Ecology, 94(10), 2013, pp. 2160–2165

Reports

Estimating transitions between states using measurements with imperfect detection: application to serological data

Rémi Choquet,[1,3] Cécile Carrié,[1] Thierry Chambert,[2] and Thierry Boulinier[1]

[1]*Centre d'Ecologie Fonctionnelle et Evolutive, campus CNRS, UMR 5175, 1919 Route de Mende, F-34293 Montpellier Cedex 5, France*
[2]*Department of Ecology, Montana State University, P.O. Box 173460, Bozeman, Montana 59717-3460 USA*

Abstract. Classifying the states of an individual and quantifying transitions between states are crucial while modeling animal behavior, movement, and physiologic status. When these states are hidden or imperfectly known, it is particularly convenient to relate them to appropriate quantitative measurements taken on the individual. This task is, however, challenging when quantitative measurements are not available at each sampling occasion. For capture–recapture data, various ways of incorporating such non-discrete information have been used, but they are either ad hoc and/or use a fraction of the available information by relying on a priori thresholds to assign individual states. Here we propose assigning discrete states based on a continuous measurement, and then modeled survival and transition probabilities based on these assignments. The main advantage of this new approach is that a more informative use of the non-discrete information is done. As an illustrative working example, we applied this approach to eco-epidemiological data collected across a series of years in which individuals of a long-lived seabird, the Black-legged Kittiwake (*Rissa tridactyla*), could either be visually detected or physically recaptured and blood sampled for subsequent immunological analyses. We discuss how this approach opens many perspectives in eco-epidemiology, but also more broadly, in population ecology.

Key words: *Black-legged Kittiwake;* Borrelia burdgorferi; *continuous measurement; eco-epidemiology; hidden Markov model; immunological assay; mixture;* Rissa tridactyla; *stable-isotope ratio.*

Introduction

Models of population dynamics are crucial for conservation, epidemiologic, and evolutionary studies. Such models are used to describe or predict the dynamic of an ecological process that involves different states across time. Transitions between states have been described using either Markov models (Zucchini and MacDonald 2009), state space models (Patterson et al. 2008), or by a system of differential equations (e.g., in epidemiology; Gandon and Day 2009). A critical issue is thus to estimate the rates of transition between states, which can be done in some situations by making use of non-discrete information available about individuals.

In the wild, it is often impossible to follow individuals exhaustively, making it difficult to determine in which conditions individuals may change states. Capture–recapture (CR) data, which result from the non-exhaustive monitoring of individuals across time and space (Williams et al. 2002), can nevertheless be used to infer ecological dynamic processes, and different categories of Markovian models (sometimes conveniently written as state space models; see Gimenez et al. 2012) can be implemented depending on the type of data available.

These models deal, in general, with discrete state data, although non-discrete information is often available about individuals at capture (e.g., mass, corticosteroid level, isotopic signature, and so on) and can be useful to consider in the analyses. These are sometimes used as covariates (Bonner et al. 2010), for instance, for studying the relation between mass and survival, but it is often convenient to relate such non-discrete measurements to a state (Conn and Cooch 2009). For instance, a given measurement obtained using a specific immunological assay could be indicative of a probable seropositive status of the individual (Rossi et al. 2010, Chambert et al. 2012). As another example, a given stable-isotope ratio in a feather tissue sample could be indicative of a probable particular winter diet or location (Powell 2004, Bearhop et al. 2005, Kelly et al. 2008, Hoye et al. 2012).

In those situations, discrete observations are built from non-discrete data using a priori thresholds, leading to potential misclassification issues. Moreover, no direct use of the full information contained in the continuous measurements has been attempted to estimate underlying states in the case of models with imperfect detection. Here, we propose to do so in a capture–recapture framework.

In human epidemiology studies, because of the comprehensive census of individuals, non-discrete mea-

Manuscript received 29 October 2012; revised 14 May 2013; accepted 4 June 2013. Corresponding Editor: J. R. Sauer.
[3] E-mail: remi.choquet@cefe.cnrs.fr

TABLE 1. Example of 10 individual encounter histories from year 2003 to year 2009.

Individual	2003	2004	2005	2006	2007	2008	2009
IIWXYY	(1, 0)	(1, 0)	(2, 0.3028)	(0, 0)	(0, 0)	(0, 0)	(0, 0)
XBBBOY	(2, 0.6062)	(1, 0)	(2, 0.2907)	(1, 0)	(0, 0)	(1, 0)	(0, 0)
XBBBRS	(1, 0)	(2, 1.299)	(1, 0)	(1, 0)	(1, 0)	(1, 0)	(1, 0)
XBBGGG	(1, 0)	(1, 0)	(2, 0.0434)	(1, 0)	(1, 0)	(0, 0)	(0, 0)
XBBGOI	(1, 0)	(2, 1.073)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)
XBBGWW	(1, 0)	(1, 0)	(1, 0)	(1, 0)	(2, 0.0989)	(1, 0)	(1, 0)
XBBIBW	(1, 0)	(2, 1.305)	(1, 0)	(1, 0)	(1, 0)	(1, 0)	(1, 0)
XBBIGW	(1, 0)	(1, 0)	(2, 0.1377)	(1, 0)	(1, 0)	(1, 0)	(0, 0)
XBBIRW	(1, 0)	(2, 0.764)	(1, 0)	(1, 0)	(1, 0)	(1, 0)	(0, 0)
XBBISI	(1, 0)	(2, 1.153)	(1, 0)	(1, 0)	(1, 0)	(1, 0)	(1, 0)

Notes: Each year, the data available for an individual Black-legged Kittiwake (*Rissa tridactyla*) is summarized by two entries (x, y), representing a mixture of observations (x) and captures/measurements (y). An individual was either observed ($x = 1$) or not ($x = 0$), and conditional on its observation, it was either captured and measured ($y =$ serological state measured as "u"), or not captured ($y = 0$). There are, therefore, three possible observational events for each individual-year: (1) visually detected but not physically captured (1, 0); (2) detected, captured, and measured (2, u); or (3) not detected at all (0, 0). Each individual was identified by a unique combination of six color and one metal bands, with each color represented by a letter code (B, blue; G, green; I, indigo; O, Orange; R, red; S, silver; W, white; X, metal; Y, yellow). For example, XBBBOY stands for the combination metal–blue–blue–blue–orange–yellow. The measurement values (following number 2, capture) correspond to optical densities (OD) of a specific anti-*Borrelia* ELISA assay conducted on the plasmas obtained from the blood samples; higher OD values correspond to higher levels of antibodies against the bacteria.

surements can be directly related to a given state through an appropriate density function (e.g., measurements conditional to a particular state may be modeled by a normal distribution). This approach has been followed in the case of captive animals (O'Connell et al. 2010) and for wild animals followed by telemetry (Patterson et al. 2008). When individuals are not followed comprehensively, a difficulty is added because values of the measurement are missing. Using an illustrative epidemiological example in free-ranging seabirds, we propose an approach based on reasonable assumptions to explicitly incorporate the modeling of the probabilistic distribution of the measurement (bird serological status) into the CR framework in order to estimate the probability to belong to different underlying serological states and probabilities of transition among them. Advantages of this approach are that no a priori thresholds have to be defined, a more parsimonious model can be used, the non-discrete information can be fully used and one can perform assessment of the goodness-of-fit of the measured quantitative variable to a specific distribution.

To illustrate our approach, we considered a longitudinal study of Black-legged Kittiwakes (*Rissa tridactyla*) in which eco-epidemiological data were not comprehensively available, as, at each sampling occasion, individuals could either be resighted visually (detected) only, or physically recaptured, and blood sampled for subsequent immunological analyses. We used this data set to estimate (1) the rates of seroconversion following natural exposure to an infectious agent and (2) a potential association between serological status and survival rate. Under normal distribution assumptions of the immunological measurements for seropositive and seronegative states, the approach we propose allowed us to estimate those quantities from the incomplete measurements.

STATISTICAL FRAMEWORK

Bivariate CR data

At each sampling occasion of the longitudinal study, three exclusive events related to each individual may occur. The event (or observation) for an individual was denoted "0" for not seen, "1" for observed but not captured physically, and "2" for captured physically and measured (i.e., a blood sample was taken and an immunological assay implemented).

This is a case where there is one discrete and one continuous valued components. A convenient approach to represent this kind of data is a bivariate series, (Y_t, U_t), where Y_t and U_t are, respectively, a discrete and a continuous variable at sampling occasion t. This approach is nonstandard in CR, which usually uses discrete univariate series (Williams et al. 2002). In Table 1, we represent 10 individual CR histories using bivariate series. By default, $u_t = 0$ when there is no physical capture ($y_t = 0, 1$). This representation of the data can easily be extended when several geographical sites are considered and to multivariate series when several measurements are taken.

Hidden Markov model for bivariate CR data

As statistical framework, we considered hidden Markov models (HMM), already used in CR in Pradel (2005) for univariate discrete series. These models are very convenient when the variable of states X_t cannot be directly observed, but rather are partially informed by the variable of observation O_t. A HMM is uniquely defined by three matrices ($\mathbf{\Pi}$, $\mathbf{A}$, $\mathbf{E}$): with $\mathbf{\Pi}_t$ being the row vector for initial distribution; $\pi_{t,i} = (\Pi)_{t,i}$ is the proportion of individuals seen for the first time in state i at occasion t. $\mathbf{A}_t$ is the matrix of transition; $a_{t,ij} = (A)_{t,ij}$ is the transition probability from state i at occasion t to state j at occasion $t + 1$. $\mathbf{E}_t$ is the matrix of event; $e_{t,j}(o) =$

$(E)_{t,j,o}$ is the joint probability/density function of $o = (y, u)$ conditional to the state j at occasion t.

In classical CR models, the observation is univariate; i.e., $O_t = (Y_t)$. Thus, the probability function of o conditional to the state j is

$$e_{t,j}(o) = e_{t,j}(y) \tag{1}$$

where $e_{t,j}(y)$ is the probability of an individual in state j at occasion t to be observed as y. This is what was proposed by Pradel (2005) to apply HMM to CR. In our case, $O_t = (Y_t, U_t)$ is a bivariate variable. Assuming contemporaneous conditional independence, the joint probability/density function is

$$e_{t,j}(o_t) = e_{t,j}(y_t) \times f_j(u_t \,|\, y_t). \tag{2}$$

For an individual physically captured (coded 2) and measured at time t with $o_t = (2, 1.073)$ then $\mathrm{e}_{t,j}(o_t) = e_{t,j}(2) \times f_j(1.073 \,|\, 2)$ and $f_j(u \,|\, 2)$ can be any density function depending on the nature of the data (a normal, a mixture of normals, and so on). When there is no measurement, $y = 0, 1$ and $f_j(u \,|\, y) = 1$. In that case, Eq. 2 simplifies to Eq. 1.

Estimation

We considered T sampling occasions and one individual history $h = (o_1, o_2, \ldots, o_T)$, with $o_t = (y_t, u_t)$, the bivariate at sampling occasions $t = 1, \ldots, T$. Let b be the time occasion of the first detection ($y_b > 0$). Let $\boldsymbol{\theta}$ be the vector of k parameters to be estimated. We considered that the Markov model is inhomogeneous, i.e., depending on time, and that individuals at one time have the same probability to survive and to be observed. From the property of a HMM (Zucchini and MacDonald 2009), the likelihood (probability P) of a history h is given by

$$P(h \,|\, \theta) = \Pi_e \times D\left(\mathbf{E}_b^0(., o_b)\right) \times \prod_{t=b+1}^{T} \left[\mathbf{A}_t \times D\left(\mathbf{E}_t(., o_t)\right)\right] \times \mathbf{1}_S \tag{3}$$

where S is the number of states, $\mathbf{1}_S$ is the S-vector of one, $\mathbf{E}_t(., o_t)$ is the vector made of $e_{t,j}(y_t) \times f_j(u_t \,|\, y_t)$, $j = 1, \ldots, S$, $\mathbf{E}_b^0(., o_b)$ is the vector made of $e_{b,j}(y_b) \times f_j(u_b \,|\, y_b)$, $j = 1, \ldots, S$ with $e_{b,j}(1) + e_{b,j}(2) = 1$, and $D(\mathbf{x})$ is the diagonal matrix of the vector $\mathbf{x}$. We made the common assumption in CR (Lebreton et al. 1992) that individuals are independent. Thus the likelihood of the set of histories H is

$$L(\theta) = \prod_{h \in H} P(h \,|\, \theta). \tag{4}$$

The maximum-likelihood estimator (MLE) denoted $\hat{\theta}$ can be estimated by minimizing the deviance $-2 \log L(\theta)$ Confidence intervals can be computed using the asymptotic normality properties of the MLE and the delta method.

Diagnostics on continuous data

Even after model selection, the question remains of whether a model is acceptable or not. For a HMM, no strict analog to a residual exists as the value of a residual depends on the state, which is unknown. However, from the estimated proportion of individuals in each state among those captured for the first time, we can compute iteratively the proportion of individuals captured in each state at the next occasions. Based on these proportions, we can estimate the number of individuals captured in each state at any time (see the Appendix for the calculation of the proportion and the number of individuals captured in each state at each occasion). Based on these quantities, we can assess the goodness-of-fit of the variable U_t with a nonparametric Kolmogorov-Smirnov (K-S) test. The K-S test compares the sample, here a set of measurements, with a reference probability distribution, here a mixture of two normal distributions. This allowed us to do a posteriori goodness-of-fit tests on the measurements. This is a very important aspect of that approach as the possibility to validate a model represents a fundamental complement to a model selection based on information criterion. This was missing in all the previous comparable CR studies on eco-epidemiological questions (Rouan 2007, Conn and Cooch 2009). It is also possible to draw quantile–quantile (QQ) plots of observations of U_t, although visual tests are less adequate than the previous test to assess goodness of fit as several draws from the mixture need to be performed.

Illustrative Working Example

Eco-epidemiology data

As in Chambert et al. (2012), the goal of this application was to (1) measure the potential effects of exposure to the tick-borne agent of Lyme disease, *Borrelia burgdorferi sensu lato* (Bbsl) on the survival of Black-legged Kittiwakes, and (2) estimate their rates of transition between seronegative and seropositive statuses. The data used for that purpose come from a CR study, where breeding birds had been repeatedly resighted and/or captured and blood sampled for immunological analyses over a series of years (see Chambert et al. [2012] and Staszewski et al. [2007] for more details). The assignment of an individual serological state for each year that a bird was physically captured was based on a quantitative measurement of antibody (Ab) level resulting from a specific anti-Borrelia enzyme-linked immunosorbent assay (ELISA) of a plasma sample. Measurements of Ab levels were taken from 315 individually marked kittiwakes surveyed from 2003 to 2009. We here considered that high Ab levels corresponded to a seropositive state, while low Ab levels indicated a seronegative state (Staszewski et al. 2007), and individuals with intermediate values were attributed to the negative or positive serological state depending on the overlap of the respective bi-modal distributions of

negative and positive individuals (Table 1). The detection of antibodies means that the animal has been exposed (infected) by the bacteria, but does not imply that the individual is still infected and/or sick (Staszewski et al. 2007). The kittiwake data set is provided in the Supplement in the form of a MATLAB array.

Eco-epidemiology model

The two serological states were seropositive (P) and seronegative (N). At any sampling occasion t, the state space was thus $\{N, P, \dagger\}$, with † denoting a bird that is dead. To illustrate the method, we used a single model based on the best model retained in Chambert et al. (2012); notably, the model with $s_{t,i}$, the probability to survive from time t to time $t+1$ being a function of the state $i = \{N, P\}$, was not retained (see *Results and diagnostics* section). To simplify, we also considered a model where initial distribution parameters and the first captured probability were independent of time. Parameters of our model were: π, the constant probability of being in state N at the first observation event; $p_{t,j}$, the probability of being observed and not captured at occasion t, conditional of being in state; $c_{t,j}$, the probability of being captured at occasion t, conditional of being in state $j = \{N, P\}$; $c_{0,j}$, the probability of being captured at the first observed event ($y_b > 0$), conditional of being in state $j \in \{N, P\}$, for $t = 1, \ldots, 5$ (As there were no physical captures in years 2008 and 2009, $c_{0,j} = 0$ for $t = 6, 7$); s_t, the probability of survival from occasions t to $t+1$; and t_{NP} and t_{PN}, respectively, the homogeneous probabilities of infection and recovery between two occasions. The last survival probability s_6 and the last probabilities of being observed $p_{7,j}$ and captured $c_{7,j}$ were demonstrated not to be separately identifiable using the method of Choquet and Cole (2012); only products $s_6 p_{7,j}$ and $s_6 c_{7,j}$ are identifiable. To get a full rank model, we thus set the following constraint: $s_5 = s_6$.

This model can easily be adapted to the case where both initial distribution and first captured probabilities are time dependent. The histogram of measurement values is close to a bimodal curve (Chambert et al. 2012), so we modeled the measurements of antibody level as a mixture of two normal distributions. We assumed that measurements were independent, identically distributed as a normal density function, conditionally to the state N or P. Let $g(u, \mu, \sigma)$ be the normal density function of mean μ and variance σ^2 for the measurement value u; $f_N(u \mid y = 2) = g(u, \mu_N, \sigma_N)$ and $f_P(u \mid y = 2) = g(u, \mu_P, \sigma_P)$ are the two densities functions, respectively, associated with states N and P. We had four additional parameters for the mixture with two components: two means μ_N, μ_P and two variances σ_N^2, σ_P^2, respectively, associated with states N and P.

Consider the history, $h = (1, 0), (2, 0.764), (1, 0)$, of an individual that is observed at occasions 1 and 3, and physically captured at occasion 2. The probability of the history h is $P(h) = P(h \mid \text{state} = N) \times \pi + P(h \mid \text{state} = P) \times (1 - \pi)$ where $P(h \mid \text{state} = i)$ represents the probability of h conditional on being in the state i at occasion 1. For this short history h, we can analytically develop the probability of each term according to the parameters previously defined. Thus,

$$
\begin{aligned}
&P(h \mid \text{state} = N) \\
&\quad = (1 - c_{0,N}) \times s_1 \\
&\qquad \times \big\{(1 - t_{NP}) \times c_{2,N} \times g(0.764, \mu_N, \sigma_N) \times s_2 \\
&\qquad\quad \times [(1 - t_{NP}) \times p_{3,N} + t_{NP} \times p_{3,P}] + t_{NP} \times c_{2,P} \\
&\qquad\quad \times g(0.764, \mu_P, \sigma_P) \times s_2 \\
&\qquad\quad \times [t_{PN} \times p_{3,N} + (1 - t_{PN}) \times p_{3,P}]\big\}
\end{aligned}
$$

and

$$
\begin{aligned}
&P(h \mid \text{state} = P) \\
&\quad = (1 - c_{0,P}) \times s_1 \\
&\qquad \times \big\{t_{PN} \times c_{2,N} \times g(0.764, \mu_N, \sigma_N) \times s_2 \\
&\qquad\quad \times [(1 - t_{NP}) \times p_{3,N} + t_{NP} \times p_{3,P}] \\
&\qquad\quad + (1 - t_{PN}) \times c_{2,P} \times g(0.764, \mu_P, \sigma_P) \times s_2 \\
&\qquad\quad \times [t_{PN} \times p_{3,N} + (1 - t_{PN}) \times p_{3,P}]\big\}.
\end{aligned}
$$

Thus, probabilities are quite complex even for this short history h. We can also define the matrices which represent the model. At the first non-null event ($y > 0$), individuals can be in states P or N. Thus, we set the initial state vector as $\Pi = (\pi_1 \; \pi_0)$. The matrix of survival-transitions between states is

$$
\mathbf{A}_t = \begin{pmatrix} s_t \times (1 - t_{NP}) & s_t \times t_{NP} & 1 - s_t \\ s_t \times t_{PN} & s_t \times (1 - t_{PN}) & 1 - s_t \\ 0 & 0 & 1 \end{pmatrix}.
$$

The matrix of events, which related the states $\{N, P, \dagger\}$ (in rows) to bivariates (in columns) is

$$
\mathbf{E}_t = \begin{pmatrix} 1 - p_{t,N} - c_{t,N} & p_{t,N} & c_{t,N} \times g(u_t, \mu_N, \sigma_N) \\ 1 - p_{t,P} - c_{t,P} & p_{t,P} & c_{t,P} \times g(u_t, \mu_P, \sigma_P) \\ 1 & 0 & 0 \end{pmatrix}
$$

with columns indices being respectively (0, 0), (1, 0), (2, u). As the model is conditional on the first non-null event, the following constraint applies to the first observation event (ocurring at time b): $P_{b,j} + c_{b,j} = 1$, for $j = N, P$ Thus, the matrix E_t^0 is defined by

$$
\mathbf{E}_b^0 = \begin{pmatrix} 0 & 1 - c_{0,N} & c_{0,N} \times g(u_b, \mu_N, \sigma_N) \\ 0 & 1 - c_{0,P} & c_{0,P} \times g(u_b, \mu_P, \sigma_P) \\ 1 & 0 & 0 \end{pmatrix}.
$$

The likelihood given by Eq. 4 was implemented in MATLAB and we used a quasi-Newton method to minimize the deviance. Some numerical difficulties may appear with fitting the continuous model when variances

were estimated close to zero. This was not the case here, but Zucchini and MacDonald (2009) discuss a way to prevent this problem.

Results and diagnostics

The most parsimonious models did not include an effect of the seropositive status on the survival rate as in Chambert et al. (2012). The likelihood ratio test (LRT) between the time-dependent model (deviance = 2368.39) and the time- and state-dependent model (deviance = 2362.01) with five degrees of freedom rejected the more complex model ($P = 0.27$). Table 2 displays the estimates and 95% confidence intervals for the time-dependent survival, transitions between states, and parameters of the mixture for the single simple model. Estimates of survival probabilities were close to the ones found by Chambert et al. (2012). The annual probability of becoming seropositive (t_{NP}) was not significantly lower than the one found by Chambert et al. (2012) 0.081 [0.025, 0.229] vs. 0.12. The annual probability of recovering (t_{PN}) was estimated to 0.024 [0.004, 0.130] vs. 0 in Chambert et al. (2012). However, (t_{PN}) was close to the annual estimated transition probability from the positive state to the intermediate state found in Chambert et al. (2012), suggesting a slow decrease of the antibody level/concentration with time.

We checked that there was no transition to a reverse status of positive individuals ($t_{PN} = 0$) by fitting the same model with $t_{PN} = 0$. The deviance was 2373.07 instead of 2368.39. Because the two models were of full rank, with only one additional parameter for the more parameterized one, we did a LRT test between the two models. In the case of a boundary parameter, the LRT test follows a mixture of χ^2 with 1 and 0 degrees of freedom, respectively (Dannemann and Holzmann 2008). The hypothesis of a null probability of transition between the positive to the negative state ($H_0 : t_{PN} = 0$) was rejected ($P = 0.015$).

We assessed the goodness of fit of the variable U_t for occasions 1 to 5 and globally. The number of accepted Kolmogorov-Smirnov tests were, respectively, 9959 ($t = 1$), 1604 ($t = 2$), 8647 ($t = 3$), 9916 ($t = 4$), 9965 ($t = 5$), and 9986 (global), using 10 000 draws. The model fits the data well at occasions $t = 1$, 4, 5, and quite well at occasion $t = 3$, but performs badly at occasion 2. But globally, we can conclude that there is a good adequacy of the normal mixture model to the measurement data.

Table 2. Estimates and confidence intervals (CIs) of model parameters: annual survival probabilities (s_t, $t = 1, \ldots, 5$), annual transition probabilities (t_{ij}) between serological states (from state i at occasion t to state j at occasion $t + 1$; $i, j \in \{N, P\}$), and parameters of the two normal density functions describing the antibody level conditional on each serological state (mean μ_i, $i \in \{N, P\}$ and standard deviation σ_i, $i \in \{N, P\}$).

Parameters	Estimates	95% CIs
s_1	0.950	[0.880, 0.980]
s_2	0.896	[0.836, 0.936]
s_3	0.917	[0.866, 0.951]
s_4	0.892	[0.835, 0.932]
s_5	0.847	[0.793, 0.889]
t_{NN}	0.919	[0.771, 0.975]
t_{NP}	0.081	[0.025, 0.229]
t_{PP}	0.976	[0.870, 0.996]
t_{PN}	0.024	[0.004, 0.130]
μ_N	0.135	[0.122, 0.148]
μ_P	0.932	[0.844, 1.020]
σ_N	0.068	[0.056, 0.082]
σ_P	0.491	[0.430, 0.560]

Note: Serological states are indicated by subscripts (*N*, seronegative, and *P*, seropositive).

Discussion

In this paper, we described an extension of univariate CR models to bivariate CR models when one variable is a continuous measurement. Applying the approach to the kittiwake data, we found that the sample of continuous measurements was well approximated by a mixture of two normal distributions. In addition to confirming the lack of an association between anti-Borrelia serological status and annual survival rate in adult kittiwakes (Chambert et al. 2012), we estimated the annual transition rate from the seronegative state to the seropositive state to be 8%, while we show that the transition rate from the seropositive state to the seronegative one is very low, but not null.

Concerning the estimation method, we used a frequentist approach and chose the direct minimization of the deviance by a quasi-Newton approach. Other alternatives exist, like the expectation–maximization algorithm (Cappé et al. 2005) or a Bayesian approach with Markov chain Monte Carlo computation (MCMC). Our approach was efficient (see the MATLAB codes provided in the Supplement). We plan to implement these kinds of models in the near future, in a more general software package for CR such as E-SURGE (Choquet et al. 2009).

If measurements can be taken more frequently, and if one wants to consider more accurately the dynamics of the state of individuals, more refined models may be used. Additional states may, for instance, be added to account for potential re-exposure to an infectious agent in an eco-epidemiology context. The time since infection may also be considered to influence the rates of changes of the immunological state of individuals. In that case, the change between serological states can no longer be described by a first-order Markov process. This could require adding a memory component because accounting for the time spent in a given state may be important, and in that case, the proposed likelihood is not adequate. Extending our approach to a semi-Markov model could be particularly relevant in that context (O'Connell et al. 2010).

The model can be easily extended to data sets with several sites/states and to multivariate cases when several measurements are taken. Other kinds of density functions can also be used; even discrete probabilities in the case of discrete measurements. Furthermore, different sources of variance can be explicitly taken into account in the measurement, e.g., the variance of the individual response

and the error of the measurement (McClintock et al. 2010). The former can be easily implemented in a frequentist framework using quadrature (Gimenez and Choquet 2010). However, for hierarchical model structures, it may be easier to implement MCMC methods. It may also be convenient to consider appropriate variance structure (e.g., autoregressive model AR(1)) when correlation exists between successive time periods.

We believe that our approach has many applications. For example, the use of stable-isotope information may allow the classification of individuals as a function of their wintering area and/or feeding behavior, and thus, allow the exploration of potential behavioral changes and relationships with annual survival rates (Powell 2004, Bearhop et al. 2005, Kelly et al. 2008, Hoye et al. 2012). In general, this approach can be used to study in a CR framework the dynamics of different individual states for which continuous covariates show different responses depending on the states.

Acknowledgments

The authors thank two anonymous reviewers for very useful comments on an earlier version of the article. R. Choquet and C. Carrié contributed equally to the work

Literature Cited

Bearhop, S., W. Fiedler, R. W. Furness, S. C. Votier, S. Waldron, J. Newton, G. J. Bowen, P. Berthold, and K. Farnsworth. 2005. Assortative mating as a mechanism for rapid evolution of a migratory divide. Science 310:502–504.

Bonner, S. J., B. J. T. Morgan, and R. King. 2010. Continuous covariates in mark-recapture-recovery analysis: a comparison of methods. Biometrics 66:1256–1265.

Cappé, O., E. Moulines, and T. Rydèn. 2005. Inference in hidden Markov models. Springer series in Statistics. Springer, New York, New York, USA.

Chambert, T., V. Staszewski, E. Lobato, R. Choquet, C. Carrié, K. D. McCoy, T. Tveraa, and T. Boulinier. 2012. Exposure of black-legged kittiwakes to Lyme disease spirochetes: dynamics of the immune status of adult hosts and effects on their survival. Journal of Animal Ecology 81:986–995.

Choquet, R., and D. J. Cole. 2012. A hybrid symbolic-numerical method for determining model structure. Mathematical Biosciences 236:117–125.

Choquet, R., L. Rouan, and R. Pradel. 2009. Program E-SURGE: A software application for fitting multievent models. Pages 845–865 *in* D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors. Modeling demographic processes in marked populations. Volume 3 of Springer Series: Environmental and Ecological Statistics. Springer, Dunedin, New Zealand.

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

Dannemann, J., and H. Holzmann. 2008. Likelihood ratio testing for hidden Markov models under non-standard conditions. Scandinavian Journal of Statistics 35:309–321.

Gandon, S., and T. Day. 2009. Evolutionary epidemiology and the dynamics of adaptation. Evolution 63:826–838.

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

Gimenez, O., J.-D. Lebreton, J.-M. Gaillard, R. Choquet, and R. Pradel. 2012. Estimating demographic parameters using hidden process dynamic models. Theoretical Population Biology 82:307–316.

Hoye, B. J., S. Hahn, B. A. Nolet, and M. Klaassen. 2012. Habitat use throughout migration: linking individual consistency, prior breeding success and future breeding potential. Journal of Animal Ecology 81:657–666.

Kelly, J. F., M. J. Johnson, S. Langridge, and M. Whitfield. 2008. Efficacy of stable isotope ratios in assigning endangered migrants to breeding and wintering sites. Ecological Applications 18:568–576.

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

McClintock, B. T., J. D. Nichols, L. L. Bailey, D. I. MacKenzie, W. L. Kendall, and A. B. Franklin. 2010. Seeking a second opinion: uncertainty in disease ecology. Ecology Letters 13:659–674.

O'Connell, J., F. Togersen, N. Friggens, P. Lovendahl, and S. Hojsgaard. 2010. Combining cattle activity and progesterone measurements using hidden semi-Markov models. Journal of Agricultural, Biological, and Environmental Statistics 16:1–16.

Patterson, T. A., L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos. 2008. State-space models of individual animal movement. Trends in Ecology and Evolution 23:87–94.

Powell, L. A. 2004. A multistate capture-recapture model using a posteriori classification to enhance estimation of movement rates. Condor 106:761–767.

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

Rossi, S., C. Toigo, J. Hars, F. Pol, J. L. Hamann, K. Depner, and M. F. Le Potier. 2010. New insights on the management of wildlife diseases using multi-state recapture models: the case of classical swine fever in wild boar. PLoS ONE 6: e24257.

Rouan, L. 2007. Apports des chaînes de Markov cachées à l'analyse de données de capture-recapture. Dissertation. University of Montpellier, Montpellier, France.

Staszewski, V., K. D. McCoy, T. Tveraa, and T. Boulinier. 2007. Interannual dynamics of antibody levels in naturally infected long-lived colonial birds. Ecology 88:3183–3191.

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations: modeling, estimation, and decision making. Academic Press, London, UK.

Zucchini, W., and I. MacDonald. 2009. Hidden Markov models for time series: an introduction using R. CRC Press, Boca Raton, Florida, USA.

Supplemental Material

Appendix

Computation of proportions of the mixture (*Ecological Archives* E094-202-A1).

Supplement

Source code and example data for fitting the kittiwake data set by maximum likelihood (*Ecological Archives* E094-202-S1).