Appendix A. Detailed description of the implementation of the model in the kfdnm R package.

Dail and Madsen (2011) first proposed a maximum likelihood approach for making inference for open *N*-mixture model parameters by using the transition kernel for the *N* process. Here, we augment their full-data likelihood with the known-fate portion of the model to obtain the known-fate open *N*-mixture full-data likelihood

$$[\boldsymbol{\beta}, \boldsymbol{\beta}^*, \boldsymbol{\rho}, \boldsymbol{\alpha} | \mathbf{n}, \mathbf{R}, \mathbf{N}] = \prod_{m=1}^{M} \prod_{i=1}^{I} \prod_{t=1}^{T} [Y_{mit} | Y_{mi,t-1}] \times \prod_{i=1}^{I} \prod_{t=1}^{T} [n_{it} | N_{it}] \, [N_{it} | N_{i,t-1}] \quad \text{(A.1)}$$

where the bold vectors represent complete collections of the associated data and abundance, $[N_{i1}|N_{i0}]=[N_{i1}]$ represents the prior or initial distribution of abundance at the first time period. Dail and Madsen (2011) use a Poisson distribution, however, we propose the scale prior $[N_{i1}]=1/N_{i1}$ as a non-informative alternative (Link 2013). Note that the full data likelihood is separable into the known-fate and open *N*-mixture model portions. The known-fate portion is simply a product of conditional Bernoulli distributions, thus, that portion of the likelihood is readily computed. The full-data portion of the open *N*-mixture likelihood must be integrated over the latent abundance process to obtain the true likelihood of that portion of the model.

Here we focus on an efficient method to calculate the true likelihood of the open *N*-mixture model integrated over the dynamic abundance processes. Using the Markov transition kernel of the abundance process, $[N_{it}|N_{i,t-1}]$, the open *N*-mixture model can be formulated as a Hidden Markov Model (HMM; see Zucchini and MacDonald 2009) from which the log-likelihood is efficiently calculated using the forward algorithm (Zucchini and MacDonald 2009; pg. 47). When combined with the backward sampling algorithm, an efficient MCMC algorithm can be devised for Bayesian inference.

In the description, we will provide the definition of the HMM forward algorithm for just a single group. The total likelihood can then be calculated by summing the individual log-

likelihoods. First, as in Dail and Madsen (2011), let K be the defined upper bound for all N_{it}. We assume that for most, if not all, applications, this can be chosen appropriately. Then, let $\boldsymbol{\eta}$ be the row vector of initial abundance probabilities, $[N_1 = j];\ j = 0, \ldots, K$ from the chosen initial distribution. Next define $\mathbf{P}(n_{it})$ to be the diagonal matrix with entries $\text{Binomial}(n_{it}|j - R_{it}, p_{it});\ j = R_{it}, \ldots, K$ and 0 for $j < R_{it}$. Finally, $\boldsymbol{\Delta}_{t-1}$ is the state transition matrix with j,k entry $[N_{it} = k|N_{i,t-1} = j]$. Now, the HMM forward algorithm for calculation of the log-likelihood, l, proceeds as follows:

(1) Set:

$$w_1 = \boldsymbol{\eta}\mathbf{P}(n_{i1})\mathbf{1}, \phi_1 = \boldsymbol{\eta}\mathbf{P}(n_{i1})/w_1, \text{and } l = \log w_1, \tag{A.2}$$

(2) For $t = 2, \ldots, T$:

$$w_t = \boldsymbol{\phi_{t-1}\Delta_{t-1}}\mathbf{P}(n_{it})\mathbf{1}, \boldsymbol{\phi}_t = \boldsymbol{\phi_{t-1}\Delta_{t-1}}\mathbf{P}(n_{it})/w_t, \text{and } l = l + \log w_t\,. \tag{A.3}$$

While the reverse-time recursion of Dail and Madsen (2011) provides a computationally efficient method to calculate the open N-mixture models likelihood, the HMM also provides efficient methods for Bayesian inference via MCMC. An outline of an MCMC routine proceeds as follows:

(1) For current parameter vector, $\boldsymbol{\theta^{curr}} = (\boldsymbol{\beta}, \boldsymbol{\beta}^*, \boldsymbol{\rho}, \boldsymbol{\alpha})$,

1. Draw, $\boldsymbol{\theta^{prop}}$, from proposal distribution $[\boldsymbol{\theta^{prop}}|\boldsymbol{\theta^{curr}}]$
2. Set $\boldsymbol{\theta^{curr}} = \boldsymbol{\theta^{prop}}$ with probability

$$\min\left(1, \exp\{l(\boldsymbol{\theta^{prop}}) - l(\boldsymbol{\theta^{curr}})\}\frac{[\boldsymbol{\theta^{prop}}]}{[\boldsymbol{\theta^{curr}}]}\cdot\frac{[\boldsymbol{\theta^{curr}}|\boldsymbol{\theta^{prop}}]}{[\boldsymbol{\theta^{prop}}|\boldsymbol{\theta^{curr}}]}\right), \tag{A.4}$$

where $l(\boldsymbol{\theta})$ is the log-likelihood evaluated at $\boldsymbol{\theta}$, calculated using the forward algorithm and $[\boldsymbol{\theta}]$ is the prior distribution of the parameters.

If an MCMC sample of the abundance vector, **N**, is desired proceed to (2), otherwise repeat step one as desired to obtain a sufficient posterior sample.

(2) Given a value, $\boldsymbol{\theta}$, from (1):

1. Run the forward algorithm and retain the $\boldsymbol{\phi}_t$ and $\boldsymbol{\Delta}_t$. The forward run from (1) can be used to avoid re-running the algorithm.
2. Draw from $[N_{gT}|\mathbf{n}] = \boldsymbol{\phi}_T$.
3. For t = *T*-1,…,1:

   Draw from $[N_{it}|N_{i,t+1}, \dots, N_{iT}, \boldsymbol{\theta}, \mathbf{n}] \propto \boldsymbol{\phi}_t \cdot \boldsymbol{\delta}_t^{N_{i,t+1}}$, where the product is element-wise and $\boldsymbol{\delta}_t^{N_{i,t+1}}$ is the $N_{i,t+1}$column $\boldsymbol{\Delta}_t$ (Zucchini and MacDonald 2009).

(3) If desired, the S_{it} and G_{it} processes can be directly sampled following updates of the N_{it} process:

1. Draw from:

   $$[S_{it}|N_{i,t-1}, N_{it}] \propto \text{Binomial}(S_{it}|N_{i,t-1} - R_{i,t-1}, \omega^*)\text{Poisson}(N_{i,t} - S_{it}|\gamma_{it}),$$

   for S_{it} in $\{0,\dots, N_{it} - R_{it}\}$, (A.5)
2. Set $G_{it} = N_{it} - S_{it}$. (A.6)

By sampling the hidden state, **N**, as a single vector versus individually given immediate neighbors, i.e., not drawing from $[N_{it}|N_{i,t-1}, N_{i,t+1}]$, within the MCMC, high autocorrelation and a slowly converging chain can be avoided. In addition, sampling of the N_{it} process by construction via serially correlated S_{it} and G_{it} samples can be avoided. Of course, if direct inference on the N_{it} process itself is not desired, step (2) is not necessary for Bayesian inference

of $\boldsymbol{\theta}$. This reduces Monte Carlo autocorrelation in $\boldsymbol{\theta}$ due to parameter updates being conditioned on the latent abundance updates.

LITERATURE CITED

Dail, D., and L. Madsen. 2011. Models for estimating abundance from repeated counts of an open metapopulation. Biometrics 67:577–587.

Link, W. A. 2013. A cautionary note on the discrete uniform prior for the binomial *N*. Ecology 94:2173–2179.

Zucchini, W., and I. L. MacDonald. 2009. Hidden Markov Models for Time Series: An Introduction Using R. CRC Press, New York. 275 pp.