

1 **Lotka-Volterra competition models for sessile organisms**

2 Matthew Spencer

3 School of Biological Sciences, University of Liverpool, Liverpool, UK

4 `m.spencer@liverpool.ac.uk`

5 Jason E. Tanner

6 SARDI Aquatic Sciences, Henley Beach, Australia and School of Earth and Environmental
7 Sciences, University of Adelaide, Adelaide, Australia

8 `Tanner.Jason@saugov.sa.gov.au`

9 Received _____; accepted _____

Short title: models for sessile organisms

ABSTRACT

10

11

Markov models are widely used to describe the dynamics of communities of sessile organisms, because they are easily fitted to field data and provide a rich set of analytical tools. In typical ecological applications, at any point in time, each point in space is in one of a finite set of states (e.g. species, empty space). The models aim to describe the probabilities of transitions between states. In most Markov models for communities, these transition probabilities are assumed to be independent of state abundances. This assumption is often suspected to be false, and is rarely justified explicitly. Here, we start with simple assumptions about the interactions among sessile organisms, and derive a model in which transition probabilities depend on the abundance of destination states. This model is formulated in continuous time and is equivalent to a Lotka-Volterra competition model. We fit this model and a variety of alternatives in which transition probabilities do not depend on state abundances to a long-term coral reef data set. The Lotka-Volterra model describes the data much better than all models we consider other than a saturated model (a model with a separate parameter for each transition at each time interval, which by definition fits the data perfectly). Our approach provides a basis for further development of stochastic models of sessile communities, and many of the methods we use are relevant to other types of community. We discuss possible extensions to spatially explicit models.

12

Subject headings: Markov models, coral reefs, Lotka-Volterra competition, maximum

13

likelihood, statistical modelling

1. Introduction

14

15 Markov models have been applied to data on a wide range of sessile communities,
16 containing organisms such as trees, mussels and corals (Usher 1979; Tanner et al. 1994;
17 Wootton 2001*a*; Hill et al. 2004), with the aim of understanding the structure and function
18 of these communities. The influential Roughgarden et al. (1985) model of an open
19 population with space-limited recruitment is also a linear Markov model in its original form,
20 although this not immediately obvious.

21 The popularity of Markov models stems from their relatively simple structure, and
22 the ease with which they can be parameterized from data obtained by repeat surveys of
23 permanent quadrats, which are a mainstay in marine ecology. There are also a wide range
24 of tools that have been developed for analysis of matrix population models (Caswell 2001),
25 many which can equally be applied to Markov models of communities. While Markov
26 models are very simple, in at least some cases they have proven to have surprisingly good
27 predictive ability (Wootton 2004).

28 In the Markov models considered here, a fixed point in space can be in one of a finite
29 set of possible states (e.g. species or groups of species) at any given time (although Markov
30 models with an infinite set of possible states also exist). Conditional on the current state,
31 future states are independent of the past. In most cases, Markov models of communities
32 are formulated in discrete time, with model time intervals based on the census intervals in
33 the observed data. The relevant data are the frequencies of transitions from each state to
34 each other state over a given time interval, and the parameters are the probabilities of these
35 transitions.

36 Several elaborations of the basic first order discrete time Markov models generally
37 used in ecology have been proposed and demonstrated. These include second-order models,
38 where transitions depend on the state over the previous two time intervals, and semi-Markov

39 models, where transitions depend on the length of time that a point has remained in its
40 current state (Tanner et al. 1996), as well as a continuous time version (Spencer and Susko
41 2005). It has often been suggested that the probabilities of transitions between states in
42 Markov models of sessile communities might depend on densities (Usher 1979; Tanner
43 et al. 1994, 1996; Hill et al. 2002). However, there have been few attempts to incorporate
44 density-dependence into these models. Acevedo (1981) studied the effects of density
45 dependence on simple models of forest dynamics. Nonlinear open-population models have
46 been developed, in which either mortality (Roughgarden et al. 1985) or growth (Muko et al.
47 2001*a*) depend on the amount of free space. There is empirical evidence for such effects in
48 some but not all populations to which they have been applied (Gaines and Roughgarden
49 1985; Hyder et al. 2001; Muko et al. 2001*b*; Svensson et al. 2004). Caswell and Cohen
50 have developed a number of nonlinear metapopulation models for two-species competition
51 (e.g. Caswell and Cohen 1995). More recently, J. E. Tanner et al. (in preparation) have
52 examined a more comprehensive density-dependent model of coral community dynamics,
53 and found that it increased coral cover at equilibrium, compared to a model without density
54 dependence.

55 Most of these density-dependent models are formulated in discrete time (exceptions
56 include Muko et al. 2001*a,b*). However, if transitions from one state to another may happen
57 at any time, the transition probabilities over a finite time interval are the net outcome of
58 all possible sequences of events in that interval. In most cases, a change in the rate of
59 any single transition will then affect all transition probabilities (Spencer 2006). We would
60 therefore expect every transition probability to be affected by the abundance of every state,
61 which leads to complicated models unless the number of states is very small. It is simpler
62 to construct continuous-time density-dependent models if the biological interactions occur
63 in continuous time. Here, we develop a continuous-time density-dependent model, based on
64 a simple probabilistic view of interspecific interactions among sessile organisms. It turns

65 out that this is a Lotka-Volterra competition model. Using maximum likelihood methods,
66 we compare the fit of continuous-time models with and without density dependence to
67 time series from a coral reef. In addition, we compare these models to time-averaged and
68 saturated discrete-time models (a saturated model is one with a separate parameter for each
69 transition probability at each time interval, which by definition fits the data perfectly). We
70 show that the density-dependent model performs much better than all but the saturated
71 model. Furthermore, this improved performance is achieved by changing the form of the
72 model, not by adding more parameters.

73 The models we evaluate are not fundamentally novel. However, in theoretical ecology,
74 the effort expended on developing new models far exceeds that expended on evaluating
75 how well these models describe real data. We agree with the idea that “the confrontation
76 between more than one model arbitrated by the data underlies science” (Hilborn and Mangel
77 1997, p. xv). As far as we are aware, this is the first attempt to formally compare the fit of
78 linear Markov models and nonlinear models to a field data set, although comparisons have
79 been made between the results of different linear models (e.g. Tanner et al. 1996).

80 2. The model

81 2.1. Assumptions

82 We describe most of the features of the model in terms of colonial organisms such as
83 corals, but the same approach can be applied to other kinds of sessile organisms such as
84 trees. We make several important assumptions:

- 85 1. That there is a fixed and finite number of possible states for a point in space. Let s
86 be the number of such states. One of these, e , is empty space, and the others may be
87 either species or groups of species that we choose not to distinguish (either because

- 88 this is too difficult or because they are of secondary interest).
- 89 2. That conditional on the current state, future states are independent of the past (the
90 Markov assumption). This is not strictly true. For example: in some species, larger
91 colonies are more likely to be dislodged by storms (Tanner et al. 1996; Madin and
92 Connolly 2006); reproduction depends on colony size in many corals (Harrison and
93 Wallace 1990); and competition between adjacent colonies may be size-dependent
94 (Lang and Chornesky 1990). Nevertheless, including historical effects in an empirical
95 model of a reef system had little effect on community dynamics (Tanner et al. 1996),
96 so violations of the Markov assumption may not be very important.
- 97 3. That the rate at which transitions occur from state j into some non-empty state i
98 depends on the availability of propagules or colonies of i to colonize or overgrow j , and
99 that this availability depends on the proportion of i in the system. This assumption
100 distinguishes our model from the usual homogeneous Markov chain, in which the
101 rates of transitions are defined by a constant generator matrix (e.g. Norris 1997,
102 section 2.1). This assumption can be tested by fitting models in which transition
103 rates are either dependent on or independent of the proportion of i , and using model
104 comparison methods to select the better model, as described in Appendix A4. Our
105 treatment of density dependence seems natural for situations where colony growth is
106 the main mechanism by which transitions to a non-empty state occur. In support
107 of this view, J. E. Tanner et al. (in preparation) show that for the same data set,
108 transition probabilities from j to i in discrete time depend on the proportion of i
109 much more frequently than on the proportion of j .
- 110 4. That the rate of clearance of points occupied by some species j is independent of
111 the proportion of empty space in the system. Clearance might occur by external
112 disturbances or because colonies of j die, and results in a transition to empty space.

113 Although J. E. Tanner et al. (in preparation) found that discrete-time transition
114 probabilities to empty space did depend on the proportion of empty space for some
115 species groups, these relationships were relatively weak. We assume that organisms do
116 not interact such that a colony of i kills a colony of j but does not occupy the resulting
117 space. Such interactions are biologically plausible, for example by allelopathy, but
118 require more complex models.

119 5. That interaction coefficients are constant over time. In particular, we are ignoring
120 seasonal effects. Seasonality may be important in some systems, and can in principle
121 be incorporated into models of the kind we develop. This might be useful because
122 some aspects of coral demography, such as annual mass-spawning (Harrison and
123 Wallace 1990), can be strongly seasonal. However, colonization of free space in
124 this system is overwhelmingly dominated by growth of existing colonies rather than
125 recruitment, and the transition rates do not separate recruitment from colony growth.
126 Mortality due to cyclones is also likely to be seasonal. However, cyclones are rare,
127 and except for extreme catastrophic events, it is difficult to distinguish cyclone effects
128 from routine mortality that occurs throughout the year (Tanner et al. 1994). From a
129 pragmatic perspective, it might be difficult to estimate seasonally varying parameters
130 because our data were collected at intervals greater than one year. Our aim here is
131 to produce a simple model which can be tested using field data, with which we can
132 evaluate the influence of state frequencies on interaction rates. However, seasonal
133 effects may be very important in other sessile assemblages, such as temperate mussel
134 communities (Wootton 2001*c*).

135 6. That the system is of infinite spatial extent and that local spatial effects are
136 unimportant. These assumptions allow us to develop a deterministic mean-field
137 differential equation model for the dynamics of the whole system, which governs the

138 transition rates in a Markov process representing the states of a finite number of
 139 sampled points in space. In reality, local spatial effects are likely to be important. In
 140 the Conclusions, we consider some ways in which they could be modelled.

141 We formulate the resulting model at the community level as a mean-field system of
 142 constant-coefficient nonlinear differential equations. Transitions between states may occur
 143 at any time. There is no reason to assume organisms only interact at fixed moments in
 144 time, unlike models of organisms with annual lifecycles, where discrete time is a natural
 145 choice. However, the properties of the system are likely to be sampled at discrete points in
 146 space and time. We therefore base our likelihood function on discrete sampling.

147 2.2. Derivation

In Appendix A1, we describe a stochastic model for the rate of transitions from state j to state i at a point in space, as a function of the number of points in state i . In the absence of detailed information on the spatial arrangement of points and dispersal distances, we then use a mean-field approximation in which the rate of transitions from j to i per unit frequency of state j is $a_{ij}x_i$, where the coefficient a_{ij} has dimensions T^{-1} , and x_i is the dimensionless frequency of state i . For transitions to empty space, we assume that there is no dependence on the frequency of empty space, and model the rate per unit frequency of j as a_{ej} (dimensions T^{-1}). We can now write an equation for the rate of change of frequency of each state, by summing the loss and gain terms over all destination and source states:

$$\frac{dx_i}{dt} = \begin{cases} - \left(a_{ei} + \sum_{k \neq e, i} a_{ki}x_k \right) x_i + \sum_{k \neq i} a_{ik}x_i x_k, & i \neq e \\ - \sum_{k \neq e} a_{ke}x_k x_e + \sum_{k \neq i} a_{ek}x_k, & i = e \end{cases} \quad (1)$$

148 with $x_i \geq 0$ and $\sum_{i=1}^s x_i = 1$.

149 For simplicity, we do not discuss facilitation in detail here, but it could be included

150 in such models. For example, if species i is better able to colonize empty space when the
 151 abundance of another species j is greater, we might include terms like $a_{ie,j}x_ix_ex_j$, where $a_{ie,j}$
 152 (dimensions T^{-1}) is the coefficient of colonization of empty space by i per unit frequency
 153 of j . If such terms were present, comparing a_{ji} with $a_{ie,j}$ would tell us about the relative
 154 importance of direct negative effects of j on i versus enhancement of colonization by i in
 155 the presence of j .

We can rewrite Eq. 1 in matrix form. Let \mathbf{A} be a matrix whose off-diagonal elements are the interaction coefficients a_{ij} and whose diagonal elements are zero. Let \mathbf{X} be a diagonal matrix with entries x_i if $i \neq e$, and 1 if $i = e$. Let \mathbf{C} be a diagonal matrix of column sums of \mathbf{XA} . Let \mathbf{x} be a column vector of probabilities of each state. Then

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= (\mathbf{XA} - \mathbf{C})\mathbf{x} \\ &= \mathbf{R}(\mathbf{x})\mathbf{x} \end{aligned} \tag{2}$$

156 where $\mathbf{R}(\mathbf{x})$ is a density-dependent rate matrix.

157 3. Relationships to other models

158 In this section, we show how the model of Section 2 is related to two well-known
 159 ecological models. First, it is a Lotka-Volterra competition model. Second, it is
 160 indistinguishable from a homogeneous continuous-time linear model (or its discrete-time
 161 equivalent) if it is at equilibrium, but will behave differently away from equilibrium and will
 162 respond differently to changes in parameters.

The general Lotka-Volterra competition model is

$$\frac{dx_i}{dt} = r_i x_i - \frac{r_i}{K_i} x_i^2 - \sum_{k \neq i} \frac{r_i}{K_i} \alpha_{ik} x_i x_k \tag{3}$$

163 (MacArthur and Levins 1967), where r_i (dimensions T^{-1}) is a per-capita population growth
 164 rate, K_i (dimensionless) is a carrying capacity, and α_{ik} (dimensionless) is an interspecific

165 competition coefficient measuring the effect of species k on the growth rate of species i .
 166 Note that as in Eq. 1, the x_i are proportions.

Because every point in the system is in one of the possible states, the proportion of points that are empty can be written as $x_e = 1 - \sum_{k \neq e} x_k$. Substituting this into Eq. 1 with $i \neq e$ and rearranging, we obtain

$$\frac{dx_i}{dt} = (a_{ie} - a_{ei})x_i - a_{ie}x_i^2 - \sum_{k \neq e, i} (a_{ki} + a_{ie} - a_{ik})x_i x_k \quad (4)$$

167 which is identical to Eq. 3 with $r_i = a_{ie} - a_{ei}$, $K_i = (a_{ie} - a_{ei})/a_{ie}$, and $\alpha_{ik} =$
 168 $(a_{ki} + a_{ie} - a_{ik})/a_{ie}$. We therefore refer to the model of Section 2 as the LV model from
 169 now on. The Lotka-Volterra competition model can also arise from completely different
 170 mechanistic assumptions, or simply as an approximation to a more complex model close to
 171 equilibrium (Schoener 1986).

If Eq. 2 is at equilibrium, $\mathbf{R}(\mathbf{x})$ does not vary over time. Let \mathbf{x}^* be an equilibrium state of Eq. 2, and set $\mathbf{Q} = \mathbf{R}(\mathbf{x}^*)$. Eq. 2 is therefore indistinguishable at equilibrium from the homogeneous continuous-time linear model with transition rates

$$\frac{d\mathbf{x}}{dt} = \mathbf{Q}\mathbf{x} \quad (5)$$

where \mathbf{Q} is a matrix whose off-diagonal elements q_{ij} (dimensions T^{-1}) are non-negative instantaneous transition rates, and whose diagonal elements q_{jj} are -1 times the column sums of off-diagonal elements. We refer to this model from now on as the linear model. Note that if the system is not at equilibrium, Eq. 2 and Eq. 5 will behave differently. Elsewhere (Spencer and Susko 2005), we discuss the relationship between this linear model and the usual discrete-time Markov models for communities of sessile organisms. Briefly, the usual formulation of a discrete-time Markov model is

$$\mathbf{x}(T + t) = \mathbf{P}(t)\mathbf{x}(T) \quad (6)$$

where $\mathbf{x}(T)$ is a vector of state probabilities at time T , and $\mathbf{P}(t)$ is a transition probability matrix whose ij th entry $p_{ij}(t)$ is the conditional probability of observing state i at time $T + t$ given that we observed state j at time T . If there is a homogeneous continuous-time process with generator \mathbf{Q} , then

$$\begin{aligned} \mathbf{P}(t) &= \sum_{m=0}^{\infty} \frac{(\mathbf{Q}t)^m}{m!} \\ &= e^{\mathbf{Q}t} \end{aligned} \tag{7}$$

172 where $e^{\mathbf{Q}t}$ is a matrix exponential. $\mathbf{P}(t)$ is a stochastic matrix, and its largest eigenvalue is
 173 1. Most models of this kind have a globally stable stationary distribution (Hill et al. 2004):
 174 the condition for this is that \mathbf{P} is regular (Kijima 1997, p. 52).

175 In many ecological analyses, the \mathbf{P} matrix is estimated by recording the identities of
 176 species at fixed points in space at a series of evenly-spaced time intervals, and aggregating
 177 the transition counts over space and time (e.g. Tanner et al. 1994; Wootton 2001c; Hill
 178 et al. 2004). To do so, we must assume either that the \mathbf{P} matrix is independent of state
 179 frequencies, or that the frequencies are close to equilibrium.

180 4. Likelihood, parameter estimation and model selection

One appealing feature of models for sessile organisms observed at discrete time intervals is that we can easily derive the likelihood of a model given the data. We can then make formal comparisons between models. Suppose we have a sequence of states y_0, y_1, \dots, y_k at a point in space observed at times t_0, t_1, \dots, t_k , where the time intervals are not necessarily equal. Under the Markov assumption, the probability of this sequence is

$$\begin{aligned} P(y_0, y_1, \dots, y_k) &= P(y_k|y_{k-1})P(y_{k-1}|y_{k-2}) \dots P(y_1|y_0)P(y_0) \\ &= P(y_0) \prod_{m=1}^k P(y_m|y_{m-1}) \end{aligned} \tag{8}$$

181 where $P(y_m|y_{m-1})$ is the probability of observing state y_m at time t_m given state y_{m-1} at
 182 time t_{m-1} and $P(y_0)$ is the probability of the initial state.

If we have a sample of sequences from a set of v independent and identically distributed (iid) points, then the likelihood L for the sequences at all the points is the product multinomial

$$\begin{aligned}
 L &= \prod_{h=1}^v P(y_{0,h}) \prod_{m=1}^k P(y_{m,h}|y_{m-1,h}) \\
 &= \left[\prod_j p_j(0)^{n_j(0)} \right] \prod_{m=1}^k \prod_{ij} p_{ij}(m, m-1)^{n_{ij}(m, m-1)}
 \end{aligned} \tag{9}$$

where $y_{m,h}$ is the state at point h at time t_m , $p_j(0)$ is the probability of state j at time 0, $p_{ij}(m, m-1)$ is the probability of state i at time t_m given state j at time t_{m-1} , $n_j(0)$ is the number of points in state j at time 0 and $n_{ij}(m, m-1)$ is the number of points in state j at time t_{m-1} and state i at time t_m . The product \prod_j is over all states and the product \prod_{ij} is over all combinations of states. In practice, it is easier to work with the log likelihood

$$l = \sum_j n_j(0) \log p_j(0) + \sum_{m=1}^k \sum_{ij} n_{ij}(m, m-1) \log p_{ij}(m, m-1) \tag{10}$$

183 In a homogeneous Markov model, initial states do not affect transition probabilities
 184 over subsequent time intervals, so the initial state is an ancillary statistic if the interaction
 185 rates or probabilities are the parameters of interest and we do not assume the process is at
 186 equilibrium. It is then usual to consider only the likelihood of the sequence conditional on
 187 the initial state (Lehmann 1986, section 10.2). However, in an LV model, initial states do
 188 affect subsequent transition probabilities, so we will use full likelihoods in all cases.

189 In deriving Eq. 10, we assumed that the sample points have negligible effects on the
 190 dynamics of the entire system. This is reasonable if the spatial extent of the system is
 191 infinite. We also assumed that the sample points are independent of each other. This
 192 is more problematic, unless sample points are far apart relative to the scale of spatial

193 interactions (which is unlikely to be strictly true for our data). However, even if the iid
194 assumption is incorrect, the model may still be useful. Since the model describes the
195 behaviour of a single point in space, parameter estimates from the marginal distribution
196 of states at a point will be asymptotically correct. If there are dependencies, it will be as
197 though there were fewer independent observations than sample points, so hypothesis tests
198 should be interpreted cautiously. In the Conclusions, we suggest some possible solutions to
199 the problem of spatial dependencies.

200 In Appendix A2, we give the transition probabilities for the LV model, the linear
201 homogeneous continuous-time model, a saturated discrete-time model with one parameter for
202 every combination of time interval, source state and destination state, and a time-averaged
203 discrete-time model with the same transition probabilities for every time interval.

204 If a model is going to help us understand the workings of a community, we have to
205 be able to estimate its parameters. A model is identifiable if all its parameters can be
206 estimated from the data. In Appendix A3, we illustrate how small singular values of the
207 Jacobian matrix of partial derivatives of expected values of each category of observations
208 with respect to each parameter can indicate identifiability problems (Catchpole and Morgan
209 1997). We apply this criterion to all the models we consider.

210 Parameter estimation requires maximizing the log likelihood as a function of the
211 parameters. For a discrete-time model, the maximum likelihood estimates of transition
212 probabilities are given in Appendix A2, Eqs. A.6 or A.7. We do not have closed-form
213 estimates for the other models, so we use numerical optimization as described in Appendix
214 A5.

215 Appendix A4 describes the methods we used to compare models. Briefly, the LV, linear,
216 and time-averaged discrete models (and all other possible models) are nested within the
217 saturated model, so we can use likelihood ratio statistics to compare each to the saturated

218 model (Hilborn and Mangel 1997, pages 153-154). We can also use Akaike’s Information
219 Criterion (AIC) to compare all four models with each other (Akaike 1992). The Akaike
220 weight for each model can be treated as an estimate of the probability that it is the best in
221 the set under consideration (Burnham and Anderson 2004).

222 5. Data

223 We fitted the models to data from a long-term study of coral community dynamics
224 at Heron Island, Great Barrier Reef, Queensland, Australia (Connell et al. 1997, 2004).
225 Data from this study have previously been analyzed using both discrete-time (Tanner et al.
226 1994, 1996; Hill et al. 2004) and continuous-time (Spencer and Susko 2005; Spencer 2006)
227 Markov models, and are available on request from JT. In parallel with the current study,
228 a discrete-time density-dependent model is also being developed, although it is not in a
229 format that allows direct comparison to the models presented here. Data were collected
230 from photographs of fixed 1m² quadrats taken at unequal intervals over 27 years, from 1962
231 to 1989. Grids of points, referenced to fixed stakes, were placed over the photographs and
232 the species present at each point recorded as described in Tanner et al. (1994). Because the
233 locations of grid points were fixed, it was possible to determine how the species present at
234 points in space changed over time, giving counts of transitions for each time interval.

235 72 species of corals and 9 species of algae were observed in the quadrats over the
236 27-year study period. In previous studies, these were grouped into eight categories based
237 on taxonomy and morphology, plus a free space state (Tanner et al. 1994, 1996). Free
238 space was usually occupied by organisms such as crustose coralline and turfing algae, but
239 is available for colonization by corals and macroalgae (J. E. Tanner et al., in preparation).
240 Initial analyses (Appendix A5) showed that the reliability of parameter estimation was
241 improved by aggregating the four original acroporid coral states into a single state, reducing

242 the number of parameters from 80 to 35. All the models presented here, including the
243 saturated model, are therefore based on six categories: three hard coral groups (acroporids,
244 massives, and pocilloporids); soft corals; algae; and free space. Aggregation is reasonable
245 because all the acroporids showed similar trends in frequency over time. Analyses of the
246 unaggregated data led to qualitatively similar conclusions. Simulation studies (Appendix
247 A5) showed that we could correctly identify the true model and obtain good parameter
248 estimates in most cases. We report results from only one intertidal site, the Protected Crest.
249 Data were also available for one other intertidal site (Exposed Crest) and one subtidal
250 site (Exposed Pools). However, these had fewer sample points in time and/or space, and
251 simulation studies like those described in Appendix A5 showed that there were too few
252 data for reliable parameter estimation. There were small numbers of missing observations
253 ($< 1\%$), which we ignore. We analyzed data from all 17 sample dates used by Tanner et al.
254 (1994). There were at least 1249 points observed per time interval.

255 6. Results and discussion

256 6.1. Model selection

257 Table 1 shows the log likelihoods l , number of parameters p , and Akaike’s Information
258 Criterion for all the models. The saturated model is much better than the LV model, which
259 in turn is much better than the linear model. The time-averaged discrete model is worst of
260 all.

261 Likelihood ratio tests reject the LV, linear and time-averaged discrete models with
262 $p < 1 \times 10^{-16}$. The small p -value is not surprising because there are 450 degrees of freedom.
263 The Akaike weight of the saturated model is > 0.9999 , indicating overwhelming support for
264 this model compared to the others.

265 The saturated model is not of much biological interest. Although it is the best
266 possible description of the data, it tells us nothing about mechanisms, and it cannot be
267 used to predict future events. Its main value is to provide a standard against which other
268 models are measured. Thus, even though our other models can be rejected as a complete
269 explanation of the data, it is still worth comparing them to each other in order to choose the
270 most promising framework for further development. Similar situations arise in modelling
271 molecular evolution: early models could often be rejected (e.g. Huelsenbeck and Crandall
272 1997, page 454), but have provided a basis for the development of more sophisticated
273 models (Sullivan and Joyce 2005, page 459).

274 The LV model has a much smaller AIC than the linear model. If the saturated model
275 is excluded from the comparison, the Akaike weight of the LV model is > 0.9999 . Thus,
276 the LV model is much better than any other non-saturated model we considered. This
277 indicates that transition probabilities are likely to depend on state frequencies, although
278 comparison with the saturated model shows that other factors must also be important.
279 J. E. Tanner et al. (in preparation) also show that discrete-time transition probabilities
280 can be strongly density-dependent in this system, with up to 74% of the variance in some
281 transition probabilities being explained by density.

282 Both continuous-time models are better than the time-averaged discrete model. It
283 is not surprising that transition probabilities depend on the length of the time interval,
284 although most previous models have not taken account of this (e.g. Tanner et al. 1994,
285 1996; Spencer and Susko 2005). However, there are plausible situations in which simple
286 continuous-time models would not work well, for example if temporal variability in
287 environmental conditions mattered more than the length of the time interval.

288 Figure 1 shows the predicted and observed frequencies of each state at each sample
289 date. Parameters are estimated from transition frequency data, but time series of abundance

290 are a good visual representation of the behaviour of each model. Predicted frequencies are
291 given by Eq. 2 for the LV model, and Eq. 5 for the linear model. For the time-averaged
292 discrete model we generated expected frequencies using Eq. 6, ignoring the variation in
293 sample intervals. The LV model generates predicted frequencies that look much more like
294 the observed data than those from the linear model. The time-averaged discrete model
295 generates frequencies similar to the linear model.

296 In the linear model, the smallest singular value of the Jacobian was 1×10^{-9} , which
297 may indicate potential identifiability problems (Appendix A3). The largest transition rate
298 was from algae to free space ($q_{63} = 147.53$), an order of magnitude larger than any other.
299 The predicted proportion of algae is low and rapidly approaches an equilibrium. Small
300 changes in other transition rates involving algae are unlikely to have much effect on this
301 behaviour. There may not be very well-defined optimal parameter estimates for the linear
302 model. However, since the linear model performs much worse than the LV model, these
303 estimates are not of much interest. Although there were some large rates in the LV model,
304 the rate matrix was not dominated by a single large rate and no singular values were less
305 than 1×10^{-5} . We can therefore be more confident that the parameter estimates for the LV
306 model have biological meaning.

307 **6.2. Interaction coefficient estimates in the LV model**

308 Parameter estimates for all the non-saturated models are given in Table 2 (LV model)
309 and Appendix: Tables A3 and A4 (linear and time-averaged discrete models). Here, we
310 briefly discuss the biological significance of the interaction coefficient estimates in the LV
311 model.

312 High coefficients for transitions into a state are not necessarily associated with high

313 abundance. For example, there are high coefficients for transitions from acroporids, massive
314 corals and free space to algae. However, there are also high coefficients for transitions from
315 algae to pocilloporid corals and free space. Algae therefore show rapid turnover but do not
316 become abundant (Figure 1). This is in accordance with the idea that algae are transient,
317 fast-colonizing species on this reef (Connell 1987).

318 There are a number of very low coefficients ($< 1 \times 10^{-9}$: 5/30 rates = 17%). In earlier
319 analyses with four separate acroporid states, almost every possible transition occurred at
320 some point during the observation period (Tanner et al. 1994), although a continuous-time
321 linear model suggested that some transitions may only have occurred indirectly (Spencer
322 and Susko 2005). When acroporids were aggregated into a single state, the proportion of
323 possible transitions that were never observed ($7/36 = 19\%$) was not dramatically different
324 from the proportion of very low coefficients in the LV model, although only three of these
325 unobserved transitions also had coefficients less than 1×10^{-9} in the LV model. When
326 some states are very transient, others persist for much longer, and the sampling intervals
327 are moderately long, the pattern of transitions that are observed may not accurately reflect
328 the events that actually occur (J.E. Tanner et al., in preparation).

329 We also investigated the long-term dynamics of the LV model. Using the estimated
330 interaction coefficients and initial conditions from Table 2, we solved the system numerically
331 for 10000 years. The state frequencies approached an equilibrium with 46% acroporids, 35%
332 free space, 11% soft corals, 8% massive corals, and no pocilloporid corals or algae. There
333 were transient oscillations, which were negligible after about 1000 years. We analyzed the
334 local stability of this equilibrium by examining the eigenvalues of the Jacobian (e.g. May
335 1971). These eigenvalues all had negative real parts. This confirms that the equilibrium
336 is locally stable (and that at this equilibrium, we could find a linear model that would be
337 indistinguishable from the LV model).

7. Conclusions

338

339 For the data set we studied, the Lotka-Volterra model performed much better than
340 two density-independent alternatives, even though all these non-saturated models had the
341 same number of parameters. Density-independent colonization is the distinguishing feature
342 of open population models (Roughgarden et al. 1985), and is typically associated with
343 long-distance larval dispersal. It is implausible that the system we studied is completely
344 closed, but it does seem likely that local abundances are important in determining transition
345 rates. This reflects the fact that colonization by recruitment of new individuals is rare
346 compared to colonization by lateral growth of existing individuals.

347 We also analyzed data from two other sites, the Exposed Crest and Exposed Pools, but
348 do not report these results here. The other two sites had shorter time series and/or fewer
349 points in space, and simulations showed that parameter estimation was less reliable at these
350 sites than at the Protected Crest. Analyses of both these sites strongly favoured the LV
351 model over the linear model, as at the Protected Crest. However, simulations showed that
352 the frequency of wrongly selecting the LV model when the linear model was the true model
353 was much higher than at the Protected Crest. As mentioned previously, a linear model
354 close to equilibrium may be difficult to distinguish from an LV model close to equilibrium,
355 and the linear models quickly approached equilibrium for parameters estimated from the
356 field data. Although this may not be a problem for the field data, because the systems were
357 not particularly close to equilibrium, we are reluctant to draw any strong conclusions from
358 these other sites. This highlights the need for long time series as a foundation for statistical
359 ecology. We hope that further studies on other long-term ecological data sets will lead to a
360 more general understanding of the situations in which linear and LV models are suitable for
361 communities of sessile organisms.

362 One of the other sites we analyzed, the Exposed Pools, was also strongly affected

363 by storm damage on several occasions (Connell et al. 1997). The LV model was able to
364 reproduce some of the observed large fluctuations in abundance of corals and free space
365 following these storms. However, in reality these fluctuations were likely to have been a
366 consequence of fluctuations in rates of transitions to free space, which were treated as
367 constant in the model. Mortality rates are likely to vary considerably over time, to depend
368 on the time since the last storm (because larger colonies may be more vulnerable to storm
369 damage), and to covary among points in space and states in the system. It therefore seems
370 unlikely that the estimated parameters from the LV model are biologically meaningful
371 when occasional major storm damage is important. It should be possible to explicitly
372 include storm damage in the models described here, by using data on storm intensities as a
373 predictor of transitions to free space (Madin and Connolly 2006).

374 Other factors may also be changing over time. At the Protected Crest site, acroporids
375 were initially abundant but showed a gradual decline, free space increased over time,
376 massive corals were moderately abundant and soft corals showed a rapid increase at the
377 end of the time series (Figure 1). These trends may be partly due to upward growth of the
378 reef and partial diversion of water flow (Connell et al. 2004).

379 Another possible improvement is making the model spatially explicit. It is likely that
380 transition rates at a point depend more strongly on the states of the immediate neighbours
381 of the point than on the average state of the system. Gratzner et al. (2004, p. 9) provide
382 an interesting discussion of the consequences of non-random distributions of individuals for
383 models of spatial interaction in forest systems, phrased in terms similar to those we used in
384 deriving the LV model (Section 2.2). Although there are spatially explicit models of sessile
385 marine communities (e.g. Wilson et al. 1996; Burrows and Hawkins 1998; Wootton 2001*b*;
386 Robles and Desharnais 2002; Langmead and Sheppard 2004; Dunstan and Johnson 2005,
387 2006), little effort has yet been expended on fitting to time series, or performance relative

388 to non-spatial models. Instead, most models have focussed on qualitative patterns. One
389 way to reduce the complexity of these models is to use moment approximations to obtain
390 more accurate differential equation models than the mean-field system we considered (e.g.
391 Bolker and Pacala 1997; Pascual and Levin 1999). Alternatively, spatially explicit models
392 are often formulated as stochastic cellular automata, which are linear Markov models with
393 a very large state space. The likelihood for a discrete-time cellular automaton is easy to
394 calculate, because the states of cells at time $t + 1$ are conditionally independent given the
395 states of their neighbourhoods at time t (Balzter et al. 1998; Craiu and Lee 2006). The
396 problem is more difficult in continuous time, because events outside the neighbourhood of
397 a cell may affect its future state through a sequence of changes to intervening cells. The
398 full generator matrix for such a model will usually be too large to store or exponentiate.
399 Possible approaches include Monte Carlo estimation of likelihoods by importance sampling
400 (Juneja and Shahabuddin 2001), pseudolikelihood methods (Besag 1975), and approximate
401 Bayesian computation (Sisson et al. 2007). Perhaps the most closely related problem is
402 that of estimating parameters for a presence-absence metapopulation from time series data.
403 If the metapopulation is not observed every year, Bayesian data augmentation methods
404 can be used to integrate over sample paths for unobserved intervals (O’Hara et al. 2002).
405 However, the problem is more difficult in continuous time because the length of the sample
406 path is unknown. An alternative approach is to treat a single observed spatial pattern as
407 a sample from the stationary distribution of a continuous-time Markov chain, as has been
408 done for models of forest gap creation and regeneration (Schlicht and Iwasa 2004). It is
409 not clear that this would be appropriate for our system, given the large variation in state
410 frequencies over time.

411 In conclusion, the models we have investigated here are simple, and have already
412 been well-studied in theoretical ecology. Our main contribution is the attempt to compare
413 the performance of these models as quantitative descriptions of long-term field data.

414 Stochastic models of real time series can both increase our understanding of the underlying
415 mechanisms, and help us to make better predictions about the future dynamics of
416 populations and communities (e.g. Foley 1994; Kendall et al. 1999; Thomas et al. 2005).

417 Acknowledgments

418 This work was supported by a grant from the University of Liverpool Research
419 Development Fund, project number 6119. We are particularly grateful to Joe Connell for
420 initiating and conducting the field study, and making the data available, and Terry Hughes
421 for his role in data collection and processing. The manuscript was improved by detailed
422 and perceptive comments from three anonymous reviewers. We are also grateful to Steve
423 Brooks and Damian Clancy for helpful suggestions, Don DeAngelis for comments on an
424 earlier draft, and Mary Thorp and Cliff Addison for technical support.

REFERENCES

425

426 Acevedo, M. F. 1981. On Horn's Markovian model of forest dynamics with particular
427 reference to tropical forests. *Theoretical Population Biology*, **19**:230–250.

428 Akaike, H. 1992. Information theory and an extension of the maximum likelihood principle.
429 In S. Kotz and N. L. Johnson, editors, *Breakthroughs in Statistics Volume I:*
430 *Foundations and Basic Theory*, pages 599–624. Springer-Verlag, New York.

431 Balzter, H., P. W. Braun, and W. Köhler. 1998. Cellular automata models for vegetation
432 dynamics. *Ecological Modelling*, **107**:113–125.

433 Besag, J. 1975. Statistical analysis of non-lattice data. *The Statistician*, **24**:179–195.

434 Bolker, B. and S. W. Pacala. 1997. Using moment equations to understand stochastically
435 driven spatial pattern formation in ecological systems. *Theoretical Population*
436 *Biology*, **52**:179–197.

437 Burnham, K. P. and D. R. Anderson. 2004. Multimodel inference: understanding AIC and
438 BIC in model selection. *Sociological Methods and Research*, **33**:261–304.

439 Burrows, M. T. and S. J. Hawkins. 1998. Modelling patch dynamics on rocky shores using
440 deterministic cellular automata. *Marine Ecology Progress Series*, **167**:1–13.

441 Caswell, H. 2001. *Matrix population models: construction, analysis, and interpretation*.
442 Sinauer, Sunderland, MA, second edition.

443 Caswell, H. and J. E. Cohen. 1995. Red, white and blue: environmental variance spectra
444 and coexistence in metapopulations. *Journal of Theoretical Biology*, **176**:301–316.

445 Catchpole, E. A. and B. J. T. Morgan. 1997. Detecting parameter redundancy. *Biometrika*,
446 **84**:187–196.

- 447 Connell, J. H. 1987. Change and persistence in some marine communities. In A. J. Gray,
448 M. J. Crawley, and P. J. Edwards, editors, *Colonization, Succession and Stability*,
449 pages 339–352. Blackwell Scientific Publications, Oxford.
- 450 Connell, J. H., T. P. Hughes, and C. C. Wallace. 1997. A 30-year study of coral abundance,
451 recruitment, and disturbance at several scales in space and time. *Ecological*
452 *Monographs*, **67**:461–488.
- 453 Connell, J. H., T. P. Hughes, C. C. Wallace, J. E. Tanner, K. E. Harms, and A. M.
454 Kerr. 2004. A long-term study of competition and diversity of corals. *Ecological*
455 *Monographs*, **74**:179–210.
- 456 Craiu, R. V. and T. C. M. Lee. 2006. Pattern generation using likelihood inference for
457 cellular automata. *IEEE Transactions on Image Processing*, **15**:1718–1727.
- 458 Dunstan, P. K. and C. R. Johnson. 2005. Predicting global dynamics from local interactions:
459 individual-based models predict complex features of marine epibenthic communities.
460 *Ecological Modelling*, **186**:221–233.
- 461 Dunstan, P. K. and C. R. Johnson. 2006. Linking richness, community variability, and
462 invasion resistance with patch size. *Ecology*, **87**:2842–2850.
- 463 Foley, P. 1994. Predicting extinction times from environmental stochasticity and carrying
464 capacity. *Conservation Biology*, **8**:124–137.
- 465 Gaines, S. and J. Roughgarden. 1985. Larval settlement rate: a leading determinant of
466 structure in an ecological community of the marine intertidal zone. *Proceedings of*
467 *the National Academy of Sciences of the United States of America*, **82**:3707–3711.
- 468 Gratzler, G., C. Canham, U. Dieckmann, A. Fischer, Y. Iwasa, R. Law, M. J. Lexer,
469 H. Sandmann, T. A. Spies, B. E. Splechna, and J. Szwagrzyk. 2004. Spatio-temporal

- 470 development of forests – current trends in field methods and models. *Oikos*,
471 **107**:3–15.
- 472 Harrison, P. L. and C. C. Wallace. 1990. Reproduction, dispersal and recruitment of
473 scleractinian corals. In Z. Dubinsky, editor, *Ecosystems of the World 25: Coral*
474 *Reefs*, pages 133–207. Elsevier, Amsterdam.
- 475 Hilborn, R. and M. Mangel. 1997. *The ecological detective*. Princeton University Press,
476 Princeton, New Jersey.
- 477 Hill, M. F., J. D. Witman, and H. Caswell. 2002. Spatio-temporal variation in Markov
478 chain models of subtidal community succession. *Ecology Letters*, **5**:665–675.
- 479 Hill, M. F., J. D. Witman, and H. Caswell. 2004. Markov chain analysis of succession in a
480 rocky subtidal community. *American Naturalist*, **164**:E46–E61.
- 481 Huelsenbeck, J. P. and K. A. Crandall. 1997. Phylogeny estimation and hypothesis testing
482 using maximum likelihood. *Annual Review of Ecology and Systematics*, **28**:437–466.
- 483 Hyder, K., P. Åberg, M. P. Johnson, and S. J. Hawkins. 2001. Models of open populations
484 with space-limited recruitment: extension of theory and application to the barnacle
485 *Chthamalus montagui*. *Journal of Animal Ecology*, **70**:853–863.
- 486 Juneja, S. and P. Shahabuddin. 2001. Fast simulation of Markov chains with small transition
487 probabilities. *Management Science*, **47**:547–562.
- 488 Kendall, B. E., C. J. Briggs, W. W. Murdoch, P. Turchin, S. P. Ellner, E. McCauley, R. M.
489 Nisbet, and S. Wood. 1999. Why do populations cycle? a synthesis of statistical and
490 mechanistic modeling approaches. *Ecology*, **80**:1789–1805.
- 491 Kijima, M. 1997. *Markov processes for stochastic modeling*. Chapman and Hall, London.

- 492 Lang, J. C. and E. A. Chornesky. 1990. Competition between scleratinian reef corals — a
493 review of mechanisms and effects. In Z. Dubinsky, editor, *Ecosystems of the World*
494 25: Coral Reefs, pages 209–252. Elsevier, Amsterdam.
- 495 Langmead, O. and C. Sheppard. 2004. Coral reef community dynamics and disturbance: a
496 simulation model. *Ecological Modelling*, **175**:271–290.
- 497 Lehmann, E. L. 1986. *Testing statistical hypotheses*. John Wiley and Sons, second edition.
498 600 pp.
- 499 MacArthur, R. and R. Levins. 1967. The limiting similarity, convergence and divergence of
500 coexisting species. *American Naturalist*, **101**:377–385.
- 501 Madin, J. S. and S. R. Connolly. 2006. Ecological consequences of major hydrodynamic
502 disturbances on coral reefs. *Nature*, **444**:477–480.
- 503 May, R. M. 1971. Stability in multispecies community models. *Mathematical Biosciences*,
504 **12**:59–79.
- 505 Muko, S., K. Sakai, and Y. Iwasa. 2001*a*. Dynamics of marine sessile organisms with
506 space-limited growth and recruitment: application to corals. *Journal of Theoretical*
507 *Biology*, **210**:67–80.
- 508 Muko, S., K. Sakai, and Y. Iwasa. 2001*b*. Size distribution dynamics for a marine sessile
509 organism with space-limitation in growth and recruitment: application to a coral
510 population. *Journal of Animal Ecology*, **70**:579–589.
- 511 Norris, J. R. 1997. *Markov Chains*. Cambridge University Press, Cambridge, England.
- 512 O’Hara, R. B., E. Arjas, H. Toivonen, and I. Hanski. 2002. Bayesian analysis of
513 metapopulation data. *Ecology*, **83**:2408–2415.

- 514 Pascual, M. and S. A. Levin. 1999. Spatial scaling in a benthic population model with
515 density-dependent disturbance. *Theoretical Population Biology*, **56**:106–122.
- 516 Robles, C. and R. Desharnais. 2002. History and current development of a paradigm of
517 predation in rocky intertidal communities. *Ecology*, **83**:1521–1536.
- 518 Roughgarden, J., Y. Iwasa, and C. Baxter. 1985. Demographic theory for an open marine
519 population with space-limited recruitment. *Ecology*, **66**:54–67.
- 520 Schlicht, R. and Y. Iwasa. 2004. Forest gap dynamics and the Ising model. *Journal of*
521 *Theoretical Biology*, **230**:65–75.
- 522 Schoener, T. W. 1986. Mechanistic approaches to community ecology: a new reductionism?
523 *American Zoologist*, **26**:81–106.
- 524 Sisson, S. A., Y. Fan, and M. M. Tanaka. 2007. Sequential Monte Carlo without likelihoods.
525 *Proceedings of the National Academy of Sciences of the United States of America*,
526 **104**:1760–1765.
- 527 Spencer, M. 2006. Sensitivity analysis of Markov models for communities of competing
528 sessile organisms. *Journal of Animal Ecology*, **75**:1024–1033.
- 529 Spencer, M. and E. Susko. 2005. Continuous-time Markov models for species interactions.
530 *Ecology*, **86**:3272–3278.
- 531 Sullivan, J. and P. Joyce. 2005. Model selection in phylogenetics. *Annual Review of Ecology*
532 *and Systematics*, **36**:445–466.
- 533 Svensson, C. J., S. R. Jenkins, S. J. Hawkins, A. A. Myers, P. Range, J. Paula, R. M.
534 O’Riordan, and P. Åberg. 2004. Models of open populations with space-limited
535 recruitment in stochastic environments: relative importance of recruitment and

- 536 survival in populations of *Semibalanus balanoides*. Marine Ecology Progress Series,
537 **275**:185–197.
- 538 Tanner, J. E., T. P. Hughes, and J. H. Connell. 1994. Species coexistence, keystone species,
539 and succession: a sensitivity analysis. Ecology, **75**:2204–2219.
- 540 Tanner, J. E., T. P. Hughes, and J. H. Connell. 1996. The role of history in community
541 dynamics: a modelling approach. Ecology, **77**:108–117.
- 542 Thomas, L., S. T. Buckland, K. B. Newman, and J. Harwood. 2005. A unified framework
543 for modelling wildlife population dynamics. Australian and New Zealand Journal of
544 Statistics, **47**:19–34.
- 545 Usher, M. B. 1979. Markovian approaches to ecological succession. Journal of Animal
546 Ecology, **48**:413–426.
- 547 Wilson, W. G., R. M. Nisbet, A. H. Ross, C. Robles, and R. A. Desharnais. 1996. Abrupt
548 population changes along smooth environmental gradients. Bulletin of Mathematical
549 Biology, **58**:907–922.
- 550 Wootton, J. T. 2001*a*. Causes of species diversity differences: a comparative analysis of
551 Markov models. Ecology Letters, **4**:46–56.
- 552 Wootton, J. T. 2001*b*. Local interactions predict large-scale pattern in empirically derived
553 cellular automata. Nature, **413**:841–844.
- 554 Wootton, J. T. 2001*c*. Prediction in complex communities: analysis of empirically derived
555 Markov models. Ecology, **82**:580–598.
- 556 Wootton, J. T. 2004. Markov chain models predict the consequences of experimental
557 extinctions. Ecology Letters, **7**:653–660.

Table 1: Log likelihoods (l), number of parameters and Akaike's Information Criterion (AIC) for Protected Crest models, ordered by increasing AIC.

Model	l	Parameters	AIC
Saturated discrete	-1.5413×10^4	485	3.1797×10^4
LV	-1.6208×10^4	35	3.2486×10^4
Linear	-1.7184×10^4	35	3.4437×10^4
Time-averaged discrete	-1.7330×10^4	35	3.4730×10^4

Table 2: Estimated \mathbf{A} matrix (years⁻¹) for LV model, Protected Crest.

	1	2	3	4	5	6
1: acroporid corals	0	0.6075	6.3557	0.0802	0.1525	0.9572
2: soft corals	0.3654	0	1.13e-54	0.8443	62.1711	0.5078
3: algae	28.9099	2.20e-25	0	11.8025	6.94e-38	25.6192
4: massive corals	0.7293	2.79e-47	0.4187	0	0	0.8297
5: pocilloporid corals	8.40e-12	4.21e-04	38.6881	1.12e-06	0	1.3899
6: free space	0.3068	0.1381	23.1526	0.4922	1.0719	0

Note: Estimated initial state frequencies were $[0.4938, 1.0441 \times 10^{-5}, 0.0174, 0.0102, 0.0175, 0.4612]^T$.

Fig. 1.— Protected Crest data (circles), LV model (solid line), linear model (dashed line), and time-averaged discrete model ignoring variation in time intervals (crosses). States are (A) acroporid corals, (B) soft corals, (C) algae, (D) massive corals, (E) pocilloporid corals, (F) free space. Time is measured in years since the first sample. The vertical axis scales are different in each panel.

