ENCOUNTER HISTORY MODELING OF JOINT MARK-RECAPTURE, TAG-RESIGHTING AND TAG-RECOVERY DATA UNDER TEMPORARY EMIGRATION

Richard J. Barker[1], Kenneth P. Burnham[2] and Gary C. White[2]

[1]*University of Otago and* [2]*Colorado State University*

Abstract: We describe a joint analysis of mark-recapture, tag-resight and tag-recovery data that directly models the encounter history of an animal. The probability of the encounter history for each animal is partitioned into survival, recapture, resighting, and recovery components, and a component for the probability that the animal is never encountered again. Temporary migration enters into the likelihood through the recapture component, and movement of marked animals in and out of the area where they are subject to capture is modeled using a Markov chain. Random temporary emigration and permanent emigration are special cases. An important feature of directly modeling the encounter histories is that covariates that are specific to individuals can be included in the analysis. The model is applied to a brown trout tagging data set and provides strong evidence of Markovian temporary emigration. The new model is needed to provide correct estimates of trout survival probabilities which are shown to depend on the length of the fish at first capture.

Key words and phrases: Band-recovery, capture-recapture, Cormack-Jolly-Seber model, mark-recapture, markovian movement, tag-recovery, tag-resighting, temporary emigration.

1. Introduction

Mark-recapture models are a fundamental tool used in the study of population size and demographic parameters of wildlife populations. For open populations (i.e., subject to births and deaths during the study period), the principal model has been the Cormack-Jolly-Seber (CJS) model (Cormack (1964)) in which recaptures of marked animals are used for inference about survival and capture processes. This model was subsequently extended by Jolly (1965) and Seber (1965) to include estimation of abundance by modeling captures of unmarked animals.

Recent mark-recapture research has led to the development of models that combine data from different sources. For example, Pollock's robust design (Pollock (1982), Kendall, Pollock and Brownie (1995)) extends the Jolly-Seber model by using information from multiple captures of animals within a capture session,

during which time the population is assumed closed. Burnham's (1993) joint live-recapture/dead recovery model extends the CJS model by including data from animals that are found dead between capture occasions. An advantage of combining data from different sources is that estimates of key parameters will usually be more precise (Lebreton, Morgan, Pradel and Freeman (1995), Catchpole, Freeman, Morgan and Harris (1998)) even if the new information can only be utilized by adding some new parameters to the model (Barker and Kavalieris (2001)).

Reducing the sampling variance of key parameter estimates is not the only benefit of combining data from different sources. It may also allow relaxation of key assumptions leading to parameter estimates that can be obtained under a wider range of conditions. The new parameters that must be introduced to the model may also be of interest in their own right (e.g., Kendall, Nichols and Hines (1997)).

In an unpublished study of a New Zealand brown trout (*Salmo trutta*) population, fish were tagged in spawning tributaries of Lake Brunner, Westland, between 1987 and 1990 (Barker (1995)). Primary interest was in estimating survival probabilities and only spawning fish were available for capture and tagging. Information on trout survival and movement is provided through later encounters of tagged animals. Fish were re-encountered either by recapture during later tagging operations or through reports of tagged fish received from anglers fishing the lake. Some of these angler-caught fish were reported killed by the anglers but many were reported as returned to the lake alive and healthy. Importantly, the fate of each fish reported caught by an angler (released or killed) was known.

To use the available information fully it is important to model the three types of re-encounter: (1) recaptured alive during tagging operations; (2) caught by an angler during the fishing season and released alive; and (3) caught by an angler during the fishing season and killed. Because not all fish spawn in any given year, and because fish may change the place and time of spawning, it is important that the model allows for animals that move in and out of the spawning population studied between samples so that a valid survival estimate can be obtained. For estimating survival probability the additional movement probabilities are often nuisance parameters. However, here they have a useful biological interpretation.

Dispersal of marked animals away from the study site between capture sessions is likely in many mark-recapture studies. This leads to a division of the marked segment alive in the population at any sampling time into those that are, and are not, at risk of capture. For these studies, movement of animals is an important issue. A key assumption in the Jolly-Seber model is that emigration is permanent (Jolly (1965) and Seber (1982)). That is, it is assumed that if a marked animal disperses away from the study site it is not available for capture for the rest of the study. However this may be unreasonable.

Burnham (1993) showed that the expected number of recaptures in each sample from a tagged cohort is identical under an alternative assumption of random temporary emigration and immigration to that expected under permanent emigration. Under random emigration the risk of capture at time i for animals at risk of capture at time $i - 1$ is the same for all animals. Burnham's (1993) result indicates that the two emigration models are indistinguishable when the data comprise only releases and live recaptures of marked animals. Depending on whether emigration is random or permanent, the survival and capture probabilities are interpreted differently. Under permanent emigration the survival probability represents the joint probability of surviving and not emigrating, and the capture probability is a true capture probability. Under random emigration, the survival probability is interpreted as a true survival probability but the capture probability is the joint probability that the animal is at risk of capture and is captured.

Burnham (1993) described the joint analysis of data in which live recaptures are augmented by data from animals that are found dead between capture occasions (dead recovery data). In this joint analysis the two emigration models can be distinguished and model selection procedures used to choose between them. Barker (1997) generalized Burnham's (1993) model to allow joint analysis of live-recapture, dead recovery and live resighting data in the special case of random temporary emigration. The live resighting data are obtained from reports of tagged animals that are resighted during the interval between capture occasions.

Random emigration (Burnham 1993) is just one possible type of movement in and out of the study site. A more general movement model considered by Kendall, Nichols and Hines (1997) for Pollock's robust design (Pollock (1982)), is Markovian temporary emigration (Markov emigration), where the probability an animal is at risk of capture at time $i + 1$ depends on whether it was at risk of capture at time i. Permanent emigration is a special case of Markov emigration where the probability of subsequent capture is zero for animals not at risk of capture at time i. Random emigration arises by constraining the Markov movement parameters so that the probability of capture at time $i + 1$ does not depend on whether the animal was at risk of capture at time i. The Markov emigration model can also be generalized to allow dependence that extends for more than one time period, for example the memory model of Brownie, Hines, Nichols, Pollock and Hestbeck (1993).

In this paper we describe a generalization of Barker's (1997) model that allows permanent and Markov emigration from the area where animals are at risk of capture as well as random emigration. This means that valid survival estimates can be obtained under a wider range of field conditions. In Barker (1997, 1999), explicit estimates of survival, capture and resighting probabilities are based on

minimal sufficient statistics derived from the encounter history for each animal assuming random emigration. Under permanent and Markov emigration, these statistics are no longer sufficient and explicit maximum likelihood estimates do not appear to exist except in restricted cases. The direct construction of the likelihood described below can be used instead to fit the model directly to the set of individual animal encounter histories.

Directly modeling the encounter histories makes it easy to include covariates that differ for each individual in the study. Clobert (1995) commented that a limitation of mark-recapture models was the inability of open population mark recapture models to incorporate individual covariates such as body mass or length. Heterogeneity among individuals is an important characteristic of animal populations. In principle, it can be modeled using individual covariates (Clobert (1995) and Lebreton (1995)). Analyses where the model is fitted to sufficient statistics formed by summarizing encounter histories across individuals preclude the use of covariates that are different for each animal in the study. Instead, the likelihood function must be constructed directly from the individual encounter histories.

2. Data, Assumptions and Model

We assume there are T occasions of live capture with L $(L \geq T)$ subsequent intervals of live resighting and dead recovery. At capture occasion i all animals in the study fall into one of three categories: (1) not caught at i, (2) caught at i and released, or (3) caught at i and removed from the population at i (Figure 1). Note that we refer to the time at which capture occasion took place as time i.

During the interval $[i, i + 1)$, each animal in the population falls into one of the three categories:
(i)   Not resighted or found dead in $[i, i + 1)$,
(ii) Not found dead but resighted alive at least once in $[i, i + 1)$,
(iii) Recovered dead in $[i, i + 1)$.

There are three things to note. First, although an animal may be resighted several times in the interval, $[i, i + 1)$, we only make use of the fact that it was resighted at least once. Second, animals may be resighted alive during $[i, i + 1)$ and then be found dead in this interval. Such animals are classified as a dead recovery and the preceding resightings are ignored. Third, an animal may be resighted alive (possibly several times) in $[i, i+1)$ and then die later in this interval without being found. The probability of this event is computed differently than the probability of a live resighting for an animal that survives the interval $[i, i+1)$.
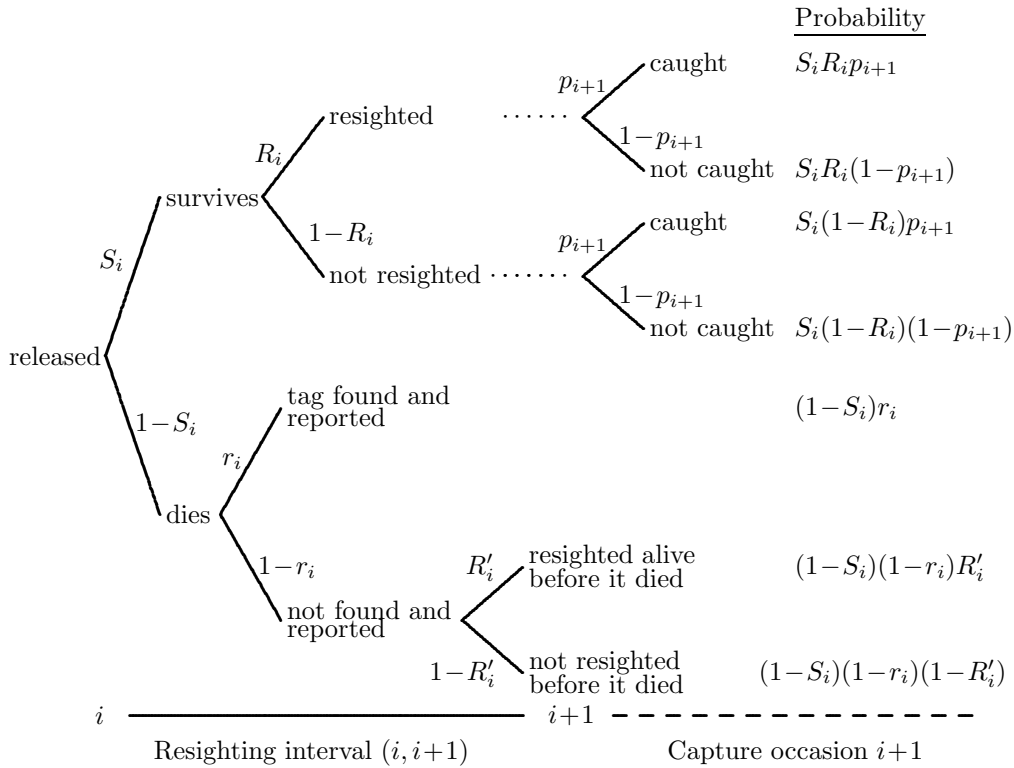
Figure 1. Possible fates for a tagged animal released following capture at occasion i assuming no temporary emigration. Note the order of events shown for animals that die and are resighted alive before they die is not the same as the chronological order. If temporary emigration is possible, then animals can be caught only if they survive and are at risk of capture at $i+1$.

Fish caught by an angler during $[i, i + 1)$ are classified as live-resightings in $[i, i+1)$ if they were released alive, and as dead-recoveries in $[i, i+1)$ if they were killed. Note that an alternative approach discussed by Barker (1997) is to first classify the fish according to whether they were caught by an angler, and then whether they were killed or released.

In addition to the usual assumptions of the Cormack-Jolly-Seber model (e.g., Seber (1965, 1982)) we assume the following.

1. At time i, all marked animals have the same probability of being resighted and the same probability of being found dead in $[i, i + 1)$.

2. All marked animals have the same probability of survival from i to $i + 1$, regardless of whether or not they were at risk of capture at time i.

3. Captures and resightings are independent events and have no influence on subsequent survival.

Assumption 3 can be relaxed to allow some dependence between captures and subsequent survival. For example, Barker (1999) introduced a short-term marking effect into the model for the special case of random emigration. Also, these assumptions can all be relaxed to allow for individual animal differences using covariates, as discussed below in Section 2.4.

2.1. Direct construction of the likelihood

Parameters in the model are of two kinds: (i) fundamental biological parameters that directly govern survival, movement, capture, and resighting processes; and (ii) derived parameters that simplify the direct construction of the likelihood function.

Fundamental Parameters: S_i is the probability that an animal alive at time i survives to $i + 1$; r_i is the probability that an animal that dies in $[i, i + 1)$ is found and reported; R_i is the probability that an animal alive at time i and $i+1$ is resighted alive in $[i, i + 1)$; R_i' is the probability that an animal alive at time i and that is dead by time $i + 1$ without being reported dead, is resighted alive in $[i, i + 1)$; p_i is the probability that an animal alive and available for capture at time i is captured at time i; F_i is the probability that an animal alive and at risk of capture at time i is at risk of capture at time $i + 1$; F_i' is the probability that an animal alive and not at risk of capture at time i is at risk of capture at time $i + 1$.

Derived parameters: α_{ij} is the probability that an animal alive at times i and j is not captured between i and j but is captured at j; β_{ij} is the probability that an animal alive at times i and j is not captured between i and j, and is not captured at j; f_i is the probability that an animal alive at time i is either resighted alive or found dead in $[i, i + 1)$; ϕ_i is the probability that an alive at time i and resighted alive in $[i, i + 1)$ is still alive at time $i + 1$.

An encounter history is recorded for each animal captured at least once in the study. We obtain the probability of this encounter history conditional on the first release, which must follow a live capture (or effective capture in a study based on natural markings). Let H_j denote the encounter history for animal j first captured at time k_j and where time l_j is the last capture occasion when this animal was known to be alive. Note that when speaking generally we omit the subscript j. The occasion l can be determined in two ways: (1) the animals was captured at l and not encountered at or after $l + 1$, or (2) it was not captured at l but it was seen (alive or dead) in $[l, l + 1)$ and was never encountered again.

If h_j is the observed value of H_j then the likelihood function is given by

$$L \propto \prod_j Pr(H_j = h_j \,|\, \text{first released at } k_j).$$

To compute $Pr(H = h \,|\,$ first released at $k)$ we partition events into those that took place in the interval $(k, l]$ and those that took place after l. Using the assumption that captures and resightings are independent and have no influence on subsequent survival, we can factor the likelihood contribution from each animal:

$$Pr(H = h \,|\, \text{first released at } k)$$
$$\propto Pr(\text{survival to } l \,|\, \text{first released at } k)$$
$$\times Pr(\text{resighting history in } [k, l) \,|\, \text{survival to } l)$$
$$\times Pr(\text{capture history in } (k, l] \,|\, \text{survival to } l)$$
$$\times Pr(\text{encounter history after } l \,|\, \text{encounter history in } [k, l]).$$

Strictly, there should be a fifth factor for the probability of the releases conditional on the captures that takes into account losses on capture. As this factor depends only on the probability that an animal is released given that it was caught, and no other parameters in the model, we omit it from further consideration.

The term $Pr(\text{survival to } l \,|\, \text{first released at } k)$ is given by $S_k S_{k+1} \cdots S_{l-1}$ for $k < l$, and 1 for $k = l$. Because we know the animal was alive between k and l, $Pr(\text{resighting history in } [k, l) \,|\, \text{survival to } l) = \prod_{i=k}^{l-1} R_i^{y_i}(1 - R_i)^{1-y_i}$, where y_i is the indicator for a live resighting in $[i, i + 1)$.

Under Markov emigration, failure to catch an animal may occur either because it was available for capture but not caught, or because it was unavailable for capture. It is expedient to express the problem in terms of a 2-state Markov chain (Schwarz and Stobo (1999)). Let state 1 represent the condition that the animal is at risk of capture and state 2 the condition the animal is not at risk of capture. Defining $\boldsymbol{p}_i = \begin{bmatrix} p_i & 0 \\ 0 & 0 \end{bmatrix}$ and $\boldsymbol{\psi}_i = \begin{bmatrix} F_i & 1 - F_i \\ F_i' & 1 - F_i' \end{bmatrix}$ then

$$Pr(\text{capture history} \mid \text{survival to } l)$$
$$= [1, 0] \prod_{i=k}^{l-1} (z_i \boldsymbol{\psi}_i \boldsymbol{p}_{i+1} + (1 - z_i)\boldsymbol{\psi}_i(\boldsymbol{I}_2 - \boldsymbol{p}_{i+1})) \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad (1)$$

where z_i is the indicator for recapture at i and $\boldsymbol{I}_2$ is a 2×2 identity matrix.

The initial $[1, 0]$ vector is needed in (1) because all animals are released following capture. For animals not captured at l, the final vector $[1, 1]'$ is needed because it is unknown whether they are at risk of capture at l. For animals captured at l a final vector $[1, \kappa]'$ is required. Because of the zeros in $\boldsymbol{p}_i$, κ can be any real number and the choice is arbitrary. The presence of the zeros in matrices and vectors above leads to considerable redundancy in the computations, however (1) can be rexpressed in a form that is more efficient computationally but less compact when written out.

The probability of the encounter history following capture occasion l, depends on how the animal was encountered in $[l, l+1)$. There are two possibilities for Pr(encounter history after l | encounter history in $[k, l)$). If the animal was last encountered by capture at l, the probability we require is Pr(never encountered after l | last captured at l). Note that if the animal was lost on capture (i.e., removed from the population by the researcher) then this term is 1. If the animal was last encountered by a live resighting in $[l, l+1)$, then the probability we require is Pr(resighting history in $[l, l+1)$ | encounter history to $l) \times Pr$(never encountered after $[l, l+1)$ | encounter history in $[k, l+1)$).

If the last encounter of the animal was a dead recovery in $[l, l+1)$ then Pr(resighting history in $[l, l+1)$ | encounter history to $l) = (1 - S_l)r_l$. If it was by a live resighting in $[l, l+1)$, then Pr(resighting history in $[l, l+1)$ | encounter history to $l) = S_l R_l + (1 - S_l)R'$.

Regardless of how the animal was last encountered, the probability that it was never encountered again, conditional on events up to and including the last encounter, can be found as the probability of the complement of the event that the animal is encountered at least once after the last encounter, as described in the appendix.

2.2. Constraints on movement

If we set $F'_i = F_i$ for all i, we obtain the random temporary emigration model in which the probability an animal is at risk of capture at $i + 1$ does not depend on whether or not it was at risk of capture at i. Under random temporary emigration, the movement probabilities are expressed in the likelihood through the derived parameters α_{ij} and β_{ij} which can be written as $\alpha_{ij} = F_{j-1}p_j \prod_{h=i}^{j-2}(1 - F_h p_{h+1})$ and $\beta_{ij} = \prod_{h=i}^{j-1}(1 - F_h p_{h+1})$. Therefore, under random emigration the pair F_i and p_{i+1} are confounded because they only appear in the product $F_i p_{i+1} = F'_i p_{i+1}$. A computational device for fitting this model is to make the constraint $F_i = 1$ for all i. The resulting capture probability estimate at time i will then be an estimate of the confounded pair $F_i p_{i+1}$. Under random emigration, and for $L = T$, the parameters $F_1 p_2, \ldots, F_{T-1}p_T$, $R_1, \ldots, R_{T-1}$, $R'_1, \ldots, R'_{T-1}$, $r_1, \ldots, r_{T-1}$, $S_1, \ldots, S_{T-1}$, f_T, and ν_T are identifiable, where $f_i = S_i R_i + (1 - S_i)\{r_i + (1 - r_i)R'_i\}$ and $\nu_i = 1 - [(1 - S_i)r_i]/f_i$. The function ν_i represents the probability that an animal was resighted alive in $[i, i+1)$ given that it was either resighted alive or found dead in $[i, i+1)$.

Under permanent emigration an animal may leave the at-risk-of-capture component of the population but never return. This is enforced by making the constraint $F'_i = 0$ for all i leading to $\alpha_{ij} = F_{j-1}p_j \prod_{h=i}^{j-2} F_h(1 - p_{h+1})$. There is not an equivalent simplification for β_{ij} but it can be expressed using a recursion

formula which is identical to that given by Burnham (1993) for his joint live-recapture/tag-recovery model. Under permanent emigration, and if $L = T$, the parameters $p_2, \ldots, p_{T-1}, R_1, \ldots, R_{T-1}, R'_1, \ldots, R'_{T-1}, r_1, \ldots, r_{T-1}, S_1, \ldots, S_{T-1},$ $F_1, \ldots, F_{T-2}, f_T, \nu_T,$ and the confounded pair $F_{T-1}p_T$ are identifiable.

Under the full Markov movement model, not all the movement parameters can be uniquely estimated. One approach is to use a computer package such as MARK (White and Burnham (1999)) that is able to find estimates of the estimable parameters despite the presence of inestimable ones in the likelihood. In MARK, the number of inestimable parameters is estimated using a singular value decomposition of the Hessian matrix. Another approach is to set constraints on the movement parameters. For example, we can assume a stationary Markov movement model, where $F_1 = F_2 = \cdots = F_{T-1} = F$, and $F'_2 = F'_3 = \cdots = F'_{T-1} = F'$. (Note that F'_1 never contributes to the likelihood under Markov emigration as all marked animals in the population immediately after sampling period 1 must have been at risk of capture at time 1). Under stationary Markov emigration, and if $L = T$, the parameters $p_2, \ldots, p_T, R_1, \ldots, R_{T-1}, R'_1, \ldots, R'_{T-1},$ $r_1, \ldots, r_{T-1}, S_1, \ldots, S_{T-1}, f_T, \nu_T, F.$ and F' are identifiable.

2.3. Resightings after the final capture period

If resightings continue beyond the interval $[T, T+1)$ we make the constraint $p_i = 0$ $(i > T)$. The parameters ν_i $(i = T+1, \ldots, L)$ and R_T are identifiable, but only certain products of resighting and survival parameters for the intervals beyond t can be estimated. These are

$$\xi_i = f_i \prod_{h=t}^{i-1} S_h(1 - R_h) \quad (i = T+1, \ldots, L),$$

$$\zeta_{ik} = \begin{cases} \phi_i f_k & j = T, \ldots, L-1; \ k = i+1 \\ \phi_i f_k \prod_{h=i+1}^{k-1} S_h(1 - R_h) & j = T, \ldots, L-1; \ k = i+2, \ldots, L \end{cases}.$$

The function ξ_i represents the probability that an animal alive at time T is next encountered (alive or dead) in the interval $[i, i+1)$, and ζ_{ik} the probability that an animal resighted alive in $[i, i+1)$ is next encountered (alive or dead) in the interval $[k, k+1)$. Derivations of these confounded parameters are given in Barker (1995).

2.4. Inclusion of individual covariates

It is straightforward to introduce individual covariates for each animal into the analysis. If a vector of covariates available for the jth animal is defined as

$\boldsymbol{x}_j$ then the parameters of the model can be expressed as a function of these covariates. In this case a time-specific parameter, say θ_i, would be re-expressed as the function $\theta_{ij} = \theta_i(\boldsymbol{x}_j)$. Catchpole et al. (2001) give an example where individual covariates are used in the joint live-recapture/dead recovery context. The inclusion of individual covariates can greatly extend the model structure away from the simple time-specific structure given here, for example to allow parameters to apply at different ages, to allow for temporary marking effects and to allow for other forms of dependencies between samples.

3. Example

Between 1987 and 1990, 3,393 brown trout were tagged in a spawning tributary of Lake Brunner, Westland, New Zealand. Recaptures of 220 trout were recorded during tagging operations, and by 1992, 278 tagged fish were reported killed by anglers (Table 2). A further 56 fish were reported captured and released at least once during a year by an angler without subsequently being reported killed by an angler in that year. To be caught during tagging operations a fish had to be passing up the spawning tributary at the time the stream was trapped. Fish that did not spawn, or that spawned outside the period in which the stream was trapped, are regarded as temporary migrants. Because angling could occur anywhere in the lake system, fish were always at risk of resighting or dead recovery.

Table 1. Definitions of summary statistics.

Statistic	Definition
N_{ic}	Number of fish caught and released in the tagging sample at time i.
n_{ic}	The number of the N_{ic} that were ever encountered again.
m_i	The number of tagged fish caught in the sample at time i.
N_{ir}	The number of fish that were reported caught and released by an angler between the samples at time i and time $i + 1$.
n_{ir}	The number of the N_{ir} that were ever encountered again.
d_i	The number of fish that were reported caught and killed by anglers between the samples at time i and time $i + 1$.

We began our analysis using the fully sex- and time-dependent model (which we denote $S_{\text{sex}*\text{time}} \, p_{\text{sex}*\text{time}} \, r_{\text{sex}*\text{time}} \, R_{\text{sex}*\text{time}} \, R'_{\text{sex}*\text{time}} \, F_{\text{sex}*\text{time}} \, F'_{\text{sex}*\text{time}}$) as the global model and fitted it using program MARK. Goodness-of-fit for the global model was assessed using a parametric bootstrap procedure available in program MARK, in which the observed deviance was compared to randomly generated values obtained using the fitted model. In 1000 bootstrap replicates the deviance for simulated data exceeded the observed deviance in 14% of cases, suggesting that the global model is adequate.

Table 2. Release, recapture, resighting and recovery statistics (see Table 1 for definitions) from brown trout tagged in spawning tributaries of Lake Brunner, Westland, New Zealand, 1987-1991.

i		Male						Female					
		N_{ic}	n_{ic}	m_i	N_{ir}	n_{ir}	d_i	N_{ic}	n_{ic}	m_i	N_{ir}	n_{ir}	d_i
	1	234	64	0	1	0	23	115	42	0	0	0	11
	2	404	49	10	1	0	20	515	115	15	0	0	24
	3	599	76	29	6	0	44	719	101	66	7	1	37
	4	516	64	38	15	4	36	512	37	62	10	1	38
	5	0	0	0	10	0	24	0	0	0	6	0	21

We next fitted models with sex effects and some time effects removed assuming stationary Markov emigration for the movement parameters. Because trapping success depends on weather, and resighting rates on angling effort, we expect time-effects on recapture and resighting probabilities a *priori*. Sex-effects may also be present, although these are likely small for the resighting probabilities as there is little reason to believe that anglers target male or female trout. Therefore, in all reduced models we included time effects on capture and resighting probabilities and screened for a sex-effect on capture probabilities. We fitted 16 reduced parameter models by fitting combinations of models with constant, sex, time, or sex- and time-effects on survival, sex- or time-effects on capture, time effects on resighting probabilities and constant or sex-effects on movement probabilities.

Model selection was carried out using AICc, a small-sample version of Akaike's Information Criterion (Burnham and Anderson (1998, p.51)). Because AICc is measured on a relative scale we report Δ AICc, the difference between AICc for each model and that for the model with the smallest observed AICc from the set of models considered. Burnham and Anderson (1998, p.323) suggest as a rough guide that models with Δ AICc ≤ 2 are considered to have substantial support and should be used for making inferences. Models having Δ AICc of about 4 to 7 have considerably less support, and models with Δ AICc > 10 have essentially no support. We also report Akaike weights w_i given by $w_i = \exp(-\Delta_i/2)/\sum_{j=1}^{R}\exp(-\Delta_j/2)$, where is the difference in AIC between model i and the model with minimum AIC from the set considered. The AIC weights express the relative support for a model on a $(0,1)$ scale and can be used to construct model-averaged parameter estimates (Burnham and Anderson (1998, p.133)).

Of the 16 models initially considered, the best one had sex- and time-effects on survival probability, time-effect (but no sex-effect) on recapture and resighting probabilities, and a sex-effect on movement probabilities (Table 3). There was

some support for a model with sex- and time-dependent capture probabilities. A further eight models were then fitted by elaborating on $S_{\text{sex}*\text{t}}p_t r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$ and $S_{\text{sex}*\text{t}}p_{\text{sex}*\text{t}}r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$ to include an effect of length (an individual covariate) at first capture on survival. These models were generated by modeling the logit of survival probability as a linear function of sex, time and length at first capture, and 2-way interactions between sex and time, and sex and length.

Table 3. Results of model selection to assess the effects of sex and time on survival and capture probability estimates and the effect of sex on movement probabilities in a joint model for the analysis of live recapture, live resighting and dead recovery data from male and female brown trout tagged in spawning tributaries of Lake Brunner, Westland, New Zealand, 1987−1991. Results are presented for the best 10 models.

Model	AICc	Δ AICc	Akaike weight	No. parameters
$S_{\text{sex}*(t+\text{length})}p_t r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$	4706.910	0.000	0.396	33
$S_{\text{sex}*(t+\text{length})}p_{\text{sex}*\text{t}}r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$	4707.821	0.911	0.251	37
$S_{t+\text{sex}*\text{length}}p_t r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$	4708.383	1.473	0.189	29
$S_{\text{sex}*t+\text{length}}p_t r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$	4710.743	3.833	0.058	32
$S_{\text{sex}*t}p_t r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$	4712.341	5.431	0.026	31
$S_t p_t r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$	4713.024	6.114	0.019	25
$S_{t+\text{sex}+\text{length}}p_t r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$	4713.204	6.294	0.017	28
$S_{\text{sex}*t}p_{\text{sex}*\text{t}}r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$	4713.389	6.479	0.016	35
$S_{t+\text{sex}*\text{length}}p_{\text{sex}*\text{t}}r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$	4713.712	6.802	0.013	33
$S_{\text{sex}*t+\text{length}}p_{\text{sex}*\text{t}}r_t R_t R'_t F_{\text{sex}}F'_{\text{sex}}$	4714.037	7.127	0.011	36

Support for sex-effect on the parameters of the stationary Markov emigration model was very strong. Of the 24 models fitted, the best 12, with a combined Akaike weight of 0.998, had different movement probabilities for males and females. The best four models, with a combined Akaike weight of 0.894, included sex-, time-, and length-effects on survival and a sex-effect on movement probabilities. In the top three models, with a combined Akaike weight of 0.836, the length-effect was different for males and females. Evidence for a sex-effect on capture was more equivocal. The best model (Akaike weight = 0.396) did not have a sex-effect on capture but it was present in the next best model (Akaike weight = 0.251).

Survival estimates for male and female trout of average length (480mm) in 1987, 1988, 1989 were model-averaged (Buckland, Burnham and Augustin (1997)) across the top four models in Table 3. Survival probabilities appear to have declined during the study and survival probabilities appear higher for females than for males (Figure 2). To assess the effect of length at first capture

on survival, estimates were averaged across the top four models in Table 3. For males, each one standard deviation in length (approximately 45mm) multiplied the odds of survival by an estimated factor of 0.96 (95% CI = 0.76, 1.22), and for females by an estimated factor of 1.60 (95% CI = 1.27, 2.03). Thus for females there is good evidence that survival probability increases with length, but the evidence is equivocal for males.
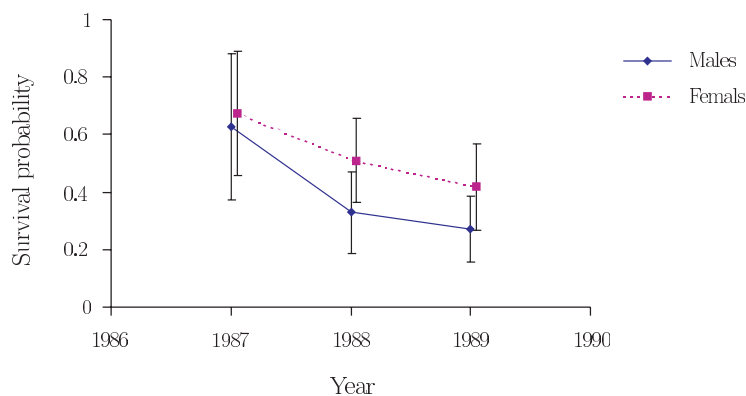

Figure 2. Model-averaged annual survival probability estimates and 95% confidence interval bars for male and female brown trout of average length (480mm) when tagged in Lake Brunner tributaries 1987−1990.

The movement parameters in the Markov emigration model have a useful biological interpretation. Averaged across the best four models, our estimates suggest that male trout spawning in year i have a lower chance of spawning the following year ($\hat{F} = 0.100, se = 0.022$) compared to trout which did not spawn in year i ($\hat{F}' = 0.350, se = 0.112$), with a 95% CI for the difference of $(0.050, 0.450)$. For females the estimated probability of spawning in year $i + 1$ for trout that spawned in year i ($\hat{F} = 0.203, se = 0.030$) was similar to the estimated probability of spawning in $i + 1$ for trout that did not spawn in year i ($\hat{F}' = 0.215, se = 0.063$) with a 95% CI for the difference of $(-0.087, 0.112)$.

4. Discussion

The models described here are new, although they bring together aspects of other published models for mark-recapture data. First, they allow data from live resighting during the open period between capture occasions, as well as live recaptures and dead recoveries, thereby extending Burnham's (1993) model. A key feature of this extension is the need to discriminate between animals that are resighted during the interval between capture occasions and survive that interval, and those animals that are resighted during the interval but, unknown to the

researcher, later die. This model was described by Barker (1997) for the special case of random temporary emigration, however the extension to allow permanent and Markovian emigration generalizes the model of Barker (1997). The modeling of temporary emigration builds on ideas represented in the multistate mark-recapture models reported by Brownie, Hines, Nichols, Pollock and Hestbeck (1993) and Schwarz, Schweigert and Arnason (1993), however differs from these models in that animals can migrate to a stratum in which no capture may occur. More importantly, live resightings of animals during the open period between capture occasions are not accommodated in either of these published multistate models.

In the Lake Brunner trout study, fish were exposed to angling throughout the lake system. Therefore, temporary migration need not be considered for the resighting component of the model. In some studies, animals might be able to move to areas where they cannot be resighted. Modeling movement in the resighting component of the model would then require additional information. This situation parallels the Cormack-Jolly-Seber model where, in the absence of extra information, temporary migration parameters are confounded with other parameters in the model. We expect however that in most studies involving resightings, animals will be available for resighting or dead recovery through most of their geographic range, but that the intensity of resighting effort may vary geographically. This situation occurs for band recovery models, where animals are assumed to be exposed to the recovery process everywhere. For band recovery models, heterogeneous recovery rates have been shown to have little influence on bias and confidence interval coverage rates for survival rate estimators provided survival probabilities do not vary between individuals (Barker (1992), Nichols et al. (1982), Pollock and Raveling (1982)). Although we have not examined the effect of heterogeneous reporting and resighting rates for our model, we anticipate a similar result.

Our analysis of the Lake Brunner trout data indicates that Markovian temporary emigration is needed in the model to estimate survival probabilities correctly. This illustrates the importance of including the live resighting and dead recovery data, as Markovian emigration can only be modeled if tag recoveries from dead animals, or resightings of live animals, are included. In addition to allowing better estimates, the expansion of the model to allow Markovian emigration provides insight into the breeding behavior of the brown trout population. We have evidence that male trout that spawn in year i have a lower chance of spawning the following year compared to trout which don't spawn in year i. For females there was no evidence of a difference, and the interval estimate suggests that any differences, if present, will not be large (< 0.1, say). This apparent asymmetry between male and female spawning behavior could be related to the

different investment that the two sexes make in each spawning effort. Female trout reach the spawning grounds before the males and take sole responsibility for nest construction (McDowall (1978, Chap. 7)). In proportion to body mass, females produce a large number of relatively large eggs in contrast to the smaller mass of milt produced by males.

This increased richness of models for analyzing mark-recapture data should be welcomed by biologists as it allows the construction of more realistic models. In our example it has led to estimates of movement parameters that provide insight into trout breeding biology. A disadvantage of this increased complexity is that model-selection will tend to be more laborious as there are many more potential models to consider. Therefore, efficient model selection algorithms, such as those based on Akaike's Information Criterion (Burnham and Anderson (1998)), are needed.

Although anyone proficient in computer programming can write code for fitting the models we have described, code that is general enough to allow full use to be made of individual covariates, and that is relatively easy for researchers to use, is time-consuming to write. The models we have described are incorporated into program MARK (White and Burnham (1999), White, Burnham and Anderson (2001)). MARK allows users to model parameters using an appropriate choice of link function and linear predictor function that can incorporate individual covariates.

The ability to include individual covariates is an important feature of the models we have described. Not only can individual covariates be used to consider interesting biological effects, they can also be used to relax key assumptions in the model. For example, it is possible to use individual covariates to introduce a short-term marking effect on survival by modeling survival as $S'_{ij} = (1, x_{ij})\boldsymbol{\beta}_i$ where S'_{ij} is the appropriately transformed survival probability for animal j between times i and $i + 1$, $\boldsymbol{\beta}_i$ is a 2-dimension parameter vector, and x_{ij} is 1 if the animal is marked and released for the first time at time i, is 0 otherwise.

Including resighting and dead recovery data greatly increases the number of parameters in the model. At first glance this may also appear to be a disadvantage. In our example, the fully sex- and time- dependent model with no individual covariates has 62 parameters. The equivalent Cormack-Jolly-Seber (CIS) model with six sampling periods has just 18 parameters. An important point about the increased number of parameters in the joint models is that they do not lead to reduced precision of survival rate estimates when compared to the equivalent live-recapture or tag-recovery model. This is true even without selection of a simplified model for the resighting and dead recovery processes. For the type of likelihood functions used in the model, the additional data cannot lead to reduced precision of survival probability estimates and should lead

to an increase (Barker and Kavalieris (2001)). This is because the equivalent live-recapture model (the CIS model) and tag-recovery models are special cases of the joint model considered here. In both models the survival probability estimates can be obtained without the additional information, although in the CIS model with permanent emigration, survival rate cannot be estimated. Therefore relative efficiency is undefined. The extra parameters are only needed to model the additional data. If this modeling is done poorly, say because there are many extra parameters, the marginal benefit of including the extra data will be small but greater than zero.

Acknowledgements

Appendix. Computation of the Probability An Animal is Not Encountered Again Following the Last Encounter

To compute the probability of never encountering an animal again we require α_{ij} and β_{ij}. These are

$$\alpha_{ij} = [1\ 0]\Big\{ \prod_{h=i}^{j-2} \boldsymbol{\psi}_h (\mathbf{1} - \boldsymbol{p}_{h+1}) \Big\} \times \boldsymbol{\psi}_{j-1} \boldsymbol{p}_j \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and}$$

$$\beta_{ij} = [1\ 0]\Big\{ \prod_{h=i}^{j-1} \boldsymbol{\psi}_h (\mathbf{1} - \boldsymbol{p}_{h+1}) \Big\} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

(a) Animal was last encountered by capture

If the animal was not released following capture at i, Pr(never encountered after $i\,|$ last captured at i) $= 1$. Otherwise, we need Pr(never encountered after $i\,|$ last captured at i and released). Following release at i, the next encounter can either be a capture at one of the times $i+1, i+2, \ldots, t$, or a resighting (alive or dead) in one of the intervals $[i, i+1), \ldots, [L, L+1)$. For a particular animal, let C_j denote the event 'next encounter is by capture at j', R_j the event that 'next encounter is by resighting (alive or dead) in $[j, j+1)$', and E_j the event 'encountered at least once at or after j'. For an animal caught and released at time i, $Pr(C_j\,|$ caught and released at $i) = \alpha_{ij} \prod_{h=i}^{j-1} S_h(1 - R_h)$ and $Pr(R_j\,|$ caught and released at $i) = f_j \beta_{ij} \prod_{h=i}^{j-1} S_h(1 - R_h)$. Therefore, if E_j^c is the complement of the event E_j:

$Pr($animal is not encountered after $i\,|$ captured at i and released$) = Pr(E_{i+1}^c\,|$ captured at i and released$)$

$$= 1 - f_i - \sum_{j=i+1}^{T} \left\{ \alpha_{ij} \prod_{h=i}^{j-1} S_i(1-R_i) \right\} - \sum_{j=i+1}^{L} \left\{ f_j\beta_{ij} \prod_{h=i}^{j-1} S_h(1-R_h) \right\}.$$

(b) Animal is last encountered by resighting or dead recovery in $[i, i+1)$

If the animal was last encountered by a dead recovery in $[i, i+1)$, the probability that it is not encountered again is 1. Otherwise, the probability that it is never encountered again depends on when it was last captured (which we index by time g). This is because the probability of later recaptures depends on whether the animal was at risk of capture at time i. For example, if the animal was at risk of capture at i, the probability that it is caught at $i+1$ is $S_iF_ip_{i+1}$ but, if it is not at risk of capture at i, the probability is $S_iF_i'p_{i+1}$. The two probabilities are equal only if emigration is random.

For an animal last encountered by resighting in $[i, i+1)$,

$Pr(E_{i+1}\,|$encounter history in $[k, i+1)$

$= Pr(E_{i+1}\,|$last captured at g, resighted in $[i, i+1))$

$= \dfrac{Pr\big(E_{i+1} \text{ and not captured between } g \text{ and } i\,|\, \text{last captured at } g, \text{ resighted in } [i, i+1)\big)}{Pr(\text{not captured between } g \text{ and } i\,|\, \text{last captured at } g, \text{ resighted in } [i, i+1))}$

$= \dfrac{Pr\big(E_{i+1} \text{ and not captured between } g \text{ and } i\,|\, \text{last captured at } g, \text{ resighted in } [i, i+1)\big)}{\beta_{gi}}.$

To compute the probability of event E_{i+1} for an animal resighted alive in $[i, i+1)$, we also require the probability that the animal is still alive at time $i+1$ given that it was resighted alive in $[i, i+1)$. This is not S_i but rather $\phi_i = (S_iR_i)/[S_iR_i + (1-S_i)(1-r_i)R_i']$.

Partitioning the compound event E_{i+1} into the set of disjoint events making up all possible re-encounter events, we find that for an animal last encountered by resighting in $[i, i+1)$:

$Pr(E_{i+1}^c\,|$ last captured at g, resighted in $[i, i+1))$

$$= 1 - \frac{\phi_i}{\beta_{gi}} \left\{ \sum_{j=i+1}^{T} \left\{ \alpha_{gi} \prod_{h=i}^{j-1} S_i(1-R_i) \right\} - \sum_{j=i+1}^{L} \left\{ f_i\beta_{gi} \prod_{h=i}^{j-1} S_i(1-R_i) \right\} \right\}.$$

References

Barker, R. J. (1992). Effect of heterogeneous survival on bird-banding model confidence interval coverage rates. *J. Wildlife Management* **56**, 111-116.

Barker, R. J. (1995). Open population mark-recapture models including ancillary sightings. Ph.D. thesis, Massey University, Palmerston North, New Zealand.

Barker, R. J. (1997). Joint modelling of live-recapture, tag-resight, and tag-recovery data. *Biometrics* **53**, 666-677.

Barker, R. J. (1999). Joint analysis of mark recapture, resighting and ring-recovery data with age- dependence and marking-effect. *Bird Study* **46** (supplement), 82-91.

Barker, R. J. and Kavalieris, L. (2001). Efficiency gain from auxiliary data requiring additional nuisance parameters. *Biometrics* **57**, 563-566.

Brownie, C., Hines, J. E., Nichols, J. D., Pollock, K. H. and Hestbeck, J. B. (1993). Capture-recapture studies for multiple strata including non-Markovian transitions. *Biometrics* **49**, 1173-1187.

Buckland, S. T., Burnham, K. P. and Augustin, N. H. (1997). Model selection: an integral part of inference. *Biometrics* **53**, 603-618.

Burnham, K. P. (1993). A theory for combined analysis of ring recovery and recapture data. In *Marked Individuals in Bird Population Studies* (Edited by J.-D. Lebreton and P. North), 199-213. Birkhauser Verlag, Basel.

Burnham, K. P. and Anderson, D. R. (1998). *Model Selection and Inference: a Practical Information Theoretic Approach.* Springer-Verlag, New York.

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

Catchpole, E. A., Freeman, S. N., Morgan, B. J. T. and Nash, W. J. (2001). Abalone I: analyzing mark-recapture-recovery data incorporating growth and delayed recovery, *Biometrics* **57**, 469-477.

Clobert, J. (1995). Capture-recapture and evolutionary ecology − a difficult wedding? *J. Appl. Stat.* **22**, 989-1008.

Cormack, R. M. (1964). Estimates of survival from the sighting of marked animals. *Biometrika* **51**, 429-438.

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

Kendall, W. L., Nichols, J. D. and Hines, J. E. (1997). Estimating temporary emigration and breeding proportions using capture-recapture with Pollock's robust design. *Ecology* **78**, 563-578.

Kendall, W. L., Pollock, K. H. and Brownie, C. (1995). A likelihood-based approach to capture-recapture estimation of demographic parameters under the robust design. *Biometrics* **51**, 293-308.

Lebreton, J.-D. (1995). The future of population dynamic studies using marked individuals: a statistician's perspective. *J. Appl. Stat.* **22**, 1009-1030.

Lebreton, J.-D. Morgan, B. J. T. Pradel, R. and Freeman, S. N. (1995). A simultaneous survival rate analysis of dead recovery and live recapture data. *Biometrics* **51**, 1418-1428.

McDowall, R. M. (1978). *New Zealand Freshwater Fishes: a Guide and Natural History.* Heinemann, Auckland.

Nichols, J. D., Stokes, S. L., Hines, J. E. and Conroy, M. J. (1982). Additional comments on the assumption of homogeneous survival rates in modern bird banding estimation models. *J. Wildlife Management* **46**, 953-962.

Pollock, K. H. (1982). A capture-recapture design robust to unequal probability of capture. *J. Wildlife Management* **46**, 757-760.

Pollock, K. H. and Raveling, D. G. (1982) Assumptions of modern band-recovery models, with emphasis on heterogeneous survival rates. *J. Wildlife Management* **46**, 88-98.

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

Schwarz, C. I. and Stobo, W. T. (1999). Estimation and effects of tag-misread rates in capture-recapture studies. *Canad. J. Fisheries and Aquatic Sci.* **56**, 551-559.

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

Seber, G. A. F. (1982). *The Estimation of Animal Abundance and Related Parameters.* Griffin, London.

White, G. C. and Burnham, K. P. (1999). Program MARK: Survival estimation from populations of marked animals. *Bird Study* **46** (Supplement), 120-139.

White, G. C., Burnham, K. P. and Anderson, D. R. (2001). Advanced features of program MARK. In *Wildlife Land and People: Priorities for the* 21*st Century* (Edited by R. Fields, R. J. Warren, H. Okarma and P. R. Sievert), 368-377. Proceedings of the Second International Wildlife Management Congress. The Wildlife Society, Bethesda, Maryland.

Department of Mathematics and Statistics, University of Otago, P.O. Box 56, Dunedin, New Zealand.

E-mail: rbarker@maths.otago.ac.nz

Colorado Cooperative Fish and Wildlife Research Unit, U.S. Geological Survey, Colorado State University, Fort Collins, Colorado 80523, U.S.A.

E-mail: kenb@lamar.colostate.edu

Department of Fishery and Wildlife Biology, Colorado State University, Fort Collins, Colorado 80523, U.S.A.

E-mail: gwhite@cnr.colostate.edu