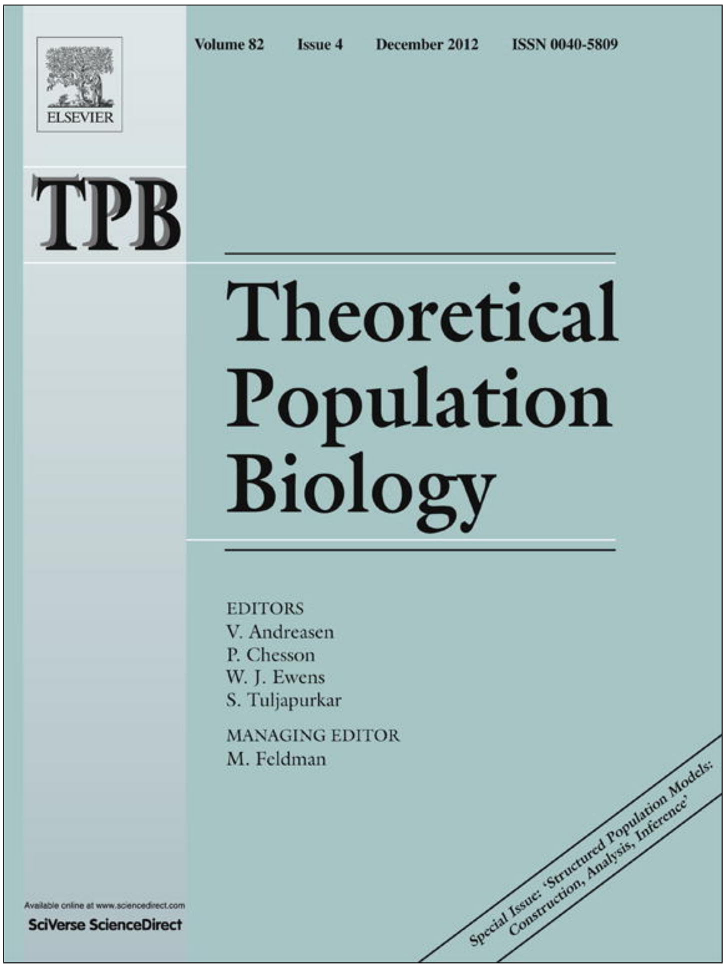

Volume 82 Issue 4 December 2012 ISSN 0040-5809

ELSEVIER

TPB

Theoretical
Population
Biology

EDITORS
V. Andreasen
P. Chesson
W. J. Ewens
S. Tuljapurkar

MANAGING EDITOR
M. Feldman

Special Issue: 'Structured Population Models:
Construction, Analysis, Inference'

Available online at www.sciencedirect.com
SciVerse ScienceDirect

Theoretical Population Biology 82 (2012) 307–316

Contents lists available at SciVerse ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

Estimating demographic parameters using hidden process dynamic models

Olivier Gimenez [a,*], Jean-Dominique Lebreton [a], Jean-Michel Gaillard [b], Rémi Choquet [a], Roger Pradel [a]

[a] *Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, CNRS, 1919 route de Mende, 34293 Montpellier Cedex 5, France*

[b] *Laboratoire de Biométrie et Biologie Evolutive, Unité Mixte de Recherche 5558, Université Claude Bernard Lyon I, 43 Bd du 11 novembre 1918, 69622 Villeurbanne Cedex, France*

ARTICLE INFO

Article history:
Available online 18 February 2012

Keywords:
Capture–recapture
Hidden Markov models
Multievent models
Multistate models
State-space models
State uncertainty

ABSTRACT

Structured population models are widely used in plant and animal demographic studies to assess population dynamics. In matrix population models, populations are described with discrete classes of individuals (age, life history stage or size). To calibrate these models, longitudinal data are collected at the individual level to estimate demographic parameters. However, several sources of uncertainty can complicate parameter estimation, such as imperfect detection of individuals inherent to monitoring in the wild and uncertainty in assigning a state to an individual. Here, we show how recent statistical models can help overcome these issues. We focus on hidden process models that run two time series in parallel, one capturing the dynamics of the true states and the other consisting of observations arising from these underlying possibly unknown states. In a first case study, we illustrate hidden Markov models with an example of how to accommodate state uncertainty using Frequentist theory and maximum likelihood estimation. In a second case study, we illustrate state-space models with an example of how to estimate lifetime reproductive success despite imperfect detection, using a Bayesian framework and Markov Chain Monte Carlo simulation. Hidden process models are a promising tool as they allow population biologists to cope with process variation while simultaneously accounting for observation error.

1. Introduction

Structured population models (Tuljapurkar and Caswell, 1997; Caswell, 2001) allow a detailed description of population change over time by accounting for the number of individuals in discrete classes (grouped according to age, stage or state). These models are widely used in plant and animal demographic studies to assess population dynamics as they allow the projection of future population size, estimation of asymptotic population growth rate, age structure, net reproductive rate, generation time and life expectancy among other descriptive statistics. Matrix models have been used to assess population viability (e.g., Ghimire et al., 2008), evaluate the effects of management strategies on life history strategies (e.g., Servanty et al., 2011) and investigate the feedback between demography and evolution (Van Tienderen, 2000 see also Gamelon et al., 2011 for a recent application).

Traditionally, structured population models are calibrated using longitudinal data collected at the individual level. These data are analyzed with statistical models to estimate demographic parameters like survival, growth, dispersal and fecundity, which are then entered into structured population models. However, several sources of uncertainty are associated with parameter estimation that need to be properly accounted for. First, individuals are difficult to monitor in the wild (mobile animals in particular) because they may or may not be observed at a given sampling occasion. This raises the issue of imperfect detection, which, if ignored, can lead to biased estimates and flawed inference about demographic parameters (e.g., Nichols, 1992; Gimenez et al., 2008). Second, there can be uncertainty related to measurement error, in particular, correctly assigning a state to an individual, even when the individual is observed or captured. This may be the gender of an individual whenever sex is determined through behavioral clues, or epidemiological status if serological tests generate false positives or negatives, or reproductive status whenever the number of offspring is visually determined by counting the number of young accompanying their parents.

To handle the complexity of the dynamics of free-ranging populations and accommodate the various sources of uncertainty associated with parameter estimation, there is a need to integrate structured population models with novel statistical tools. In this paper, we aim to review and demonstrate how recent statistical models can help bring theoretical models (mechanisms) and the analysis of empirical data (patterns) together. We will focus our attention on hidden process models (see glossary) that have recently received much attention in the literature (Buckland et al., 2004, 2007; Newman et al., 2006). In hidden process models, two time series are run in parallel: one captures the dynamics of the true (latent) states via a stochastic process model; the other is

---

[*] Corresponding author.
E-mail address: olivier.gimenez@cefe.cnrs.fr (O. Gimenez).

0040-5809/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.tpb.2012.02.001

static and describes the observations that are made from these underlying but possibly unknown states.

In the first section, we review standard multistate capture–recapture models (see glossary; Lebreton et al., 2009) that are used to estimate key demographic parameters, which are then fed to structured population models. We pay particular attention to the imperfect detection of individuals in the wild for analyzing individual longitudinal data. In the second and third sections, we reformulate longitudinal multistate models as hidden process models, either as hidden Markov models (HMMs; see glossary, Pradel, 2005) or state-space models (SSMs; see glossary, Gimenez et al., 2007). In both sections, case studies are used to illustrate these models. In the second section, we show how hidden-Markov models can be used to deal with uncertainty in the assignment of reproductive status and its importance when assessing life-history trade-offs. In the third section, we show how state-space models can be used to estimate lifetime reproductive success (sensu Clutton-Brock, 1988) while accounting for a detection probability less than one. Finally, we discuss the limitations of hidden process models as well as several extensions of these models.

2. Overview of multistate capture–recapture models

Capture–recapture (CR; see glossary) models have become a central tool in population biology for estimating demographic parameters under imperfect detection of individuals (Lebreton et al., 1992). CR models were originally designed to estimate population size, with demographic parameters such as survival probabilities progressively incorporated in these models, first as nuisance parameters, and eventually as the main focus (for the early history see Lebreton et al., 2009 p. 93 ff). The standard Cormack–Jolly–Seber model (see glossary; Cormack, 1964; Jolly, 1965; Seber, 1965) is based on the embedding of two distinct processes: survival (with survival probabilities denoted as ϕ) and detection (with recapture probabilities denoted as p) over discrete time occasions. These parameters are indexed appropriately depending on the type of variation of interest (e.g., time, age or sex; Lebreton et al., 1992).

A notable feature of the original development of CJS models is that the survival and the detection processes were viewed as intimately entangled. Adequately representing how they were imbricated was indeed at the core of early methodological developments to obtain a likelihood that could be used to properly estimate the parameters. For instance, under time-dependence, the probability that an individual is never seen again after occasion i, denoted as χ_i, is obtained through a backward recurrence relationship (Cormack, 1964), accounting for the fact that an individual may die between occasion i and $i+1$ or survive and not be detected:

$$\chi_i = 1 - \phi_i + \phi_i(1 - p_{i+1})\chi_{i+1}$$

with $\chi_{\text{last occasion}} = 1$.

However, when comparing various types of detection, it clearly appeared that the survival–mortality process was a background process distinct from the detection process. For instance for the same time-dependent survival process, one must be able to accommodate several situations illustrated in Fig. 1 (see also Lebreton et al., 1999): (i) complete detection, as might occur in developed countries for the human population (with administrative declaration of deaths), (ii) incomplete detection of live individuals, i.e., live recaptures or resightings, to be analyzed with CJS models or (iii) incomplete detection of deaths, as in 'dead recoveries' of banded birds (usually by the general public) to be analyzed with so-called dead-recovery models (Brownie et al., 1985).

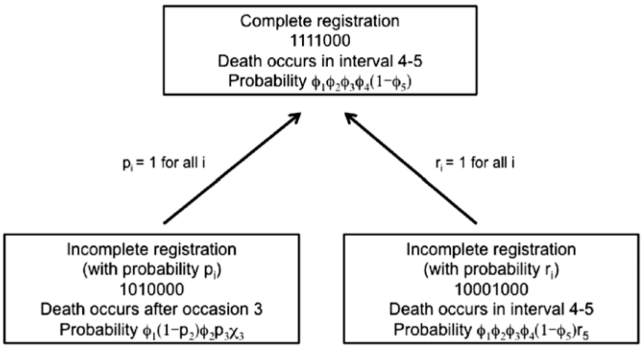

Fig. 1. Three types of registration of a same discrete time survival process. Top: follow-up data, bottom left: CJS models, bottom right: dead recovery models (see text for details). The complete registration is a particular case of each of the two other ones.

In the survival process, death or survival between two sampling occasions is treated as strictly conditional on the fact that the individual is alive at the onset of the target interval: the system has no memory, i.e. the survival process is a Markov chain. Hence, CJS models are implicitly made up of a Markov chain on top of which is an imperfect observation process (e.g., Lebreton et al., 1999). In other words, CJS models are hidden process models where the demographic process is represented by a (first-order) Markov chain (see next sections) that is only partially observed due to imperfect detection. Conveniently, the explicit representation of the survival process in CJS models as a two-state Markov chain (alive/dead) readily provides the probability of individual histories, without having to use the complicated χ parameters mentioned above (Caswell and Fujiwara, 2004, pp. 477–478).

However, representation of the survival process as a Markov chain did not capture much attention for some time for a number of reasons. First, CR models were considered to be specific and exotic statistical tools, in particular because of the historical focus on population size estimation and the associated difficulties in producing an adequate fully stochastic model (Jolly, 1965). The focus on the estimation of demographic flows developed only in the late 80s (Burnham et al., 1987; Clobert and Lebreton, 1987). Second, while the succession of events such as survival and death in CR models was clearly Markovian, historically, the statistical point of views on CR models and Markov chains were radically different: when considering n individuals moving over T time steps according to a Markov chain, the development of stochastic process theory for Markov chains was centered on asymptotic results when T tends to infinity, such as ergodicity for instance; on the contrary, CR models asymptotic results, such as the optimality of maximum likelihood estimators for example, are derived for n tending to infinity for a fixed value of T. Third, parameter estimation for Markov chains under partial observation was not fully developed at that time (e.g., Lawless and McLeish, 1984) and could not offer solutions to the CR situation.

The initially slow development of the Markov chain approach to analyzing individual histories is illustrated in Fig. 2. The first papers explicitly using this point of view were published in the early 1990s, and from there on, parallel growth in human and animal population studies took place at a common rate of ~9% a year.

Modeling survival as a hidden process has several strong advantages. First, the process model can easily be made more complex or detailed to address new biological questions. Existing CR extensions, such as multisite (Arnason, 1972, 1973; Hestbeck et al., 1991) and recruitment (Clobert et al., 1994) models, entered this common framework; new generalizations could easily be produced, such as multisite recruitment models (Lebreton et al., 2003) and multistate CR models (reviewed in detail by Lebreton

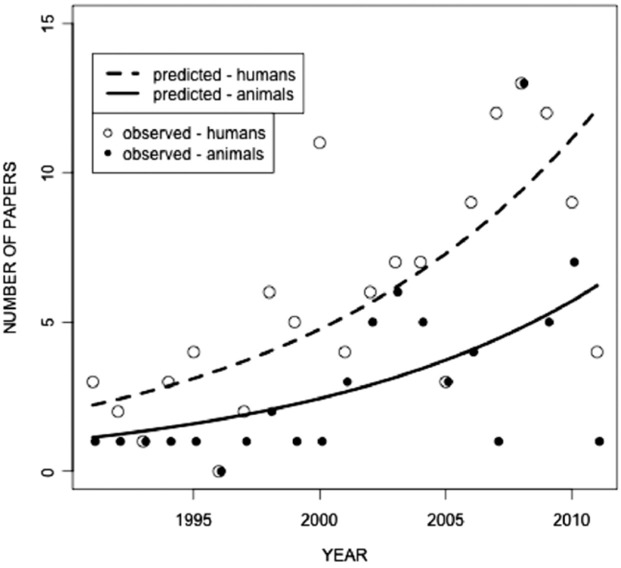

Fig. 2. Number of papers about Markov models (keywords 'Markov' or 'Markovian') in human and animal population biology (keywords 'man or human' and 'animal or mammals or birds' respectively) extracted from the Web of Science (search completed on the 9th of August, 2011). Predicted trends were obtained from a Poisson regression in which both the year and the type of population biology (human vs. animal) effects were statistically significant on an additive scale.

et al., 2009, and Williams et al., 2002 pp. 454–468). Second, it linked CR statistical models with parallel developments in population dynamics models and notably matrix population models (Caswell, 2001 p. 37). These models are based on a reasonable representation of differences among individuals, through a finite number of so-called '*i*-state variables' that summarize relevant information on the past life of an individual to determine its future in terms of, e.g., individual physiology, reproduction or death (Diekmann and Heesterbeek, 2000). The same idea is clearly inherent in multistate CR models. The empirical use of these models requires that individual data should translate into parameter estimates, i.e. estimates of probabilities of the transition from, say, life to death, rosette to flowering plant or non-breeder to breeder. Several papers have addressed estimation and the use of estimates in multistate dynamics models (e.g., Caswell and Fujiwara, 2004). Third, the observation process can also be made more complex and tuned to specific situations. Dupuis (1995) was among the first to distinguish clearly between the two processes in CR models, considering the transitions between states as a Markov chain and imperfect detection as an independent process generating 'missing data'. Soon, however, the detection process was considered to be potentially more complex, e.g., with the development and systematic exploration of the use of unobservable states for modeling temporary emigration (Kendall and Bjorkland, 2001) and recruitment (Pradel and Lebreton, 1999). Specific models for 'mixtures of information' (Burnham, 1993) considered detection of individuals either as alive or at the time of death. They appear as specific multistate CR models based on the two states alive/just dead (Lebreton et al., 1999). The generalization of the observation process currently includes models that account for uncertainty in the assessment of states (Kendall et al., 2003). The observation process then considers a probabilistic relationship between the underlying states (such as breeder vs. non-breeder) and 'events' (such as seen at nest or seen out of a nest). Hence, an individual seen out of a nest may be in either state. Earlier approaches were specific to permanent states (such as sex, Nichols et al., 2004), but the more general models are multistate CR models that account for state uncertainty, so-called multievent models (Pradel, 2005). The movement towards understanding and unifying CR models as hidden process models (Fig. 3) has followed rapidly with the development of specific statistical tools for hidden process models, either hidden-Markov models or state-space models. We review these developments in the following next two sections.

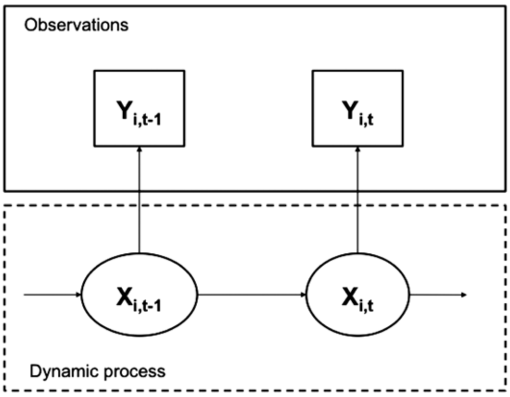

Fig. 3. Schematic representation of a hidden process model for an individual i between two sampling occasions $t-1$ and t. The first layer is a succession of hidden states (solid circles) or latent states of individual i at time t ($X_{i,t}$) depending on its states at time $t-1$ ($X_{i,t-1}$). The dynamics of the states is driven by transition probabilities, e.g., survival probability if states are alive or dead. The second layer corresponds to observation or not of individual i at time t ($Y_{i,t}$) conditional on individual i being in state $X_{i,t}$. These events are driven by some probabilities, e.g., detection probabilities if observations are being captured or not. The observation process is the visible part of the state process. Note that this representation holds for both hidden-Markov models and state-space models.

3. Hidden Markov models: dealing with uncertainty in state assignment

3.1. Hidden-Markov modeling

3.1.1. HMM for capture–recapture data

In the above multistate CR models, it is assumed that the state of an individual is ascertained upon encounter. This poses no difficulty if the state is a geographical location (e.g., wintering site or breeding colony), but is less obvious when the state is reproductive status (breeder or non-breeder) or a physiological state like diseased vs. healthy. Diagnosis is typically attempted based upon some criterion like the presence of a brood patch or nest-building activity (for a breeder) or antibody level (for a potentially diseased individual). The distinction between the criterion and the state leads to models that distinguish the visible layer (the criterion) from the hidden layer (the state). This distinction between what is observed and what is inferred is key to hidden process models. Because the hidden layer is a first order Markov chain, this naturally leads to the introduction of a HMM for analyzing CR data. The key step is to recognize that whether an individual is observed or not may itself be considered as belonging to the 'observable' layer. This conceptual leap led to the definition of multievent CR models (see glossary; Pradel, 2005), which both generalize multistate CR models to accommodate uncertainty in state assignment and firmly establish that CR models are a particular type of HMMs. However, unlike in HMMs, the recorded information in multievent models is named 'events' rather than 'observations', because the record can be a 'non-observation'.

3.1.2. Review of HMM applications to estimate demographic parameters

HMMs have a tremendous potential with applications in all domains of ecology (Pradel, 2009), but they historically have been seldom used. An obvious application is in epidemiology (Conn and Cooch, 2009). Another is sex identification in monomorphic or weakly polymorphic species where clues are mainly derived from behavior (Nichols et al., 2004; Pradel et al., 2010). A less obvious application is investigating the role of memory in movements among a discrete set of sites. When movements between the successive occasions are determined not only by the position occupied at a particular occasion but also by some previous positions, the previous positions must enter in the definition of the state. Then, when an individual is encountered, it may happen that because the previous relevant positions have not been observed, the exact state is unknown. This invalidates the use of multistate CR models but HMMs can be used (Rouan et al., 2009a). The most elaborate application to date makes use of information from nest inspection to study the probability of skipping behavior and nest changing in a colony of Cory's shearwater (Sanz-Aguilar et al., 2011). A more methodological example is the incorporation of mark loss in mark-recapture models during a double-mark study (Juillet et al., 2011). Another interesting application copes with the issue of keeping track of the number of breeding episodes when detectability is less than one (Desprez et al., 2011; Pradel et al., 2011). It is also possible to account for tag loss and different causes of mortality (Tavecchia et al., 2012). However, by far, the greatest number of applications relates to class heterogeneity whereby a given individual belongs to a particular unknown class for the duration of the study. Cubaynes et al. (2010) emphasize the importance of explicitly modeling capture heterogeneity to avoid potentially large biases in the estimation of population size (see Ebert et al., 2010 for a review) and Crespin et al. (2008) point out that heterogeneity in capture rate may mirror social hierarchy. Véran et al. (2007) and Oliver et al. (2010) used capture heterogeneity to correct for unequal sampling over space. Péron et al. (2010a) went further by modeling class heterogeneity in survival, detection and movement simultaneously and assessed its influence on the study of senescence. Péron et al. (2010b) analyzed in further depth the significance of movement class heterogeneity within a population. Given all these examples, we were surprised to not find studies of reproduction that account for uncertainty in the breeding status. Identification of reproductive status is not always obvious and failing to acknowledge this problem may lead to mistakes or at least to a loss of useful information. This issue is particularly important in studies of life-history trade-offs. Here we develop an example of how such data could be handled.

3.2. *Case study 1: dealing with uncertain states*

3.2.1. *Data and models*

Data were collected between 1940 and 1957 by Lance Richdale on Sooty shearwaters *Puffinus griseus* (titis hereafter) and were reanalyzed using multistate CR models by Scofield et al. (2001) (see also Gimenez et al., 2005).

Following the way the data were collected, four states were originally considered: (1) Breeder (B); (2) Keeping company with another bird in a burrow; (3) Alone in a burrow; (4) On the surface. Because of numerical issues, Scofield et al. (2001) pooled 2–3–4 together in a Non-Breeder state (NB) that includes failed breeders (birds that had bred previously but were currently skipping reproduction or divorced) and pre-breeders (birds that had not yet bred). Note that because burrows were not checked before hatching, some birds in the category NB might have bred and failed before being checked. We therefore regard those birds in the *B* state as successful breeders, and those in the NB state as a mixture of non-breeders, pre-breeders and failed breeders.

We start by fitting a model with state-specific survival, detection and transition probabilities between states, with all parameters constant over time. The states were {non-breeder, breeder, dead} = {NB, B, D}, and the events were {non-detected, detected as not a successful breeder, detected as a successful breeder} = {0, 1, 2}. Upon its first encounter, an individual has a probability π^B to be a breeder and the complement probability to be a non-breeder. Then, it moves among the states according to a first-order Markov chain that can be described by the state-specific survival probabilities ϕ^{NB} and ϕ^B and the transition probabilities conditional on survival $\psi^{NB\to B}$ and $\psi^{B\to NB}$. The state process can be represented by the transition matrix with departure states in rows and arrival states in columns:

$$\begin{bmatrix} \phi^{NB}\psi^{NB\to NB} & \phi^{NB}\psi^{NB\to B} & 1-\phi^{NB} \\ \phi^{B}\psi^{B\to NB} & \phi^{B}\psi^{B\to B} & 1-\phi^{B} \\ 0 & 0 & 1 \end{bmatrix}.$$

Equivalently, it can be represented as the product of a survival matrix by a conditional transition matrix:

$$\begin{bmatrix} \phi^{NB} & 0 & 1-\phi^{NB} \\ 0 & \phi^{B} & 1-\phi^{B} \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \psi^{NB\to NB} & \psi^{NB\to B} & 0 \\ \psi^{B\to NB} & \psi^{B\to B} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The event process conditional on underlying states is described via the detection probabilities p^{NB} and p^B. This event can be summarized by a row-stochastic matrix (i.e., probabilities in a row sum up to one) with states in rows and events in columns:

$$\begin{bmatrix} 1-p^{NB} & p^{NB} & 0 \\ 1-p^{B} & 0 & p^{B} \\ 1 & 0 & 0 \end{bmatrix}.$$

For instance, in row 1 column 2 is the probability that a non-breeder is detected as not a successful breeder p^{NB}.

To demonstrate how uncertain detection of reproductive status could be analyzed, we generated uncertainty in the assignment of Non-Breeder vs. Breeder as follows. For each individual at each detection occasion, a Non-Breeder was assumed to be judged a Non-Breeder with probability 0.2 (and assigned to event 1), and not judged with the complementary probability 0.8 (and assigned to the ambiguous event 3), while a Breeder was assumed to be judged a Breeder with probability 0.7 (and then assigned to event 2), and not judged with the complementary probability 0.3 (and assigned to event 3). This procedure was implemented in R (R Development Core Team, 2011; see Appendix A.2). To analyze these imperfect data, only the event process changes. In addition to the detection probabilities, we add state-specific probabilities of breeding state assignment δ^{NB} and δ^B. With the same presentation as before, the event process becomes:

$$\begin{bmatrix} 1-p^{NB} & p^{NB}\delta_{NB} & 0 & p^{NB}(1-\delta_{NB}) \\ 1-p^{B} & 0 & p^{B}\delta_{B} & p^{B}(1-\delta_{B}) \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Note that the new event 'breeding state not ascertained' (column 4) can arise for non-breeders (first row) as well as breeders (second row), though with potentially different probabilities. This matrix can also be written as a product of two matrices, highlighting the successive processes of detection and breeding state ascertainment:

$$\begin{bmatrix} 1-p^{NB} & p^{NB} & 0 \\ 1-p^{B} & 0 & p^{B} \\ 1 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \delta_{NB} & 0 & 1-\delta_{NB} \\ 0 & 0 & \delta_{B} & 1-\delta_{B} \end{bmatrix}.$$

Note that when the animal is first encountered, the capture process is not modeled because an animal must be encountered at least once to enter the dataset. But the reproductive state ascertainment remains valid. Models with or without imperfect assignment of reproductive state were implemented in the freely downloadable program E-SURGE (Choquet et al., 2009; see Appendix A.1 for details).

Table 1
Parameter maximum likelihood estimates (MLEs) with standard errors (SE) for the hidden-Markov model fitted to the titis data, with and without uncertainty in state assignment. All parameters are probabilities. Note that the probability of remaining non-breeder and the probability of making the transition from state breeder to state non-breeder are not displayed but can be obtained as a complement.

Parameter	Reproductive state uncertainty ignored	Reproductive state uncertainty modeled
Being a breeder upon first encounter π^B	0.704 (0.014)	0.704 (0.028)
Survival of non-breeders ϕ^{NB}	0.801 (0.016)	0.814 (0.016)
Survival of breeders ϕ^B	0.849 (0.016)	0.837 (0.019)
Transition non-breeder to breeder $\psi^{NB\to B}$	0.249 (0.021)	0.219 (0.025)
Transition breeder to breeder $\psi^{B\to B}$	0.761 (0.021)	0.774 (0.049)
Detection of non-breeders p^{NB}	0.597 (0.028)	0.565 (0.028)
Detection of breeders p^B	0.563 (0.027)	0.598 (0.032)
Non-breeder state assignment δ^{NB}	NA	0.188 (0.014)
Breeder state assignment δ^B	NA	0.738 (0.055)

3.2.2. *Results*

The confidence intervals for the assignment probabilities include the true value, despite the considerable loss of information due to uncertain state, in particular for the non-breeders which are recognized as such only 20% of the time (Table 1). The estimates of other parameters remain similar to those of the original model. The most notable and logical change is a loss of precision. However, in a real study, there would probably be on the contrary a gain of precision because the data chosen for analysis with multistate models are those deemed most reliable. In contrast, the model with uncertainty would be run on the larger dataset that includes the less reliable data. For this reason, precision should be enhanced by explicitly dealing with breeding state uncertainty.

4. State-space models: estimating lifetime reproductive success

4.1. *State-space modeling*

4.1.1. *SSM for capture–recapture data*

Parallel to the HMM formulation of CR models, there has been increasing use of state-space models (SSM) to handle data on marked individuals. State-space models are a generalization of hidden Markov models in that the states are not necessarily governed by a Markov process. In practice, parameters of SSMs have been estimated in a Bayesian framework (using Markov Chain Monte Carlo [MCMC] simulation of the joint posterior distribution of parameters; see Section 4.2.3) while HMMs have been fit in a Frequentist framework using maximum likelihood methods. CR models were reformulated as SSMs by several authors (Gimenez et al., 2007; Royle, 2008; Schofield and Barker, 2008) to distinguish the underlying demographic process from the observation process.

4.1.2. *Use of SSMs to estimate demographic parameters data—a survey*

The SSM formulation provides a great deal of flexibility and has led to substantial progress in modeling key demographic parameters. Link and Barker (2005) showed how to estimate correlations between demographic parameters such as survival and reproductive rates. Servanty et al. (2010) used this approach to estimate cause-specific mortalities and test for additivity in mortality. Other applications include comparisons among species (Papadatou et al., 2011), spatial structure (Péron et al., 2011) and the study of migration (Calvert et al., 2009; O'Hara et al., 2009). Demographic parameters can be estimated for species with complex life cycles using SSMs (Clark et al., 2005; Buoro et al., 2010). Like HMM, important developments included the incorporation of individual heterogeneity via random effects (Schofield and Barker, 2011, Royle, 2008, Marzolin et al., 2011; see Gimenez and Choquet, 2010 for an implementation of individual random effects in HMM). In addition, Gimenez et al. (2006, 2009) used SSM to estimate selection gradients and visualize fitness surfaces based on capture–recapture data, while Papaïx et al. (2010) showed how to estimate components of phenotypic variance that can be attributed to genetic factors, environmental factors and other unknown factors, hence quantifying heritability of demographic parameters. All these applications have used SSM to estimate demographic parameters. However, SSM also allows the estimation of the unknown latent states. To our knowledge, this has rarely been attempted, although it could be useful in reconstructing individuals' fates and population dynamics (Gimenez et al., in preparation) . We develop an example of this below.

4.2. *Case study 2: measuring lifetime reproductive success in the wild*

4.2.1. *Estimating fitness in the wild can be difficult*

Lifetime Reproductive Success (LRS), the number of young measured at some stage by an individual during its lifespan is a single-generation measure of the individual's contribution to future generations (Clutton-Brock, 1988; Newton, 1989) and thus provides a reliable measure of individual fitness (see Brommer et al., 2004 for an empirical demonstration). The estimation of LRS requires information on the number of offspring an individual has produced and successfully raised at some stage (e.g., born, weaned or fledged, first reproducing) at each reproductive attempt. In theory then, exhaustive monitoring of individuals over their whole reproductive life is needed to estimate LRS. In the field, however, estimating LRS is difficult, because individuals are typically observed (or recaptured) only a fraction of their lifetime. In years when a given individual is not observed, whether it gives birth or not and how many offspring it produces at birth or at the end of the parental care period if it does breed remain unknown. In the following section, we show that, by considering two states (breeder vs. non-breeder) with the associated number of offspring produced in a SSM, reliable estimates of LRS can be obtained while accounting for detection probability less than one. A similar approach was proposed by Rouan et al. (2009b) who used HMM to compute LRS.

4.2.2. *Data and models*

We illustrate our approach using the reproductive histories of 211 female roe deer that were monitored in the Trois Fontaines forest, North-East France, between 1976 and 2006 (see e.g. McLoughlin et al., 2007 for further details about this monitoring). We developed a CR model to analyze these live recaptures of female roe deer. In line with recent work by Gimenez et al. (2007), Royle (2008) and Schofield and Barker (2008), we used a state-space formulation of CR models that explicitly separates the demographic process of interest (survival status in a particular reproductive state) from the observations (the data). We considered five states: alive non-breeder (NB), alive and weaning one fawn (B1), alive and weaning two fawns (B2), alive and weaning three fawns (B3) and dead (D). We denoted $X_{i,t}$ the random state vector taking values $(1, 0, 0, 0, 0)$, $(0, 1, 0, 0, 0)$, $(0, 0, 1, 0, 0)$, $(0, 0, 0, 1, 0)$, $(0, 0, 0, 0, 1)$ if, at time t, individual i is in state NB, B1, B2, B3 or D, respectively. Based on the roe deer life history (see e.g. Gaillard et al., 1998), we considered several possible observations that were generated from these underlying states: the animal is not seen (coded 0), the animal is seen without any fawn (coded 1), the animal is seen with one fawn (coded 2), the animal is seen with two fawns (coded 3) and the animal is seen with three fawns (coded 4) in the dataset. We denoted $Y_{i,t}$ the random observation vector taking values $(1, 0, 0, 0, 0)$, $(0, 1, 0, 0, 0)$, $(0, 0, 1, 0, 0)$, $(0, 0, 0, 1, 0)$, $(0, 0, 0, 0, 1)$ if, at time t, individual i is not seen or seen with zero, one, two or three fawns respectively.

The model has several parameters. We modeled the probability $\phi_{i,t}^{k}$ that an individual i alive at time t in state $k = \text{NB}$ (nonbreeder) or B (breeder whatever the number of fawns weaned) survives to time $t + 1$. Age-dependence in $\phi_{i,t}^{k}$ was modeled using three classes: yearling (1–2 years), prime-aged (2–5 years) and old adults (6 years and older) according to previous survival analyses (see e.g. Gaillard et al. (1993). We also modeled the probability $\psi_{i,t}^{k\to k'}$ that an individual i makes a transition between states k and $k'(k, k' = \text{NB, B1, B2 or B3})$ between occasion t and $t + 1$. Finally, we considered the probability $p_{i,t}^{k}$ that an individual i is recaptured at time t in state $k = \text{NB}$ or B. Note that, although it is possible, we do not consider individual effects on these parameters, so that the index i is dropped hereafter.

The state-space model relies on a combination of two sets of equations: the state equations, which specify the state of individuals at time $t + 1$ given their state at time t, and the observation equations, which specify the observation of individuals at time t given their state at time t. The state-space formulation of the CR model (Gimenez et al., 2007) is given in Box I. In the state equation, the matrix gathers the conditional probabilities of being in a state at a particular time occasion (columns) given the state at the previous time occasion (rows). Given that an individual is alive in a given state, it can survive and move to another state, survive and remain in the same state or die. To ensure that these probabilities are within the interval [0, 1] and sum to 1, we used a generalized (or multinomial) logit link function for the transition probabilities (Choquet, 2008). Consider for example the four transition probabilities from state NB to one of the four states NB, B1, B2 and B3. We considered the probability of remaining in the same state as a reference (one minus the sum of the three other transition probabilities) and used the transformation $\log\left(\frac{\psi_t^{NB\to k}}{\psi_t^{NB\to NB}}\right)$ for $k = \text{B1, B2 or B3}$ (see Appendix B for technical details). In the observation equation, the matrix gathers the conditional probability of being observed or not at a particular time occasion (columns) given the state at this current occasion (rows). Given that an individual is alive in state non-breeder or breeder, it can be recaptured or not in state non-breeder or breeder respectively. Given that an individual is dead, it is missed with certainty. We had also to specify the probability of being in a given state when first captured. Because first reproduction starts at 2 years in roe deer (Gaillard et al., 1992) and females are marked in their first winter at about 8 months of age (hence as non-breeder, Gaillard et al. (1993)), we considered the probability of being in the state non-breeder when first captured to be unity, and null for any other state.

4.2.3. Estimation of LRS using MCMC methods in a Bayesian framework

We used MCMC simulations in a Bayesian context (McCarthy, 2007; King et al., 2009; Kéry and Schaub, 2011; Royle and Dorazio, 2008; Link and Barker, 2010) to fit the above SSM. Prior distributions need to be specified for the parameters of a given model. Bayes' theorem is then used to update the priors using the likelihood to obtain posterior probabilities of all model parameters. The MCMC algorithm then generates values from a Markov chain whose stationary distribution is the joint posterior distributions of all model parameters. We used empirical means and standard deviations to summarize these posterior parameter distributions.

A nice feature of MCMC algorithms is that reproductive states can be treated as parameters to be estimated, just like survival, transition or recapture probabilities (see Newman et al., 2009; Toni et al., 2009 for alternatives to MCMC). We produced posterior distributions for breeding states of individuals, which were used to calculate LRS. Specifically, for each MCMC iteration, we calculated for each individual the number of fawns it weaned per time step and then computed an average over all (post burn-in) iterations for each of the females. Summarizing across individuals, we obtained the frequency distribution of LRS. We quantified uncertainty around the LRS estimate by computing a 'mean' standard deviation which was obtained as an average, for a particular value of reproductive output, of all the individual LRS standard deviation for individuals having this specific number of offspring (see Appendix B for more details).

We provided vague prior distributions for all model parameters to induce little prior knowledge. Specifically, we chose uniform distributions on [0, 1] for the recapture and survival probabilities, and normal distributions with mean zero and variances 0.01 for the logit of the transition probabilities. We conducted a sensitivity analysis by modifying these priors using beta distributions with informative moments and posterior estimates of parameters were not changed.

Two MCMC chains with over-dispersed initial values were generated using 50,000 iterations with 10,000 burn-in iterations. Convergence was assessed using the Gelman and Rubin statistic, which compares the within- to the between-variability of chains started at different and dispersed initial values (Gelman, 1996). Mixing was checked visually by inspecting the chains. The simulations were performed using JAGS (Plummer, 2003) and the R package rjags was used to call JAGS, export results in R and calculate LRS. The code is available in Appendix B.

4.2.4. Results

Posterior mean estimates for all parameters of the model are provided in Table 2. There was strong heterogeneity in detection probability according to reproductive status, with breeder less prone to recapture than non-breeders. Survival decreased with increasing age, regardless of reproductive status.

Much of the density of the distribution of LRS falls at zero (Fig. 4): there were many individuals that did not successfully wean any fawns over their entire lifetime. We also observed a long tail in the LRS distribution, suggesting that some females had very high reproductive output.

5. Discussion

Multistate CR models are a valuable tool for estimating demographic parameters in the wild when detection probability is less than one. Their extension to hidden process models allows improved estimation of demographic parameters by accounting for variation in the state process (e.g., demographic stochasticity) while dealing with observation error. Using the same data that one would analyze with multistate models, SSM allows the estimation of demographic parameters as well as unknown states with imperfect detection (a feature we used to estimate LRS in Section 4.2). By optimizing the use of all available information including those data that could not be used in multistate analyses,

$$X_{i,t+1}|X_{i,t} \sim \text{multinomial}\left(1, X_{i,t}\begin{bmatrix} \phi_t^{NB}\psi_t^{NB\to NB} & \phi_t^{NB}\psi_t^{NB\to B1} & \phi_t^{NB}\psi_t^{NB\to B2} & \phi_t^{NB}\psi_t^{NB\to B3} & 1-\phi_t^{NB} \\ \phi_t^{B}\psi_t^{B1\to NB} & \phi_t^{B}\psi_t^{B1\to B1} & \phi_t^{B}\psi_t^{B1\to B2} & \phi_t^{B}\psi_t^{B1\to B3} & 1-\phi_t^{B} \\ \phi_t^{B}\psi_t^{B2\to NB} & \phi_t^{B}\psi_t^{B2\to B1} & \phi_t^{B}\psi_t^{B2\to B2} & \phi_t^{B}\psi_t^{B2\to B3} & 1-\phi_t^{B} \\ \phi_t^{B}\psi_t^{B3\to NB} & \phi_t^{B}\psi_t^{B3\to B1} & \phi_t^{B}\psi_t^{B3\to B2} & \phi_t^{B}\psi_t^{B3\to B3} & 1-\phi_t^{B} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}\right)$$

and

$$Y_{i,t}|X_{i,t} \sim \text{multinomial}\left(1, X_{i,t}\begin{bmatrix} 1-p_t^{NB} & p_t^{NB} & 0 & 0 & 0 \\ 1-p_t^{B} & p_t^{B} & 0 & 0 & 0 \\ 1-p_t^{B} & 0 & p_t^{B} & 0 & 0 \\ 1-p_t^{B} & 0 & 0 & p_t^{B} & 0 \\ 1-p_t^{B} & 0 & 0 & 0 & p_t^{B} \end{bmatrix}\right)$$

where the two equations above are the state and the observation equations, respectively.

Box I.

Table 2
Monte Carlo parameter numerical summaries (posterior means with standard deviations) from the state-space model fitted to the roe deer data. All parameters are probabilities. Note that the probability of remaining in the same reproductive state is not displayed but can be obtained as a complement.

Parameter	Posterior mean estimate (standard deviation)
Transition non-breeder to weaning 1 fawn $\psi^{NB\to B1}$	0.160 (0.033)
Transition non-breeder to weaning 2 fawns $\psi^{NB\to B2}$	0.149 (0.031)
Transition non-breeder to weaning 3 fawns $\psi^{NB\to B3}$	0.208 (0.036)
Transition weaning 1 fawn to non-breeder $\psi^{B1\to NB}$	0.274 (0.067)
Transition weaning 1 to 2 fawns $\psi^{B1\to B2}$	0.163 (0.068)
Transition weaning 1 to 3 fawns $\psi^{B1\to B3}$	0.316 (0.086)
Transition weaning 2 fawns to non-breeder $\psi^{B2\to NB}$	0.229 (0.059)
Transition weaning 2 to 1 fawns $\psi^{B2\to B1}$	0.159 (0.059)
Transition weaning 2 to 3 fawns $\psi^{B2\to B3}$	0.302 (0.082)
Transition weaning 3 fawns to non-breeder $\psi^{B3\to NB}$	0.261 (0.048)
Transition weaning 3 to 1 fawns $\psi^{B3\to B1}$	0.156 (0.045)
Transition weaning 3 to 2 fawns $\psi^{B3\to B2}$	0.224 (0.052)
Detection of breeders p^B	0.341 (0.032)
Detection of non-breeders p^{NB}	0.821 (0.110)
Survival of yearling non-breeders ϕ^{NB}(y)	0.941 (0.027)
Survival of prime-age non-breeders ϕ^{NB}(pa)	0.843 (0.031)
Survival of old non-breeders ϕ^{NB}(o)	0.648 (0.057)
Survival of prime-age non-breeders ϕ^{B}(pa)	0.934 (0.019)
Survival of old non-breeders ϕ^{NB}(o)	0.831 (0.030)

HMM allow us to cope with uncertainty in state assignment (Section 3.2).

Traditionally, in the CR literature, HMM are fit in a Frequentist framework while SSM are fitted in a Bayesian framework. This does not need to be the case (de Valpine, 2009), and examples of Frequentist analyses of SSMs are becoming available (Lele et al., 2007; de Valpine, 2011). We advocate pragmatism when making the decision of one framework over the other (Gimenez, 2008). This decision can be made based on one's programming skills or familiarity with software. However, there is a risk of using a sledgehammer to crack a nut by fitting unnecessarily complex models that come with difficulties (Bolker, 2009; Cressie et al., 2009) like non-identifiability for example (Gimenez et al., 2003; Gimenez et al., 2009; Luo et al., 2009; Rouan et al., 2009a). We encourage practitioners to think first of the biological question to be addressed, and resort to the Bayesian or Frequentist approach accordingly. For example, in the roe deer example (Section 4.2), we aimed at estimating LRS. We opted for SSM fit in a Bayesian framework using MCMC methods because it allowed the simultaneous estimation of unknown reproductive states and demographic parameters while accounting for parameter estimation error (Fig. 2). We could have used Frequentist alternatives to compute LRS (as shown by Rouan et al., 2009b).

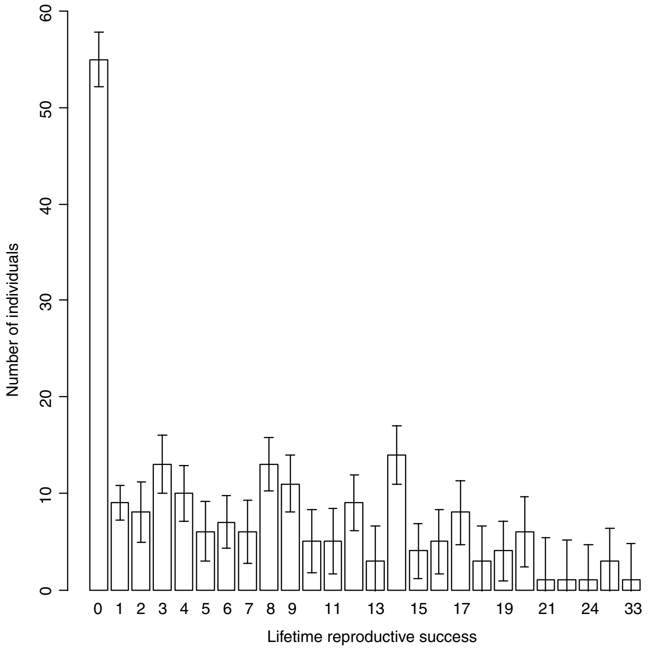

Fig. 4. Frequency distribution of lifetime reproductive success in roe deer. Error bars are +/−1 'mean' standard deviation where 'mean' standard deviation was obtained as an average, for a particular value of reproductive output, of all the individual LRS standard deviation for individuals having this specific number of weaned offspring (see Appendix B for more details).

In the roe deer data analysis (Section 4.2), we assumed that uncertainty in assessing the reproductive status was due to the issue of detection only: if a female was seen in the field, its reproductive status corresponded to the observation made. This was a fair assumption since fawns closely follow their mother during the period considered (Gaillard et al., 2000). However, if this assumption were to be violated, one could easily incorporate uncertainty in state assignment in the calculation of LRS by adapting the model used in the titis data analysis (Section 3.2).

We reviewed two types of hidden process models, both having a discrete space of states. Continuous hidden process models have never been used to our knowledge but could be useful in fitting an alternative to matrix models known as integral projection models developed by Easterling et al. (2000) for plant populations. In contrast with matrix models, which require discretizing individuals' states into classes, demographic parameters are modeled as continuous functions of individuals' states in integral projection models (e.g., size in plants or body mass in animals). Continuous hidden process models could be used as a unifying framework to combine the analysis of individual data, the construction of integral projection models and the projection of population fate by considering relevant states as being continuous.

The extension of hidden process models to the analysis of other sources of information than individual data also holds promise. In particular, Besbeas et al. (2002) developed an integrated framework to jointly analyze individual data using CR models and count data collected at the population level. So-called integrated population models use hidden process models to build the likelihood of count data (e.g., Schaub et al., 2007), while standard product-multinomials are used to form the likelihood of capture–recapture data. The formulation of integrated population models in a unifying framework using HMM or SSM for both count and individual data would certainly benefit from the flexibility of hidden process models.

Acknowledgments

The authors warmly thank Mathieu Buoro for his help with Fig. 3, Lauriane Rouan for discussion about LRS calculation using capture–recapture data, Paul Scofield and David Fletcher for sharing the titis data, Jean-Michel Marin and Christian Robert for a discussion about the distinction between HMM and SSM. We also thank the participants to the 2009 and 2010 Montpellier workshops on the analysis of capture–recapture data for their feedback on the titis analyses. This research was supported by a grant (reference: ANR-08-JCJC-0028-01) from the 'Jeunes Chercheuses et Jeunes Chercheurs' program of the French ANR to work on capture–recapture mixed models.

Appendix A

A.1. Instructions for fitting the HMM to titis data in program E-SURGE

To fit the model without uncertainty in E-SURGE, three states are considered {alive non-breeder, alive breeder, dead} = {NB, B, D}, and the events are three as well {non-detected, seen non-successful breeder, seen successful breeder} = {0, 1, 2}.

In GEPAT, declare two steps for the transitions (survival then conditional transitions).

Initial state: $\Pi = \begin{bmatrix} \pi & * & - \end{bmatrix}$.

Transition 1 (survival): $S = \begin{bmatrix} \phi & - & * \\ - & \phi & * \\ - & - & * \end{bmatrix}$.

Transition 2 (conditional transition): $T = \begin{bmatrix} * & \psi & - \\ \psi & * & - \\ - & - & * \end{bmatrix}$.

Event: $C = \begin{bmatrix} * & \beta & - \\ * & - & \beta \\ * & - & - \end{bmatrix}$.

In GEMACO, the syntax is:
For Initial State: IS $= i$
For Transition:
S (i.e. step 1) $= f$
T (i.e. step 2) $= f$
For Event: C = firste + nexte.f

To fit the model with uncertainty in E-SURGE, the same three states are required {alive non-breeder, alive breeder, dead} = {NB, B, D}, but the events are four {non-detected, seen and ascertained as non-breeder, seen and ascertained as breeder, not ascertained} = {0, 1, 2, 3}. In GEPAT, declare two steps for the transitions (survival then conditional transitions) and two steps for the events (capture then ascertainment).

Initial state: $\Pi = \begin{bmatrix} \pi & * & - \end{bmatrix}$.

Transition 1 (survival): $S = \begin{bmatrix} \phi & - & * \\ - & \phi & * \\ - & - & * \end{bmatrix}$.

Transition 2 (conditional transition): $T = \begin{bmatrix} * & \psi & - \\ \psi & * & - \\ - & - & * \end{bmatrix}$.

Event 1 (capture): $C = \begin{bmatrix} * & \beta & - \\ * & - & \beta \\ * & - & - \end{bmatrix}$.

Event 2 (ascertainment): $A = \begin{bmatrix} * & - & - & - \\ - & \delta & - & * \\ - & - & \delta & * \end{bmatrix}$.

In GEMACO, the syntax is:
For Initial State: IS $= i$
For Transition:
S (i.e. step 1) $= f$
T (i.e. step 2) $= f$
For Event:
C (i.e. step 1) = firste + nexte.f
A (i.e. step 2) $= f$

A.2. R Code for generating uncertainty in titis data

To artificially generate uncertainty on both the states non-breeder and breeder, we used the R script below to alter the raw capture–recapture data

```
# original data are stored in R matrix titi with individuals in rows
and years in columns
# 1 seen as non-breeder
# 2 seen as breeder
# 0 not seen

# nb of capture occasions
ny <- ncol(titi)
# nb of individuals
nind <- nrow(titi)

# the amended data with uncertainty are stored in R matrix titi2
titi2 <- titi
for (i in 1:nind)
{
for (j in 1:ny){
# 1 seen and ascertained Non-Breeder (with probability.2)
# 2 seen and ascertained Breeder (with probability.7)
# 3 seen but not ascertained (Non-Breeders with probability.8 +
Breeders with probability.3)
# 0 not seen

# Non-Breeders are ascertained with probability.2
if (titi[i, j] == 1)
{
temp <- rbinom(1,size=1,prob=.2)
if (temp == 1) titi2[i, j] <- 1 # if ascertained NB, event = 1
if (temp == 0) titi2[i, j] <- 3 # if not ascertained, event = 3
}
# Breeders are ascertained with probability.7 (event = 1),
# or not ascertained with probability.3 (event = 2)
if (titi[i, j] == 2)
{
temp <- rbinom(1,size=1,prob=.7)
if (temp == 1) titi2[i, j] <- 2 # if ascertained B, event = 2
if (temp == 0) titi2[i, j] <- 3 # if not ascertained, event = 3
}

}
}
```

Appendix B. Supplementary data

Supplementary material related to this article can be found online at doi:10.1016/j.tpb.2012.02.001.

References

Arnason, A.N., 1972. Parameter estimates from mark-recapture experiments on two populations subject to migration and death. Researches on Population Ecology 13, 97–113.

Arnason, A.N., 1973. The estimation of population size, migration rates and survival in a stratified population. Researches on Population Ecology 15, 1–8.

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.

Bolker, B., 2009. Learning hierarchical models: advice for the rest of us. Ecological Applications 19, 588–592.

Brommer, J.E., Gustafsson, L., Pietiäinen, H., Merilä, J., 2004. Single-generation estimates of fitness as proxies for long-term genetic contribution. American Naturalist 163, 505–517.

Brownie, C., Anderson, D.R., Burnham, K.P., Robson, D.S., 1985. Statistical Inference from Band Recovery Data—A Handbook. US Fish and Wildlife Service Editions.

Buckland, S.T., Newman, K.B., Fernandez, C., Thomas, L., Harwood, J., 2007. Embedding population dynamics models in inference. Statistical Science 22, 44–58.

Buckland, S.T., Newman, K.B., Thomas, L., Koesters, N., 2004. State-space models for the dynamics of wild animal populations. Ecological Modelling 171, 157–175.

Buoro, M., Prévost, E., Gimenez, O., 2010. Investigating evolutionary trade-offs in wild populations of atlantic salmon (salmo salar): incorporating detection probabilities and individual heterogeneity. Evolution 64, 2629–2642.

Burnham, K.P., 1993. A theory for combined analysis of ring recovery and recapture data. In: Lebreton, J.D., North, P.M. (Eds.), Marked Individual in the Study of Bird Population. Birkhäuser Verlag, Basel, Switzerland, pp. 199–213.

Burnham, K.P., Anderson, D.R., White, G.C., Brownie, C., Pollock, K.H., 1987. Design and analysis methods for fish survival experiments based on release–recapture. American Fisheries Society Monograph 5, 437.

Calvert, A., Bonner, S., Jonsen, I., Flemming, J., Walde, S., Taylor, P., 2009. A hierarchical Bayesian approach to multi-state mark-recapture: simulations and applications. Journal of Applied Ecology 46, 610–620.

Caswell, H., 2001. Matrix Population Models: Construction, Analysis and Interpretation, second ed. Sinauer Associates, Sunderland, Massachusetts.

Caswell, H., Fujiwara, M., 2004. Beyond survival estimation: mark-recapture, matrix population models, and population dynamics. Animal Biodiversity and Conservation 27 (1), 471–488.

Choquet, R., 2008. Automatic generation of multistate capture–recapture models. The Canadian Journal of Statistics 36, 1–15.

Choquet, R., Rouan, L., Pradel, R., 2009. Program E—SURGE: a software application for fitting multievent models. In: Thomson, D.L., Cooch, E.G., Conroy, M.J. (Eds.), Modeling Demographic Processes in Marked Populations. Springer, Berlin, Germany, pp. 845–865.

Clark, J.S., Ferraz, G., Oguge, N., Hays, H., Dicostanzo, J., 2005. Hierarchical bayes for structured, variable populations: from recapture data to life-history prediction. Ecology 86, 2232–2244.

Clobert, J., Lebreton, J.-D., 1987. About recent research in models for mark-recapture and mark-resighting data. A reply to C. Brownie. Biometrics 43, 1019–1022.

Clobert, J., Lebreton, J.-D., Allaine, D., Gaillard, J.M., 1994. The estimation of age-specific breeding probabilities from recaptures or resightings of marked animals. II longitudinal models. Biometrics 50, 375–387.

Clutton-Brock, T.H., 1988. Reproductive Success. Studies of Individual Variation in Contrasting Breeding Systems. University of Chicago Press, Chicago.

Conn, P.B., Cooch, E.G., 2009. Multistate capture–recapture analysis under imperfect state observation: an application to disease models. Journal of Applied Ecology 46, 486–492.

Cormack, R.M., 1964. Estimates of survival from the sighting of marked animals. Biometrika 51, 429–438.

Crespin, L., Choquet, R., Lima, M.A., Merritt, J.F., Pradel, R., 2008. Is heterogeneity of catchability in capture–recapture studies a mere sampling artifact or a biologically relevant feature of the population? Population Ecology 50, 247–256.

Cressie, N., Calder, C.A., Clark, J.S., Hoef, J.M.V., Wikle, C.K., 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modelling. Ecological Applications 19, 553–570.

Cubaynes, S., Pradel, R., Choquet, R., Duchamp, C., Gaillard, J.-M., Lebreton, J.-D., Marboutin, E., Miquel, C., Reboulet, A.-M., Poillot, C., Taberlet, P., Gimenez, O., 2010. Importance of accounting for detection heterogeneity when estimating abundance: the case of French wolves. Conservation Biology 24, 621–626.

de Valpine, , 2009. Shared challenges and common ground for Bayesian and classical analysis of hierarchical statistical models. Ecological Applications 19, 584–588.

de Valpine, P., 2011. Frequentist analysis of hierarchical models for population dynamics and demographic data. Journal of Ornithology doi:10.1007/s10336-010-0642-5.

Desprez, M., Pradel, R., Cam, E., Monnat, J.-Y., Gimenez, O., 2011. Now you see him, now you don't: experience, not age, is related to reproduction in Kittiwakes. Proceedings of the Royal Society B: Biological Sciences 278, 3060–3066.

Diekmann, O., Heesterbeek, J.A.P., 2000. Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation. Wiley, New York.

Dupuis, J.A., 1995. Bayesian estimation of movement and survival probabilities from capture–recapture data. Biometrika 82, 761–772.

Easterling, M.R., Ellner, S.P., Dixon, P.M., 2000. Size-specific sensitivity: applying a new structured population model. Ecology 81, 694–708.

Ebert, C., Knauer, F., Storch, I., Hohmann, U., 2010. Individual heterogeneity as a pitfall in population estimates based on non-invasive genetic sampling: a review and recommendations. Wildlife Biology 16, 225–240.

Gaillard, J.M., Delorme, D., Boutin, J.M., Van Laere, G., Boisaubert, B., Pradel, R., 1993. Roe deer survival patterns – A comparative analysis of contrasting populations. Journal of Animal Ecology 62, 778–791.

Gaillard, J.M., Andersen, R., Delorme, D., Linnell, J., 1998. Family effects on growth and survival of juvenile roe deer. Ecology 79, 2878–2889.

Gaillard, J.M., Festa-Bianchet, M., Delorme, D., Jorgenson, J., 2000. Body mass and individual fitness in female ungulates: bigger is not always better. Proceedings of the Royal Society of London, Series B 267, 471–477.

Gaillard, J.M., Sempéré, A.J., Boutin, J.M., Van Laere, G., Boisaubert, B., 1992. Effects of age and body weight on the proportion of females breeding in a population of roe deer (Capreolus capreolus). Canadian Journal of Zoology 70, 1541–1545.

Gamelon, M., Besnard, A., Gaillard, J.-M., Servanty, S., Baubet, E., Brandt, S., Gimenez, O., 2011. High hunting pressure selects for earlier birth data: wild boar as a case study. Evolution 65, 3100–3112.

Gelman, A., 1996. Inference and monitoring convergence. In: Gilks, W.R., Richardson, S., Spiegelhalter, D.J. (Eds.), Markov Chain Monte Carlo in Practice. Chapman and Hall, London, UK, pp. 131–143.

Ghimire, S., Gimenez, O., Pradel, R., McKey, D., Aumeeruddy-Thomas, Y., 2008. Demographic variation and population viability in a threatened Himalayan medicinal and aromatic herb (Nardostachys grandiflora): matrix modelling of harvesting effects in two contrasting habitats. Journal of Applied Ecology 45, 41–51.

Gimenez, O., 2008. Discussion: towards a Bayesian analysis template? The Canadian Journal of Statistics 36, 21–24.

Gimenez, O., Choquet, R., 2010. Incorporating individual heterogeneity in studies on marked animals using numerical integration: capture–recapture mixed models. Ecology 91, 951–957.

Gimenez, O., Charmantier, A., Gaillard, J.-M., A game of hide and seek: Estimating fitness in the wild (in preparation).

Gimenez, O., Choquet, R., Lebreton, J.-D., 2003. Parameter redundancy in multistate capture-recapture models. Biometrical Journal 45, 704–722.

Gimenez, O., Choquet, R., Lamor, L., Scofield, P., Fletcher, D., Lebreton, J.-D., Pradel, R., 2005. Efficient profile-likelihood confidence intervals for capture–recapture models. Journal of Agricultural, Biological, and Environmental Statistics 10, 184–196.

Gimenez, O., Covas, R., Brown, C.R., Anderson, M.D., Bomberger Brown, M., Lenormand, T., 2006. Nonparametric estimation of natural selection on a quantitative trait using mark-recapture data. Evolution 60, 460–466.

Gimenez, O., Morgan, B.J.T., Brooks, S.P., 2009. Weak identifiability in models for mark-recapture–recovery data. In: Thomson, D.L., Cooch, E.G., Conroy, M.J. (Eds.), Modeling Demographic Processes in Marked Populations. In: Environmental and Ecological Statistics, vol. 3. Springer, pp. 1055–1067.

Gimenez, O., Rossi, V., Choquet, R., Dehais, C., Doris, B., Varella, H., Vila, J.-P., Pradel, R., 2007. State-space modelling of data on marked individuals. Ecological Modelling 206, 431–438.

Gimenez, O., Viallefont, A., Charmantier, A., Pradel, R., Cam, E., Brown, C.R., Anderson, M.D., Bomberger Brown, M., Covas, R., Gaillard, J.-M., 2008. The risk of flawed inference in evolutionary studies when detectability is less than one. The American Naturalist 172, 441–448.

Hestbeck, J.B., Nichols, J.D., Malecki, R., 1991. Estimates of movement and site fidelity using mark-resight data of wintering Canada geese. Ecology 72, 523–533.

Jolly, G.M., 1965. Explicit estimates from capture–recapture data with both death and immigration-stochastic model. Biometrika 52, 225–247.

Juillet, C., Choquet, R., Gauthier, G., Pradel, R., 2011. A capture–recapture model with double-marking, live and dead encounters, and heterogeneity of reporting due to auxiliary mark loss. Journal of Agricultural, Biological, and Environmental Statistics 16, 88–104.

Kendall, W.L., Bjorkland, R., 2001. Using open robust design models to estimate temporary emigration from capture–recapture data. Biometrics 57, 1113–1122.

Kendall, W.L., Hines, J.E., Nichols, J.D., 2003. Adjusting multistate capture–recapture models for misclassification bias: manatee breeding proportions. Ecology 84, 1058–1066.

Kéry, M., Schaub, M., 2011. Bayesian Population Analysis Using WinBUGS—a Hierarchical Perspective. Academic Press.

King, R., Morgan, B.J.T., Gimenez, O., Brooks, S.P., 2009. Bayesian Analysis for Population Ecology. Chapman and Hall.

Lawless, J.F., McLeish, D.L., 1984. The information in aggregate data from Markov chains. Biometrika 71, 419–430.

Lebreton, J.-D., Almeras, T., Pradel, R., 1999. Competing events, mixtures of information and multistrata recapture models. Bird Study 46, 39–46.

Lebreton, J.-D., Burnham, K.P., Clobert, J., Anderson, D.R., 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs 62, 67–118.

Lebreton, J.-D., Hines, J.E., Pradel, R., Nichols, J.D., Spendelow, J.A., 2003. Estimation by capture–recapture of recruitment and dispersal over several sites. Oikos 101, 253–264.

Lebreton, J.-D., Nichols, J.D., Barker, R.J., Pradel, R., Spendelow, J.A., 2009. Modeling individual animal histories with multistate capture–recapture models. In: Caswell, H. (Ed.), Advances in Ecological Research, vol. 41. Academic Press, Burlington, MA, pp. 87–173.

Lele, S.R., Dennis, B., Lutscher, F., 2007. Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. Ecology Letters 10, 551–563.

Link, W.A., Barker, R.J., 2005. Modeling association among demographic parameters in open population capture–recapture data. Biometrics 61, 46–54.

Link, W.A., Barker, R.J., 2010. Bayesian Inference with Ecological Applications. Academic Press.

Luo, Y.Q., Wend, E.S., Wu, X.W., Gao, C., Zhou, X.H., Zhang, L., 2009. Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models. Ecological Applications 19, 571–574.

Marzolin, G., Charmantier, A., Gimenez, O., 2011. Frailty in state-space models: application to actuarial senescence in the dipper. Ecology 92, 562–567.

McCarthy, M.A., 2007. Bayesian Methods for Ecology. Cambridge University Press, Cambridge, UK.

McLoughlin, P.D., Gaillard, J.-M., Boyce, M.S., Bonenfant, C., Messier, F., Duncan, P., Delorme, D., Van Moorter, B., Said, S., Klein, F., 2007. Lifetime reproductive success and composition of the home range in a large herbivore. Ecology 88, 3192–3201.

Newman, K.B., Buckland, S.T., Lindley, S.T., Thomas, L., Fernandez, C., 2006. Hidden process models for animal population dynamics. Ecological Applications 16, 74–86.

Newman, K.B., Fernández, C., Buckland, S.T., Thomas, L., 2009. Monte Carlo inference for state-space models of wild animal populations. Biometrics 65, 572–583.

Newton, I. (Ed.), 1989. Lifetime Reproduction in Birds. Academic Press, London.

Nichols, J.D., 1992. Capture–recapture models. Bioscience 42, 94–102.

Nichols, J.D., Kendall, W.L., Hines, J.E., Spendelow, J.A., 2004. Estimation of sex-specific survival from capture–recapture data when sex is not always known. Ecology 85, 3192–3201.

O'Hara, R., Lampila, S., Orell, M., 2009. Estimation of rates of births, deaths, and immigration from mark-recapture data. Biometrics 65, 275–281.

Oliver, L.J., Morgan, B.J.T., Durant, S.M., Pettorelli, N., 2010. Individual heterogeneity in recapture probability and survival estimates in cheetah. Ecological Modelling 222, 776–784.

Papadatou, E., Pradel, R., Schaub, M., Dolch, D., Geiger, H., Ibañez, C., Kerth, G., Popa-Lisseanu, A., Schorcht, W., Teubner, J., Gimenez, O., 2011. Comparing survival among species with imperfect detection using multilevel analysis of mark-recapture data: a case study on bats. Ecography. doi:10.1111/j.1600-0587.2011.07084.x.

Papaïx, J., Cubaynes, S., Buoro, M., Charmantier, A., Perret, P., Gimenez, O., 2010. Combining capture–recapture data and pedigree information to assess heritability of demographic parameters in the wild. Journal of Evolutionary Biology 23, 2176–2184.

Péron, G., Crochet, P.A., Choquet, R., Pradel, R., Lebreton, J.-D., Gimenez, O., 2010a. Capture–recapture models with heterogeneity to study survival senescence in the wild. Oikos 119, 524–532.

Péron, G., Ferrand, Y., Gossmann, F., Bastat, C., Guenezan, M., Gimenez, O., 2011. Nonparametric spatial regression of survival probability: visualization of population sinks in Eurasian Woodcock. Ecology. doi:10.1890/10-2224.1.

Péron, G., Lebreton, J.D., Crochet, P.A., 2010b. Breeding dispersal in black-headed gull: the value of familiarity in a contrasted environment. Journal of Animal Ecology 79, 317–326.

Plummer, M., 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd International Workshop on Distributed Statistical Computing, DSC 2003, Vienna, Austria, March 20–22. ISSN: 1609-395X.

Pradel, R., 2005. Multievent: an extension of multistate capture–recapture models to uncertain states. Biometrics 61, 442–447.

Pradel, R., 2009. The stakes of capture–recapture models with state uncertainty. In: Thomson, D.L., Cooch, E.G., Conroy, M.J. (Eds.), Modeling Demographic Processes in Marked Populations. Springer, Berlin, Germany, pp. 781–795.

Pradel, R., Choquet, R., Béchet, A., 2011. Breeding experience might be a major determinant of breeding probability in long-lived species: the case of the greater flamingo. http://hal.archives-ouvertes.fr/hal-00655507/.

Pradel, R., Choquet, R., Lima, M.A., Merritt, J.F., Crespin, L., 2010. Estimating population growth rate from capture–recapture data in presence of capture heterogeneity. Journal of Agricultural, Biological, and Environmental Statistics 15, 248–258.

Pradel, R., Lebreton, J.-D., 1999. Comparison of different approaches to study the local recruitment of breeders. Bird Study 46, 74–81.

R Development Core Team, 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN: 3-900051-07-0. URL: http://www.R-project.org/.

Rouan, L., Choquet, R., Pradel, R., 2009a. A general framework for modeling memory in capture–recapture data. Journal of Agricultural, Biological, and Environmental Statistics 14, 338–355.

Rouan, L., Gaillard, J.-M., Guédon, Y., Pradel, R., 2009b. Estimation of lifetime reproductive success when reproductive status cannot always be assessed. In: Thomson, D.L., Cooch, E.G., Conroy, M.J. (Eds.), Modeling Demographic Processes in Marked Populations. Springer, Berlin, Germany, pp. 867–879.

Royle, J.A., 2008. Modeling individual effects in the Cormack–Jolly–Seber model: a state-space formulation. Biometrics 64, 364–370.

Royle, J.A., Dorazio, R.M., 2008. Hierarchical Modelling and Inference in Ecology. Academic Press.

Sanz-Aguilar, A., Tavecchia, G., Genovart, M., Igual, J.M., Oro, D., Rouan, L., Pradel, R., 2011. Studying the reproductive skipping behavior in long-lived birds by adding nest-inspection to individual-based data. Ecological Applications 21, 555–564.

Schaub, M., Gimenez, O., Sierro, A., Arlettaz, R., 2007. Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. Conservation Biology 21, 945–955.

Schofield, M.R., Barker, R.J., 2008. A unified capture–recapture framework. Journal of Agricultural, Biological, and Environmental Statistics 13, 458–477.

Schofield, M.R., Barker, R.J., 2011. Full open population capture–recapture models with individual covariates. Journal of Agricultural, Biological, and Environmental Statistics 16, 253–268.

Scofield, P., Fletcher, D., Robertson, C., 2001. Titi (sooty shearwaters) on Whero Island: an analysis of historic banding data using modern capture–recapture techniques. Journal of Agricultural, Biological, and Environmental Statistics 6, 268–280.

Seber, G.A.F., 1965. A note on the multiple-recapture census. Biometrika 52, 249–259.

Servanty, S., Choquet, R., Baubet, E., Brandt, S., Gaillard, J.-M., Schaub, M., Toïgo, C., Lebreton, J.-D., Buoro, M., Gimenez, O., 2010. Assessing whether mortality is additive using marked animals: a Bayesian state-space modeling approach. Ecology 91, 1916–1923.

Servanty, S., Gaillard, J.-M., Ronchi, F., Focardi, S., Baubet, E., Gimenez, O., 2011. Influence of harvesting pressure on demographic tactics: implications for wildlife management. Journal of Applied Ecology 48, 835–843.

Tavecchia, G., Adrover, J., Muñoz Navarro, A., Pradel, R., 2012. Modelling mortality causes in longitudinal data in the presence of tag loss: application to raptor poisoning and electrocution. Journal of Applied Ecology 49, 297–305.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., Stumpf, M.P.H., 2009. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. Journal of the Royal Society Interface 6, 187–202.

Tuljapurkar, S., Caswell, H., 1997. Structured-Population Models in Marine, Terrestrial, and Freshwater Systems. Chapman and Hall, New York.

Van Tienderen, P.H., 2000. Elasticities and the link between demographic and evolutionary dynamics. Ecology 81, 666–679.

Véran, S., Gimenez, O., Flint, E., Kendall, W.L., Doherty, P.F., Lebreton, J.-D., 2007. Quantifying the impact of longline fisheries on adult survival in the black-footed albatross. Journal of Applied Ecology 44, 942–952.

Williams, B.K., Nichols, J.D., Conroy, M.J., 2002. Analysis and Management of Animal Populations. Academic Press, San Diego.

Glossary

Capture–recapture (CR): A generic term embedding several protocols for monitoring (mainly) plants and animals in the wild using marks to uniquely identify individuals. The analysis of these data allows the estimation of, e.g., survival, dispersal and abundance while explicitly accommodating for detection probability less than one.

Cormack–Jolly–Seber model (CJS): A capture–recapture model that is used for open populations to estimate time-dependent survival and detection probabilities. The Cormack–Jolly–Seber model is easily formulated as a hidden process model by considering two states (alive or dead) and two observations or events (detected or not). The dynamic of states is governed by a Markov model parameterized with survival probabilities while the observations are made from these states according to detection probabilities.

Hidden Markov model (HMM): A particular case of state-space model in which the states are Markovian, i.e. the next state depends only on the current state and not on the sequence of states that occurred before.

Hidden process model: A generic term referring to either a state-space model or a hidden Markov model.

Multievent model: A particular type of hidden Markov model. Multievent models correspond to extensions of multistate models that allow accommodating uncertainty in state assignment.

Multistate model: An extension of the Cormack–Jolly–Seber model that allows estimation of transition between states. Examples of such states are geographical sites, behavioral or physiological conditions like, e.g., breeding vs. non-breeding or healthy vs. diseased.

State-space model (SSM): A model that runs two time series in parallel, one captures the dynamic of the true states (latent) and the other consists of observations that are made from these underlying but possibly unknown states.