J. Andrew Royle and Marc Kery. 2007. A Bayesian state-space formulation of dynamic occupancy models. ***Ecology*** **88:1813–1823.**

Appendix A: Development of an MCMC algorithm by Gibbs sampling for a simple model.

Ecological Archives **E088-108-A1.**

.

Appendix A: Full conditional distributions for $z(i,t)$

We adopt conventional uniform priors for the model parameters, so that ϕ_t, γ_t and ψ are all distributed as Uniform(0,1) random variables: $\phi_t \sim \text{Uniform}(0,1)$ and $\gamma_t \sim \text{Uniform}(0,1)$ for each t, and $\psi_1 \sim \text{Uniform}(0,1)$. The full conditionals for model parameters are developed conditional on $\{z(i,t)\}$, i.e., as if each $z(i,t)$ were known. We require summary statistics, the number of survival events, say $n_{s,t}$. This is the number of times $z(i,t+1) = 1$ given that $z(i,t) = 1$. We also must tabulate the number of potential survival events, $N_{s,t}$, i.e. the number of $z(i,t) = 1$. Similarly, we tabulate the number of colonization events, $n_{c,t}$ and the number of potential colonization events, $N_{c,t}$ (the frequency of $z(i,t) = 0$). Note that $\sum_i N_{s,t} + N_{c,t} = R \times (T-1)$. As an example of the calculation of these intermediate summaries, with 4 sites and 4 replicates, we have:

Table A.1: Example of calculating summary statistics for a case in which there are 4 sample sites and $T = 4$. The summary statistics are: $n_{c,t}$ = the number of colonization events during the interval $t-1$ to t, $N_{c,t}$ = the number of unoccupied sites at time $t-1$, $n_{s,t}$ = the number of survival events during the interval $t-1$ to t, $N_{s,t}$ = the number of occupied sites at time $t-1$.

	Site				$n_{s,t}$	$N_{s,t}$	$n_{c,t}$	$N_{c,t}$
	1	2	3	4				
t=1	1	0	0	1	–	–	–	–
t=2	1	0	1	1	2	2	1	2
t=3	0	0	1	1	2	3	0	1
t=4	1	0	1	1	2	2	1	2

Given these summary statistics, the full-conditional distributions for each of the structural model parameters are:

(1) $\phi_t|\cdot \sim \text{Beta}(n_{s,t}{+}1, N_{s,t}{-}n_{s,t}{+}1)$

(2) $\gamma_t|\cdot \sim \text{Beta}(n_{c,t}{+}1, N_{c,t}{-}n_{c,t}{+}1)$

(3) $\psi_1|\cdot \sim \text{Beta}(1{+}\sum_i z(i,1), 1{+}R{-}\sum_i z(i,1))$

These have beta full-conditionals because the prior distributions are uniform, which is a beta(1,1), the conjugate prior for the binomial likelihood.

Obtaining the full conditionals for $z(i,t)$ is somewhat more involved, however, analytic forms exist. For $1 < t < T$ the full conditional of $z(i,t)$ is proportional to the product

$$z(i,t)|y(i,t), z(i,t-1), z(i,t+1) \propto [y(i,t)|z(i,t)][z(i,t)|z(i,t-1), z(i,t+1)] \qquad \text{(A.1)}$$

Dependence on dynamical parameters has been suppressed. The problem can be simplified by considering first the conditional distribution $[z(i,t)|z(i,t+1), z(i,t-1)]$ (i.e., in the absence of data). For clarity, we omit the i subscript in this development, and suppose that $t = 1, 2, 3$. The full conditionals can be developed algebraically for the 4 possible combinations of $z(i,t-1)$ and $z(i,t+1)$. Consider the case $T = 3$, and calculation of the full conditional for z_2. The 4 cases, corresponding to possible values of $z(1)$ and $z(3)$ are $(1, z_2, 1)$, $(0, z_2, 1)$, $(1, z_2, 0)$ and $(0, z_2, 0)$. Given $(1, z_2, 1)$, there are 2 possible sequences $(1, 1, 1)$ and $(1, 0, 1)$. The probability of the first is $\phi_1\phi_2$ since $z_2 = 1$ implies that there was a survival event between $t = 1$ and $t = 2$ and also between $t = 2$ and $t = 3$. The probability of obtaining the second is $(1 - \phi_1)\gamma_2$. Thus,

$$\psi_{condl} = Pr(z_2 = 1|z_1 = 1, z_3 = 1) = \frac{\phi_1\phi_2}{\phi_1\phi_2 + (1 - \phi_1)\gamma_2} \tag{A.2}$$

(the denominator being the sum of the 2 possibilities). Repeating this for the other 3 cases, we see that ψ_{condl} has 4 potential values depending on the previous and subsequent values of $z(i, t)$, those being:

$$\begin{aligned}
(1, X, 1) : \psi_{condl} &= \frac{\phi_{t-1}\phi_t}{\phi_{t-1}\phi_t + (1 - \phi_{t-1})\gamma_t} \\
(0, X, 1) : \psi_{condl} &= \frac{\gamma_{t-1}\phi_t}{\gamma_{t-1}\phi_t + (1 - \gamma_{t-1})\gamma_t} \\
(1, X, 0) : \psi_{condl} &= \frac{\phi_{t-1}(1 - \phi_t)}{\phi_{t-1}(1 - \phi_t) + (1 - \phi_{t-1})(1 - \gamma_t)} \\
(0, X, 0) : \psi_{condl} &= \frac{\gamma_{t-1}(1 - \phi_t)}{\gamma_{t-1}(1 - \phi_t) + (1 - \gamma_{t-1})(1 - \gamma_t)}
\end{aligned}$$

This yields the full-conditional for $z(i, t)$ given $z(i, t + 1)$ and $z(i, t - 1)$ (i.e., this is the 2nd component on the RHS of Eq. (A.1). That is, $z(i, t)$ is Bernoulli with parameter ψ_{condl}, which is dependent on z_1 and z_3 (or z_{t-1} and z_{t+1}, in general).

We now apply Bayes' Rule recursively, using this Bernoulli distribution as the prior distribution for updating the likelihood $[y(i, t)|z(i, t)]$ (the first term on the RHS of Eq.

(A.1)). Basic probability calculus yields that

$$Pr(z(i,t)=1|y(i,t)=0) = \frac{(1-p)^T \psi_{condl}}{(1-p)^T \psi_{condl} + (1-\psi_{condl})}.$$

Finally, updating the first and last values of $z(i,t)$ requires only a slight modification to these expressions, simplifying Eq. (A.2) to account for the fact that there is only one value of z being conditioned on – that being $z(i,2)$ in the case of updating $z(i,1)$, and $z(i,T-1)$ in the case of updating $z(i,T)$. In both cases, there are only 2 possible values of ψ_{condl} instead of four.