Cajo J. F. ter Braak and Rampal S. Etienne. 2003. Improved Bayesian analysis of metapopulation data with an application to a tree frog metapopulation. *Ecology* 84: 231-241.

Appendix A. The jumping distribution for missing data

In this appendix we describe our Metropolis-Hastings (MH) algorithm for updating missing patch data. First we introduce some notation. While updating, the model parameters Θ are kept fixed. For notational convenience we therefore drop the dependence on Θ. Furthermore, let the proposal for patch i in year t be denoted by $X_i^*(t)$ to distinguish it from the current state $X_i(t)$. Of course non-missing data are not updated, so that for these patches the proposed state and current state are identical. By $X_{-i}(t)$ we mean the states, at time t, of all patches except patch i, and by X_{-i} we mean the states of all patches in all years, except $X_i(t)$. Our mechanism to generate a proposal is motivated by the Gibbs sampler for updating a single patch (O'Hara et al. 2002). The proposal for a single patch must, of course, be either 1 or 0. The ratio f_i of the probabilities of these two states for patch i conditional on X_{-i} is

$$f_i = \frac{P[X_i^*(t) = 1|X_{-i}]}{P[X_i^*(t) = 0|X_{-i}]}. \tag{A-1}$$

By the definition of conditional probabilities this becomes

$$f_i = \frac{P[X_i^*(t) = 1, X_{-i}]}{P[X_i^*(t) = 0, X_{-i}]} \tag{A-2}$$

and using the properties of the metapopulation model expressed in (3) and (2) in the main text, we can simplify this to

$$f_i = \frac{P[X_i^*(t) = 1|X(t-1)]\ P[X(t+1)|X_i^*(t) = 1, X_{-i}(t)]}{P[X_i^*(t) = 0|X(t-1)]\ P[X(t+1)|X_i^*(t) = 0, X_{-i}(t)]}, \tag{A-3}$$

so that we, in fact, condition only on the two adjacent years $t-1$ and $t+1$.

Because the numerator and the denominator in (A-1) must sum to 1, we get

$$p_i = P[X_i^*(t) = 1|X_{-i}] = \frac{f_i}{1+f_i}. \tag{A-4}$$

Proposing $X_i^*(t) = 1$ with probability p_i (and leaving the other patches unchanged) yields a Gibbs sampler, because (A-4) is a full conditional distribution, and the acceptance ratio is equal to 1.

The Gibbs proposal is costly to compute because each probability in (A-3) involving $X(t+1)$ is already a product of N probabilities (see equation (2) in the main text). The proposal can be simplified considerably by keeping $X_i^*(t)$ at its current state $X_i(t)$ when calculating the transition probabilities for $X_j(t+1)$ for $j \neq i$ in the numerator and denominator of (A-3). Then $2(N-1)$ probabilities cancel out and we obtain, instead of (A-3),

$$\widetilde{f}_i = \frac{P[X_i^*(t)=1|X(t-1)]\; P[X_i(t+1)|X_i^*(t)=1, X_{-i}(t)]}{P[X_i^*(t)=0|X(t-1)]\; P[X_i(t+1)|X_i^*(t)=0, X_{-i}(t)]}, \tag{A-5}$$

and, instead of (A-4),

$$\widetilde{p}_i = \frac{\widetilde{f}_i}{1+\widetilde{f}_i}. \tag{A-6}$$

In our MH-algorithm, we calculate (A-6) for each missing patch in year t and then propose, for each patch i independently, $X_i^*(t)=1$ with probability $\widetilde{p}_i$. The jumping probability from $X(t)$ to $X^*(t)$ at step u is thus

$$J_u[(X^*(t)|X(t)] = \prod_i \widetilde{p}_i^{X_i^*(t)}(1-\widetilde{p}_i)^{(1-X_i^*(t))}. \tag{A-7}$$

By interchanging the role of $X^*(t)$ and $X(t)$ in (A-5) - (A-7), we also obtain $J_u[(X(t)|X^*(t)]$ so that with

$$f = \frac{P[X^*(t)|X(t-1)]\; P[X(t+1)|X^*(t)]}{P[X(t)|X(t-1)]\; P[X(t+1)|X(t)]} \tag{A-8}$$

the acceptance ratio r is

$$r = f\frac{J_u[X(t)|X^*(t)]}{J_u[X^*(t)|X(t)]}. \tag{A-9}$$

Updating missing values in the last year (T) is not covered by the foregoing, but it can be done by Gibbs sampling, because $P[X_i(T)|X(T-1), X_{-i}(T)] = P[X_i(T)|X(T-1)]$ which can be calculated explicitly.

It is instructive to see which proposals are generated when E_i and C_i would be constant over time (i.e. when the connectivity of a patch remains the same over time). If $X_i(t-1)=0$ and $X_i(t+1)=0$, then $\widetilde{f}_i = C_iE_i/(1-C_i)^2$ which will be a small value if C_i and/or E_i are small, so that $\widetilde{p}_i$ is also close to 0, so that with only a small probability a 1 is proposed in year t. Similar considerations show that if $X_i(t-1)=1$ and $X_i(t+1)=1$, $\widetilde{p}_i$ is close to 1 so that with a large probability a 1 is proposed in year t. If $X_i(t-1) \neq X_i(t+1)$, $\widetilde{f}_i = (1-E_i)/(1-C_i)$ so that if C_i and E_i are equal or both small, $\widetilde{p}_i$ is about one half. All of these proposals make intuitive sense.

In the tree frog case study, our proposal mechanism accepted more than half of the proposals for a year. In simulated data, with lower turnover rates, acceptance probabilities of over 0.9 were observed. Our algorithm is expected to outperform the Gibbs sampler when many or all patches in a particular year have missing data, as is the case in our pre-years. However, if there is only one patch missing in a particular year, the Gibbs sampler is the

most efficient and in that case we use it in our computer software. In the case study with its 202 patches and 25 pre-years, our MH-algorithm was about 10 times faster than an efficient implementation of the Gibbs sampler.

Literature cited

O'Hara, R.B., E. Arjas, H. Toivonen and I. Hanski. 2002. Bayesian analysis of metapopulation data. *Ecology* 83: 2408-2415.