A UNIFIED FRAMEWORK FOR MODELLING WILDLIFE POPULATION DYNAMICS[†]

LEN THOMAS[1*], STEPHEN T. BUCKLAND[1],
KEN B. NEWMAN[2] AND JOHN HARWOOD[3]

University of St Andrews and University of Idaho

Summary

This paper proposes a unified framework for defining and fitting stochastic, discrete-time, discrete-stage population dynamics models. The biological system is described by a state–space model, where the true but unknown state of the population is modelled by a state process, and this is linked to survey data by an observation process. All sources of uncertainty in the inputs, including uncertainty about model specification, are readily incorporated. The paper shows how the state process can be represented as a generalization of the standard Leslie or Lefkovitch matrix. By dividing the state process into subprocesses, complex models can be constructed from manageable building blocks. The paper illustrates the approach with a model of the British Grey Seal metapopulation, using sequential importance sampling with kernel smoothing to fit the model.

Key words: auxiliary particle filter; ecology; Grey Seals; *Halichoerus grypus*; metapopulation; nonlinear stochastic matrix models; sequential importance sampling; state–space models; wildlife conservation and management.

1. Introduction

Wildlife populations are highly structured stochastic systems, about which there is usually incomplete information. Nevertheless it is often important to make inferences about the system, for example when the population is of economic or conservation concern and a management or policy decision must be made. Constructing and fitting mathematical models of the population can greatly aid the decision-making process, first by allowing hypotheses about the populations to be tested against the observed data, and second by predicting the consequences of the possible decisions, and uncertainty about these consequences.

Traditional approaches to wildlife population modelling involve constructing a model based on assumptions about the system and then fixing the model parameters based on available information. Model outputs are then determined analytically or via computer simulation.

Analytic methods can be applied only to relatively simple (often deterministic) models, but have the advantage that they can produce results that help understanding of the general behaviour of the model (e.g. whether the population cycles or is chaotic). Simulation can be applied to more realistic models, but it produces only specific results (e.g. predicted trends in abundance or time to extinction). The two approaches are therefore often seen as complementary.

A major limitation of these approaches is that there is no formal way of accounting for uncertainty in the inputs, or representing uncertainty about the conclusions. Uncertainty arises from two sources. First, the input parameter values are rarely known absolutely — instead, they come from separate analyses of survey data when these are available, or expert opinion when they are not. Second, the model structure is rarely completely based on proven mechanism — rather, it represents an attempt to construct a parsimonious, robust and tractable characterization of the system under study. In the traditional approaches, parameter uncertainty is addressed through ad hoc 'sensitivity analysis', which quantifies the influence on outputs of varying the input values. Model uncertainty, when considered at all, is treated in a similar way (by trying different models and observing the changes in results). However, failure to adequately account for uncertainty can have drastic consequences (e.g. Ludwig, Hilborn & Walters, 1993; Wood & Thomas, 1999).

Here, we propose a framework for embedding stochastic population dynamics models into statistical inference. Determining model outputs then becomes a statistical prediction problem, where uncertainties in inputs translate into confidence limits (or other measures) on outputs. All sources of uncertainty, including model uncertainty (Buckland, Burnham & Augustin, 1997; Hoeting *et al.*, 1999), can be readily incorporated.

Our approach extends the population projection matrix models described by Caswell (2001, 2005), in three respects. First, to a model for population dynamics we add an explicit observation model which connects measurements on the population to the underlying state of the population. We then define a joint likelihood which is used to fit both models simultaneously. Second, we specify the population dynamics model using a separate model for each population process (e.g. birth, death, migration, age incrementation, etc.). These separate subprocess models can be chained together to produce the overall population dynamics model. This separation into subprocesses provides a simple framework for building complex models. Third, the subprocess models can be represented in matrix form, as expectations with attached stochastic errors, but they can also be represented as probability density functions (pdfs). The matrix form is more familiar to many biological scientists and this representation can be used to facilitate model building. However, in some cases it is only approximate, and the pdf form is always used in fitting the models.

The proposed framework is made possible by recent advances in computer-intensive statistical inference (e.g. Doucet, de Freitas & Gordon, 2001; Liu, 2001), and builds on previous developments in this direction in the fields of fisheries management (e.g. Mendelssohn, 1988; Schnute, 1994; Punt & Hilborn, 1997; Newman, 1998; Meyer & Millar, 1999), marine mammals (Raftery, Givens & Zeh, 1995), deer (Trenkel, Elston & Buckland, 2000) and birds (Besbeas *et al.*, 2002; Besbeas, Lebreton & Morgan, 2003; Besbeas, Freeman & Morgan, 2005).

In Section 2, we describe the modelling framework, and outline a method of fitting the models. We then give an example: that of the British Grey Seal metapopulation. Lastly, we discuss some future areas of research. Buckland *et al.* (2004) give more mathematical details, and outlines of example applications.

2. Modelling framework

We describe the system using a state–space model (Harvey, 1989; Durbin & Koopman, 2001) which divides the system into two linked processes: a state process and an observation process. The state process describes the true, but unknown, state of the population at successive time steps, while the observation process links the unknown states to data on the population recorded during surveys or experiments. The advantage of this framework is that it allows modelling of biologically plausible population processes and estimation of the key biological parameters, while at the same time explicitly recognizing the uncertainties involved in the data collection.

2.1. State process

In its simplest form, the state process is simply a deterministic population projection matrix, such as a Leslie matrix (Leslie, 1945, 1948; Caswell, 2001),

$$\boldsymbol{n}_t = \boldsymbol{L}\boldsymbol{n}_{t-1} \,,$$

where $\boldsymbol{n}_t$ is a vector representing numbers of animals by category in year t, and $\boldsymbol{L}$ is the Leslie matrix that updates $\boldsymbol{n}_{t-1}$ to $\boldsymbol{n}_t$. The standard Leslie matrix handles just the population processes of birth (through birth rates on the first row) and death (through survival probabilities on the sub-diagonal) in addition to incrementing ages, although it can readily be extended to incorporate other processes such as movement (Caswell, 2001). In simple models, elements of $\boldsymbol{n}$ correspond to numbers of animals by age class and perhaps sex, but in more complex models n can include numbers in different populations of a metapopulation, different species, different genotypes, etc. If animals are categorized by developmental stage, rather than age, the matrix is usually referred to as a Lefkovitch matrix (Lefkovitch, 1965; Caswell, 2001).

To extend this formulation to allow a stochastic state process we write

$$\mathrm{E}(\boldsymbol{n}_t \mid \boldsymbol{n}_{t-1}) = \boldsymbol{L}\boldsymbol{n}_{t-1} \,,$$

where $\mathrm{E}(\boldsymbol{n}_t \mid \boldsymbol{n}_{t-1})$ is the expected value of $\boldsymbol{n}_t$, given $\boldsymbol{n}_{t-1}$, and $\boldsymbol{L}$ now describes the long-term average effect of a set of stochastic processes.

For complex models, it is often helpful to split the state process into a series of subprocesses, each of which is defined as a separate matrix. The process model can then be written as the matrix product of the subprocess matrices. By selecting appropriate subprocess matrices and their ordering, a wide range of biological processes can be modelled. Partitioning the model in this way also allows attention to be focussed on one particular subprocess in isolation, allowing alternative theories for the subprocess structure to be more readily generated and tested.

As an example, consider a stage-structured metapopulation in which the principal subprocesses are breeding in the spring, summer survival, migration among populations in the autumn, overwinter survival and finally possible incrementation to the next stage. These can be represented by the matrices $\boldsymbol{B}$, $\boldsymbol{S}$, $\boldsymbol{M}$, $\boldsymbol{W}$ and $\boldsymbol{A}$ respectively. Assume that there is an early winter harvest each year of a known number of individuals, $\boldsymbol{c}_t$ (assumed to occur between the migration and winter survival subprocesses). The state process can be summarized as

$$\mathrm{E}(\boldsymbol{n}_t \mid \boldsymbol{n}_{t-1}) = \boldsymbol{A}\boldsymbol{W}(\boldsymbol{M}\boldsymbol{S}\boldsymbol{B}\boldsymbol{n}_{t-1} - \boldsymbol{c}_t) \,.$$

In this matrix formulation, the expectation $\mathrm{E}(\boldsymbol{n}_t \mid \boldsymbol{n}_{t-1})$ is exact only when the expected states of the population after each subprocess are a linear combination of the states after the previous subprocess — e.g. $\boldsymbol{B}\boldsymbol{n}_{t-1}$ should have elements each of which is a linear combination of the elements of $\boldsymbol{n}_{t-1}$. Otherwise, the expectation is an approximation. This does not detract from its use as a tool to aid model definition and explanation.

The exact representation of the state process is given by defining each stochastic subprocess as a pdf, with a known distribution. The evolution from $\boldsymbol{n}_{t-1}$ to $\boldsymbol{n}_t$ is then described by a series of linked pdfs, where the input to one pdf is the output from the previous one. From the above example, the full state process distribution, $\boldsymbol{n}_t \stackrel{\mathrm{d}}{=} H_t(\boldsymbol{n}_{t-1})$, comprises the following subprocesses:

$$\boldsymbol{u}_{bt} \stackrel{\mathrm{d}}{=} H_{bt}(\boldsymbol{n}_{t-1}), \qquad \boldsymbol{u}_{st} \stackrel{\mathrm{d}}{=} H_{st}(\boldsymbol{u}_{bt}), \qquad \boldsymbol{u}_{mt} \stackrel{\mathrm{d}}{=} H_{mt}(\boldsymbol{u}_{st})$$
$$\boldsymbol{u}_{wt} \stackrel{\mathrm{d}}{=} H_{wt}(\boldsymbol{u}_{mt} - \boldsymbol{c}_t) \qquad \boldsymbol{n}_t \stackrel{\mathrm{d}}{=} H_{at}(\boldsymbol{u}_{wt}),$$

where $\boldsymbol{u}_{xt}$ is a realization of the state vector after subprocess x at time t and $H_{xt}(\boldsymbol{u}_{yt})$ is the distribution defining subprocess x, given the input state $\boldsymbol{u}_{yt}$.

2.2. Observation process

In general, we have some observational data on the population of interest which can be used to make inferences about the state of the population. These data may come from a variety of sources, and are not necessarily collected on a regular basis. The observation process is a stochastic function that maps the data onto the underlying, unknown states. In matrix form,

$$\mathrm{E}(\boldsymbol{y}_{xt} \mid \boldsymbol{u}_{xt}) = \boldsymbol{O}_x \boldsymbol{u}_{xt},$$

where $\boldsymbol{y}_{xt}$ represents data collected after subprocess x has occurred and $\boldsymbol{O}_x$ is the observation matrix. Generally, when there is only one measurement per time period, the state process is set up so that this measurement coincides with the time period t. The observation process usually is stochastic, but can be deterministic. For example, if a complete census is taken at time t, without attempting to distinguish components of the population, then the observation process simply sums all the components of $\boldsymbol{n}_t$.

2.3. Fitting the models

The complete state–space model can be written as a set of three pdfs:

$$g_0(\boldsymbol{n}_0 \mid \Theta) \qquad \text{initial state distribution}$$
$$g_t(\boldsymbol{n}_t \mid \boldsymbol{n}_{t-1}, \Theta) \qquad \text{state process distribution}$$
$$f_t(\boldsymbol{y}_t \mid \boldsymbol{n}_t, \Theta) \qquad \text{observation process distribution},$$

where Θ, the vector of model parameters, was not written explicitly in the previous sections, and is often implicit in what follows. The state process distribution is first-order Markov: the distribution of $\boldsymbol{n}_t$ is dependent only on $\boldsymbol{n}_{t-1}$ (and the model parameters), and not on the state in any previous time periods. The formulation readily extends to higher orders.

We wish to make inferences about $\boldsymbol{n}_t$ or Θ or both, given the data. For example, for in-season management of a harvested population, we may wish to make statements about the current state of the population given all of the data up to the current time period, $g(\boldsymbol{n}_t \mid \boldsymbol{y}^t)$, where $\boldsymbol{y}^t = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t)$. For retrospective assessment, we may wish to make statements

about the past state of the population given all of the data up to the current time period, $g(\boldsymbol{n}_s \mid \boldsymbol{y}^t)$ where $s < t$. Lastly, we may want to evaluate future management scenarios by predicting the future state of the population, $g(\boldsymbol{n}_s \mid \boldsymbol{y}^t)$ where $s > t$. These three pdfs are referred to as filtered, smoothed and predicted distributions, respectively.

Evaluating these pdfs requires integration — in particular, integrating over all or some portion of the state process distribution (Pitt & Shephard, 1999). In some special cases, analytic solutions are available. Specifically, if both the state and observation processes are linear in the parameters and have normal densities then the Kalman filter algorithm can be used (Harvey, 1989). This approach has been applied with success to several ornithological time series by Besbeas *et al.* (2002, 2003). However, many models of interest in wildlife and conservation are neither linear nor normal. In these cases, Monte Carlo inference procedures can be used. These procedures can be divided into two general classes: Markov chain Monte Carlo (MCMC; Gilks, Richardson & Spiegelhalter, 1996) and sequential importance sampling (SIS; Doucet *et al.*, 2001), although there is considerable overlap between the classes (Liu, 2001). Here we focus on SIS and briefly describe the basic algorithm, followed by some refinements used in the example.

2.4. Sequential importance sampling

The basic SIS algorithm is the weighted bootstrap of Smith & Gelfand (1992). We begin by defining a joint prior distribution on the parameters and initial states. We simulate a large number, J, of initial parameter and state vectors from the prior. Each pair of state and parameter vectors $(\boldsymbol{n}_0^{[j]}, \Theta^{[j]})$ is called a 'particle'. We then stochastically project each particle forward to the first time period using the state process distribution, so that $\boldsymbol{n}_1^{[j]} \overset{\mathrm{d}}{=} H_1(\boldsymbol{n}_0^{[j]}, \Theta^{[j]})$. The particle density then gives an estimate of the predicted state distribution $g(\boldsymbol{n}_1 \mid \boldsymbol{y}^0)$. To estimate the filtered state distribution $g(\boldsymbol{n}_1 \mid \boldsymbol{y}^1)$, we use the observation process distribution to calculate a likelihood weight

$$w_1^{[j]} = \frac{f_1(\boldsymbol{y}_1 \mid \boldsymbol{n}_1^{[j]}, \Theta^{[j]})}{\sum_{i=1}^{J} f_1(\boldsymbol{y}_1 \mid \boldsymbol{n}_1^{[i]}, \Theta^{[i]})} \,, \tag{1}$$

and then take a weighted resample from the particles. This process is then repeated for subsequent time periods, first using the state process distribution to project forward to the next time period and then 'correcting' the resulting predicted state distribution using the weighted resample, with weights calculated according to the observation process. This yields estimates of the filtered state and parameter density at each time point. The parameter density at the final time point is an estimate of posterior parameter density given all the data.

In practice, the basic algorithm is unreliable, because of a phenomenon known as 'particle depletion'. In the weighted bootstrap step, particles with relatively large weights tend to be chosen many times, while those with small weights tend not to be chosen at all. Therefore, over time, a few of the initial particles come to dominate the particle set, resulting in a very inaccurate representation of the posterior state and parameter densities. Particle depletion becomes worse as the time series of observations gets longer, and within each time step the problem is worse when the variance of the bootstrap weights is high.

Many strategies are available to help circumvent this particle depletion problem (see papers in Doucet *et al.*, 2001; Liu, 2001). We used three strategies in the example presented in this paper, and we briefly outline them below.

First, we used kernel smoothing of the parameter vectors at each time step, as proposed by West (1993) and Trenkel *et al.* (2000). This adds a small perturbation to the parameter values of each particle selected at the resampling stage. The effect is to increase the diversity of parameter values in the vicinity of parameter space supported by the data. The degree of perturbation is controlled by a smoothing parameter, a, where $a = 0$ corresponds to no smoothing (i.e. the non-parametric weighted bootstrap of Smith & Gelfand, 1992) and $a = 1$ corresponds to a parametric weighted bootstrap, where the parametric form is determined by the choice of kernel. We found that moderate smoothing of $a = 0.7$ produced consistently reliable results using a multivariate normal kernel with shrinkage to preserve the mean and variance–covariance structure of the parameter estimates.

Second, we used the auxiliary particle filter of Pitt & Shephard (1999). Here, an initial 'auxiliary' resample is taken from the population at time t, with weights calculated according to the *expected* likelihood of the states at time $t + 1$, given the data at $t + 1$. This resampled set of particles is then projected forward from time t to time $t + 1$, and 'corrected' using likelihood weights just as with the bootstrap filter, except that the likelihood weights must take account of the auxiliary resampling stage. The effect of the auxiliary resampling is to boost the number of particles that are expected to be projected into the vicinity of parameter and state space supported by the data. This reduces the variance of the likelihood weights, and reduces the effect of particle depletion. Auxiliary particle filtering combined with kernel smoothing of parameters was first suggested by Liu & West (2001).

Third, we used residual sampling (Liu & Chen, 1998) rather than simple random sampling at the weighted resampling stages. Residual sampling gives the same expected distribution of particles as simple random sampling, but has smaller Monte Carlo variance.

3. Example: British Grey Seals

The British population of the Grey Seal (*Halichoerus grypus*) provides an interesting case study for scientific management. The species breeds colonially, mostly on offshore islands to the north and west of Scotland. Grey Seals spend over 80% of their time at sea (McConnell *et al.*, 1999), and 90% of this time underwater (Thompson *et al.*, 1991) — as a result it is extremely difficult to survey the entire population. The only component that can be readily counted are pups during their first three weeks of life, when they spend almost all of their time ashore at the breeding colonies. All of the major breeding colonies have been surveyed from the air in almost every year since 1962, although survey methods were changed some-what after the tragic death of the field crew in a plane crash in 1983. The aerial surveys are used to produce annual colony-specific estimates of pup production (Hiby & Duck, in press). Estimated pup production has increased at an approximately exponential rate of 6% per year since records began, although individual breeding colonies have exhibited diverse dynamics, including fluctuations around exponential or logistic increases, decreases to extinction and a few new colonizations (Hiby & Duck, in press). Since seals eat fish, the overall increase is of great concern to Scottish fishermen and others.

Here, we present a preliminary model of the Grey Seal metapopulation dynamics for the period 1984–2000. In this model, we have aggregated the colonies into four geographi-cally distinct regions: North Sea (4 colonies), Inner Hebrides (19 colonies), Outer Hebrides (11 colonies) and the Orkneys (22 colonies); these are sometimes indexed 1–4 below. Our first objectives were to find out if important population parameters could be estimated based on pup production data alone, and to test whether a model for density-dependent migration

could account for the observed pattern of metapopulation dynamics. Such a model has been proposed by Ruxton & Rohani (1999) and there is some genetic evidence that it may apply to British Grey Seals (Gaggiotti *et al.*, 2002). Adult female seals show strong philopatry (Pomeroy, Twiss & Redman, 2000), so any density-dependent migration would most likely occur when young females recruit to the breeding population. Our ultimate objective is to predict the long-run stationary distribution of seals by colony.

3.1. Model formulation

Grey Seals breed colonially in the autumn, when breeding females produce a single offspring. After weaning, both pups and adults disperse and may travel large distances during the rest of the year to forage (McConnell *et al.*, 1999). Female offspring recruit to the breeding population at about five years old. Grey Seals are polygynous, with mating occurring at the breeding colonies towards the end of the breeding season.

The time step for the process model is 1 year, beginning just after the breeding season. The state vector is of length 7 for each region, representing pups (n_0), age 1 females (n_1) through to age 5 females (n_5), and then age 6 and older females grouped into a single category (n_{6+}). In general, n_{art} represents the number of seals of age a in region r at time t.

The process model is divided into four subprocesses: survival, age incrementation, movement and breeding. These can be represented in matrix form by the matrices $\boldsymbol{S}_t$, $\boldsymbol{A}$, $\boldsymbol{M}_t$ and $\boldsymbol{B}$, respectively (Table 1). The process model can be written

$$\mathrm{E}(\boldsymbol{n}_t \mid \boldsymbol{n}_{t-1}) = \boldsymbol{B}\boldsymbol{M}_t\boldsymbol{A}\boldsymbol{S}_t\boldsymbol{n}_{t-1}.$$

The expectation is approximate because the expected pup survival is a nonlinear function of the number of pups born in the previous year (see below). Multiplying the subprocess matrices together yields the generalized Leslie matrix shown in Table 2.

We assume that survival of adults during the non-breeding season is density independent, with binomial survival probability ϕ_a. Survival of pups is assumed to be a density-dependent function of the number of pups born in that region in the previous year,

$$u_{s0rt} \stackrel{\mathrm{d}}{=} \mathrm{Bi}(n_{0r,t-1}, \phi_{prt}),$$

where u_{s0rt} is the number of pups surviving and ϕ_{prt} is the probability of survival, given by

$$\phi_{prt} = \frac{\phi_p^*}{1 + \beta_r n_{0r,t-1}}.$$

This is equivalent to the Beverton–Holt stock recruitment model commonly used in fisheries (Quinn & Deriso, 1999). Here ϕ_p^* is the expected survival rate of pups when the number of pups is low, and β_r determines the rate at which density dependence reduces productivity in a region with increasing numbers of pups. The β_r parameter can be thought of as being related to the region-specific carrying capacity: smaller values of this parameter mean that the population in a region can grow larger before density-dependent reduction in productivity slows the population growth.

The age incrementation subprocess includes sexing of pups. Since we model only adult females, the number of age 1 females is a binomial random variable, $u_{a1rt} \stackrel{\mathrm{d}}{=} \mathrm{Bi}(u_{s0rt}, 0.5)$.

TABLE 1

Component matrices for the British Grey Seal process model. Only the parts of the matrices representing the first and last of the four regions are shown, the rest being indicated by dots ($\cdots$). S_t is the survival matrix, A is age incrementation and sexing of pups, M_t is movement of recruiting females between colonies, and B is breeding. Notation of parameters within each matrix is given in the text.

$$
S_t = \begin{bmatrix}
\phi_{p1t} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \phi_a & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \phi_a & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \phi_a & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \phi_a & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \phi_a & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \phi_a & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \phi_{p4t} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & \phi_a & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \phi_a & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \phi_a & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \phi_a & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \phi_a & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \phi_a
\end{bmatrix}
$$

$$
A = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.5 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 & 1
\end{bmatrix}
$$

$$
M_t = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \rho_{1\to1,t} & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \rho_{4\to1,t} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & 0 & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \rho_{1\to4,t} & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \rho_{4\to4,t} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

$$
B = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & \alpha & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \alpha \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

For all other ages, age incrementation is deterministic, with animals ageing by one year except for age 6+ females, who remain in the same age class.

To model movement, we assume that only females in their last year before breeding may move — once a female has started breeding she remains faithful to that region. We assume that movement is density dependent, such that females will not move unless the expected survival of their future offspring is higher elsewhere, in which case probability of movement is proportional to the expected survival difference. In addition, we assume that females are more likely to move among regions that are relatively close together, and that females show some degree of site fidelity — that is, they may not move even if conditions for their offspring

TABLE 2

Generalized Leslie matrix for the British Grey Seal process model. Only the parts of the matrices representing the first and last of the four regions are shown, the rest being indicated by dots $(\cdots)$.

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & \phi_a\alpha & \phi_a\alpha & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.5\phi_{p1t} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \phi_a & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \phi_a & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \phi_a & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \phi_a\rho_{1\to1,t} & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \phi_a\rho_{4\to1,t} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \phi_a & \phi_a & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \phi_a\alpha & \phi_a\alpha \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0.5\phi_{p4t} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & \phi_a & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \phi_a & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \phi_a & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \phi_a\rho_{1\to4,t} & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \phi_a\rho_{4\to4,t} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \phi_a & \phi_a
\end{bmatrix}
$$

are better elsewhere. We model movement from each region as a multinomial random variable,

$$
(u_{m5,r\to1,t}, \ldots, u_{m5,r\to4,t}) \stackrel{\mathrm{d}}{=} \mathrm{Mn}_4(u_{a5rt}, \rho_{r\to1,t}, \ldots, \rho_{r\to4,t}),
$$

where $u_{m5,r\to i,t}$ is the number of age 5 females moving from region r to region i, u_{a5rt} is the number of age 5 females at region r after age incrementation but before movement, and $\rho_{r\to i,t}$ is the probability of movement from region r to region i.

$$
\rho_{r\to i,t} = \frac{\theta_{r\to i,t}}{\sum_{j=1}^{4}\theta_{r\to j,t}},
$$

where

$$
\theta_{r\to i,t} = \begin{cases} \exp(\gamma_{\mathrm{sf}}) & r = i, \\ \exp(\gamma_{\mathrm{dd}}\max([\phi_{pit} - \phi_{prt}], 0) - \gamma_{\mathrm{dist}}d_{ri}) & r \neq i \end{cases}
$$

and γ_{sf} is the parameter regulating the strength of the site fidelity, γ_{dd} regulates the effect of density dependence on movement and γ_{dist} regulates the effect of distance between regions, d_{ri}, on movement.

We model breeding by assuming that the number of pups produced is a density independent function of the number of breeding females in the region, with binomial probability α.

For the observation process, we assume that pup production estimates follow a normal distribution, where the variance is a linear function of the true pup production squared:

$$
y_{0rt} \stackrel{\mathrm{d}}{=} \mathrm{N}(n_{0rt}, \psi^2 n_{0rt}^2). \tag{2}
$$

The parameter ψ determines the coefficient of variation (CV) of the observations. This model was found to fit the observed regional-level data well, but is likely to be a poor description of the observation process at the colony level (see Discussion).

TABLE 3

*Prior distributions used for parameters in the British Grey Seal model
(see also Figure 2 for plot of priors). Each parameter was initialized independently.*

Parameter	Prior	Mean
ϕ_a	Be (22.05, 1.15)	0.95
ϕ_p^*	Be (14.53, 6.23)	0.7
β_1	Ga (4, 2.07×10^{-4})	8.29×10^{-4}
β_2	Ga (4, 2.96×10^{-4})	1.18×10^{-3}
β_3	Ga (4, 7.40×10^{-5})	2.96×10^{-4}
β_4	Ga (4, 5.76×10^{-5})	2.30×10^{-4}
γ_{sf}	Ga (2.25, 1.33)	3
γ_{dd}	Ga (2.25, 1.33)	3
γ_{dist}	Ga (2.25, 1.33)	3
α	Be (22.05, 1.15)	0.95
ψ	Ga (4, 0.025)	0.10

The state–space model therefore has 11 parameters: $\{\phi_a, \phi_p^*, \beta_1, \dots, \beta_4, \gamma_{\text{sf}}, \gamma_{\text{dd}}, \gamma_{\text{dist}}, \alpha, \psi\}$. We fit the model using the SIS algorithm described earlier, with 250 000 particles. Prior distributions used for the parameters are shown in Table 3. These priors were selected after discussion with Grey Seal biologists, from the Sea Mammal Research Unit at St Andrews, and simulation exercises, where we simulated from the process model using a range of parameter values. Expected values were chosen to represent values considered most likely by the biologists, and variances set so that the middle 95% of the distribution of each parameter covered all values considered feasible. The results are likely to be sensitive to choice of at least some of the priors.

Priors for the states were generated using the estimated pup production in 1984 in conjunction with the priors for the parameters, as follows. The initial number of pups in each particle, $n_{0r0}^{[j]}$ (where j denotes the particle number, $j = 1, \dots, 250\,000$), was generated by 'reversing' the observation equation, i.e. by sampling from

$$n_{0r0}^{[j]} \stackrel{\text{d}}{=} \text{N}(y_{0r0}, \psi^{2[j]} y_{0r0}^2),$$

where y_{0r0} is the estimated pup production in region r in 1984 (year 0) and $\psi^{[j]}$ is the value of the ψ parameter sampled from the prior for particle j. It is more important that the distribution of initial values encompasses all likely values than that the distribution of initial values has minimum variance, so in all cases we further dispersed the initial values by resampling each one from a uniform distribution with bounds given by the sampled value $\div 1.3$ and the sampled value $\times 1.3$. (The value 1.3 was chosen by trial and error.) We then generated initial values for age 1 females from the survival subprocess, i.e. by sampling from

$$n_{1r0}^{[j]} \stackrel{\text{d}}{=} \text{Bi}(n_{0r0}^{[j]}, 0.5\phi_{pr0}^{[j]})$$

(the 0.5 is because only half of the pups are expected to be female), and ages 2 to 5 females from

$$n_{ar0}^{[j]} \stackrel{\text{d}}{=} \text{Bi}(n_{a-1,r0}^{[j]}, \phi_a^{[j]}),$$

where $a = 2, \dots, 5$. Again, we resampled the values from a uniform distribution to further disperse them. Lastly, we generated initial values for age 6+ females by 'reversing' the fecundity subprocess, i.e. by sampling from

$$n_{6r0}^{[j]} \stackrel{\text{d}}{=} \text{Nb}(n_{0r0}^{[j]}, \alpha^{[j]}) + n_{0r0}^{[j]}$$
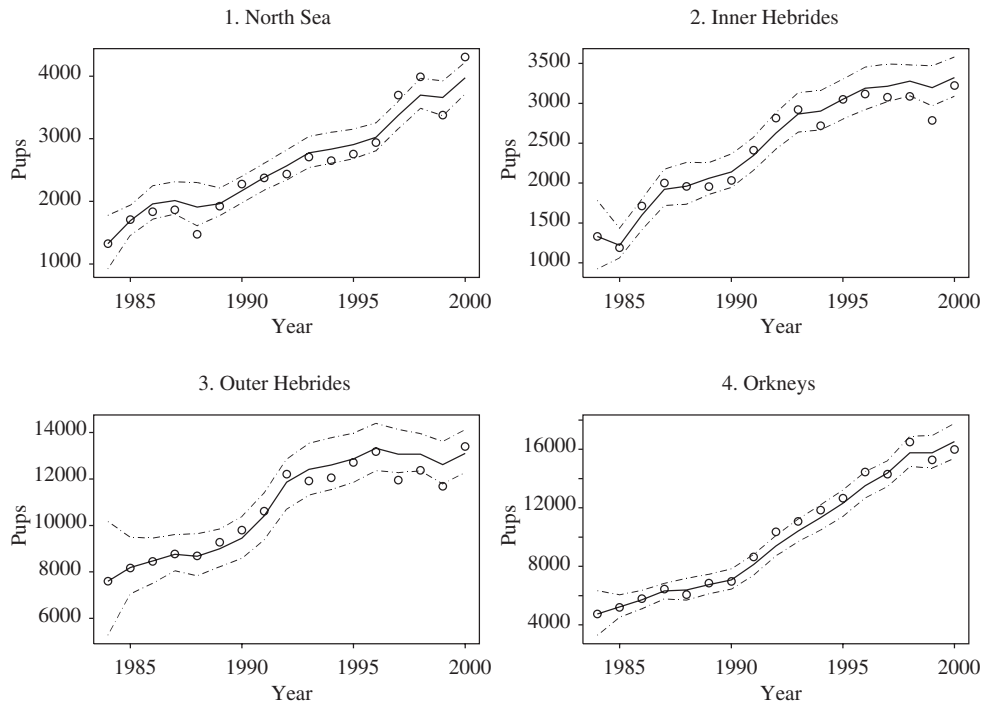
Figure 1. Pup production data (circles) and filtered estimates of pup production (lines) from the British Grey Seal model. First year estimates are starting values. Solid lines show the mean of the particle values, dashed lines show 2.5th and 97.5th percentiles.

and again resampling from a uniform distribution. This procedure produced initial values for all of the states.

Because we used the 1984 values to generate the initial values, we started the fitting algorithm with the 1985 data. The SIS algorithm requires calculation of likelihood weights (1), the form of which is specified by the observation process model. In this example, given the observation model (2), the likelihood is a product of normal densities:

$$f_t(\mathbf{y}_t \mid \mathbf{n}_t^{[j]}, \Theta^{[j]}) = \prod_{r=1}^{4} \left(\frac{1}{\sqrt{2\pi}\psi^{[j]}n_{0rt}^{[j]}} \exp\left(\frac{-(y_{0rt} - n_{0rt}^{[j]})^2}{2\psi^{2[j]}n_{0rt}^{2[j]}} \right) \right).$$

Program code (in S-PLUS 6.1) is available from the journal web site.

3.2. Results

Filtered estimates of pup production for the four regions are shown in Figure 1, together with the pup production data used in fitting the model. The fit appears to be very good, with the algorithm quickly homing in from the wide starting bounds to provide a tight bracket on the observed counts. There appears to have been a sudden drop in pup production in 1999 which the filtered estimates struggled to track.

Posterior marginal parameter estimates are shown in Figure 2, with the corresponding priors. Some posterior distributions have changed markedly from the priors, particularly adult survival, ϕ_a and the observation CV parameter ψ. Other parameters, such as maximum
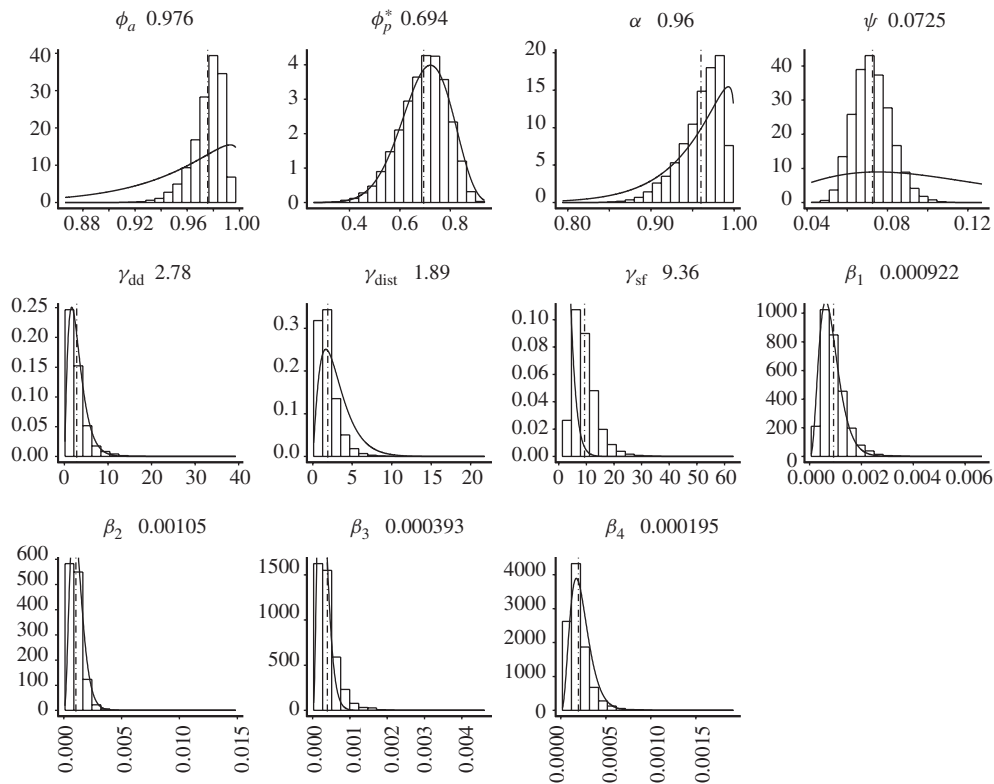
Figure 2. Posterior parameter estimates (histograms) and priors (solid lines) from the British Grey Seal model. Vertical line shows posterior mean, and this value is given in the title of each plot after the parameter name.

pup survival ϕ_p^* and the movement parameters, γ_{sf}, γ_{dd} and γ_{dist}, have virtually unchanged marginal distributions. Given the model structure and available data, we had expected some parameters to be strongly confounded. Scatterplots of the posterior parameter estimates indicate some dependencies among the posterior distributions, but less than we had expected (highest correlation is between ϕ_a and fecundity, α, with $r = -0.48$, but the next highest is between ϕ_a and ϕ_p^*, with $r = -0.09$).

4. Discussion

4.1. Grey Seal model

It is not surprising that these data contain strong information about adult survival, given that the species is long-lived and productive into old age: small errors in estimated adult survival would therefore lead to large errors in pup production predictions, which would be penalized by the likelihood weights. We were not expecting the data to contain much information on animal movement, so we were surprised to see the site fidelity parameter move so far from its prior. The posterior estimate is high enough that there is almost no movement of recruiting females among regions, regardless of differences in survival or distance between regions. This implies that either some part of the model is substantially incorrect, or that there

is indeed little movement of animals between regions. We plan to try a number of alternative models, as well as moving to a colony-level analysis before drawing conclusions.

Our model can be improved in a number of other respects. Constant survival and fecundity rates could be replaced by time-varying rates in a hierarchical or random-effects framework (e.g. Newman, 2000), where for example fecundity rate is sampled from a hyper-distribution that is slowly varying through time. This would help, for example, to fit the 1999 data, where pup production was unexpectedly low. Our observation model was adequate for the regional level, where observations have approximately the same order of magnitude. However, the assumption of constant CV in estimated pup production is not appropriate at the colony level, where production estimates of near zero at some colonies may be quite inaccurate (because little emphasis is placed on counting pups at colonies that account for a very small part of the overall population). A mixture model for the observation process may be appropriate in this case. Ultimately, we would like to develop an observation model for the raw counts taken from the repeated aerial survey photographs, rather than the estimated pup production for the season. This may allow us to estimate the pre-weaning mortality of pups on the colonies and, under the assumption that this is correlated with first year survival, improve our estimates of pup survival.

British Grey Seals are relatively well studied, and there are many other sources of data that could provide information for the model, either to improve the priors or to be incorporated explicitly through additional observation models. For example, there are mark–recapture data from two intensive study sites, some genetic data on relatedness (Gaggiotti *et al.*, 2002), information on body condition of pups that could be extracted from the aerial surveys, and other information that could be used to provide covariates for pup-carrying-capacity of colonies. Any available covariates are easily incorporated into the framework, by modifying the model parameters to make them functions of the covariates.

Another area of interest is in determining which data collection methods may in the future yield the most information about parameters of interest. For example, performing the aerial survey only every second year may free up resources for an extensive mark–recapture exercise aimed at estimating survival and migration rates. Fitting the above models to simulated data can help in evaluating the trade-offs between different possible allocations of survey effort.

4.2. Future methodological research

Given that the fitting method introduces Monte Carlo error into the posterior estimates, it is important to check the reliability of the results. Of the 250 000 independent particles that started the simulation presented here, only 18 790 survived the first year and only 39 were represented by descendants at the final time period. A large amount of particle depletion is typical in the first year, as generating the particles from independent priors produces many that have an impossible combination of parameters and/or states. Having only 39 of the original ancestral particles represented at the end is not necessarily a problem either, since kernel smoothing of the parameters means that each particle generated from the ancestral particle has a slightly different set of parameters. Nevertheless, it is important that the remaining particles span the entire region of parameter and state space that has posterior density. There is a need to develop diagnostic statistics that give an indication of the 'mixing' of particles generated from different ancestral particles.

Once past the first few time periods, severe particle depletion at one time step indicates a data value that is not supported by the model — because either the data value is an outlier or

the model is incorrect. A useful diagnostic in model checking therefore is the level of particle depletion at each time step.

A crude measure of reliability is to re-run the simulations, drawing a new set of starting values. For the model presented here, three independent runs produced almost identical results. It may be possible to use information from within and between repeat runs to estimate the Monte Carlo error rates on estimates efficiently.

The state estimates presented here are filtered estimates, that is estimates of $g(\boldsymbol{n}_t \mid \boldsymbol{y}^t)$ for $t = 1, \ldots, T$. It can be argued that if we are interested in estimating past states, we want to use all the data available to us, i.e. we are interested in smoothed estimates $g(\boldsymbol{n}_t \mid \boldsymbol{y}^T)$ for $t = 1, \ldots, T$. Smoothing can be done analytically in linear normal models, using the backwards Kalman smoother (Harvey, 1989), and is feasible (although very computer intensive) in nonlinear, non-normal models where the state process density can be evaluated (Doucet, Godsill & West, 2000). However, in the case of most wildlife population dynamics process models, where the state process density cannot be easily evaluated, an efficient and general smoothing algorithm for SIS has yet to be developed.

Several aspects of the fitting algorithm need further research. For example, local kernel smoothing may prove better than global kernel estimation when the relationship between parameters is nonlinear. The amount of kernel smoothing to perform is also not well understood (although see Trenkel *et al.*, 2000). There are many strategies in addition to those mentioned here to reduce the effects of particle depletion (e.g. papers in Doucet *et al.*, 2001; Liu, 2001) and these should be investigated in the context of wildlife population dynamics models. There is some hope for an omnibus suite of methods that could be applied semi-automatically once the state and process models have been formulated. Lastly, the potential for fitting the models outlined here using MCMC needs further exploration. Smoothed estimates of the states is an automatic by-product of MCMC output.

The framework outlined here is an ideal vehicle for model selection and model averaging. If a number of models are specified *a priori*, and these are given prior weights, then the frequencies of the different models at the end is an estimate of their posterior probability given the data. The likelihood weights in the correction step would need to be replaced by some form of penalized likelihood weights (e.g. Akaike Information Criterion (AIC) weights). This approach becomes impractical when there are a very large number of candidate models — for example, in multiple regression with many potential covariates. An algorithm for SIS in this circumstance remains to be developed (although algorithms for MCMC methods exist; see Green, 2003).

Relaxing the requirement for a fully parametric specification of the models would be an advantage when the form of the model is not known, as is commonly the case. Semi-parametric differential equation models have been developed by Wood (1999, 2001), and extending this into the framework described here would have wide utility.

As we have shown, the state–space framework can be seen as an extension of the matrix population methods comprehensively documented by Caswell (2001). Many of the concepts associated with those methods can also be extended into the state–space framework. For example, perturbation analysis is used to predict which life-history components have the largest potential impact on outputs of interest, such as asymptotic growth rate (Caswell, 2005). Such analyses, based on the elasticities and sensitivities of matrix models, have been criticized because they usually do not incorporate demographic and environmental stochasticity as well as density dependence in the analysis (Benton & Grant, 1999). Using SIS, sensitivities and/or

elasticities could be calculated for each particle, and the resulting distribution of sensitivities/elasticities used to perform perturbation analyses that incorporate all of these features, as well as model uncertainty if required.

In conclusion, this framework is applicable to a wide array of current problems in wildlife management and conservation. Although there are several methodological issues still to be resolved and improvements to be made, we anticipate that these methods will receive widespread use in the future.

References

BENTON, T.G. & GRANT, A. (1999). Elasticity analysis as an important tool in evolutionary and population ecology. *TREE* **14**, 467–471.

BESBEAS, P., FREEMAN, S.N. & MORGAN, B.J.T. (2005). The potential of integrated population modelling. *Aust. N. Z. J. Stat.* **47**, 33–46.

BESBEAS, P., LEBRETON, J-D. & MORGAN, B.J.T. (2003). The efficient integration of abundance and demographic data. *Appl. Statist.* **52**, 95–102.

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.

BUCKLAND, S.T., BURNHAM, K.P. & AUGUSTIN, N.H. (1997). Model selection: an integral part of inference. *Biometrics* **53**, 603–618.

BUCKLAND, S.T., NEWMAN, K.B., THOMAS, L. & KÖSTERS, N. (2004). State–space models for the dynamics of wild animal populations. *Ecological Modelling* **171**, 157–175.

CASWELL, H. (2001). *Matrix Population Models: Construction, Analysis and Interpretation*, 2nd edn. Sunderland, Massachusetts: Sinauer Associates, Inc.

CASWELL, H. (2005). Sensitivity analysis of the stochastic growth rate: three extensions. *Aust. N. Z. J. Stat.* **47**, 73–83.

DOUCET, A., DE FREITAS, N. & GORDON, N. (Eds) (2001). *Sequential Monte Carlo Methods in Practice.* New York: Springer-Verlag.

DOUCET, A., GODSILL, S.J. & WEST, M. (2000). Monte Carlo filtering and smoothing with application to time-varying spectral estimation. In *Proceedings of the 2000 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. II. Institute of Electrical and Electronics Engineers.

DURBIN, J. & KOOPMAN, S.J. (2001). *Time Series Analysis by State Space Methods.* Oxford, UK: Oxford University Press.

GAGGIOTTI, O.E., JONES, F., LEE, W.M., AMOS, W., HARWOOD, J. & NICHOLS, R.A. (2002). Patterns of colonization in a metapopulation of grey seals. *Nature* **416**, 424–427.

GILKS, W.R., RICHARDSON, S. & SPIEGELHALTER, D.J. (Eds) (1996). *Markov Chain Monte Carlo in Practice.* London: Chapman and Hall.

GREEN, P.J. (2003). Trans-dimensional Markov chain Monte Carlo. In *Highly Structured Stochastic Systems*, eds P.J. Green, N.L. Hjort & S. Richardson. Oxford: Oxford University Press.

HARVEY, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge, UK: Cambridge University Press.*

HIBY, L.R. & DUCK, C.D. (in press). Point and interval estimates of the size of the British grey seal *Halichoerus grypus* population and their implications for management. *J. Appl. Ecol.*

HOETING, J.A., MADIGAN, D., RAFTERY, A.E. & VOLINSKY, C.T. (1999). Bayesian model averaging: a tutorial. *Statist. Sci.* **14**, 382–417.

LEFKOVITCH, L.P. (1965). The study of population growth in organisms grouped by stages. *Biometrics* **21**, 1–18.

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.S. (2001). *Monte Carlo Strategies in Scientific Computing.* New York: Springer-Verlag.

LIU, J.S. & CHEN, R. (1998). Sequential Monte Carlo methods for dynamic systems. *J. Amer. Statist. Assoc.* **93**, 1032–1044.

LIU, J. & WEST, M. (2001). Combining parameter and state estimation in simulation-based filtering. In *Sequential Monte Carlo Methods in Practice*, eds A. Doucet, N. de Freitas & N. Gordon, pp. 197–223. New York: Springer-Verlag.

LUDWIG, D., HILBORN, R. & WALTERS, C. (1993). Uncertainty, resource exploitation, and conservation: lessons from history. *Science* **260**, 36.

MCCONNELL, B.J., FEDAK, M.A., LOVELL, P. & HAMMOND, P.S. (1999). Movements and foraging areas of grey seals in the North Sea. *J. Appl. Ecol.* **36**, 573–590.

MENDELSSOHN, R. (1988). Some problems in estimating population sizes from catch-at-age data. *Fish. Bull.* **86**, 617–630.

MEYER, R. & MILLAR, R.B. (1999). Bayesian stock assessment using a state–space implementation of the delay difference model. *Canad. J. Fish. Aquat. Sci.* **56**, 37–52.

NEWMAN, K.B. (1998). State–space modeling of animal movement and mortality with application to salmon. *Biometrics* **54**, 274–297.

NEWMAN, K.B. (2000). Hierarchic modeling of salmon harvest and migration. *J. Agric. Biol. Envir. S.* **5**, 430–455.

PITT, M.K. & SHEPHARD, N. (1999). Filtering via simulation: auxiliary particle filters. *J. Amer. Statist. Assoc.* **94**, 590–599.

POMEROY, P.P., TWISS, S.D. & REDMAN, P. (2000). Philopatry, site fidelity and local kin associations within grey seal breeding colonies. *Ethology* **105**, 899–919.

PUNT, A.E. & HILBORN, R. (1997). Fisheries stock assessment and decision analysis: a review of the Bayesian approach. *Rev. Fish Biol. Fish.* **7**, 35–63.

QUINN, T.J. II & DERISO, R.B. (1999). *Quantitative Fish Dynamics.* New York: Oxford University Press.

RAFTERY, A.E., GIVENS, G.H. & ZEH, J.E. (1995). Inference from a deterministic population dynamics model for bowhead whales (with Discussion). *J. Amer. Statist. Assoc.* **90**, 402–430.

RUXTON, G.D. & ROHANI, P. (1999). Fitness dependent dispersal in metapopulations and its consequences for persistence and synchrony. *J. Anim. Ecol.* **67**, 530–539.

SCHNUTE, J.T. (1994). A general framework for developing sequential fisheries models. *Canad. J. Fish. Aquat. Sci.* **42**, 414–429.

SMITH, A.F.M. & GELFAND, A.E. (1992). Bayesian statistics without tears: a sampling-resampling perspective. *Amer. Statist.* **46**, 84–88.

THOMPSON, D., HAMMOND, P.S., NICHOLAS, K.S. & FEDAK, M.A. (1991). Movements, diving and foraging behaviour of grey seals *Halichoerus grypus*. *J. Zool.* **224**, 223–232.

TRENKEL, V.M., ELSTON, D.A. & BUCKLAND, S.T. (2000). Calibrating population dynamics models to count and cull data using sequential importance sampling. *J. Amer. Statist. Assoc.* **95**, 363–374.

WEST, M. (1993). Mixture models, Monte Carlo, Bayesian updating and dynamic models. *Comp. Sci. Statist.* **24**, 325–333.

WOOD, S.N. (1999). Semi-parametric population models. *Aspects of Applied Biology* **53**, 41–49.

WOOD, S.N. (2001). Partially specified ecological models. *Ecol. Monog.* **71**, 1–25.

WOOD, S.N. & THOMAS, M.B. (1999). Super-sensitivity to structure in biological models. *Proc. Roy. Soc. London Ser. B* **266**, 565–570.