



# State-space models for the dynamics of wild animal populations

S.T. Buckland<sup>a,\*</sup>, K.B. Newman<sup>b</sup>, L. Thomas<sup>a</sup>, N.B. Koesters<sup>a</sup>

<sup>a</sup> Centre for Research into Ecological and Environmental Modelling, University of St Andrews,  
The Observatory, Buchanan Gardens, St. Andrews KY16 9LZ, Scotland, UK

<sup>b</sup> Division of Statistics, University of Idaho, Moscow, ID 83844, USA

Received 27 February 2003; received in revised form 9 July 2003; accepted 12 August 2003

## Abstract

We develop a unified framework for jointly defining population dynamics models and measurements taken on a population. The framework is a state-space model where the population processes are modelled by the state process and measurements are modelled by the observation process. In many cases, the expected value for the state process can be represented as a generalisation of the standard population projection matrix: each sub-process within the state process may be modelled by a separate matrix and the product of these matrices is a generalised Leslie matrix. By selecting appropriate matrices and their ordering, a wide range of models may be specified. The method is fully flexible for allowing stochastic variation in the processes. Process parameters may themselves be modelled as functions of covariates. The structure accommodates effects such as density dependence, competition and predator–prey relationships, and metapopulations are readily modelled. Observations on the population enter through an observation process model, and we show how likelihood functions can be built that reflect both demographic stochasticity (which appears in the state process) and stochastic errors in the observations. Parameter estimation and estimation of state process variables can be conducted using sequential Monte Carlo procedures.

© 2003 Elsevier B.V. All rights reserved.

**Keywords:** Generalised Leslie matrix; Open population models; Population dynamics models; Recursive filtering algorithms; State-space models

## 1. Introduction

To date, wildlife assessment methods have been developed without the benefit of any underlying general framework. Borchers et al. (2002) show how methods for closed populations may be unified. They propose a general likelihood-based approach with two model components, which we label here a state process model and an observation process model. In Chapter 13 of Borchers et al. (2002), a state-space framework is pro-

posed for open populations, in which the state process specifies the population dynamics model (PDM). In this paper, we develop this suggestion, and show how a wide range of models may be specified and fit.

The developments of this paper were motivated by the need for a general approach for modelling the dynamic processes of managed wildlife populations. Several concurrent projects being developed by the authors required different models, but some components of those models were common across projects, suggesting a modular approach to model building and fitting. The primary motivating projects are a metapopulation of UK grey seals *Halichoerus grypus*, for which estimates of future abundance are

\* Corresponding author. Tel.: +44-1334-461841;

fax: +44-1334-461800.

E-mail address: [steve@mcs.st-and.ac.uk](mailto:steve@mcs.st-and.ac.uk) (S.T. Buckland).

required (Thomas et al., in press); introduced sika deer *Cervus nippon* in Scotland, where the interest is in predicting future increase in range, and interactions (competition and hybridisation) with native red deer; and assessment of demersal fish stocks in the North Sea. In each case, we wish to model the dynamics of the population, using survey data to fit our models. For the grey seals, the survey data are counts of pups at breeding colonies; for the sika deer, they are line transect data on dung pellet groups, from which population abundance is estimated by forest; and for the North Sea fisheries, the data are landing and discarding rates aboard sampled fishing vessels.

Often, change in wildlife populations is assessed by estimating trends from survey data empirically (e.g. Fewster et al., 2000). For many managed populations, we have sufficient data to allow us to model demographic variation, while smoothing out stochastic fluctuation in the survey data (e.g. Raftery et al., 1995; Newman, 1998, 2000; Meyer and Millar, 1999; Millar and Meyer, 2000; Trenkel et al., 2000; Besbeas et al., 2002). Raftery et al. (1995) used deterministic Leslie matrix models (Leslie, 1945, 1948), whereas the other authors used state-space models with stochastic errors. We show how to generalise that state-space structure to encompass a wide range of process and observation models.

The inferential framework uses Bayesian methods. A model is parameterised, and prior distributions are specified for the parameters, and for the initial state of the population. These priors reflect knowledge of the system prior to the start of the study of interest, and may be based on information from past studies, or expert knowledge, or may simply be vague priors, reflecting a lack of knowledge. Parameter sets and populations are simulated from these priors, and computer-intensive methods are used to generate samples from the joint posterior distribution of the parameters and the population. The posterior distribution updates the prior distribution, using the data from the study of interest. Inference on the population parameters, including assessment of uncertainty, is drawn from the posterior sample. Model uncertainty can be readily accommodated by specifying more than one model, and specifying a prior model distribution. Typically, this would be an uninformative distribution; if  $m$  models were specified, each model would be given a prior probability of  $1/m$ .

Many published models for population dynamics could be implemented using the general framework we develop. Features of our approach that are frequently lacking from published models include: the ability to specify a number of processes, both deterministic and stochastic, in a single model; an inferential framework for fitting population dynamics models, given observational data; the ability to allow for model uncertainty; an ability to quantify risk effectively using management models based on our methodology, which allow for the three major sources of uncertainty (demographic stochasticity, observational error and model uncertainty); complex models can be specified in a simple, modular way, in which simple sub-processes can be assembled like building blocks, to create a wide range of models. Recently published models that could utilise our approach include ones for Florida Key deer *Odocoileus virginianus clavium* (Peterson et al., 2003), grey wolves *Canis lupus* (Miller et al., 2002), short-tailed shearwaters *Puffinus tenuirostris* (Yearsley et al., 2003), black-capped vireos *Vireo atricapillus* (Parysow and Tazik, 2002), large blue butterflies *Maculinea arion* (Griebeler and Seitz, 2002), and prickly acacia *Acacia nilotica* (Kriticos et al., 2003).

The models of this paper largely fit into the framework for matrix models laid down by Caswell (2001). We depart from that framework in two major respects. First we put population processes into the state process and observations into the observation process, to give a more general framework for model fitting, in which the observations are fully integrated into the modelling. Model definition is the other main departure from Caswell; we separate the processes (stochastic and deterministic), so that each individual process can easily be defined, and combined to give an overall model with the required structure. This makes the task of model-building relatively straightforward, in marked contrast with formulating complex models in a single step.

As an example of our ‘modular’ approach to model definition, consider a population with three age classes, updated using a simple Leslie matrix model (Caswell, 2001):

$$\begin{bmatrix} n_{0,t} \\ n_{1,t} \\ n_{2,t} \end{bmatrix} = \begin{bmatrix} \lambda & \lambda & \lambda \\ \phi_1 & 0 & 0 \\ 0 & \phi_2 & \phi_2 \end{bmatrix} \begin{bmatrix} n_{0,t-1} \\ n_{1,t-1} \\ n_{2,t-1} \end{bmatrix}$$

where  $n_{i,t}$  is the number of animals of age  $i$  in year  $t$ ,  $i = 0, 1$ ,  $n_{2,t}$  is the number of animals of age  $\geq 2$  in year  $t$ ,  $\lambda$  is the mean number of young produced per year per animal, and  $\phi_i$  is the annual probability of survival of an animal of age  $i$ .

Now suppose we rewrite this model as

$$\begin{bmatrix} n_{0,t} \\ n_{1,t} \\ n_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_0 & 0 & 0 \\ 0 & \phi_1 & 0 \\ 0 & 0 & \phi_2 \end{bmatrix} \begin{bmatrix} 0 & \lambda & \lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} n_{0,t-1} \\ n_{1,t-1} \\ n_{2,t-1} \end{bmatrix}$$

Multiplying the three matrices together, we obtain:

$$\begin{bmatrix} n_{0,t} \\ n_{1,t} \\ n_{2,t} \end{bmatrix} = \begin{bmatrix} \lambda\phi_0 & \lambda\phi_0 & \lambda\phi_0 \\ \phi_1 & 0 & 0 \\ 0 & \phi_2 & \phi_2 \end{bmatrix} \begin{bmatrix} n_{0,t-1} \\ n_{1,t-1} \\ n_{2,t-1} \end{bmatrix}$$

Note that this is slightly different from above: the term  $\lambda\phi_0$  ensures that only newly-born animals that survive their first year contribute to the population at the end of that first year. The previous formulation modelled ‘recruits’ (number of animals entering the population at the end of their first year) rather than births.

Consider the three matrices separately:

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

This handles the aging process: animals of age 0 become aged 1 at the end of their first year, while 1-year-old animals are amalgamated with animals of age 2+.

$$\begin{bmatrix} 0 & \lambda & \lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This generates new births. The 1’s on the diagonal leave animals aged 1 or 2+ unaltered, while the first row creates the new age 0 animals.

$$\begin{bmatrix} \phi_0 & 0 & 0 \\ 0 & \phi_1 & 0 \\ 0 & 0 & \phi_2 \end{bmatrix}$$

This matrix reduces each age class to those that survive to the next year.

The above formulation is deterministic. In the development below, we extend it to be fully stochastic.

## 2. State-space models: structure and inference

A state-space model describes the evolution of two time series running in parallel, one referred to as the state process and the other as the observation process. The state process, denoted  $\mathbf{n}_t$ ,  $t = 0, 1, \dots, T$ , is an unobservable vector. In this paper,  $\mathbf{n}_t$  represents a vector of abundances of an animal population, in which the elements are numbers of animals by category, for example, number of age 1 females. The observation process, denoted  $\mathbf{y}_t$ ,  $t = 1, \dots, T$ , is a completely observable vector that is a function of the state process. A complete realisation of the observation process will be denoted  $\mathbf{y}^T = \mathbf{y}_1, \dots, \mathbf{y}_T$ . The observation vector  $\mathbf{y}_t$  has components that correspond to some or all state components, or aggregated subsets of them, and may be measured with or without error.

Let  $g(\cdot)$  and  $f(\cdot)$  represent either probability density or mass functions for the state and observation processes, respectively. Throughout the paper we refer to such functions as pdf’s whether the probabilities are defined with respect to Lebesgue or counting measures; for simplicity, when writing integrals we write them as if they were density functions. A state-space model can be written as a set of three pdfs:

$$g_0(\mathbf{n}_0; \Theta) \quad \text{Initial state distribution} \quad (1)$$

$$g_t(\mathbf{n}_t|\mathbf{n}_{t-1}; \Theta) \quad \text{State process distribution} \quad (2)$$

$$f_t(\mathbf{y}_t|\mathbf{n}_t; \Theta) \quad \text{Observation process distribution} \quad (3)$$

where  $t = 1, 2, \dots, T$  and  $\Theta$  is a vector of parameters. The state process is assumed to be first-order Markov, namely,

$$g(\mathbf{n}_t|\mathbf{n}_{t-1}, \dots, \mathbf{n}_0; \Theta) = g(\mathbf{n}_t|\mathbf{n}_{t-1}; \Theta)$$

Statistical inferences for a state-space model can be categorised as inferences about the state  $\mathbf{n}_t$  and inferences about the parameter  $\Theta$  conditional on the observation process. All inferences can be viewed as the result of integration. For example, to find the maximum likelihood estimate for  $\Theta$ , the likelihood must be evaluated. This involves integrating or summing the joint distribution of  $\mathbf{n}_t$  and  $\mathbf{y}_t$  over  $\mathbf{n}_t$ ,

$t = 0, 1, \dots, T$ :

$$p(\mathbf{y}^T | \Theta) = \int_{\mathbf{n}_0} \dots \int_{\mathbf{n}_T} \left\{ \prod_{t=1}^T f_t(\mathbf{y}_t | \mathbf{n}_t; \Theta) g_t(\mathbf{n}_t | \mathbf{n}_{t-1}; \Theta) d\mathbf{n}_t \right\} g_0(\mathbf{n}_0; \Theta) d\mathbf{n}_0 \tag{4}$$

Another example: given  $\Theta$  and  $\mathbf{y}^T$ , the expectation of the state vector is found again by integrating over the states:

$$E[\mathbf{n}_t | \mathbf{y}^T; \Theta] = \frac{\int_{\mathbf{n}_0} \dots \int_{\mathbf{n}_T} \mathbf{n}_t \left\{ \prod_{t=1}^T f_t(\mathbf{y}_t | \mathbf{n}_t; \Theta) g_t(\mathbf{n}_t | \mathbf{n}_{t-1}; \Theta) d\mathbf{n}_t \right\} g_0(\mathbf{n}_0; \Theta) d\mathbf{n}_0}{p(\mathbf{y}^T | \Theta)} \tag{5}$$

The above calculation about  $\mathbf{n}_t$  given the entire realisation of the observation process is known as smoothing. Inferences about  $\mathbf{n}_t$  given  $\mathbf{y}^{t-1}$  and  $\mathbf{y}^t$  have been referred to as one-step ahead prediction and filtering, respectively (West and Harrison, 1990). A third example: Bayesian inference centres on the posterior distribution for  $\Theta$ , which again involves integration:

$$p(\Theta | \mathbf{y}^T) = \frac{p(\mathbf{y}^T | \Theta) p(\Theta)}{f(\mathbf{y}^T)} \tag{6}$$

where  $p(\Theta)$  is the prior pdf for  $\Theta$  and  $f(\mathbf{y}^T)$  can be found by integrating  $p(\mathbf{y}^T | \Theta) p(\Theta)$  over  $\Theta$ .

### 2.1. Modularisation of the state process

The state process is a stochastic representation of a PDM, many of which have several sub-processes in common, in particular, survival, birth and movement. We believe that it is advantageous to modularise the state process pdf in terms of separate sub-process pdfs. Attention can be directed to one particular sub-process in isolation, alternative theories for a sub-process can perhaps be more readily generated, and this in turn provides a structure for hypothesis testing. For example, in a multi-agency management setting, there might be general agreement amongst biologists on the nature of survival and movement processes, but disagreement over whether births are density independent or dependent.

Suppose that there are three sub-processes, survival, movement and birth, occurring in sequence that generate  $\mathbf{n}_t$  from  $\mathbf{n}_{t-1}$ . This might correspond for example to winter survival, movement in spring, and births in early summer. We assume that processes may be discretised in this way, although there is in principle no limit on the number of sub-processes, so for example

a sub-process for summer survival or autumn movement may be added, if required. The evolution from  $\mathbf{n}_{t-1}$  to  $\mathbf{n}_t$  is described by a series of three linked pdfs, where the input to one pdf is output from a previous pdf. The survival process has pdf  $g_{s,t}$ , the movement process has pdf  $g_{m,t}$ , and the birth process has pdf  $g_{b,t}$ . The realisations of the successive sub-processes are denoted by vectors  $\mathbf{u}_{s,t}$ ,  $\mathbf{u}_{m,t}$ , and  $\mathbf{u}_{b,t}$ , where  $\mathbf{u}_{b,t} = \mathbf{n}_t$ . The distributions of the sub-processes in year  $t$  can be written as follows:

$$\begin{aligned} \mathbf{u}_{s,t} &\sim H_{s,t}(\mathbf{n}_{t-1}, \Theta) \\ \mathbf{u}_{m,t} &\sim H_{m,t}(\mathbf{u}_{s,t}, \Theta) \\ \mathbf{n}_t &\sim H_{b,t}(\mathbf{u}_{m,t}, \Theta) \end{aligned}$$

where each  $H$  is a distribution corresponding to the appropriate pdf.

The cost of the added flexibility of modularising the state process is an increase in the complexity of the state pdf and the likelihood. To evaluate the state process pdf, we must integrate over the sub-processes appropriately, where appropriate integration means that the spaces over which we integrate are ones constrained by the values of  $\mathbf{n}_{t-1}$  and  $\mathbf{n}_t$ . Given the above three sub-processes, and assuming that each sub-process is first-order Markov, the state pdf is

$$\begin{aligned} &g_t(\mathbf{n}_t | \mathbf{n}_{t-1}; \Theta) \\ &= \int_{\mathbf{u}_{m,t}} \int_{\mathbf{u}_{s,t}} g_t(\mathbf{n}_t, \mathbf{u}_{m,t}, \mathbf{u}_{s,t}, | \mathbf{n}_{t-1}; \Theta) d\mathbf{u}_{s,t} d\mathbf{u}_{m,t} \\ &= \int_{\mathbf{u}_{m,t}} \int_{\mathbf{u}_{s,t}} g_{s,t}(\mathbf{u}_{s,t} | \mathbf{n}_{t-1}; \Theta) \\ &\quad \times g_{m,t}(\mathbf{u}_{m,t} | \mathbf{u}_{s,t}; \Theta) g_{b,t}(\mathbf{n}_t | \mathbf{u}_{m,t}; \Theta) d\mathbf{u}_{s,t} d\mathbf{u}_{m,t} \end{aligned}$$

More generally, suppose there are  $k + 1$  sub-processes occurring in sequence that generate  $\mathbf{n}_t$  from  $\mathbf{n}_{t-1}$ . The

state process pdf can be written as follows:

$$g_t(\mathbf{n}_t|\mathbf{n}_{t-1}; \Theta) = \int_{\mathbf{u}_{k,t}} \dots \int_{\mathbf{u}_{1,t}} g_{1,t}(\mathbf{u}_{1,t}|\mathbf{n}_{t-1}; \Theta) \times \left[ \prod_{i=2}^k g_{i,t}(\mathbf{u}_{i,t}|\mathbf{u}_{i-1,t}; \Theta) \right] \times g_{k+1,t}(\mathbf{n}_t|\mathbf{u}_{k,t}; \Theta) d\mathbf{u}_{1,t} \dots d\mathbf{u}_{k,t}$$

As later examples will show, the complexity of such likelihood evaluations makes sequential Monte Carlo procedures that avoid calculating the likelihood in the first place very attractive.

We have adopted the following conventions. In cases where age of animal is relevant, advancement to the next age class is assumed to occur just prior to the birth process. This simplifies the matrix representation of the sub-processes discussed in the next section. A second convention is that the time of advancement from time period  $t - 1$  to  $t$  coincides with the time of observation under the assumption that just one survey occurs in a given time interval  $(t - 1, t]$ . In fact, the methods readily extend to variable numbers of surveys per period (including none), and to irregular time periods between surveys. A third convention is that, unless stated otherwise,  $\mathbf{n}_0$  will be the assumed starting point, rather than any of the intermediate states between  $\mathbf{n}_0$  and  $\mathbf{n}_1$ .

### 2.2. Matrix representations

Under certain conditions, the conditional expectation of the states at time  $t$ , given the states at time  $t - 1$ , which we denote by  $E_{\mathbf{n}_t|\mathbf{n}_{t-1}}[\mathbf{n}_t]$ , can be written as a product of matrices and  $\mathbf{n}_{t-1}$ . When possible, this makes clear the sequencing of sub-processes and their average effect. Suppose there are three sub-processes, i.e.  $\mathbf{u}_{3,t} = \mathbf{n}_t$ .

$$E_{\mathbf{n}_t|\mathbf{n}_{t-1}}[\mathbf{n}_t] = E_{\mathbf{u}_{1,t}|\mathbf{n}_{t-1}}\{E_{\mathbf{u}_{2,t}|\mathbf{u}_{1,t}}[E_{\mathbf{n}_t|\mathbf{u}_{2,t}}(\mathbf{n}_t)]\}$$

If all the sub-processes are first-order Markov, and if the expectation of any sub-process, say  $\mathbf{u}_{i,t}$ , conditional on  $\mathbf{u}_{i-1,t}$  is a linear function of  $\mathbf{u}_{i-1,t}$ ,  $E_{\mathbf{n}_t|\mathbf{n}_{t-1}}[\mathbf{n}_t]$  can be written in terms of matrix products, i.e.

$$E_{\mathbf{n}_t|\mathbf{n}_{t-1}}[\mathbf{n}_t] = M_{k+1}M_kM_{k-1} \dots M_1\mathbf{n}_{t-1}$$

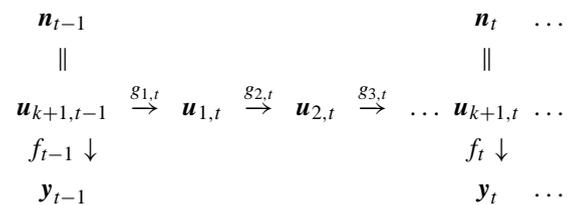
where each matrix  $M$  corresponds to a single sub-process. In such cases, the product of the matrices corresponds to a generalised form of a Leslie matrix (Caswell, 2001, pp. 8–34) or Lefkovich matrix (Caswell, 2001, pp. 56–62). By splitting such a matrix into its components, we obtain a convenient framework for defining and explaining model structure.

Certain sub-processes are common to several of the examples given later and will be denoted as follows.  $S$  is a matrix of survival parameters and may be further subscripted by time or age, for example.  $B$  is a matrix of birth or recruitment rates.  $G$  comprises parameters that assign sex to newly-born animals.  $X$  contains parameters that reflect a change in the state of an animal arising from an experiment (e.g. removal in a removal experiment, or marking in a mark–recapture experiment).

We also define an advancement matrix  $A$ . For age-structured populations, this typically advances the age of an animal by one at the appropriate time point, and so is deterministic. For stage-structured populations, the matrix will typically represent stochastic processes, such as a probability of an immature animal maturing in a given time period.

### 2.3. The observation process

The place of the observation process in this structure is represented schematically as follows:



As some of the examples will demonstrate, the components of the observation vector are sometimes exactly equal to components of the state vector or the components are summations of state vector components with no stochastic error. The effect on inference appears in the likelihood (Eq. (4)) where the term  $f(\mathbf{y}_t|\mathbf{n}_t)$  equals 1 for  $\mathbf{n}_t$  that map exactly or correctly into  $\mathbf{y}_t$  and  $f(\mathbf{y}_t|\mathbf{n}_t)$  equals 0, otherwise. Evaluation of the integral for the likelihood then involves integrating the state process pdf alone over a particular

subspace of  $\mathbf{n}_0, \dots, \mathbf{n}_T$ . More concisely, let  $m(\mathbf{n}_t)$  be the mapping of  $\mathbf{n}_t$  to  $\mathbf{y}_t$  and define the sets  $\eta_t$ , where

$$\eta_t = \{\mathbf{n}_t : m(\mathbf{n}_t) = \mathbf{y}_t\} \tag{7}$$

Then

$$p(\Theta|y^T) = \int_{\mathbf{n}_0} \int_{\eta_1} \dots \int_{\eta_T} \left\{ \prod_{t=1}^T g(\mathbf{n}_t|\mathbf{n}_{t-1}; \Theta) d\mathbf{n}_t \right\} \times g(\mathbf{n}_0; \Theta) d\mathbf{n}_0 \tag{8}$$

(Note that  $\mathbf{n}_0$  is free of such constraints, as the first observation,  $y_1$ , occurs at time point 1.)

### 3. Example populations

We illustrate the flexibility of the methods through a series of examples. Each is simpler than most real examples are likely to be, but together they serve to convey the power of the framework. The first example might be appropriate for a small, isolated population of a species of bird, which breeds at age one, and for which an approximate census can be conducted. The second example might be suitable for a population of large mammals, for which females do not all mature at the same age, and on which a removal experiment is being conducted. The third example shows how mark–recapture fits into the general framework. In the fourth example, density dependence is incorporated in recruitment, and survival rates depend on weather and age. A metapopulation, comprising three colonies, is considered in example five, and the final example shows how a predator–prey system can be modelled.

#### 3.1. Two age classes; census with errors

Animals are categorised as juveniles and adults, where a juvenile becomes an adult at 1 year of life. The number of juveniles in year  $t$  present after birth is denoted  $n_{j,t}$  and the number of adults at the same time is  $n_{a,t}$ . The annual survival rates for juveniles and adults, denoted  $\phi_j$  and  $\phi_a$ , cover the time just after birth to the beginning of the next birthing period (nearly a year later). The probability of an animal surviving is assumed independent of any other animal; i.e. the number surviving in a given category is binomially

distributed. Births occur annually in the spring; the number born is distributed as Binomial( $n_{a,t}, b$ ) (thus density independent). A census of juveniles and adults is attempted each year just after the breeding season. The number seen in each category is modelled by a Normal( $n_{i,t}, \sigma_i^2$ ), where  $n_{i,t}$  refers to the number of animals in category  $i$  in year  $t$ , and  $\sigma_i^2$  is the variance. The initial abundance  $\mathbf{n}_0$  is an unknown vector.

In this example, there are two stochastic sub-processes, survival and birth. To make the connection with the matrix formulation clearer, an additional non-stochastic sub-process, namely advancement in age, is inserted between survival and birth. The distributions of the sub-processes are as follows:

$$\begin{aligned} \mathbf{u}_{s,t} &\sim H_{s,t}(\mathbf{n}_{t-1}) : \begin{bmatrix} u_{j,s,t} \sim \text{Binomial}(n_{j,t-1}, \phi_j) \\ u_{a,s,t} \sim \text{Binomial}(n_{a,t-1}, \phi_a) \end{bmatrix} \\ \mathbf{u}_{a,t} &\sim H_{a,t}(\mathbf{u}_{s,t}) : \begin{bmatrix} u_{j,a,t} = 0 \\ u_{a,a,t} = u_{j,s,t} + u_{a,s,t} \end{bmatrix} \\ \mathbf{n}_t &\sim H_{b,t}(\mathbf{u}_{a,t}) : \begin{bmatrix} n_{j,t} \sim \text{Binomial}(u_{a,a,t}, b) \\ n_{a,t} = u_{a,a,t} \end{bmatrix} \end{aligned}$$

The initial state vector in this case could be  $\mathbf{u}_{a,0}$  or simply the number of adults,  $u_{a,a,0}$ .

The pdf for the state process is found by integrating the joint pdfs for the survival and birth sub-processes over the survival process with the constraint that the number of adults at time  $t$  equals the sum of the number of surviving juveniles and the number of surviving adults from time  $t - 1$ :

$$\begin{aligned} &g(\mathbf{n}_t|\mathbf{n}_{t-1}; \Theta) \\ &= \int_{\mathbf{u}_{a,t}} \int_{\mathbf{u}_{s,t}} g_{b,t}(\mathbf{n}_t|\mathbf{u}_{a,t}; \Theta) g_{a,t}(\mathbf{u}_{a,t}|\mathbf{u}_{s,t}; \Theta) \\ &\quad \times g_{s,t}(\mathbf{u}_{s,t}|\mathbf{n}_{t-1}; \Theta) d\mathbf{u}_{s,t} d\mathbf{u}_{a,t} \\ &= \sum_{u_{a,s,t}=\max(0, n_{a,t}-n_{j,t-1})}^{\min(n_{a,t-1}, n_{a,t})} \binom{n_{a,t}}{n_{j,t}} b^{n_{j,t}} (1-b)^{n_{a,t}-n_{j,t}} \\ &\quad \times \binom{n_{j,t-1}}{n_{a,t}-u_{a,s,t}} \phi_j^{n_{a,t}-u_{a,s,t}} \\ &\quad \times (1-\phi_j)^{(n_{j,t-1}-n_{a,t}+u_{a,s,t})} \binom{n_{a,t-1}}{u_{a,s,t}} \phi_a^{u_{a,s,t}} \\ &\quad \times (1-\phi_a)^{(n_{a,t-1}-u_{a,s,t})} \end{aligned}$$

The observation process pdf is the product of two independent normal pdfs, corresponding to the counts of juveniles and adults:

$$f(\mathbf{y}_t | \mathbf{n}_t, \Theta) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left[-\frac{0.5(y_{j,t} - n_{j,t})^2}{\sigma_j^2}\right] \times \frac{1}{\sqrt{2\pi\sigma_a^2}} \exp\left[-\frac{0.5(y_{a,t} - n_{a,t})^2}{\sigma_a^2}\right]$$

where  $\Theta = (\phi_j, \phi_a, b, \sigma_j^2, \sigma_a^2)$ .

As the state and observation process pdf's indicate, the likelihood (Eq. 14) is involved to evaluate. Generally, estimation procedures which do not require evaluation of the likelihood are preferable.

$E_{\mathbf{n}_t | \mathbf{n}_{t-1}}[\mathbf{n}_t]$  can be written as a product of matrices.

$$\begin{aligned} E_{\mathbf{n}_t | \mathbf{n}_{t-1}}[\mathbf{n}_t] &= E_{\mathbf{u}_{s,t} | \mathbf{n}_{t-1}} \{ E_{\mathbf{u}_{a,t} | \mathbf{u}_{s,t}} [ E_{\mathbf{n}_t | \mathbf{u}_{a,t}}(\mathbf{n}_t) ] \} \\ &= E_{\mathbf{u}_{s,t} | \mathbf{n}_{t-1}} \left\{ E_{\mathbf{u}_{a,t} | \mathbf{u}_{s,t}} \left( \begin{bmatrix} 0 & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_{j,a,t} \\ u_{a,a,t} \end{bmatrix} \right) \right\} \\ &= E_{\mathbf{u}_{s,t} | \mathbf{n}_{t-1}} \left\{ \begin{bmatrix} 0 & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u_{j,s,t} \\ u_{a,s,t} \end{bmatrix} \right\} \\ &= \begin{bmatrix} 0 & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \phi_j & 0 \\ 0 & \phi_a \end{bmatrix} \begin{bmatrix} n_{j,t-1} \\ n_{a,t-1} \end{bmatrix} \end{aligned}$$

or more compactly

$$E_{\mathbf{n}_t | \mathbf{n}_{t-1}}[\mathbf{n}_t] = \mathbf{BAS}_{\mathbf{n}_{t-1}}$$

Making the analogy with a single Leslie matrix, the product of the three matrices is

$$\begin{bmatrix} b\phi_j & b\phi_a \\ \phi_j & \phi_a \end{bmatrix}$$

### 3.2. Five age/stage/sex classes; removal experiment

Males are categorised as juveniles or adults (age  $\geq 1$ ), and females as juveniles, immature or mature. Juvenile females become immature with certainty at age 1, but then become mature with probability  $\beta$  at subsequent birthdays. Once mature, they remain mature for the duration of their life. Mature females produce either 0, 1 or 2 young with probability  $\pi_0, \pi_1$  and  $\pi_2$ , respectively, where  $\sum_{i=0}^2 \pi_i = 1$ . A newly-born animal is female with probability 0.5.

Juveniles are born at the beginning of summer and survive to the autumn with probability  $\phi_j$ . Survival rate the subsequent winter and spring (i.e. up to their first birthday) is  $\phi_w$ . The annual survival rate for animals aged  $\geq 1$  is  $\phi_a$ ; mortality during the summer is negligible.

At the beginning of autumn, a single sample of a removal experiment (Seber, 1982, pp. 309–315) is carried out, in which each animal has the same probability  $P$  of being removed. Animals are neither sexed nor aged on removal.

The state vector  $\mathbf{n}_t$  represents the state of the population immediately following the removal experiment and has 10 components:

$$\mathbf{n}_t = \begin{bmatrix} n_{m,0,u,t} \\ n_{m,1+,u,t} \\ n_{f,0,u,t} \\ n_{f,i,u,t} \\ n_{f,b,u,t} \\ n_{m,0,r,t} \\ n_{m,1+,r,t} \\ n_{f,0,r,t} \\ n_{f,i,r,t} \\ n_{f,b,r,t} \end{bmatrix}$$

where the first subscript is for sex ( $m$  or  $f$ ), the second is for age (0 or 1+ for males) or stage (0,  $i$  for immature, or  $b$  for mature breeder for females), the third is for removal status ( $r$  for removed and  $u$  for un-removed), and the fourth indicates year ( $t$ ).

Seven sub-processes were chosen to model this population: post-removal survival (winter and spring), maturation of some immatures, age advancement of juveniles, birth, sexing of young, summer survival, and removal. Fewer could be used, but convolutions of random variables then occur. For example, the birth and sexing sub-processes could be combined, and modelled as a multinomial with six categories, rather than the simpler strategy of one trinomial and one binomial distribution.

The pdf linkages of the sub-processes are as follows, beginning with post-removal survival. Age incrementing occurs at the end of the spring, when all surviving juveniles become age 1. The dimensioning of the intermediate random vectors, the  $\mathbf{u}_{i,t}$ , varies,

with components sure to be zero omitted. First, the post-removal survival,  $\mathbf{u}_{1,t} \sim H_{1,t}(\mathbf{n}_{t-1})$ :

$$\left[ \begin{array}{l} u_{1,m,0,u,t} \sim \text{Binomial}(n_{m,0,u,t-1}, \phi_w) \\ u_{1,m,1+,u,t} \sim \text{Binomial}(n_{m,1+,u,t-1}, \phi_a) \\ u_{1,f,0,u,t} \sim \text{Binomial}(n_{f,0,u,t-1}, \phi_w) \\ u_{1,f,i,u,t} \sim \text{Binomial}(n_{f,i,u,t-1}, \phi_a) \\ u_{1,f,b,u,t} \sim \text{Binomial}(n_{f,b,u,t-1}, \phi_a) \end{array} \right]$$

Advancement of some immature females to maturity occurs next,  $\mathbf{u}_{2,t} \sim H_{2,t}(\mathbf{u}_{1,t})$ :

$$\left[ \begin{array}{l} u_{2,m,0,u,t} = u_{1,m,0,u,t} \\ u_{2,m,1+,u,t} = u_{1,m,1+,u,t} \\ u_{2,f,0,u,t} = u_{1,f,0,u,t} \\ u_{2,f,i,u,t} \sim \text{Binomial}(u_{1,f,i,u,t}, 1 - \beta) \\ u_{2,f,b,u,t} = u_{1,f,b,u,t} + u_{1,f,i,u,t} - u_{2,f,i,u,t} \end{array} \right]$$

Then age advancement of juveniles,  $\mathbf{u}_{3,t} \sim H_{3,t}(\mathbf{u}_{2,t})$ :

$$\left[ \begin{array}{l} u_{3,m,1+,u,t} = u_{2,m,0,u,t} + u_{2,m,1+,u,t} \\ u_{3,f,i,u,t} = u_{2,f,0,u,t} + u_{2,f,i,u,t} \\ u_{3,f,b,u,t} = u_{2,f,b,u,t} \end{array} \right]$$

(Note that  $H_{3,t}(\mathbf{u}_{2,t})$  is a degenerate distribution, in which  $\mathbf{u}_{3,t}$  given  $\mathbf{u}_{2,t}$  is known with certainty.) Next come births,  $\mathbf{u}_{4,t} \sim H_{4,t}(\mathbf{u}_{3,t})$ :

$$\left[ \begin{array}{l} u_{4,\cdot,0,u,t} = x_1 + 2x_2 \text{ with } (x_0, x_1, x_2) \\ \quad \sim \text{Trinomial}(u_{3,f,m,u,t}, \pi_0, \pi_1, \pi_2) \\ u_{4,m,1+,u,t} = u_{3,m,1+,u,t} \\ u_{4,f,i,u,t} = u_{3,f,i,u,t} \\ u_{4,f,b,u,t} = u_{3,f,b,u,t} \end{array} \right]$$

Next young are partitioned into sexes,  $\mathbf{u}_{5,t} \sim H_{5,t}(\mathbf{u}_{4,t})$ :

$$\left[ \begin{array}{l} u_{5,m,0,u,t} \sim \text{Binomial}(u_{4,\cdot,0,u,t}, 0.5) \\ u_{5,m,1+,u,t} = u_{4,m,1+,u,t} \\ u_{5,f,0,u,t} = u_{4,\cdot,0,u,t} - u_{5,m,0,u,t} \\ u_{5,f,i,u,t} = u_{4,f,i,u,t} \\ u_{5,f,b,u,t} = u_{4,f,b,u,t} \end{array} \right]$$

Then summer survival of the juveniles,  $\mathbf{u}_{6,t} \sim H_{6,t}(\mathbf{u}_{5,t})$ :

$$\left[ \begin{array}{l} u_{6,m,0,u,t} \sim \text{Binomial}(u_{5,m,0,u,t}, \phi_j) \\ u_{6,m,1+,u,t} = u_{5,m,1+,u,t} \\ u_{6,f,0,u,t} \sim \text{Binomial}(u_{5,f,0,u,t}, \phi_j) \\ u_{6,f,i,u,t} = u_{5,f,i,u,t} \\ u_{6,f,b,u,t} = u_{5,f,b,u,t} \end{array} \right]$$

Lastly, removal,

$$\mathbf{n}_t = \left[ \begin{array}{l} n_{m,0,u,t} \sim \text{Binomial}(u_{6,m,0,u,t}, 1 - p) \\ n_{m,1+,u,t} \sim \text{Binomial}(u_{6,m,1+,u,t}, 1 - p) \\ n_{f,0,u,t} \sim \text{Binomial}(u_{6,f,0,u,t}, 1 - p) \\ n_{f,i,u,t} \sim \text{Binomial}(u_{6,f,i,u,t}, 1 - p) \\ n_{f,b,u,t} \sim \text{Binomial}(u_{6,f,b,u,t}, 1 - p) \\ n_{m,0,r,t} = u_{6,m,0,u,t} - n_{m,0,u,t} \\ n_{m,1+,r,t} = u_{6,m,1+,u,t} - n_{m,1+,u,t} \\ n_{f,0,r,t} = u_{6,f,0,u,t} - n_{f,0,u,t} \\ n_{f,i,r,t} = u_{6,f,i,u,t} - n_{f,i,u,t} \\ n_{f,b,r,t} = u_{6,f,b,u,t} - n_{f,b,u,t} \end{array} \right]$$

Rather than using  $\mathbf{n}_0$  directly as the initiating state, which requires specifying the numbers removed and not removed by sex, age and maturity, it would perhaps be more convenient to begin with  $\mathbf{u}_{6,1}$ , the vector of abundances just prior to removal in year 1.

Assuming that the removals are counted without error, given the state vector, the observation vector (a scalar) is simply the number of animals removed in year  $t$ , which is deterministic:

$$y_t = f(\mathbf{n}_t) = n_{m,0,r,t} + n_{m,1+,r,t} + n_{f,0,r,t} + n_{f,i,r,t} + n_{f,b,r,t}$$

Calculation of the likelihood is extremely involved given the number of sub-processes to integrate over. The observation vector's effect on the likelihood is to constrain the integration over  $\mathbf{n}_t$  such that the sum of removed animals equals  $y_t$ . Again simulation-based integration is required. Furthermore, to avoid identifiability problems, some of the parameters ( $\Theta = (\phi_a, \phi_w, \beta, \pi_0, \pi_1, \phi_j, p)$ ) will either have to be known, or informative priors used, or additional data collected and the observation process expanded accordingly. For example, given that only annual observations are made, the birth rate parameters  $\pi_0$  and  $\pi_1$  are confounded with the juvenile survival probability  $\phi_j$ .

The conditional expectations,  $E_{\mathbf{n}_t|\mathbf{n}_{t-1}}[\mathbf{n}_t]$ , can be easily expressed as a product of matrices representing each of the sub-processes:

$$E_{\mathbf{n}_t|\mathbf{n}_{t-1}}[\mathbf{n}_t] = XS_jGBA_2A_1S_w\mathbf{n}_{t-1}$$

where

$$X = \begin{bmatrix} 1-p & 0 & 0 & 0 & 0 \\ 0 & 1-p & 0 & 0 & 0 \\ 0 & 0 & 1-p & 0 & 0 \\ 0 & 0 & 0 & 1-p & 0 \\ 0 & 0 & 0 & 0 & 1-p \\ p & 0 & 0 & 0 & 0 \\ 0 & p & 0 & 0 & 0 \\ 0 & 0 & p & 0 & 0 \\ 0 & 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 & p \end{bmatrix} \quad S_j = \begin{bmatrix} \phi_j & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \phi_j & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad G = \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & \pi_1 + 2\pi_2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1-\beta & 0 \\ 0 & 0 & 0 & \beta & 1 \end{bmatrix}$$

$$S_w = \begin{bmatrix} \phi_w & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \phi_a & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \phi_w & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \phi_a & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \phi_a & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Again making the analogy with a single Leslie matrix, the product of the seven matrices is

$$\begin{bmatrix} 0 & 0 & 0 & 0.5\beta(1-p)(\pi_1 + 2\pi_2)\phi_j\phi_a & 0.5(1-p)(\pi_1 + 2\pi_2)\phi_j\phi_a & 0 & 0 & 0 & 0 & 0 \\ (1-p)\phi_w & (1-p)\phi_a & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5\beta(1-p)(\pi_1 + 2\pi_2)\phi_j\phi_a & 0.5(1-p)(\pi_1 + 2\pi_2)\phi_j\phi_a & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & (1-p)\phi_w & (1-\beta)(1-p)\phi_a & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta(1-p)\phi_a & (1-p)\phi_a & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5\beta p(\pi_1 + 2\pi_2)\phi_j\phi_a & 0.5p(\pi_1 + 2\pi_2)\phi_j\phi_a & 0 & 0 & 0 & 0 & 0 \\ p\phi_w & p\phi_a & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5\beta p(\pi_1 + 2\pi_2)\phi_j\phi_a & 0.5p(\pi_1 + 2\pi_2)\phi_j\phi_a & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & p\phi_w & (1-\beta)p\phi_a & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta p\phi_a & p\phi_a & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The advantage of separating the sub-processes is readily apparent.

3.3. Two age classes; mark–recapture survey

To simplify this example, we model the female population only. Animals have a constant survival rate  $\phi_j$  in their first year and an annual survival rate  $\phi_a$  subsequently. Mortality occurs primarily in the winter and spring. Breeding first occurs at age 1 and the number of births per adult is distributed as a Bernoulli random variable with probability of success  $\pi$ .

Births occur in summer and are followed by a single sample of a mark–recapture experiment in the autumn. It is assumed that there is no mortality in the period between births and the mark–recapture sample, although the model could be extended (by adding another sub-process) to include such mortality. Juvenile animals can be distinguished from adults on capture. In year  $t$ , juveniles are caught with probability  $p_{jt}$  and adults with probability  $p_{at}$ ,  $t = 1, \dots, T$ . (Age-dependent probabilities of capture commonly arise in practice, for example because young animals are inexperienced and more prone to capture.) Note that age will be known only for animals marked as juveniles. Again, such a model might be appropriate for some large mammal populations, and it can readily be extended for wider applicability.

The state vector  $\mathbf{n}_t$  contains the abundances of animals immediately following the mark–recapture experiment. Although conventional mark–recapture formulations tally capture histories of marked animals, both alive and dead, the state vector needs only to include animals alive at time  $t$ . A capture history of “110” indicates an animal that was caught in the previous 2 years, survives to the current year, but is not caught; animals that died in the preceding year are not included. The length of  $\mathbf{n}_t$  can equal  $2 \times 2^t$ , where  $2^t$  is the number of distinct capture histories at time  $t$ , and the doubling is for making and retaining the distinction between animals first captured as juveniles and as adults. If a distinction between males and females was made as well, the length would double again. However, a reduced length  $\mathbf{n}_t$  is shown here. The distinction between animals first captured as juveniles and those first captured as adults is retained for just 1 year following initial capture, sufficient to account for the differences in winter and spring survival between juveniles and adults. The subscripts  $j$ ,  $y$  and  $a$  are for juvenile, young and adult. Marked juveniles at time  $t$  that survive to autumn of the year  $t + 1$  are labelled

“young” (because we need to identify such animals on recapture, to estimate juvenile survival), while unmarked juveniles at time  $t$  that survive to  $t + 1$  are advanced to the adult category (because they are unobserved and therefore do not contribute to estimation of  $\phi_j$ ).

The state vectors for three time periods are as follows:

$$\mathbf{n}_1 = \begin{bmatrix} n_{j0,1} \\ n_{j1,1} \\ n_{a0,1} \\ n_{a1,1} \end{bmatrix} \rightarrow \mathbf{n}_2 = \begin{bmatrix} n_{j0,2} \\ n_{j1,2} \\ n_{y10,2} \\ n_{y11,2} \\ n_{a00,2} \\ n_{a01,2} \\ n_{a10,2} \\ n_{a11,2} \end{bmatrix} \rightarrow \mathbf{n}_3 = \begin{bmatrix} n_{j0,3} \\ n_{j1,3} \\ n_{y10,3} \\ n_{y11,3} \\ n_{a000,3} \\ n_{a001,3} \\ n_{a010,3} \\ n_{a011,3} \\ n_{a101,3} \\ n_{a110,3} \\ n_{a111,3} \end{bmatrix}$$

Thus the subscripts show capture history, preceded by a label indicating age class. Year is indicated by the final subscript. Capture histories vary in length because some animals were born after the study started, and only have entries corresponding to those years in which they were alive.

The observation vector  $\mathbf{y}_t$  is a subset of  $\mathbf{n}_t$  and contains abundances of animals captured or recaptured at time  $t$ . Examples of the observation vector are shown below for three time periods:

$$\mathbf{y}_1 = \begin{bmatrix} n_{j1,1} \\ n_{a1,1} \end{bmatrix} \rightarrow \mathbf{y}_2 = \begin{bmatrix} n_{j1,2} \\ n_{y11,2} \\ n_{a01,2} \\ n_{a11,2} \end{bmatrix} \rightarrow \mathbf{y}_3 = \begin{bmatrix} n_{j1,3} \\ n_{y11,3} \\ n_{a001,3} \\ n_{a011,3} \\ n_{a101,3} \\ n_{a111,3} \end{bmatrix}$$

Note the absence of capture histories ending in “0”; with this formulation, we do not observe such animals in year  $t$ , and have no need to include them, although the living ones do appear in the state vector.

Four sub-processes, three stochastic and one deterministic (advancement of juveniles to either the young or the adult stage and advancement of young to adult), generate  $\mathbf{n}_t$  from  $\mathbf{n}_{t-1}$ . The sequencing is winter and

spring survival (following mark and recapture), advancement, births, and mark and recapture. The initial state vector consists of unmarked juveniles and unmarked adults in the autumn, before the first marking. The sequencing is detailed for time period 2. Note that for  $t > 2$ , the “young” animals (age 1 marked animals) would be included in the survival sub-process and would be advanced to adults in the age incrementation step. First, the winter-spring survival,  $\mathbf{u}_{1,2} \sim H_{1,2}(\mathbf{n}_1)$ :

$$\begin{bmatrix} u_{1,j0,2} \sim \text{Binomial}(n_{j0,1}, \phi_j) \\ u_{1,j1,2} \sim \text{Binomial}(n_{j1,1}, \phi_j) \\ u_{1,a0,2} \sim \text{Binomial}(n_{a0,1}, \phi_a) \\ u_{1,a1,2} \sim \text{Binomial}(n_{a1,1}, \phi_a) \end{bmatrix}$$

Then age incrementation of juveniles,  $\mathbf{u}_{2,2} \sim H_{2,2}(\mathbf{u}_{1,2})$ :

$$\begin{bmatrix} u_{2,y1,2} = u_{1,j1,2} \\ u_{2,a0,2} = u_{1,j0,2} + u_{1,a0,2} \\ u_{2,a1,2} = u_{1,a1,2} \end{bmatrix}$$

Then births,  $\mathbf{u}_{3,2} \sim H_{3,2}(\mathbf{u}_{2,2})$ :

$$\begin{bmatrix} u_{3,j0,2} \sim \text{Binomial}(u_{2,y1,2} + u_{2,a0,2} + u_{2,a1,2}, \pi) \\ u_{3,y1,2} = u_{2,y1,2} \\ u_{3,a0,2} = u_{2,a0,2} \\ u_{3,a1,2} = u_{2,a1,2} \end{bmatrix}$$

Lastly, marking and recapturing occurs,  $\mathbf{n}_2 \sim H_{4,2}(\mathbf{u}_{3,2})$ :

$$\begin{bmatrix} n_{j0,2} \sim \text{Binomial}(u_{3,j0,2}, 1 - p_{j2}) \\ n_{j1,2} = u_{3,j0,2} - n_{j0,2} \\ n_{y10,2} \sim \text{Binomial}(u_{3,y1,2}, 1 - p_{a2}) \\ n_{y11,2} = u_{3,y1,2} - n_{y10,2} \\ n_{a00,2} \sim \text{Binomial}(u_{3,a0,2}, 1 - p_{a2}) \\ n_{a01,2} = u_{3,a0,2} - n_{a00,2} \\ n_{a10,2} \sim \text{Binomial}(u_{3,a1,2}, 1 - p_{a2}) \\ n_{a11,2} = u_{3,a1,2} - n_{a10,2} \end{bmatrix}$$

As in the previous example, the likelihood is the integral of the joint distribution of  $\mathbf{n}_t$ ,  $t = 1, \dots, T$ , constrained by the observation vectors.

The conditional expectation for  $\mathbf{n}_t$  can be expressed by products of matrices:

$$E_{\mathbf{n}_t | \mathbf{n}_{t-1}}[\mathbf{n}_t] = X_t B_t A_t S_t \mathbf{n}_{t-1}$$

The details for  $t = 2$  are shown as follows:

$$S_2 = \begin{bmatrix} \phi_j & 0 & 0 & 0 \\ 0 & \phi_j & 0 & 0 \\ 0 & 0 & \phi_a & 0 \\ 0 & 0 & 0 & \phi_a \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} \pi & \pi & \pi \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$X_2 = \begin{bmatrix} 1 - p_{j2} & 0 & 0 & 0 \\ p_{j2} & 0 & 0 & 0 \\ 0 & 1 - p_{a2} & 0 & 0 \\ 0 & p_{a2} & 0 & 0 \\ 0 & 0 & 1 - p_{a2} & 0 \\ 0 & 0 & p_{a2} & 0 \\ 0 & 0 & 0 & 1 - p_{a2} \\ 0 & 0 & 0 & p_{a2} \end{bmatrix}$$

### 3.4. Density-dependent, harvested population with environmental and age effects; line transect survey

Again for simplicity, we model just the female population. First-year survival is assumed to be related to accumulated day degrees frost (i.e. number of degrees below freezing, summed over days on which there was a frost) in year  $t$ ,  $d_t$ , as  $\phi_j(d_t) = \exp(\alpha + \beta d_t) / [1 + \exp(\alpha + \beta d_t)] = 1 / [1 + \exp(-\alpha - \beta d_t)]$ , where  $\beta < 0$ . (Here and elsewhere, we assume a logistic link function when modelling probabilities, although other link functions suitable for proportions may also be used.) Subsequently, survival is assumed to be the following function of age  $a$  until  $a = 3$ :  $\phi_a = 1 / [1 + \exp(-\gamma - \delta \log a)]$ , where  $\delta > 0$ . After age 3, the survival rate stays constant at the value for  $a = 3$ . Births per adult of age 3 or more in year  $t$  are assumed to be Bernoulli random variables with parameter  $\pi(t) = 1 / [1 + \exp(-\mu - \nu N_t)]$ , where  $N_t$  is the total number of animals in the population just before the breeding season in year  $t$ , and  $\nu < 0$ . A line transect survey just after the breeding season yields independent lognormally distributed estimates of the numbers of animals aged 1+ and young; the lognormal variance parameter is assumed constant for each category and is denoted  $\sigma^2$ . A harvest takes place immediately after the survey, and numbers of animals removed are recorded by age (with animals of age 3 or more grouped together). The vector of input parame-

ters in this case is

$$\Theta = (\alpha \ \beta \ \gamma \ \delta \ \mu \ \nu \ \sigma^2)'$$

The state vector consists of the abundances distinguished by age (0, 1, 2 and 3+) at the time the line transect survey is carried out:

$$\mathbf{n}_t = \begin{bmatrix} n_{0,t} \\ n_{1,t} \\ n_{2,t} \\ n_{3+,t} \end{bmatrix}$$

Four sub-processes, two deterministic and two stochastic, are used to describe the transition between  $\mathbf{n}_{t-1}$  and  $\mathbf{n}_t$ . Harvest occurs first and is viewed as deterministic, so that numbers harvested in age category  $i$  ( $c_{i,t}$ ) are simply subtracted from the states. Survival is next, and is followed by age advancement and birth. First, harvest,  $\mathbf{u}_{1,t} \sim H_{1,t}(\mathbf{n}_{t-1})$ :

$$\begin{bmatrix} u_{1,0,t} = n_{0,t-1} - c_{0,t} \\ u_{1,1,t} = n_{1,t-1} - c_{1,t} \\ u_{1,2,t} = n_{2,t-1} - c_{2,t} \\ u_{1,3+,t} = n_{3+,t-1} - c_{3+,t} \end{bmatrix}$$

Then survival,  $\mathbf{u}_{2,t} \sim H_{2,t}(\mathbf{u}_{1,t})$ :

$$\begin{bmatrix} u_{2,0,t} \sim \text{Binomial} \left( u_{1,0,t}, \frac{1}{1 + \exp(-\alpha - \beta d_t)} \right) \\ u_{2,1,t} \sim \text{Binomial} \left( u_{1,1,t}, \frac{1}{1 + \exp(-\gamma)} \right) \\ u_{2,2,t} \sim \text{Binomial} \left( u_{1,2,t}, \frac{1}{1 + \exp(-\gamma - \delta \log(2))} \right) \\ u_{2,3+,t} \sim \text{Binomial} \left( u_{1,3+,t}, \frac{1}{1 + \exp(-\gamma - \delta \log(3))} \right) \end{bmatrix}$$

Followed by age advancement,  $\mathbf{u}_{3,t} \sim H_{3,t}(\mathbf{u}_{2,t})$ :

$$\begin{bmatrix} u_{3,1,t} = u_{2,0,t} \\ u_{3,2,t} = u_{2,1,t} \\ u_{3,3+,t} = u_{2,2,t} + u_{2,3+,t} \end{bmatrix}$$

Lastly, births,  $\mathbf{n}_t \sim H_{4,t}(\mathbf{u}_{3,t})$ :

$$\begin{bmatrix} n_{0,t} \sim \text{Binomial} \left( u_{3,3+,t}, \frac{1}{1 + \exp(-\mu - \nu N_t)} \right) \\ n_{1,t} = u_{3,1,t} \\ n_{2,t} = u_{3,2,t} \\ n_{3+,t} = u_{3,3+,t} \end{bmatrix}$$

where  $N_t = u_{3,1,t} + u_{3,2,t} + u_{3,3+,t}$ .

The observation process vector estimates the number of age 1 and older animals and the number of age 0. The probability distribution for each component is assumed Lognormal( $\theta_i, \sigma^2$ ), where  $\theta_i$  corresponds to the log of state vector sub-totals; i.e.

$$y_{a,t} \sim \text{Lognormal}(\log(n_{1,t} + n_{2,t} + n_{3+,t}), \sigma^2) \\ y_{0,t} \sim \text{Lognormal}(\log(n_{0,t}), \sigma^2)$$

Note that we could explicitly include estimation of abundance from the line transect survey in the methods of this paper, by including parameters of the detection function in our parameter vector, and modifying the observation vector and corresponding likelihood appropriately. The approach illustrated here has the advantage of simplicity.

In this example, the state equation involves non-linear functions of states. Because the expectation of a function of a random variable is not in general the function of the expectation when the function is non-linear, we can only represent  $E_{\mathbf{n}_t|\mathbf{n}_{t-1}}[\mathbf{n}_t]$  approximately using matrices:

$$E_{\mathbf{n}_t|\mathbf{n}_{t-1}}[\mathbf{n}_t] \approx B_t A S_t (\mathbf{n}_{t-1} - \mathbf{c}_t)$$

where

$$S_t = \begin{bmatrix} \phi_j(d_t) & 0 & 0 & 0 \\ 0 & \phi_1 & 0 & 0 \\ 0 & 0 & \phi_2 & 0 \\ 0 & 0 & 0 & \phi_3 \end{bmatrix} \\ A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad B_t = \begin{bmatrix} 0 & 0 & \pi(N_t) \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The vector  $\mathbf{c}_t$  tallies the harvest by state component and  $\phi_j(d_t)$ ,  $\phi_1$ ,  $\phi_2$ ,  $\phi_3$ , and  $\pi(N_t)$  are the binomial parameters in the respective survival and birth sub-processes.

### 3.5. Metapopulation; adult female population census

Only females from three different populations are modelled. Animals have a constant survival rate  $\phi_j$  in their first year, and an annual survival rate  $\phi_a$  subsequently. Females first breed at age 1, and female births per adult female per year are distributed as a Bernoulli random variable with probability of success  $\pi$ .

After winter mortality but before breeding, movement of young females only is assumed to occur between colonies. There are three colonies, and colony  $i$  has size  $N_{it}$  breeding-age female animals and habitat suitability  $H_{it}$  in year  $t$ . ('Habitat suitability' might be measured in a number of ways. For example, it could be simply the amount of suitable habitat present in a colony, or a measure of the resources available to a colony, or a habitat quality measure.) The distance between colonies  $i$  and  $j$  is  $d_{ij}$ . For a given colony  $i$ , the proportions of young (first-year) animals staying or moving to another colony in year  $t$  are modelled according to a Dirichlet distribution with parameters:

$$\lambda_{i \rightarrow i}(t) = \exp(\alpha)$$

$$\lambda_{i \rightarrow j}(t) = \exp(\alpha + \beta N_{it} + \gamma H_{it} + \delta N_{jt} + \mu H_{jt} + \rho d_{ij}), \quad j \neq i$$

where  $\beta$  and  $\mu$  are positive and  $\gamma, \delta$  and  $\rho$  are negative. Observations are assumed to be independent lognormally distributed counts of females of breeding age at colonies made just after the birthing period; the lognormal variance parameter is assumed constant for all colonies and is denoted  $\sigma^2$ . Thus

$$\Theta = (\phi_j \quad \phi_a \quad \pi \quad \alpha \quad \beta \quad \gamma \quad \delta \quad \mu \quad \rho \quad \sigma^2)'$$

The state vector consists of the abundances of juvenile and adult females at three colonies just after the birthing period:

$$\mathbf{n}_t = \begin{bmatrix} n_{1,j,t} \\ n_{1,a,t} \\ n_{2,j,t} \\ n_{2,a,t} \\ n_{3,j,t} \\ n_{3,a,t} \end{bmatrix}$$

where the first subscript denotes the colony, the second age class, and the third year.

There are four sub-processes linking  $\mathbf{n}_{t-1}$  to  $\mathbf{n}_t$ . Winter mortality occurs first, then movement of juveniles, then age advancement, and finally birth. Beginning with winter mortality,  $\mathbf{u}_{1,t} \sim H_{1,t}(\mathbf{n}_{t-1})$ :

$$\begin{bmatrix} u_{1,1,j,t} \sim \text{Binomial}(n_{1,j,t}, \phi_j) \\ u_{1,1,a,t} \sim \text{Binomial}(n_{1,a,t}, \phi_a) \\ u_{1,2,j,t} \sim \text{Binomial}(n_{2,j,t}, \phi_j) \\ u_{1,2,a,t} \sim \text{Binomial}(n_{2,a,t}, \phi_a) \\ u_{1,3,j,t} \sim \text{Binomial}(n_{3,j,t}, \phi_j) \\ u_{1,3,a,t} \sim \text{Binomial}(n_{3,a,t}, \phi_a) \end{bmatrix}$$

Then movement of juveniles,  $\mathbf{u}_{2,t} \sim H_{2,t}(\mathbf{u}_{1,t})$ :

$$\begin{bmatrix} u_{2,1,j,t} = u_{2,1 \rightarrow 1,j,t} + u_{2,2 \rightarrow 1,j,t} + u_{2,3 \rightarrow 1,j,t} \\ u_{2,1,a,t} = u_{1,1,a,t} \\ u_{2,2,j,t} = u_{2,1 \rightarrow 2,j,t} + u_{2,2 \rightarrow 2,j,t} + u_{2,3 \rightarrow 2,j,t} \\ u_{2,2,a,t} = u_{1,2,a,t} \\ u_{2,3,j,t} = u_{2,1 \rightarrow 3,j,t} + u_{2,2 \rightarrow 3,j,t} + u_{2,3 \rightarrow 3,j,t} \\ u_{2,3,a,t} = u_{1,3,a,t} \end{bmatrix}$$

where

$$(u_{2,1 \rightarrow 1,j,t}, u_{2,1 \rightarrow 2,j,t}, u_{2,1 \rightarrow 3,j,t}) \sim u_{1,1,j,t} \times \text{Dirichlet}(\lambda_{1 \rightarrow 1}(t), \lambda_{1 \rightarrow 2}(t), \lambda_{1 \rightarrow 3}(t))$$

$$(u_{2,2 \rightarrow 1,j,t}, u_{2,2 \rightarrow 2,j,t}, u_{2,2 \rightarrow 3,j,t}) \sim u_{1,2,j,t} \times \text{Dirichlet}(\lambda_{2 \rightarrow 1}(t), \lambda_{2 \rightarrow 2}(t), \lambda_{2 \rightarrow 3}(t))$$

$$(u_{2,3 \rightarrow 1,j,t}, u_{2,3 \rightarrow 2,j,t}, u_{2,3 \rightarrow 3,j,t}) \sim u_{1,3,j,t} \times \text{Dirichlet}(\lambda_{3 \rightarrow 1}(t), \lambda_{3 \rightarrow 2}(t), \lambda_{3 \rightarrow 3}(t))$$

Thus the abundances of young at each colony after movement are convolutions. The abundance of breeding-age females,  $N_{it}$ , used in the calculation of  $\lambda_{i \rightarrow j}(t)$ , are  $u_{2,i,a,t}$ ,  $i = 1, 2, 3$ . Age advancement follows,  $\mathbf{u}_{3,t} \sim H_{3,t}(\mathbf{u}_{2,t})$ :

$$\begin{bmatrix} u_{3,1,a,t} = u_{2,1,j,t} + u_{2,1,a,t} \\ u_{3,2,a,t} = u_{2,2,j,t} + u_{2,2,a,t} \\ u_{3,3,a,t} = u_{2,3,j,t} + u_{2,3,a,t} \end{bmatrix}$$

Lastly, births occur,  $\mathbf{n}_t \sim H_{4,t}(\mathbf{u}_{3,t})$ :

$$\begin{bmatrix} n_{1,j,t} \sim \text{Binomial}(u_{3,1,a,t}, \pi) \\ n_{1,a,t} = u_{3,1,a,t} \\ n_{2,j,t} \sim \text{Binomial}(u_{3,2,a,t}, \pi) \\ n_{2,a,t} = u_{3,2,a,t} \\ n_{3,j,t} \sim \text{Binomial}(u_{3,3,a,t}, \pi) \\ n_{3,a,t} = u_{3,3,a,t} \end{bmatrix}$$

The components of the observation vector (counts of breeding females at each colony) are assumed independently lognormally distributed:

$$\begin{bmatrix} y_{1,t} \sim \text{Lognormal}(\log(n_{1,a,t}), \sigma^2) \\ y_{2,t} \sim \text{Lognormal}(\log(n_{2,a,t}), \sigma^2) \\ y_{3,t} \sim \text{Lognormal}(\log(n_{3,a,t}), \sigma^2) \end{bmatrix}$$

The conditional expectation of the state can be approximated by matrix products (the approximation

arising, as before, because the state equation includes non-linear functions of states):

$$E_{n_t|n_{t-1}}[n_t] \approx BAP_t S n_{t-1}$$

where

$$S = \begin{bmatrix} \phi_j & 0 & 0 & 0 & 0 & 0 \\ 0 & \phi_a & 0 & 0 & 0 & 0 \\ 0 & 0 & \phi_j & 0 & 0 & 0 \\ 0 & 0 & 0 & \phi_a & 0 & 0 \\ 0 & 0 & 0 & 0 & \phi_j & 0 \\ 0 & 0 & 0 & 0 & 0 & \phi_a \end{bmatrix}$$

$$P_t = \begin{bmatrix} p_{1 \rightarrow 1}(t) & 0 & p_{2 \rightarrow 1}(t) & 0 & p_{3 \rightarrow 1}(t) & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ p_{1 \rightarrow 2}(t) & 0 & p_{2 \rightarrow 2}(t) & 0 & p_{3 \rightarrow 2}(t) & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ p_{1 \rightarrow 3}(t) & 0 & p_{2 \rightarrow 3}(t) & 0 & p_{3 \rightarrow 3}(t) & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} \pi & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \pi \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$p_{i \rightarrow j}(t) = \frac{\lambda_{i \rightarrow j}(t)}{\sum_{k=1}^3 \lambda_{i \rightarrow k}(t)}$$

where  $p_{i \rightarrow j}(t)$  are the expected values for the Dirichlet vector.

### 3.6. Predator–prey system; strip transect survey

Again for simplicity, populations are assumed to be exclusively female. Births occur in summer. The number of young per prey animal is Poisson with rate  $\lambda(t) = \exp(\kappa + \omega Q_t)$ , where  $Q_t$  is the number of prey present at the start of the breeding season in year  $t$ , and  $\omega < 0$ . The number of young per predator is Poisson with rate  $\mu(t) = \exp(\eta + \zeta R_t)$ , where  $R_t$  is the number of predators present at the start of the breeding season in year  $t$ , and  $\zeta < 0$ . Survival of predator and prey are Bernoulli processes with different rates. The logit of the rate for prey is  $\alpha + \beta u_{1,q,t} + \gamma u_{1,r,t}$ , where

$\beta < 0$ ,  $\gamma < 0$  and  $u_{1,r,t}$  and  $u_{1,q,t}$  are the pre-survival process abundances of predator and prey at year  $t$ . The logit of the rate for predators is  $\delta + \nu u_{1,q,t} + \rho u_{1,r,t}$ , where  $\nu > 0$  and  $\rho < 0$ . Independent lognormally distributed estimates of the sizes of the prey and predator populations are made just before breeding, with variance parameters  $\sigma_q^2$  and  $\sigma_r^2$ . The estimates are based on strip transect surveys which cover an area  $a$  out of total area  $A$  for both predator and prey and all time periods. We now have

$$\Theta = (\alpha \ \beta \ \gamma \ \kappa \ \omega \ \delta \ \nu \ \rho \ \eta \ \zeta \ \sigma_q^2 \ \sigma_r^2)'$$

The state vector consists of the abundances of prey and predators after mortality occurs but prior to birthing:

$$n_t = \begin{bmatrix} Q_t \\ R_t \end{bmatrix}$$

There are only two sub-processes, birth, then survival. Beginning with birth,  $u_{1,t} \sim H_{1,t}(n_{t-1})$ :

$$\begin{bmatrix} u_{1,q,t} \sim Q_{t-1} + \text{Poisson}(\exp(\kappa + \omega Q_{t-1})) \\ u_{1,r,t} \sim R_{t-1} + \text{Poisson}(\exp(\eta + \zeta R_{t-1})) \end{bmatrix}$$

where the left-hand terms are the sum of the existing abundance and the randomly generated progeny. Then survival,  $n_t \sim H_{2,t}(u_{2,t})$ :

$$\begin{bmatrix} Q_t \sim \text{Binomial}(u_{1,q,t}, \phi_t) \\ R_t \sim \text{Binomial}(u_{1,r,t}, \psi_t) \end{bmatrix}$$

where

$$\phi_t = \frac{1}{1 + \exp(-\alpha - \beta u_{1,q,t} - \gamma u_{1,r,t})}$$

$$\psi_t = \frac{1}{1 + \exp(-\delta - \nu u_{1,q,t} - \rho u_{1,r,t})}$$

The observation vector comprises separate counts of predators and prey on sample strips covering a proportion  $a/A$  of the study area. Its distribution is given by

$$\begin{bmatrix} y_{q,t} \sim \text{Lognormal} \left( \log \left( \frac{aQ_t}{A} \right), \sigma_q^2 \right) \\ y_{r,t} \sim \text{Lognormal} \left( \log \left( \frac{aR_t}{A} \right), \sigma_r^2 \right) \end{bmatrix}$$

The approximate conditional expectation is simple:

$$E[\mathbf{n}_t | \mathbf{n}_{t-1}] \approx S_t B_t \mathbf{n}_{t-1}$$

where

$$B_t = \begin{bmatrix} 1 + \exp(\kappa + \omega Q_{t-1}) & 0 \\ 0 & 1 + \exp(\eta + \zeta R_{t-1}) \end{bmatrix}$$

$$S_t = \begin{bmatrix} \phi_t & 0 \\ 0 & \psi_t \end{bmatrix}$$

#### 4. Fitting the models

For many animal populations, the primary inference problem is to calculate estimates of the SSM’s unknown parameters  $\Theta$ , while inferences of the unknown states  $\mathbf{n}_t$  will often be of secondary importance. The focus in this section is on  $\Theta$  but inferences about  $\mathbf{n}_t$  will often be byproducts. The topic of inference for SSMs is an enormous subject (see, for example, Harvey, 1989; Doucet et al., 2001a). Here we only sketch some of the basic ideas of inference procedures for SSMs: recursive filtering, analytic and simulation-based solutions to the recursion, and the complication of sub-processes. In a companion paper (Newman et al., in preparation) we discuss inference procedures for the types of SSMs shown previously and expand upon some of the issues raised below (such as alternative proposal distributions for simulation-based methods).

##### 4.1. Recursive filtering algorithms

As Eqs. (4) and (6) indicate, the likelihood and the posterior pdf for  $\Theta$  can be a high-dimensional integral. Analytic solutions are rare; a notable exception is the Kalman filter for the special case of normal dynamic linear models (NDLMs; West and Harrison, 1990). For the general SSM, however, the more feasible solutions are simulation-based methods that yield estimates of the likelihood or simulated samples from the posterior. A common feature of most analytic and simulation-based approaches is their recursive nature.

One algorithm for carrying out the integration that yields the likelihood (or the posterior for  $\Theta$ ) is the recursive filtering algorithm. Output from the al-

gorithm includes the conditional distribution for  $\mathbf{n}_t$  given the observation vector up to and including time  $t$ ,  $\mathbf{y}^t$ . The algorithm uses the Markov characteristics of the SSM to change the problem of evaluating a high-dimensional integral over  $T$  time intervals into a problem of carrying out a sequence of at least  $T$ , lower dimensional, integrations. The result of integration at step  $t$  becomes input for integration at step  $t + 1$ . Additional details of the algorithm can be found in Section 3.7.3 of Harvey (1989).

The recursive filtering algorithm involves two steps: prediction and updating. Prediction yields the conditional distribution for  $\mathbf{n}_t$  given  $\mathbf{y}^{t-1}$ ,  $t > 1$ , which can be written as a function of the state process pdf and the updated conditional distribution for the state. Dependence on  $\Theta$  is implicit below:

$$g(\mathbf{n}_t | \mathbf{y}^{t-1}) = \int_{\mathbf{n}_{t-1}} g(\mathbf{n}_t, \mathbf{n}_{t-1} | \mathbf{y}^{t-1}) d\mathbf{n}_{t-1}$$

$$= \int_{\mathbf{n}_{t-1}} g_t(\mathbf{n}_t | \mathbf{n}_{t-1}) g(\mathbf{n}_{t-1} | \mathbf{y}^{t-1}) d\mathbf{n}_{t-1} \quad (9)$$

For time  $t = 1$ , the starting point is the initial state,  $\mathbf{n}_0$  and the predictive distribution for  $\mathbf{n}_1$  is  $g(\mathbf{n}_1) = \int_{\mathbf{n}_0} g_1(\mathbf{n}_1 | \mathbf{n}_0) g(\mathbf{n}_0) d\mathbf{n}_0$ .

Updating yields the conditional distribution for  $\mathbf{n}_t$  given  $\mathbf{y}^t$ , which is a function of the observation process pdf and the predicted conditional distribution. This is done via Bayes’ theorem:

$$g(\mathbf{n}_t | \mathbf{y}^t) = \frac{g(\mathbf{n}_t, \mathbf{y}^t)}{f(\mathbf{y}^t)} = \frac{g(\mathbf{n}_t, \mathbf{y}_t | \mathbf{y}^{t-1}) f(\mathbf{y}^{t-1})}{f(\mathbf{y}_t | \mathbf{y}^{t-1}) f(\mathbf{y}^{t-1})}$$

$$= \frac{f(\mathbf{y}_t | \mathbf{n}_t, \mathbf{y}^{t-1}) g(\mathbf{n}_t | \mathbf{y}^{t-1})}{f(\mathbf{y}_t | \mathbf{y}^{t-1})}$$

$$= \frac{f_i(\mathbf{y}_t | \mathbf{n}_t) g(\mathbf{n}_t | \mathbf{y}^{t-1})}{f(\mathbf{y}_t | \mathbf{y}^{t-1})} \quad (10)$$

where the denominator is the result of an integration,

$$f(\mathbf{y}_t | \mathbf{y}^{t-1}) = \int_{\mathbf{n}_t} f_i(\mathbf{y}_t | \mathbf{n}_t) g(\mathbf{n}_t | \mathbf{y}^{t-1}) d\mathbf{n}_t \quad (11)$$

The likelihood is a by-product of the filtering algorithm. The conditional distribution for the observation at time  $t$  given prior observations (Eq. (11)) is used to construct the unconditional distribution for  $\mathbf{y}^T$ :

$$f(\mathbf{y}^T) = \prod_{t=1}^T f(\mathbf{y}_t | \mathbf{y}^{t-1}) \quad (12)$$

Given the likelihood and a prior for  $\Theta$ , the posterior distribution for both  $\Theta$  and the states can be evaluated:

$$\begin{aligned} p(\Theta, \mathbf{n}^T | \mathbf{y}^T) &= \frac{p(\mathbf{y}^T, \mathbf{n}^T | \Theta) p(\Theta)}{f(\mathbf{y}^T)} \\ &= \frac{\prod_{t=1}^T [f_t(\mathbf{y}_t | \mathbf{n}_t, \Theta) g_t(\mathbf{n}_t | \mathbf{n}_{t-1})] g_0(\mathbf{n}_0) p(\Theta)}{f(\mathbf{y}^T)} \end{aligned}$$

4.2. Solutions to the integrations

The recursive filtering algorithm breaks down the high-dimensional integration over  $T$  intervals into  $T$  separate integrations, but the integration over a single time interval can still be high-dimensional and quite difficult.

In the special case of a NDLM, the Kalman filter provides an analytic solution to each of the single time period integrals. Due to the special structure of the normal linear case, the prediction and update pdfs are normal and the Kalman algorithm is a recursive procedure for calculating the corresponding mean vectors and covariance matrices; thus explicit integration is not required. Because of the Kalman algorithm’s relative simplicity, approximating a non-linear and/or non-normal SSM for animal population dynamics with an NDLM has been attractive (Sullivan, 1992; Newman, 1998; Besbeas et al., 2002). The matrix representation of the conditional expectations of  $\mathbf{n}_t$  provide one component of the approximation, namely the mean vector for the state process. The same can be done relatively easily for the observation process mean. The covariance matrices for the state and observation processes will be somewhat more involved. Newman et al. (in preparation) examine the accuracy of such approximations for particular SSMs.

For cases where an NDLM approximation may not be sufficiently accurate, simulation-based methods can be used to carry out the integrations of the recursive filtering algorithm. There are many such simulation methods; the volume edited by Doucet et al. (2001a) describes several. We describe only one method, not necessarily the most computationally efficient, but one of the simplest to explain. The method is a variation on the description of the bootstrap filter given by Doucet et al. (2001b). We deviate from their description in that the focuss is on estimating the likelihood for a

given value of  $\Theta$ , and we include estimation of the prediction and update pdfs evaluated at particular  $\mathbf{n}_t$ , along with generation of samples from the smoothed pdf. In many situations, evaluation of the pdfs may not be that much more informative than samples generated from the pdfs.

1. *Initialisation*,  $t = 0$ . A sample of size  $M$  is simulated from  $g_0(\mathbf{n}_0; \Theta)$  with results denoted  $\mathbf{n}_0^{**i}$ ,  $i = 1, \dots, M$ . Increment to  $t = 1$  and proceed to step 2.
2. *Prediction*:
  - (a) *pdf evaluation*. The prediction pdf is estimated as follows:

$$\hat{g}(\mathbf{n}_t | \mathbf{y}^{t-1}) = \frac{1}{M} \sum_{i=1}^M g_t(\mathbf{n}_t | \mathbf{n}_{t-1}^{**i}) \tag{13}$$

- (b) *prediction sample*. A sample of size  $M$  is simulated from  $g_t(\mathbf{n}_t | \mathbf{n}_{t-1}^{**i})$  and the results are denoted  $\mathbf{n}_{t|t-1}^{*i}$ .
3. *Update*:
  - (a) *pdf evaluation*. The update pdf is estimated as follows:

$$\hat{g}(\mathbf{n}_t | \mathbf{y}^t) = \frac{f_t(\mathbf{y}_t | \mathbf{n}_t) \hat{g}(\mathbf{n}_t | \mathbf{y}^{t-1})}{\hat{f}(\mathbf{y}_t | \mathbf{y}^{t-1})} \tag{14}$$

where

$$\hat{f}(\mathbf{y}_t | \mathbf{y}^{t-1}) = \frac{1}{M} \sum_{i=1}^M f_t(\mathbf{y}_t | \mathbf{n}_{t|t-1}^{*i}) \tag{15}$$

- (b) *update sample*. A sample of size  $M$  is simulated by resampling  $\mathbf{n}_{t|t-1}^{*i}$  with weights  $w_i$  defined as follows:
- $$w_i = \frac{f_t(\mathbf{y}_t | \mathbf{n}_{t|t-1}^{*i})}{M \hat{f}(\mathbf{y}_t | \mathbf{y}^{t-1})} \tag{16}$$

The results are denoted  $\mathbf{n}_t^{**i}$ .

4. *Increment  $t$  to  $t + 1$* . If  $t \leq T$ , return to step 2.

The estimates of the prediction pdf (Eq. (13)) and the update pdf (Eq. (14)) are consistent by the law of large numbers.

The likelihood is estimated using Eq. (12) with the estimates of each term coming from Eq. (15). It is computationally inefficient to generate new  $\mathbf{n}^{**}$  for each value of  $\Theta$  in order to find the maximum likelihood estimate. Instead, as Geyer (1996) shows, a

single sample of  $\mathbf{n}^{**}$  can be generated based on a particular value of  $\Theta$ , say  $\Theta = \Psi$ , and this sample can be reused for other  $\Theta$  by adjusting the objective function and using importance weights. For example, if the original objective function was  $\log(L(\Theta))$ , use  $\log(L(\Theta)/L(\Psi))$  instead and use importance weights  $g(\mathbf{n}^{**}|\Theta)/g(\mathbf{n}^{**}|\Psi)$ .

An alternative to maximum likelihood estimation is described by Liu and West (2001). Their method yields a sample from the posterior distribution for  $\Theta$  given  $\mathbf{y}^T$  as well as samples from the smoothed state pdf. The procedure is an extension of the bootstrap filter in which they also simulate  $\Theta$  from a prior distribution for  $\Theta$ , resample simulated  $\Theta$ 's sequentially to yield samples from the posterior at time  $t$ , and then kernel smooth to compensate for the ‘‘particle depletion’’, i.e. the reduction in the number of unique parameter values, that results from resampling the parameters. Trenkel et al. (2000) use a similar procedure.

#### 4.3. Sub-processes

Sub-processes complicate the evaluation of the state pdf. In the case of an NDLM approximation to the SSM, the complication is lessened in that the conditional expectation and covariance of the state and observation processes can be written, at least approximately, as a result of matrix operations. In the more general case, however, where simulation methods are required, simulation from the state process and all intermediate sub-processes avoids evaluation of the state pdf. ‘‘Proposal’’ distributions other than the state process pdfs can be used, but this leads to calculation of importance sampling weights that do require evaluating the state process.

### 5. Discussion

The approach of this paper gives a general framework for integrating the modelling of the dynamics of open populations with the analysis of survey data on those populations. The models are calibrated using the survey data, and the parameter sets obtained at the final step can be used to project into the future under a range of management strategies.

When multiple sub-processes exist, it may be advantageous to sample some of the sub-processes,

which leads to an expansion of the observation process. There may be two or more independent surveys. The components from different sources may simply be multiplied together when independence can be assumed, so that this is readily incorporated. Especially when a sub-process is assumed to have the same parameter values throughout the system and through time, observations on that sub-process may alternatively be included through the use of informative priors.

The approach outlined relies on discretising the system. Thus animals are classified into groups, and time and space are divided into discrete units. The division of time into discrete units is not a serious limitation. For example, if the year is the basic time unit, we can break it into smaller units, as was done in Population 1. However, if breeding for example occurs through much of the year, there is a limit to how well discrete models can approximate this. Whether this matters depends on the purpose of the modelling.

We can move away from discrete-space movement models while still retaining the same general approach by including locational covariates in the model. Movement of an animal could then be handled by specifying a model for its location at the next time point, given its location (and other covariate values) at the current time point.

In principle, there is no reason why the state equation and the observation equation cannot be expanded to have one row per animal in the population. This would allow individual-based models, with covariates recorded by individual. The use of stochastic simulation for model fitting means that we do not need to manipulate the matrices used to define the models, so that it becomes feasible to allow these matrices to become very large. Because individual covariates would typically not be known for most animals, we would need to model the distribution of covariates in the population, to generate values by stochastic simulation.

When animals are removed from the population as a result of harvest, and harvest is in turn modelled, perhaps as a function of some measure of harvesting effort, then the number harvested should be modelled stochastically, in contrast with the deterministic treatment given in the exploited population example (Section 3.4). The situation is analogous to the re-

moval experiment example (Section 3.2) if numbers harvested are known without error, in that harvested animals are just a type of removal. If harvest is estimated, however, then the observation process vector includes estimated harvest, and it will include a stochastic element.

In many situations, the parameters of the state process evolve in time according to some stochastic process. The state-space formulation can be expanded to include a third process that describes the temporal variation of parameters. In other words, a hierarchical state-space model (Newman, 2000) can be formulated:

$h(\Theta_t)$  Parameter process

$g_0(\mathbf{n}_0; \Theta_0)$  Initial state distribution

$g_t(\mathbf{n}_t | \mathbf{n}_{t-1}; \Theta_t)$  State process distribution

$f_t(\mathbf{y}_t | \mathbf{n}_t; \Theta_t)$  Observation process distribution

Models for the parameter process include random effect models where independence exists between any two points in time and random walk models where  $\Theta_t$  is a random perturbation of  $\Theta_{t-1}$ . More scientifically interesting models are ones where explanatory covariates are used to model the parameters of the state process model, i.e. the hyperparameters are functions of covariates.

We have not addressed in detail the practical difficulties in carrying out inference about the states and unknown parameters. Thomas et al. (in press) provide more details for a specific population dynamics problem, the British grey seal metapopulation. They use a variation of Liu and West's (2001) approach, which combines auxiliary particle filtering (Pitt and Shephard, 1999) with kernel smoothing of the parameters. Newman et al. (in preparation) address: (a) the estimation of parameters or generation of parameter sets, in particular contrasting sequential particle filtering with Markov chain Monte Carlo; (b) the use of the time series trajectory of filtered estimates of the parameters (when using sequential particle filtering) as a model diagnostic and measure of model mis-specification, and of particle depletion as a diagnostic for model mis-specification; and (c) quantifying model uncertainty with AIC- or BIC-type measures. Much remains to be done on the topic of efficient implementation of these methodologies.

## Acknowledgements

We thank Carmen Fernández, David Elston and Steve Brooks for helpful discussions of the ideas presented. We also thank a referee for detailed and helpful comments.

## References

- Besbeas, P., Freeman, S.N., Morgan, B.J.T., Catchpole, E.A., 2002. Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. *Biometrics* 58, 540–547.
- Borchers, D.L., Buckland, S.T., Zucchini, W., 2002. *Estimating Animal Abundance: Closed Populations*. Springer-Verlag, London.
- Caswell, H., 2001. *Matrix Population Models*. Sinauer Associates, Sunderland, MA.
- Doucet, A., de Freitas, N., Gordon, N., 2001a. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York.
- Doucet, A., de Freitas, N., Gordon, N., 2001b. An introduction to sequential Monte Carlo methods. In: Doucet, A., de Freitas, N., Gordon, N. (Eds.), *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, pp. 3–14.
- Fewster, R.M., Buckland, S.T., Siriwardena, G.M., Baillie, S.R., Wilson, J.D., 2000. Analysis of population trends for farmland birds using generalized additive models. *Ecology* 81, 1970–1984.
- Geyer, C., 1996. Estimation and optimization of functions. In: Gilks, W.R., Richardson, S., Spiegelhalter, D.J. (Eds.), *Markov Chain Monte Carlo in Practice*. Chapman and Hall, New York, pp. 241–258.
- Griebeler, E.M., Seitz, A., 2002. An individual based model for the conservation of the endangered Large Blue Butterfly, *Maculinea arion* (Lepidoptera: Lycaenidae). *Ecol. Model.* 156, 43–60.
- Harvey, A.C., 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge, UK.
- Kriticos, D.J., Brown, J.R., Maywald, G.F., Radford, I.D., Nicholas, D.M., Sutherst, R.W., Adkins, S.W., 2003. SPAnDX: a process-based population dynamics model to explore management and climate change impacts on an invasive alien plant, *Acacia nilotica*. *Ecol. Model.* 163, 187–208.
- Leslie, P.H., 1945. On the use of matrices in certain population mathematics. *Biometrika* 33, 183–212.
- Leslie, P.H., 1948. Some further notes on the use of matrices in population mathematics. *Biometrika* 35, 213–245.
- Liu, J., West, M., 2001. Combining parameter and state estimation in simulation-based filtering. In: Doucet, A., de Freitas, N., Gordon, N. (Eds.), *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, pp. 197–223.
- Meyer, R., Millar, R.B., 1999. Bayesian stock assessment using a state-space implementation of the delay difference model. *Can. J. Fish. Aquat. Sci.* 56, 37–52.

- Millar, R.B., Meyer, R., 2000. Non-linear state space modelling of fisheries biomass dynamics by using Metropolis–Hastings within-Gibbs sampling. *Appl. Stat.* 49, 327–342.
- Miller, D.H., Jensen, A.L., Hammill, J.H., 2002. Density dependent matrix model for grey wolf population projection. *Ecol. Model.* 151, 271–278.
- Newman, K.B., 1998. State-space modeling of animal movement and mortality with application to salmon. *Biometrics* 54, 1290–1314.
- Newman, K.B., 2000. Hierarchic modeling of salmon harvest and migration. *J. Agric. Biol. Environ. Stat.* 5, 430–455.
- Newman, K.B., Fernández, C., Buckland, S.T., in preparation. Alternative procedures for parameter estimation and filtering state-space models for animal population dynamics.
- Parysow, P., Tazik, D.J., 2002. Assessing the effect of estimation error on population viability analysis: an example using the black-capped vireo. *Ecol. Model.* 155, 217–229.
- Peterson, M.N., Grant, W.E., Lopez, R.R., Silvy, N.J., 2003. Implications of fetal sex ratio hypotheses in endangered populations: simulated dynamics of Florida Key deer, Florida, USA. *Ecol. Model.* 165, 209–220.
- Pitt, M.K., Shephard, N., 1999. Filtering via simulation: auxiliary particle filters. *J. Am. Stat. Assoc.* 94, 590–599.
- Raftery, A.E., Givens, G.H., Zeh, J.E., 1995. Inference from a deterministic population dynamics model for bowhead whales (with discussion). *J. Am. Stat. Assoc.* 90, 427–430.
- Seber, G.A.F., 1982. *The Estimation of Animal Abundance and Related Parameters*. Griffith, London.
- Sullivan, P., 1992. A Kalman filter approach to catch-at-length analysis. *Biometrics* 48, 237–257.
- Thomas, L., Buckland, S.T., Newman, K.B., Harwood, J., in press. A unified framework for modelling wildlife population dynamics. *Aust. New Zealand J. Stat.*
- Trenkel, V.M., Elston, D.A., Buckland, S.T., 2000. Fitting population dynamics models to count and cull data using sequential importance sampling. *J. Am. Stat. Assoc.* 95, 363–374.
- West, M., Harrison, J., 1990. *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York.
- Yearsley, J.M., Fletcher, D., Hunter, C., 2003. Sensitivity analysis of equilibrium population size in a density-dependent model for short-tailed shearwaters. *Ecol. Model.* 163, 119–129.