4, \( F_{i0}(t) > F_{i0}^d(t) \) for all 0 < t < \infty. Thus \( E_i[T_0] < E_i[T_0^d] \).

**Theorem 6** Let \( S_0 = \begin{bmatrix} 1 & 0^T \\ r & T \end{bmatrix} \) be a jump matrix in which state 0 is absorbing.

In the chain defined by \( S_0 \), the number of jumps \( J_0 \) from state \( i \) to state 0 has expectation \( E_i[J_0] = e_i N_1 + E_i[J_0] - (E_i[J_0])^2 \), where \( N = (I - T)^{-1} \) and \( e_i \) is a vector with a 1 in position \( i \) and 0 elsewhere [10, pp. 80-81].

**Proof** The generating function of a non-negative discrete random variable \( X \) is
\[ g_X(z) = \sum_{n=0}^{\infty} p_n z^n \quad (41) \]
where \( p_n = P(X = n) \). If \(|z| < 1\), the series converges and we can differentiate with respect to \( z \), giving
\[ g_X'(z) = \sum_{n=1}^{\infty} np_n z^{n-1} \quad (42) \]
Thus
\[ g_X'(1) = \sum_{n=1}^{\infty} np_n = E[X] \quad (43) \]
Differentiating again with respect to \( z \),
\[ g_X''(z) = \sum_{n=1}^{\infty} n^2 p_n z^{n-2} - np_n z^{n-2} \quad (44) \]
When \( z = 1 \), we have
\[ g_X''(1) = \sum_{n=1}^{\infty} n^2 p_n - np_n = E[X^2] - E[X] \quad (45) \]
Thus
\[ V[X] = E[X^2] - (E[x])^2 = g_X''(1) + g_X'(1) - (g_X'(1))^2 \] (46)

We now apply these ideas to the jump chain.

Let \( f(n) \) be a vector of probabilities \( f_{io}(n) \) that, starting from any state \( i \neq 0 \), the jump chain first reaches state 0 after \( n \) jumps. The \( n \)-jump transition probabilities are

\[ S_n^0 = \begin{bmatrix} 1 & 0^T \\ r_n & T^n \end{bmatrix} \] (47)

where \( r_n = (I+T+\ldots T^{n-1})r \). Because state 0 is absorbing, \( f(n) = r_n - r_{n-1} = T^{n-1}r \). The vector of generating functions of \( f_{io}(n) \) is

\[ f^s(z) = z(I - zT)^{-1}r, \quad |z| < 1 \] (48)

[10, Theorem 2.19]. Differentiating Equation 48 with respect to \( z \) and setting \( z = 1 \),

\[ f''(1) = (I - T)^{-1}r + (I - T)^{-1}T(I - T)^{-1}r \]
\[ = (I - T)^{-1}[I + T(I - T)^{-1}]r \] (49)

Because state 0 is absorbing and can be reached from any state \( i \neq 0 \), \( T^n \to 0 \) componentwise as \( n \to \infty \). Then \( (I - T)^{-1} = \sum_{n=0}^{\infty} T^n \) [10, Lemma A.1]. Thus \( T(I - T)^{-1} = T \sum_{n=0}^{\infty} T^n = \sum_{n=1}^{\infty} T^n \), so \( I + T(I - T)^{-1} = \sum_{n=0}^{\infty} T^n = (I - T)^{-1} \). Also, because the row sums of \( S_0 \) are 1, \( r = (I - T)1 \). Then Equation 49 becomes

\[ f''(1) = (I - T)^{-2}r \]
\[ = (I - T)^{-1}1 \]
\[ = N1 \] (50)

The \( i \)th element of this vector is \( e_i N1 = E_i[J_0] \), from Equation 43.
Differentiating Equation 48 twice with respect to $z$ and setting $z = 1$,

\[ f^{*''}(1) = (I - T)^{-1}[T(I - T)^{-1} + (I - T)^{-1}T(I - T)^{-1} + T(I - T)^{-1}T(I - T)^{-1}]r \]

\[ = 2(I - T)^{-2}(I - T)^{-1} - I)r \]

\[ = 2N^2[N - I]r \quad (51) \]

Then using Equation 46,

\[ V_i[J_0] = 2e_iN^2(N - I)r + E_i[J_0] - (E_i[J_0])^2 \quad (52) \]
Appendix B: estimating the continuous-time model

Assume that each of a set of independent fixed points in space has a state determined by the same finite-state homogeneous Markov chain. If the interval between observations is $t$ for all points, the maximum likelihood estimate of a transition probability $p_{ij}$ is $\hat{p}_{ij} = n_{ij}/n_i$, where $n_{ij}$ is the number of points in state $j$ at time $\tau + t$ that were in state $i$ at time $\tau$, and $n_i$ is the number of points in state $i$ at time $\tau$ [17, p. 135].

We used data from Table 2 in [4], derived from annual photographic observations of a rocky subtidal community over nine years. Including bare rock, there were $s = 15$ distinguishable states, labelled 1 to 15. Note that in [4], the transition probability matrix is transposed with respect to our notation. The rows of the matrix in Table 2 of [4], which we refer to as $\tilde{P}(t)$, did not exactly sum to 1, presumably because of rounding errors. We therefore rescaled all elements by the row sums to obtain the estimate $\hat{P}(t)$:

$$\hat{p}_{ij} = \tilde{p}_{ij} / \sum_{k=1}^{s} \tilde{p}_{ik}$$

(53)

If $P(t) = e^{Qt}$, an obvious estimate of the continuous-time rate matrix $Q$ is $\tilde{Q} = 1/t \log \tilde{P}(t)$, where log is the matrix logarithm. This is the maximum likelihood estimate if $P(t)$ was generated by a homogeneous continuous-time Markov chain [18]. However, this estimate often has negative off-diagonal entries, either because of sampling error or because the transition probabilities were not generated by a homogeneous continuous-time Markov chain. If $\tilde{Q}$ has negative off-diagonals, there will be some time intervals for which $e^{\tilde{Q}t}$ also has negative entries, so this cannot be a valid rate matrix. If the original data are available, constrained maximum likelihood [19,11] or Markov Chain Monte Carlo [18] would be appropriate methods.
methods can be used. However, if only $\hat{P}(t)$ is available, as in this case, we could set any negative off-diagonal entries in $\hat{Q}$ to zero and adjust the diagonal elements to maintain zero row sums, giving the estimate $\hat{Q}$:

$$
\hat{q}_{ij} = \begin{cases} 
\max(\tilde{q}_{ij}, 0) & i \neq j \\
\tilde{q}_{ii} + \sum_{k \neq i} \min(\tilde{q}_{ik}, 0) & i = j
\end{cases}
$$

(54)

[20]. Other approaches are possible when only $\hat{P}(t)$ is available, such as finding a rate matrix whose eigenspectrum matches that of the transition probability matrix as closely as possible [21]. However, for the data we used, $e^{\hat{Q}t}$ was very close to $\hat{P}(t)$, and gave almost indistinguishable results for the distribution of first passage times to empty space. $\hat{Q}$ is therefore a satisfactory estimate of the continuous-time rate matrix for our purposes. All the discrete-time results we report are based on $e^{\hat{Q}t}$, so that differences between discrete- and continuous-time results are not artefacts of the estimation method for the continuous-time model. Table 1 shows the estimate $\hat{Q}$. 

26
Table 1


<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.2925</td>
<td>0.1206</td>
<td>0.0170</td>
<td>0.0030</td>
<td>0.0137</td>
<td>0</td>
<td>0.0423</td>
<td>0.0193</td>
<td>1.54e-04</td>
<td>0.0279</td>
<td>0.0012</td>
<td>0.0011</td>
<td>0.0081</td>
<td>0.0069</td>
<td>0.0312</td>
</tr>
<tr>
<td>2</td>
<td>0.1821</td>
<td>-0.5889</td>
<td>0.0391</td>
<td>0.0128</td>
<td>0.0332</td>
<td>0.0052</td>
<td>0.0479</td>
<td>0.0599</td>
<td>0.0147</td>
<td>0.0275</td>
<td>0.0172</td>
<td>0.0103</td>
<td>0.0071</td>
<td>0.0062</td>
<td>0.1260</td>
</tr>
<tr>
<td>3</td>
<td>0.0537</td>
<td>0.0216</td>
<td>-0.3553</td>
<td>0.0032</td>
<td>0.0202</td>
<td>0.0043</td>
<td>2.52e-06</td>
<td>0.0078</td>
<td>0.0361</td>
<td>0</td>
<td>0.0051</td>
<td>0.0081</td>
<td>0.0016</td>
<td>0.0281</td>
<td>0.1655</td>
</tr>
<tr>
<td>4</td>
<td>0.0098</td>
<td>0.0510</td>
<td>0.0049</td>
<td>-0.1776</td>
<td>0.0010</td>
<td>0</td>
<td>7.61e-04</td>
<td>0.0084</td>
<td>0.0100</td>
<td>0.0042</td>
<td>0.0076</td>
<td>0.0148</td>
<td>0</td>
<td>0.0086</td>
<td>0.0565</td>
</tr>
<tr>
<td>5</td>
<td>0.1290</td>
<td>0.3239</td>
<td>0.0455</td>
<td>2.50e-04</td>
<td>-0.9734</td>
<td>0.0093</td>
<td>0.0941</td>
<td>0.0834</td>
<td>0.0177</td>
<td>0</td>
<td>9.51e-04</td>
<td>0.0017</td>
<td>0.0083</td>
<td>0.0031</td>
<td>0.2561</td>
</tr>
<tr>
<td>6</td>
<td>0.0039</td>
<td>0.0105</td>
<td>0.0131</td>
<td>0</td>
<td>0.0219</td>
<td>-0.1595</td>
<td>0</td>
<td>0.0117</td>
<td>0.0212</td>
<td>0.0032</td>
<td>0.0048</td>
<td>0.0088</td>
<td>0</td>
<td>0</td>
<td>0.0603</td>
</tr>
<tr>
<td>7</td>
<td>0.5416</td>
<td>0.3596</td>
<td>0.1207</td>
<td>0.0326</td>
<td>0.2628</td>
<td>0.0645</td>
<td>-2.5204</td>
<td>0.2331</td>
<td>0.0263</td>
<td>0.1729</td>
<td>0.0864</td>
<td>0</td>
<td>0.0644</td>
<td>0.1467</td>
<td>0.4147</td>
</tr>
<tr>
<td>8</td>
<td>0.3483</td>
<td>0.4066</td>
<td>0.0609</td>
<td>0.0327</td>
<td>0.2681</td>
<td>0.0075</td>
<td>0.2215</td>
<td>-2.0042</td>
<td>0.0278</td>
<td>0.0966</td>
<td>0.0317</td>
<td>0.0530</td>
<td>0.0218</td>
<td>0.0522</td>
<td>0.3754</td>
</tr>
<tr>
<td>9</td>
<td>0.0346</td>
<td>0.1750</td>
<td>0.0305</td>
<td>0.0122</td>
<td>0.0169</td>
<td>0.0068</td>
<td>5.70e-04</td>
<td>0.0612</td>
<td>-0.5435</td>
<td>0.0472</td>
<td>0.0053</td>
<td>0.0036</td>
<td>0</td>
<td>0.0133</td>
<td>0.1364</td>
</tr>
<tr>
<td>10</td>
<td>0.6604</td>
<td>0.4312</td>
<td>0.0428</td>
<td>0.0123</td>
<td>0.0049</td>
<td>0.0065</td>
<td>0.2063</td>
<td>0.1055</td>
<td>0.0149</td>
<td>-1.8940</td>
<td>0.0205</td>
<td>0.0111</td>
<td>0.0370</td>
<td>0.0090</td>
<td>0.3316</td>
</tr>
<tr>
<td>11</td>
<td>0.0245</td>
<td>0.3324</td>
<td>0.0315</td>
<td>0.0077</td>
<td>0.0500</td>
<td>0</td>
<td>0.0601</td>
<td>0.0349</td>
<td>0.0060</td>
<td>0</td>
<td>-0.6977</td>
<td>0</td>
<td>0.0068</td>
<td>0</td>
<td>0.1438</td>
</tr>
<tr>
<td>12</td>
<td>0.0201</td>
<td>0</td>
<td>0.0213</td>
<td>0.0420</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0169</td>
<td>5.94e-04</td>
<td>0</td>
<td>0</td>
<td>-0.8073</td>
<td>0</td>
<td>0</td>
<td>0.7063</td>
</tr>
<tr>
<td>13</td>
<td>0.6096</td>
<td>0.2839</td>
<td>0.0230</td>
<td>0</td>
<td>0.1546</td>
<td>0</td>
<td>0.0936</td>
<td>0.0347</td>
<td>0</td>
<td>0.0540</td>
<td>0.0342</td>
<td>0</td>
<td>-1.4746</td>
<td>0</td>
<td>0.1870</td>
</tr>
<tr>
<td>14</td>
<td>0.2815</td>
<td>1.9948</td>
<td>0</td>
<td>0.0488</td>
<td>0.4633</td>
<td>0.0075</td>
<td>0.0938</td>
<td>0</td>
<td>0.0547</td>
<td>0.2226</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-3.9231</td>
<td>0.7561</td>
</tr>
<tr>
<td>15</td>
<td>0.0281</td>
<td>0.6355</td>
<td>0.0191</td>
<td>0.0060</td>
<td>0.1204</td>
<td>0.0015</td>
<td>0.2224</td>
<td>0.0080</td>
<td>0.0177</td>
<td>0.1023</td>
<td>0.0285</td>
<td>0.0365</td>
<td>0.0298</td>
<td>0.0062</td>
<td>-1.4418</td>
</tr>
</tbody>
</table>
Table 2
Mean \(E_i[T_{15}]\) and variance \(V_i[T_{15}]\) of first passage time in years to state 15 (bare rock) from each other current state \(i\) in the time-reversed stationary continuous-time Markov chain based on the data in [4].

<table>
<thead>
<tr>
<th>Current state (i)</th>
<th>(E_i[T_{15}])</th>
<th>(V_i[T_{15}])</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (\textit{Hymedesmia} sp. 1)</td>
<td>10.60</td>
<td>90.33</td>
</tr>
<tr>
<td>2 (\textit{Crisia eburnea})</td>
<td>7.49</td>
<td>78.09</td>
</tr>
<tr>
<td>3 (\textit{Myxilla fimbriata})</td>
<td>10.54</td>
<td>91.08</td>
</tr>
<tr>
<td>4 (\textit{Mycale lingua})</td>
<td>13.00</td>
<td>112.38</td>
</tr>
<tr>
<td>5 (\textit{Filograna implexa})</td>
<td>7.04</td>
<td>76.04</td>
</tr>
<tr>
<td>6 (\textit{Urticina crassicornis})</td>
<td>13.34</td>
<td>117.97</td>
</tr>
<tr>
<td>7 (\textit{Ascidia callosa})</td>
<td>6.32</td>
<td>74.63</td>
</tr>
<tr>
<td>8 (\textit{Aplidium pallidum})</td>
<td>7.13</td>
<td>78.49</td>
</tr>
<tr>
<td>9 (\textit{Hymedesmia} sp. 2)</td>
<td>9.48</td>
<td>89.84</td>
</tr>
<tr>
<td>10 (\textit{Idmidronea atlantica})</td>
<td>6.98</td>
<td>78.89</td>
</tr>
<tr>
<td>11 (coralline algae)</td>
<td>7.67</td>
<td>78.40</td>
</tr>
<tr>
<td>12 (\textit{Metridium senile})</td>
<td>7.18</td>
<td>79.47</td>
</tr>
<tr>
<td>13 (\textit{Parasmittina jeffreysi})</td>
<td>7.09</td>
<td>78.28</td>
</tr>
<tr>
<td>14 (\textit{Spirorbis spirorbis})</td>
<td>5.41</td>
<td>70.31</td>
</tr>
</tbody>
</table>
Table 3
Mean ($E_i[J_{15}]$) and variance ($V_i[J_{15}]$) of number of state changes (dimensionless) separating each other current state $i$ from state 15 (bare rock) in the time-reversed stationary continuous-time Markov chain based on the data in [4].

<table>
<thead>
<tr>
<th>Current state $i$</th>
<th>$E_i[J_{15}]$</th>
<th>$V_i[J_{15}]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ($Hymedesmia$ sp. 1)</td>
<td>5.42</td>
<td>17.64</td>
</tr>
<tr>
<td>2 ($Crisia eburnea$)</td>
<td>4.35</td>
<td>17.38</td>
</tr>
<tr>
<td>3 ($Myxilla fimbriata$)</td>
<td>5.45</td>
<td>17.91</td>
</tr>
<tr>
<td>4 ($Mycale lingua$)</td>
<td>5.35</td>
<td>17.80</td>
</tr>
<tr>
<td>5 ($Filograna implexa$)</td>
<td>4.63</td>
<td>17.54</td>
</tr>
<tr>
<td>6 ($Urticina crassicornis$)</td>
<td>5.42</td>
<td>17.68</td>
</tr>
<tr>
<td>7 ($Ascidia callosa$)</td>
<td>4.41</td>
<td>17.41</td>
</tr>
<tr>
<td>8 ($Aplidium pallidum$)</td>
<td>4.86</td>
<td>17.76</td>
</tr>
<tr>
<td>9 ($Hymedesmia$ sp. 2)</td>
<td>5.23</td>
<td>18.13</td>
</tr>
<tr>
<td>10 ($Idmidronea atlantica$)</td>
<td>4.68</td>
<td>17.79</td>
</tr>
<tr>
<td>11 (coralline algae)</td>
<td>4.68</td>
<td>17.50</td>
</tr>
<tr>
<td>12 ($Metridium senile$)</td>
<td>4.26</td>
<td>17.12</td>
</tr>
<tr>
<td>13 ($Parasmittina jeffreysi$)</td>
<td>4.73</td>
<td>17.83</td>
</tr>
<tr>
<td>14 ($Spirorbis spirobis$)</td>
<td>3.97</td>
<td>16.54</td>
</tr>
</tbody>
</table>
Fig. 1. Cumulative distributions of first passage time \(T_{15}\), years to state 15 (bare rock) from state 14 (\textit{Spororbis spirorbis}) in the time-reversed stationary discrete–time Markov chains based on the data in [4], with observation intervals 1 (solid line), 2 (broken line) and 5 (dash-dot line) years.
Fig. 2. Cumulative distributions of first passage time ($T_{15}$, years) to state 15 (bare rock) from each current state $i$ (panel numbers) in the time-reversed stationary continuous-time (broken lines) and discrete-time (solid lines) Markov chains based on the data in [4]. See Table 2 for state names.
Fig. 3. Relationships between (a) stationary probability ($\pi_i$, dimensionless) and expected first passage time ($E_i[T_{15}]$, years) to state 15 (bare rock), (b) rate of leaving ($-q_{ii}$, years$^{-1}$) and expected first passage time to state 15, (c) stationary probability and rate of leaving, for all states $i$ other than bare rock in the time-reversed stationary continuous-time Markov chain based on the data in [4].
Fig. 4. Successional patterns where: (a) differences in successional rank result only from differences in the rate of colonization of bare rock (state \( j \)); and (b) differences in successional rank result from differences in the number of events separating a state from bare rock.

![Diagram a](image-a)

![Diagram b](image-b)
Fig. 5. Relationship between expected number of jumps ($E_i[J_{15}]$, dimensionless) and expected time ($E_i[T_{15}]$, years) to state 15 (bare rock), for all states $i$ other than bare rock, in the time-reversed continuous-time Markov chain based on data in [4].