

MARK-RECAPTURE WITH TAG LOSS

by

Laura Louise Elizabeth Cowen

M.Math (Statistics), University of Waterloo, 2000

B.Sc.(Biological Science), Simon Fraser University, 1997

A THESIS SUBMITTED IN PARTIAL FULFILMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
in the Department
of
Statistics and Actuarial Science

© Laura Louise Elizabeth Cowen 2005
SIMON FRASER UNIVERSITY
Spring, 2005

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without the permission of the author.

APPROVAL

Name: Laura Louise Elizabeth Cowen

Degree: Doctor of Philosophy

Title of thesis: Mark-Recapture with Tag Loss

Examining Committee: Dr. Richard Lockhart
Chair

Dr. Carl J. Schwarz
Senior Supervisor

Dr. Brad McNeney
Supervisory Committee

Dr. Tim B. Swartz
Supervisory Committee

Dr. Boxin Tang
Internal External Examiner

Dr. John R. Skalski
External Examiner
University of Washington

Date Approved: _____

Abstract

Mark-recapture studies are used to estimate population parameters such as abundance, survival and recruitment. Briefly, animals are captured, marked with an individually identifiable tag and released. First capture provides information about abundance. Subsequent recaptures provided survival information about the individual.

One of the fundamental assumptions in mark-recapture studies is that tags are not lost. If this assumption is violated, parameter and standard error estimates are biased. This thesis deals with the analysis of 3 mark-recapture experiments under tag-loss.

The second chapter looks at premature radio failure in radio-telemetry studies. Radio-tags, because of their high detectability, are often used in capture-recapture studies. A key assumption is that radio-tags do not cease functioning during the study. Radio-tag failure before the end of a study can lead to underestimates of survival rates. We develop a model to incorporate secondary radio-tag failure data. This model was applied to chinook smolts (*Oncorhynchus tshawytscha*) on the Columbia River, WA.

The third chapter incorporates tag loss into the Jolly-Seber model. Tag loss in the Jolly-Seber model has only been dealt with in an *ad hoc* manner. We develop methodology to estimate population sizes and tag-retention in double-tagging mark-recapture experiments. We apply this methodology to the study of walleyes (*Stizostedion vitreum*) in Mille Lacs, Minnesota.

Finally, in the fourth chapter, we develop a Poisson migration model incorporating tag loss. This model is applied to the study of yellowtail flounder (*Limanda ferruginea*) on the Grand Banks of Newfoundland.

*To Jennie Burritt who taught me how to count and
Vivian Cowen who encourages me to keep counting.*

“Keep on swimming,

Keep on swimming,

Keep on swimming,

swimming,

swimming.”

— Dory the Fish

In the film *Finding Nemo*, ©Disney/Pixar 2003, by permission.

Acknowledgments

First and foremost I would like to thank my senior supervisor Dr. Carl Schwarz. He continued to push me beyond what I thought were my limits, and offered advice academically and personally whenever needed. Carl's commitment to his students is genuine and his open-door policy is extremely appreciated.

I am grateful for the support that came from the faculty and staff of the Department of Statistics and Actuarial Science of Simon Fraser University. In particular, Dr. Tim Swartz offered help with many of my Bayesian queries and Dr. Richard Lockhart helped sift through the details of conditional likelihood theory. I am indebted to Dr. Rick Routledge as he encouraged me first as an undergraduate, and later as a graduate student, to continue my studies in statistics.

LGL Limited assisted in providing funding in the form of an NSERC Industrial Scholarship. The entire staff at LGL is first-rate and a pleasure to work with. I am thankful to have had the opportunity to exercise my consulting skill with them.

Many friends and fellow graduate students encouraged me along the way through different means. In particular I would like to recognize the following people: Diena Abdurahman, H el ene Bolduc, Christian Boudreau, Malcolm Coupe, Gail Fraser, Qing Guo, Hugh Knechtel, Jason Loepky, Wen Lu, Michelle Lurkins, Tara McFarlane, Trity Pourbahrami, Deena Rubuliak, John Ryder, Victor Sasan, Susie Silversides, Christine Stoddard, Elly Tompe, Steve Quattrochi, Rick Witzke and Gord Witzke.

I am indebted to my parents Emily and Kirk Cowen along with the rest of my family. Their ongoing encouragement enabled me to complete this degree.

Lastly, I would like to express my appreciation towards Alf Wilson whose patience and respect are exemplary.

Contents

Approval	ii
Abstract	iii
Dedication	iv
Quotation	v
Acknowledgments	vi
Contents	vii
List of Tables	xi
List of Figures	xv
1 Capture-Recapture Methodology	1
1.1 Background	1
1.2 Experimental Protocol, Data, Assumptions and Models	2
1.2.1 Experimental Protocol	2
1.2.2 Data	3
1.2.3 Assumptions	4
1.2.4 Models	5
1.3 Model Specification, Fitting, Selection, Goodness of Fit, and Overdispersion.	8

1.3.1	Model Specification	8
1.3.2	Model fitting	14
1.3.3	Model Selection	16
1.3.4	Goodness of Fit	19
1.3.5	Over-dispersion	21
1.4	Poisson versus Multinomial Models	22
1.5	Bayesian Methods	23
1.6	Tag Loss	24
1.7	Outline of Thesis	25
1.7.1	Chapter 2	26
1.7.2	Chapter 3	26
1.7.3	Chapter 4	26
1.7.4	Appendices	27
2	Premature Radio-Tag Failure	28
2.1	Design Protocol	29
2.2	Notation	30
2.3	Classical Model Development	32
2.3.1	Likelihood Development	32
2.3.2	Parameter Estimation	34
2.3.3	Standard Error Estimation	35
2.3.4	Model Selection	35
2.3.5	Goodness of Fit	36
2.4	Bayesian Model Development	36
2.4.1	Complete Data Likelihood	36
2.4.2	Priors	37
2.4.3	Full Conditionals and Latent Data	38
2.5	Example	41
2.6	Simulation	49
2.7	Discussion	55

3	The Jolly-Seber Model with Tag Loss	56
3.1	Design Protocol	57
3.2	Notation	57
3.3	Model Development	60
3.3.1	Likelihood Development	60
3.3.2	Parameter Estimation	62
3.3.3	Per Unit Time	67
3.3.4	Goodness of Fit	67
3.3.5	Model Selection	68
3.4	Example	69
3.5	Discussion	74
4	A Migration Model with Tag Loss	76
4.1	Notation	77
4.2	Design Protocol	78
4.3	Model Development	79
4.4	Parameter and Standard Error Estimates	80
4.4.1	Goodness of Fit	81
4.4.2	Over-Dispersion	81
4.4.3	Model Selection	82
4.5	Example	83
4.6	Discussion	92
5	Conclusions and Future work	94
5.1	Summary	94
5.2	Future Work	97
A	Conditional Likelihood Approach	100
A.1	2 Dam Case	101
A.2	k Dam Case	103
A.3	Simulation	106

B	Expectations of Latent Variables	108
B.1	$E(a_{ij})$	108
B.2	$E(a_{ij}a_{i,j+1})$	109
B.3	$E(a_{ij}g_{ij+})$	109
B.4	$E(a_{i,j+1}g_{ijd}g_{i,j+1,d})$	110
B.5	$E(g_{ijd}a_{i,j+1})$	110
C	JSTL Simulation Results	112
D	Mille Lac Data	114
	Bibliography	118

List of Tables

1.1	The m_{ij} -array representation for data from a 4 period capture recapture study. Taken from Williams et al. 2001, p.149. R_i is the number of tagged fish released at time i and m_{ij} is the number of tagged fish released at time i , next seen (captured or resighted) at time j	4
1.2	Capture histories and associated probabilities for the 3-sample Cormack-Jolly-Seber model.	7
1.3	Capture histories and associated probabilities for the 3-sample Jolly-Seber model.	8
1.4	Common link functions for a parameter θ used in mark-recapture analyses.	12
1.5	The 2×2 contingency table for the first component of the goodness-of-fit test described by Pollock et al. 1990.	19
1.6	A 2×2 contingency table for the second component of the goodness-of-fit test described by Pollock et al. 1990.	20
2.1	A sample of the capture histories and cumulative time for the following sampling points: Rock Island tailrace, Wanapum Dam, Wanapum tailrace, Priest Rapids dam, Priest Rapids tailrace, Hanford Reach 1. Complete data are available at the Biometrics website.	30
2.2	Radio-tag survival data. Note that the last failure (*) was changed to a censored time so that radio-survival estimates would not go to zero.	45

2.3	Classical parameter estimates and bootstrap standard errors(SE) (100 replicates) for the radio-tag failure model, Bayesian mean estimates and posterior standard deviations (SD), survival estimates assuming 100% detectability, and Cormack-Jolly-Seber (CJS) parameter estimates and standard errors.	46
2.4	Parameter estimates and standard errors (SE) for the classical tag-life adjusted model, Bayesian mean estimates and posterior standard deviations (SD), survival estimates assuming 100% detectability and Cormack-Jolly-Seber (CJS) estimates and standard errors when probability of recapture was simulated at 80%.	50
2.5	Small sample size effects of the number of radio-tags (m=25 or 100) used to estimate radio-tag survival and the number of fish (n=100 or 400) released. Actual parameter values used to simulate the data were $\phi_1 = 0.9$, $p_2 = 0.8$, $\phi_2 p_3 = 0.72$, $\mu_1 = \mu_2 = 5$, $\lambda_1 = \lambda_2 = 25$, $\nu = 10$, $\sigma^2 = 25$. Parameter estimates and bootstrap standard errors (SE) of 100 replicates are shown for the classical radio-tag failure model. . . .	51
2.6	Simulation to determine the effect of changing the mean radio-tag failure time ($\nu = 10, 15, 20, 25$). The number of fish released was 1000, the number of radio-tags used to create the failure curve was 25. Other parameters included: $\phi_1 = 0.9$, $p_2 = 0.8$, $\phi_2 p_3 = 0.72$, $\mu_1 = \mu_2 = 5$, $\lambda_1 = \lambda_2 = 5$, $\sigma^2 = 25$. Parameter estimates and bootstrap standard errors (SE) of 100 replicates are shown for the classical radio-tag failure model (RTF). Cormack-Jolly-Seber (CJS) parameter estimates and standard errors are provided for comparison. Standard errors are shown below associated estimates.	53

2.7	Sensitivity analysis for the violation of the assumption that fish survival and travel time is independent. The number of fish released was 10000, the number of radio-tags used to create the failure curve was 1000. Other parameters included: $p_2 = p_3 = 0.8$, $\nu = 10$, $\sigma = 5$, $\mu_1 = \mu_2 = 5$, $\xi = 1$. Percent bias of parameter estimates for the model are shown for 5 different distributions of survival, $\phi(t)$, where the approximate average survival rate was held at 0.5.	54
3.1	Examples of statistic and latent variables for double-tagged fish in a 3 sample experiment.	63
3.2	Summary statistics and standardized sample times for the 6 periods of the Mille Lacs data.	71
3.3	The number of double-tagged fish released and subsequent recaptures with either double (first row) or single (second row) tags present. Cumulative tag-loss rates based on the methods presented in Seber and Felton (1981) for 2 indistinguishable tags are provided (third row). . .	72
3.4	Log-likelihood (l), Akaike information criterion (AIC), the number of model parameters (P), Δ AIC and, population estimates (\hat{N}) and standard errors for several different models. Both \hat{N}_i and $SE(\hat{N}_i)$ are reported in thousands of fish.	73
4.1	Tag type classifications used in the yellowtail flounder tagging study.	78
4.2	The number of yellowtail flounder released and recovered by tag type, period and strata. Tag types are as follows: 1) single-tagged fish, 2) double-tagged fish recovered with a single tag, 3) double-tagged fish, 4) single-tagged fish with a high reward tag. Years are divided into 2 periods, January-June and July-December. Fish are released in the first period of each year (1,3,5,7,and 9). Note that there were no double-tagged flounder released in period 9 (2004).	86
4.3	Results of the instantaneous tagging mortality experiments where yellowtail flounder were kept in cages for 10 days (19 days in 2001). . . .	88

4.4	Summary of models fit to the yellowtail flounder data. Over-dispersion estimates $\hat{\psi}$ and $\tilde{\psi}$ are estimated from squared Pearson residuals divided by $n - K$ and deviance divided by $n - K$ respectively (where n is the degrees of freedom of the saturated model, the number of observed counts). The number of parameters in the model is K ; QAIC (adjusted for over-dispersion using $\hat{\psi} = 2.010$) and Δ_i are also shown.	89
4.5	Maximum likelihood estimates and standard errors inflated by the square root of the over-dispersion estimate $\hat{\psi} = 2.010$ for model $[p_s, \phi_{st}, \Lambda, Q_s, \lambda]$. The unit for the parameters is the 6-month periods.	90
4.6	Outlying residuals found in plots of the Pearson residual versus expected recovery for model $[p_s, \phi_{st}, \Lambda, Q_s, \lambda]$	90
A.1	Conditional and unconditional maximum likelihood estimates of λ (the mean of the exponential distribution), the estimated standard error and the coefficient of variation for 2, 3, and 4 dams, 10, 30, and 100 radios, and 100, 500, and 1000 fish. The true value of $\lambda = 0.067$	107
C.1	Simulation results for a 3 sample experiment where $p = 0.8$, $\phi = 0.8$, $\Lambda = 0.8$, $\nu = 0.1$, and $b_0^* = 0.5, b_1^* = 0.2$. Note that p_1 and p_3 were fixed to 0.8 due to confounding.	113

List of Figures

1.1	Probabilities associated with a 2-sample experiment where S_1 is the probability of survival from time 1 to time 2 and p_2 is the capture probability at time 2.	6
2.1	Map of the Columbia River showing Rock Island dam, Wanapum dam, Priest Rapids dam and Hanford Reach. Source ©LGL Limited, Sidney, British Columbia, by permission.	42
2.2	Diagram of the experimental design where fish are ‘released’ (radio-tags detected) at sampling points 1, 3, and 5 and recaptured at sampling points 2, 4, and 6. The time since release is shown as t_1, t_2, \dots, t_6	44
2.3	Probability plots to test the goodness of fit of the inverse gaussian distribution modelling the travel time of fish from a) dam 1 to dam 2 (release 1), b) dam 3 to dam 4 (release 2), c) dam 5 to dam 6 (release 3). $F(time)$ (the cumulative distribution function of the inverse gaussian evaluated at the radio-tag failure estimates) is plotted against the empirical cumulative distribution function (ECDF) defined as $\frac{i-0.5}{n}$, $i = 1, 2, \dots, n$, where n is the number of fish released.	48
4.1	Map of the strata boundaries and release points for yellowtail flounder on the Grand Banks of Newfoundland. Release points 26-31 were added in 2001. Release points 32-34 were not included in this analysis. Map ©Fisheries and Oceans Canada by permission, provided by Stephen J. Walsh, Northwest Atlantic Fisheries Center, Newfoundland.	84

4.2	Frequency of release and recovery by month from January 2000 to December 2004.	85
4.3	Pearson residuals versus expected recoveries for model $[p_s, \phi_{st}, \Lambda, Q_s, \lambda]$	91

Chapter 1

A Brief Review of Capture-Recapture Methodology

1.1 Background

Capture-recapture methods are used in the estimation of demographic parameters such as population size, birth rates, survival rates, migration rates and exploitation rates. The problem of estimating population size dates back at least as far as 1786 when Laplace (1786) attempted to estimate the population of France (Buckland, Goudie and Borchers 2000; Chapman, 1948).

The simplest experiment involves the use of the Peterson estimator. First, n_1 animals are sampled, marked and released back into the population. After a certain amount of time, a second sample of size n_2 is drawn and the number of marked animals m_2 is counted. The marked fraction in the second sample allows for the estimation of the initial population size:

$$\hat{N} = \frac{n_1 n_2}{m_2} \tag{1.1}$$

The Peterson estimator (1.1) does not account for mortality between sampling times (if there is only mortality \hat{N} estimates the initial population size, whereas if there are only births \hat{N} estimates the final population size), nor is there provision to

estimate other population parameters such as survival between time periods.

Capture-recapture models are broken down into two major categories – open and closed population models. Closed population models assume that population size does not change over the study period. For a thorough discussion of closed population models see Otis et al, (1978), Williams, Nichols and Conroy (2001), and Chao (2001). As I only deal with open population models in this thesis, the remainder of this review will consider only such models.

Open population models do not assume that population size is constant over the study. There are several approaches to modeling open populations. Cormack (1964), Jolly (1965) and Seber (1965) introduced a multinomial framework for modelling multiple-sample capture-recapture when population size is not of interest but removals from the population (deaths) can occur. This is known as the Cormack-Jolly-Seber (CJS) model. For a thorough overview of CJS models see Lebreton et al. (1992), Pollock et al. (1990), and Williams et al. (2001). The Cormack-Jolly-Seber model is a subset of the more general Jolly-Seber (JS) model formalized by Jolly (1965) and Seber (1965). The Jolly-Seber model incorporates estimates of recruitment of new animals to the population to estimate population size.

1.2 Experimental Protocol, Data, Assumptions and Models

1.2.1 Experimental Protocol

Capture-recapture experiments occur over k distinct sampling occasions. At each occasion animals are captured, tagged with unique identifiers and released. Subsequent recaptures can occur either via physical recapture or through observation. When animals are recaptured their tag information is recorded and if alive, the animals are released. Ring-recovery models deal with experiments where animals are tagged and released, but recovered dead (due to hunting or harvesting). As there is reliance on the harvester to return tags, reporting rates are considered in ring-recovery models.

It is assumed that reporting rates are independent of the time of release. As reporting is confounded with exploitation, it is not estimable in itself without the use of reward bands (Brownie et al. 1985). Another key difference between ring-recovery and capture-recapture models is the assumption of instantaneous sampling. Ring-recovery models assume recoveries over a period of time; capture-recapture models assume instantaneous sampling.

1.2.2 Data

The basic unit of data is the capture history for each animal. This is a k -element vector where at each sampling occasion the component is given a ‘1’ if the animal was captured or a ‘0’ if the animal was neither observed nor captured. For example, in a 3-sample experiment, an animal with a capture history of {101} was captured, tagged and released at time 1, not observed at time 2, and recaptured at time 3.

If migration is involved in the model, then stratification is used to separate the areas of migration. Here the history vector is generalized by giving each component a number corresponding to the stratum where it was captured or a ‘0’ if it was not observed. For example, an animal with a capture history of {302} would be captured, tagged and released in stratum 3 at time 1, not captured at time 2 and recaptured in stratum 2 at time 3.

If possible, capture histories are summarized such that each capture history is paired with a frequency of occurrence. A common convention is to use negative frequencies to indicate that the animal was not released following its last capture. This typically occurs due to tagging mortality or harvest.

It is possible to further summarize the data into the number of releases at each sample occasion and the number of recaptures at each period. This representation is known as the m_{ij} -array representation and is shown in Table 1.1 (Lebreton et al. 1992; Williams et al. 2001). For example, the history {10110} would contribute to the R_1 and m_{13} values; the R_3 and m_{34} values and the R_4 values. Be that as it may, recent modelling efforts deal directly with capture histories rather than try to reduce to sufficient statistics. This has occurred for two reasons. First, models have become

Table 1.1: The m_{ij} -array representation for data from a 4 period capture recapture study. Taken from Williams et al. 2001, p.149. R_i is the number of tagged fish released at time i and m_{ij} is the number of tagged fish released at time i , next seen (captured or resighted) at time j .

	Recapture Period j		
Releases in Period i	2	3	4
R_1	m_{12}	m_{13}	m_{14}
R_2		m_{23}	m_{24}
R_3			m_{34}

so complex that the capture histories are the sufficient statistic; second, using capture histories is more intuitive when describing likelihoods for models. Consequently, the use of summary statistics such as the m_{ij} array has been replaced by direct consideration of the capture histories. One of the earliest examples is when Cormack (1989) modelled capture histories with log-linear models.

1.2.3 Assumptions

Assumptions for the Cormack-Jolly-Seber model are as follows (Williams et al. 2001).

- 1** All animals present at sample time i have the same probability of being captured.
- 2** All animals present immediately following sample time i have the same probability of surviving to sample time $i + 1$.
- 3** No tags are lost and all tags are correctly identified.
- 4** Sampling occurs instantaneously and animals are released immediately.
- 5** Emigration from the sample area is permanent; such emigration is indistinguishable from death.
- 6** The survival and capture of an individual is independent of the survival and capture of all other animals.

The Jolly-Seber model also assumes:

7 Marked and unmarked animals have the same probability of survival and capture.

This thesis is concerned with the violation of assumption 3, in particular, tag loss.

1.2.4 Models

The Cormack-Jolly-Seber Model

The Cormack-Jolly-Seber and Jolly-Seber models are developed from a multinomial framework. For a simple 2-sample CJS experiment, possible capture histories are $\{11\}$, $\{10\}$ and $\{01\}$. The likelihood is constructed conditional upon first release and is given by

$$L \propto P(11)^{n_{11}}P(10)^{n_{10}} \quad (1.2)$$

where $P(11)$ is the probability of observing capture history $\{11\}$, and n_{11} is the number of animals with capture history $\{11\}$. Similar definitions can be formed for $P(10)$ and n_{10} . Because of the conditioning on first release, history $\{01\}$ provides no information in the CJS model for survival rates and is omitted.

At this point it is convenient to introduce some notation.

ϕ_i is the probability that an animal alive at time i survives to time $i + 1$.

p_i is the probability that an animal is captured at time i given that it is alive.

ω_i is an indicator variable given a ‘1’ if the animal was captured and a ‘0’ if the animal was not captured.

b_i is the fraction of the total number of animals that enter the system between time i and $i + 1$.

χ_i is the probability that an animal is not observed after time i , given it was alive at time i .

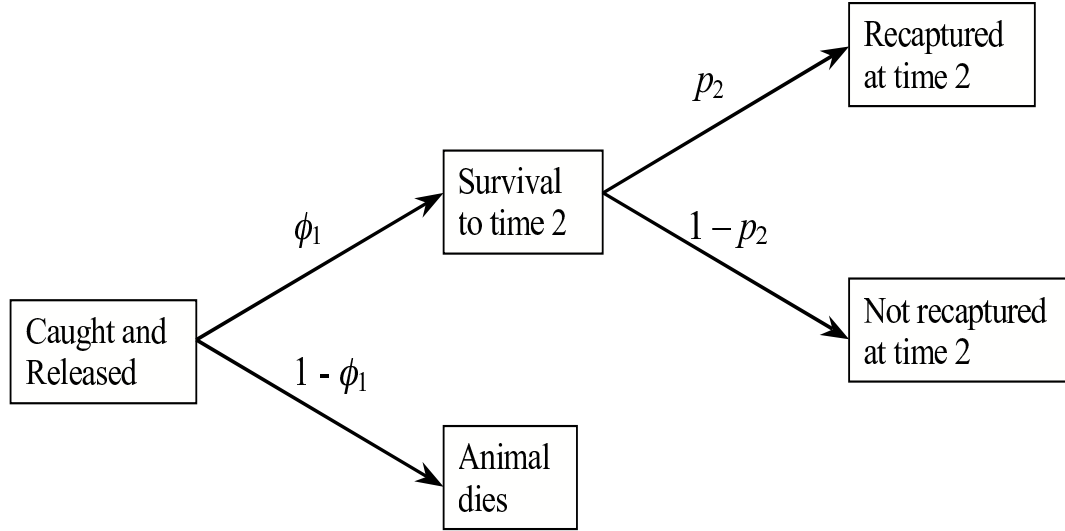


Figure 1.1: Probabilities associated with a 2-sample experiment where S_1 is the probability of survival from time 1 to time 2 and p_2 is the capture probability at time 2.

Ψ_i is the probability that an animal is not captured before time i .

f_j is the first time animal j was observed.

l_j is the last time animal j was observed.

In order to obtain $P(11)$ and $P(10)$ we look at possible outcomes for observing these capture histories (Figure 1.1). The 2-sample CJS likelihood is

$$L(\phi, p) \propto (\phi_1 p_2)^{n_{11}} (1 - \phi_1 p_2)^{n_{10}} \quad (1.3)$$

Table 1.2 shows the capture histories and their respective probabilities for a 3-sample experiment. The CJS likelihood for multiple sample occasions can be written as

$$L(\phi, p) \propto \prod_{\omega} \chi_l \prod_{i=f}^{l-1} \phi_i \prod_{j=i+1}^l p_j^{\omega_j} (1 - p_j)^{1-\omega_j} \quad (1.4)$$

Table 1.2: Capture histories and associated probabilities for the 3-sample Cormack-Jolly-Seber model.

Capture History	Probability
111	$\phi_1 p_2 \phi_2 p_3 \chi_3$
101	$\phi_1 (1 - p_2) \phi_2 p_3 \chi_3$
110	$\phi_1 p_2 \chi_2$
100	χ_1
011	$\phi_2 p_3 \chi_3$
010	χ_2
001	$\chi_3 = 1$

$$\chi_i = \begin{cases} 1 - \phi_i + \phi_i (1 - p_{i+1}) \chi_{i+1} & \text{if } i = 1, 2, \dots, k - 1 \\ 1 & \text{if } i = k \end{cases}$$

The Jolly-Seber Model

The Jolly-Seber model extends the Cormack-Jolly-Seber model to incorporate birth and/or immigration of animals. It can also be extended to incorporate loss on capture (as can the CJS model). The likelihood can be separated into three parts as shown in Darroch (1959), Jolly (1965), Seber (1965) and Schwarz and Arnason (1996). These include $L_1 = P(\text{ first capture })$, $L_2 = P(\text{ recapture } | \text{ capture })$ and $L_3 = P(\text{ loss on capture } | \text{ capture })$ where $L = L_1 \times L_2 \times L_3$. Schwarz and Arnason (1996) use a superpopulation approach following the development of Crosbie and Manly (1985) where animals are born out of a superpopulation into the experiment. They describe the problems with the alternative approaches of Darroch (1959), Jolly (1965), Seber (1965), Cormack (1989) and Burnham (1991). For the rest of this overview, we do not discuss modelling loss on capture any further; details are provided by Schwarz and Arnason (1996).

Once again we are concerned with modelling $P(\omega_i)$. Capture histories and their corresponding probabilities are given in Table 1.3 for a 3-sample Jolly-Seber model. Unlike the CJS model, the JS model is not conditional on first capture. This difference results in the addition of a term (Ψ) representing the probability of first capture; Ψ is a recursive function shown in Table (1.3).

Table 1.3: Capture histories and associated probabilities for the 3-sample Jolly-Seber model.

Capture History	Probability
111	$\Psi_1 p_1 \phi_1 p_2 \phi_2 p_3 \chi_3$
101	$\Psi_1 p_1 \phi_1 (1 - p_2) \phi_2 p_3 \chi_3$
110	$\Psi_1 p_1 \phi_1 p_2 \chi_2$
100	$\Psi_1 p_1 \chi_1$
011	$\Psi_2 p_2 \phi_2 p_3 \chi_3$
010	$\Psi_2 p_2 \chi_2$
001	$\Psi_3 p_3 \chi_3$

$$\Psi_{i+1} = \begin{cases} \Psi_i(1 - p_i)\phi_i + b_i & \text{if } i + 1 > 1 \\ b_0 & \text{if } i + 1 = 1 \end{cases}$$

In order to make use of the capture histories, we first model the number of observed histories (n_{obs}) as a binomial:

$$L^A = \binom{N}{n_{obs}} P(00\dots 00)^{N-n_{obs}} (1 - P(00\dots 00))^{n_{obs}} \quad (1.5)$$

where N is the superpopulation size. Then we model the number of each observed history ($n_{\omega_1}, n_{\omega_2}, \dots, n_{\omega_m}$) as a multinomial conditional on being observed:

$$L^B = \binom{n_{obs}}{n_{\omega_1}, n_{\omega_2}, \dots, n_{\omega_m}} \frac{P(\omega_1)^{n_{\omega_1}} P(\omega_2)^{n_{\omega_2}} \dots P(\omega_m)^{n_{\omega_m}}}{(1 - P(00\dots 00))^{n_{obs}}} \quad (1.6)$$

The likelihood now becomes $L = L^A \times L^B$.

1.3 Model Specification, Fitting, Selection, Goodness of Fit, and Overdispersion.

1.3.1 Model Specification

So far discussion has focussed on simple mark-recapture experiments, with k -sample occasions and one release group. If there are several release groups, parameters are

given an extra subscript to indicate release group (e.g. ϕ_{gi}). Groups correspond to a fixed unchanging attribute of an animal, such as sex. In some cases, parameters may be common across groups; for instance, capture probabilities may depend upon time, but not upon sex.

Lebreton et al. (1992) developed a notation for specifying the relationships among parameters across time and across groups. Parameters are given a subscript representing a variable they are dependent on (e.g. time= t , sex= s , age= a , group= g , etc). A notation similar to that used in specifying models in a generalized linear model framework is then used. Thus the CJS model where survival and capture vary independently over time (but are equal across groups) is referred to as model $[\phi_t, p_t]$. Models where effects of time and groups are additive for a parameter (i.e. effects vary over time, but there is a constant difference at each time point between parameters across groups) are denoted as ϕ_{g+t} . Models where parameters are equal across all time periods and groups are denoted without subscripts, as in ϕ . Finally, models where parameters vary over time independently for each group (commonly referred to as an interaction between group and time effects) are denoted as $\phi_{g,t}$.

A parallel can be drawn with analysis of variance models. For the model with $p_{g,t}$, the corresponding analysis of variance model would be $E(\text{transformation}(p_{gt})) = \mu + \alpha_g + \beta_t + \gamma_{gt}$ (Lebreton et al. 1992). As there is this parallel with analysis of variance modelling, capture-recapture can make use of the idea of design matrices when specifying models.

While the notation of Lebreton et al. (1992) is compact, it is insufficiently general for model specification. For example, there is no compact way to specify a model where the capture rates may be equal at times 1 and 2, but then change over time for the rest of the periods. The mark-recapture community has adopted a more general approach to specifying models using a technique called Parameter Identification Matrices (PIM). This method is the *de facto* standard used by a common computer package for the analysis of mark-recapture experiments (MARK) written by White and Burnham (1999).

The use of PIM will be illustrated for the capture-rates with 2 groups ($G = 2$) and 4 sampling occasions ($K = 4$) in a CJS model. Similar developments occur for the

survival rates in the CJS model and for the JS parameters.

First, to allow for heterogeneity in catchability depending upon when the animal was last captured, the capture rates p are indexed by (i, t, g) where i refers to the last time an animal was captured, t to the current capture time, and g to the group membership. For example, the capture history (1011) in group 1 in a CJS model, where probabilities are conditioned upon release, would have probability $\phi_{111}(1 - p_{121})\phi_{121}p_{131}\phi_{331}p_{341}$ of occurring. The parameters can be arranged into a set of triangular arrays.

$$\begin{array}{ccc|ccc} p_{121} & p_{131} & p_{141} & p_{122} & p_{132} & p_{142} \\ & p_{231} & p_{241} & & p_{232} & p_{242} \\ & & p_{341} & & & p_{342} \end{array}$$

If no structure is to be applied to these parameters, there are a total of 12 separate parameters. Consequently, this model would be specified as:

$$\begin{array}{ccc|ccc} 1 & 2 & 3 & 4 & 5 & 6 \\ & 7 & 8 & & 9 & 10 \\ & & 11 & & & 12 \end{array}$$

Note that neither the actual numbers used to identify the parameters, nor the order of the numbers is important. All that is important is that there are 12 separate numbers specified.

Simpler models are specified by reusing some of the numbers. For example, if capture rates are homogeneous across the time when last released but heterogeneous across groups, the PIM could be specified as:

$$\begin{array}{ccc|ccc} 1 & 2 & 3 & 4 & 5 & 6 \\ & 2 & 3 & & 5 & 6 \\ & & 3 & & & 6 \end{array} \quad (\text{PIM set 1})$$

A model where there are no group effects, but are time effects upon the capture rates could be specified as:

Table 1.4: Common link functions for a parameter θ used in mark-recapture analyses.

Name	Link Function
logit	$\log(\theta/(1 - \theta))$
log	$\log(\theta)$
sin	$\sin(\theta)$
identity	θ

$$\begin{bmatrix} \text{logit}(p_{121}) \\ \text{logit}(p_{131}) \\ \text{logit}(p_{141}) \\ \text{logit}(p_{231}) \\ \text{logit}(p_{241}) \\ \text{logit}(p_{341}) \\ \text{logit}(p_{122}) \\ \text{logit}(p_{132}) \\ \text{logit}(p_{142}) \\ \text{logit}(p_{232}) \\ \text{logit}(p_{242}) \\ \text{logit}(p_{342}) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}$$

In the JS model, we cannot formulate such a general model as for the CJS model because we do not condition on first release. Because of the need to model first capture, JS models traditionally assume homogeneity of capture and survival regardless of when last released. Furthermore, it is not clear how birth parameters would fit into a triangular PIM structure. Thus the PIM for the JS model are single rows.

Consider a 4-sample JS experiment with 2 groups. If we want to use the model $[\phi_t, p_g, b_{g,t}]$ where survival varies over time, capture varies over release group, and birth varies over group and time, the PIM would be:

$$\begin{array}{cccc|cccc}
 \phi : & 1 & 2 & 3 & 1 & 2 & 3 & \\
 p : & 4 & 4 & 4 & 4 & 5 & 5 & 5 & 5
 \end{array}$$

$$b: 6 \ 7 \ 8 \ 9 \mid 10 \ 11 \ 12 \ 13$$

where the left hand vector specifies group 1 and the right hand vector specifies group 2. Note that the columns correspond to the sample times. The corresponding design matrices are:

$$\begin{bmatrix} \text{logit}(\phi_{11}) \\ \text{logit}(\phi_{21}) \\ \text{logit}(\phi_{31}) \\ \text{logit}(\phi_{12}) \\ \text{logit}(\phi_{22}) \\ \text{logit}(\phi_{32}) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

$$\begin{bmatrix} \text{logit}(p_{11}) \\ \text{logit}(p_{21}) \\ \text{logit}(p_{31}) \\ \text{logit}(p_{41}) \\ \text{logit}(p_{12}) \\ \text{logit}(p_{22}) \\ \text{logit}(p_{32}) \\ \text{logit}(p_{42}) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_4 \\ \beta_5 \end{bmatrix}$$

$$\begin{bmatrix} \text{logit}(b_{11}) \\ \text{logit}(b_{21}) \\ \text{logit}(b_{31}) \\ \text{logit}(b_{41}) \\ \text{logit}(b_{12}) \\ \text{logit}(b_{22}) \\ \text{logit}(b_{32}) \\ \text{logit}(b_{42}) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_6 \\ \beta_7 \\ \beta_8 \\ \beta_9 \\ \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \end{bmatrix}$$

Model $[\phi_{g+t}]$ cannot be specified using a PIM, it must be specified using design matrices.

1.3.2 Model fitting

In simple cases, closed form parameter estimates and standard error estimates are available in terms of sufficient statistics from the m_{ij} -array representation. Pollock et al. (1990) provide the maximum likelihood estimates and variance estimates in terms of the sufficient statistics. In more complex models we must rely on numerical methods such as the Newton-Raphson algorithm. We find the estimates of the standard parameters by computing $X\beta$ and taking the anti-logits. The estimated standard errors are found using repeated applications of the delta-method (Taylor series expansion) using the estimated covariance matrix of the beta parameters.

An alternative to the direct maximization of the likelihood or when missing data or latent variables (see later chapters in the thesis) exist we can use an EM algorithm. The EM algorithm has two steps: expectation and maximization. Furthermore, in capture-recapture studies the EM algorithm may decompose the maximization problem into independent subproblems. This becomes important when models have many parameters and maximization routines become unstable. VanDeusen (2002) outlines the EM algorithm for the CJS model.

In the CJS model, the alive status of the animal is a latent variable for sample occasions after the last time observed. For example, capture history $\{110\}$ could correspond to an alive status vector of $\{111\}$ or $\{110\}$. Here a ‘1’ signifies that the animal is alive and ‘0’ signifies that the animal is dead. If we were to know the alive status of all individuals at all sample times, as well as the capture histories, this would constitute complete data for the CJS model. As such, VanDeusen (2002) gives the complete-data likelihood for the CJS model.

$$L(\phi, \mathbf{p})_{complete\ data} = \prod_{i=1}^n \prod_{t=f_i+1}^k \phi_{t-1}^{a_{i,t-1}a_{it}} (1 - \phi_{t-1})^{a_{i,t-1}(1-a_{it})} p_t^{a_{it}\omega_{it}} (1 - p_t)^{a_{it}(1-\omega_{it})} \quad (1.7)$$

where a_{it} is the alive status of individual i at time t and ω_{it} is the capture history element for individual i at time t . Thus the capture component depends on whether the animal was alive, otherwise the animal cannot be captured. Similarly, for the survival component the animal must be alive at time $t - 1$ in order to survive to time

t .

The E step of the EM algorithm requires the expectation of the complete-data log-likelihood. Hence $E(a_{i,t-1}a_{it})$ and $E(a_{it})$ are required. VanDeusen (2002) shows the derivation of the $E(a_{it})$. For $E(a_{i,t-1}a_{it})$ we note that $E(a_{i,t-1}a_{it}) = E(a_{it})$. If $t \leq l_i$, $E(a_{it}) = a_{it}$ the observed alive status at time t for animal i . Let u_i be the event that the animal has been unseen since time l_i . As such, a_{it} can take on the value of 1 or 0, and:

$$\begin{aligned}
 E(a_{it}) &= 1P(a_{it} = 1|u_i) + 0P(a_{it} = 0|u_i) \\
 &= \frac{P(a_{it} = 1, u_i)}{P(u_i)} \\
 &= \frac{\chi_t \prod_{j=l_i+1}^t \phi_{j-1}(1 - p_j)}{\chi_{l_i}}
 \end{aligned} \tag{1.8}$$

The denominator is the probability that an animal has not been seen since time l_i , which is the definition of χ_{l_i} . The numerator is the probability of not being seen after time l_i and being alive at time t .

Maximization, the M-step, occurs separately for each component as they appear in the complete-data likelihood. First derivative vectors and information matrices of the complete-data likelihood are required to perform the Newton-Raphson algorithm in the M-step. VanDeusen (2002) models these parameters using the logit link function and finds the score and information matrices for the capture component to be:

$$U(\beta_p) = \sum_{i=1}^n \sum_{t=f_{i+1}}^k E(a_{it})(\omega_{it} - p_t)x'_t \tag{1.9}$$

$$I(\beta_p) = \sum_{i=1}^n \sum_{t=f_{i+1}}^k \frac{e^{x_t \beta_p}}{(1 + e^{x_t \beta_p})^2} x'_t x_t E(a_{it}) \tag{1.10}$$

and for the survival component to be:

$$U(\beta_\phi) = \sum_{i=1}^n \sum_{t=f_{i+1}}^k \{E(a_{it}) - E(a_{i,t-1})\phi_{t-1}\} z'_t \tag{1.11}$$

$$I(\beta_\phi) = \sum_{i=1}^n \sum_{t=f_{i+1}}^k \frac{e^{z_t \beta_\phi}}{(1 + e^{z_t \beta_\phi})^2} z_t' z_t E(a_{i,t-1}) \quad (1.12)$$

where x_t and z_t are appropriate components of the design matrix for p and ϕ .

Standard errors are obtained by inverting the information matrix of the original likelihood. Alternatively, they can be found by inverting the information matrix from the complete-data likelihood after an adjustment is made to account for the estimation of latent variables (Louis, 1982). For numerical methods when the information matrix is not available, the information matrix must be estimated using central differencing equations. Again, we first obtain the covariance of the β parameters and convert these to covariances of the standard parameters by multiplying by the design matrix, and then invoke the delta method to multiply by the derivative of the link function ($diag(\frac{dg}{dX\beta})XCov(\beta)X'diag(\frac{dg}{dX\beta})$). The square root of the diagonal of this resulting covariance matrix are the standard errors of the standard capture and survival parameters.

Alternatively, standard errors of parameter estimates can be obtained from bootstrap methodology (Buckland and Garthwaite, 1991). Here we sample the original data several times and compute parameter estimates for each sample. The standard deviation of the estimates over the bootstrap samples provide an estimate of the standard error of a parameter (Efron, 1982).

1.3.3 Model Selection

Modelling in some sense remains an art. There is no “correct” model to be found. Nonetheless, tools exist to lead us in making good decisions around which model or models to select to describe the data. All models are approximations to truth and it is possible to find several models that represent the data fairly well.

Current model selection methods in mark-recapture studies use Akaike model selection techniques (Akaike, 1973). Likelihood ratio tests are the standard alternative, requiring that one model be nested within another more general model. Specifically, it is used to test a constraint on a model such as testing $\theta_1 = \theta, \theta_2 = \theta$ (model 1) versus θ_1, θ_2 (model 2). Likelihood ratio tests are not useful for models that have

completely different parameter sets.

Akaike (1973) described a relationship between the Kullback-Leibler information and maximum likelihood (Burnham and Anderson, 2001). The Kullback-Leibler information is the negative of Boltzmann's entropy and describes the distance between two models (conceptual reality and the approximating model) (Burnham and Anderson, 2002, p.51).

Akaike's Information Criterion (AIC) is defined as:

$$AIC = -2\log(L(\hat{\theta}|y)) + 2K \quad (1.13)$$

where K is the number of parameters in the model and $\log(L(\hat{\theta}|y))$ is the log-likelihood evaluated at the maximum likelihood estimate $\hat{\theta}$. AIC takes into consideration the log-likelihood as well as the number of parameters in the model. Hence it takes into consideration the principle of parsimony, penalizing models with many parameters. The use of AIC does not necessitate the need for nested models allowing for more flexibility and comparison between models. It is important to note that AIC cannot be used to compare models of different data sets (Burnham and Anderson, 2002, p.80).

AIC is a relative scale (Burnham and Anderson, 2002, p.70). Burnham and Anderson (2002, p.71) advocate the use of AIC differences over all candidate models in a set. The Δ_i values allow for quick comparisons and ranking of models. The best model is that with $\Delta_i = 0$ (Burnham and Anderson, 2002, p.71).

$$\Delta_i = AIC_i - AIC_{min} \quad (1.14)$$

where Δ_i is the difference between the AIC value of model i and the AIC value of the model with the smallest AIC value. Burnham and Anderson (2002, p.71) suggest that models with $\Delta_i > 10$ have essentially no support and those with $\Delta_i \leq 2$ have substantial support.

Akaike (1978) developed Akaike weights:

$$w_i = \frac{e^{-\frac{1}{2}\Delta_i}}{\sum_{r=1}^R e^{-\frac{1}{2}\Delta_r}} \quad (1.15)$$

The idea behind the weights is that given a set of R models, the likelihood of model g_i given the data is proportional to $L(g_i|x) \propto e^{-\frac{1}{2}\Delta_i}$. If we normalize this likelihood over all R models, we obtain the Akaike weights for each model (Burnham and Anderson, 2002, p.75).

If there are too many parameters in relation to the size of the sample, AIC may perform poorly. In order to deal with small samples, Hurvich and Tsai (1989) introduced the AIC_c where there is a small-sample bias-correction factor of $\frac{n}{n-K-1}$ multiplied by the penalty term.

$$AIC_c = -2\log(L(\hat{\theta})) + 2K \left(\frac{n}{n-K-1} \right) \quad (1.16)$$

where n is the size of the sample. Burnham and Anderson (2002, p.66) recommend the use of AIC_c over AIC when n/K is small.

Information-theoretic model selection allows us to use a set of models for inference purposes. The AIC weights are used to scale models by their relative plausibility. Inference is no longer conditional on one particular model, but the full set of models. This is what is deemed model averaging (Burnham and Anderson, 2002, p.149).

Parameter estimates are obtained by summing over the product of individual model estimates and weights (Burnham and Anderson, 2002, p. 150).

$$\hat{\theta} = \sum_{i=1}^R w_i \hat{\theta}_i \quad (1.17)$$

where $\hat{\theta}$ denotes the model averaged estimate of θ and $\hat{\theta}_i$ denotes the model estimate of θ . Variance estimates are given by the unconditional variance estimator (Burnham and Anderson, 2002, p. 162).

$$\widehat{Var}(\hat{\theta}) = \left\{ \sum_{i=1}^R w_i \sqrt{Var(\hat{\theta}_i|g_i) + (\hat{\theta}_i - \hat{\theta})^2} \right\}^2 \quad (1.18)$$

where $Var(\hat{\theta}_i|g_i)$ is the sample variance conditional on model g_i . This estimator incorporates the uncertainty associated with model selection as well as the variability in the data. It is also possible to estimate the variance of $\hat{\theta}$ using Monte Carlo simulations or using the bootstrap (Burnham and Anderson, 2002, p.154).

Table 1.5: The 2×2 contingency table for the first component of the goodness-of-fit test described by Pollock et al. 1990.

	First captured before time i .	First captured at time i
Captured at time i and recaptured		
Captured at time i and not recaptured.		

1.3.4 Goodness of Fit

The fit of a model to the observed data must be checked using a goodness-of-fit measure. The most traditional goodness-of-fit test statistic for count data is the sum of squared Pearson residuals or

$$X^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (1.19)$$

where O_i is the observed count and E_i is the expected count for observation i . X^2 is asymptotically χ^2 distributed with degrees of freedom equal to n (the number of unique capture histories) minus the number of parameters (K) in the model producing the expected counts.

The squared Pearson residuals behave similarly to the residual deviance function which is twice the difference between the log-likelihood of the saturated and the fitted models (McCullagh and Nelder, 1989, p. 118). The deviance is asymptotically χ^2 distributed with $n - K$ degrees of freedom, where n is the number of counts and K is the number of fitted parameters.

Pollock et al. (1990) describe a goodness-of-fit test based on a series of linked contingency tables for a χ^2 test of the JS model derived from the residual distribution after factoring the minimal sufficient statistics from the likelihood. There are 2 components involved in this goodness-of-fit test, the first of which includes $k - 2$ contingency tables. An example of one of the tables is shown in Table 1.5.

The columns are a count of the animals captured at time i and a pooling of counts of

Table 1.6: A 2×2 contingency table for the second component of the goodness-of-fit test described by Pollock et al. 1990.

	First captured before time $i - 1$.	First captured at time $i - 1$.
Recaptured at time i		
Recaptured after time i		

animals captured before time i . The rows deal with the animals that are captured at time i and whether or not they are recaptured at some time after time i .

The second component of the goodness-of-fit test includes $k - 3$ contingency tables. Again animals are pooled to determine the number of animals before time $i - 1$ and all the animals captured at time $i - 1$. The rows deal with whether an animal was next captured at time i or at some later time after time i . Table 1.6 shows this second contingency table.

If any of the expected cell counts are less than 2 for any table Pollock et al. (1990) suggest omitting this table from the test. Individual X^2 values are calculated for each table with 1 degree of freedom. These values are summed to produce a total X^2 value with degrees of freedom equal to the number of tables summed. Finally, an overall X^2 value is determined which is the sum of the 2 total X^2 values for each of the 2 components of the test. Similarly the overall degrees of freedom is the sum of the total degrees of freedom for the 2 components. Pollock et al. (1990) provide an example of the use of this goodness-of-fit test.

These goodness-of-fit tests perform differently under different assumption violations. The first component will detect heterogeneity of capture probabilities as well as a negative relationship between capture and survival probabilities. The second component detects heterogeneous capture or survival probabilities. It is sensitive to direct relationships between capture and survival. The overall goodness-of-fit test detects tag loss as a function of time and is sensitive to a positive relationship between survival and capture probabilities. This test will not detect a permanent trap response of capture probabilities or a permanent lowering of survival due to tagging,

if marking has a long term effect. Further, it has low power for small populations, low capture rates and low survival rates (Pollock et al. 1990).

Burnham et al. (1987) do not pool the m_{ij} -array, rather they compare 2 rows of counts to compare the proportion of counts across columns. If the proportions do not differ significantly then it is said that survival and capture probabilities show no evidence of differing between the 2 rows of counts (Burnham et al. 1987).

1.3.5 Over-dispersion

Over-dispersion occurs when the sampling variance exceeds the theoretical variance based on a given model. Count data is often over-dispersed and tends not to conform to model assumptions. Over-dispersion will arise if there are small violations in model assumptions such as a lack of independence (e.g. animal behaviour governs groups rather than individuals) or parameter heterogeneity (e.g. individuals have unique survival probability) (Burnham and Anderson, 2002, p.67).

In over-dispersed Poisson models, the point estimates from the generalized estimating equations (GEE) accounting for over-dispersion are the same as the maximum likelihood estimates (MLE) ignoring over-dispersion. But the covariance of the estimates needs to be increased. Often a variance inflation factor (c) is used to inflate the estimated variance from the simple MLE fit (i.e. $Var(\hat{\theta})_{GEE} = cVar(\hat{\theta})_{MLE}$). Hence the theoretical variance is multiplied by a variance inflation factor to estimate the over-dispersed variance.

One estimate of c is based on the goodness-of-fit chi-square statistic of the saturated model and its degrees of freedom (Burnham and Anderson, 2002, p.68),

$$\hat{c} = \frac{X^2}{df}. \quad (1.20)$$

Alternatively, over-dispersion can be estimated by the deviance divided by the degrees of freedom (McCullagh and Nelder, 1989; p.126). Burnham and Anderson (2002, p.68) note that if an over-dispersion parameter is estimated, one parameter must be added to K (the number of parameters in the model). Large values of \hat{c} indicate gross lack of fit of the model (Lebreton et al. 1992).

Over-dispersion also comes into play during model selection and the AIC criterion requires modification. Now, QAIC is computed for model selection (1.21).

$$QAIC = -2 \frac{\log(L(\hat{\theta}))}{\hat{c}} - 2K \quad (1.21)$$

Similar to AIC, QAIC can be modified for small samples (Lebreton et al. (1992); Burnham and Anderson, 2002, p.69):

$$QAIC_c = QAIC + \frac{2K(K+1)}{n-K-1} \quad (1.22)$$

1.4 Poisson versus Multinomial Models

Counts of each observed capture history can be modelled as independent Poisson random variables rather than multinomial random variables (Sandland and Cormack, 1984). These models were also proposed independently by Cormack (1979) and Jolly (1979). The relationship between these models is best studied with an example.

If N is the number of fish tagged and released, and X is the number of tagged fish recovered, then $(N - X)$ is the number of unobserved tagged individuals. If we model these counts as a multinomial with P the probability of recovery, we find

$$L = \binom{N}{x} P^x (1-P)^{N-x} \quad (1.23)$$

However, we can treat X and $Y = (N - X)$ as Poisson random counts with $E(X) = NP$ and $E(Y) = N(1 - P)$. Then the Poisson model of the observed and unobserved counts would be

$$\begin{aligned} L &= \frac{e^{-NP} (NP)^x}{x!} \frac{e^{-N(1-P)} (N(1-P))^y}{y!} \\ &= \frac{e^{-NP} (NP)^x}{x!} \frac{e^{-N(1-P)} (N(1-P))^{N-x}}{(N-x)!} \end{aligned} \quad (1.24)$$

Sanderland and Cormack (1984) showed that Poisson models give the same point estimates as multinomial models; however, the covariances of estimators from the Poisson model differ by a small added constant. Further, if we rearrange the Poisson model and condition on N we indeed obtain the multinomial model.

A Poisson model is often used when the probability of recovery of individual animals is low ($P \rightarrow 0$) (Hilborn, 1990). In this case only the counts of the animals seen are modelled and $L = \frac{e^{-NP}(NP)^x}{x!}$ with the term for animals never seen after release ignored. This latter term often has a very complicated expression which is difficult to compute but is close to 1. Thus if P is small, $(1 - P)$ is close to 1 and we find the Poisson model of just the observed counts is a good approximation to the multinomial model. Using the Poisson model simplifies the problem as we are no longer concerned with the unobserved fish and their expected counts.

1.5 Bayesian Methods

Bayesian modelling techniques have also been used in the mark-recapture literature. Brooks, Catchpole and Morgan (2000) find Bayesian implementation to be straightforward and convenient for model averaging. Moreover, asymptotic techniques are not required for standard error estimation. Bayesian techniques are enticing as they provide a natural method to include prior knowledge (Brooks et al. 2000a).

Vounatsou and Smith (1995) outline a Bayesian analysis for ring-recovery models using Markov chain Monte Carlo (MCMC) simulation methods. In particular they provide details of Gibbs sampling and the Hastings-Metropolis algorithm. Brooks et al. (2000b) look at the same data sets as Vounatsou and Smith (1995) and discuss the issue of sensitivity of parameter estimates to the choice of prior distributions. Additionally they found that when the likelihood surface has a completely flat ridge Bayesian estimators are precise due to prior information.

Brooks et al. (2000a) describe a Bayesian analysis for CJS models and they provide an overview of Bayesian methods in general. They speculate that while prior distributions have a profound effect on ring-recovery models, they will have little effect on recapture models (Brooks et al. 2000b). As probabilities are typically

the parameters of interest in mark-recapture experiments, priors tend to be Beta or uniform(0,1) distributed (see Brooks et al. 2000a; Brooks et al. 2000b; Poole, 2002). Note that a uniform(0,1) is a special case of a Beta distribution.

Poole (2002) discusses Bayesian implementation of the CJS model using the Metropolis Hastings algorithm. Poole (2002) also considers the case where missing data occurs. Missing data or latent parameters are often treated as extra parameters in the Bayesian context, estimated along with regular parameters (Poole, 2002). Similar to the EM algorithm, the complete-data likelihood is required when dealing with missing data.

1.6 Tag Loss

One of the crucial assumptions of mark-recapture experiments is that animals do not lose their tags. Arnason and Mills (1981) investigated the effects of homogeneous tag loss: when marked animals in the population have equal probability of tag loss between sample times i and $i + 1$ regardless of when they were tagged. In their paper, Arnason and Mills (1981) state that “it is never safe to assume that tag loss is negligible in a fish mark-recapture experiment.”

Tag loss affects JS models in various ways depending on the type of tag loss and the type of model. For birth and death models under homogeneous tag loss, survival, birth and standard error estimates for birth are biased (Arnason and Mills, 1981). For the death-only model (no births), only the standard error estimate of population size is unbiased. Homogeneous tag loss also causes a loss in precision in all estimates (Arnason and Mills, 1981). McDonald, Amstrup, and Manly (2003) investigate the effects of tag loss under several tag-loss models that differ from Arnason and Mills (1981), such as tag-induced mortality and immediate tag loss. McDonald et al. (2003) found that tag loss can bias population size estimates especially when coupled with low capture probabilities.

Multiple methods have been used to estimate tag loss. Tag-loss rates are traditionally obtained from holding experiments (where fish are held in tanks for the duration of the experiment) or double-tagging experiments. Robson and Regier (1966) offer a

method to estimate tag-retention rates for experiments where one of the double tags is a permanent mark. Wetherall (1982) develops estimates of instantaneous tag-loss and time-dependent tag-loss rates. Barrowman and Myers (1996) use maximum likelihood methods to estimate tag-loss rates in double-tagging experiments. Seber and Felton (1981) provide methods for estimating tag loss (1.25). If tags are indistinguishable and Λ is the probability of tag retention,

$$\tilde{\Lambda} = \frac{2n_{11,11}}{n_{11,01} + 2n_{11,11}} \quad (1.25)$$

where $n_{11,11}$ is the number of double-tagged animals with 2 tags and $n_{11,01}$ is the number of double-tagged animals with a single tag. Notice that in experiments where animals are both double- or single-tagged, we make use of a modified capture history or tag history. A history of $\{11\ 01\}$ signifies an animal who is captured and double tagged at time 1 and captured with 1 tag present at time 2; a history $\{01\ 01\}$ represents a single-tagged animal captured at times 1 and 2.

To date, the solution to these bias problems induced by tag loss has been to perform *ad hoc* corrections to estimates of abundance. Seber and Felton (1981) provide methods for correcting estimates of population size using their tag-retention estimates for 2-sample experiments. *Ad hoc* corrections are unsatisfactory and this thesis explores modelling tag loss along with other population parameters in several mark-recapture models.

1.7 Outline of Thesis

This thesis looks at 3 different cases where the assumption of no tag loss is violated for open populations. Chapters 2, 3 and 4 are written in a manner similar to that for publication. A reduced version of chapter 2 has been accepted for publication in *Biometrics*, and a modified version of chapter 3 is in review. Chapter 4 is in preparation for submission. As such, some of the introductory material found in this chapter is reproduced in later chapters.

1.7.1 Chapter 2

Radio-telemetry is a growing form of tagging in mark-recapture experiments. Due to their high detectability radio-tags are reliant on battery power and premature radio failure is a form of tag loss. We develop an extension to the CJS model that incorporates auxiliary information regarding radio-tag failure. This auxiliary information comes from a second experiment where the distribution of the time to radio-tag failure is estimated.

Both a classical and a Bayesian model are developed and applied to an example involving chinook salmon smolts on the Columbia River, Washington.

1.7.2 Chapter 3

Tag loss has been known to occur in many tagging studies (see Hubert, 1975; Alt, McLaughlin and Pollock, 1985; Stobo and Horne, 1993; Diefenbach and Alt, 1998; Bradshaw, Barker and Davis, 2000; and Pistorius et al. 2000). Loss of tags affects parameter and standard error estimates. Previously, *ad hoc* corrections to estimates of abundance have been used to deal with tag loss (Seber and Felton, 1981), though this is unsatisfactory.

The JS model is extended to incorporate tag-retention parameters. This methodology requires some of the animals to be initially double tagged. We employ the EM algorithm to apply this model to estimate abundance of walleye in Mille Lacs, Minnesota.

1.7.3 Chapter 4

Quantifying the migration of animals is of interest to biological scientists studying movement patterns. Mark-recapture studies can be used to model migration between different strata. Hilborn (1990) introduces a Poisson model which models movement, mortality and exploitation. The model is extended using a matrix formulation similar to that of Schwarz, Schweigert and Arnason (1993), to model migration, natural survival, tag-loss, exploitation and reporting rate. An example illustrates the use of

this model as applied to yellowtail flounder off the Grand Banks of Newfoundland.

1.7.4 Appendices

Four appendices accompany this thesis. The first looks at a problem a reviewer brought up regarding chapter 2. It was suggested that there was some loss of information when we estimated radio failure outside of the full likelihood. Appendix A investigates the amount of information lost and the impact on parameter estimates.

Appendix B presents the expectations needed to complete the EM algorithm as applied to the model in Chapter 3. Appendix C provides selected results of a simulation study to test the code for the Jolly-Seber-Tag-Loss model found in chapter 3. Finally, appendix D contains the data for the Mille Lac walleye example found in chapter 3. Data from the Columbia River example will be available from the Biometrics web site. The yellowtail flounder data is summarized within chapter 4. As the raw data is confidential, it is only available from Fisheries and Oceans Canada.

Chapter 2

Capture-Recapture Studies with Premature Radio-Tag Failure

Radio-tags are often used in mark-recapture studies because of their very high detectability. This leads to a reduction in the number of animals needed to be marked to obtain a specified precision, which is particularly important when studying threatened populations. However, radio-tags may fail prior to the end of the study which violates a key assumption of mark-recapture studies of no tag loss. For example, consider a capture-recapture study using radio-tagged fish to estimate survival of fish as they pass through a system of dams and reservoirs as outlined by Burnham et al. (1987). Ideally, batteries should fail after the study ends. If fish move downriver slower than expected (e.g., due to slower river flow rates), increased battery failure is experienced which will negatively bias survival estimates. In such cases, it would be desirable to be able to adjust survival estimates using some function of the rate of radio-tag failure.

Pollock et al. (1995) combined the use of capture-recapture models with Kaplan-Meier estimators to introduce a capture-recapture survival model for radio-tagged animals. The model required observations of both live and dead marked animals and a key assumption was that radio-tags did not fail.

Alternatively, English et al. (2001b) suggested censoring observations based on the percentage of fish detected over time. The censoring time was determined by

looking for changes in the detection patterns. This introduced another parameter into the likelihood (i.e. probability of being censored) and it ignored information collected after the censoring time. In other words, fish that were detected after the censor time were not included in the study; this was a *post hoc* decision. Further, the censoring time was defined after the study was complete which could incorporate bias into the survival estimates because the censoring mechanism was not independent of the radio-survival time. Another problem with this method is that the fish that are never recaptured could still have radio failures regardless of the chosen censoring time.

In this paper, we develop methodology for capture-recapture experiments incorporating auxiliary information on the distribution of radio-failure times to adjust survival estimates. We apply this methodology to data from the 2001 season on the Columbia River where radio-tagged chinook salmon (*Oncorhynchus tshawytscha*) were used to estimate dam and reservoir mortality as smolts migrate to sea.

2.1 Design Protocol

There are two studies designed to collect different types of information. The first study is based on capture-recapture methods and is designed to estimate fish passage survival in a river system (Burnham et al. 1987). Briefly, fish are radio-tagged and released at a number of sampling points along a river. Detection of functioning radio-tags also takes place at a number of sampling points along the river. These may be, but are not necessarily the same points where fish are released. If a fish is detected at a sampling point, the time since release is recorded. Note that in a classical Cormack-Jolly-Seber design, the sampling points are synchronous with time (e.g. sampling point 1= week 1, sampling point 2= week 2, etc.); whereas in this design, sampling points do not occur at a set point in time. Rather, recoveries at sampling point 2 could occur at time 5, 7 or 10 days after release, depending on the individual fish. From these data, capture-history vectors and time-from-release vectors are produced. An example of these vectors is given in Table 2.1 for the Columbia River example described below.

The second study consists of holding a batch of radio-tags and observing the time

Table 2.1: A sample of the capture histories and cumulative time for the following sampling points: Rock Island tailrace, Wanapum Dam, Wanapum tailrace, Priest Rapids dam, Priest Rapids tailrace, Hanford Reach 1. Complete data are available at the Biometrics website.

Tag Number	Capture History	Time Vector (days)					
6001	1 1 0 1 0 1	0	4	-	10	-	11
6002	1 1 0 1 0 1	0	5	-	6	-	7
6003	1 1 0 0 0 0	0	17	-	-	-	-
6004	1 1 0 0 0 0	0	6	-	-	-	-
6005	1 0 0 0 0 0	0	-	-	-	-	-
7001	0 0 1 1 0 1	-	-	0	3	-	4
7002	0 0 1 1 0 1	-	-	0	4	-	5
7003	0 0 1 1 0 0	-	-	0	3	-	-
7004	0 0 1 1 0 1	-	-	0	3	-	4
7005	0 0 1 1 0 0	-	-	0	12	-	-
8001	0 0 0 0 1 1	-	-	-	-	0	1
8002	0 0 0 0 1 1	-	-	-	-	0	1
8003	0 0 0 0 1 1	-	-	-	-	0	1
8004	0 0 0 0 1 1	-	-	-	-	0	1
8005	0 0 0 0 1 1	-	-	-	-	0	1
1062 additional histories not shown.							

to failure of each tag. This data is used to estimate the failure-time curve of the radio-tags.

2.2 Notation

Statistics:

k number of sampling points.

ω capture history vector $\omega = (\omega_1, \omega_2, \dots, \omega_k)$ for each fish where

$$\omega_i = \begin{cases} 1 & \text{if the fish was detected at sampling point } i \\ 0 & \text{if the fish was not detected at sampling point } i \end{cases}$$

- f first sampling point where $\omega_i = 1$.
- l last sampling point where $\omega_i = 1$.
- \mathbf{t} time vector $\mathbf{t} = (t_1, \dots, t_f, t_f + 1, \dots, t_l, \dots, t_k)$ where t_i is the relative sample time from initial release to sampling point i with $t_f = 0$ for each release; $i = 1, 2, 3, \dots, k$; $0 < t_i < \infty$. If a fish is not detected at a sampling point, the corresponding component of \mathbf{t} is missing. t_i is undefined for $i < f$; t_i is unknown for $i > l$.
- Δt_{ij} change in time from sampling point i to j ; $i = 1, 2, \dots, k - 1$, $j = i + 1, i + 2, \dots, k$.
- n_r number of radio-tags used to produce the radio-failure curve.
- n_f number of fish in the study.
- z_i time that the radio-tag either fails or was removed from the study (censored), $i = 1, 2, \dots, n_r$.
- $\delta_i = \begin{cases} 1 & \text{radio-tag not censored} \\ 0 & \text{radio-tag removed from the study prior to failure (censored)} \end{cases}$

Parameters:

- p_i the conditional probability that a fish is recaptured at sampling point i , given that it is alive and the radio is working; $i = 2, 3, \dots, k$.
- ϕ_i the conditional probability that a fish alive at sampling point i , regardless of radio status, survives to sampling point $i + 1$, $i = 1, 2, \dots, k - 1$.
- $\chi_i(t_i)$ the probability that a fish is not seen after sampling point i , after time t_i from release. Note that $\chi_i(t_i)$ is a complicated function of ϕ , p , g , and S .
- $S(z)$ the probability of radio-tag survival to time z ; $0 < z < \infty$.
- $g_{ij}(\Delta t_{ij})$ the probability that a fish travels from sampling point i to sampling point j in $t_j - t_i = \Delta t_{ij}$ amount of time; $i = 1, 2, \dots, k - 1$, $j = i + 1, i + 2, \dots, k$.
- μ_i the expected travel time between sampling point i and $i + 1$.
- λ_i the scale parameter of the travel-time distribution.
- ξ a constant $\frac{\lambda_i}{\mu_i^2}$.

Latent Variables:

a alive status vector $\mathbf{a} = (a_1, a_2, \dots, a_k)$ for each fish where

$$a_i = \begin{cases} 1 & \text{if the fish was alive at sampling point } i \\ 0 & \text{if the fish was not alive at sampling point } i \end{cases}$$

$a_i = 1$ for $f \leq i \leq l$, but a_i is unknown for $i > l$ for each fish.

z_i^\dagger actual time that the radio-tag fails for an individual fish; $i = 1, 2, \dots, n_f$.

It is only known that $z_i^\dagger \geq t_l$ for any fish.

2.3 Classical Model Development

2.3.1 Likelihood Development

The likelihood consists of two parts –that for the capture histories and that for radio failures. The data needed to develop the likelihood for the first component are the individual capture histories of each fish along with the time vector for each fish.

Assumptions are the same as those outlined in Burnham et al. (1987). In addition, we assume that travel time and radio failure are independent. We further assume that fish survival between sampling points is independent of travel time. This is likely to be the case when the time duration of the experiment is short compared to fish lifetime.

The likelihood for the capture-recapture study is constructed by the product of likelihoods for each fish with capture history $\boldsymbol{\omega}$ and time vector \mathbf{t} : (the subscript $\boldsymbol{\omega}$ on the f and l variables has been dropped for clarity)

$$L^{fish} = \prod_{\boldsymbol{\omega}} \left\{ \chi_l(t_l) \frac{S(t_l)}{S(t_f)} \prod_{i=f}^{l-1} \phi_i \prod_{j=f+1}^l p_j^{\omega_j} (1 - p_j)^{1-\omega_j} \times \prod_{\forall i, j, i < j, \omega_i=1, \omega_{i+1}=0, \omega_{i+2}=0, \dots, \omega_j=1} g_{ij}(\Delta t_{ij}) \right\} \quad (2.1)$$

where

$$\chi_l(t_l) = \begin{cases} 1 - \phi_l + \\ \phi_l(1 - p_{l+1}) \int_{t_l}^{\infty} g_{l,l+1}(u - t_l) \frac{S(u)}{S(t_l)} \chi_{l+1}(u) du + \\ \phi_l \int_{t_l}^{\infty} g_{l,l+1}(u - t_l) \frac{S(t_l) - S(u)}{S(t_l)} du & \text{if } l = 1, 2, \dots, k - 1 \\ 1 & \text{if } l = k \end{cases}$$

and

$$g_{ij}(\Delta t_{ij}) = \begin{cases} g_{ij}(t_j - t_i) & \text{if } j = i + 1 \\ \int_{u_{i+1}} \int_{u_{i+2}} \dots \int_{u_{j-1}} g_{i,i+1}(u_{i+1} - t_i) g_{i+1,i+2}(u_{i+2} - u_{i+1}) \dots \\ \quad \times g_{j-1,j}(t_j - u_{j-1}) du_{i+1} du_{i+2} \dots du_{j-1} & \text{if } j > i + 1 \end{cases}$$

The first component of the likelihood ($\chi_l(t_l)$) is similar to the classical Cormack-Jolly-Seber (CJS) (Cormack, 1964; Jolly, 1965; Seber, 1965) model but includes fish that die, fish that are alive but not recaptured with radio-tags still working, and fish who are alive with failed radio-tags. The $\frac{S(t_l)}{S(t_f)}$ term deals with the survival of the radio-tag and specifically is the probability of the radio tag surviving beyond time t_l given it survived to time t_l . The third component of the likelihood containing the ϕ_i terms is similar to the usual CJS survival component and the $p_j^{\omega_j} (1 - p_j)^{1 - \omega_j}$ component corresponds to the CJS recapture component. The last component of the likelihood containing the $g_{ij}(\Delta t_{ij})$ deals with the travel time of the fish as discussed below.

We used the inverse gaussian ($IG(\mu_i, \lambda_i)$) distribution to model the travel-time distribution, g_{ij} . The parameter μ_i is the mean of the travel time from point i to point $i + 1$ and λ_i is the scale parameter. The inverse gaussian distribution describes the first passage time of a Brownian motion and is often used as a lifetime model (Chhikara and Folks, 1989).

A useful property of the inverse gaussian distribution is that if $\frac{\lambda_i}{\mu_i^2} = \xi$ is constant then the sum of inverse gaussian distributed random variables is also inverse gaussian. We make use of this property by first assuming $\frac{\lambda_i}{\mu_i^2} = \xi$ is constant and then the distribution of the travel times between sampling point i and sampling point m is

an $IG(\sum_{j=i}^{m-1} \mu_j, \xi(\sum_{j=i}^{m-1} \mu_j)^2)$. This implies it is no longer necessary to evaluate the multidimensional integral for $g_{ij}(\Delta t_{ij})$.

The likelihood for the second component is based on the radio survival times.

$$L^{radio} = \prod_{i=1}^{n_r} \{-S'(z_i)\}^{\delta_i} S(z_i)^{1-\delta_i} \quad (2.2)$$

The parametric form of $S(z_i)$ can be chosen appropriately. The combined likelihood is the product of the likelihood for the fish and the radio survival likelihood.

$$L = L^{fish} \times L^{radio} \quad (2.3)$$

2.3.2 Parameter Estimation

As shown in (2.3) the likelihood can be factored into two parts: one pertaining to radio survival and the other pertaining to fish survival. L^{radio} involves parameters of S only, whereas L^{fish} involves parameters of S, ϕ, p, μ and ξ . Thus, by factoring the likelihood in this way we can first estimate parameters of S using L^{radio} . Then by substituting our estimate of the parameters for S and hence of the function S into L^{fish} , we estimate the remaining parameters by maximizing L^{fish} holding our estimates of S fixed (Sanathanan, 1977).

There was concern that by not maximizing the full likelihood, information on radio survival would be lost. However, after investigating this issue via a simulation (with a parametric radio survival model) and looking at the score equations (Appendix A), we find that for the configurations of our experiment, our approach produces efficient estimators of radio survival. Radio failure and death of an animal cannot be distinguished in the mark-recapture experiment. This further reduces the information about radio failure from released animals. Moreover, maximizing the likelihood with this method, may be the only way to do so when radio survival is estimated in a nonparametric framework.

For the radio-failure curve, we estimated $S(z)$ using the Kaplan-Meier estimator $\widehat{S}(z) = \prod_{j|z_{(j)} \leq z} (1 - \frac{d_j}{r_j})$ where $z_{(j)}$ are the ordered radio-failure times, r_j is the number of radio-tags at risk of failure at time $z_{(j)}$ and d_j is the number of radio-tags failing at

time $z_{(j)}$ (Lawless, 1982, p.72). Alternatively, a parametric model such as the Weibull distribution could be fit.

Evaluation of the integrals that contain the travel-time and survival distributions in the χ terms was done by discretizing the observed range of t_i into 1 day intervals. The integral is then replaced by a summation within the integral boundaries:

$$\int_{t_i}^{\infty} g_{i,i+1}(u - t_i)S(u)du \approx \sum_{u=t_i}^{\infty} g_{i,i+1}(u - t_i)S(u) \quad (2.4)$$

There are no closed forms for the parameter estimators and numerical methods must be used.

2.3.3 Standard Error Estimation

Standard errors of parameter estimates were estimated using a bootstrap estimator (Efron, 1982). Both the radio survival times and the capture histories were independently resampled for this nonparametric bootstrap (Buckland and Garthwaite, 1991). Therefore, variability is accounted for from both the fish histories as well as from the radio data.

2.3.4 Model Selection

As we are able to formulate a likelihood, we can use the Akaike Information Criterion (AIC) to perform model selection (Burnham and Anderson, 2002). AIC is considered more appropriate than likelihood-ratio tests as it takes into account both the value of the likelihood as well as the number of parameters in the model.

Ideally we would like to perform model selection on L ; however, model selection need only address L^{fish} since L^{radio} is usually held fixed for all models. The number of parameters in L^{radio} is constant as well as the $\widehat{S}(t)$ parameters, no matter what model is selected for the capture-recapture likelihood. Thus, the contribution of L^{radio} to AIC can be ‘ignored’ as it is only the relative AIC between models that is relevant in model selection.

2.3.5 Goodness of Fit

Goodness of fit can be evaluated for the travel-time distribution of the first sampling point after release. Between these points, we assume that radio-tags are usually functioning because most studies are designed with some latitude for unforeseen events (such as low river-flow rates) which would have to be extremely severe for radios not to last at least as long as the time to travel past the next sampling point. This leaves the travel-time distribution nearly untangled from the radio-failure distribution in that radio-tag failure will not occur leaving most travel times observed. The travel times of these fish are representative of all fish because censoring did not occur prior to the first two sampling points. Probability plots for the inverse gaussian distribution can be produced for this part of the likelihood.

2.4 Bayesian Model Development

We also develop a Bayesian implementation of the model. Because of the intractability of the likelihood, we used a Markov Chain Monte Carlo (MCMC) method after generating latent variables for the complete set of travel times (between sampling points when the fish was not observed), the alive status (after the last sampling occasion where the fish was observed) and the actual radio failure time (after the last sampling occasion when the fish was observed).

2.4.1 Complete Data Likelihood

If all latent variables are known for a particular fish, the likelihood is easy to specify. This requires that all travel times, the radio-failure time, the capture histories and whether or not the fish is alive at each sampling point (alive status) are assumed known for each fish. We model the system as radios moving downstream; attached to the radios are fish. Only if the fish is alive and if the radio is functioning can the radio be detected. The radios continue to ‘move’ through the system even after the death of a fish or radio failure. The likelihood for an individual radio/fish is

$$\begin{aligned}
L_m^{fish} &= \prod_{i=f}^{k-1} g_{i,i+1}(\Delta t_{i,i+1}) \prod_{i=f}^{k-1} \phi_i^{a_i a_{i+1}} (1 - \phi_i)^{a_i(1-a_{i+1})} \times \\
&\quad \prod_{i=f+1}^k p_i^{a_i \omega_i I_{[z^\dagger \geq t_i]}} (1 - p_i)^{a_i(1-\omega_i) I_{[z^\dagger \geq t_i]}} f_{radio}(z^\dagger)
\end{aligned} \tag{2.5}$$

and

$$L^{fish} = \prod_{m=1}^{n_f} L_m^{fish} \tag{2.6}$$

As all travel times are known, the travel time components are for all sampling points, regardless of the capture history ω or the alive status \mathbf{a} . The radio-failure time for each fish is used in the component for the capture probabilities ($I_{[z^\dagger \geq t_i]}$) as well as the probability that the radio failed at time z^\dagger ($f_{radio}(z^\dagger)$). The alive status of the fish \mathbf{a} is used in the survival and capture probability components.

As before the travel time distribution g_{ij} is modelled using an inverse gaussian distribution.

2.4.2 Priors

Beta priors were chosen for the survival parameters and the recapture parameters as these parameters are constrained to the interval (0,1) and form a conjugate prior with the complete-data likelihood. Priors for the travel time parameters lend to some difficulty as no conjugate prior exists if both μ and ξ are unknown. However, if μ is known the natural conjugate prior for ξ is a gamma distribution. Even if ξ is known, there is no conjugate prior for μ . As μ is strictly positive and unimodal, a gamma prior was chosen for μ . When employing Metropolis-Hastings algorithms, parameters are known from one iteration to be used in the next iteration. Thus we are able to use the conjugate prior for ξ given μ .

2.4.3 Full Conditionals and Latent Data

The Metropolis-Hastings algorithm is used to generate candidate values. We followed the usual Bayesian paradigm, and treated the latent variables as parameters except that they do not have associated prior distributions. These latent variables are imputed using the full conditional distributions.

In the derivation of the conditional distributions for the latent variable for each fish, we drop the subscript m . It is also convenient to define the probability of not seeing a fish after sampling point i , after time t_i from release with $\Delta t_{i,i+1}, \Delta t_{i+1,i+2}, \dots, \Delta t_{k-1,k}$ and z^\dagger (2.7).

$$\chi_i(t_i, \Delta t_{i,i+1}, \Delta t_{i+1,i+2}, \dots, \Delta t_{k-1,k}, z^\dagger) = \begin{cases} 1 & \text{if } i = k \\ 1 & \text{if } z^\dagger < t_i + \Delta t_{i,i+1} \\ 1 - \phi_i + \\ \phi_i(1 - p_{i+1})\chi_{i+1}(t_i + \Delta t_{i,i+1}) & \text{else} \end{cases} \quad (2.7)$$

The missing travel times are generated depending on whether they occur before or after the last observed travel time, t_l . If they occur before t_l , such as between sampling points 2 and 3 in the capture history 100100, then we generate from the distribution with probability density function (pdf) $f(\Delta t_{i,i+1}|v)$ where v is the travel time between sampling point i and the next observed sampling point, s . The pdf of this distribution is given by:

$$f(\Delta t_{i,i+1}|v) = \frac{\sqrt{\frac{\xi}{2\pi}}\mu_i\Delta t_{i,i+1}^{-3/2}e^{\left(\frac{-\xi(\Delta t_{i,i+1}-\mu_i)^2}{2\Delta t_{i,i+1}}\right)} \times \sqrt{\frac{\xi}{2\pi}}\mu_{i+1,s}(v - \Delta t_{i,i+1})^{-3/2}e^{\left(\frac{-\xi(v-\Delta t_{i,i+1}-\mu_{i+1,s})^2}{2(v-\Delta t_{i,i+1})}\right)}}{\sqrt{\frac{\xi(\mu_i+\mu_{i+1,s})^2}{2\pi}}v^{-3/2}e^{\left(\frac{-\xi(v-\mu_i-\mu_{i+1,s})^2}{2v}\right)}} \quad (2.8)$$

where $\Delta t_{i,i+1} > 0$, $v > 0$, $\mu_i > 0$, $\mu_{i+1,s} > 0$, and $\xi > 0$. Further, μ_i is the mean travel

time from sampling point i to sampling point $i + 1$, and $\mu_{i+1,s}$ is the mean travel time from sampling point $i + 1$ to s , with $\mu_{i+1,s} = \mu_{i+1} + \mu_{i+2} + \dots + \mu_{s-1}$.

After generating $\Delta t_{i,i+1}$, the process repeats until all latent travel times between the sampling points where the fish was observed are complete.

If the missing travel times occur after the last observed travel time, such as with a capture history after sampling point 4 in capture history 100100, the full conditional distribution is found to be (after integrating over the alive latent variable):

$$[\Delta t_{l,l+1} | \cdot] \propto g_{l,l+1}(\Delta t_{l,l+1}) \chi_l(t_l, \Delta t_{l,l+1}, \Delta t_{l+1,l+2}, \dots, \Delta t_{k-1,k}, z^\dagger) \quad (2.9)$$

Each interpoint travel time is generated in turn.

As neither of the full conditionals for travel times are densities that are easy to generate from, we implemented the Metropolis-Hastings algorithm with an inverse gaussian distribution as the candidate distribution in both cases. In the second case, where travel times are missing after the last time seen, the Metropolis-Hastings ratio simply becomes the ratio of the χ terms at the proposed and old values of the interpoint travel times.

The full conditional for z_m^\dagger is given by:

$$[z_m^\dagger | \cdot] \propto f_{radio}(z^\dagger) I_{[z^\dagger \geq t_l]} \quad (2.10)$$

In our implementation, $f_{radio}(z^\dagger) = \frac{1}{n_r}$ at the observed radio-failure times and 0 elsewhere. The set of observed radio-failure times were sampled from with replacement for each iteration of the Markov Chain. We are able to integrate over the latent variable $a_{l+1}, a_{l+2}, \dots, a_k$ to give

$$f_{radio}(z^\dagger) \propto \frac{1}{n_r} \chi_l(t_l, \Delta t_{l,l+1}, \Delta t_{l+1,l+2}, \dots, \Delta t_{k-1,k}, z^\dagger) I_{[z^\dagger \geq t_l]} \quad (2.11)$$

Finally, it is convenient to fully generate the alive status latent variables for sampling points f to k . For technical reasons it is easier to consider the latent variable for death = $1 - a_i$, and

$$[1 - a_i | \cdot] \propto \begin{cases} 1 & \text{if } a_{i-1} = 0 \\ 0 & \text{if } f \leq i \leq l \\ \frac{1 - \phi_{i-1}}{\chi_{i-1}(t_{i-1}, \Delta t_{i-1, i}, \Delta t_{i, i+1}, \dots, \Delta t_{k-1, k}, z^\dagger)} & \text{else} \end{cases} \quad (2.12)$$

Following generation of the latent variables, the full conditional distributions for the survival and capture probabilities are beta distributed. However, the capture probabilities only consider cases where the radio is working.

$$[\phi_i | \cdot] \sim \text{Beta} \left(\sum_{m=1}^{n_f} a_{m,i} a_{m,i+1} + \alpha_1, \sum_{m=1}^{n_f} a_{m,i} (1 - a_{m,i+1}) + \beta_1 \right) \quad (2.13)$$

$$[p_i | \cdot] \sim \text{Beta} \left(\sum_{m=1}^{n_f} a_{m,i} \omega_{m,i} I_{[z_m^\dagger \geq t_{m,i}]} + \alpha_2, \sum_{m=1}^{n_f} a_{m,i} (1 - \omega_{m,i}) I_{[z_m^\dagger \geq t_{m,i}]} + \beta_2 \right) \quad (2.14)$$

The full conditional distributions for the ξ and μ parameters are:

$$[\xi | \cdot] \sim \text{Gamma} \left(\sum_{i=1}^{k-1} \frac{n_j}{2} + \zeta, \left[\frac{1}{2} \sum_{i=1}^{k-1} \sum_{m=1}^{n_k} \frac{(\Delta t_{m,i,i+1} - \mu_i)^2}{\Delta t_{m,i,i+1}} + \frac{1}{\tau} \right]^{-1} \right) \quad (2.15)$$

and

$$[\mu_i | \cdot] \propto \mu_i^{n_i + \eta - 1} \exp \left(\xi n_i \mu_i - \frac{\xi \mu_i^2}{2} \sum_{m=1}^{n_i} \frac{1}{\Delta t_{m,i,i+1}} - \frac{\mu_i}{\kappa} \right) \quad (2.16)$$

As there is no easy way to generate from $[\mu_i | \cdot]$, a Metropolis-Hastings step was implemented. The candidate distribution was a Normal, centered at the value of μ_i from the previous iteration, with variance $\frac{S_{i,i+1}^2}{n_f}$ where $S_{i,i+1}^2 = \frac{1}{n_f - 1} \sum_{m=1}^{n_f} (\Delta t_{m,i,i+1} - \Delta \bar{t}_{i,i+1})^2$.

Hence the full implementation proceeds as follows.

1. (a) Generate initial values for the values of ϕ_i , p_i , μ_i and ξ . We used the mean of the respective prior distributions.
- (b) Generate initial values for the latent variables for each fish in the following order

- i. for unknown $\Delta t_{i,i+1}$ between f and l (eg. for $i, i+1 = 2, 3$ in capture history 100100), set $\Delta t_{i,i+1} = \mu_i \sum_{j=a}^b \frac{(t_b - t_a)}{\mu_j}$ where b, a 'bracket' $i+1$.
 - ii. for unknown $\Delta t_{i,i+1}$ for $i \geq l$, set $\Delta t_{i,i+1} = \mu_i$
 - iii. generate z^\dagger given $\Delta t_{1,2}, \Delta t_{2,3}, \dots, \Delta t_{k-1,k}, \boldsymbol{\omega}$ using (2.11).
 - iv. generate $a_{f+1}, a_{f+2}, \dots, a_k$ given $\Delta t_{1,2}, \Delta t_{2,3}, \dots, \Delta t_{k-1,k}, z^\dagger, \boldsymbol{\omega}$ using (2.12)
2. Repeat for a large number of iterations:
 - (a) Generate latent variables for each fish in the same order as in 1(b) using equations 2.8, 2.9, 2.11, and 2.12.
 - (b) Generate new values for ϕ_i, p_i, μ_i and ξ using equations 2.13-2.16. Save the new value after each iteration
 3. Use the results for 2(b) to check for stationarity, burn in, and autocorrelation among consecutive iterates. Choose appropriate thinning values and then look at the posterior distribution after thinning. The thinned series represents the posterior distribution of the parameters.
 4. Calculate means and standard deviations of the thinned series for final point and error estimates.

2.5 Example

The Columbia river originates in the Selkirk Mountains in British Columbia and runs to Astoria, Oregon. Throughout Washington state, 11 hydro-electric dams have been built along the main waterway of the Columbia River (Figure 1). Chinook smolts migrating to the ocean must pass through these hydro dams. Interest lies in the survival rate of these smolts as they pass through the dams and reservoir systems.

Traditionally, Passive Integrated Transponder (PIT) tags have been used to obtain smolt survival estimates in the Columbia River (Skalski et al. 1998). Recapture rates tend to be low (Skalski et al. 2001b) which means high sample sizes must be used to get precise estimates of survival. Often a 100 fold increase in fish is needed compared

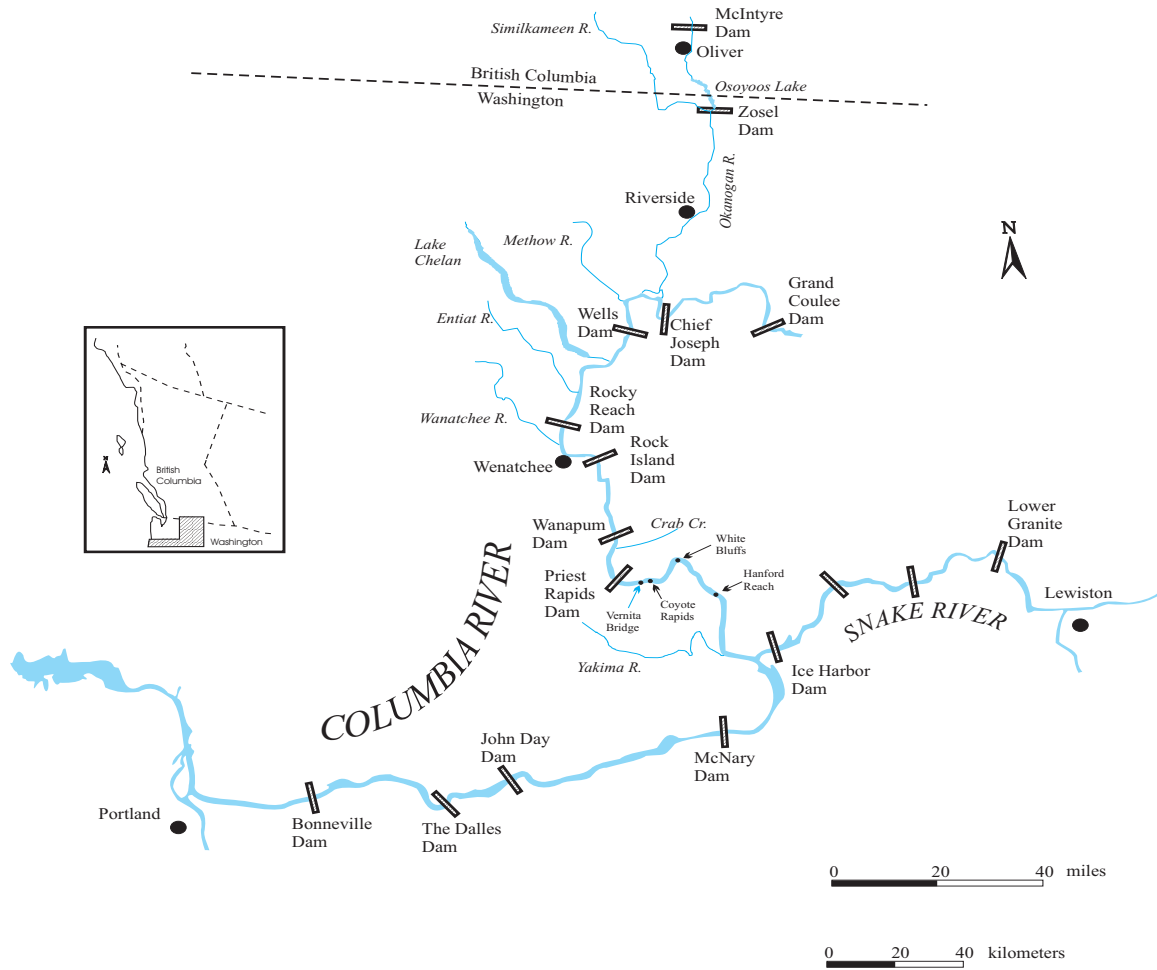


Figure 2.1: Map of the Columbia River showing Rock Island dam, Wanapum dam, Priest Rapids dam and Hanford Reach. Source ©LGL Limited, Sidney, British Columbia, by permission.

to radio-tagged fish to get an accurate and precise estimate of survival when PIT-tags are used (Skalski et al. 2001a). Recently, radio-tags have come into use since recapture rates are much higher and sample sizes can be substantially lower (English et al. 2001a; Skalski et al. 2001a).

From 25 April to 30 May 2001, 324 radio-tagged chinook smolts were released from Rock Island dam tailrace in replicates of 36. A second release group of 360 radio-tagged smolts were released from Wanapum dam tailrace in replicates of 36. Finally, 393 radio-tagged smolts were released from Priest Rapids dam tailrace in replicates of 36 (except for the last replicate which had 33 fish). Smolts were detected above the Wanapum dam, Priest Rapids dam and Hanford Reach 1 where possible. A schematic of the design is shown in Figure 2.

Releases were staggered such that the 1st release group's arrival at Wanapum dam would coincide with the release of the 2nd release group and the arrival of the 2nd group at Priest Rapids would coincide with the release of the 3rd release group. With 3 release groups, the estimates of survival through Rock Island tailrace (ϕ_1), over the Wanapum dam (ϕ_2), through Wanapum tailrace (ϕ_3), over the Priest Rapids dam (ϕ_4), through Priest Rapids tailrace (ϕ_5) and through Handford Reach 1 (ϕ_6) are obtainable. Selected sample data for each release group is shown in Table 2.1. The complete data set is available at the Biometrics web site.

Fish in release group 1 are released at time 0 into the tailrace of the first dam (Figure 2). These fish travel downstream to the second dam. Fish in the second release group are released in the tailrace of the second dam and travel with the first release group through the dam pool to the next sampling point. It is not possible to detect fish at sampling points where fish are released (sampling points 3 and 5), thus corresponding capture probabilities are constrained to zero. Note that fish are held for 48 hours before release to recover from tagging.

Twenty radio-tags were used to estimate the radio-survival curve, one of which failed at time 0. This was removed from the radio-tag data as no fish would be released with a failed tag; leaving 19 tags. The tags were attached to plastic vanes and submerged in a water vessel. Failure times for the radio-tags are provided in Table 2.2.

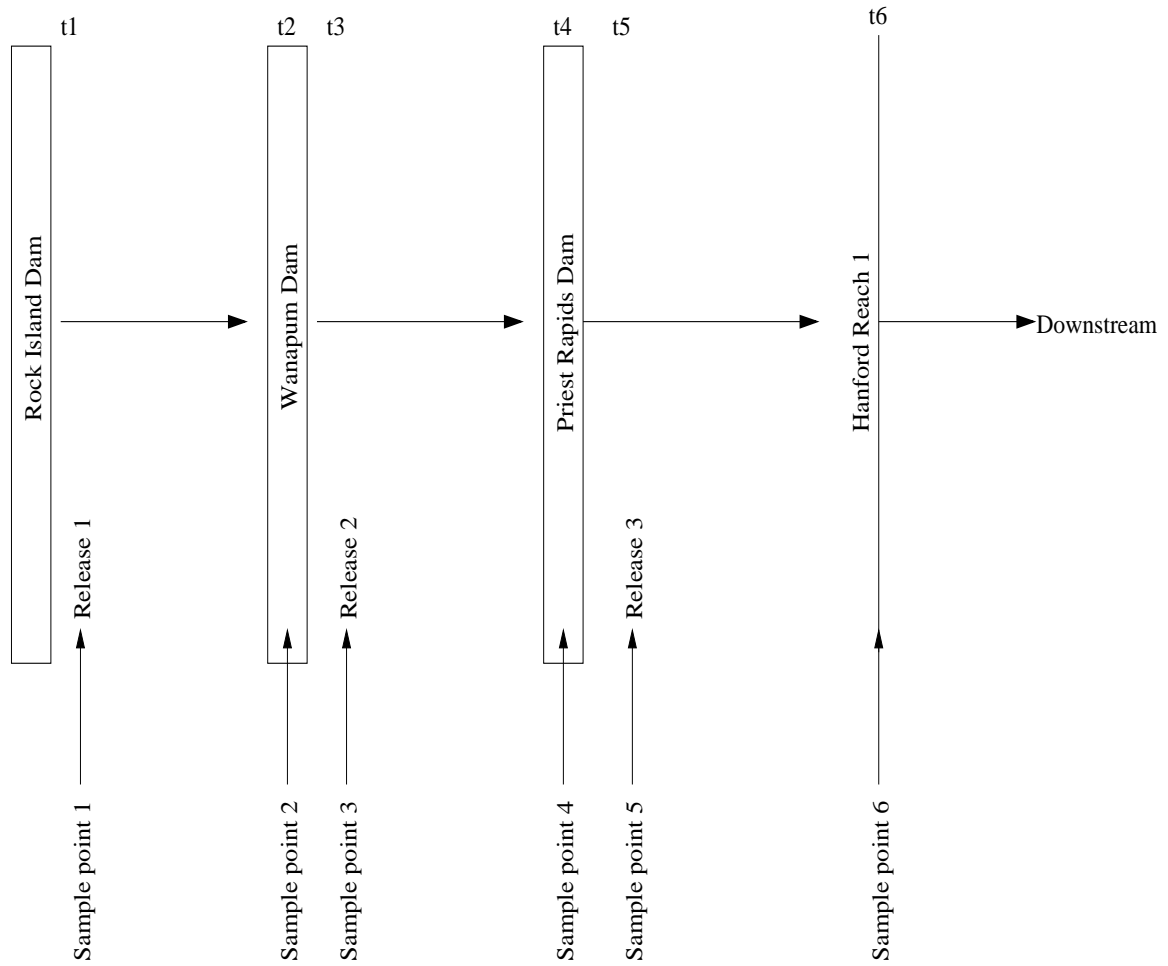


Figure 2.2: Diagram of the experimental design where fish are ‘released’ (radio-tags detected) at sampling points 1, 3, and 5 and recaptured at sampling points 2, 4, and 6. The time since release is shown as t_1, t_2, \dots, t_6 .

Table 2.2: Radio-tag survival data. Note that the last failure (*) was changed to a censored time so that radio-survival estimates would not go to zero.

Day	Number failed	K-M Survival
9	1	0.947
20	1	0.895
21	6	0.790
22	5	0.474
23	3	0.211
24	2	0.053
25	1	*0.053

As there were only 19 radio-tags from which to build a failure curve, it happened there were some fish who were observed after estimated radio survival was zero. To alleviate this problem, the 19th radio-failure was considered a censored time so that radio survival did not reach zero. Alternatively, we could have added a small constant to each radio-survival estimate.

Estimates of survival, recapture, and the movement parameters (μ_i and ξ) are provided in Table 2.3. A total of 100 bootstrap samples were taken to obtain standard error estimates. For the Bayesian analysis, we generated 100000 estimates, discarded the first 1000 (which was a conservative burn-in period) and then sampled every 100th estimate ending up with a final 990 Bayesian estimates from which posterior means and standard deviations are provided in Table 2.3.

Because there were no capture histories where fish were detected, not detected, then detected again, detectability was estimated as 100%. This is quite common with radio-tag studies. In other words, fish were either detected at a sampling point, or they were never seen again. This means that estimates of recapture were all 1, except at release sites where there were no antenna arrays and recapture probabilities were constrained to zero (p_3 and p_5).

Kaplan-Meier (KM) point estimates of fish survival based upon marked fish were also compared with the current estimates from the method of this chapter (Table 2.3).

Table 2.3: Classical parameter estimates and bootstrap standard errors(SE) (100 replicates) for the radio-tag failure model, Bayesian mean estimates and posterior standard deviations (SD), survival estimates assuming 100% detectability, and Cormack-Jolly-Seber (CJS) parameter estimates and standard errors.

Parameter	Classical		Bayesian		Survival estimate assuming 100% detectability \ddagger	CJS	
	estimate	SE	estimate	SD		estimate	SE
ϕ_1	0.95	0.022	0.92	0.019	0.87	0.87	0.018
ϕ_2	0.88	0.020	0.90	0.028	0.81	0.81	0.028
ϕ_3	1.00	0.000	0.96	0.011	0.96	0.96	0.011
ϕ_4	0.91	0.012	0.92	0.014	*0.89	0.89	0.014
** ϕ_5	1.00	0.000	0.99	0.004	0.99	1.00	0.003
p_2	1.00	0.000	1.00	0.001	.	1.00	0.000
$\dagger p_3$	0.00	0.000	0.04	0.012	.	0.00	0.000
p_4	1.00	0.000	1.00	0.001	.	1.00	0.000
$\dagger p_5$	0.00	0.000	0.02	0.006	.	0.00	0.000
** p_6	1.00	0.000	1.00	0.002	.	1.00	0.001
μ_1	9.76	0.047	10.61	0.326	.	.	.
μ_2	0.91	0.024	0.22	0.088	.	.	.
μ_3	3.68	0.015	4.09	0.144	.	.	.
μ_4	0.54	0.030	0.05	0.021	.	.	.
μ_5	0.46	0.070	0.89	0.041	.	.	.
ξ	0.18	0.011	0.39	0.033	.	.	.

* Pooled estimate from 2 release groups.

\dagger No detection possible as these release points have recapture rate set to zero.

\ddagger These were computed using a Kaplan-Meier type estimate.

** Note that p_6 and ϕ_5 are not individually estimable and p_6 is constrained to 1.

These estimates assume no battery failure and thus 100% detectability. It can be seen that the KM estimates are less than those of the current estimates and underestimate fish survival because some apparent deaths are due to radio failure. Standard errors were not computed for the KM estimates as too many assumptions would have to be made to do so and these estimates were only used for comparison purposes.

Cormack-Jolly-Seber estimates were computed using MARK (White and Burnham, 1999) and are also shown in Table 2.3. Here radio failure is not taken into account and the model only makes use of the capture-history data. Again survival estimates are lower than the current estimates because they do not distinguish between tag-failure and death which leads to underestimates of fish survival.

The radio-tag failure model estimates are comparable to the KM and CJS estimates. The current survival estimates were higher than or equal to the corresponding KM or CJS estimates and have comparable standard errors.

Probability plots to assess goodness of fit are presented in Figure 2.3 for the 3 release groups. The first and second release groups (Figure a and b) fit the model fairly well. However, the third release group (Figure c) does not fit the model. For this release group all but one travel time was 1 day. Part of the problem with the fit of the model is that the data were collected in hour time units but was analysed in day time units for simplicity. Further, travel times had to be a minimum of 1 day. The last release group made it from sampling point 5 to sampling point 6 in various portions of 1 day. However for analysis purposes, almost all travel times for that group were rounded up to 1 day, allowing for almost no variation in the travel time for that group.

The observed sample mean travel times from sampling points 1-2, 3-4 and 5-6 were 9.42, 3.68, and 1.24 days respectively; these sample means are very similar to estimates of μ_1 , μ_3 , and μ_5 . Mean radio-failure time of the 19 radio-tags was 21.45 days (SE=0.73).

For this example we did not implement AIC model selection as survival parameter estimates are clearly unique (Table 2.3). As such, a reduced model where some or all survival parameters are constant over sampling points would not fit the data. Thus no model selection was performed for this example.

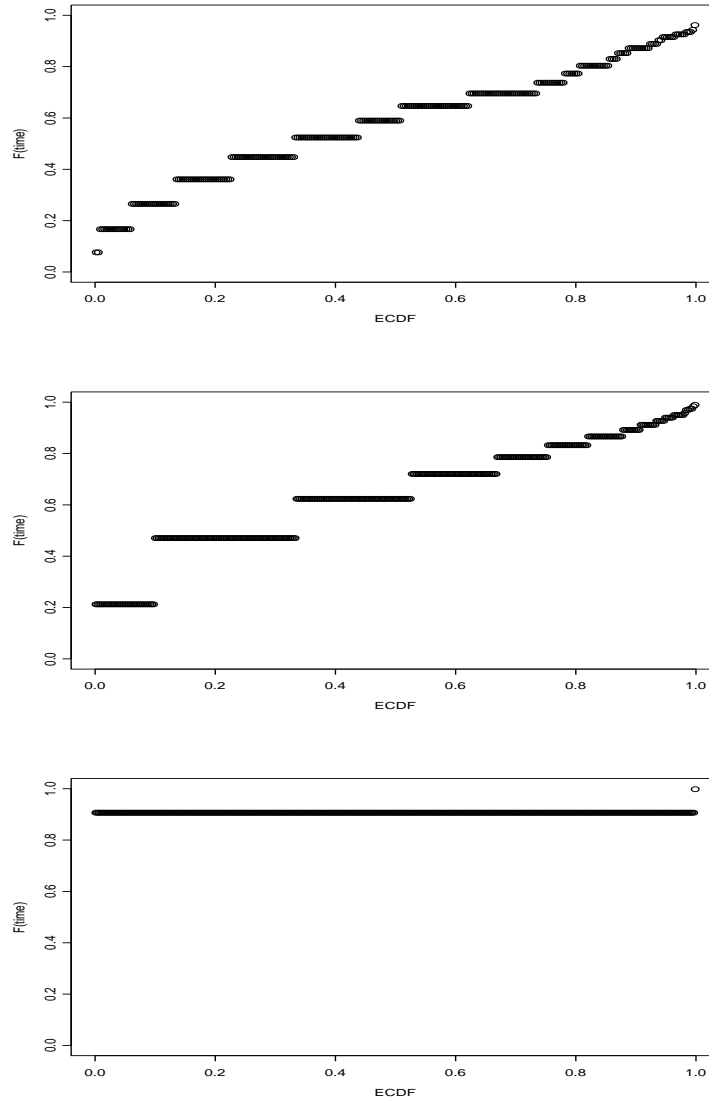


Figure 2.3: Probability plots to test the goodness of fit of the inverse gaussian distribution modelling the travel time of fish from a) dam 1 to dam 2 (release 1), b) dam 3 to dam 4 (release 2), c) dam 5 to dam 6 (release 3). $F(time)$ (the cumulative distribution function of the inverse gaussian evaluated at the radio-tag failure estimates) is plotted against the empirical cumulative distribution function (ECDF) defined as $\frac{i-0.5}{n}$, $i = 1, 2, \dots, n$, where n is the number of fish released.

Bayesian survival and capture estimates are similar to the classical estimates as are their respective error estimates. Again, Bayesian survival estimates are equal to or higher than the KM or the CJS estimates and they have comparable error estimates.

The travel-time parameters of the classical and Bayesian estimates are comparable. When travel times were short (as was the case from sampling point 2 to 3), mean travel times differ. For the SAS program implementing the classical analysis, travel times were rounded to integer days whereas for the Fortran program implementing the Bayesian analysis did not require this rounding. This could account for differences in the estimates of the travel-time parameters. As these are nuisance parameters, this is not of concern.

2.6 Simulation

Several small simulation studies were conducted to investigate the performance of the estimators. First, in order to investigate how the estimates would perform when detectability was less than 100%, we randomly changed 20% of detections to non-detections in the data set. The first (release point) time seen was not changed. Results are summarized in Table 2.4.

The current model survival estimates were generally higher than the corresponding KM and CJS survival estimates. The standard errors for the current estimates are again comparable to the CJS errors. Similarly, Bayesian survival estimates tend to be higher than the KM and CJS estimates; however, error estimates are also higher. Bayesian survival estimates are lower than classical radio-tag failure estimates, again with larger error estimates.

Mean travel-time estimates for long travel times are comparable. The dissimilarity in shorter mean travel-times (μ_2, μ_4 and μ_5) and ξ parameter estimates may be due to rounding error in the classical analysis.

A second simulation study was done to investigate small sample properties of our estimators. Survival and recapture probabilities were set to 0.9 and 0.8 respectively. Fish travel times between dams were generated from an inverse gaussian distribution with $\mu = 5$ days and $\lambda = 25$. Radio-tag data were generated from a Normal($\nu = 10$,

Table 2.4: Parameter estimates and standard errors (SE) for the classical tag-life adjusted model, Bayesian mean estimates and posterior standard deviations (SD), survival estimates assuming 100% detectability and Cormack-Jolly-Seber (CJS) estimates and standard errors when probability of recapture was simulated at 80%.

Parameter	Classical		Bayesian		Survival estimate assuming 100% detectability \ddagger	CJS	
	estimate	SE	estimate	SD		estimate	SE
ϕ_1	1.00	0.015	0.92	0.026	0.75	0.92	0.021
ϕ_2	0.89	0.012	0.87	0.038	0.78	0.77	0.031
ϕ_3	1.00	0.000	0.96	0.017	0.78	0.98	0.012
ϕ_4	0.91	0.012	0.88	0.032	*0.95	0.87	0.026
** ϕ_5	1.00	0.000	0.85	0.040	0.99	1.00	0.155
p_2	0.76	0.040	0.78	0.029		0.81	0.026
$\dagger p_3$	0.00	0.000	0.00	0.001		0.00	0.000
p_4	0.78	0.034	0.81	0.018		0.81	0.018
$\ddagger p_5$	0.00	0.000	0.00	0.000		0.00	0.000
** p_6	1.00	0.000	0.98	0.038		1.00	0.156
μ_1	9.50	0.016	10.28	0.315			
μ_2	0.83	0.011	0.10	0.082			
μ_3	4.39	0.008	4.00	0.137			
μ_4	0.56	0.016	0.03	0.017			
μ_5	0.52	0.025	0.88	0.038			
ξ	0.17	0.008	0.42	0.034			

* Pooled estimate from 2 release groups.

\dagger No detection possible as these release points have recapture rate set to zero.

\ddagger These were computed using a Kaplan-Meier type estimate.

** Note that p_6 and ϕ_5 are not individually estimable and p_6 is constrained to 1.

$\sigma^2 = 25$) distribution truncated at 0. Only the 3 dam case with one release group was simulated and estimates of ϕ_1 , p_2 , $\phi_2 p_3$, μ_1 , μ_2 and ξ were obtained. We cannot separate the $\phi_2 p_3$ estimate into its components due to non-identifiability problems. Two different sample sizes for the radio-failure phase ($m=25, 100$) and two different sample sizes of fish release groups ($n=100, 400$) were examined. Table 2.5 shows the estimate and bootstrap standard errors for the 4 different combinations of sample sizes.

Table 2.5: Small sample size effects of the number of radio-tags ($m=25$ or 100) used to estimate radio-tag survival and the number of fish ($n=100$ or 400) released. Actual parameter values used to simulate the data were $\phi_1 = 0.9$, $p_2 = 0.8$, $\phi_2 p_3 = 0.72$, $\mu_1 = \mu_2 = 5$, $\lambda_1 = \lambda_2 = 25$, $\nu = 10$, $\sigma^2 = 25$. Parameter estimates and bootstrap standard errors (SE) of 100 replicates are shown for the classical radio-tag failure model.

Parameter	Radio=25		Radio=100	
	Fish=100 (SE)	Fish=400 (SE)	Fish=100 (SE)	Fish=400 (SE)
ϕ_1	0.75 (0.10)	0.83 (0.08)	0.84 (0.07)	0.93 (0.04)
p_2	0.85 (0.07)	0.85 (0.03)	0.85 (0.07)	0.85 (0.03)
$\phi_2 p_3$	0.77 (0.12)	0.72 (0.09)	0.92 (0.12)	0.86 (0.08)
μ_1	3.78 (0.38)	3.92 (0.22)	4.03 (0.31)	4.14 (0.18)
μ_2	4.99 (0.60)	4.76 (0.26)	5.25 (0.50)	5.01 (0.30)
ξ	0.48 (0.15)	0.64 (0.09)	0.43 (0.12)	0.57 (0.08)

Increasing the number of fish released by a factor of 4 improves estimates and reduces standard error estimates by approximately $\frac{1}{\sqrt{4}}$. Increasing the number of radio-tags used to build the radio-failure curve had little effect as the mean time to radio failure was equal to the mean travel time. Only a few radios failed in the ‘experiment’ and so there was little loss of information.

The optimal allotment of radio-tags between the capture-recapture and radio-failure components is complicated. If the average radio failure is occurring at 50 days and the fish survival experiment is over in 10 days, then few resources need to be put into the radio-failure curve. On the other hand, if the average time to radio failure is

5 days, with the mean time to complete the experiment 20 days, then the the model will fail as the situation is hopeless as virtually no recaptures will be recorded.

The third simulation study investigated the effect of changing the number of radio-tags on fish that failed during the experiment. We generated 1000 fish travel-time vectors in a 3 dam design using an inverse gaussian ($\mu = 5$, $\lambda = 5$) distribution. Survival and recapture probabilities were set at 0.9 and 0.8 respectively. Radio-tag data was simulated from a Normal distribution with varying mean ($\nu = 10, 15, 20, 25$) and $\sigma^2 = 25$. These varying means translated into radio-tag failure rates of 45%, 22%, 9% and 5% on the tagged fish respectively (Table 2.6).

The RTF estimates of survival are within ± 2 standard errors of the true values but have large standard errors particularly with large amounts of radio failure. The CJS estimates underestimate fish survival and have lower standard errors when the failure rate is high because the model assumes radio-tag failures are deaths. The recapture parameter p_2 is fairly robust for both models with similar standard errors.

Estimates of mean travel time increase with increased mean radio-failure time. With many radios failing, only the faster fish are observed. As the mean radio failure time increases, slower travelling fish are recorded, bringing the estimates of μ_1 and μ_2 closer to their true value.

Lastly, the assumption that survival estimates are independent of travel time may not be tenable in some studies. A simulation study with 3 sampling points was conducted where survival was dependent on travel time. Parameters other than survival were held constant ($\mu_1 = \mu_2 = 5$, $\xi = 1$, $p = 0.8$, $\nu = 10$, $\sigma^2 = 25$) with 10000 fish and 1000 radios generated (Table 2.7). Percent bias was calculated for a parameter θ with estimate $\hat{\theta}$ as $\frac{(\hat{\theta} - \theta)}{\theta} \times 100$.

A small positive bias is seen for ϕ_1 . This bias is consistent with the effects of heterogeneity in CJS models (Williams et al. 2001; p.405). Larger positive bias is seen for $\phi_2 p_3$. We would expect larger bias for $\phi_2 p_3$ as the independence assumption is more likely to be violated between sampling points 2 and 3. Even when survival was constant over time, $\phi_2 p_3$ still contained a positive bias. Estimates of mean travel times have large negative biases as only the fast fish tend to make it to the subsequent sampling points. The ξ parameter is inversely related to the variance of the travel

Table 2.6: Simulation to determine the effect of changing the mean radio-tag failure time ($\nu = 10, 15, 20, 25$). The number of fish released was 1000, the number of radio-tags used to create the failure curve was 25. Other parameters included: $\phi_1 = 0.9$, $p_2 = 0.8$, $\phi_2 p_3 = 0.72$, $\mu_1 = \mu_2 = 5$, $\lambda_1 = \lambda_2 = 5$, $\sigma^2 = 25$. Parameter estimates and bootstrap standard errors (SE) of 100 replicates are shown for the classical radio-tag failure model (RTF). Cormack-Jolly-Seber (CJS) parameter estimates and standard errors are provided for comparison. Standard errors are shown below associated estimates.

Parameter	Mean radio-failure time (ν)							
	10		15		20		25	
	RTF	CJS	RTF	CJS	RTF	CJS	RTF	CJS
ϕ_1	0.95	0.71	0.95	0.84	0.91	0.87	0.91	0.88
	0.06	0.02	0.04	0.02	0.02	0.01	0.02	0.01
p_2	0.82	0.82	0.80	0.80	0.80	0.80	0.80	0.80
	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
$\phi_2 p_3$	0.78	0.53	0.76	0.62	0.79	0.70	0.78	0.72
	0.06	0.02	0.03	0.02	0.03	0.02	0.02	0.02
μ_1	3.84		4.30		4.25		4.44	
	0.22		0.23		0.18		0.13	
μ_2	4.06		4.66		4.65		5.17	
	0.25		0.20		0.14		0.11	
ξ	0.15		0.12		0.13		0.11	
	0.02		0.01		0.01		0.01	
Number Radio	447		222		92		46	
Failures	45%		22%		9%		5%	

Table 2.7: Sensitivity analysis for the violation of the assumption that fish survival and travel time is independent. The number of fish released was 10000, the number of radio-tags used to create the failure curve was 1000. Other parameters included: $p_2 = p_3 = 0.8$, $\nu = 10$, $\sigma = 5$, $\mu_1 = \mu_2 = 5$, $\xi = 1$. Percent bias of parameter estimates for the model are shown for 5 different distributions of survival, $\phi(t)$, where the approximate average survival rate was held at 0.5.

Parameter	Distribution				
	(1)	(2)	(3)	(4)	(5)
ϕ_1	1.00	0.55	0.72	2.07	-0.13
p_2	0.75	0.68	0.10	0.63	0.62
$\phi_2 p_3$	8.75	26.87	12.70	9.93	16.82
μ_1	-16.27	-30.42	-28.16	-17.13	-26.99
μ_2	-0.63	-32.79	-15.49	-2.87	-10.16
ξ	-35.94	9.91	10.22	33.99	36.41

- (1) $\phi(t) = 0.5$
- (2) $\phi(0 < t < 7.1) = 0.9$, $\phi(7.1 \leq t) = 0.1$
- (3) $\phi(0 < t < 6.1) = 0.8$, $\phi(6.1 \leq t) = 0.3$
- (4) $\phi(0 < t < 7.5) = 0.55$, $\phi(7.5 \leq t) = 0.45$
- (5) $\phi(t) = 1 - F(t)$, $t \sim \text{lognormal}(\mu = 2, \sigma = 1)$

times. Thus as more fish survive, the variance of the travel time increases and ξ decreases.

2.7 Discussion

Radio failure is a form of ‘tag loss’ which is ordinarily indistinguishable from mortality. Consequently survival estimates that do not account for tag loss can be negatively biased, particularly if radio-tag failure rates exceed 10%. Our model can account for this ‘tag loss’ if auxiliary information on radio-failure rates is available. A similar methodology can be used for any study where the ‘tag loss’ can be estimated; for example from returns of double-tagged fish.

We examined two methods for dealing with the problem. In the frequentist approach, we used maximum pseudo-likelihood estimation to obtain parameter estimates. The numerical integration to estimate the χ term was the most computationally intensive part of the estimation procedure. Two obvious alternatives to finding estimates present themselves. The EM-algorithm could have been used (VanDeusen, 2002); however, first and second degree derivatives are needed in the M step to implement the Newton-Raphson maximization and the expectation step was intractable.

Alternatively, we implemented a Bayesian approach using a Markov Chain Monte Carlo method and implemented the Metropolis-Hastings sampling technique as outlined by Poole (2002) for mark-recapture studies. This approach is computationally intensive; however, this method has some advantages over the frequentist approach. For example, models where survival times are now a function of the travel time are intractable in the classical approach, but should be relatively easy to implement in the Bayesian approach during the latent variable generation phase. On the other hand, model selection is much more complicated in the Bayesian approach.

Future work should include the modelling of the radio-tag survival curve using a parametric approach.

Chapter 3

The Jolly-Seber Model with Tag Loss

Mark-recapture studies are used to estimate population parameters such as abundance and survival. Individuals are marked with unique tags and recaptured at later sample times. The Jolly-Seber (JS) model is used for open populations when the parameter of interest is the population size. Recaptures provide information about survival rates and recapture rates. Capture of unmarked individuals provides information on abundance and births.

A crucial assumption of JS models is that animals do not lose their tags. This assumption is often violated and tag loss can severely bias JS estimates (McDonald, Amstrup and Manly, 2003; Arnason and Mills, 1981). In some studies, individuals are double tagged so that a tag-loss rate may be estimated (see Robson and Regier, 1966; Seber and Felton, 1981; Wetherall, 1982; Barrowman and Myers, 1996). Often an *ad hoc* correction to \hat{N} is made based on the estimated tag-loss rate.

We extend the JS model to incorporate tag loss and estimate tag-retention parameters. This model is applied to walleyes (*Stizostedion vitreum*) that were both double and single tagged in Mille Lacs, Minnesota.

3.1 Design Protocol

The design protocol follows that of traditional JS studies for single-tagged fish with the exception that some fish receive two tags so that tag-retention rates can be estimated. The assumptions for the JS model are outlined in MacDonald et al. (2003). For the Jolly-Seber-Tag-Loss model (JSTL) we do not assume tags are retained, however we do assume that tag loss on a double-tagged individual is independent between tags. Our model is developed based on the super-population approach proposed by Schwarz and Arnason (1996).

We must make an additional assumption concerning animals that lose all of their tags and are recaptured. We will assume that either these animals can be recognised upon recapture (via a permanent mark as in Diefenbach and Alt (1998); or via scarring as in Bradshaw, Barker and Davis (2000)) and are not re-tagged, or there are only a small number of incidences of these fish and they are treated as new fish. Of course, even if scarring is recognized, the tag numbers of fish that lost both tags are unknown. If either is the case, we argue that these fish will have a small effect on our parameter estimates.

3.2 Notation

Statistics:

- k number of sample times.
- m number of unique tag histories.
- i index for a tag history, $i = 0, 1, 2, \dots, m$; $i = 0$ is used to refer to history $\{00 \ 00 \ \dots \ 00\}$, i.e. for animals never caught.
- nt_j number of tags on the animal at sample time j , $j = 1, 2, \dots, k$.
- n_j total animals captured at time j .
- ω_i^* capture history vector $\omega_i^* = (\omega_{i1}^*, \omega_{i2}^*, \dots, \omega_{ik}^*)$, where

$$\omega_{ij}^* = \begin{cases} 1 & \text{if animal } i \text{ was captured for the first time at sample time } f_i. \\ 1 & \text{if animal } i \text{ was captured at sample time } f_i < j \text{ with at least} \\ & \text{one tag present.} \\ 0 & \text{if animal } i \text{ was not captured at sample time } j. \end{cases}$$

ω_i tag history vector $\omega_i = (\omega_{i,1,1}, \omega_{i,1,2}, \omega_{i,2,1}, \omega_{i,2,2}, \dots, \omega_{i,k,1}, \omega_{i,k,2})$,
 $i = 0, 1, \dots, m$, where

$$\omega_{ijd} = \begin{cases} 1 & \text{if animal } i \text{ was captured and tag } d \text{ was present at sample time } j. \\ 0 & \text{if animal } i \text{ was captured and tag } d \text{ was not present at sample time } j. \\ 0 & \text{if animal } i \text{ was not captured at sampling time } j. \end{cases}$$

n_{ω_i} number of ω_i tag histories.

f_i first sample time where $\omega_{ij}^* = 1$.

l_{id} last sample time where tag d was present.

l_i last sample time where $\omega_{ij}^* = 1$; $l_i = \max_d(l_{id})$.

v_j number of animals lost on capture at sample time j .

q_{id} first sample time where tag d was known to be missing for animal i ;
 e.g. for history $\{11\ 00\ 01\}$, $q_{11} = 3$ as tag 1 is not known to be missing at sample time 2.

$$z_i = \begin{cases} 1 & \text{if animals with tag history } i \text{ were injected into the population} \\ & i=1,2,\dots,m. \\ 0 & \text{no injection occurred.} \end{cases}$$

where an injection is defined as an animal entering the population by human means. The number of injections is known by the experimenter and they do not provide information about births.

Parameters:

p_j the conditional probability that an animal is recaptured at sample time j given that the animal is alive, $j = 1, 2, \dots, k$.

- ϕ_j the conditional probability that an animal survives and remains in the population between sample time j and $j + 1$ given it was alive and in the population at time j , $j = 1, 2, \dots, k - 1$.
- b_j fraction of the total number of animals that enter the system between sample times j and $j + 1$, $j = 0, 1, \dots, k - 1$. b_0 is the fraction of the animals alive just prior to the first sample time. Note that $\sum_{j=0}^{k-1} b_j = 1$.
- ν_j the probability that an animal captured at sample time j will be lost on capture, $j = 1, 2, \dots, k$.
- $\Lambda_{f_i, j}$ the conditional probability that an animal first tagged at time f_i retains its tag between sample time j and $j + 1$ given that it is alive.
- N total super-population size, i.e. all animals that ever enter the population of interest during the study.

Functions of Parameters:

- b_j^* the fraction of the population remaining to enter the population that enter between sample times j and $j + 1$, $j = 0, 1, \dots, k - 1$. Note that $b_0^* = b_0$, $b_j^* = \frac{b_j}{\sum_{u=j}^{k-1} b_u}$ and $b_{k-1}^* = 1$.
- $\chi_{f_i, l_i, nt}$ the probability that animal i first seen at f_i , is not seen after sample time l_i , with nt tags. Note that $\chi_{f_i, l_i, nt}$ is a recursive function of ϕ , p , and Λ . If $f_i = 0$, this indicates animals not yet captured but alive at time l_i .

$$\chi_{(0, j, 0)} = \begin{cases} 1 - \phi_j + \phi_j(1 - p_{j+1})\chi_{(0, j+1, 0)} & \text{if } j < k \\ 1 & \text{if } j = k \end{cases}$$

$$\chi_{(f_i, j, 1)} = \begin{cases} 1 - \phi_j + \phi_j(1 - p_{j+1})\Lambda_{f_i, j}\chi_{(f_i, j+1, 1)} + \phi_j(1 - \Lambda_{f_i, j}) & \text{if } j < k \\ 1 & \text{if } j = k \end{cases}$$

$$\chi_{(f_i, j, 2)} = \begin{cases} 1 - \phi_j + \phi_j(1 - p_{j+1})\Lambda_{f_i, j}^2\chi_{(f_i, j+1, 2)} + \\ 2\phi_j(1 - p_{j+1})\Lambda_{f_i, j}(1 - \Lambda_{f_i, j})\chi_{(f_i, j+1, 1)} + \phi_j(1 - \Lambda_{f_i, j})^2 & \text{if } j < k \\ 1 & \text{if } j = k \end{cases}$$

- Ψ_j probability that an animal enters the population, is still alive and is not seen before time j , $j = 1, 2, \dots, k$. $\Psi_1 = b_0$, and $\Psi_{j+1} = \Psi_j(1 - p_j)\phi_j + b_j$ or $\Psi_1 = b_0^*$, and $\Psi_{j+1} = \Psi_j(1 - p_j)\phi_j + \sum_{u=0}^{j-1} (1 - b_u^*)b_j^*$.
- N_j population size at time j . $N_1 = B_0$, $N_{j+1} = (N_j - N_j p_j v_j + \sum_{i=1}^{n_{obs}} I[f_i = j] z_i n_{\omega_i})\phi_j + B_j$.
- B_j number of animals that enter the population after sample time j and survive to sample time $j + 1$; $j = 0, 1, \dots, k - 1$. B_0 is the number of animals alive just prior to the first sample time. $B_j = N b_j$

Latent variables

- \mathbf{a}_i the alive status vector for animal i , $a_i = \{a_{i1}, a_{i2}, \dots, a_{ik}\}$, $i = 0, 1, \dots, m$, where

$$a_{ij} = \begin{cases} 1 & \text{if animal } i \text{ is alive at sample time } j. \\ 0 & \text{if animal } i \text{ is dead at sample time } j. \end{cases}$$

- \mathbf{g}_i the tag status vector for animal i , $g_i = \{g_{i11}, g_{i12}, g_{i21}, g_{i22}, \dots, g_{ik1}, g_{ik2}\}$, $i = 0, 1, \dots, m$, where

$$g_{ijd} = \begin{cases} 1 & \text{if tag } d \text{ is present on animal } i \text{ at sample time } j. \\ 0 & \text{if tag } d \text{ is not present on animal } i \text{ at sample time } j. \end{cases}$$

$$g_{ij+} = \begin{cases} 1 & \text{if at least one tag is present animal } i \text{ at sample time } j. \\ 0 & \text{if no tags are present on animal } i \text{ at sample time } j. \end{cases}$$

3.3 Model Development

3.3.1 Likelihood Development

As is commonly done in mark-recapture models, the likelihood can be partitioned into components dealing with first capture, subsequent recaptures, and loss on capture. Schwarz and Arnason (1996) split the likelihood into 3 parts— $P(\text{first recapture}|\{p_j\}, \{\phi_j\}, \{b_j\})$, $P(\text{recapture}|\{p_j\}, \{\phi_j\})$ and $P(\text{loss on capture}|\{\nu_j\})$. Rather than separate the first two components, we combine them so that the notation makes use

of the tag histories. We first model the total number of observed tag histories ($n_{obs} = \sum_{i=1}^m n_{\omega_i}$ and $N - n_{obs} = n_{\omega_0}$) by

$$\begin{aligned}
L_1^A &= \binom{N}{n_{obs}} P(\{00\dots 00\})^{(N-n_{obs})} (1 - P(\{00\dots 00\}))^{n_{obs}} \\
&= \binom{N}{n_{obs}} \left\{ \sum_{j=0}^{k-1} b_j (1 - p_{j+1}) \chi_{(0,j+1,0)} \right\}^{(N-n_{obs})} \times \\
&\quad \left\{ 1 - \sum_{j=0}^{k-1} b_j (1 - p_{j+1}) \chi_{(0,j+1,0)} \right\}^{n_{obs}}
\end{aligned} \tag{3.1}$$

This can be thought of as modelling the number of observed tag histories (animals) in a super-population of size N . $P(\{00\dots 00\})$ is the probability of never being seen.

Then we model the number of each observed tag history as a multinomial conditional upon being seen at least once such that

$$\{n_{\omega_1}, n_{\omega_2}, \dots, n_{\omega_m}\} \sim \text{Multinomial}(n_{obs}, P(\omega_1), P(\omega_2), \dots, P(\omega_m)) \times \frac{1}{(1 - P(\{00\dots 00\}))^{n_{obs}}}.$$

$$\begin{aligned}
L_1^B &= \binom{n_{obs}}{n_{\omega_1} n_{\omega_2} \dots n_{\omega_m}} \frac{P(\omega_1)^{n_{\omega_1}} P(\omega_2)^{n_{\omega_2}} \dots P(\omega_m)^{n_{\omega_m}}}{(1 - P(\{00\dots 00\}))^{n_{obs}}} \\
&= \binom{n_{obs}}{n_{\omega_1} n_{\omega_2} \dots n_{\omega_m}} \prod_{i=1}^m \left[\Psi_{f_i} \left\{ \prod_{j=f_i}^{l_i} p_j^{\omega_{ij}^*} (1 - p_j)^{(1 - \omega_{ij}^*)} \right\} \left\{ \prod_{j=f_i}^{l_i-1} \phi_j \right\} \times \right. \\
&\quad \left. \prod_{d=1}^2 \left\{ \left(\prod_{j=f_i}^{l_{id}-1} \Lambda_{f_{ij}} \right) \left(1 - \prod_{j=l_{id}}^{q_{id}-1} \Lambda_{f_{ij}} \right)^{I(l_{id} \neq l_i)} \right\} \times \chi_{(f_i, l_i, n_{l_i})} \right]^{n_{\omega_i}} \\
&\quad \left\{ 1 - \sum_{j=0}^{k-1} b_j (1 - p_j) \chi_{(0,j+1,0)} \right\}^{-n_{obs}}
\end{aligned} \tag{3.3}$$

For single-tagged animals, $\omega_{i,j,2}$ will be zero for all j and they will only contribute one $\Lambda_{f_i,j}$ to the likelihood.

Loss on capture is modelled as a binomial as done by Schwarz and Arnason (1996) such that

$$L_3 = \prod_{j=1}^k \left\{ \binom{n_j}{v_j} (1 - \nu_j)^{v_j} \nu_j^{(n_j - v_j)} \right\} \quad (3.4)$$

The likelihood for the JSTL model is the product of the components.

$$L = L_1^A \times L_1^B \times L_3 \quad (3.5)$$

3.3.2 Parameter Estimation

Estimates could be found using standard numerical methods. However, direct maximization could become numerically unstable with large numbers of parameters. Alternatively, Van Deusen(2002) showed that the EM method leads to a decomposition of the maximization problem into independent subproblems. The same occurs in this study. The (complete data) likelihood reduces to separate independent components for catchability, survival, new entrants, and tag-retention.

The EM algorithm is broken into two steps. The E step computes the expectation of the log-likelihood given the current parameter estimates and the observed data. The M step maximizes the expected complete-data likelihood. Convergence of the algorithm to maximum likelihood estimates (MLE) occurs under mild regularity conditions (Van Deusen, 2002).

The complete data for the JSTL model involves the observed variables and the latent (unobservable) variables: the capture histories ω^* ; the tag histories for both known and missing data ω ; the alive status of the animal \mathbf{a} ; the tag status of the animal \mathbf{g} ; the injection status z ; and whether or not the animal was lost on capture. The injection and loss on capture status are known exactly. Examples of statistics and latent variables are shown in Table 3.1. Note that we must distinguish between capture histories and tag histories as different tag histories can have the same capture history (Table 3.1).

The complete-data likelihood for the JSTL model is:

Table 3.1: Examples of statistic and latent variables for double-tagged fish in a 3 sample experiment.

Capture History	Tag History	Tag Status	Alive Status	Injection Status	Loss on Capture
ω^*	ω	g	a	z	l
111	11 11 11	1 1 1 1 1 1	1 1 1	0	0
101	11 00 11	1 1 1 1 1 1	1 1 1	0	0
101	11 00 01	1 1 g_1 1 0 1	1 1 1	0	0
101	11 00 10	1 1 1 g_2 1 0	1 1 1	0	0
110	11 11 00	1 1 1 1 $g_1 g_2$	1 1 a_3	0	0
010	00 11 00	0 0 1 1 $g_1 g_2$	a_1 1 0	0	1
110	11 10 00	1 1 1 0 1 g_2	1 1 a_3	0	0
101	11 00 01	1 1 g_1 1 0 1	1 1 1	1	0
100	11 00 00	1 1 $g_1 g_2 g_1 g_2$	1 $a_2 a_3$	1	0

$$\begin{aligned}
L_{complete} \propto & \prod_{i=0}^m \left\{ \prod_{j=1}^{f_i-1} p_j^{a_{ij}\omega_{ij}^*} (1-p_j)^{a_{ij}(1-\omega_{ij}^*)} \right\}^{n_{\omega_i}} \times \\
& \left\{ \prod_{j=f_i}^k p_j^{a_{ij}g_{ij}+\omega_{ij}^*} (1-p_j)^{a_{ij}g_{ij}+(1-\omega_{ij}^*)} \right\}^{n_{\omega_i}} \times \\
& \left\{ \prod_{j=1}^{k-1} \phi_j^{a_{ij}a_{i,j+1}} (1-\phi_j)^{a_{ij}(1-a_{i,j+1})} \right\}^{n_{\omega_i}} \times \\
& \left\{ \prod_{j=f_i}^{k-1} \Lambda_{f_{ij}}^{\sum_{d=1}^2 a_{i,j+1}g_{ij}d} g_{i,j+1,d} (1-\Lambda_{f_{ij}})^{\sum_{d=1}^2 a_{i,j+1}g_{i,j,d}(1-g_{i,j+1,d})} \right\}^{n_{\omega_i}} \times \\
& \left\{ \prod_{j=0}^{k-2} b_j^{*(1-a_{ij})a_{i,j+1}} (1-b_j^*)^{\sum_{u=j+1}^{k-1} (1-a_u)a_{u+1}} \right\}^{n_{\omega_i}} \times \\
& \left\{ \prod_{j=1}^k \binom{n_j}{v_j} (1-\nu_j)^{v_j} \nu_f^{n_j-v_j} \right\} \tag{3.6}
\end{aligned}$$

Note that the unknown population size (N) is captured by the n_{ω_0} term. The complete-data likelihood is a product of binomial likelihoods determined by the alive and tag status variables.

The E step of the EM algorithm involves computing the expected complete-data log-likelihood (3.7).

$$\begin{aligned}
E(\ln(L)) \propto & \sum_{i=0}^m n_{\omega_i} \sum_{j=1}^{f_i-1} \{E(a_{ij})\omega_{ij}^* \ln(p_j) + E(a_{ij})(1 - \omega_{ij}^*) \ln(1 - p_j)\} + \\
& n_{\omega_i} \sum_{j=f_i}^k \{E(a_{ij}g_{ij+})\omega_{ij}^* \ln(p_j) + E(a_{ij}g_{ij+})(1 - \omega_{ij}^*) \ln(1 - p_j)\} + \\
& n_{\omega_i} \sum_{j=1}^{k-1} \{E(a_{ij}a_{i,j+1}) \ln(\phi_j) + (E(a_{ij}) - E(a_{ij}a_{i,j+1})) \ln(1 - \phi_j)\} + \\
& n_{\omega_i} \sum_{j=f_i}^{k-1} \left\{ \sum_{d=1}^2 E(a_{i,j+1}g_{ij,d}g_{i,j+1,d}) \right\} \ln(\Lambda_{f_i,j}) + \\
& \left\{ \sum_{d=1}^2 E(a_{i,j+1}g_{i,j,d}) - E(a_{i,j+1}g_{i,j,d}g_{i,j+1,d}) \right\} \ln(1 - \Lambda_{f_i,j}) + \\
& n_{\omega_i} \sum_{j=0}^{k-2} \{E(a_{i,j+1}) - E(a_{ij}a_{i,j+1}) \ln(b_j^*) + \\
& \left(\sum_{u=j+1}^{k-1} E(a_{u+1}) - E(a_u a_{u+1}) \right) \ln(1 - b_j^*) \} + \\
& \sum_{j=1}^k \left\{ \ln \binom{n_j}{v_j} v_j \ln(1 - \nu_j) + (n_j - v_j) \ln(\nu_j) \right\} \tag{3.7}
\end{aligned}$$

Van Deusen (2002) illustrates how to estimate $E(a_{ij})$ for the Cormack-Jolly-Seber model. We use similar methods to derive the expected statistics in (3.7) in Appendix B.

The M step involves maximizing the complete-data log-likelihood over the observed and missing data. This is commonly done using the Newton-Raphson method (NR). Each parameter set (e.g. the p 's, ϕ 's, b^* 's, and the Λ 's) can be maximized separately conditional on the current parameter values and the complete data as the complete-data likelihood is completely separable.

Unknown probabilities such as p_j , ϕ_j , b_j^* , and $\Lambda_{f_i,j}$ are often modelled using link

functions and a reduced parameter set with design matrices allowing for implementation of flexible models. The logit-link function is often used as it constrains probabilities to be between 0 and 1 (Lebreton et al. 1992) although other functions are possible. Here we let $h(\beta^p) = \ln\left(\frac{p_j}{1-p_j}\right) = X_{ij}^p \beta^p$, $h(\beta^\phi) = \ln\left(\frac{\phi_j}{1-\phi_j}\right) = X_{ij}^\phi \beta^\phi$, $h(\beta^{b^*}) = \ln\left(\frac{b_j^*}{1-b_j^*}\right) = X_{ij}^{b^*} \beta^{b^*}$ and $h(\beta^\Lambda) = \ln\left(\frac{\Lambda_{fj}}{1-\Lambda_{fj}}\right) = X_{ij}^\Lambda \beta^\Lambda$ as outlined by Van Deusen(2002). We use parameter identification matrices (PIM) as done in MARK (White and Burnham, 1999) to specify the various models for the underlying parameters.

This leads to maximizing for the β parameters rather than the standard p , ϕ , b^* and Λ parameters. The score vector (using the expected complete-data log-likelihood) for the j^{th} capture component using a link function is $E\left(\frac{dl}{d\beta_j^p}\right) = E\left(\frac{dl}{dp_j} \frac{dp_j}{dh} \frac{dh}{d\beta_j^p}\right)$. The j^{th} capture component for the Fisher information matrix of the expected complete-data log-likelihood is $E\left(-\frac{d^2l}{d\beta_j^{p2}}\right) = E\left\{-\frac{dh}{d\beta_j^p} \frac{dp_j}{dh} \frac{d}{dp_j} \left(\frac{dl}{dp_j} \frac{dp_j}{dh} \frac{dh}{d\beta_j^p}\right)\right\}$.

The EM algorithm can be summarized as:

- 1 Determine initial values for the parameters.
- 2 Compute the χ values.
- 3 Compute the Ψ values.
- 4 Compute $\hat{P}(\{00\dots 00\})$, \hat{N} and $\hat{n}_{\{00\dots 00\}}$ using $\hat{N} = \frac{n_{obs}}{1-\hat{P}(\{00\dots 00\})}$ and $\hat{n}_{\{00\dots 00\}} = \hat{N} - n_{obs}$.
- 5 E-step: Determine expectations of unknown latent variables using Appendix B.
- 6 M1-step: Maximize the capture component.
- 7 M2-step: Maximize the survival component.
- 8 M3-step: Maximize the birth component.
- 9 M4-step: Maximize the tag-retention component.
- 10 Compute $\ln L$ and check the convergence criterion compared to the previous iteration. If convergence has not been met, goto step 2, otherwise stop.

Standard errors of parameter estimates can be computed via a numerical delta method. First we use central differencing equations (Abramowitz and Stegun, 1972, p. 884) to find second order partial derivatives of the likelihood (3.5) with respect to the β parameters. The central differencing equations provided by Abramowitz and Stegun (1972, p. 884) are

$$\frac{d^2l}{d\beta_i^2} = -\frac{\{l(\beta_i + 2\epsilon_{1i}) + 16l(\beta_i + \epsilon_{1i}) - 30l(\beta_i - \epsilon_{1i}) + 16l(\beta_i - 2\epsilon_{1i})\}}{12\epsilon_{1i}^2} \quad (3.8)$$

and

$$\begin{aligned} \frac{d^2l}{d\beta_i d\beta_j} = & -\frac{\{l(\beta_i + \epsilon_{1i}, \beta_j + \epsilon_{2j}) - l(\beta_i + \epsilon_{1i}, \beta_j - \epsilon_{2j}) - l(\beta_i - \epsilon_{1i}, \beta_j + \epsilon_{2j})\}}{4\epsilon_{1i}\epsilon_{2j}} + \\ & \frac{\{l(\beta_i - \epsilon_{1i}, \beta_j - \epsilon_{2j})\}}{4\epsilon_{1i}\epsilon_{2j}} \end{aligned} \quad (3.9)$$

where $\epsilon_{1i} = \beta_i/1000$ and $\epsilon_{2j} = \beta_j/1000$; $l(\beta_i + \epsilon_{1i})$ is the cumulative log-likelihood evaluated with $\hat{\beta}_i$ adjusted by ϵ_{1i} . Similarly, $l(\beta_i + \epsilon_2, \beta_j + \epsilon_2)$ is the cumulative log likelihood with $\hat{\beta}_i$ adjusted by ϵ_{1i} and $\hat{\beta}_j$ adjusted by ϵ_{2j} . Central differencing evaluates the likelihood at 4 points around the MLE and uses the differences in the likelihoods to estimate partial derivatives (Thisted, 1988, p.198). This provides covariance estimates for the β parameters. Then we use the delta method to obtain covariance estimates for the standard parameters.

Derived estimates (e.g. \hat{N}_i) are computed in the usual fashion and estimated standard errors are found using the delta method (Schwarz and Arnason, 1996).

As with the JS model, care must be taken to deal with confounded parameters. For example, only the products $\phi_{k-1}p_k$ and Nb_0p_1 can be estimated. In JSTL models, all tag-retention parameters are identifiable. Confounded parameters can be detected using the singular value decomposition as outlined in White and Burnham (1999).

FORTRAN 95 computer code to implement parameter estimation as described above is available. Extensive simulation studies were used to exercise the code to validate the implementation (Appendix C). The program allows a great deal of flexibility in fitting tag-loss models. Through the use of PIMs, tag-loss models are fit in a similar way to survival models in the CJS model.

3.3.3 Per Unit Time

It is sometimes the case that sampling periods are not evenly distributed and an estimate of a parameter (e.g. survival) per unit time rather than for an entire sample period is desired. For instance, if our sample times occurred at $\{1, 4, 9\}$, then the time periods (Δt_i) would be $\{3, 5\}$. ϕ_1 would represent 3 units of time and ϕ_2 would represent 5 units of time. In these instances, survival rates are not easily compared. We could estimate survival per unit time such that $\phi_1 = \phi_1^{\dagger 3}$ and $\phi_2 = \phi_2^{\dagger 5}$ (i.e. $\phi_i = \phi_i^{\dagger \Delta t_i}$). In this way, ϕ_1^\dagger and ϕ_2^\dagger are estimates of survival per unit time.

When modelling parameters per unit time, a few changes are needed. First, Δt_i values need to be calculated. Second, the score vectors are affected by per unit time such that $\frac{dl}{d\beta^\dagger} = \frac{dl}{d\phi} \frac{d\phi}{d\phi^\dagger} \frac{d\phi^\dagger}{dg(\phi^\dagger)} \frac{dg(\phi^\dagger)}{d\beta^\dagger}$. Third, the information matrix is also adjusted resulting in $\frac{d^2 l}{d\beta^{\dagger 2}} = \frac{dg(\phi^\dagger)}{d\beta^\dagger} \frac{d\phi^\dagger}{dg(\phi^\dagger)} \frac{d\phi}{d\phi^\dagger} \frac{d}{d\phi} \left\{ \frac{dl}{d\phi} \frac{d\phi}{d\phi^\dagger} \frac{d\phi^\dagger}{dg(\phi^\dagger)} \frac{dg(\phi^\dagger)}{d\beta^\dagger} \right\}$. Finally, the covariances are also affected by the transformation to per unit time. We initially find the covariances of β^\dagger using the central differencing equations as before. This leads to finding $Cov(g(\phi^\dagger)) = Cov(X\beta^\dagger)$. Notice that $Cov(\phi) = Cov(\phi^{\dagger \Delta t})$ which makes the delta method to find the $Cov(\phi)$ become:

$$Cov(\phi) = \left\{ \frac{d\phi}{d\phi^\dagger} \frac{\phi^\dagger}{dg(\phi^\dagger)} \right\}' Cov(X\beta^\dagger) \left\{ \frac{d\phi}{d\phi^\dagger} \frac{\phi^\dagger}{dg(\phi^\dagger)} \right\} \quad (3.10)$$

where $\frac{d\phi}{d\phi^\dagger} = \Delta t_i \phi_i^{(\Delta t_i - 1)} = \Delta t_i \phi_i^{1 - \frac{1}{\Delta t_i}}$.

3.3.4 Goodness of Fit

The goodness-of-fit test involves the likelihood of the model of interest conditional on being observed. These conditional probabilities are found by normalizing the probability of each history (after adjustment for the probability of being tagged with one or two tags when released) by the probability of being observed somewhere in the experiment. We compare this fitted model to the saturated multinomial model where every ω_i history has probability $\frac{n_{\omega_i}}{n_{obs}}$:

$$L_{fitted} \propto \left\{ \prod_{i=1, double}^m \left(\frac{P(\omega_i)}{1 - P(00..00)} \right)^{n_{\omega_i}} P(double) \right\} \times \left\{ \prod_{i=1, single}^m \left(\frac{P(\omega_i)}{1 - P(00..00)} \right)^{n_{\omega_i}} (1 - P(double)) \right\} \quad (3.11)$$

$$L_{saturated} \propto \prod_{i=1}^m \left(\frac{n_{\omega_i}}{n_{obs}} \right)^{n_{\omega_i}} \quad (3.12)$$

Twice the difference in the log-likelihood of the saturated model and the model of interest is asymptotically χ^2 distributed. The degrees of freedom is the difference in parameters between the two models. The number of parameters in the saturated model is the number of unique tag histories minus 1. The number of parameters in the fitted model includes those for catchability, survival, new entrants, and tag-retention plus those k parameters for losses on capture and k parameters for the $P(\text{double tagging})$. The goodness-of-fit statistic can be used to derive a variance inflation factor in the usual way (Burnham and Anderson, 2002, p. 68).

3.3.5 Model Selection

Many different models can be specified for the parameter sets. Using notation similar to Lebreton et al. (1992), common models include time-varying parameters and parameters equal over time. Of particular interest are models for the tag-retention rates as these can vary across release groups and across time. Common models are:

- $\Lambda_{.,t}$ Tag-retention rates vary across time but are homogeneous across release group. This was considered by Arnason and Mills (1981) who showed that in the ordinary JS model, estimated survival ($\hat{\phi}_j$), estimated births and standard error of estimated births are biased under a model open to recruitment and that estimated abundance (\hat{N}_j), $\hat{\phi}_j$, and $SE(\hat{\phi}_j)$ are biased under a death only model.
- $\Lambda_{.,T}$ Tag-retention rate is a function of time since application, but homogeneous across release groups.
- $\Lambda_{.,1+t}$ Tag-retention rate has an initial acute tag loss, but are then homogeneous across release group at subsequent intervals.
- $\Lambda_{t,t}$ fully general model for tag loss where every release has a different tag retention rate in subsequent intervals and over time.

The Akaike information criterion (AIC) and related methods (model averaging) can be applied as outlined in Burnham and Anderson (2002).

3.4 Example

In 2002, the Minnesota Department of Natural Resources began a multi-year mark-recapture program to estimate the number of walleye greater than 14 inches in length in Mille Lacs, Minnesota. In each year, fish were tagged in the spring on the spawning grounds. During the summer, members of the team also tagged fish that were captured on launches and boats. Recoveries of tags also occurred from a spring tribal harvest and a fall gillnet assessment survey.

T-bar anchor tags with individual serial numbers were used to mark between 20,000 and 30,000 fish per year. Approximately 6% of the fish were double tagged. Complete details of the experimental protocol are found in Bruesewitz (2002).

This example uses a subset of the data collected between 2002 and 2004 to illustrate the methodology. These do not represent the official estimates of population size due to a number of technical issues with the study (see Schwarz, 2004).

Each of the 3 years was divided into 2-sample intervals, roughly corresponding

to the spring tagging and tribal harvests, and the summer fishing seasons for each year. This makes for 6-sample intervals in the experiment. The mid-point of each interval was taken as the time of sampling. The lower bound for the fish length to delineate members of the target population was raised in each season based on growth profiles from recaptured fish to limit recruitment into the population from fish growing during the season - see Schwarz (2003) for details. This allows N_1 to be estimated in all models as b_1 was fixed to 0.

The full set of individual capture histories is available in Appendix D; summary statistics are shown in Table 3.2. Note that these statistics are not sufficient statistics to estimate the parameters of the models, the individual capture histories were used.

A summary of the tag-loss statistics from the double-tagged fish is shown in Table 3.3. Tag loss appears to happen primarily over the winter months. The cumulative tag-loss rates also appear to stabilize over time. This could be due to high initial tag loss followed by high tag retention. The homogeneous tag-loss model seems inappropriate.

Several models were fit to the dataset as summarized in Table 3.4 including a simple Jolly-Seber model based on a single tag only. This simple model cannot be compared to the other models via AIC as it uses a different dataset (the second tag history was ignored). However, we can compare the models that do involve the tag-retention parameters (Λ) and we can compare abundance estimates for all models.

Some care must be taken when trying to analyse the data from a double-tagging experiment using a single-tag Jolly-Seber model. It is not sufficient to simply ignore the second tag. Some accounting must be made for fish that lost tags and are subsequently recaptured; these fish look like unmarked fish, but are merely ‘recycled’.

The simple JS model produced abundance estimates that are substantially different from any of the tag-loss models. Estimates of abundance in the first-tag periods are substantially reduced among tag-loss models because of the high tag-loss experience after the second-tag periods. Because of the large amount of data, these simple tag-loss models appear to be inappropriate.

Based on the AIC, the full JSTL model $(p_t, \phi_t, b_t, \Lambda_{t,t})$ is the most appropriate model of this set. However, the JSTL model with no trends in tag loss has poor fit

Table 3.2: Summary statistics and standardized sample times for the 6 periods of the Mille Lacs data.

Period ^a	Time ^b	Fish Captured n	Fish without marks u	Fish with marks m	Fish tagged and released R	Fish released, later recaptured r	Fish captured before and after but not at period z
1	0.00	49,242	49,242	0	19,750	3,302	-
2	0.17	13,877	13,644	233	11,991	872	3,069
3	1.00	64,573	61,884	2,689	29,412	3,962	1,252
4	1.17	63,86	5,815	571	4,573	275	4,643
5	2.00	51,940	47,128	4,812	19,457	84	106
6	2.17	21,15	1,925	190	882	-	-

^a Periods 1, 3, and 5 occur during spawning in 2002, 2003, and 2004 respectively. Periods 2, 4, and 6 occur during the angling season in 2002, 2003, and 2004 respectively.

^b Time is measured in years.

Table 3.3: The number of double-tagged fish released and subsequent recaptures with either double (first row) or single (second row) tags present. Cumulative tag-loss rates based on the methods presented in Seber and Felton (1981) for 2 indistinguishable tags are provided (third row).

Period	Released	Recaptured				
		2	3	4	5	6
1	1364	15	98	4	50	1
		0	111	7	52	1
		0.00	0.36	0.47	0.34	0.33
2	1137		23	6	10	0
			28	6	22	1
			0.38	0.33	0.52	1.00
3	1934			28	142	2
				1	115	3
				0.02	0.29	0.43
4	479				13	2
					7	1
					0.21	0.20
5	1263					5
						2
						0.17

Table 3.4: Log-likelihood (l), Akaike information criterion (AIC), the number of model parameters (P), Δ AIC and, population estimates (\hat{N}) and standard errors for several different models. Both \hat{N}_i and $SE(\hat{N}_i)$ are reported in thousands of fish.

Model ^a	P	l^b	AIC ^b	Δ AIC	N_1	N_2	N_3	N_4	N_5
JS	20	-	-	-	905	876	379	344	322
					15	15	9	9	40
$b_t^* p_t \phi_t \Lambda$	21	322	-601	855	725	695	377	341	306
					14	14	10	18	42
$b_t^* p_t \phi_t \Lambda_{.,t}$	25	348	-646	810	715	684	385	348	288
					16	16	9	9	31
$b_t^* p_t \phi_t \Lambda_{w,t}$	24	343	-638	818	713	683	383	347	301
					16	16	9	9	33
$b_t^* p_t \phi_t \Lambda_{t,.}$	25	641	-1231	225	657	627	372	333	303
					14	14	10	18	42
$b_t^* p_t \phi_t \Lambda_{.,1+t}$	22	725	-1407	49	572	524	322	284	290
					12	28	8	10	35
$b_t^* p_t \phi_t \Lambda_{s,t}$	25	737	-1424	32	577	481	340	298	305
					12	34	10	19	37
$b_t^* p_t \phi_t \Lambda_{t,t}$	35	763	-1456	0	545	481	321	272	209
					17	30	12	15	29

JS	Jolly-Seber model.
$b_t^* p_t \phi_t \Lambda$	Constant tag loss.
$b_t^* p_t \phi_t \Lambda_{.,t}$	Homogeneous across release group but variable over time.
$b_t^* p_t \phi_t \Lambda_{w,t}$	Homogeneous across release group, with constant summer tag loss but variable winter tag loss.
$b_t^* p_t \phi_t \Lambda_{t,.}$	Constant over time, variable over initial tagging group.
$b_t^* p_t \phi_t \Lambda_{.,1+t}$	Initial tag loss, followed by constant tag loss, homogeneous across group.
$b_t^* p_t \phi_t \Lambda_{s,t}$	Tag loss constant over sampling period (s), but variable over time.
$b_t^* p_t \phi_t \Lambda_{t,t}$	No trends in tag loss, different parameters for each tagging group, sampling period and across time.

(a) Because of the gradual raising of the size limit in each year to account for growth, recruitment and immigration was minimal within each year. We were able to set b_1^* to 0 to enable estimation of the initial population size.

(b) To simplify the table, 1582000 was subtracted from the log-likelihood values reported and 3164000 was added to the AIC values reported.

($\chi^2 = 5657, df = 126, p = 0.0$).

The Mille Lac data has heterogeneity in catchability. Fish tagged on the spawning grounds are mostly male while those in the launch phase are mostly female. This results in estimates of abundance for the 3 year study (6 sample times) being quite different from those produced from one year of data (2 sample times). When males are randomly discarded to make the sex-ratio equal in both phases, estimates of the 6-sample experiment are similar to those of the 2-sample experiments. The fit of the model also substantially improves but is not entirely satisfactory. There is clearly a need for a more complex model (potentially stratified by area of release) to capture this heterogeneity. Despite the problems with the fit of the model, the message remains clear - tag loss is not homogenous over time and the estimates differ substantially across models. Tag loss cannot be ignored.

3.5 Discussion

By incorporating tag-retention parameters into the Jolly-Seber model, we have eliminated the need to make *ad hoc* corrections to estimate population size in double-tagging experiments. Furthermore, structuring tag-loss rates as a function of both release group and time allows great flexibility in modelling tag loss.

There are still difficulties. When fish lose all of their tags and are recaptured, they are in a sense ‘recycled’ and appear to be immigrants. Consequently, estimates of recruitment and abundance will be biased upwards. If the number of such fish is small (e.g. high tag-retention or low recapture rates) this bias will be small. Alternatively, if tag-retention rates were low and recapture rates were high, this bias could be large. If this was the case, it would be better to redesign the study to reduce tag loss.

A few modifications remain for future work. The JSTL model can be extended to multiple groups (e.g. males and females) as has been done for the JS model (Schwarz and Arnason, 1996). Incorporation of different rates of tag loss for individual tags would deal with animals given two types of tags with possibly different tag-loss rates (e.g. permanent tatoos and ear clips). Finally, such a model could handle dependent tag-loss parameters, such as the case of black bears who’s rate of tag loss of the second

tag depends on that of the first (Diefenbach and Alt, 1998).

Chapter 4

A Migration Model with Tag Loss

The northern limit of yellowtail flounder (*Limanda ferruginea*) inhabits the Grand Banks off the eastern coast of Newfoundland (Walsh et al. 2001). It was thought that yellowtail flounder are sedentary. However, there is evidence to suggest that seasonal migration exists in the Gulf of St. Lawrence stock where fish move to warmer deeper waters in the winter (Poirier et al. 1997). Walsh et al. (2001) found similar migration in the Grand Banks stock.

Stratification in mark-recapture experiments can be useful when modelling migration. Animals are captured, tagged with unique tags and released in one stratum and recaptured at a later time in the same or different strata. Stratification is a form of the more general multi-state models that permit stochastic transitions among states. Animals move stochastically from one stratum to another. Included in these models are transition or migration probabilities along with the standard survival and capture parameters. These models are a first order Markov process as an animal in state $i + 1$ is stochastically determined as a function of its state at time i (Williams et al. 2001; p.454).

Hilborn (1990) presents a general framework using a Poisson model for the analysis of movement, mortality and harvest from tagging. They do not consider tag-loss and catchability and movement parameters are assumed constant over time.

Schwarz, Schweigert and Arnason (1993) extend Hilborn's model so that catchability and movement are dependent on time. To do so, they introduce a matrix

formulation, however; tag-loss and reporting rates are not modelled separately from recovery rates.

In order to get an estimate for tag-reporting rates, tags of differing value are used in the experiment. It is assumed that all high-reward tags will be returned if recovered. The return rate of all other tags are compared to those from high-reward tags. Brattey and Cadigan (1998), and Cadigan and Brattey (1999), developed such a system to estimate tag-reporting rates. They proposed a simple migration model to estimate exploitation rates of cod.

We develop a model whose primary purpose is to model exploitation rates. Furthermore, we incorporate migration, natural survival, tag-loss, and reporting rates into the model. We apply this model to yellowtail flounder off the Grand Banks of Newfoundland.

4.1 Notation

Matrix Operations and Special Matrices:

\mathbf{I} the identity matrix.

$D(X)$ transforms the column vector X into a matrix by placing the elements of X along the diagonal.

$$\text{If } N = \begin{bmatrix} 100 \\ 100 \end{bmatrix}, \text{ then } D(N) = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}$$

Statistics:

a the number of strata in which releases and recoveries take place.

k the number of sample periods.

m the number of tag types.

\mathbf{N}_i^g an $a \times 1$ vector whose s^{th} element N_i^{gs} represents the number of animals tagged and released in stratum s in period i with type g tags.

\mathbf{R}_{ij}^g an $a \times a$ matrix whose $(s, t)^{th}$ element R_{ij}^{gst} is the number of animals released in stratum s at time i that are recovered in stratum t at time j with tag type g .

Table 4.1: Tag type classifications used in the yellowtail flounder tagging study.

Tag Type	Description
1	Single tag, low reward
2	Double tag, recovered as single tag, low reward.
3	Double tag, recovered as double tag, low reward.
4	Single tag, high reward.

Parameters:

- ϕ_i an $a \times a$ diagonal matrix whose $(s, s)^{th}$ element ϕ_i^{ss} represents the probability that an animal alive in stratum s at period i will survive to period $i + 1$.
- \mathbf{P}_i an $a \times a$ diagonal matrix whose $(s, s)^{th}$ element p_i^{ss} represents the probability that an animal present in stratum s is exploited at period i .
- λ_i^g a $a \times a$ diagonal matrix whose $(s, s)^{th}$ element λ_i^{gss} represents the tag reporting rate of an animal captured in stratum s at time i with tag type g .
- Λ_i a $a \times a$ diagonal matrix whose $(s, s)^{th}$ element Λ_i^{ss} represents the tag-retention rate of an animal alive in stratum s at time i until time $i + 1$.
- \mathbf{Q}_i a $a \times a$ matrix whose $(s, t)^{th}$ element q_i^{st} represents the probability that an animal tagged in stratum s , will migrate to stratum t between time i and $i + 1$. Note that the rows of \mathbf{Q}_i sum to 1 such that $\sum_{t=1}^a q_i^{st} = 1$.

4.2 Design Protocol

The study area is broken into a strata. Upon capture, animals are tagged with one of the tag types. For the flounder study, tag types are shown in Table 4.1. Tag types are a convenient mechanism for modelling tag loss. Obviously, an animal is not initially designated as losing a tag; however, as will be seen later, this provides a convenient mechanism for computing the expected number of such recoveries. At the end of the study, tag-type 2 indicates that a double-tagged animal has lost a tag.

Release and recovery must occur in all strata in order to estimate movement

parameters (Schwarz et al. 1993). Recoveries occur when fish are harvested. Tags are removed from the animal and exchanged for some reward. High reward tags are assumed to be reported at a rate of 100%.

The data is in the form of tag histories, similar to that of chapter 3. For example, a tag history of {33 00 20} would indicate that an animal was first caught and double tagged in stratum 3 at time 1, not captured at time 2 and captured at time 3 with only 1 tag in stratum 2. Because fish are never released when recovered, tag histories can be summarized in terms of the number of fish released (\mathbf{N}_i^g) and the number of fish recovered (\mathbf{R}_{ij}^g). Moreover, because recovery rates are low, we model the number of fish recovered R_{ij}^{gst} as independent Poisson random variables following Hilborn (1990).

Assumptions for the Poisson models are the same as for the Jolly-Seber models with respect to instantaneous sampling and emmigration. However, we no longer assume that there is no tag loss. Further, we assume the following:

- Every marked animal present in stratum s at sample time i has the same probability of capture.
- Every marked animal immediately following period i has the same probability of surviving until period $i + 1$.
- Every marked animal in stratum s , period i has the same probability of migration to stratum t by period $i + 1$.
- Every marked animal immediately following period i has the same probability of tag retention until period $i + 1$.

4.3 Model Development

The number of tag recoveries is modelled by a Poisson likelihood:

$$L(\theta) = \prod_{g=1}^m \prod_{i=1}^{k-1} \prod_{j=i+1}^k \prod_{s=1}^a \prod_{t=1}^a \frac{e^{-\theta_{ij}^{gst}} \theta_{ij}^{gst R_{ij}^{gst}}}{R_{ij}^{gst}!} \quad (4.1)$$

where $\theta_{ij}^{gst} = E(R_{ij}^{gst})$. Thus the flexibility of the model comes in describing the expected recoveries. For single-tagged animals ($g = 1$ or 4), $E(\mathbf{R}_{ij}^g)$ is modelled as

$$E(\mathbf{R}_{ij}^g) = D(\mathbf{N}_i)\Lambda_i\phi_i\mathbf{Q}_i(\mathbf{I} - \mathbf{P}_{i+1})\dots\Lambda_{j-1}\phi_{j-1}\mathbf{Q}_{j-1}\mathbf{P}_j\lambda_j^g \quad (4.2)$$

Thus the expected recoveries are the number of released animals at time i (\mathbf{N}_i) that retained their tag (Λ_i), survived (ϕ_i), migrated (\mathbf{Q}_i) from i to $i + 1$, and were not captured ($\mathbf{I} - \mathbf{P}_{i+1}$) at time i . Further, they retained their tag, survived, migrated and were not captured up to time $j - 1$ where they retained their tag, survived, migrated and finally were captured and the tag was reported at time j . Similarly, for double-tagged animals that retain both of their tags, we find the $E(\mathbf{R}_{ij}^3)$ is

$$E(\mathbf{R}_{ij}^3) = D(\mathbf{N}_i)\Lambda_i\Lambda_i\phi_i\mathbf{Q}_i(\mathbf{I} - \mathbf{P}_{i+1})\dots\Lambda_{j-1}\Lambda_{j-1}\phi_{j-1}\mathbf{Q}_{j-1}\mathbf{P}_j\lambda_j^3 \quad (4.3)$$

Complexity lies in modelling animals that were double tagged and recovered with single tags as it is unknown when animals lose tags. For instance, for animals that were tagged in period 1 and recovered in period 3, the $E(\mathbf{R}_{ij}^2)$ is

$$E(\mathbf{R}_{13}^2) = D(\mathbf{N}_1) \{ \Lambda_1\Lambda_1\phi_1\mathbf{Q}_1(\mathbf{I} - \mathbf{P}_2)\Lambda_2(\mathbf{I} - \Lambda_2) + \Lambda_1(\mathbf{I} - \Lambda_1)\phi_1\mathbf{Q}_1(\mathbf{I} - \mathbf{P}_2)\Lambda_2 \} \times \phi_2\mathbf{Q}_2\mathbf{P}_3\lambda_3^2 \quad (4.4)$$

Thus in general we need to account for all the possible periods where a single tag loss could have occurred.

4.4 Parameter and Standard Error Estimates

Closed form maximum likelihood estimators are not available so numerical methods are employed to obtain parameter estimates. We used a Newton-Raphson algorithm to obtain maximum likelihood estimates.

As in chapter 3, we used design matrices to allow for different model specifications. Design matrices provide the link between standard parameters (p, ϕ, Λ , etc) and β parameters. For instance, if we use a logit link function, we find $\text{logit}(\mathbf{p}) = \mathbf{X}\boldsymbol{\beta}$. The design matrices offer the possibility of making some parameters constant and reduce the number of parameters to be maximized. The logit link function transforms design

matrices (\mathbf{X}) and standard parameters into the β parameters which are used in the maximization procedure.

Maximum likelihood estimates of the β parameters are transformed back to standard parameters using the anti-logit. Standard errors of the β parameters were obtained using a numerical central differencing method as in chapter 3. Standard error estimates of the back-transformed estimates are obtained using the delta method.

Parameters \mathbf{p}_k and $\boldsymbol{\lambda}_k$ are confounded for the last time period k unless different tag types are used to differentiate the $\boldsymbol{\lambda}_k^g$ values. Further when releases do not occur in each time period, survival and migration parameters are confounded.

R computer code (R Development Core Team, 2004) to implement parameter estimation as described above is available. The program allows for flexibility in fitting several models through the use of design matrices.

4.4.1 Goodness of Fit

The goodness-of-fit test is similar to that of chapter 3. The log-likelihood is evaluated for the fitted and saturated models. Twice the difference in these likelihoods is approximately χ^2 distributed with degrees of freedom being the difference in the number of parameters ($n - K$). The saturated model is one where $\theta_{ij}^{gst} = R_{ij}^{gst}$ (the expected number of recoveries is the observed number recovered).

Pearson residuals were computed as $\frac{(R-E(R))}{\sqrt{E(R)}}$.

When fitting a model whose mean is unaffected but the variance is inflated by an unknown factor, the use of the deviance statistic as a goodness of fit test has a limiting distribution of $\psi\chi_{n-K}^2$ rather than χ_{n-K}^2 as per usual (McCullagh and Nelder, 1989, p. 126).

4.4.2 Over-Dispersion

Over-dispersion exists when the variance in the observed counts exceeds the variance given by the Poisson assumption where the variance is proportional to the mean. McCullagh and Nelder (1989, p.90) suggest that with count data, one should assume

over-dispersion is present unless prior information indicates otherwise. One mechanism of over-dispersion is clustering in the population (i.e. schooling of fish) which leads to the usual over-dispersed Poisson model .

$$V(R_{ij}^{gst}) = \psi\theta_{ij}^{gst} \quad (4.5)$$

Myers and Hoenig (1997) consider over-dispersion when implementing a binomial model. They modelled over-dispersion using a scale factor for the variances as described by McCullagh and Nelder (1989, p. 126). This over-dispersion factor can be estimated as the residual sum of squares scaled by the degrees of freedom ($n - K =$ number of counts- number of parameters), $\sum \frac{(R-\theta)^2}{(n-K)\theta}$. The alternative estimator for over-dispersion is the deviance divided by the degrees of freedom

$$\frac{D}{(n - K)} = \sum_{g=1}^m \sum_{i=1}^{k-1} \sum_{j=i+1}^k \sum_{s=1}^a \sum_{t=1}^a \frac{2}{(n - K)} \left\{ R_{ij}^{gst} \ln \left(\frac{R_{ij}^{gst}}{\theta_{ij}^{gst}} \right) - (R_{ij}^{gst} - \theta_{ij}^{gst}) \right\} \quad (4.6)$$

4.4.3 Model Selection

Many models can be implemented using the design matrices. Notation is similar to that of Lebreton et al. (1992). For instance, the simplest model has constant parameters over strata, time and tag type $[p, \phi, \Lambda, Q, \lambda]$. Other models include but are not limited to: parameters vary across strata but are constant over time $[p_s, \phi_s, \Lambda_s, Q_s, \lambda_s]$, parameters vary across strata and time $[p_{st}, \phi_{st}, \Lambda_{st}, Q_{st}, \lambda_{st}]$, and parameters vary across strata and time and reporting rate varies over tag type $[p_{st}, \phi_{st}, \Lambda_{st}, Q_{st}, \lambda_{stg}]$, etc.

Akaike's Information Criterion (AIC) can be used to select between competing models, and model averaging of several candidate models can be employed (Burnham and Anderson, 2002). When over-dispersion enters the model, QAIC which accounts for over-dispersion should be used (Burnham and Anderson, 2002).

4.5 Example

In 2000, the Atlantic Fisheries Centre and Fishery Products International started a 5 year tagging program to study the Grand Banks yellowtail flounder. The stock area was partitioned into 2 strata (Figure 4.1). The majority of the Grand Banks' stock is found within these strata (Walsh et al. 2001). Flounder were captured with a commercial otter trawl within one of the strata and released at one of 25 release positions. In 2001, the strata were extended northward adding 6 release positions.

Flounder were tagged with either a single or a double Petersen disc tag. At each release point, 20 fish were released with high reward tags, 90 fish were released with single tags and 90 fish were released with double tags. In 2004, only single-tag low-reward tagged fish were released. Figure 4.2a shows the frequency of flounder released by month.

Recoveries occurred when flounder were caught by otter trawl and their tags were returned by the fishers. The fishery was closed from June 15 to July 31 of each year as this is the spawning season for yellowtail flounder. Figure 4.2b shows the frequency of recovery by month. Each year was broken into two 6-month periods: January to June, July to December. Releases occurred in the first period of each year; recoveries occurred in both periods.

Table 4.2 shows the number of flounder released and recovered by tag type, period and strata. As indicated previously, the capture history can be summarized into the N_i^g and R_i^g matrices (Table 4.2). A total of 29663 fish were released between 2000 and 2004, 2440 of which have been recovered.

Instantaneous tagging mortality was estimated by putting flounder into cages for 10 days in 2001, 2002, 2003 and 2004. Table 4.3 shows the numbers of animals in the cages and the number that died. Both single and double tagging were evaluated. In 2001, flounders were left for 19 days as opposed to 10 days and it was found that when caged, flounders bite each other's tails, possibly causing increased probability of infection. This problem was rectified for later experiments and initial tagging mortality rate was estimated as 0% (S. Walsh, 2005 pers. comm.).

We examined several models in the model selection process. The goodness-of-fit

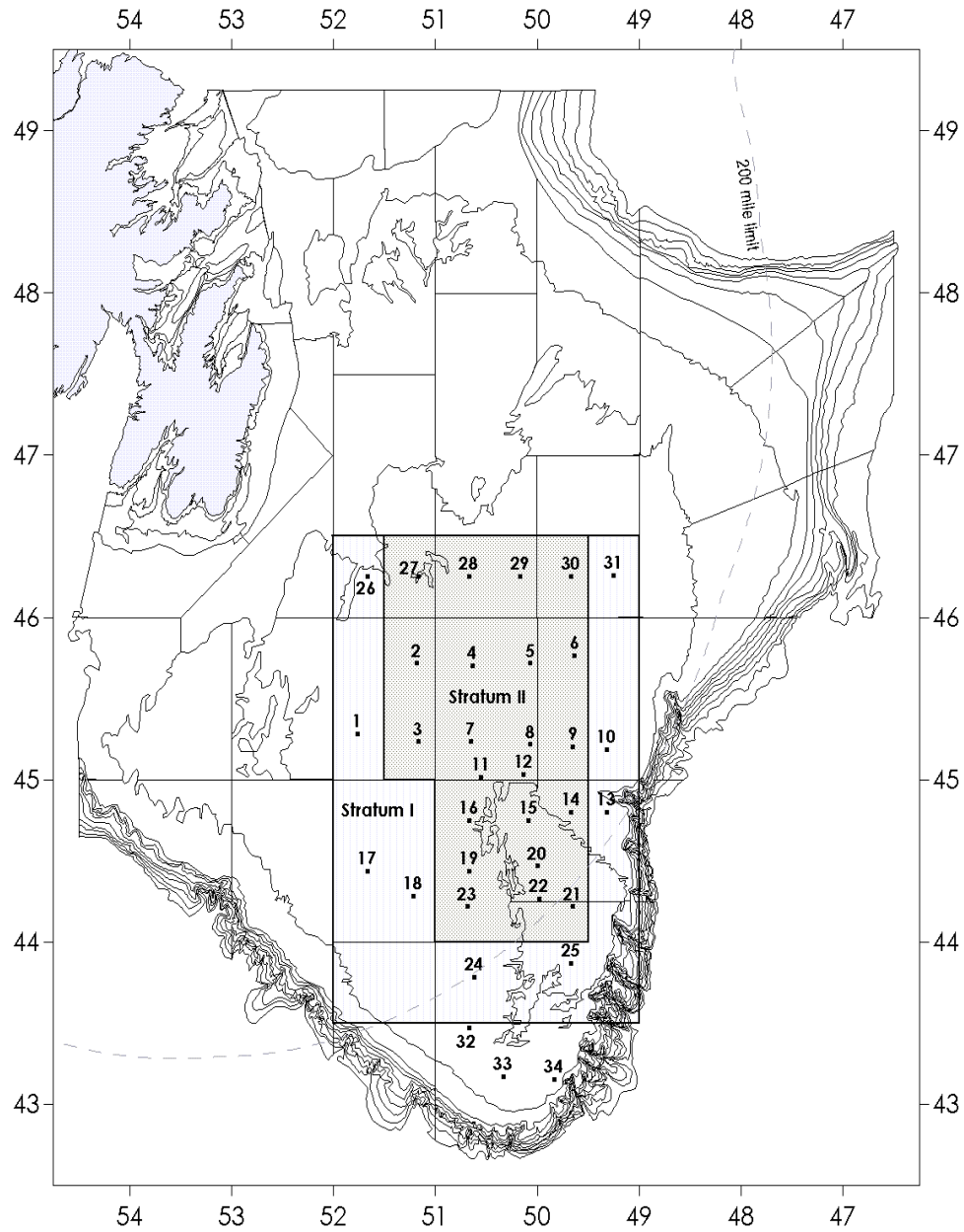


Figure 4.1: Map of the strata boundaries and release points for yellowtail flounder on the Grand Banks of Newfoundland. Release points 26-31 were added in 2001. Release points 32-34 were not included in this analysis. Map ©Fisheries and Oceans Canada by permission, provided by Stephen J. Walsh, Northwest Atlantic Fisheries Center, Newfoundland.

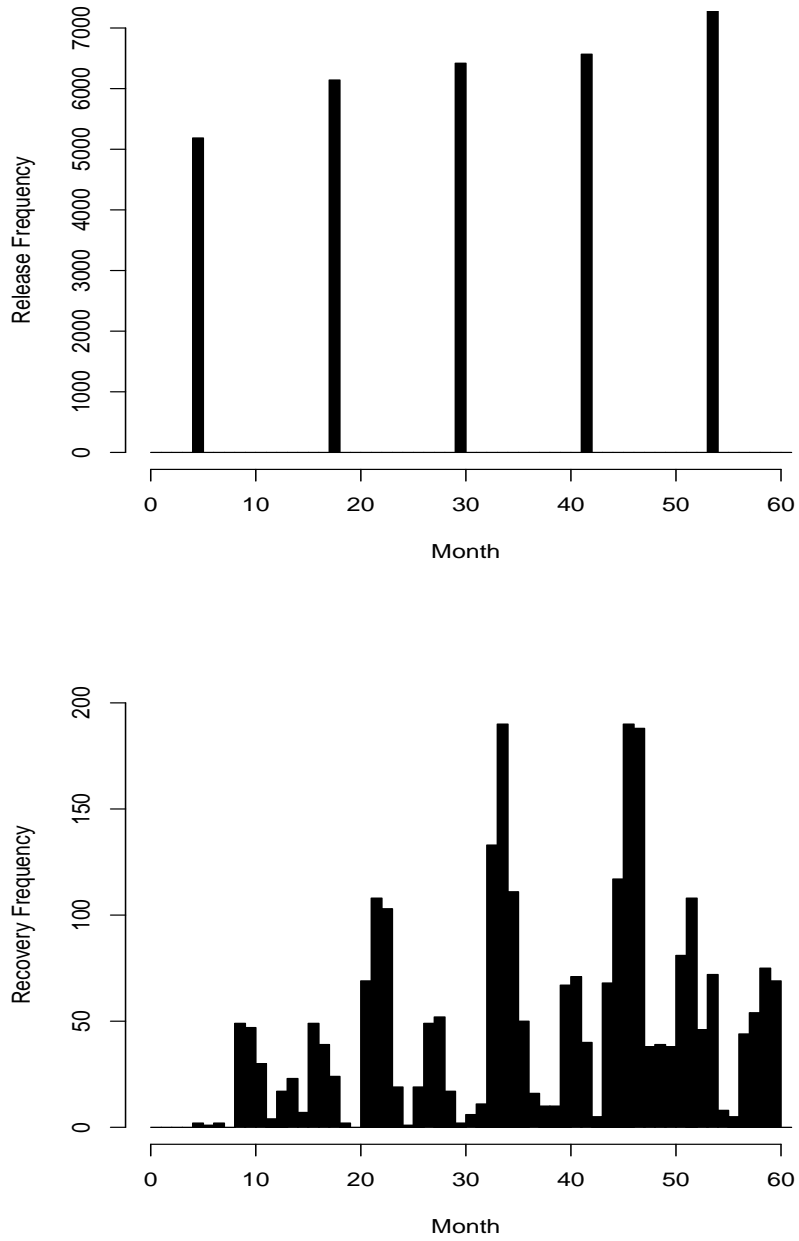


Figure 4.2: Frequency of release and recovery by month from January 2000 to December 2004.

Table 4.2: The number of yellowtail flounder released and recovered by tag type, period and strata. Tag types are as follows: 1) single-tagged fish, 2) double-tagged fish recovered with a single tag, 3) double-tagged fish, 4) single-tagged fish with a high reward tag. Years are divided into 2 periods, January-June and July-December. Fish are released in the first period of each year (1,3,5,7,and 9). Note that there were no double-tagged flounder released in period 9 (2004).

Tag Type	Release			Recovery Period and Stratum																	
	Period	Stratum	Number	2		3		4		5		6		7		8		9		10	
				1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
1	1	1	625	1	4	11	5	7	3	0	2	3	7	2	3	5	2	1	5	2	1
		2	1608	0	44	0	49	1	28	0	22	2	37	1	22	4	24	3	24	0	7
	3	1	671				4	3	0	7	2	4	3	3	5	3	1	3	2	2	
		2	1957			2	60	0	39	0	58	2	22	6	31	1	31	2	7		
	5	1	808								6	5	4	2	2	7	1	5	6	6	
		2	1983								2	86	1	37	4	47	0	42	0	10	
	7	1	807												8	10	2	12	3	3	
		2	1978												4	68	0	39	1	19	
	9	1	1798																	23	5
		2	4400																	2	72
2	1	1		0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
		2		0	1	0	1	1	2	0	1	0	0	0	0	0	2	0	0	0	0
	3	1					0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
		2				0	4	0	0	0	4	0	2	0	2	0	3	0	0		
	5	1									0	1	0	0	0	0	0	1	0	0	
		2									1	10	0	7	0	5	0	8	0	2	
	7	1													0	0	0	0	0	0	
		2													0	2	0	4	1	0	
	9	1																		0	0
		2																		0	0

(Table 4.2 Continued)

Tag Type	Release			Recovery																		
	Period	Stratum	Number	2		3		4		5		6		7		8		9		10		
				1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	
3	1	1	624	1	2	11	11	10	1	0	1	3	4	0	4	2	2	2	2	3	2	
		2	1612	1	50	1	50	1	25	0	16	0	26	1	11	2	13	0	3	0	5	
	3	1	659					7	6	0	4	2	4	1	2	5	3	0	4	1	1	
		2	1952					2	75	0	25	2	41	0	13	4	33	2	16	1	7	
	5	1	804									8	12	3	2	3	9	3	9	3	1	
		2	1977									3	83	2	38	2	37	0	33	2	12	
	7	1	809													14	10	0	10	3	4	
		2	1980													3	75	1	43	0	14	
	9	1	0																		0	0
		2	0																			0
4	1	1	138	0	0	3	0	3	0	0	1	1	1	0	0	1	0	0	2	0	0	
		2	355	0	10	0	8	0	4	0	7	0	8	0	2	3	8	0	5	0	1	
	3	1	149					3	2	0	0	0	1	0	0	1	2	0	2	0	0	
		2	435					0	13	0	8	0	11	0	3	2	11	0	5	0	0	
	5	1	178									0	3	0	1	0	1	0	1	0	0	
		2	438									0	20	0	8	1	9	1	13	1	4	
	7	1	179													3	4	0	1	0	0	
		2	439													0	12	0	8	0	1	
	9	1	200																		0	0
		2	100																			0

Table 4.3: Results of the instantaneous tagging mortality experiments where yellowtail flounder were kept in cages for 10 days (19 days in 2001).

Tag		Year			
		2001	2002	2003	2004
Single	Total	9	11	4	13
	Dead	7	0	0	0
Double	Total	9	11	4	13
	Dead	8	0	0	0
None	Total	9	10	4	16
	Dead	6	0	0	0

test had p-values equal to 0 for all models. This is likely due to over-dispersion. Table 4.4 shows the estimates of over-dispersion, the associated QAIC, and the Δ_i values.

From the QAIC and Δ_i values, model $[p_s, \phi_{st}, \Lambda, Q_s, \lambda]$ is chosen to represent the yellowtail flounder data. As the Δ_i values are quite large, we did not implement model averaging. Parameter estimates and standard errors inflated for over-dispersion are given in Table 4.5 for model $[p_s, \phi_{st}, \Lambda, Q_s, \lambda]$ which had the lowest Δ_i value.

The residual plot of Pearson residuals versus the expected recoveries is shown in Figure 4.3. We note that there are 6 outliers seen in the residual plot for model $[p_s, \phi_{st}, \Lambda, Q_s, \lambda]$ with values greater than $|\pm 4|$. These are shown in Table 4.5. For three, the residuals are large due to division by a small number. For the other 3, these are true outliers that exist in the data. There were 18 residuals greater than $|\pm 3|$

Initial estimates of tag-reporting rates obtained from the raw data were 104% for double-tagged fish and 77% single-tagged fish. These were obtained by summing up the number of animals released and recovered for each tag type over all strata and time periods. The fraction of animals recovered over released for each tag type was compared to the fraction recovered over released from the high reward tags (i.e. for tag type 1, $\frac{R^1}{N^1} / \frac{R^4}{N^4}$). The observed rate of return for double-tagged fish was higher than those of high reward-tagged fish which leads to the estimate of 100% tag reporting rate (see Table 4.5).

Table 4.4: Summary of models fit to the yellowtail flounder data. Over-dispersion estimates $\hat{\psi}$ and $\tilde{\psi}$ are estimated from squared Pearson residuals divided by $n - K$ and deviance divided by $n - K$ respectively (where n is the degrees of freedom of the saturated model, the number of observed counts). The number of parameters in the model is K ; QAIC (adjusted for over-dispersion using $\hat{\psi} = 2.010$) and Δ_i are also shown.

Model	$\hat{\psi}$	$\tilde{\psi}$	K	QAIC	Δ_i
$p_{st}, \phi_{st}, \Lambda_{st}, Q_{st}, \lambda_{st}$	2.645	2.355	50	-4497	184
$p_s, \phi_s, \Lambda, Q_{st}, \lambda_k$	2.237	2.071	19	-4630	51
$p_s, \phi_s, \Lambda, Q_s, \lambda_k$	2.192	2.114	11	-4633	48
$p_s, \phi_s, \Lambda_s, Q_s, \lambda_s$	2.367	2.108	10	-4633	48
$p_s, \phi, \Lambda, Q_s, \lambda$	2.220	2.128	7	-4636	45
$p_s, \phi, \Lambda, Q_s, \lambda_k$	2.186	2.109	10	-4636	45
$p_{st}, \phi_s, \Lambda, Q_s, \lambda$	2.095	1.886	16	-4672	9
$p_s, \phi_{st}, \Lambda, Q_s, \lambda_{k12}$	2.201	1.861	18	-4677	4
$p_s, \phi_{st}, \Lambda, Q_s, \lambda_s$	2.016	1.856	17	-4679	2
$p_s, \phi_{st}, \Lambda, Q_s, \lambda$	2.010	1.850	16	-4681	0

s varies over stratum.

t varies over year.

k varies over tag type.

$k12$ tag type 1 and tag type 2 are equal.

Table 4.5: Maximum likelihood estimates and standard errors inflated by the square root of the over-dispersion estimate $\hat{\psi} = 2.010$ for model $[p_s, \phi_{st}, \Lambda, Q_s, \lambda]$. The unit for the parameters is the 6-month periods.

Parameter	MLE	SE
p^1	0.013	0.002
p^2	0.039	0.002
ϕ_1^1	0.908	0.166
ϕ_1^2	0.970	0.091
ϕ_2^1	0.984	0.124
ϕ_2^2	0.818	0.054
ϕ_3^1	0.706	0.080
ϕ_3^2	0.951	0.051
ϕ_4^1	1.000	0.000
ϕ_4^2	0.950	0.048
ϕ_5^1	0.755	0.099
ϕ_5^2	0.606	0.038
Λ	0.976	0.004
q^{12}	0.145	0.019
q^{21}	0.052	0.011
λ	1.000	0.000

Table 4.6: Outlying residuals found in plots of the Pearson residual versus expected recovery for model $[p_s, \phi_{st}, \Lambda, Q_s, \lambda]$.

Tag Type	Period		Stratum		Recovery		Pearson Residual
	Release	Recover	Release	Recover	Observed	Expected	
2	5	6	2	2	10	1.63	6.55
2	5	6	1	2	1	0.31	5.54
3	5	6	2	1	12	3.05	5.12
4	1	8	1	2	3	0.33	4.65
1	1	9	2	2	24	9.80	4.54
3	1	10	1	1	3	0.41	4.07

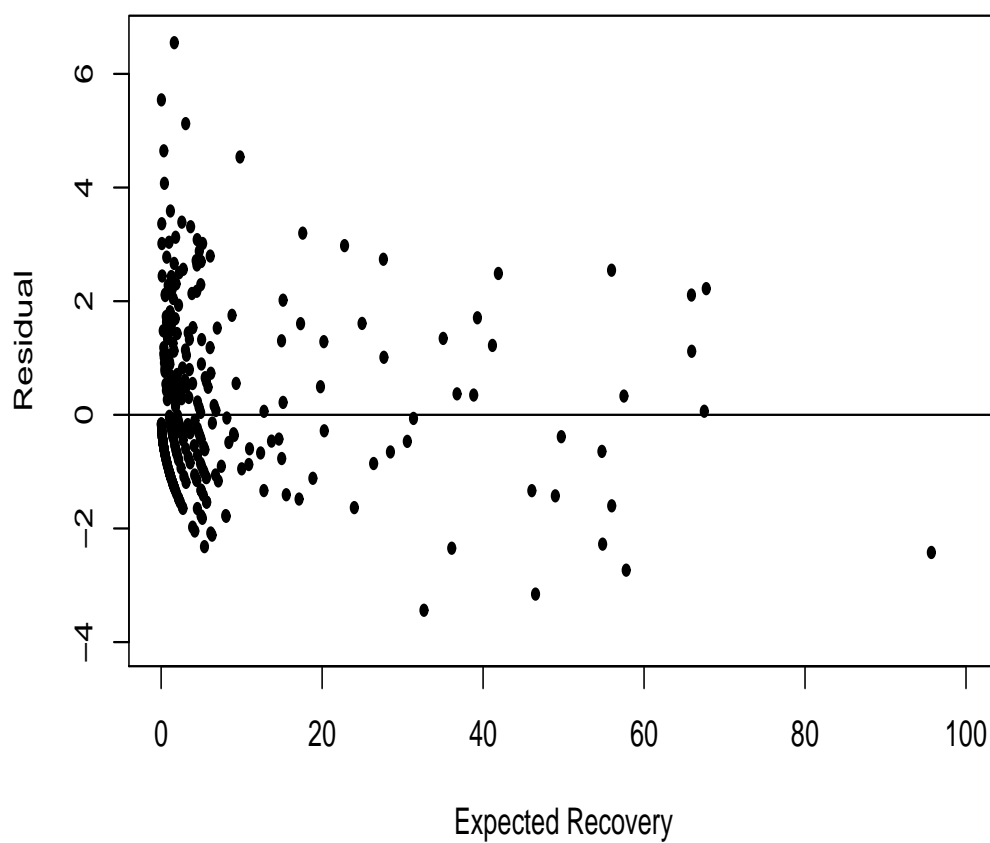


Figure 4.3: Pearson residuals versus expected recoveries for model $[p_s, \phi_{st}, \Lambda, Q_s, \lambda]$.

Tag-retention estimates were high for all models and there was no indication that tag loss varied across year or stratum. The number of double-tagged flounder reported as having single tags was only 73, whereas there were 932 reported as double tagged of the 10417 double-tagged flounder released. The gross tag-retention rate from the raw data was 96% (using Seber and Felton, 1981).

Migration rates varied over stratum. Estimated migration from stratum 1 to stratum 2 was a factor of 3 times larger than migration from stratum 2 to stratum 1. Estimated exploitation rates varied over stratum; exploitation in stratum 2 was higher than exploitation in stratum 1 by a factor of 3. Estimated survival estimates varied over stratum and time. There is no trend as to which stratum had higher survival rates over the years.

4.6 Discussion

Capture-recapture models are used in fishery studies to estimate movement, mortality and exploitation. When sample sizes are large and recovery rates small, it is convenient to implement a Poisson framework. Migration can be studied by stratifying the study area. As tag loss is still a concern within these studies, it needs to be considered in order to reduce bias in exploitation estimates.

The yellowtail flounder data provided an example for migration under tag loss. One potential problem with the analysis was that one of the assumptions made was that capture is instantaneous. However, fish were captured and tagged once a year, yet aside from a short closing of the fishery for spawning, flounder were recovered all year long. Looking at the graphs of releases and recoveries (Figure 4.2), this assumption is met for the releases. However, it is not met for the recoveries. Fish that were caught early in the sampling period would not have the same amount of time to survive and migrate as those caught late. The same may be true for tag loss. This leads to heterogeneity in survival, migration and tag retention which may be a source of over-dispersion.

Another problem that could affect the estimates of migration is heterogeneity in the distance from the boundary of the stratum. It would be more feasible for fish that

were released near the stratum boundary to migrate than it would for fish released farther away. Accordingly, future work should consist of developing a model that takes heterogeneity of movement into consideration. For instance, migration could depend on closest distance to a stratum boundary. It would be useful to run a simulation study to determine the effects of migration heterogeneity on other parameter estimates.

With the current modification of the R program code, we do not allow for the implementation of tag loss by age. It is conceivable that tag loss may increase over time from tagging. This modification has been left for future work which can easily be done by adding an additional subscript to the tag-retention matrices indicating the time of tagging. We would then have matrix Λ_{ij} with elements Λ_{ij}^{ss} that would be the probability of tag retention between time j and $j + 1$ in stratum s when initially tagged at time i .

In addition, there may be a problem with non-mixing of tagged fish with untagged fish before recovery. If non-mixing occurs, newly tagged animals may have different survival rates than previously tagged animals. Hoenig et al. (1998) and Brownie et al. (1985) dealt with this by allowing animals that were freshly tagged and not yet mixed into the population to have a different parameter to represent fishing mortality rate (F_j^*) than animals already tagged and mixed in the population (F_j) for the same year j of recovery.

Finally, the flounder data was separated into 6-month time periods. We initially thought that parameters for all time periods were estimable. This would be true if releases occurred in each period along with recoveries. However, for the flounder data, releases occurred once a year or every other period. When there are no releases, survival and migration are confounded in a complicated fashion. By forcing parameters to be equal in the two periods within a year, the confounding problem was resolved. Unfortunately, we can no longer detect any seasonal migration that may be occurring as suggested by Walsh et al. (2001).

Chapter 5

Conclusions and Future work

5.1 Summary

This thesis consisted of 3 papers that develop methodology to deal with tag loss in various capture-recapture studies. As tag loss can have a profound impact on estimates of survival and population size, it is best to either develop a study that incorporates and estimates tag-loss rates, or design the study to minimize tag loss.

In chapter 2, tag loss was in the form of untimely battery failure in radio-tags. Radio-telemetry experiments are designed to have the batteries fail after the experiment is complete. If radio-tags may fail prior to the end of the study survival estimates can be severely biased. If a second study is performed that estimates the radio failure-time distribution, this auxiliary information can be used to modify survival estimates.

We developed methodology for extending the Cormack-Jolly-Seber model to incorporate auxiliary radio-tag failure-time information. The likelihood was separated into two components— one for the information from the mark-recapture study and one for the radio-tag study. Travel time of the animal had to be modelled as this was convoluted with the radio-failure time. We used the inverse-gaussian distribution to model travel time as it has properties that simplified the likelihood.

The likelihood of the radio-tags was maximized first and survival estimates were obtained using the Kaplan-Meier estimator. The remaining parameters were then

estimated by maximizing the component of the likelihood corresponding to the mark-recapture data, holding the radio-tag survival parameters fixed. Appendix A motivates the use of this methodology. Both classical and Bayesian frameworks were used to obtain parameter estimates. Numerical maximum likelihood estimates were obtained using a Newton-Raphson algorithm and Bayesian posterior mean estimates were obtained using Metropolis-Hastings steps within a Markov chain Monte Carlo algorithm. Standard error estimates were obtained using bootstrap methodology for the classical estimator. Bayesian standard error estimates were obtained in the usual fashion from the standard deviation of the Markov chain.

We applied this methodology to data from the Columbia River, Washington where Chinook smolts migrate downriver passing several antenna where they are ‘recaptured’. Normally batteries fail well after a study; however, in 2000 a drought caused fish to travel downriver slowly, lengthening the time of the study. This increased the chances of batteries failing prior to the fish traveling past the last antenna. Our model provides higher survival estimates than the standard Cormack-Jolly-Seber model.

Simulation studies were used to look at several aspects of this model. As the Columbia River data had near 100% recapture rates, we investigated the performance of the model when recaptures occurred 80% of the time. Again, our model provides higher survival estimates than the standard Cormack-Jolly-Seber model.

A second simulation study explored the effect of sample size. When designing a study, it is useful to know how many resources should be put into estimating the distribution of the radio-tag survival time. Increasing the number of fish released improves parameter estimates and reduces standard errors. On the other hand, increasing the number of radio-tags had little effect.

Additionally, we examined our model subject to increasing the number of radio-tags failing during the study. Under high levels of failure, our model outperforms the Cormack-Jolly-Seber model when estimating survival. Although there is a loss in precision in the survival estimates as the number of radio-tags failing increases, the radio-tag-failure model produces reasonable point estimates. The recapture estimates are fairly robust for both models when battery failure is high.

Finally, we considered the assumption that survival estimates are independent of

travel time. When this assumption is violated, survival estimates tend to be positively biased. Estimates of mean travel times are negatively biased as only the fast fish are observed at the later sampling occasions.

The methodology developed in this chapter is shown to produce better point estimates than the standard Cormack-Jolly-Seber model. It is best applied when radio-tags fail before the end of a study and auxiliary radio-failure time data is available.

In chapter 3 we extended the Jolly-Seber model to estimate tag-retention rates within the model. As outlined in the introduction, traditionally, tag-loss rates have been estimated using *ad hoc* methods. The experimental design incorporates both single and double-tagged fish requiring traditional capture histories to become tag histories. We broadened the Jolly-Seber model outlined by Arnason and Schwarz (1996) using tag-history notation rather than sufficient statistics and incorporated tag-retention probabilities within the model. We maximized the likelihood using the EM algorithm. As the EM algorithm makes use of the complete-data likelihood, we expanded the work of VanDeusen (2002) to produce expected values of latent variables for the Jolly-Seber model (see Appendix B). As in chapter 2, model selection employed the Akaike information criterion.

As an example of the use of the Jolly-Seber-Tag-Loss model, we estimated the number of walleye in Mille Lacs, Minnesota. We compared the estimates from our model to those from the simple Jolly-Seber model. The Jolly-Seber model produced estimates that were quite different from our model. High tag loss was experienced which increased the Jolly-Seber abundance estimate compared to the Jolly-Seber-Tag-Loss estimate.

There was enough complexity in the data due to heterogeneity in catchability that the Jolly-Seber-Tag-Loss model is not sufficient to model the Mille Lac data. This idea is supported by poor model fit. Still, the Jolly-Seber-Tag-Loss model would likely perform better when these types of model assumptions are not violated.

This method can be applied when some double-tagged animals are in the study. Otherwise tag loss cannot be estimated. There are difficulties that exist in the sense of animals that are “recycled”. These are animals that lose all of their tags and are recaptured, mistaken as new individuals. These animals could bias estimates

profoundly if their numbers were large due to high tag-loss rates coupled with high recapture rates. If this is the case, the JSTL will not perform well and we suggest redesigning the study to reduce tag loss.

In chapter 4 we consider a Poisson model that incorporates migration and tag loss. We extend Hilborn's (1990) model to incorporate tag loss and variable exploitation rates over time. As migration is modelled through stratification of the study area, probabilities are in the form of matrices. As in chapter 3 both double- and single-tagged animals are required to estimate tag-loss rates. Tag rewards are offered so that tag-reporting rates are estimable with the assumption that high-reward tags will be returned 100% of the time.

We obtained numerical maximum likelihood estimates in the usual fashion and incorporated over-dispersion estimates into the model selection procedure (via QAIC) and into standard error estimates as per standard methodology.

We applied this model to estimate exploitation rates of yellowtail flounder on the Grand Banks of Newfoundland. Over-dispersion was found to be present in the model contributing to the lack of fit. We had difficulty implementing several different models. It may be due to data sparsity. Nonetheless, this model is beneficial to fishery scientists interested in modelling migration under tag loss.

5.2 Future Work

In chapter 2 we discuss the impact of the assumption that survival is independent of travel time. It is clear that in some studies this assumption will not be viable. For example, fish that travel downriver slowly, may have lower survival rates due to increased exposure to predation for example. We would like to be able to model fish survival as a function of travel time.

Additionally, the use of the Kaplan-Meier estimator causes difficulty in the maximization procedure due to the mixing of parametric and non-parametric likelihoods. It is not clear how the full likelihood could be maximized unconditionally. If it were suitable, a parametric model such as a Weibull could be implemented for the radio-tag survival distribution. Along with simplicities associated with parametric modelling

come increased model assumptions. One must check the validity of the assumptions made under a parametric framework. It may be the case that a fully parametric likelihood can be maximized but lack-of-fit of the parametric model exists.

There are some extensions to be made to the Jolly-Seber-Tag-Loss model. First we would like to incorporate multiple groups (such as males and females). We would use methods similar to those outlined by Arnason and Schwarz (1996) where they incorporated groups into the Jolly-Seber model by adding an additional index to the parameters that could be group dependent. This would increase the complexity of the model as the number of parameters increases.

Additionally, we would like to build on the JSTL model such that we could model different types of tags. The yellowtail flounder study from chapter 4 used two different types of tags (Peterson-disk tags and data-logger tags). It is conceivable that tags of different types would be lost at different rates, especially if one tag was a permanent marking. We would like to be able to model separate tag-loss rates for each type of tag used. Note that we did not use the data-logger tag data in chapter 4 due to sparsity concerns.

When an animal's behaviour affects tag loss, independence of tags on double-tagged animals is unlikely. Bradshaw et al. (2000) found that New Zealand fur seals were more likely to lose a second tag if their first tag was lost. In these cases, the assumption of independent tag loss would be violated. Thus we would extend the JSTL model to incorporate dependent tag loss.

The migration model from chapter 4 could incorporate similar advances as the JSTL model. We have already incorporated tag-types in the form of single and double tagging, but it would be a simple extension to apply this to several types of tags (e.g. Peterson-disk and data-logger tags) by defining tag types appropriately. In chapter 4, we described 4 tag types which were some combination of single and double Peterson-disk tags; this could be extended to include the data-logger tags. Similarly, we would like to extend the migration model to include multiple groups as described for the JSTL model.

Additionally, fish that are released close to a stratum boundary would be more

likely to migrate into neighbouring strata. Migration between stratum may be dependent on the distance from release to the stratum boundary. Thus the migration parameter could be modeled as a function of distance, dealing with the heterogeneity in distance to stratum boundaries. Heterogeneity of many kinds is a common problem in mark-recapture studies and future work needs to develop more methodology to deal with heterogeneity problems.

As it stands now, the migration model will not model tag loss as a function of time from release. This again would be fairly straightforward to implement and would likely be a valuable tool for users of the model. We could do the same for the survival parameter. This would provide the capability of dealing with the problem of non-mixing of tagged fish with untagged fish as described by Hoenig et al. (1998). Through design matrices, animals that are initially tagged could have different survival rates from those who were already tagged. It may be conceivable that after one time period animals could be considered mixed. This assumption would reduce the complexity of the non-mixing model.

Finally, this thesis deals with tag loss in three types of open population models. Tag loss is a problem that transcends the boundaries between open and closed population models. Some work exists for closed populations (Seber, 1982; p.94-96). Future work should involve investigating the impact of tag loss on closed-population models and develop methodology to incorporate tag-loss estimation into the modelling process.

The End

Appendix A

Conditional Likelihood Approach for the Radio-Tag Failure Model

The method used for maximizing the likelihood in chapter 2 is somewhat non-standard as we did not maximize the full likelihood. The likelihood was broken up into two parts: the likelihood for the fish and the likelihood for the radios. This corresponded to the two experiments: the mark-recapture experiment and the radio-failure time experiment.

When passage times past the recording points (i.e. travel time) are not recorded, it is well known that radio failure is indistinguishable from death or emigration. This is not strictly true if you know travel time, but fish survival (ϕ), travel time (t) and radio survival (S) always appear in a complicated function within χ . Some information on radio survival is contained in the first part of the likelihood in the form of censored times of the observed individuals, i.e. you know that the radio failed after the last time a fish was seen, but there is no information on exact radio-failure times. We argue that the amount of information regarding radio-failure in L^{fish} is small compared to L^{radio} .

If $p = 1$ such that recapture rates were 100%, radio failure, travel time and survival are all bound together making it difficult to separate these terms out. Thus, we implemented the conditional likelihood approach in the sense of Sanathanan (1977) where the likelihood is first broken into two parts, one of which is based on a smaller

set of parameters. We were convinced that L^{radio} contained most of the information about radio-failure.

A.1 2 Dam Case

Consider the 2 dam case:

The only possible capture histories are $\{11\}$ and $\{10\}$. There are n_{11} of the former with travel time t_i and n_{10} of the later with no observed travel time. Due to there being only 2 dams, only the parameter $\phi_1 p_2$ can be identified.

Let $g(t) = g_{12}(t)$ be the distribution of travel time between dam 1 and 2 which depends on parameter Φ . Further, let $S(z)$ be the survival function for radio tags which depends on parameter β .

The likelihood function would then be

$$L = \prod_{\text{all fish with history } \{11\}} \phi_1 p_2 g(t_i) S(t_i) \left[1 - \phi_1 p_2 \int_0^\infty g(u) S(u) du \right]^{n_{10}} \times \prod_{\text{all radios}} [-S'(z_i)] \quad (\text{A.1})$$

and the log-likelihood function would be

$$l = \sum_{\text{all fish with history } \{11\}} \ln(\phi_1 p_2) + \ln(g(t_i) S(t_i)) n_{10} \ln \left(1 - \phi_1 p_2 \int_0^\infty g(u) S(u) du \right) + \sum_{\text{all radios}} \ln(-S'(z_i)) \quad (\text{A.2})$$

The first partial derivatives— $\frac{\delta l}{\delta \phi_1 p_2}$, $\frac{\delta l}{\delta \beta}$, $\frac{\delta l}{\delta \Phi}$ are

$$\frac{\delta l}{\delta \phi_1 p_2} = \sum_{\text{all fish with history } \{11\}} \frac{1}{\phi_1 p_2} - \frac{n_{10} \int_0^\infty g(u) S(u) du}{1 - \phi_1 p_2 \int_0^\infty g(u) S(u) du} \quad (\text{A.3})$$

$$\frac{\delta l}{\delta \beta} = \sum_{\text{all fish with history } \{11\}} \frac{g(t_i)}{g(t_i) S(t_i)} \frac{\delta S}{\delta \beta} - \frac{n_{10} \phi_1 p_2 \frac{\delta}{\delta \beta} \int_0^\infty g(u) S(u) du}{1 - \phi_1 p_2 \int_0^\infty g(u) S(u) du} + \sum_{\text{all radios}} \frac{\frac{\delta}{\delta \beta} S'(z_i)}{S'(z_i)} \quad (\text{A.4})$$

$$\frac{\delta l}{\delta \Phi} = \sum_{\text{all fish with history } \{11\}} \frac{S(t_i)}{g(t_i)S(t_i)} \frac{\delta g}{\delta \Phi} - \frac{n_{10}\phi_1 p_2 \frac{\delta}{\delta \Phi} \int_0^\infty g(u)S(u)du}{1 - \phi_1 p_2 \int_0^\infty g(u)S(u)du} \quad (\text{A.5})$$

If we solve (A.3) for $\phi_1 \hat{p}_2$ we find:

$$\phi_1 \hat{p}_2 = \frac{n_{11}}{(n_{11} + n_{10}) \int_0^\infty g(u)S(u)du} \quad (\text{A.6})$$

and if we substitute this into (A.5) we find:

$$\begin{aligned} 0 &= \sum_{\text{all fish with history } \{11\}} \frac{S(t_i)}{g(t_i)S(t_i)} \frac{\delta g}{\delta \Phi} - \frac{n_{10}\phi_1 \hat{p}_2 \frac{\delta}{\delta g} \frac{\delta g}{\delta \Phi} \int_0^\infty g(u)S(u)du}{1 - \phi_1 \hat{p}_2 \int_0^\infty g(u)S(u)du} \\ &= \sum_{\text{all fish with history } \{11\}} \frac{S(t_i)}{g(t_i)S(t_i)} \frac{\delta g}{\delta \Phi} - \frac{n_{10} \frac{n_{11}}{(n_{11} + n_{10}) \int_0^\infty g(u)S(u)du} \frac{\delta}{\delta g} \frac{\delta g}{\delta \Phi} \int_0^\infty g(u)S(u)du}{1 - \frac{n_{11}}{(n_{11} + n_{10}) \int_0^\infty g(u)S(u)du} \int_0^\infty g(u)S(u)du} \\ &= \sum_{\text{all fish with history } \{11\}} \frac{\delta}{\delta \Phi} \ln(g(t_i)S(t_i)) - \frac{n \frac{\delta}{\delta g} \frac{\delta g}{\delta \Phi} \int_0^\infty g(u)S(u)du}{\int_0^\infty g(u)S(u)du} \\ &= \sum_{\text{all fish with history } \{11\}} \frac{\delta}{\delta \Phi} \ln(g(t_i)S(t_i)) - n \frac{\delta}{\delta \Phi} \ln\left(\int_0^\infty g(u)S(u)du\right) \\ &= \sum_{\text{all fish with history } \{11\}} \frac{\delta}{\delta \Phi} \ln\left(\frac{g(t_i)S(t_i)}{\int_0^\infty g(u)S(u)du}\right) \end{aligned} \quad (\text{A.7})$$

which can be thought of as the conditional density of fish that move to dam 2 with radios working at time t_2 .

Similarly if we substitute (A.6) into (A.4) we find

$$0 = \sum_{\text{all fish with history } \{11\}} \frac{\delta}{\delta \beta} \ln\left(\frac{g(t_i)S(t_i)}{\int_0^\infty g(u)S(u)du}\right) + \sum_{\text{all radios}} \frac{\delta}{\delta \beta} \ln(-S'(z_i)) \quad (\text{A.8})$$

whose first term can be thought of as the conditional density of fish moving to dam 2 with radios working and whose second term as the information from the radios followed.

Hence the maximum likelihood estimators for Φ and β are found by solving (A.7) and (A.8) which both take the form of a ‘weighted’ contribution of a censored observation for those animals with history $\{11\}$ at time t_2 .

Notice that these are only dependent upon capture histories $\{11\}$ (animals that are recaptured); animals with capture histories $\{10\}$ do not contribute any information. These animals do not tell you whether the radio died, the animal died or you simply did not recapture the animal again.

A.2 k Dam Case

When we look at the general k dam case we write the likelihood in terms of the $m_{ij}^{t_i}$ array where $m_{ij}^{t_i}$ is the number of animals released at dam i at time t_i and next recovered at dam j at time t_j ; $m_{i0}^{t_i}$ is the number of animals released at dam i at time t_i and never recovered.

Each capture history can be broken into a series of $m_{ij}^{t_i}$ events; for example history $\{1010010\}$ gives rise to $m_{13}^{t_1}$, $m_{36}^{t_3}$, $m_{60}^{t_6}$ entries.

$$P(m_{ij}^{t_i} = 1 | t_i t_j) = \phi_i (1 - p_{i+1}) \phi_{i+1} (1 - p_{i+2}) \dots \phi_{j-1} p_j \frac{S(t_j)}{S(t_i)} g_{ij}(t_j - t_i) \quad (\text{A.9})$$

and

$$P(m_{i0}^{t_i} = 1 | t_i) = \chi_i(t_i) \quad (\text{A.10})$$

The probability of being seen after dam j , some time after t_i is

$$1 - \chi_i(t_i) = \sum_{j=i+1}^k \phi_i (1 - p_{i+1}) \phi_{i+1} (1 - p_{i+2}) \dots \phi_{j-1} p_j \frac{1}{S(t_i)} \int_0^\infty g_{ij}(u - t_i) S(u) du \quad (\text{A.11})$$

Again we let β be the underlying parameter of $S(x)$ (the survival distribution) and Φ be the underlying parameter of $g(x)$ (the movement distribution). The full likelihood becomes:

$$\begin{aligned}
L &= \prod_i \prod_{t_i \in i} \prod_{j=i+1}^k \left[\phi_i(1-p_{i+1}) \dots \phi_{j-1} p_j \frac{S(t_j)}{S(t_i)} g_{ij}(t_j - t_i) \right]^{m_{ij}^{t_i}} [\chi_i(t_i)]^{m_{i0}^{t_i}} \times L^{radio} \\
&= \prod_i \prod_{t_i \in i} \left\{ \prod_{j=i+1}^k \left[\frac{\phi_i(1-p_{i+1}) \dots \phi_{j-1} p_j \frac{S(t_j)}{S(t_i)} g_{ij}(t_j - t_i)}{\int_{t_i}^{\infty} \phi_i(1-p_i) \dots \phi_j p_j \frac{S(u)}{S(t_i)} g_{ij}(u - t_i) du} \right]^{m_{ij}^{t_i}} \right. \\
&\quad \left. \left[\int_{t_i}^{\infty} \phi_i(1-p_i) \dots \phi_j p_j \frac{S(u)}{S(t_i)} g_{ij}(u - t_i) du \right]^{m_{ij}^{t_i}} \right\} [\chi_i(t_i)]^{m_{i0}^{t_i}} \times L^{radio} \\
&= \prod_i \prod_{t_i \in i} \left\{ \prod_{j=i+1}^k \left[\frac{g_{ij}(t_j - t_i) S(t_j) / S(t_i)}{\int_{t_i}^{\infty} g_{ij}(u - t_i) S(u) / S(t_i) du} \right]^{m_{ij}^{t_i}} \right\} \times \\
&\quad \prod_i \left\{ \prod_{t_i \in i} \prod_{j=i+1}^k \left[\int_{t_i}^{\infty} \phi_i(1-p_i) \dots \phi_j p_j \frac{S(u)}{S(t_i)} g_{ij}(u - t_i) du \right]^{m_{ij}^{t_i}} \chi_j(t_i)^{m_{i0}^{t_i}} \right\} \times L^{radio}
\end{aligned} \tag{A.12}$$

Note that $t_i \in i$ is all times of fish released at dam i .

Here the first term represents the conditional distribution of seeing a fish move from dam i to dam j from t_i to t_j . This term does not depend on ϕ or p , it only depends on β and Φ . The second term is a multinomial distribution for release at dam i at time t_i . The log-likelihood then becomes

$$\begin{aligned}
l &= \sum_i \sum_{t_i \in i} \left\{ \sum_{j=i+1}^k m_{ij}^{t_i} \ln \left[\frac{g_{ij}(t_j - t_i) S(t_j) / S(t_i)}{\int_{t_i}^{\infty} g_{ij}(u - t_i) S(u) / S(t_i) du} \right] \right\} + \\
&\quad \sum_i \sum_{t_i \in i} \left\{ \sum_{j=i+1}^k m_{ij}^{t_i} \ln \left[\int_{t_i}^{\infty} g_{ij}(u - t_i) S(u) / S(t_i) du \right] \right\} + \\
&\quad \sum_i \sum_{t_i \in i} m_{i0}^{t_i} \ln[\chi_i(t_i)] + \ln[L^{radio}]
\end{aligned} \tag{A.13}$$

Again if we look at the score equation for β we find

$$\begin{aligned}
\frac{\delta l}{\delta \beta} &= \sum_i \sum_{t_i \in i} \left\{ \sum_{j=i+1}^k m_{ij}^{t_i} \frac{\delta}{\delta \beta} \ln \left[\frac{g_{ij}(t_j - t_i) \frac{S(t_j)}{S(t_i)}}{\int_{t_i}^{\infty} g_{ij}(u - t_i) \frac{S(u)}{S(t_i)} du} \right] \right\} + \\
&\sum_i \sum_{t_i \in i} \left\{ \sum_{j=i+1}^k m_{ij}^{t_i} \frac{\delta}{\delta \beta} \ln \left[\int_{t_i}^{\infty} g_{ij}(u - t_i) \frac{S(u)}{S(t_i)} du \right] \right\} + \\
&\sum_i \sum_{t_i \in i} \left\{ m_{i0}^{t_i} \frac{\delta}{\delta \beta} \ln \left[1 - \sum_{j=i+1}^k \phi_i(1 - p_{i+1}) \dots \phi_j p_j \int_0^{\infty} g_{ij}(u - t_i) \frac{S(u)}{S(t_i)} du \right] \right\} + \\
&\frac{\delta}{\delta \beta} \ln(L^{radio}) \\
&= \sum_i \sum_{t_i \in i} \left\{ \sum_{j=i+1}^k m_{ij}^{t_i} \frac{\delta}{\delta \beta} \ln \left[\frac{g_{ij}(t_j - t_i) \frac{S(t_j)}{S(t_i)}}{\int_{t_i}^{\infty} g_{ij}(u - t_i) \frac{S(u)}{S(t_i)} du} \right] \right\} + \\
&m_{ij}^{t_i} \frac{\frac{\delta}{\delta \beta} \int_{t_i}^{\infty} g_{ij}(u - t_i) \frac{S(u)}{S(t_i)} du}{\int_{t_i}^{\infty} g_{ij}(u - t_i) \frac{S(u)}{S(t_i)} du} - \\
&m_{i0}^{t_i} \frac{\sum_{j=i+1}^k \phi_i(1 - p_{i+1}) \dots \phi_j p_j \frac{\delta}{\delta \beta} \int_0^{\infty} g_{ij}(u - t_i) \frac{S(u)}{S(t_i)} du}{1 - \sum_{j=i+1}^k \phi_i(1 - p_{i+1}) \dots \phi_j p_j \int_0^{\infty} g_{ij}(u - t_i) \frac{S(u)}{S(t_i)} du} + \frac{\delta}{\delta \beta} \ln(L^{radio}) \\
&= \sum_i \sum_{t_i \in i} \left\{ \sum_{j=i+1}^k m_{ij}^{t_i} \frac{\delta}{\delta \beta} \ln \left[\frac{g_{ij}(t_j - t_i) \frac{S(t_j)}{S(t_i)}}{\int_{t_i}^{\infty} g_{ij}(u - t_i) \frac{S(u)}{S(t_i)} du} \right] \right\} + \frac{\delta}{\delta \beta} \ln(L^{radio}) \\
&\sum_i \sum_{t_i \in i} \sum_{j=i+1}^k \left[\frac{m_{ij}^{t_i}}{\int_0^{\infty} g_{ij}(u - t_i) \frac{S(u)}{S(t_i)} du} - \frac{m_{i0}^{t_i} \phi_i(1 - p_{i+1}) \dots \phi_j p_j}{\chi_i(t_i)} \right] \times \\
&\frac{\delta}{\delta \beta} \int_0^{\infty} g_{ij}(u - t_i) \frac{S(u)}{S(t_i)} du
\end{aligned} \tag{A.14}$$

The first term gives information on censored radio times ($S(u)$) only. In some sense, this must be spread over all $u > t_i$ to give information on the specific radio failures which is expected to be small. The second term is the contribution from the observed radio failures and has good information on the radio failure times. The third term contains a difference of observed and expected counts which is expected to be small. In the 2 dam case all fish were released at dam 1, $t_1 = 0$, so the term reduces to zero.

A.3 Simulation

We ran simulations to look at the 2, 3 and 4 dam case using an exponential distribution (parameterized by $\lambda = 0.067$) for the radio failure time. We estimated λ using a maximum likelihood estimate based only on the observed radios (as proposed in the manuscript) and a maximum likelihood estimate based on the full likelihood.

Table A.1 looks at the results of simulations with 10, 30 and 100 radios and 100, 500, and 1000 fish. Shown in the table are the maximum likelihood estimate of λ , the estimated standard error and the coefficient of variation for each case.

The two estimates of λ are very similar as are the standard errors suggesting there is little loss of information when using the conditional likelihood method.

Table A.1: Conditional and unconditional maximum likelihood estimates of λ (the mean of the exponential distribution), the estimated standard error and the coefficient of variation for 2, 3, and 4 dams, 10, 30, and 100 radios, and 100, 500, and 1000 fish. The true value of $\lambda = 0.067$.

Number of Dams	Number of Radios	Number of Fish	Conditional Likelihood			Unconditional Likelihood		
			λ	SE	CV	λ	SE	CV
2	10	100	0.054	0.017	0.316	0.054	0.014	0.251
		500	0.071	0.022	0.312	0.067	0.014	0.213
		1000	0.058	0.018	0.316	0.055	0.015	0.275
	30	100	0.063	0.012	0.183	0.063	0.013	0.200
		500	0.097	0.018	0.183	0.096	0.016	0.164
		1000	0.074	0.014	0.183	0.073	0.012	0.170
	100	100	0.055	0.005	0.100	0.052	0.005	0.095
		500	0.058	0.006	0.100	0.058	0.006	0.111
		1000	0.064	0.006	0.100	0.064	0.010	0.154
3	10	100	0.054	0.017	0.316	0.053	0.013	0.246
		500	0.096	0.030	0.316	0.095	0.007	0.074
		1000	0.058	0.018	0.316	0.070	0.044	0.063
	30	100	0.071	0.013	0.183	0.072	0.015	0.207
		500	0.066	0.012	0.183	0.067	0.012	0.185
		1000	0.073	0.013	0.183	0.067	0.012	0.174
	100	100	0.076	0.008	0.100	0.076	0.030	0.393
		500	0.072	0.007	0.100	0.072	0.051	0.712
		1000	0.063	0.006	0.100	0.063	0.006	0.098
4	10	100	0.050	0.016	0.316	0.053	0.018	0.335
		500	0.076	0.024	0.316	0.083	0.087	1.05
		1000	0.079	0.025	0.316	0.062	0.017	0.266
	30	100	0.065	0.012	0.183	0.064	0.010	0.152
		500	0.047	0.009	0.183	0.048	0.008	0.175
		1000	0.112	0.020	0.183	0.088	0.005	0.061
	100	100	0.054	0.005	0.100	0.054	0.007	0.127
		500	0.065	0.007	0.100	0.065	0.009	0.132
		1000	0.065	0.007	0.100	0.067	0.030	0.445

Appendix B

Expectations of Latent Variables for the Jolly-Seber Tag-Loss Model

The EM algorithm requires the evaluation of the expectations of the complete-data log-likelihood at the E-step. This reduces to finding the expectations (conditional upon the observed data) of functions of the latent variables. Van Deusen (2002) illustrates this for the Cormack-Jolly-Seber model; similar methods are used below.

For simplicity, the conditioning on the observed data is implicitly assumed in all expressions below.

B.1 $E(a_{ij})$

Time (j)	Expected value
$j < f_i$	$\frac{\Psi_j \prod_{t=j}^{f_i-1} (1-p_t) \phi_t}{\Psi_{f_i}}$
$f_i \leq j \leq l_i$	1
$l_i < j$	$\frac{[\prod_{t=l_i}^{j-1} \phi_t] \xi(f_i, l_i, j, nt_i)}{\chi(f_i, l_i, nt_i)}$

where

$$\xi(f_i, t, j, d) = \begin{cases} (1 - \Lambda_{f_i t}) + \Lambda_{f_i t}(1 - p_{t+1})\chi_{(f_i, t+1, 1)} & \text{if } d = 1 \text{ and } t + 1 = j \\ (1 - \Lambda_{f_i t}) + \Lambda_{f_i t}(1 - p_{t+1})\xi(f_i, t + 1, j, 1) & \text{if } d = 1 \text{ and } t + 1 < j \\ (1 - \Lambda_{f_i t})^2 + \Lambda_{f_i t}^2(1 - p_{t+1})\chi_{(f_i, t+1, 2)} & \text{if } d = 2 \text{ and } t + 1 = j \\ + 2\Lambda_{f_i t}(1 - \Lambda_{f_i t})(1 - p_{t+1})\chi_{(f_i, t+1, 1)} & \\ (1 - \Lambda_{f_i t})^2 + \Lambda_{f_i j}^2(1 - p_{t+1})\xi(f_i, t + 1, j, 2) & \text{if } d = 2 \text{ and } t + 1 < j \\ + 2(1 - p_{t+1})\Lambda_{f_i t}(1 - \Lambda_{f_i t})\xi(f_i, t + 1, j, 1) & \end{cases}$$

B.2 $E(a_{ij}a_{i,j+1})$

Time (j)	Expected value
$j < f_i$	$E(a_{ij})$ from above.
$f_i \leq j < l_i$	1
$l_i \leq j$	$E(a_{i,j+1})$ from above.

B.3 $E(a_{ij}g_{ij+})$

Time (j)	Expected value
$j < f_i$	0 as no tags are present.
$f_i \leq j \leq l_i$	1
$l_i < j$	$\frac{[\prod_{t=l_i}^{j-1} \phi_t(1-p_{t+1})]\xi_2(f_i, l_i, j, nt_i)}{\chi(f_i, l_i, nt_i)}$

where

$$\xi_2(f_i, t, j, d) = \begin{cases} \chi(f_i, j, d) & \text{if } t = j \\ \Lambda_{f_i t}\xi_2(f_i, t + 1, j, 1) & \text{if } d = 1 \text{ and } t < j \\ \Lambda_{f_i t}^2\xi_2(f_i, t + 1, j, 2) + & \text{if } d = 2 \text{ and } t < j \\ 2\Lambda_{f_i t}(1 - \Lambda_{f_i t})\xi_2(f_i, t + 1, j, 1) & \end{cases}$$

B.4 $E(a_{i,j+1}g_{i,jd}g_{i,j+1,d})$

Time (j)	Expected value
$j < f_i$	0 as no tags are present.
$f_i \leq j < l_i^d$	1
$l_i \neq l_i^d, l_i^d \leq j < q_{id}$	$\frac{\prod_{t=l_i^d}^j \Lambda_{f_i t} - \prod_{t=l_i^d}^{q_{id}-1} \Lambda_{f_i t}}{1 - \prod_{t=l_i^d}^{q_{id}-1} \Lambda_{f_i t}}$
$l_i \neq l_i^d, q_{id} < j + 1$	0 as tag d is not present
$l_i = l_i^d \leq j$	$\frac{\prod_{t=l_i}^j \phi_t(1-p_{t+1})\Lambda_{f_i t}\xi_3(f_i, l_i, j, nt_{l_i})}{\chi(f_i, l_i, nt_{l_i})}$

where

$$\xi_3(f_i, t, j, d) = \begin{cases} \chi(f_i, t, d) & \text{if } d = 1 \text{ and } t = j + 1 \\ \xi_3(f_i, t + 1, j, 1) & \text{if } d = 1 \text{ and } t < j + 1 \\ \Lambda_{f_i t}\xi_3(f_i, t + 1, j, 2) + & \text{if } d = 2 \text{ and } t < j + 1 \\ (1 - \Lambda_{f_i t})\xi_3(f_i, t + 1, j, 1) & \end{cases}$$

B.5 $E(g_{i,jd}a_{i,j+1})$

Time (j)	Expected value
$j < f_i$	0 as no tags are present.
$f_i \leq j < l_i^d$	1
$l_i \neq l_i^d, l_i^d \leq j < j + 1 < q_{id}$	$\frac{\prod_{t=l_i^d}^{j-1} \Lambda_{f_i t} - \prod_{t=l_i^d}^{q_{id}-1} \Lambda_{f_i t}}{1 - \prod_{t=l_i^d}^{q_{id}-1} \Lambda_{f_i t}}$
$l_i \neq l_i^d, q_{id} \leq j$	0 as tag d is not present.
$l_i = l_i^d < j + 1$	if $d = 1$, $\frac{\prod_{t=l_i}^{j-1} \phi_t(1-p_{t+1})\Lambda_{f_i t}\phi_j\{\Lambda_{f_i j}(1-p_{j+1})\chi(f_i, j+1, 1) + (1-\Lambda_{f_i j})\}}{\chi(f_i, l_i, 1)}$ if $d = 2$, $\frac{\prod_{t=l_i}^{j-1} \phi_t(1-p_{t+1})\Lambda_{f_i t}\phi_j\xi_4(f_i, l_i, j, 2)}{\chi(f_i, l_i, 2)}$

where

$$\xi_4(f_i, t, j, d) = \begin{cases} \Lambda_{f_i t}(1 - p_{t+1})\chi(f_i, t + 1, 1) & \text{if } d = 1 \text{ and } t = j \\ \xi_4(f_i, t + 1, j, 1) & \text{if } d = 1 \text{ and } t < j \\ \Lambda_{f_i t}^2(1 - p_{t+1})\chi(f_i, t + 1, 2) & \text{if } d = 2 \text{ and } t = j \\ + 2(1 - \Lambda_{f_i t})\Lambda_{f_i t}(1 - p_{t+1})\chi(f_i, t + 1, 1) & \\ + (1 - \Lambda_{f_i t})^2 & \\ \Lambda_{f_i t}\xi_4(f_i, t + 1, j, 2) + (1 - \Lambda_{f_i t})\xi_4(f_i, t + 1, j, 1) & \text{if } d = 2 \text{ and } t < j \end{cases}$$

Appendix C

Sample simulation result for the Jolly-Seber-Tag-Loss Model

To test the model and the FORTRAN 95 code we ran several simulations studies. We include the results of one such study. For a 3-sample experiment with 1000000 fish initially released we set $p = 0.8$, $\phi = 0.8$, $\Lambda = 0.8$, $\nu = 0.1$, and $b_0^* = 0.5, b_1^* = 0.2$. Table C shows the parameter estimates for this simulation. Due to confounding with birth parameters, p_1 was fixed to 0.8. Similarly, due to confounding with ϕ_2 , p_3 was fixed to 0.8.

Table C.1: Simulation results for a 3 sample experiment where $p = 0.8$, $\phi = 0.8$, $\Lambda = 0.8$, $\nu = 0.1$, and $b_0^* = 0.5$, $b_1^* = 0.2$. Note that p_1 and p_3 were fixed to 0.8 due to confounding.

Parameter	Estimate	Standard Error
p_1	0.800	0.000
p_2	0.799	0.001
p_3	0.800	0.000
ϕ_1	0.799	0.002
ϕ_2	0.803	0.003
Λ_{11}	0.802	0.002
Λ_{12}	0.797	0.002
Λ_{22}	0.797	0.002
b_0^*	0.499	0.001
b_1^*	0.200	0.001
ν_1	0.100	0.000
ν_2	0.101	0.001
ν_3	0.100	0.000

Appendix D

Mille Lac Data

This is the data used in the Mille Lac example which was provided by the Department of Natural Resources, Minnesota, U.S.A. There is one capture history that is not possible and its frequency is coded with a †. This fish was single tagged with tag 1 and was recaptured and recorded as having a tag 2. This capture history was removed from the analysis. Negative frequencies indicate loss on capture.

Tag History	Frequency	Tag History	Frequency
000000000010	561	000000001111	-2
000000000010	-1166	000000010000	4
000000000011	197	000000010000	-1
000000000011	-1	000000100000	3421
000000000100	6	000000100000	-1725
000000000100	-6	000000100010	14
000000000100	15190	000000100010	-4
000000000100	-30599	000000101000	81
000000000101	38	000000101000	-85
000000000101	-26	000000101010	1
0000000001100	1205	000000110000	456
0000000001100	-51	000000110001	1
0000000001101	2	000000110011	1
0000000001111	3	000000110011	-1

Tag History	Frequency	Tag History	Frequency
000000110100	1	000011001100	87
000000110100	-2	000011001100	-53
000000111000	2	000011001111	1
000000111000	-2	000011010100	-1
000000111100	4	000011110000	20
000000111100	-9	000011110000	-7
000001000000	6	000011111100	1
000001000000	-2	000100000000	4
000001000100	1	001000000000	9888
000010000000	22371	001000000000	-1869
000010000000	-34324	001000000010	21
000010000010	20	001000000010	-3
000010000010	-17	001000001000	77
000010001000	1758	001000001000	-137
000010001000	-1161	001000100000	130
000010001010	3	001000100000	-4
000010001010	-2	001000100010	1
000010100000	186	001000101000	6
000010100000	-64	001000101000	-1
000010101000	25	001010000000	158
000010101000	-10	001010000000	-180
000011000000	1584	001010000010	-1
000011000000	-62	001010001000	13
000011000001	1	001010001000	-7
000011000001	-2	001010100000	6
000011000011	1	001010101000	1
000011000100	48	001100000000	1050
000011000100	-35	001100000100	5
000011001000	18	001100000100	-10
000011001000	-13	001100000101	-1

Tag History	Frequency	Tag History	Frequency
001100001000	-1	100001000000†	-1
001100001100	4	100010000000	1176
001100001100	-5	100010000000	-512
001100010000	4	100010000010	1
001100110000	6	100010000010	-2
001101000000	7	100010001000	249
001101000000	-11	100010001000	-62
001101000100	2	100010001010	2
001101000100	-2	100010001010	-2
001101010000	2	100010100000	16
001110000000	1	100010100000	-6
001110000000	-2	100010101000	5
001110001000	1	101000000000	166
001111000000	13	101000000000	-16
001111000000	-9	101000001000	3
001111001100	-1	101000001000	-3
010000000000	1	101010000000	16
010000000000	-1	101010000000	-10
100000000000	15394	101010001000	4
100000000000	-29473	110000000000	1053
100000000010	10	110000000000	-18
100000000010	-2	110000000100	17
100000001000	486	110000000100	-5
100000001000	-202	110000001000	7
100000100000	38	110000001000	-1
100000100000	-5	110000001100	24
100000100010	1	110000001100	-7
100000101000	7	110000010000	2
100000101000	-6	110000010000	-1
100000101010	-1	110000100000	3

Tag History	Frequency
110000110000	3
110001000000	41
110001000000	-20
110001000100	14
110001010000	1
110010000000	24
110010000000	-7
110010001000	2
110010111100	1
110011000000	53
110011000000	-22
110011000010	1
110011000011	1
110011000100	4
110011001000	1
110011001100	10
110011001100	-6
111100000000	11
111100000000	-1
111100001100	2
111101000100	1

Bibliography

- Abramowitz, M. and Stegun, I. A. editors. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover.
- Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory*, B. N. Petrov and F. Csaki (eds), 267-281. Budapest: Akademiai Kiado.
- Akaike, H. (1978) A Bayesian analysis of the minimum AIC procedure. *Annals of the Institute of Statistical Mathematics* **30**, 9-14.
- Alt, G. L., McLaughlin, C. R. and Pollock, K. H. (1985). Ear Tag loss by black bears in Pennsylvania. *Journal of Wildlife Management* **49**, 316-320.
- Arnason, A. N. and Mills, K. H. (1981). Bias and loss of precision due to tag loss in Jolly-Seber estimates for mark-recapture experiments. *Canadian Journal of Fisheries and Aquatic Science* **38**, 1077-1095.
- Barrowman, N. J. and Myers, R. A. (1996). Estimating tag-shedding rates for experiments with multiple tag types. *Biometrics* **52**, 1410-1416.
- Bradshaw, C. J. A., Barker, R. J. and Davis, L. S. (2000). Modeling tag loss in New Zealand fur seal pups. *Journal of Agricultural, Biological, and Environmental Statistics* **5**, 475-485.
- Bratthey, J. and Cadigan, N. (1998). Exploitation rates and population size of cod in Placentia Bay (Subdivision 3Ps): estimates from a new mark-recapture study. *Canadian Stock Assessment Secretariat Research Document* **98/20**, 1-27.

- Brooks, E. A., Catchpole, B. J. and Morgan, B. J. T. (2000a). Bayesian animal survival estimation. *Statistical Science* **15**, 357-376.
- Brooks, E. A., Catchpole, B. J., Morgan, B. J. T. and Barry, S. C. (2000b). On the bayesian analysis of ring-recovery data. *Biometrics* **56**, 951-956.
- Brownie, C., Anderson, D. R., Burnham, K. P. and Robson, D. S. (1985). *Statistical Inference from Band Recovery Data- a Handbook, 2nd edition*. United States Fish and Wildlife Service, Resource Publication No. 156, Washington, D.C.: United States Department of the Interior.
- Bruesewitz, R. (2002). *A tagging study of Mille Lacs Lake walleye*. Minnesota Department of Natural Resources, Fisheries, St. Paul, Minnesota. Project Proposal F-29-R(p) Study IV, Job 619.
- Buckland, S. T. and Garthwaite, P. H. (1991). Quantifying precision of mark-recapture estimates using the bootstrap and related methods. *Biometrics* **47**, 255-268.
- Buckland, S. T., Goudie, I. B. J. and Borchers, D. L. (2000). Wildlife population assessment: past developments and future directions. *Biometrics* **56**, 1-12.
- Burnham, K. P. (1991). On a unified theory for release-resampling of animal populations. In *Proceedings of 1990 Taipei Symposium in Statistics*, M. T. Chao and P. E. Cheng (eds), 11-36. Taipei, Taiwan: Institute of Statistical Science, Academia Sinica.
- Burnham, K. P. and Anderson, D. R. (2001). Kullback-Leibler information as a basis for strong inference in ecological studies. *Wildlife Research* **28**, 111-119.
- Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach, 2nd edition* New York: Springer-Verlag.
- Burnham, K. P., Anderson, D. R., White, G. C., Brownie, C. and Pollock, K. H. (1987). *Design and Analysis Methods for Fish Survival Experiments Based on Release-Recapture*. Bethesda, Maryland: American Fisheries Society.

- Cadigan, N. and Brattey, J. (1999). Estimation of 3PSc and 3Kl coarse exploitation and migration rates in 1997 and 1998 from tagging experiments in these regions. *Zonal Cod Assessment* **WP(99/26)**, 1-34.
- Chao, A. (2001). An overview of closed capture-recapture models. *Journal of Agricultural, Biological, and Environmental Statistics* **6**, 158-175.
- Chapman, D. G. (1948). A mathematical study of confidence limits of salmon populations calculated from sample tag ratios. *International Pacific Salmon Fisheries Commission Bulletin II*, 69-85.
- Chhikara, R. S. and Folks, J. L. (1989). *The Inverse Gaussian Distribution Theory, Methodology, and Applications*. New York: Marcel Dekker, Inc.
- Cormack, R. M. (1964). Estimates of survival from the sighting of marked animals. *Biometrika* **51**, 429-438.
- Cormack, R. M. (1979). Models for capture-recapture. In *Sampling Biological Populations*, R. M. Cormack, G. P. Patil and D. S. Robson (eds), 217-255. Fairland: International Co-operative Publishing House.
- Cormack, R. M. (1989). Log-linear models for capture-recapture. *Biometrics* **45**, 395-413.
- Crosbie, S. F. and Manly, B. F. (1985). Parsimonious modeling of capture-mark-recapture studies. *Biometrics* **41**, 385-398.
- Darroch, J. N. (1959). The multiple recapture census. II. Estimation when there is immigration or death. *Biometrika* **46**, 336-359.
- Diefenbach, D.R. and Alt, G.L. (1998). Modeling and evaluation of ear tag loss in black bears. *Journal of Wildlife Management* **62**, 1292-1300.
- Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia: Society for Industrial and Applied Mathematics.

- English, K. K., Skalski, J. R., Lady, J., Koski, W. R., Nass, B. L. and Sliwinski, C. (2001a). *An Assessment of Project, Pool, and Dam Survival for Run-of-River Steelhead Smolts at Wanapum and Priest Rapids Projects Using Radio-telemetry Techniques, 2000*. Public Utility District No. 2 of Grant County, Washington. Draft report.
- English, K. K., Skalski, J. R., Lady, J., Robichaud, D., Nass, B. L., Sliwinski, C. and Blakley, A. (2001b). *Survival of Hatchery Chinook Smolts assessed at Wanapum and Priest Rapids Projects using Radio-telemetry Techniques*. Public Utility District No. 2 of Grant County, Washington. Draft report.
- Hilborn, R. (1990). Determination of fish movement patterns from tag recoveries using maximum likelihood estimators. *Canadian Journal of Fisheries and Aquatic Sciences* **47**, 635-643.
- Hoenig, J. M., Barrowman, N. J., Pollock, K. H., Brooks, E. N., Hearn, W. S. and Polacheck, T. (1998). Models for tagging data that allow for incomplete mixing of newly tagged animals. *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 1477-1483.
- Hubert, G. F. (1975). Ear tag loss in red foxes. *Journal of Wildlife Management* **41**, 164-167.
- Huvich, C. M. and Tsai, C-L. (1989). Regression and time series model selection in small samples. *Biometrika* **76**, 297-307.
- Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika* **52**, 225-247.
- Jolly, G. M. (1979). A unified approach to mark-recapture stochastic models exemplified by a constant survival rate model. In *Sampling Biological Populations*, R. M. Cormack, G. P. Patil and D. S. Robson (eds), 277-282. Fairland: International Co-operative Publishing House.

- Laplace, P. S. (1786). Sur les naissances, les mariages et les morts. *Histoire de l'Academie Royale des Sciences Annee 1783*. Paris.
- Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. New York: John Wiley and Sons.
- Lebreton, J.D., Burnham, K.P., Clobert, J. and Anderson, D.R. (1992). Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs* **62**, 67-118.
- Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society B* **44**, 226-233.
- McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models, 2nd edition*. Boca Raton: Chapman and Hall/CRC.
- McDonald, T. L., Amstrup, S. C. and Manly, B. F. J. (2003). Tag loss can bias Jolly-Seber capture-recapture estimates. *Wildlife Society Bulletin* **31**, 814-822.
- Myers, R. A. and Hoenig, J. M. (1997). Direct estimates of gear selectivity from multiple tagging experiments. *Canadian Journal of Fisheries and Aquatic Sciences* **54**, 1-9.
- Otis, D. L., Burnham, K. P., White, G. C. and Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs* **62**.
- Pistorius, P. A., Bester, M. N., Kirkman, S. P. and Boveng, P. L. (2000). Evaluation of age- and sex-dependent rates of tag loss in southern elephant seals. *Journal of Wildlife Management* **64**, 373-380.
- Poirier, G., Morin, R. and Swain, D. (1997). The status of yellowtail flounder in NAFO division 4T in 1996. *DFO Canadian Stock Assessment Research Document* **97/70**.
- Pollock, K. H., Nichols, J. D., Brownie, C. and Hines, J. E. (1990). Statistical inference in capture-recapture experiments. *Wildlife Monographs* **107**.

- Pollock, K. H., Bunck, C. M., Winterstein, S. R. and Chen, C. (1995). A capture-recapture survival analysis model for radio-tagged animals. *Journal of Applied Statistics* **22**, 661-672.
- Poole, D. (2002). Bayesian estimation of survival from mark-recapture data. *Journal of Agricultural, Biological and Environmental Statistics* **7**, 264-276.
- R Development Core Team (2004). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Robson, D. S. and Regier H. A. (1966). Estimates of tag loss from recoveries of fish tagged and permanently marked. *Transactions of the American Fisheries Society* **95**, 56-59.
- Sanathanan, L. (1977). Estimating the size of a truncated sample. *Journal of the American Statistical Association* **72**, 669-672.
- Sandland, R. L. and Cormack, R. M. (1984). Statistical inference for Poisson and multinomial models for capture-recapture experiments. *Biometrika* **71** 27-33.
- Seber, G. A. F. (1965). A note on the multiple recapture census. *Biometrika* **52**, 249-259.
- Seber, G. A. F. and Felton, R. (1981). Tag loss and the Petersen mark-recapture experiment. *Biometrika* **68**, 211-219.
- Seber, G. A. F. (1982). *The Estimation of Animal Abundance and Related Parameters*. New York: John Wiley and Sons.
- Schwarz, C. J. (2003). *Estimating the Number of Walleye in Mille Lacs, Minnesota*. Report prepared for the Minnesota Department of Natural Resources.
- Schwarz, C. J. (2004). *Estimating the Number of Walleye in Mille Lacs, Minnesota-Final Report*. Report prepared for Minnesota Department of Natural Resources.

- Schwarz, C. J. and Arnason, A. N. (1996). A general methodology for the analysis of capture-recapture experiments in open populations. *Biometrics* **52**, 860-873.
- Schwarz, C. J., Schweigert, J. F. and Arnason, A. N. (1993). Estimating migration rates using tag-recovery data. *Biometrics* **49**, 177-193.
- Skalski, J. R., Lady, J., Townsend, K., Giorgi, A. E., Stevensen, J. R., Peven, C. M. and McDonald, R. D. (2001). Estimating in-river survival of migrating salmonoid smolts using radio-telemetry. *Canadian Journal of Fisheries and Aquatic Sciences* **58**, 1987-1997.
- Skalski, J. R., Smith, S. G., Iwamoto, R. N., Williams, J. G. and Hoffmann, A. (1998). Use of passive integrated transponder tags to estimate survival of migrant juvenile salmonoids in the Snake and Columbia rivers. *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 1484-1493.
- Skalski, J. R., Stevenson, J. R., Lady, J., Townsend, R., Giorgi, A. E., Miller, M. and English, K. (2001). *An Assessment of Project, Pool, and Dam Survival for Chinook and Steelhead Smolts at Rocky Reach and Rock Island Dams Using Radiotelemetry and PIT-tag Techniques, 2000*. Prepared for Public Utility District No. 1 of Chelan Country, Wenatchee, Washington.
- Stobo, W. T. and Horne, J. K. (1994). Tag loss in grey seals (*Halichoerus grypus*) and potential effects on population estimates. *Canadian Journal of Zoology* **72**, 555-561.
- Thisted, R. A. (1988). *Elements of Statistical Computing: Numerical Computation*. New York: Chapman and Hall.
- VanDeusen, P. C. (2002). An EM algorithm for capture-recapture estimation. *Environmental and Ecological Statistics* **9**, 151-165.
- Vounatsou, P. and Smith, A. F. M. (1995). Bayesian analysis of ring-recovery data via Markov chain Monte Carlo simulation. *Biometrics* **51**, 687-708.

- Walsh, S. J., Morgan, M. J. Brodie, W. B., Dwyer, K. S. and Mansfield, L. (2001). *A New Tagging Program for Yellowtail Flounder on the Grand Bank, NAFO Divs. 3LNO*. NAFO SCR Doc. 01/53, Serial No. N4331, 1-21.
- Wetherall, J. A. (1982). Analysis of double-tagging experiments. *Fishery Bulletin* **80**, 687-701.
- White, G. C. and Burnham, K. P. (1999). Program MARK- Survival estimation from populations of marked animals. *Bird Study* **46** Supplement, 120-138.
- Williams, B. K., Nichols, J. D. and Conroy, M. J. (2001). *Analysis and Management of Animal Populations. Modeling, Estimation, and Decision Making*. San Diego: Academic Press.