



STATE-SPACE MODELLING OF DATA ON MARKED INDIVIDUALS

Olivier Gimenez, Vivien Rossi, Rémi Choquet, Camille Dehais, Blaise Doris,
Hubert Varella, Jean-Pierre Vila, Roger Pradel

► To cite this version:

Olivier Gimenez, Vivien Rossi, Rémi Choquet, Camille Dehais, Blaise Doris, et al.. STATE-SPACE MODELLING OF DATA ON MARKED INDIVIDUALS. Ecological Modelling, 2007, 206 (3-4), pp.431-438. 10.1016/j.ecolmodel.2007.03.040 . hal-03515273

HAL Id: hal-03515273

<https://hal.science/hal-03515273>

Submitted on 6 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

STATE-SPACE MODELLING OF DATA ON MARKED INDIVIDUALS

Olivier GIMENEZ^{1,2,6}, Vivien ROSSI^{3,4,5}, Rémi CHOQUET¹, Camille DEHAIS¹,
Blaise DORIS^{1,3}, Hubert VARELLA^{1,3}, Jean-Pierre VILA³ and Roger PRADEL¹

¹*Centre d'Ecologie Fonctionnelle et Evolutive/CNRS - UMR 5175*

1919 Route de Mende

34293 Montpellier - FRANCE

²*Centre for Research into Ecological and Environmental Modelling*

University of St Andrews, St Andrews

The Observatory, Buchanan Gardens, KY16 9LZ - SCOTLAND

³*Laboratoire d'Analyse des Systèmes et Biométrie - UMR 729*

INRA/ENSAM

2 Place Pierre Viala, 34060 Montpellier - FRANCE

⁴*Institut de Modélisation Mathématiques de Montpellier - UMR 5149*

Université Montpellier 2

CC051, Place Eugène Bataillon, 34095 Montpellier Cedex 5 - FRANCE

⁵*Laboratoire de Biostatistique*

Institut Universitaire de Recherche Clinique

641 Avenue du Doyen Gaston Giraud, 34093 Montpellier Cedex 5 - FRANCE

⁶*E-mail: olivier.gimenez@cefe.cnrs.fr*

Abstract

State-space models have recently been proposed as a convenient and flexible framework for specifying stochastic models for the dynamics of wild animal populations. Here we focus on the modelling of data on marked individuals which is

frequently used in order to estimate demographic parameters while accounting for imperfect detectability. We show how usual models to deal with capture-recapture and ring-recovery data can be fruitfully written as state-space models. An illustration is given using real data and a Bayesian approach using MCMC methods is implemented to estimate the parameters. We give important future developments facilitated by the SSM formulation. Overall, the general state-space modelling framework is shown to have a great potential in population ecology modelling.

1 Introduction

The estimation of animal survival is essential in population biology to investigate population dynamics, with important applications in the understanding of ecological, evolutionary, conservation and management issues for wild populations (Pollock, 1991; Williams et al., 2002). While the time to event is known in medical, social or engineering sciences (death, marriage and failure respectively), models for estimating wild animal survival must incorporate nuisance parameters to account for incomplete detectability in monitoring individuals (Schwarz et Seber, 1999). Typically, individuals are captured, marked and can be resighted or recaptured (encountered thereafter) to construct encounter histories which consist of sequences of 1's and 0's according to whether a detection occurs or not. The likelihood for such data arises from products of multinomial distributions whose cell probabilities are complex functions of survival probabilities - parameters of primary interest - and encounter probabilities - nuisance parameters (Cormack, 1964; Jolly, 1965; Seber, 1965 - CJS thereafter).

In this note, we show how the population process can be fruitfully disentangled, by distinguishing the underlying demographic process, i.e. the survival (as well as transitions between sites/states if needed), from its observation, i.e. the detectability.

1 This leads us to consider a natural formulation for capture-recapture models using state-
2 space models (SSMs). Our contribution is in line with a recent paper by Buckland et
3 al. (2004) who have proposed to adopt SSMs as a convenient and flexible framework
4 for specifying stochastic models for the dynamics of wild animal populations.

5 Thus far, SSMs have been mainly used for time series of animal counts (de Valpine,
6 2004; Millar and Meyer, 2000) or animal locations (Anderson-Sprecher and Ledolter,
7 1991) to allow true but unobservable states (the population size or trajectory) to be
8 inferred from observed but noisy data (see Clark et al., 2005 and Wang, 2006 for
9 reviews). The novelty of our approach lies in the use of SSMs to fit capture-recapture
10 models to encounter histories.

11 In Section 2, we discuss how to express the CJS model under the form of a SSM. The
12 implementation details are provided, and real data are presented to compare parameter
13 estimates as obtained using the standard product-multinomial and the SSM approaches.
14 In Section 3, the flexibility of the state-space modeling approach is demonstrated by
15 considering two widely used alternative schemes for collecting data on marked animals.
16 Finally, Section 4 discusses important developments of capture-recapture models facil-
17 itated by the SSM formulation. We emphasize that this general framework has a great
18 potential in population ecology modelling.

19 **2 State-space modelling of capture-recapture data**

20 We focus here on the CJS model for estimating animal survival based on capture-
21 recapture data, as this model is widely used in the ecological and evolutionary literature
22 (e.g. Lebreton et al., 1992).

1 2.1 Likelihood

2 We first define the observations and then the states of the system. We assume that
 3 n individuals are involved in the study with T encounter occasions. Let $X_{i,t}$ be the
 4 binary random variable taking values 1 if individual i is alive at time t and 0 if it is
 5 dead at time t . Let $Y_{i,t}$ be the binary random variable taking values 1 if individual i is
 6 encountered at time t and 0 otherwise. Note that we consider the encounter event as
 7 being physically captured or barely observed. The parameters involved in the likelihood
 8 are $\phi_{i,t}$, the probability that an animal i survives to time $t + 1$ given that it is alive
 9 at time t ($t = 1, \dots, T - 1$), and $p_{i,t}$ the probability of detecting individual i at time t
 10 ($t = 2, \dots, T$). Let finally e_i be the occasion where individual i is encountered for the
 11 first time. A general state-space formulation of the CJS model is therefore given by:

$$12 \quad Y_{i,t}|X_{i,t} \sim \text{Bernoulli}(X_{i,t}p_{i,t}), \quad t > e_i \quad (1)$$

$$13 \quad X_{i,t+1}|X_{i,t} \sim \text{Bernoulli}(X_{i,t}\phi_{i,t}), \quad t \geq e_i \quad (2)$$

14 where Equation (1) and Equation (2) are the observation and the state equations re-
 15 spectively. This formulation naturally separates the nuisance parameters (the encounter
 16 probabilities) from the parameters of actual interest i.e. the survival probabilities, the
 17 latter being involved exclusively in the state Equation (2). Such a clear distinction be-
 18 tween a demographic process and its observation makes the description of a biological
 19 dynamic system much simpler and allows complex models to be fitted (Pradel, 2005;
 20 Clark et al., 2005). We will refer to this formulation as the individual state-space CJS
 21 model (individual SSM CJS hereafter). The rationale behind the above formulation is
 22 as follows. We give the full details for the observation Equation (1) only, as a similar
 23 reasoning easily leads to Equation (2). If individual i is alive at time t , then it has prob-

1 ability $p_{i,t}$ of being encountered and probability $1 - p_{i,t}$ otherwise, which translates into
 2 $Y_{i,t}$ is distributed as *Bernoulli*($p_{i,t}$) given $X_{i,t} = 1$. Now if individual i is dead at time t ,
 3 then it cannot be encountered, which translates into $Y_{i,t}$ is distributed as *Bernoulli*(0)
 4 given $X_{i,t} = 0$. Putting together those two pieces of reasoning, the distribution of the
 5 observation $Y_{i,t}$ conditional on the state $X_{i,t}$ is given by Equation (1).

6 Statistical inference then requires the likelihood of the state-space model specified
 7 above. Assuming independence of individuals, the likelihood is given by the product
 8 of all individual likelihood components. The likelihood component for individual i is
 9 the probability of the vector of observations $\mathbf{Y}_i^T = (Y_{i,e_i}, \dots, Y_{i,T})$ which gathers the
 10 information set up to time T for this particular individual. The likelihood component
 11 corresponding to individual i is therefore given by (e.g. Harvey, 1989)

$$12 \quad \int_{X_{i,0}}, \dots, \int_{X_{i,T}} \left\{ \prod_{t=e_i+1}^T [Y_{i,t}|X_{i,t}][X_{i,t}|X_{i,t-1}] dX_{i,t} \right\} [X_{i,e_i}] dX_{i,e_i} \quad (3)$$

13 where $[X]$ denotes the distribution of X and X_{i,e_i} the initial state of individual i which
 14 is assumed to be alive. Because we deal with binary random vectors, we used the
 15 counting measure instead of the Lebesgue measure.

16 In its original formulation, the CJS makes important assumptions regarding indi-
 17 viduals. First, all individuals share the same parameters, which means that the survival
 18 and detection probabilities depend on the time index only. In mathematical notation,
 19 we have $\phi_{it} = \phi_t$ and $p_{it} = p_t$ for all $i = 1, \dots, n$, so Equation (1) and Equation (2)
 20 become $X_{i,t+1}|X_{i,t} \sim \text{Bernoulli}(X_{i,t}\phi_t)$ and $Y_{i,t}|X_{i,t} \sim \text{Bernoulli}(X_{i,t}p_t)$ respectively.
 21 Second, the CJS model also assumes independence between individuals. By using sim-
 22 ple relationships between Bernoulli and Binomial distributions, one can finally fruitfully

1 formulates the original CJS model under the following state-space model:

$$2 \quad Y_t | X_t^- \sim \text{Bin}(X_t^-, p_t) \quad (4)$$

$$3 \quad X_{t+1}^- | X_t^+ \sim \text{Bin}(X_t^+, \phi_t) \quad (5)$$

4 where X_t^+ is the number of survivors of marked individuals from time $t-1$, X_t^- , plus the
5 newly marked individuals at time t , and Y_t is the total number of marked individuals en-
6 countered at time t . We will refer to this formulation as the population state-space CJS
7 model (population SSM CJS hereafter). Interestingly, specifying the system under a
8 state-space formulation now requires much less equations than the individual SSM CJS
9 model, which may avoid the computational burden. Nevertheless, while the individual
10 SSM CJS involves parameters for every single individual and sampling occasion, the
11 population SSM CJS model makes the strong assumptions that all individuals behave
12 the same as well as independently, which may be of little relevance from the biological
13 point of view. To cope with this issue, in-between modelling can be achieved by consid-
14 ering age effects or groups classes in the population SSM model (Lebreton et al., 1992).
15 Finally, covariates can be incorporated in order to assess the effect of environment such
16 as climate change, most conveniently by writing the relationship between the target
17 probabilities and the predictors on the logit scale (Pollock, 2002).

18 **2.2 Implementation**

19 Fitting capture-recapture models in a state-space framework is complicated due to
20 the high-dimensional integral involved in the individual likelihood Equation (3). To
21 circumvent this issue, several techniques have been proposed including Kalman filtering,
22 Monte-Carlo particle filtering (such as sequential importance sampling) and MCMC (see

Clark et al., 2005 and Wang, 2006 for reviews). Our objective here is not to discuss these different methods. For our purpose, we adopt the MCMC technique which is now widely used in biology (Ellison, 2004; Clark, 2005), in particular for estimating animal survival (Seber et Schwarz, 1999; Brooks et al., 2000). Besides, this is to our knowledge the only methodology which comes with an efficient and flexible program to implement it, which, in our case, will allow biologists to easily and rapidly adopt our approach.

2.3 Illustration

We consider capture-recapture data on the European dipper (*Cinclus cinclus*) that were collected between 1981 and 1987 (Lebreton et al., 1992). The data consists of marking and recaptures of 294 birds ringed as adults in eastern France. We applied standard maximum-likelihood estimation (Lebreton et al. 1992) and MCMC techniques (Brooks et al. 2000) using the product-multinomial likelihood and the state-space likelihood of Equation (3) in combination with Equation (1) and Equation (2). We ran two overdispersed parallel MCMC chains to check whether convergence was reached (Gelman, 1996). We used 10,000 iterations with 5,000 burned iterations for posterior summarization. We used uniform flat priors for both survival and detection probabilities. The simulations were performed using WinBUGS (Spiegelhalter et al., 2003). The R (Ihaka and Gentleman, 1996) package R2WinBUGS (Sturtz et al., 2005) was used to call WinBUGS and export results in R. This was especially helpful when converting the raw encounter histories into the appropriate format, generating initial values and calculate posterior modes. The programs are available in Appendix A. Posterior summaries for encounter and survival probabilities are given in Table 1, along with their posterior probability distributions that are displayed in Figure 1.

[Table 1 about here.]

[Figure 1 about here.]

Survival estimates were uniformly similar whatever the method used (Table 1). In particular, there is a clear decrease in survival 1982-1983 and 1983-1984, corresponding to a major flood during the breeding season in 1983 (Lebreton et al., 1992).

In contrast, posterior medians of detection probabilities using the CJS SSM approach are quite different from the classical maximum likelihood estimates, but more similar to the posterior medians obtained with the product-multinomial likelihood approach (Table 1). These discrepancies are no longer present when posterior modes are examined, as expected (recall that we use non-informative uniform distributions as priors for all parameters).

The last survival probability as well as the last detection probability are estimated with high variability (Table 1 and Figure 1). The fact that these two parameters cannot be separately estimated is not surprising since the CJS model is known to be parameter-redundant (Catchpole and Morgan, 1997). Also, the first survival probability and the first detection probability are weakly identifiable, due to the fact that very few individuals were marked at the first sampling occasion (approximately 7% of the full data set).

3 Further state-space modelling

3.1 Multistate capture-recapture models

Multistate capture-recapture models (Arnason, 1973; Schwarz et al., 1993; AS hereafter) are a natural generalization of the CJS model in that individuals can move between states, according to probabilities of transition between those states. States can be either geographical sites or states of categorical variables like reproductive status or size class

1 (Lebreton and Pradel, 2002). We provide here a state-space modelling formulation
 2 of the AS model (Dupuis, 1995; Newman, 1998; Clark et al., 2005). Without loss of
 3 generality, we consider 2 states. Let $X_{i,t}$ be the random state vector taking values
 4 $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ if, at time t , individual i is alive in state 1, 2 or dead
 5 respectively. Let $Y_{i,t}$ be the random observation vector taking values $(1, 0, 0)$, $(0, 1, 0)$
 6 and $(0, 0, 1)$ if, at time t , individual i is encountered in state 1, 2 or not encountered.
 7 Parameters involved in the modelling include $\phi_{i,t}^{rs}$, the probability that an animal i
 8 survives to time $t + 1$ given that it is alive at time t ($t = 1, \dots, T - 1$) and makes the
 9 transition between state r and state s over the same interval ($r, s = 1, 2$), as well as $p_{i,t}^r$
 10 the probability of detecting individual i at time t in state r ($t = 2, \dots, T$, $r = 1, 2$). A
 11 state-space formulation for the multistate AS model is given by:

$$Y_{i,t}|X_{i,t} \sim \text{Multinomial} \left(1, X_{i,t} \begin{bmatrix} p_{i,t}^1 & 0 & 1 - p_{i,t}^1 \\ 0 & p_{i,t}^2 & 1 - p_{i,t}^2 \\ 0 & 0 & 1 \end{bmatrix} \right) \quad (6)$$

$$X_{i,t+1}|X_{i,t} \sim \text{Multinomial} \left(1, X_{i,t} \begin{bmatrix} \phi_{i,t}^{11} & \phi_{i,t}^{12} & 1 - \phi_{i,t}^{11} - \phi_{i,t}^{12} \\ \phi_{i,t}^{21} & \phi_{i,t}^{22} & 1 - \phi_{i,t}^{21} - \phi_{i,t}^{22} \\ 0 & 0 & 1 \end{bmatrix} \right) \quad (7)$$

14 where Equation (6) and Equation (7) are the observation and the state equations
 15 respectively. This formulation has similarities with that of Pradel (2005) who used
 16 hidden-Markov models to extend multistate models to cope with uncertainty in state
 17 assignment. Again, it should be noted that the state-space formulation allows the de-
 18 mographic parameters to be distinguished from the nuisance parameters. A similar
 19 reasoning to that adopted for the CJS model leads to Equations (6) and (7). As ex-
 20 pected, Equation (6) and Equation (7) reduce to Equation (1) and Equation (2) if one

1 considers a single state. Making similar assumptions as in the CJS model leads to the
 2 population AS SSM.

3 **3.2 Ring-recovery models**

4 The capture-recapture models presented above deals with apparent survival, which
 5 turns out to be true survival if emigration is negligible. When marks of individuals
 6 (or individuals themselves) are actually recovered, true survival probabilities can be
 7 estimated using ring-recovery models (Brownie et al., 1985; RR models hereafter). Let
 8 $X_{i,t}$ be the binary random variable taking values 1 if individual i is alive at time t and
 9 0 if it is dead at time t . Let $Y_{i,t}$ be the binary random variable taking values 1 if mark
 10 of individual i is recovered at time t and 0 otherwise. The parameters involved in the
 11 likelihood are $\phi_{i,t}$, the probability that an animal i survives to time $t + 1$ given that
 12 it is alive at time t ($t = 1, \dots, T - 1$), and $\lambda_{i,t}$ the probability of recovering the mark
 13 of individual i at time t ($t = 2, \dots, T$). A general state-space formulation of the RR
 14 model is therefore given by:

$$15 \quad Y_{i,t} | X_{i,t}, X_{i,t-1} \sim \text{Bernoulli}((X_{i,t-1} - X_{i,t})\lambda_{i,t}) \quad (8)$$

$$16 \quad X_{i,t+1} | X_{i,t} \sim \text{Bernoulli}(X_{i,t}\phi_{i,t}) \quad (9)$$

17 where Equation (8) and Equation (9) are the observation and the state equations re-
 18 spectively. While the state Equation (9) is the same as that in the individual SSM
 19 CJS, the observation Equation (8) deserves further explanation. If individual i , alive
 20 at time t , does not survive to time $t + 1$, then its mark has probability $\lambda_{i,t}$ of being
 21 recovered and probability $1 - \lambda_{i,t}$ otherwise, which translates into $Y_{i,t}$ is distributed as
 22 $\text{Bernoulli}(\lambda_{i,t})$ given $X_{i,t-1} = 1$ and $X_{i,t} = 0$ i.e. $X_{i,t-1} - X_{i,t} = 1$. Now if individual i

1 is in a given state (dead or alive) at time t and remains in this state till time $t + 1$, then
 2 its mark cannot be recovered, which translates into $Y_{i,t}$ is distributed as *Bernoulli*(0)
 3 given $X_{i,t-1} = 0$ and $X_{i,t} = 0$ or $X_{i,t-1} = 1$ and $X_{i,t} = 1$ i.e. $X_{i,t-1} - X_{i,t} = 0$. The dis-
 4 tribution of the observation $Y_{i,t}$ conditional on the combination of states $X_{i,t-1} - X_{i,t}$ is
 5 thus given by Equation (8). Similar comments to that of previous sections can be made
 6 here as well. Finally, we note that because the probability distribution of the current
 7 observation does not only depend on the current state variable, the model defined by
 8 Equation (8) and Equation (9) does not exactly matches the definition of a state-space
 9 model but can be rewritten as such (see Appendix B).

10 4 Discussion

11 We have shown that, by separating the demographic process from its observation, CR
 12 models for estimating survival can be expressed as SSMs. In particular, the SSM
 13 formulation of the CJS model competes well with the standard method when applied to
 14 a real data set. Bearing this in mind, we see at least two further promising developments
 15 to our approach.

16 First, by scaling down from the population to the individual level while modelling
 17 the survival probabilities, random effects can readily be incorporated to cope with
 18 heterogeneity in the detection probability (Huggins, 2001) and deal with a frailty in
 19 the survival probability (Vaupel and Yashin, 1985). Second, the combination of various
 20 sources of information which has recently received a growing interest, (e.g. recovery
 21 and recapture data, Catchpole et al., 1998; recovery and census data, Besbeas et al.,
 22 2002; Besbeas et al., 2003) can now be operated/conducted in a unique SSM framework
 23 and hence benefits from the corpus of related methods. Of particular importance, we

1 are currently investigating the robust detection of density-dependence by accounting
2 for error in the measurement of population size using the combination of census data
3 and data on marked individuals.

4 Because most often, data collected in population dynamics studies give only a noisy
5 output of the demographic process under investigation, the SSM framework provides
6 a flexible and integrated framework for fitting a wide range of models which, with
7 widespread adoption, has the potential to advance significantly ecological statistics
8 (Buckland et al., 2004; Thomas et al., 2005).

1 5 Acknowledgments

2 O. Gimenez's research was supported by a Marie-Curie Intra-European Fellowship
3 within the Sixth European Community Framework Programme. This project was
4 funded by the Action Incitative Régionale BioSTIC-LR 'Modélisation intégrée en dy-
5 namique des populations : applications à la gestion et à la conservation'.

6 6 Literature Cited

- 7 Anderson-Sprecher, R., Ledolter, J., 1991. State-space analysis of wildlife telemetry
8 data. *Journal of the American Statistical Association.* 86, 596–602.
- 9 Arnason, A.N., 1973. The estimation of population size, migration rates and survival
10 in a stratified population. *Research in Population Ecology.* 15, 1–8.
- 11 Besbeas, P., Freeman, S.N., Morgan, B.J.T., Catchpole, E.A., 2002. Integrating mark-
12 recapture-recovery and census data to estimate animal abundance and demographic
13 parameters. *Biometrics.* 58, 540–547.
- 14 Besbeas, P., Lebreton, J.D., Morgan, B.J.T., 2003. The efficient integration of abun-
15 dance and demographic data. *Journal of the Royal Statistical Society Series C - Applied*
16 *Statistics.* 52, 95–102.
- 17 Brooks, S. P., Catchpole, E. A., Morgan, B. J. T., 2000. Bayesian animal survival
18 estimation. *Statistical Science.* 15, 357–376.
- 19 Brownie, C., Anderson, D.R., Burnham, K.P., Robson, D.S., 1985. Statistical infer-
20 ence from band recovery data - a handbook. United States Fish and Wildlife Service,
21 Washington.
- 22 Buckland, S.T., Newman, K.B., Thomas, L., Koesters, N.B., 2004. State-space models
23 for the dynamics of wild animal populations. *Ecological Modelling.* 171, 157–175.

1 Catchpole, E.A., Freeman, S.N., Morgan, B.J.T., Harris, M.P., 1998. Integrated recovery/recapture data analysis. *Biometrics*. 54, 33–46.

3 Choquet, R., Reboulet, A.-M., Pradel, R., Gimenez, O., Lebreton, J.-D., 2004. M-SURGE: new software specifically designed for multistate capture-recapture models. *Animal Biodiversity and Conservation*. 27, 207–215.

6 Clark, J.S., 2005. Why environmental scientists are becoming Bayesians. *Ecology Letters*. 8, 2–14.

8 Clark, J.S., Ferraz, G.A., Oguge, N., Hays, H., DiCostanzo, J., 2005. Hierarchical Bayes for structured, variable populations: From recapture data to life-history prediction. *Ecology*. 86, 2232–2244.

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

13 De Valpine, P., 2004. Monte Carlo state-space likelihoods by weighted posterior kernel density estimation. *Journal of the American Statistical Association*. 99, 523–536.

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

17 Ellison, A.M., 2004. Bayesian inference in ecology. *Ecology Letters*. 7, 509–520.

18 Gelman, A., 1996. Inference and monitoring convergence. In Gilks, W. R., Richardson, S., Spiegelhalter, D. J. (eds), *Markov chain Monte Carlo in practice*, pp 131–143. Chapman and Hall, London.

21 Harvey, A.C., 1989. *Forecasting, structural time series models and the Kalman Filter*. Cambridge University Press, Cambridge.

23 Huggins, R., 2001. A note on the difficulties associated with the analysis of capture-recapture experiments with heterogeneous capture probabilities. *Statistics and Probability Letters*. 54, 147–152.

1 Ihaka, R., Gentleman R., 1996. R: a language for data analysis and graphics. *Journal*
2 *of Computational and Graphical Statistics.* 5, 299–314.

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

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

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

10 Lebreton, J.D., Pradel, R., 2002. Multistate recapture models: modelling incomplete
11 individual histories. *Journal of Applied Statistics.* 29, 353–369.

12 Millar, R.B., Meyer, R., 2000. Non-linear state space modelling of fisheries biomass
13 dynamics by using Metropolis-Hastings within-Gibbs sampling. *Journal of the Royal*
14 *Statistical Society Series C - Applied Statistics.* 49, 327–342.

15 Newman, K. B. 1998. State-space modeling of animal movement and mortality with
16 application to salmon. *Biometrics.* 54, 1290–1314.

17 Pollock, K.H., 1991. Modeling capture, recapture, and removal statistics for estimation
18 of demographic parameters for fish and wildlife populations - past, present, and future.
19 *Journal of the American Statistical Association.* 86, 225–238.

20 Pollock, K. H., 2002. The use of auxiliary variables in capture-recapture modelling: an
21 overview. *Journal of Applied Statistics.* 29, 85–102.

22 Pradel, R., 2005. Multievent: An extension of multistate capture-recapture models to
23 uncertain states. *Biometrics.* 61, 442–447.

24 Schwarz, C.J., Schweigert, J.F., Arnason, A.N., 1993. Estimating migration rates using
25 tag-recovery data. *Biometrics.* 49, 177–193.

1 Schwarz, C. J., Seber, G. A. F., 1999. Estimating animal abundance: review III.
2 Statistical Science. 14, 427–56.

3 Seber, G. A. F., 1965. A note on the multiple-recapture census. Biometrika. 52,
4 249–259.

5 Spiegelhalter, D. J., Thomas, A., Best, N. G., Lunn, D., 2003. WinBUGS user manual
6 - Version 1.4. MRC Biostatistics Unit, Cambridge, UK.

7 Sturtz, S., Ligges, U., Gelman, A., 2005. R2WinBUGS: a package for running Win-
8 BUGS from R. Journal of Statistical Software. 12, 1–16.

9 Thomas, L., Buckland, S.T., Newman, K.B., Harwood, J., 2005. A unified framework
10 for modelling wildlife population dynamics. Australian and New Zealand Journal of
11 Statistics. 47, 19–34.

12 Vaupel, J.W., Yashin, A.I., 1985. Heterogeneity ruses - some surprising effects of
13 selection on population dynamics. American Statistician. 39, 176–185.

14 Wang, G., 2006. On the latent state estimation of nonlinear population dynamics using
15 Bayesian and non-Bayesian state-space models. Ecological Modelling. doi: 10.1016/
16 j.ecolmodel.2006.09.004.

17 Williams, B. K., Nichols, J. D., Conroy, M. J., 2002. Analysis and management of
18 animal populations. Academic Press, San Diego.

19 **Appendix A: WinBUGS code for fitting the CJS** 20 **model using the SSM formulation**

```
21 #####
22 # MODEL #
23 # State-space formulation of the Cormack-Jolly-Seber model #
```

```

1 # observations = 0 (non-encountered) and 1 (encountered)      #
2 # states = 0 (dead) and 1 (alive)                             #
3 #####
4 model
5 {
6 # Define the priors for survival phi and detectability p
7 p[1] <- 1
8 phi[1] <- 1
9 for (j in 2:K)
10 {
11 phi[j] ~ dbeta(1,1)
12 p[j] ~ dbeta(1,1)
13 }
14 # Define the SYSTEM PROCESS
15 for (i in 1:n)
16 {
17 # if first capture
18 PrX[i,e[i]] <- 1 # Pr(alive | first capture) = 1
19 X[i,e[i]+1] ~ dbern(PrX[i,e[i]]) # alive (a 1 is generated with certainty)
20 Pr0[i,e[i]] <- X[i,e[i]+1] # detection probability at initial detection is 100%
21 # otherwise
22   for (j in (e[i]+1):K)
23   {
24     PrX[i,j] <- phi[j] * X[i,j]
25     X[i,j+1] ~ dbern(PrX[i,j])

```

```

1      PrO[i,j] <- p[j] * X[i,j+1]
2      }
3      # fullfil the remaining cells with zeros
4      for (j in 1:(e[i]-1))
5      {
6          PrX[i,j] <- 0
7          X[i,j] <- 1
8          PrO[i,j] <- 0
9      }
10     }
11 # Define the OBSERVATION PROCESS
12 for (h in 1:nx)
13 {
14     data[h,3] ~ dbern(PrO[data[h,1],data[h,2]])
15 }
16 }
17 #####
18 # DATA
19 # 'K' is the number of encounter occasions
20 # 'n' is the number of individuals
21 # 'nx' is 'K' times 'n'
22 # 'e' is the vector of first encounters ('n' components)
23 # 'data' is a matrix with dimensions 'nx' times 3 where
24 #   the first column gives the current individual (1,...,'n'),
25 #   the second column gives the current encounter occasion (1,...,'K'),

```

1 # the third column gives the observation (= 1 if detection, = 0 otherwise)
2 # corresponding to the current individual and current encounter occasion
3 #####

4 Appendix B

5 Let $Z_{i,t} = [X_{i,t-1}, X_{i,t}]$ be a bivariate random vector where its two components are
6 denoted $Z_{i,t}^1$ and $Z_{i,t}^2$. Equation (8) becomes

$$7 \quad Y_{i,t}|Z_{i,t} \sim \text{Bernoulli}\left((Z_{i,t}^1 - Z_{i,t}^2)\lambda_{i,t}\right) \quad (10)$$

8 and Equation (9) becomes

$$9 \quad Z_{i,t+1}|Z_{i,t} = \begin{cases} (Z_{i,t+1}^1|Z_{i,t}) = Z_{i,t}^2 \\ Z_{i,t+1}^2|Z_{i,t} \sim \text{Bernoulli}(Z_{i,t}^2\phi_{i,t}). \end{cases} \quad (11)$$

10 The system defined by Equation (10) and Equation (11) is a state-space model and it
11 is equivalent to the model defined by Equation (8) and Equation (9).

12 Note that an alternative state-space formulation can be adopted using a multistate
13 formulation of the RR model (Lebreton et al., 1999) and Section 3.1.

Figure 1

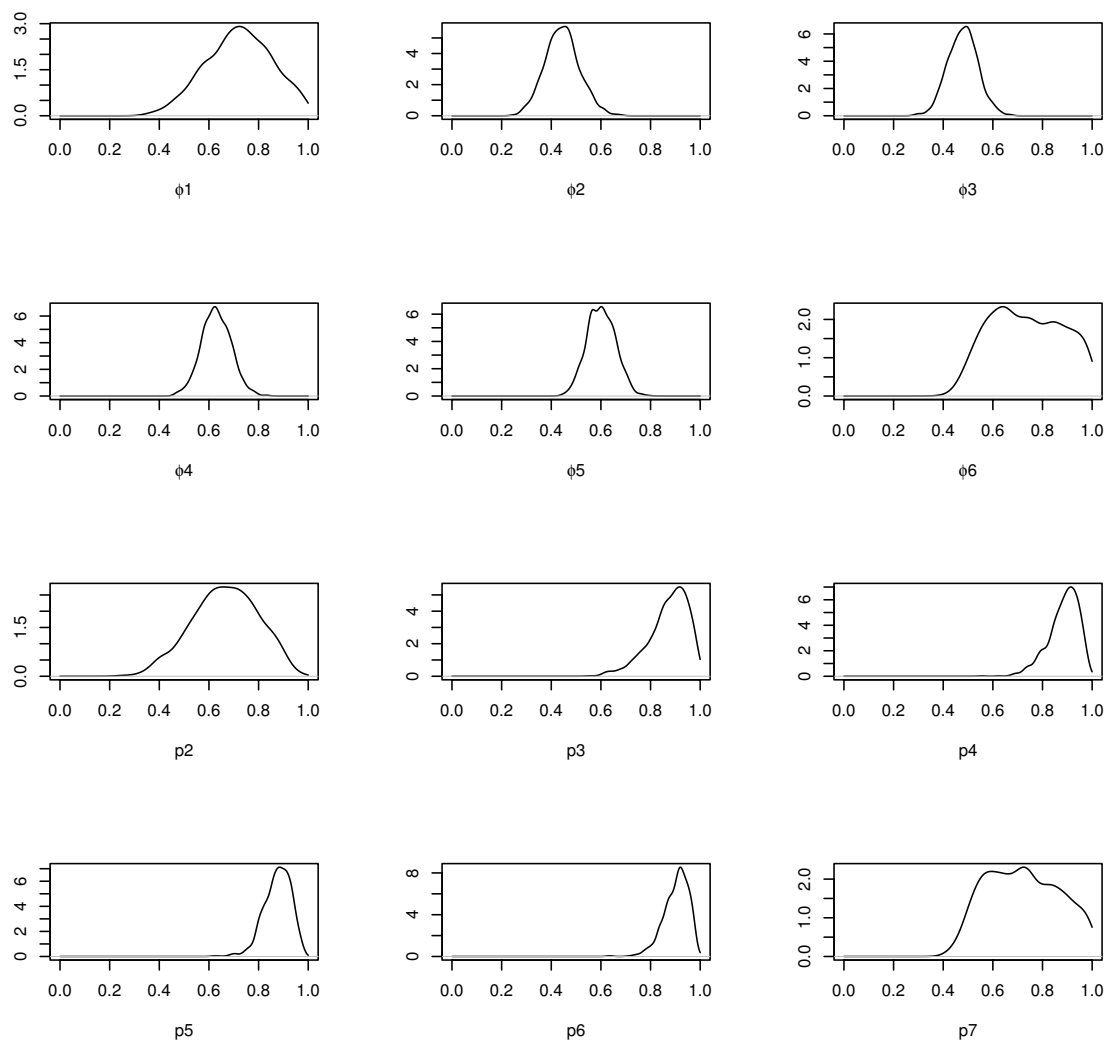


Figure 1: Posterior distributions for the survival and detection probabilities for the Dipper data using the CJS model applied to the Dipper data set as estimated by the state-space model and MCMC methods.

Table 1: Estimated survival and detection probabilities for the Dipper data using the CJS model and three different methods, the state-space model (SSM) using a Monte Carlo Markov Chain (MCMC) method, the product-multinomial model (PMM) using a MCMC method and the PMM using a maximum-likelihood (ML) method. The two first methods were implemented using program WinBUGS (Spiegelhalter et al., 2003), while program M-SURGE (Choquet et al., 2004) was used to implement the last one.

Parameter	SSM using MCMC	PMM using MCMC	PMM using ML
	Posterior median/mode (SD)	Posterior median/mode (SD)	MLE (SE)
ϕ_1	0.721/0.722 (0.132)	0.723/0.693 (0.132)	0.718 (0.156)
ϕ_2	0.448/0.456 (0.071)	0.448/0.460 (0.071)	0.435 (0.069)
ϕ_3	0.480/0.493 (0.060)	0.480/0.476 (0.061)	0.478 (0.060)
ϕ_4	0.628/0.624 (0.061)	0.627/0.616 (0.060)	0.626 (0.059)
ϕ_5	0.602/0.601 (0.057)	0.602/0.607 (0.057)	0.599 (0.056)
ϕ_6	0.713/0.640 (0.142)	0.720/0.628 (0.143)	- (-)*
p_2	0.671/0.658 (0.134)	0.670/0.691 (0.134)	0.696 (0.166)
p_3	0.883/0.918 (0.083)	0.883/0.904 (0.083)	0.923 (0.073)
p_4	0.888/0.914 (0.063)	0.889/0.912 (0.063)	0.913 (0.058)
p_5	0.882/0.885 (0.057)	0.883/0.904 (0.057)	0.901 (0.054)
p_6	0.913/0.920 (0.052)	0.912/0.935 (0.051)	0.932 (0.046)
p_7	0.735/0.724 (0.142)	0.727/0.648 (0.143)	- (-)*

* Non-identifiability detected.