Partial Year Tagging Models:

Accounting for Changing Tag Visibility and Delayed Mixing

_____

A Thesis

Presented to

The Faculty of the School of Marine Science

The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Science

_____

by

Lynn Waterhouse

2010

APPROVAL SHEET

This thesis is submitted in partial fulfillment of

the requirements for the degree of

Master of Science

_____
Lynn Waterhouse

Approved, August 2010

_____
John M. Hoenig, Ph.D.
Co-Advisor

_____
Mary Fabrizio, Ph.D.
Co-Advisor

_____
Romuald N. Lipcius, Ph.D.

_____
Ross Iaci, Ph.D.
Department of Mathematics
College of William and Mary in Virginia

_____
Clay Porch, Ph.D.
Southeast Fisheries Science Center
National Marine Fisheries Service
Miami, FL

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I am sincerely thankful to my co-advisors, Drs. John M. Hoenig and Mary Fabrizio, and committee members, Drs. Clay Porch, Romuald N. Lipcius, and Ross Iaci for their guidance and support. In particular, thanks to Dr. Hoenig: I appreciate his patience with the struggles encountered in my project and the flexibility to adapt to changes while staying focused on producing a scholarly product and to further my career; to Dr. Fabrizio: her willingness to become a mentor and her ability to stay focused on the main goal; to Dr. Porch: for an analytical look at my thesis work and striving to make it of the upmost quality; to Dr. Lipcius: for wholeheartedly jumping in on my original project and continuing to support me when the direction changed and to Dr. Iaci: for patiently dealing with my lengthy answers and providing a rigorous statistical viewpoint to my work.

This project was accomplished with help and support from many colleagues, friends, and family. The striped bass group: Lydia Goins, Phil Sadler, and Dan Gonzales allowed me to gain field experience outside of my research and helping with my short-lived field project. I also thank Dr. Graves and Dr. Latour for providing guidance and support, as well as being excellent teachers.

Many thanks to all of my family members who supported me through the many ups and downs of this process, especially to my parents Lorraine and Buzz Waterhouse and my brother Brian (the only member of the family with a "real" job), my grandmother Ruth Bohrer, for being a true inspiration, both intellectually and spiritually, your prayers and advice gave me strength and my grandparents Ellen and Jack Waterhouse. I am thankful for the friends I made while here and the ones who stayed with me from my past: Sarah Barbo, Noelle Relles, Aaron Relles, Matt Smith, Asheley Supick, Kersey Sturdivant, Mark Henderson, Dan Dutton, Daniel Vogt, among many others. You and your dogs made my VIMS experience truly wonderful.

LIST OF TABLES

APPENDIX B

LIST OF FIGURES

Figure	**Page**

Partial Year Tagging Models:

Accounting for Changing Tag Visibility and Delayed Mixing

ABSTRACT

The "Brownie" tagging models are commonly used for estimating survival rates from multi-year tagging studies in which cohorts of tagged animals are released at the start of each year. Brownie models can be reparameterized in terms of instantaneous rates of fishing and natural mortality, and these models are referred to as instantaneous rates models. Typically the recaptures of tagged animals are tabulated on the same periodicity as that of tagging, typically a year. This thesis shows two situations in which tabulating the tag recaptures by part of the year rather than full year proves to be advantageous: a Brownie-type model for the case when tag visibility of newly tagged cohorts is different from previously tagged cohorts (referred to as model 0') and an instantaneous rates model for delayed mixing of newly tagged animals with previously tagged animals that lasts part of the year (referred to as delayed pyt model).

Model 0' allows for newly tagged animals to have a different tag recovery rate than previously tagged ones. It makes use of a known fouling time (or change in visibility time), the time it takes for newly tagged animals to have the same visibility as previously tagged animals, to divide the year into two parts. During the first part of the year, newly tagged animals are more visible than previously tagged ones while in the second part all tagged animals have the same visibility. Dividing the year into parts and recording recaptures in each part avoids the failure of the assumption that the reporting rate is constant for all tagged animals, achieves greater precision (smaller standard errors), and provides estimates of the survival rate at the end of the 2nd year instead of after the 3rd year. The superiority of model 0' is demonstrated through Monte Carlo simulation.

Hoenig et al. described instantaneous rates models that assume full mixing and others that allow for the newly tagged population to become fully mixed in less than one year. The delayed pyt model divides the year into parts and tag returns are tabulated by parts of the year rather than a full year. This is beneficial when there is delayed mixing because it achieves greater precision and provides estimates of the instantaneous rate of fishing mortality in the first year, which cannot be estimated when tag returns are tabulated by full year. The new model can be used at little or no extra cost. The superiority of the delayed pyt model is demonstrated through Monte Carlo simulation.

Chapter 1

Introduction and Motivation

Introduction

One way to evaluate the status of a stock of exploited animals is to estimate the survival rate (or its complement, mortality rate, which is equal to one minus the survival rate). The survival rate can then be explained in terms of its components, the exploitation rate (fishing mortality) and natural death rate (mortality due to all other causes). These can be used to solve for the instantaneous rates of fishing and natural mortality using information on how the fishery functions during the course of the year. Instantaneous rates of fishing mortality may also be used as a means of evaluating the status of fish stocks. Survival rates and the components of mortality can be estimated from a tagging study in which animals are captured, tagged or marked with unique identifying numbers, and released. From the recovery of tags from the harvest (fishery) one can estimate survival rates and components of mortality. This thesis describes two advancements in tagging theory involving tabulation of tag returns by part of the year rather than full year. The first is a model that addresses the problem when the tag visibility for the first part of the first year a tagged cohort is at liberty is different from the tag visibility of previously released tagged cohorts. This model is called model 0'. The second model addresses delayed mixing of the newly tagged cohort with the population at large that lasts the first part of the first year a tagged cohort is at liberty. This model is called the partial-year tabulation for delayed mixing model (delayed pyt model).

Definitions

Survival rate, represented by *S*, is defined as the fraction of the population alive at the start of a time period (typically a one-year interval) that is still alive at the end of the

time period. The relationship between survival rate and the total instantaneous mortality rate, Z, is:

(1) $\quad S = e^{-Zt}$,

which implies

(2) $\quad Zt = -\ln S$,

where t denotes the length of a time interval. Here, Z has units of time^{-1} and, if not specified, t is assumed to be one time unit. Total instantaneous mortality rate refers to mortality due to all causes combined. The product $-ZtN$ represents the amount by which a population would change due to deaths in a short period of time, t, where N is the population size at the start of the time interval.

The instantaneous total mortality rate can be split into two components, natural mortality, M, and harvest mortality, F (in the case of fisheries, fishing mortality) (e.g., Ricker 1975). Thus,

(3) $\quad Z = F + M$,

where fishing and natural mortality are both instantaneous rates. The relationship between annual survival rate ($t = 1$) and the instantaneous rate of total mortality can be rewritten as:

(4) $\quad S = e^{-(F+M)}$.

The annual mortality rate A is the fraction of the population alive at the start of the year that dies during the year. It is the complement of the annual survival rate S and thus can be specified as:

(5) $\quad A = 1 - S$,

implying

(6) $S = 1 - A$.

Note that A and S are finite rates and are unitless. The death rate can be split into the finite death rate from fishing, or the exploitation rate, denoted by u, and the finite death rate due to natural causes, denoted by v. Thus:

(7) $A = u + v$,

where u is the exploitation rate, which is the fraction of the population alive at the start of year that dies due to harvest during the year and v is the fraction of the population alive at the start of the year that dies due to natural causes during the year.

It is necessary to know the timing of the fishery during the year in order to relate the exploitation rate to the instantaneous rates of fishing and natural mortality. Fisheries can be conveniently described as either a type I or type II fishery (e.g., Ricker 1975), depending on the placement of fishing mortality during the year. A type I fishery is a pulse fishery, meaning that all the fishing effort is concentrated at the start of the fishing year, and during the rest of the year no fishing mortality takes place. A type II fishery is a continuous fishery in which fishing mortality takes place during the entire year at a constant intensity. More generally, a characteristic of a type II fishery is that the ratio of fishing to natural mortality is constant over the entire year. Depending on the type of fishery, the exploitation rate and death rate due to natural causes can be expressed in terms of instantaneous rates of fishing and natural mortality.

For a pulse fishery, the year is divided into two parts: during the fishery and during the rest of the year. During the fishery u and v can be expressed as:

(8a) $u = 1 - e^{-F}$

and

(8b) $v = 0$;

during the rest of the year

(9a) $u = 0$

and

(9b) $v = 1 - e^{-M}$.

For a continuous fishery:

(10) $u = \dfrac{F}{F + M}\left(1 - e^{-(F+M)}\right)$

and

(11) $v = \dfrac{M}{F + M}\left(1 - e^{-(F+M)}\right)$ (see Ricker 1975).

The exploitation rate, u, can also be expressed for an arbitrary fishing pattern over the

year (Hoenig et al. 1998a).

Tagging models

Multi-year, single recapture tagging studies provide a method of estimating

survival rates, components of mortality, and the annual exploitation rates (e.g., Brownie

et al. 1978; Pollock et al. 1991; Hoenig et al. 1998a). Another method is to perform a

capture-recapture study (e.g. Jolly 1965; Seber 1965), where it is possible to have

multiple resightings of a marked individual. This work will focus on tagging studies in

which recaptures come from harvested animals, thus only one resighting (i.e., the harvest)

of each marked animal is possible. The methods of Brownie et al. (1978, 1985) are

widely used to estimate survival rates from such studies. Brownie models are

parameterized in terms of annual survival rates, but Hoenig et al. (1998a,b) provided

formulations that permit estimation of instantaneous rates of fishing and natural mortality from the Brownie models. All of these models tabulate tag recaptures over periods which correspond to the intervals between tag releases, which are typically a year. This means that estimates of survival rates, for Brownie et al. (1978, 1985) models, and components of natural mortality, in the Hoenig et al (1998a,b) formulations, cannot be calculated until the end of a full time period.

Brownie Models

Model 1

Brownie et al. (1978, 1985) described a suite of models that enables the user to estimate age- and year-specific survival rates from multi-year tagging studies from which tag recoveries are recorded. The basic age-invariant model is known as model 1. At the start of each time period, here assumed to be a year, a sample of the population is captured, tagged, and released for a given number of years. It is assumed that the tagged sample is representative of the population of interest. The tagged sample is termed a cohort; during each additional year a new cohort is tagged and released. Throughout the year, the animals experience mortality due to natural causes and from harvest. A fraction of the tagged animals that are harvested will be reported. It is assumed that the tag reporting rate will not vary among cohorts within one year. In a two year tagging study, recaptures during the second year of the cohorts tagged in the first and second year can be used to estimate survival during the first year. The number of recaptures in the second year from the cohort tagged in the first year is the number tagged in year one multiplied by the fraction that survived the first year, multiplied by the fraction that is caught and

reported in the second year. Similarly, the number of recaptures in the second year from the cohort tagged in the second year will be the number tagged in the second year multiplied by the fraction that is caught and reported. Thus, the fraction of tags recovered from the first cohort should equal the fraction from the second cohort except for the fact that the first cohort has experienced an extra year of mortality, thereby reducing the number of potential tag returns from the first cohort. More formally, the expected recaptures in year two from animals tagged in year one, $E(R_{12})$, can be modeled as:

$$E(R_{12}) = N_1 S_1 f_2 ,$$

where N is the number tagged, S_1 is the survival rate in year one and f is the tag recovery. In contrast, the expected recaptures in year two from animals tagged in year two, $E(R_{22})$, can be modeled as:

$$E(R_{22}) = N_2 f_2 .$$

Thus the ratio of recaptures is an estimate of S_1 by the method of moments. The complete mathematical formulation for Model 1 is presented in Chapter 2.

Model 2

If survival is constant across years, model 2 may be used (Brownie et al. 1978, 1985). Because the model has fewer parameters to estimate, model 2 may be more precise. However, if survival varies from year to year then model 2 will be biased. Note that model 2 is nested within model 1, because $S_1 = S_2 = \ldots = S_{J-1}$, where J is the number of years of recoveries.

Model 0

Model 0 is a more generalized version of model 1 (Brownie et al. 1978, 1985). This model allows tag-recovery rates, *f*, to vary between newly tagged and previously tagged animals and thus affords some protection from bias. Note that model 1 is nested within model 0.

New Model: Model 0'

One way in which the tag-recovery rates for newly tagged and previously tagged animals differs is as a result of different tag visibilities. For a variety of tagging studies, tag fouling over time (Dicken et al. 2006; Lowry and Suthers 1998; Tarbath 1999; Verweij and Nagelkerken 2007) has been reported to be prominent and this tag fouling could affect tag visibility and, thus, reporting rate (Figure 1). The problem of differing tag visibility between a newly tagged and previously tagged cohort in a given year applies to more general situations than just tag fouling. For example, tagging studies with visible implant elastomer tags have reported skin thickening resulting in reduced visibility (Curtis 2006; Reeves and Buckmeier 2009).

If tag fouling or some other mechanism affects tag visibility, and thus reporting rates, model 0 might be more appropriate than model 1 by accounting for this difference in reporting rates for the first year a tagged cohort is at liberty. But, model 0 can be considered an inefficient model, as it requires three years of observations before the first estimate of survival rate can be made, compared to model 1 which takes two. Model 0 also has more parameters than model 1 which tends to lead to less precision (larger standard errors). Because tagging studies can be expensive and time consuming, and

10

management agencies can be anxious for results, models which make better use of the data are desirable. If tag visibility changes quickly, and the lapsed time is obtainable, one can make better use of the data by dividing the year into two parts. As will be seen in Chapter 2, this allows one to obtain estimates of the survival rates a year sooner than with model 0. In terms of fisheries and wildlife management, such as setting catch limits and evaluating stock status, the ability to produce estimates in the first part of a year, instead of waiting a full year, could prove advantageous. I will investigate the benefits of the application of a tagging model, called model 0', for which the recaptures are tabulated over two parts of the year, rather than tabulating over an entire year.

Instantaneous Rates Models

Fully Mixed Model

The Brownie models can be re-expressed in terms of instantaneous rates of fishing and natural mortality given an arbitrary pattern of fishing effort over the year (Hoenig et al. 1998a,b). The basic models assume that the tagged population mixes completely with the population at large prior to the start of the fishery.

Delayed Mixing Model

The assumption of immediate full mixing is not necessary and instead an extra parameter can be included in the model to account for a delay in mixing. This is accomplished by allowing newly tagged animals to have an abnormal fishing mortality rate, i.e., one that is different from that experienced by previously tagged animals. Hoenig et al. (1998b) described a model that allowed for the tagged animals to be fully

11

mixed with the untagged population after an entire year at liberty, as well as a model that

allowed for the tagged animals to be fully mixed after part of a year has elapsed. A delay

in mixing could result in the newly tagged cohorts experiencing different fishing

mortality than previously tagged cohorts within a given year. The problem of delayed

mixing of newly tagged cohorts may be one of the largest problems facing tagging

studies as it is hard to ensure tagged cohorts fully mix into the population at large prior to

the start of fishing. Because recaptures are tabulated by year, there are confounded

parameters and the model parameterization does not allow for an estimate of the

instantaneous rate of fishing mortality in the first year.

New Model: partial-year tabulation for delayed mixing model

I will investigate extensions of the instantaneous rates models with partial year

tabulation, which accounts for delayed mixing lasting less than a year, and this model is

referred to as the partial-year tabulation for delayed mixing model (delayed pyt model).

Partial year tabulation allows for: parameter estimates after only part of the year has

passed, increased precision, and additional parameter estimation.

Maximum Likelihood Estimation

Maximum likelihood estimation (MLE) provides a general method to obtain

estimates for unknown model parameters based on the observed data. MLE methods

consist of writing an expression, called the likelihood function, for the probability of

having obtained the observed data; the likelihood is written in terms of the unknown

parameters of interest. The values of the parameters which maximize the likelihood

function are called the maximum likelihood estimates. Maximum likelihood estimators have the following desirable asymptotic properties under very general regularity conditions (particularly that the likelihood function is twice differentiable, which is the case for this thesis): unbiased, unique, consistent, minimum variance, and normally distributed (e.g., Hogg et al. 2005; Wackerly et al. 2002).

The freely available statistical package R (R Development Core Team 2008) will be used to calculate maximum likelihood estimates. Both Brownie models and instantaneous rates models can be expressed as the product of independent multinomial distributions of tag returns over time. For each tagged cohort, the cell counts of recaptures and tags never seen again are assumed to follow a multinomial distribution. The general form of the likelihood function Λ for product multinomial models can be expressed as:

$$(26) \quad \Lambda \propto \prod_{i=1}^{I} \left(\prod_{j=i}^{J} P_{ij}^{r_{ij}} \right) \left(1 - \sum_{j=i}^{J} P_{ij} \right)^{N_i - \sum_{j=i}^{J} r_{ij}},$$

where the symbol $\propto$ means "is proportional to," I is the number of years during which tagging occurs, J is the number of years for which recaptures are observed, P_{ij} is the cell probability of observing an animal tagged in year i and recaptured in year j, N_i is the number tagged in year i, and r_{ij} is the observed recaptures in year j from the cohort tagged in year i. The P_{ij}'s are functions of S and f for Brownie models and M and F for the instantaneous rates models. These functions are substituted for the P_{ij}'s in the likelihood function and the likelihood is maximized with respect to these substituted parameters. The likelihood function is proportional to the product of all cell probabilities

for year i raised to the power of observed recaptures in the corresponding cell over all years during which tagging occurs ($i = 1, …, I$). The right hand portion of the equation involving one minus a sum is the expression for the animals in a cohort that are never seen again after tagging.

The variance-covariance matrix can be estimated using the inverse of the Fisher information matrix (Burnham and Anderson 2002; Hogg et al. 2005; and Seber 2002). When there are k parameters $\theta_1, \theta_2, ..., \theta_k$, the Fisher information matrix, I, is the k x k matrix with elements equal to the negative of the expectation of the mixed partial second derivatives of the natural logarithm of the likelihood function (given in equation 26). Thus, the resulting matrix is:

$$(27)\ I = -\begin{bmatrix} \mathrm{E}\left[\dfrac{\partial^2 \ln \Lambda}{\partial \theta_1{}^2}\right] & \cdots & \mathrm{E}\left[\dfrac{\partial^2 \ln \Lambda}{\partial \theta_1 \partial \theta_k}\right] \\ \vdots & \ddots & \vdots \\ \mathrm{E}\left[\dfrac{\partial^2 \ln \Lambda}{\partial \theta_k \partial \theta_1}\right] & \cdots & \mathrm{E}\left[\dfrac{\partial^2 \ln \Lambda}{\partial \theta_k{}^2}\right] \end{bmatrix},$$

where E[] denotes the expectation operator, the symbol ∂ denotes the partial derivative, and Λ is the likelihood function (given in equation 26). Estimates of the variance-covariance matrix can be made using the observed information. The vector of maximum likelihood estimates ($\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k$) is substituted for the vector of parameters ($\theta_1, \theta_2, ..., \theta_k$) and then the inverse of the matrix is computed to obtain the variance-covariance matrix. When this substitution is made, the expectation is no longer taken and instead the expression uses the observed information.

Simulation Methods:

Program SURVIV enables users to create, and maximize, likelihood functions for product multinomial models (White 1983, 1992) and is available from the USGS Patuxent Wildlife Research Center website (www.mbr-wrc.usgs.gov/software/survive.html). The online version of SURVIV will be used for preliminary evaluation of the performance of model 0'.

The statistical package R will be used for model evaluation since it offers more flexibility than SURVIV. To evaluate the performance of model 0', simulated tag-recapture data will be generated and each of the three models (one, zero, and zero prime) will be fit to the data. For the instantaneous rates delayed mixing models, data will be generated and then fit with the Hoenig et al. (1998b) partial year model and the delayed pyt model, the new model in which tag-recaptures are tabulated by part of the year. Additionally, lake trout (*Salvelinus namaycush*) tag-recapture data from Cayuga Lake, New York (Youngs and Robson 1975), will be modified and used as an example to demonstrate how one would select between models allowing for changing tag visibility and delayed mixing (Thesis Appendix A and Thesis Appendix B, respectively). For the simulations, since the true parameter values are known, the parameter estimates will be compared in terms of:

1) bias,

2) bias of the estimated standard error, and

3) root mean squared error (RMSE).

Thesis Structure

This thesis will evaluate the performance of these partial year tabulation tagging

models. The structure of the thesis is as follows:

Chapter I.   Introduction

Chapter II.   Tagging models for estimating survival rates when tag visibility
     changes over time: partial year tabulation of recaptures
     a. Model specifications
     b. Simulation in R (comparison to models 1 and 0)
     c. Discussion

Chapter III.   Instantaneous rates tagging models allowing for delayed mixing of
     newly tagged cohorts: partial year tabulation of recaptures
     a. Model specifications
     b. Simulation in R (comparison to Hoenig et al. (1998b) model)
     c. Simulation in R (comparison to Hoenig et al. (1998a) fully mixed
         model)
     d. Discussion

Appendix A.   Choosing between models 0, 0', and 1: Lake trout example

Appendix B.   Choosing between an instantaneous rates model with full mixing
     and delayed mixing: Lake trout example

References

Brownie, C. , Anderson, D. R. , Burnham, K. P. , and D. S. Robson. 1978. Statistical inference from band recovery data: a handbook. 2nd Ed. U.S. Fish and Wildlife Service Resource Publication 131. Washington, D.C.

Brownie, C. , Anderson, D. R. , Burnham, K. P. , and D. S. Robson. 1985. Statistical inference from band recovery data: a handbook. U.S. Fish and Wildlife Service Resource Publication 156. Washington, D.C.

Burnham, K. P. and D. R. Anderson. 2002. Model Selection and Mutimodel inference. A practical Information-Theoretic Approach. 2nd Ed. Springer Science and Business Media, LLC. USA.

Curtis, J.M.R. 2006. Visible Implant Elastomer Color Determination, Tag Visibility, and Tag Loss: Potential Sources of Error for Mark–Recapture Studies. North American Journal of Fisheries Management. 26: 327-337.

Dicken, M. L. , Booth, A. K. , and M. J. Smale. 2006. Preliminary observations of tag shedding, tag reporting, tag wounds, and tag biofouling for raggedtooth sharks (Carcharias taurus) tagged off the east coast of South Africa. ICES Journal of Marine Science. 63: 1640-1648.

Hoenig, J. M. , Barrowman, N. J. , Hearn, W. S. , and K. H. Pollock. 1998a. Multiyear Tagging Studies Incorporating Fishing Effort Data. Canadian Journal of Fisheries and Aquatic Sciences. 55:1466-1476.

Hoenig, J. M. , Barrowman, N. J. , Pollock, K. H. , Brooks, E. N. , Hearn, W. S. and T. Polacheck. 1998b. Models for Tagging Data That Allow for Incomplete Mixing

of Newly Tagged Animals. Canadian Journal of Fisheries and Aquatic Sciences. 55:1477-1483.

Hogg, R. V., McKean, J. W. , and A. T. Craig. 2005. Introduction to Mathematical Statistics. 6th Ed. Pearson Prentice Hall. Upper Saddle River, New Jersey.

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

Lowry, M. B. and I. M. Suthers. 1998. Home range, activity and distribution patterns of a temperate rocky-reef fish, Cheilodactylus fuscus. Marine Biology. 132(4):569-578.

Pollock, K. H. , Hoenig, J. M. , and C. M. Jones. 1991. Estimation of fishing and natural mortality when a tagging study is combined with a creel survey or port sampling. American Fisheries Society Symposium. 12: 423.434.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available: http://www.R-project.org. (July 2010).

Reeves, K. S. and D. L. Buckmeier. 2009. Mortality, Predation, and Tag Visibility of Fish Marked with Visible Implant Elastomer Tags. North American Journal of Fisheries Management. 29: 323-329.

Ricker, W. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin of the fisheries research board of Canada. 191. Department of environment fisheries and marine service, Ottawa, Canada.

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

Seber, G. A. F. 2002. The Estimation of Animal Abundance and related parameters. 2nd Ed. The Blackburn Press. Caldwell, NJ.

Tarbath, D. B. 1999. Estimates of growth and natural mortality of the blacklip abalone (Haliotis rubra) in Tasmania. Marine Research Laboratories, Tasmanian

Verweij, M. C. and I. Nagelkerken. 2007. Short and long term movement and site fidelity of juvenile Haemulidae in back-reef habitats of a Caribbean embayment. Hydrobiologia. 592(1):257-270.

Wackerly, D. D. , Mendenhall, W. , R. L. Scheaffer. 2002. Mathematical Statistics with Applications. 6th Ed. Duxbuy of Thomson Learning. Pacific Grove, CA.

White, G. C. 1983. Numerical Estimation of Survival Rates from Band-Recovery and Biotelemetry Data. The Journal of Wildlife Management. 47(3):716-728.

White, G. C. 1992. PC SURVIV User's Manual. Dept. of Fishery and Wildlife Biology, Colorado State University, Fort Collins, CO.

Youngs, W. D. and D. S. Robson. 1975. Estimating survival rates from tag returns: model test and sample size determination. Journal of the Fisheries Research Board of Canada. 32: 2365-2371.

Figure 1. Tagged blacklip abalone (*Haliotis rubra*) from Tasmania (photo provided by David Tarbath, Tasmanian Aquaculture and Fisheries Institute, University of Tasmania). The arrow highlights tag position.
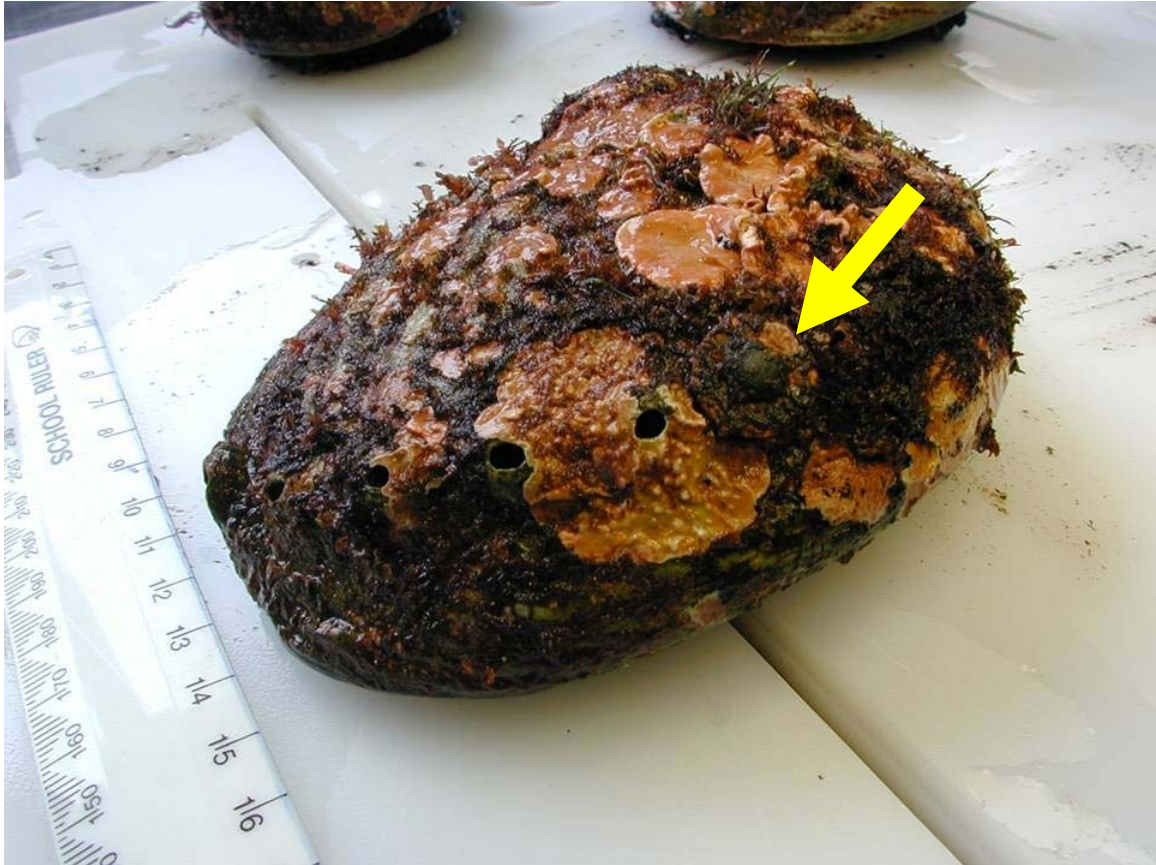
Chapter 2

Tagging Models for Estimating Survival Rates when Tag Visibility Changes Over Time:

partial year tabulation of recaptures

Abstract

The "Brownie" tagging models are commonly used for estimating survival rates from multiyear tagging studies. The basic model, model 1, assumes that all tags have the same tag recovery rate. An alternative, model 0, allows for newly tagged animals to have a different tag recovery rate than previously tagged animals. This might be necessary because new tags are less fouled and more visible than previously applied tags and thus the new tags have a higher tag reporting rate. Model 0 accommodates this problem through the use of an additional parameter which leads to less precision (larger standard errors) than model 1. Model 0', a new model, also allows for newly tagged animals to have a different tag recovery rate than previously tagged ones. It makes use of a known fouling time (or change in visibility time), the time it takes for newly tagged animals to have the same visibility as previously tagged animals, to divide the year into two parts. During the first part of the year, newly tagged animals are more visible than previously tagged ones while in the second part all tagged animals have the same visibility. Dividing the year into parts and recording recaptures in each part avoids the failure of the assumption that the reporting rate is constant for all tagged animals, achieves greater precision (smaller standard errors), and provides estimates of the survival rate at the end of the 2nd year instead of after the 3rd year (as in model 0). The superiority of model 0' over models 0 and 1 is demonstrated for a number of important cases using Monte Carlo simulation.

Introduction

   In multi-year tagging studies, a sample of the population, termed a cohort, is captured, tagged, and released at the start of each of several years. Brownie et al. (1978, 1985) described a suite of models that enables the user to estimate age- and year- specific survival rates from tag recoveries which are tabulated by year. Annual survival rate, represented by *S*, is defined as the fraction of the population alive at the start of the year that is still alive at the end of the year. Additionally, Brownie models enable one to get estimates of the fraction of tagged animals that are caught and reported, termed the tag-recovery rate and denoted by *f*.

   The basic age-invariant and year-specific model is known as model 1. Alternative Brownie models enable the user to impose year-specific constraints on the parameters *f* and *S* and to allow for newly tagged animals to have a different tag-recovery rate than previously tagged animals.

   The assumptions of Brownie models are well documented (e.g., Brownie et al. 1978, 1985; Pollock and Raveling1982) and for model 1 include:

1. The tagged sample is representative of the target population;

2. There is no long-term tag loss;

3. The long-term survival is not affected by tagging or handling processes;

4. The fate of each tagged animal is independent of other tagged animals (no pseudo-replication);

5. All animals within a tagged cohort experience the same *S* and *f* within a time period; this is known as homogeneous survival and tag-recovery rates;

6. The tag-recovery rate*, f,* does not vary among cohorts within a given year.

Various models have been proposed which allow for violations of model 1's assumptions. For example, Hoenig et al. (1998b) described a model for delayed mixing of newly tagged animals – a violation of assumption (1) and thus assumption (5). Brownie et al. (1978) described a model, model 0, which allows for the violation of assumption (6), that is, when a newly tagged cohort has a different tag-recovery rate, f, than previously tagged cohorts.

Tagged cohorts might not be subject to the same tag-recovery rate, f, during a time period for a variety of reasons: if tag-reporting varies among cohorts because tags become unreadable or are shed after some years, or, if exploitation varies among cohorts within one time period (this also affects survival rates, see Hoenig et al. 1998b).

The tag-recovery rate, f, can be expressed as a product of its components (Pollock et al. 1991; Hoenig et al. 1998a):

(1) $f = \phi \, \lambda \, u$,

where ϕ is a composite factor representing the effective number of tags released, here defined as a combination of short term survival rate from tagging and short term probability of tag retention; λ is the tag-reporting rate, or the probability that a tag will be reported if the fish is recaptured; and u is the exploitation rate, or the expected fraction of the population alive at the start of the year that dies due to harvesting during the year. The tag-reporting rate, λ, can further be thought of as a combination of a variety of factors which affect the probability a tag is reported if the animal is recaptured, including the visibility of the tag.

If the visibility of the tag is constant over time and does not vary within cohorts, then visibility is of little interest except inasmuch as low visibility may cause a low rate

of tag returns. If tag visibility changes with time-at-liberty, this will affect the tag-reporting rate differently among cohorts, and thus the tag-recovery rate, which introduces bias into parameter estimates. One way in which tag visibility could change over time is through tag fouling. In a variety of tagging studies, tag fouling has been reported to be prominent (Dicken et al. 2006; Lowry and Suthers 1998; Tarbath 1999; Verweij and Nagelkerken 2007). The use of antifouling materials to prevent fouling on tags may alleviate the problem of reduced tag visibility but such materials could introduce additional problems. Antifouling materials, which could be used to coat the tags, are costly and may potentially harm the tagged animal. Other studies have reported issues with tag visibility over time. Tagging programs using visible implant elastomer tags have reported diminishing tag visibility over time as a result of thickening of the skin overlying the tags (Curtis 2006; Reeves and Buckmeier 2009). The problem of tag visibility varying with time-at-liberty may be more prevalent than discussed in the literature for a variety of reasons including poor communication between fishers and scientists, and researchers not knowing how to incorporate the change of tag visibility into the models and choosing simply to ignore it, or not understanding the bias it may introduce in the parameter estimates.

Brownie et al. (1978, 1985) introduced a model, model 0, that can deal with the tag-recovery rate of new tags being different from older tags for the entire first year each cohort is at liberty (note this is equivalent to the tag visibility being different for the entire first year a cohort is at liberty). Under this model, all cohorts at liberty for more than one year have the same tag reporting rate (within a given year), and thus the same tag visibility. If it takes new tags less than one year to have the same visibility as previously

tagged animals, better use can be made of the data by partitioning the year into parts.

The tag returns can be tabulated by portions of the year, which coincide with this change

in visibility. I present such a model, model 0', and study its properties by Monte Carlo

simulation. Model 0' is compared with model 1 and model 0 in order to provide

guidance as to which model(s) should be applied based on the availability of information

on tag fouling. Additionally, the importance of the degree to which tag visibility (and

thus tag-reporting) affects model performance is evaluated in order to highlight the

benefit of including fouling in the model when it occurs.

Brownie Models

Model 1

The age-invariant and year-specific model described by Youngs and Robson

(1975) and Brownie et al. (1978, 1985) is known as model 1. The data consist of an

upper triangular or trapezoidal array made up of observed r_{ij}'s, which are the realizations

of R_{ij}'s, the random variable representing the number of animals tagged in year i (i = 1, 2

,…, I) and recaptured in year j, j = i, ..., J with $J \geq I$. The expected recaptures of

animals tagged in year i and recaptured in year j is

$$(2) \; E(R_{ij}) = \begin{cases} N_i f_j & , i = j \\ N_i f_j \left(\displaystyle\prod_{k=i}^{j-1} S_k \right) & , i < j \leq J \end{cases}$$

where the expression $E(\cdot)$ denotes the expected value of the variable within the

parentheses, R_{ij} is as defined above, N_i is the number tagged and released at the start of

year i, i = 1, ..., I, S_j is the fraction of the population alive at the start of year j that is

still alive at the end of year j, for j = 1, ..., J-1, and f_j is the tag-recovery rate in year j,

for $j = 1, \ldots, J$. Note that there is an implicit category for all animals of a cohort that are never seen again, denoted Y_i. This can be expressed as:

$$(3) \quad Y_i = N_i - \sum_{k=i}^{J} r_{ik} .$$

Note that $J \geq I$, and that if $J = I$ the data form a triangular array.

Maximum likelihood estimation (MLE) is frequently used to obtain estimates for unknown model parameters based on the observed data. MLE methods consist of writing an expression, in terms of the unknown parameters, for the probability of having obtained the observed data and this is called the likelihood function. Then, the parameter values which maximize the value of the likelihood function are found, and these parameter values are the maximum likelihood estimates. Maximum likelihood estimators have the following desirable asymptotic properties (under regularity conditions): unbiased, unique, consistent, minimum variance, and normally distributed (e.g., Hogg et al. 2005; Wackerly et al. 2002).

Brownie type models can be expressed as the product of independent multinomial distributions of tag returns over time, with each tagged cohort giving rise to a multinomial distribution. The general form of the likelihood function Λ for product multinomial models can be expressed as

$$(4) \quad \Lambda \propto \prod_{i=1}^{I} \left(\prod_{j=i}^{J} P_{ij}^{r_{ij}} \right) \left(1 - \sum_{j=i}^{J} P_{ij} \right)^{Y_i} ,$$

where the symbol $\propto$ means "is proportional to," P_{ij} is the cell probability of recovering a tagged animal in year j given that it was tagged in year i, that is,

$$(5) \quad P_{ij} = {E(R_{ij})} \big/ {N_i} \,,$$

and the other symbols are as defined previously. For model 1, the cell probabilities, P_{ij}, are found by substituting equation (2) into equation (5). The likelihood function is proportional to the product of the cell probabilities raised to the power of the number of observed recaptures. The right hand portion of the equation involving one minus a sum is the expression for the animals in a cohort that are never seen again after tagging.

Model 0

Brownie et al. (1978, 1985) described a generalized version of model 1 known as model 0 (Table 1). It pertains to the case where the tag-recovery rate is different for the first year a cohort is at liberty compared with previously tagged cohorts within a given time period. The model incorporates an additional parameter f_j* for each cohort, which is the tag-recovery rate in year j, for newly tagged animals, with the * indicating the fraction reported is for a cohort in its first year at liberty during which the tag visibility is greater than the tag visibility of previously released cohorts. Allowing the tag-recovery rates to be different for the first year a cohort is at liberty leads to more parameters than model 1, which leads to less precision than the use of model 1. However, the unequal recovery rates, $f*$ and f, for newly tagged and previously tagged cohorts, respectively, means model 0 affords protection from bias due to model misspecification. Note that model 1 is a special case of model 0.

In order to estimate the survival rate in the first year, three years of tag returns are needed, instead of just two years of tag returns needed for model 1. If recoveries are made for J years, then for the cohort tagged in year i, there will be J-i-1 moment

estimates of S_i. For example for the cohort tagged in year one, an estimate of S_1 is not

possible in the second year (there will be three equations and four unknowns). This can

be seen by looking at the moment estimate formed by the ratio of expectations (using R_{12}

and R_{22} compared to R_{13} and R_{23})

$$(6a)\ \frac{E(R_{12})}{E(R_{22})} = \frac{S_1 f_2}{f_2{}^*},$$

$$(6b)\ \frac{E(R_{13})}{E(R_{23})} = \frac{S_1 S_2 f_3}{S_2 f_3} = S_1.$$

Thus, for an example with three years of tagging and four years of recoveries, there will

be $J\text{-}2 = 2$ estimates of S_1, survival during the first year formed by the ratio in (6b) and by

the ratio of expectations of R_{14}/R_{24}.

In practice, maximum likelihood estimates are found by maximizing equation (4)

with appropriate cell probabilities obtained by substituting values in Table 1 for the

values in equation (5).

New Model: Model 0'

The year can be divided into two parts; in the first part, part (a), newly tagged

cohorts have new and highly visible tags, and in the second part, part (b), the tags are

fouled and have the same visibility as fouled tags from animals released in previous

years. The tag returns can be tabulated separately for parts (a) and (b) of the year (Table

2).

For previously tagged cohorts there is a tag-recovery rate parameter for each part

of the year, i.e., f_{ja} and f_{jb}, for parts (a) and (b), respectively of year j. Similar to model 0,

model 0' has an additional parameter in the form of an $f_{ia}{}^*$ for the first part of the first

year that cohort i is at liberty. In the second part of the first year a cohort is at liberty, the tag-recovery rate is the same as for previously tagged cohorts, i.e., there is no * on the recovery parameter (Table 3). Since the year is divided into two parts, an estimate of survival can be made at the end of the second year as seen with the ratio of expectations

$$(7) \quad \frac{E(R_{12b})}{E(R_{22b})} = \frac{S_1 S_{2a} f_{2b}}{S_{2a} f_{2b}} = S_1.$$

This is in contrast to the situation for model 0 where an estimate can only be obtained after the end of the third year (equation 6b).

The likelihood is constructed as before using the cell probabilities from Table 3. The recapture cell representing the tagged animals which are never seen again can be given by

$$(8) \quad W_i = N_i - \sum_{j=i}^{J} \sum_{k \in \{a,b\}} r_{ijk}$$

where r_{ijk} are the observed number of recaptures of animals tagged in year i ($i=1,2,\ldots,I$) and recaptured in part k, $k \in \{a,b\}$, of year j, $j = i$, ..., J with $J \geq I$. Thus the likelihood can be expressed

$$(9) \quad \Lambda \propto \prod_{i=1}^{I} \left(\prod_{j=i}^{J} \prod_{k \in \{a,b\}} P_{ijk}^{r_{ijk}} \right) \left(1 - \sum_{j=i}^{J} \sum_{k \in \{a,b\}} P_{ijk} \right)^{W_i},$$

where P_{ijk} is the cell probability of recovering a tagged animal in part k of year j given that it was tagged in year i, that is, $P_{ijk} = \dfrac{E(R_{ijk})}{N_i}$ and the other parameters are as defined previously. Estimates for the parameters can be found by maximizing equation (9).

Model evaluation by simulation

To evaluate the performance of model 0', I used Monte Carlo simulation to generate data repeatedly under models 1 (no fouling), 0 (fouling that takes a year), and 0' (fouling that takes only part of the year) and then fit all three models to each dataset. The simulations consisted of three years of tagging and four full years of recaptures, that is, recaptures for periods 1a, 1b, 2a, 2b, 3a, 3b, 4a, and 4b for model 0' and periods 1, 2, 3, and 4 for model 1 and model 0. Ten thousand datasets were simulated for each scenario. Computations were done using the statistical language R (R Development Core Team 2008) as described below. Parallel computations were done for some scenarios using program SURVIV (White 1983, 1992) to check for program errors and numerical problems. Program SURVIV is available on the internet (http://www.mbr-pwrc.usgs.gov/software/ survive.html).

The function 'rmultinomial' (Chasalow 2005) in the R package 'combinat' was used to generate multinomial datasets with specified sample sizes and cell probabilities. The function 'nlm' was used to minimize the negative log likelihood functions (R Development Core Team 2008). Standard errors were estimated by inverting the Hessian matrix using the R function 'solve'. The true standard error was determined from the variability of the 10,000 estimates of each parameter. The output from the Monte Carlo simulations includes: estimates of the parameters for each simulated dataset as well as the bias, % bias of the average estimate (referred to as % bias), standard errors, and bias of the estimated standard errors. Additionally, the root mean square error (RMSE) for each parameter was calculated using

31

$$(10)\ RMSE = \sqrt{bias^2 + variance} = \sqrt{\frac{\sum\limits_{i=1}^{T}\left(\hat{\theta}_i - \theta\right)^2}{T}}$$

(see Hogg et al. 2005) where T is the number of simulated datasets (10,000) and $\hat{\theta}_i$ is the ith estimate of the parameter whose true value is θ.

If the combined short-term survival rate from tagging and short-term probability of tag retention, ϕ, is assumed to be 1, then equation (1) becomes

$$(11)\ f = \lambda\, u.$$

With the year split in two parts, there is a tag-recovery parameter for each part of the year, given by f_a and f_b for model 1 and $f_a{*}$ and f_b for model 0'. These tag-recovery parameters can be modeled as:

$$(12)\ f_a = \lambda_f\, u_a,$$

$$(13)\ f_a{*} = \lambda_c\, u_a,$$

and

$$(14)\ f_b = \lambda_f\, u_b$$

where λ_f is the tag-reporting rate when the tags are fouled, when the tags are clean and becoming fouled, u_a is the exploitation rate during part (a) of the year, λ_c is the tag-reporting rate during part (a) of the first year each cohort is at liberty, and u_b is the exploitation rate during part (b) of the year. Additionally, the exploitation rates are constrained by

$$(15)\ u_a + u_b < 1 - S.$$

By varying the values of f_a and f_b, a multitude of situations are simulated. Since f_b is a product of λ_f and u_b, a high u_b may correspond to a variety of situations (Figure 1). For example, a high exploitation rate in part (b) of the year, u_b, is possible when the tag-recovery rate, f_b, is high. Another situation that leads to a high u_b is when the tag-recovery rate is low and the tag-reporting rate is also low. That is, if u_b is high, and λ_f is also high, f_b would be high, but f_b would be low if λ_f is low.

Data generated under model 0': base scenario

The parameters for the first set of simulations were loosely patterned after data from the queen conch (*Strombus gigas* L.) fishery of the Turks and Caicos Islands, British West Indies. The queen conch fishery was chosen because tag fouling is known to be a problem and the possibility of a tagging program was being explored. The exploitation rate, $u = 0.2$, was based on the ratio of the annual harvest plus local consumption to the estimated biomass at the start of the year, as determined from a surplus production model (Kathy Lockhart, The Department of Environment and Coastal Resources, Turks and Caicos Islands, British West Indies, personal communication). For these simulations it was assumed that tagging would occur during the summer when the conch fishery is closed. The tags were assumed to take six months to foul completely, thus the year was split in half, with part (a) running from July to December, and part (b) running from January to June. The percentage of the fishing effort occurring in part (a) was set at fifty percent and fifty percent in (b) to reflect the seasonal distribution of the harvest.

The survival rates were calculated from the exploitation rate, $u = 0.2$, and a natural mortality rate of $M = 0.3$ (SEDAR 2007), using Baranov's catch equation which relates the exploitation rate, u, to the components of instantaneous mortality rates, F and M:

$$(16)\ u = \frac{F}{F + M}\left(1 - e^{-(F+M)}\right)\ ,$$

and

$$(17)\ S = e^{-(F+M)}.$$

The survival rate in the first year, $S_1 = 0.57$, comes from solving for F given M using equation (15) and then calculating S from (16). The survival rate in the second year, $S_2 = 0.62$, was altered from the survival rate in the first year in order to allow the years to have different survival rates.

A preliminary tag fouling study was conducted in the Turks and Caicos Islands on queen conch. The conch were tagged with custom-made tags (Hallprint custom code T6230) and vinyl tubing (spaghetti) tags (Floytag) secured around the spires of the conch. Fouling was observed on both tag types. Conch fishermen typically free-dive in less than 10m of water working off small boats with 50-65hp engines and the conch are collected by hand (see Medley and Ninnes 1999). The meats are removed from the shell by the boat driver while the divers continue collecting conch; the shell is knocked (hammered) in the same location as the tag placement on the shell. Based on the nature of the fishery and the proposed tag reward, US$5 for the return of a tag to a fish processing plant, the tag-reporting rate for newly tagged animals, λ_c, was set at 1.0. Information on the tag-reporting rate for fouled animals, λ_f, was unavailable so the initial value of 0.5 was used for the first example and then varied in later examples.

In order to compare model 0' to model 1 and model 0 it was necessary to generate data for two parts of each year, part (a) and (b), and then combine the data from parts of the year into full years to analyze datasets with model 1 and model 0. All parameters were initially defined for parts of the year, and then collapsed for use in other models (Table 4), using the relationships:

(18) $f_1 = f_{1a} + S_{1a} f_{1b}$,

(19) $f_2 = f_{2a} + S_{2a} f_{2b}$,

(20) $f_3 = f_{3a} + S_{3a} f_{3b}$,

and

(21) $S_3 f_4 = S_3 f_{4a} + S_3 S_{4a} f_{4b}$.

Note that some parameters are confounded and cannot be estimated; rather, the products are estimated (see Table 4).

The various models' performance are described in terms of the bias, root mean squared error, true standard error, and mean estimated standard error for the survival rates during the first and second year and the recovery rates in the second and third year. These four parameters were chosen for comparison because they are the only parameters in common when data are analyzed under all three models (the fifth parameter in common is $S_3 f_4$, a confounded parameter).

In order to determine what effect the tag-reporting rate for fouled animals, λ_f, has on model performance, the effect of tag fouling on tag visibility was varied in additional scenarios. In all cases, the tag-reporting rate for clean, newly tagged animals, λ_c, was kept at 1.0. Simulations were run using values for λ_f of 0.25, 0.5 (base scenario), and 0.75.

Data generated under model 1

The purpose of this scenario is to evaluate the penalty for applying model 0' when model 1 is correct. Data were generated using parameters based on the queen conch fishery in the Turks and Caicos Islands (Table 4), assuming fouling was not a problem; that is, the tag-reporting rate for newly tagged animals was equal to that of previously tagged animals.

Data generated under model 0

Under model 0, it takes a full year for newly tagged animals to have the same tag visibility (and thus reporting rate) as previously tagged animals. For this simulation the parameters were again based on information from the queen conch fishery of the Turks and Caicos Islands (see description of data from model 0' and Table 4). The tag-reporting rate for clean, newly tagged animals, λ_c, remained 1.0, and the tag-reporting rate for fouled animals, λ_f, was 0.5. In other words, the tag-recovery rate of fouled animals was equal to half the value of the tag recovery rate of clean animals, $0.5(f_j) = f_j^*$.

Additional simulations were conducted to evaluate the models' performance when the effect of fouling on tag visibility, and thus tag-reporting rate, varies. Simulations were run with the tag-reporting rate for fouled tags, λ_f, set at 0.25, 0.50 (base case) and 0.75.

Simulation Results

Data generated under model 0': base scenario

As expected, using data generated when tag fouling affects tag-reporting and takes less than a year to fully foul (the case of model 0' being appropriate), parameter estimates for S_1, S_2, f_2, and f_3 were essentially unbiased whether estimated by model 0 (all parameters with % bias < 2.50%) or model 0' (all parameters with % bias < 1%) (Table 5). Analysis with model 1 produced biased estimates for all four parameters: survival rates are underestimated (% bias for S_1 = -18% and % bias for S_2 = -26%) and the fraction caught and reported in years two and three, f_2 and f_3, are overestimated (% bias for f_2 = 37% and % bias for f_3 = 61%). The estimates from analysis with model 0' have the lowest standard error and root mean squared error. Model 1 has the lowest estimated standard error for all four parameters but is not an attractive estimator because of the high bias and thus high RMSE.

When tag fouling has a large effect on tag visibility, making the reporting rate of previously tagged animals much lower than that of newly tagged animals, model 0' is the most appropriate model (Table 6). As the effect of change in visibility on reporting rate decreases, model 1 begins to yield smaller root mean squared errors than model 0' for S_1 and f_2 (Figure 2); this is because the estimates from model 1 are becoming closer to being essentially unbiased (Table 6, λ_f = 0.75, % bias for S_1 is -5%, % bias for S_2 is -13%, % bias for f_2 is 9%, and % bias for f_3 is 23%) and have better precision than model 0' as a result of having less parameters.

Data generated under model 1

Using data generated with tag fouling having no effect on the tag-reporting rate (model 1 is appropriate) yields essentially unbiased estimates of S_1, S_2, f_2, and f_3 under all three models (Table 7). Model 1 produces estimates with the smallest standard error of the estimate and root mean squared error. Model 0 produces estimates with the largest RMSEs. Thus model 0 performs the worst.

Data generated under model 0

When data are generated under model 0, where tag fouling affects tag-reporting rate and it takes one full year for the tags to become fouled, and the tag-reporting rate for fouled animals is half that of clean animals, only model 0 produces essentially unbiased results for estimates of S_1, S_2, f_2, and f_3 (Table 8). Model 0 has the lowest RMSE for every parameter. Model 0' outperforms model 1 by producing estimates with both a smaller bias and RMSE.

Since the value of tag-reporting rate for clean tags, λ_c, was set at 1.0, discussion of the value of tag-reporting rate for fouled tags, λ_f, or the ratio of fouled tags to clean tags represents the same idea. As expected when the reporting rate for fouled tags is 0.25 model 0 produces estimates of S_1, S_2, f_2, and f_3 with the lowest biases and RMSE. Model 0' outperforms model 1 because model 0' results in estimates of S_1, S_2, f_2, and f_3 with lower RMSEs and lower biases than the estimates from model 1 (Table 9). For estimating S_1 and S_2, the most appropriate model changes from model 0 to model 0' as the ratio of reporting rates (fouled:clean) gets large (Figure 3). Note that model 0 produces estimates of f_2 and f_3 with the lowest RMSEs.

38

Discussion

There are three important factors to consider when choosing between models 1, 0, and 0'. First, the time period that must elapse for new tags to have the same tag-reporting rate as older tags, that is, the length of part (a), must be bounded, i.e., known to be less than some specified period of time. When part (a) is neither very short nor close to a full year, I recommend model 0'. As part (a) of the year becomes shorter in duration, I recommend using model 1 and as part (a) becomes longer (close to 1 year) then I recommend applying model 0. If part (a) takes longer than one year, then a new model should be parameterized to account for this.

Second, the timing of the fishery relative to the timing of the fouling (which determines part (a) and part (b) of the year) must be known. When fishing occurs in both parts (a) and (b) of the year, model 0' should remain appropriate (given fouling lasting less than a year). However, if all the fishing effort takes place in period (b), then the recaptures for all (a) periods will be zero and model 1 will be a more appropriate model. If all the fishing effort takes place in period (a), then model 0 should be the most appropriate model.

Finally, the magnitude of the change in visibility affects which model is the most appropriate. If the change in visibility is extremely small, that is tag visibility does not greatly vary between cohorts in a given part of a year, then the proportion of recaptured individuals reported in period (a) should be equal to the proportion of recaptured individuals reported in period (b). Thus, when the change in visibility is approximately zero, model 1 is more appropriate than model 0' or model 0.

If tag fouling is known to occur in much less than a year and affects the reporting rate greatly, then model 0' will outperform model 1 (Figure 2), and model 0' and model 0 will both provide unbiased estimates. Even when it takes a full year for the tags to foul model 0' can outperform model 0 (in terms of lower RMSE) as the influence of tag fouling on tag-reporting rate becomes smaller, making λ_f closer in value to λ_c (Figure 3). Furthermore, when fouling takes less than a year model 0' can outperform model 1 and model 0.

As the change in visibility due to tag fouling becomes smaller, model 1 produces smaller root mean squared errors when tag fouling takes a year (model 0) and when tag fouling takes less than a year (model 0'). This emphasizes the importance of considering the magnitude of change in visibility when selecting a model. As the ratio of the tag-reporting rate for fouled tags to new tags becomes closer to one, model 1 becomes the more appropriate model (Figures 2 and 3).

If fouling is known to affect reporting rate, but the time necessary for a tag to become fouled has not been determined, then model 0' and model 0 are valid candidates since they provide unbiased estimates when tag fouling takes less than a year. If possible, a study should be done to determine the time to tag fouling. Such a study may be inexpensive and require only modest effort, and may provide key information in choosing between model 0 and 0'.

The problem of changing tag visibility is similar to that of tag loss. Model 0' is parameterized such that the change in tag visibility takes an appreciable amount of time, which is less than a year, and then the tag visibility remains constant over time. Thus, one can think of two time periods: when the visibility is constant over time and when it is

not. In contrast, tag loss is of two types. Type I, or short-term, which occurs so rapidly

that it occurs before fishing begins (Beverton and Holt 1957). Short-term tag loss

essentially modifies the effective number tagged. Type II, or long-term, tag loss is

similar to changing visibility over time except that it is usually described as occurring

progressively, rather than leveling off. If it can be assumed that the rate of tag loss

declines to zero in less than a year, then model 0' would be appropriate. However, most

of the literature supports the idea of progressive tag loss.

Tabulating the recaptures by part of a year rather than a full year should not be a

problem. It has been shown by Pollock and Raveling (1982) that when conducting a

tagging study it is important to determine the date of tag recapture since even for

Brownie model 1 it is important to know the year of recapture for each tag return to avoid

biased estimates. Thus knowing whether the tag return is from part (a) or part (b) of the

year should add little or no additional cost to the study.

For cases where tag fouling occurs, takes less than a year, and causes a change in

the tag-reporting rate, model 0' can greatly improve the efficiency of the tagging study

(in terms of smaller standard errors). Furthermore, model 0' provides the first estimate of

survival during the first year, S_1, at the end of the second year which is a full year before

model 0 provides an estimate.

This is the first work to demonstrate the value of tabulating tag returns with a

greater periodicity than the periodicity of tagging (e.g., tabulating by parts of the year

when tagging occurs annually). Another example of the value of tabulating recaptures

this way is given in Waterhouse (2010, chapter three) where partial year tabulation is

used for dealing with delayed mixing of newly tagged animals with the population at

large. It is yet to be seen what additional applications can be found for tabulating tag
returns with greater periodicity than that of tagging, and what generalizations can be
made from the model.

Acknowledgements

References

Beverton, J. H. and S. J. Holt. 1957. On the dynamics of exploited fish populations.
Chapman and Hall. London, Great Britain.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson. 1978. Statistical
inference from band recovery data: a handbook. 2nd Ed. U.S. Fish and Wildlife
Service Resource Publication 131. Washington, D.C.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson. 1985. Statistical
inference from band recovery data: a handbook. U.S. Fish and Wildlife Service
Resource Publication 156. Washington, D.C.

Chasalow, S. 2005. combinat: combinatorics utilities. R package version 0.0-6.
Available: http://cran.r-project.org/web/packages/combinat/index.html. (July
2010).

Curtis, J.M.R. 2006. Visible Implant Elastomer Color Determination, Tag Visibility,
and Tag Loss: Potential Sources of Error for Mark–Recapture Studies. North
American Journal of Fisheries Management. 26: 327-337.

Dicken, M.L., A.K. Booth, and M.J. Smale. 2006. Preliminary observations of tag
shedding, tag reporting, tag wounds, and tag biofouling for raggedtooth sharks
(*Carcharias taurus*) tagged off the east coast of South Africa. ICES Journal of
Marine Science. 63: 1640-1648.

Hoenig, J.M., N.J. Barrowman, W.S. Hearn, and K.H. Pollock. 1998a. Multiyear
Tagging Studies Incorporating Fishing Effort Data. Canadian Journal of
Fisheries and Aquatic Sciences. 55:1466-1476.

Hoenig, J. M. , Barrowman, N. J. , Pollock, K. H. , Brooks, E. N. , Hearn, W. S. , and T. Polacheck. 1998b. Models for Tagging Data That Allow for Incomplete Mixing of Newly Tagged Animals. Canadian Journal of Fisheries and Aquatic Sciences. 55:1477-1483.

Hogg, R. V. , McKean, J. W. , and A. T. Craig. 2005. Introduction to Mathematical Statistics. 6th Ed. Pearson Prentice Hall. Upper Saddle River, New Jersey.

Lowry, M. B. and I. M. Suthers. 1998. Home range, activity and distribution patterns of a temperate rocky-reef fish, *Cheilodactylus fuscus*. Marine Biology. 132(4): 569-578.

Medley, P. A. H. and C. H. Ninnes. 1999. A stock assessment for the conch (*Strombus gigas* L.) fishery in the Turks and Caicos Islands. Bulletin of Marine Science. 64(3): 399-406.

Pollock, K. H. , Hoenig, J. M. , and C. M. Jones. 1991. Estimation of Fishing and Natural Mortality Rates when a Tagging Study is Combined with a Creel Survey or Port Sampling. American Fisheries Society Symposium 12:423-434.

Pollock, K. H. and D. G. Raveling. 1982. Assumptions of Modern Band-Recovery Models, with Emphasis on Heterogeneous Survival Rates. The Journal of Wildlife Management. 46(1):88-98.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available: http://www.R-project.org. (July 2010).

Reeves, K. S. and D. L. Buckmeier. 2009. Mortality, Predation, and Tag Visibility of Fish Marked with Visible Implant Elastomer Tags. North American Journal of Fisheries Management. 29: 323-329.

SEDAR (Southeast Data, Assessment, and Review). 2007. Caribbean Queen Conch. SEDAR 14, Stock Assessment Report 3. North Charleston, South Carolina.

Tarbath, D. B. 1999. Estimates of growth and natural mortality of the blacklip abalone (*Haliotis rubra*) in Tasmania. Marine Research Laboratories, Tasmanian Aquaculture and Fisheries Institute. Technical Report No. 3, University of Tasmania, Tasmania. Available: http://www.tafi.org.au/publications/TechnicalReport_3.pdf. (July 2010).

Verweij, M. C. and I. Nagelkerken. 2007. Short and long term movement and site fidelity of juvenile Haemulidae in back-reef habitats of a Caribbean embayment. Hydrobiologia. 592(1):257-270.

Wackerly, D. D. , Mendenhall, W. , and R. L. Scheaffer. 2002. Mathematical Statistics with Applications. 6th Ed. Duxbuy of Thomson Learning. Pacific Grove, CA.

Waterhouse, L.W. 2010. Partial Year Tagging Models: accounting for tag fouling and delayed mixing, Master's thesis. Virginia Institute of Marine Science, The College of William and Mary.

White, G. C. 1983. Numerical Estimation of Survival Rates from Band-Recovery and Biotelemetry Data. The Journal of Wildlife Management. 47(3):716-728.

White, G. C. 1992. PC SURVIV User's Manual. Dept. of Fishery and Wildlife Biology, Colorado State University, Fort Collins, CO. Available: http://www.mbr-pwrc.usgs.gov/software/survive.html. (July 2010).

Table 1. Expected recaptures under Brownie model 0, where N_i is the number tagged and released at the start of year i, $i = 1, \ldots, 3$; S_j is the fraction of the population alive at the start of year j that is still alive at the end of year j, for $j = 1, \ldots, J\text{-}1$ (J=4); f_j is the expected fraction of the tagged population (at large at the start of year j) that is caught and reported during year j, for $j = 1, \ldots, 4$; and f_j* is the expected fraction of the newly tagged animals (that is tagged in year j) that is caught and reported in year j, with the * indicating the fraction reported is different for the first year a cohort is at liberty (than for previously tagged cohorts) as the tags are still new and unfouled.

		Expected # recaptured in year			
Year	**# tagged**	**1**	**2**	**3**	**4**
1	N_1	$N_1 f_1$*	$N_1 S_1 f_2$	$N_1 S_1 S_2 f_3$	$N_1 S_1 S_2 S_3 f_4$
2	N_2	--	$N_2 f_2$*	$N_2 S_2 f_3$	$N_2 S_2 S_3 f_4$
3	N_3	--	--	$N_3 f_3$*	$N_3 S_3 f_4$

Table 2. Observed recaptures under the new model 0'. The year is divided into two parts, (a) and (b). In the first part of the year, part (a), newly tagged animals are more visible than previously tagged animals but in the second part of the year, part (b), the visibility of new and previously tagged animals is equal. The observations consist of counts, r_{ijk}, realizations (i.e., observations) of the random variables R_{ijk}, i.e., the observed number of recaptures of those animals tagged in year i, $i = 1,\ldots,3$, that were recaptured in year j, $j=1,..,4$ and period k, $k \in \{a,b\}$.

		Observed # recaptured in period							
Year	**# Tagged**	**1a**	**1b**	**2a**	**2b**	**3a**	**3b**	**4a**	**4b**
1	N_1	r_{11a}	r_{11b}	r_{12a}	r_{12b}	r_{13a}	r_{13b}	r_{14a}	r_{14b}
2	N_2	--	--	r_{22a}	r_{22b}	r_{23a}	r_{23b}	r_{24a}	r_{24b}
3	N_3	--	--	--	--	r_{33a}	r_{33b}	r_{34a}	r_{34b}

Table 3. Expected recaptures under the new model, model 0'. Where N_i is the number tagged and released at the start of year i, $i = 1,$

....., 3; S_j is the fraction of the population alive at the start of year j that is still alive at the end of year j, for $j = 1, \ldots, J\text{-}1$ ($J = 4$); S_{jk}

is the fraction of the population alive at the start of year j that is still alive at the end of year j, for $j = 1, \ldots, J\text{-}1$ ($J = 4$), and k is the

part of the year, with $k \in \{a,b\}$; f_{jk} is the fraction of the tagged population (at large at the start of year j) that is caught and reported

during year j, for $j = 1, \ldots, 4$, and k is the part of the year, with $k \in \{a,b\}$; and $f_{ja}*$ is the fraction of the newly tagged animals that is

caught and reported in year j, with the * indicating the fraction reported is different for part (a) of the first year a cohort is at liberty.

Expected # recaptured in period

Year	# Tagged	1a	1b	2a	2b	3a	3b	4a	4b
1	N_1	$N_1 f_{1a}*$	$N_1 S_{1a} f_{1b}$	$N_1 S_1 f_{2a}$	$N_1 S_1 S_{2a} f_{2b}$	$N_1 S_1 S_2 f_{3a}$	$N_1 S_1 S_2 S_{3a} f_{3b}$	$N_1 S_1 S_2 S_3 f_{4a}$	$N_1 S_1 S_2 S_3 S_{4a} f_{4b}$
2	N_2	--	--	$N_2 f_{2a}*$	$N_2 S_{2a} f_{2b}$	$N_2 S_2 f_{3a}$	$N_2 S_2 S_{3a} f_{3b}$	$N_2 S_2 S_3 f_{4a}$	$N_2 S_2 S_3 S_{4a} f_{4b}$
3	N_3	--	--	--	--	$N_3 f_{3a}*$	$N_3 S_{3a} f_{3b}$	$N_3 S_3 f_{4a}$	$N_3 S_3 S_{4a} f_{4b}$

Table 4. Parameter values used in the simulations. All three models were parameterized with the year split into two parts, in order to make analysis with model 0' possible, but for models zero and one the parameters shown in the table are collapsed back into their form for a full year, which is as they appear in the actual model parameterizations. Some parameters are confounded with others and therefore cannot be estimated on their own. For example, in model 0 and model 1, survival in the third year and the tag-recovery rate in the fourth year, S_3f_4, are confounded. Parameters that are confounded appear as a product in the table.

<u>**Parameters & Values Used for Data Generation Under:**</u>

Model 1		Model 0'		Model 0	
f_1	0.092	$f_{1a}*$	0.100	f_1*	0.178
S_1	0.570	$S_{1a}f_{1b}$	0.040	S_1	0.57
f_2	0.091	S_1	0.570	f_2	0.091
S_2	0.620	f_{2a}	0.050	S_2	0.62
f_3	0.086	$S_{2a}f_{2b}$	0.041	f_3	0.086
S_3f_4	0.050	S_2	0.620	S_3f_4	0.05
		f_{3a}	0.048	f_2*	0.17
		$S_{3a}f_{3b}$	0.038	f_3*	0.186
		S_3f_{4a}	0.028		
		$S_3S_{4a}f_{4b}$	0.023		
		$f_{2a}*$	0.090		
		$f_{3a}*$	0.11		

49

Table 5. Simulation results for estimating survival rates in the first two years, S_1 and S_2,

and the expected fraction that is caught and reported in years two and three, f_2 and f_3,

under the base scenario when tag fouling occurs and takes half a year (data generated

with model 0'). Fouled tags have visibility that is 50% of new tags. Values for all

parameters appear in Table 4. The smallest values for bias, % bias, standard error of the

estimates, and root mean squared error (RMSE) are in bold. Mean $\hat{SE}$ refers to the mean

of the 10,000 estimated standard errors which come from the square root of the variance,

calculated by solving the inverse of the hessian. SE of estimates refers to the true

standard error (i.e., the standard deviation) of the 10,000 estimates of each parameter.

Model Fitted	Parameter	True Value	Mean estimate	Bias	% Bias	Mean $\hat{SE}$	SE of estimates	RMSE
0'	S_1	0.57	0.575	**5.19E-03**	**0.91%**	0.082	0.082	**0.083**
1	S_1	0.57	0.467	-1.03E-01	-18.15%	0.053	**0.053**	0.116
0	S_1	0.57	0.577	7.01E-03	1.23%	0.102	0.102	0.103
0'	S_2	0.62	0.626	**5.83E-03**	**0.94%**	0.098	0.099	**0.099**
1	S_2	0.62	0.459	-1.61E-01	-26.01%	0.051	**0.050**	0.169
0	S_2	0.62	0.632	1.24E-02	2.00%	0.133	0.135	0.135
0'	f_2	0.091	0.092	**5.97E-04**	**0.66%**	---	0.014	**0.014**
1	f_2	0.091	0.125	3.38E-02	37.19%	0.010	**0.010**	0.035
0	f_2	0.091	0.093	1.94E-03	2.13%	0.021	0.021	0.022
0'	f_3	0.086	0.087	**7.19E-04**	**0.84%**	---	0.013	**0.013**
1	f_3	0.086	0.139	5.28E-02	61.40%	0.010	**0.010**	0.054
0	f_3	0.086	0.088	1.82E-03	2.11%	0.020	0.020	0.020

Table 6. Simulation results for estimating survival rates in the first two years, S_1 and S_2, and the expected fraction caught and reported in years two and three, f_2 and f_3, when data are generated for tag fouling which takes half a year (data generated with model 0'). The smallest values of bias, % bias, standard error of the estimate, and root mean squared error (RMSE) are in bold. Mean $\hat{SE}$ denotes the mean of the 10,000 estimated standard errors which come from the square root of the variance, calculated by solving the inverse of the hessian. SE of estimates refers to true standard error (i.e., standard deviation) of the 10,000 estimates of each parameter. Tag-reporting rate for fouled tags is given by λ_f.

Model Fitted	Parameter	True Value	Mean estimate	Bias	% Bias	Mean $\hat{SE}$	SE of estimates	RMSE
$\lambda_f = 0.25$								
0'	S_1	0.57	0.579	**0.009**	**1.61%**	0.118	0.119	**0.120**
1	S_1	0.57	0.329	-0.241	-42.33%	0.051	**0.052**	0.247
0	S_1	0.57	0.582	0.012	2.11%	0.143	0.147	0.147
0'	S_2	0.62	0.633	**0.013**	**2.17%**	0.140	0.141	**0.142**
1	S_2	0.62	0.343	-0.277	-44.68%	0.049	**0.048**	0.281
0	S_2	0.62	0.644	0.024	3.90%	0.193	0.198	0.199
0'	f_2	0.043	0.044	**0.001**	**1.85%**	---	0.010	**0.010**
1	f_2	0.043	0.101	0.058	134.22%	0.009	**0.010**	0.058
0	f_2	0.043	0.045	0.002	5.05%	0.015	0.015	0.015
0'	f_3	0.047	0.048	**0.001**	**1.08%**	---	**0.010**	**0.010**
1	f_3	0.047	0.122	0.075	158.71%	0.010	0.010	0.075
0	f_3	0.047	0.049	0.002	4.53%	0.016	0.016	0.017
$\lambda_f = 0.75$								
0'	S_1	0.57	0.573	**0.003**	**0.54%**	0.064	0.065	0.065
1	S_1	0.57	0.541	-0.029	-5.04%	0.049	**0.049**	**0.057**
0	S_1	0.57	0.575	0.005	0.88%	0.080	0.080	0.081
0'	S_2	0.62	0.625	**0.005**	**0.76%**	0.076	0.076	**0.076**
1	S_2	0.62	0.540	-0.080	-12.94%	0.049	**0.049**	0.094
0	S_2	0.62	0.627	0.007	1.08%	0.105	0.106	0.106
0'	f_2	0.134	0.135	**0.001**	**0.48%**	---	0.017	0.017
1	f_2	0.134	0.146	0.012	9.24%	0.010	**0.010**	**0.016**
0	f_2	0.134	0.136	0.002	1.19%	0.025	0.025	0.025
0'	f_3	0.135	0.135	**0.000**	**0.24%**	---	0.016	**0.016**
1	f_3	0.135	0.166	0.031	22.91%	0.011	**0.011**	0.033
0	f3	0.135	0.137	0.002	1.58%	0.025	0.026	0.026

Table 7. Simulation results for estimating survival rates in the first two years, S_1 and S_2, and the expected fraction that is caught and reported in years two and three, f_2 and f_3, when tag fouling does not affect visibility (data generated with model 1). Values for all parameters appear in Table 4. The smallest values for bias, % bias, standard error of the estimates, and root mean squared error (RMSE) are in bold. Mean $\hat{SE}$ refers to the mean of the 10,000 estimated standard errors which come from the square root of the variance, calculated by solving the inverse of the hessian. SE of estimates refers to the true standard error (i.e., the standard deviation) of the 10,000 estimates of each parameter.

Model Fitted	Parameter	True Value	Mean estimate	Bias	% Bias	Mean $\hat{SE}$	SE of estimates	RMSE
0'	S_1	0.57	0.575	0.005	0.91%	0.082	0.082	0.083
1	S_1	0.57	0.574	**0.004**	**0.65%**	0.067	**0.068**	**0.068**
0	S_1	0.57	0.577	0.007	1.23%	0.102	0.102	0.103
0'	S_2	0.62	0.626	0.006	0.94%	0.098	0.099	0.099
1	S_2	0.62	0.624	**0.004**	**0.72%**	0.077	**0.077**	**0.077**
0	S_2	0.62	0.632	0.012	2.00%	0.133	0.135	0.135
0'	f_2	0.091	0.092	0.001	0.66%	---	0.014	0.014
1	f_2	0.091	0.091	**0.000**	**-0.11%**	0.008	**0.008**	**0.008**
0	f_2	0.091	0.093	0.002	2.13%	0.021	0.021	0.022
0'	f_3	0.086	0.087	0.001	0.84%	---	0.013	0.013
1	f_3	0.086	0.086	**0.000**	**-0.10%**	0.008	**0.008**	**0.008**
0	f_3	0.086	0.088	0.002	2.11%	0.020	0.020	0.020

Table 8. Simulation results for estimating survival rates in the first two years, S_1 and S_2, and the expected fraction that is caught and reported in years two and three, f_2 and f_3, for the situation where tag fouling occurs and takes a full year (data generated with model 0). Fouled tags have visibility that is 50% of that of new tags. Values for all parameters appear in Table 4. The smallest values for bias, % bias, standard error of the estimates, and root mean squared error (RMSE) are in bold. Mean $\hat{SE}$ refers to the mean of the 10,000 estimated standard errors which come from the square root of the variance, calculated by solving the inverse of the hessian. SE of estimates refers to the true standard error (i.e., the standard deviation) of the 10,000 estimates of each parameter.

Model Fitted	Parameter	True Value	Mean estimate	Bias	% Bias	Mean $\hat{SE}$	SE of estimates	RMSE
0'	S_1	0.57	0.438	-0.132	-23.19%	0.059	0.059	0.145
1	S_1	0.57	0.395	-0.175	-30.76%	0.043	**0.043**	0.181
0	S_1	0.57	0.577	**0.007**	**1.23%**	0.102	0.102	**0.103**
0'	S_2	0.62	0.477	-0.143	-23.08%	0.067	0.066	0.158
1	S_2	0.62	0.404	-0.216	-34.83%	0.042	**0.041**	0.220
0	S_2	0.62	0.632	**0.012**	**2.00%**	0.133	0.135	**0.135**
0'	f_2	0.091	0.138	0.047	52.11%	---	0.019	0.051
1	f_2	0.091	0.159	0.068	74.95%	0.011	**0.011**	0.069
0	f_2	0.091	0.093	**0.002**	**2.13%**	0.021	0.021	**0.022**
0'	f_3	0.086	0.137	0.051	59.19%	---	0.018	0.054
1	f_3	0.086	0.173	0.087	100.94%	0.011	**0.012**	0.088
0	f_3	0.086	0.088	**0.002**	**2.11%**	0.020	0.020	**0.020**

Table 9. Simulation results for estimating survival rates in the first two years, S_1 and S_2, and the expected fraction caught and reported in years two and three, f_2 and f_3, when it takes one year for the tags to fully foul (data generated with model 0). Tag-reporting rate for fouled tags is represented by λ_f. The smallest values for bias, % bias, standard error of the estimates, and root mean squared error (RMSE) are in bold. Mean $\hat{SE}$ refers to the mean of the 10,000 estimated standard errors which come from the square root of the variance, calculated by solving the inverse of the hessian. SE of estimates refers to the true standard error (i.e., standard deviation) of the 10,000 estimates of each parameter.

Model Fitted	Parameter	True Value	Mean estimate	Bias	% Bias	Mean $\hat{SE}$	SE of estimates	RMSE
$\lambda_f = 0.25$								
0'	S_1	0.57	0.288	-0.282	-49.51%	0.054	0.055	0.287
1	S_1	0.57	0.230	-0.340	-59.62%	0.035	**0.036**	0.342
0	S_1	0.57	0.583	**0.013**	**2.33%**	0.149	0.153	**0.153**
0'	S_2	0.62	0.348	-0.272	-43.89%	0.063	0.061	0.279
1	S_2	0.62	0.256	-0.364	-58.71%	0.035	**0.034**	0.366
0	S_2	0.62	0.645	**0.025**	**4.10%**	0.198	0.205	**0.206**
0'	f_2	0.043	0.120	0.077	179.73%	---	0.021	0.080
1	f_2	0.043	0.158	0.115	268.10%	0.011	**0.012**	0.116
0	f_2	0.043	0.045	**0.002**	**5.28%**	0.015	0.016	**0.016**
0'	f_3	0.043	0.116	0.073	170.19%	---	0.018	0.075
1	f_3	0.043	0.173	0.130	303.43%	0.012	**0.012**	0.131
0	f_3	0.043	0.045	**0.002**	**4.83%**	0.015	0.016	**0.016**
$\lambda_f = 0.75$								
0'	S_1	0.57	0.518	-0.052	-9.10%	0.057	0.058	**0.078**
1	S_1	0.57	0.503	-0.067	-11.75%	0.045	**0.045**	0.081
0	S_1	0.57	0.575	**0.005**	**0.86%**	0.080	0.081	0.081
0'	S_2	0.62	0.574	-0.046	-7.40%	0.068	0.068	**0.082**
1	S_2	0.62	0.519	-0.101	-16.35%	0.046	**0.046**	0.111
0	S_2	0.62	0.627	**0.007**	**1.09%**	0.105	0.106	0.106
0'	f_2	0.134	0.158	0.024	17.82%	---	0.019	0.031
1	f_2	0.134	0.164	0.030	22.46%	0.011	**0.011**	0.032
0	f_2	0.134	0.136	**0.002**	**1.26%**	0.025	0.025	**0.025**
0'	f_3	0.133	0.155	0.022	16.78%	---	0.018	0.029
1	f_3	0.133	0.178	0.045	33.49%	0.011	**0.011**	0.046
0	f_3	0.133	0.135	**0.002**	**1.57%**	0.025	0.025	**0.025**

Figure 1. The relationship between u_b, the exploitation rate in part (b) of year, and λ_f, the tag-reporting rate for fouled tags, and the resultant recovery rate, f_b, in part (b) of the year.
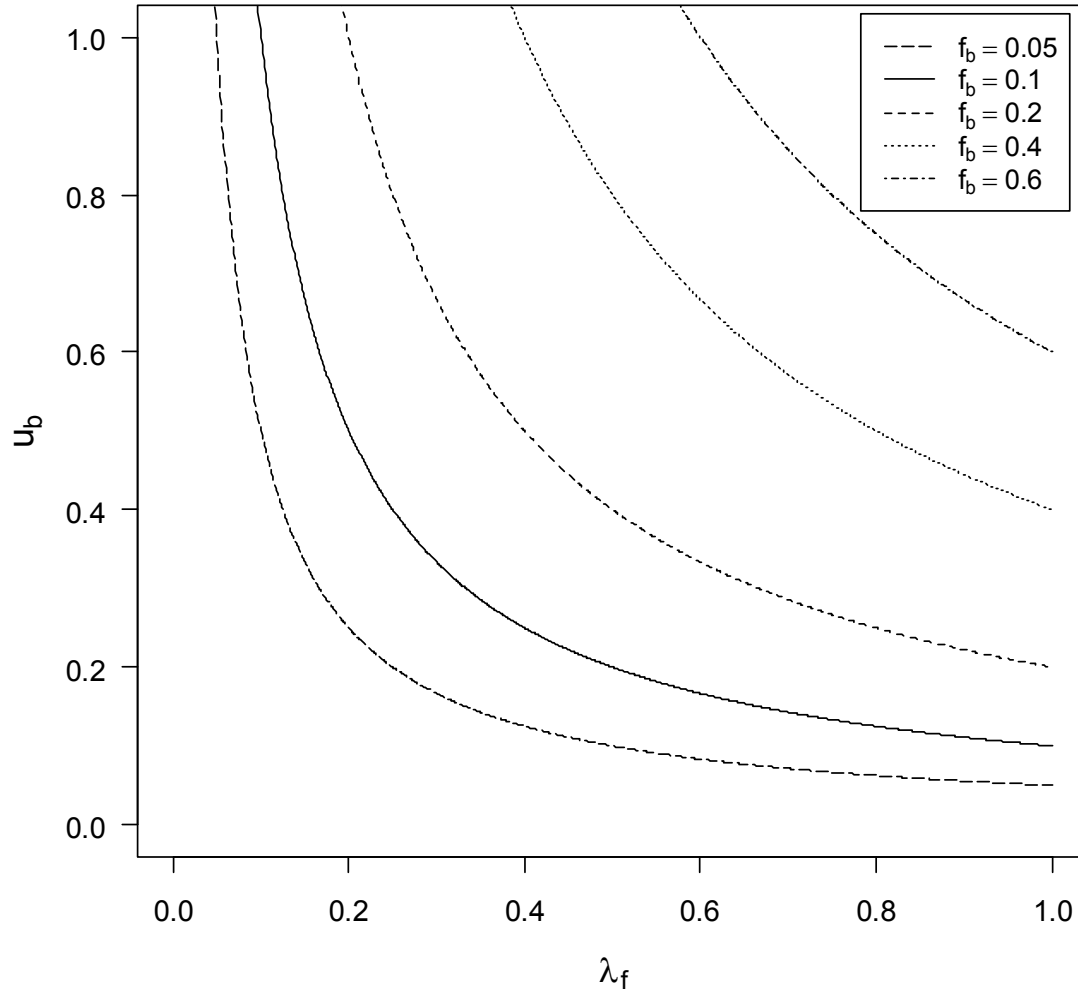
Figure 2. The root mean squared errors (RMSE) for the estimates of survival during the

first year, S_1. Data are generated for the situation where tag fouling affects visibility and

it takes a half year for tags to become fouled (data generated under model 0'). The tag-

reporting rate for clean tags was 1.0 and for fouled tags it was modeled at 0.25, 0.50, and

0.75. Note that when the tag-reporting rate for fouled tags is 0.75, model 1 has a lower

RMSE than both model 0 and model 0'.

RMSE for S_1 for changing visibility effect

Data generated with model zero prime
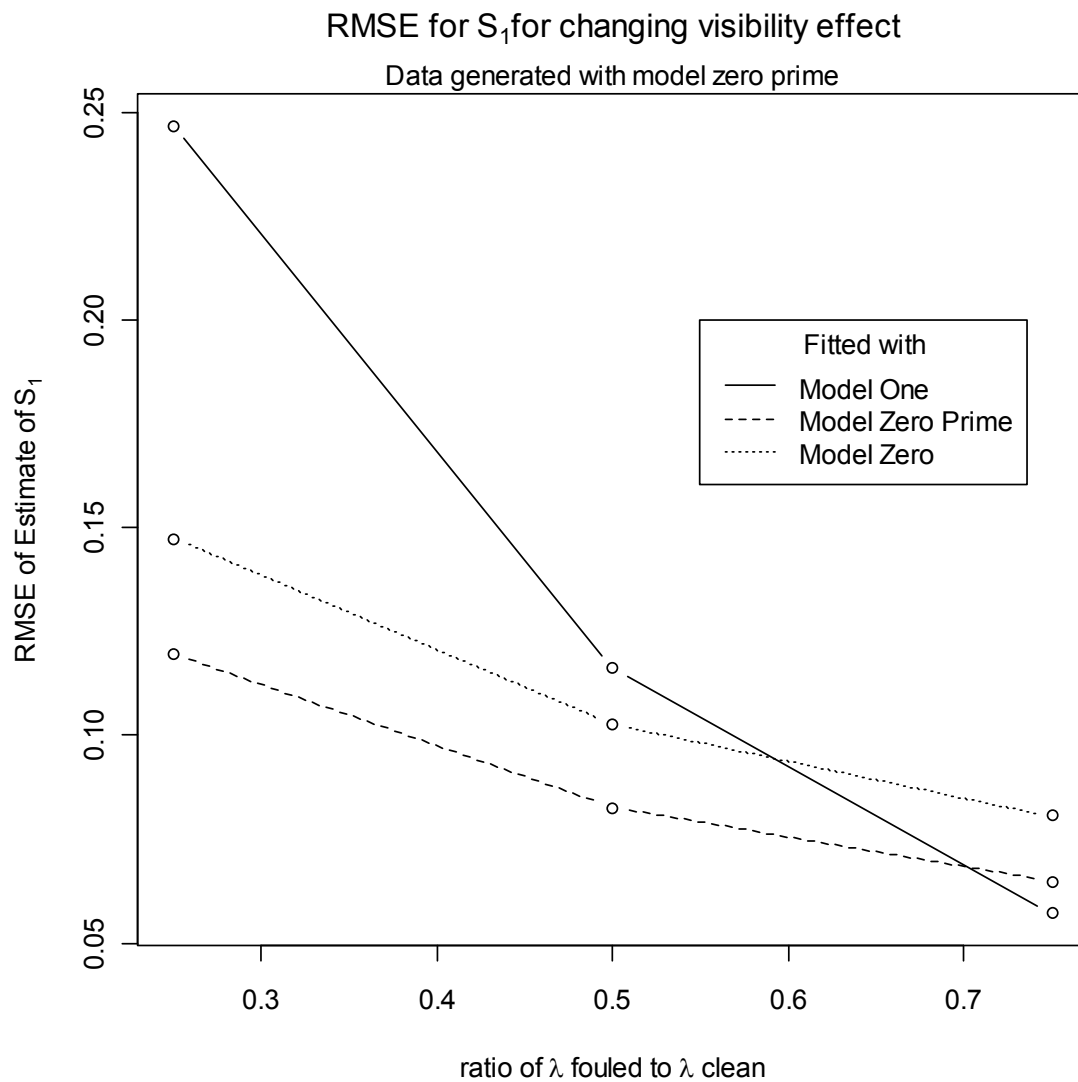
Figure 3. The root mean squared errors (RMSE) for the estimates of survival during the

second year, S_2. Data are generated for the situation where tag fouling affects visibility

and it takes a full year for tags to become fouled (data generated under model 0). The

tag-reporting rate for clean tags was 1.0 and for fouled tags it was modeled at 0.25, 0.50,

and 0.75. Note that when the tag-reporting rate for fouled tags is 0.75, model 0' has a

lower RMSE than model 0.

RMSE for S_2 for changing visibility effect
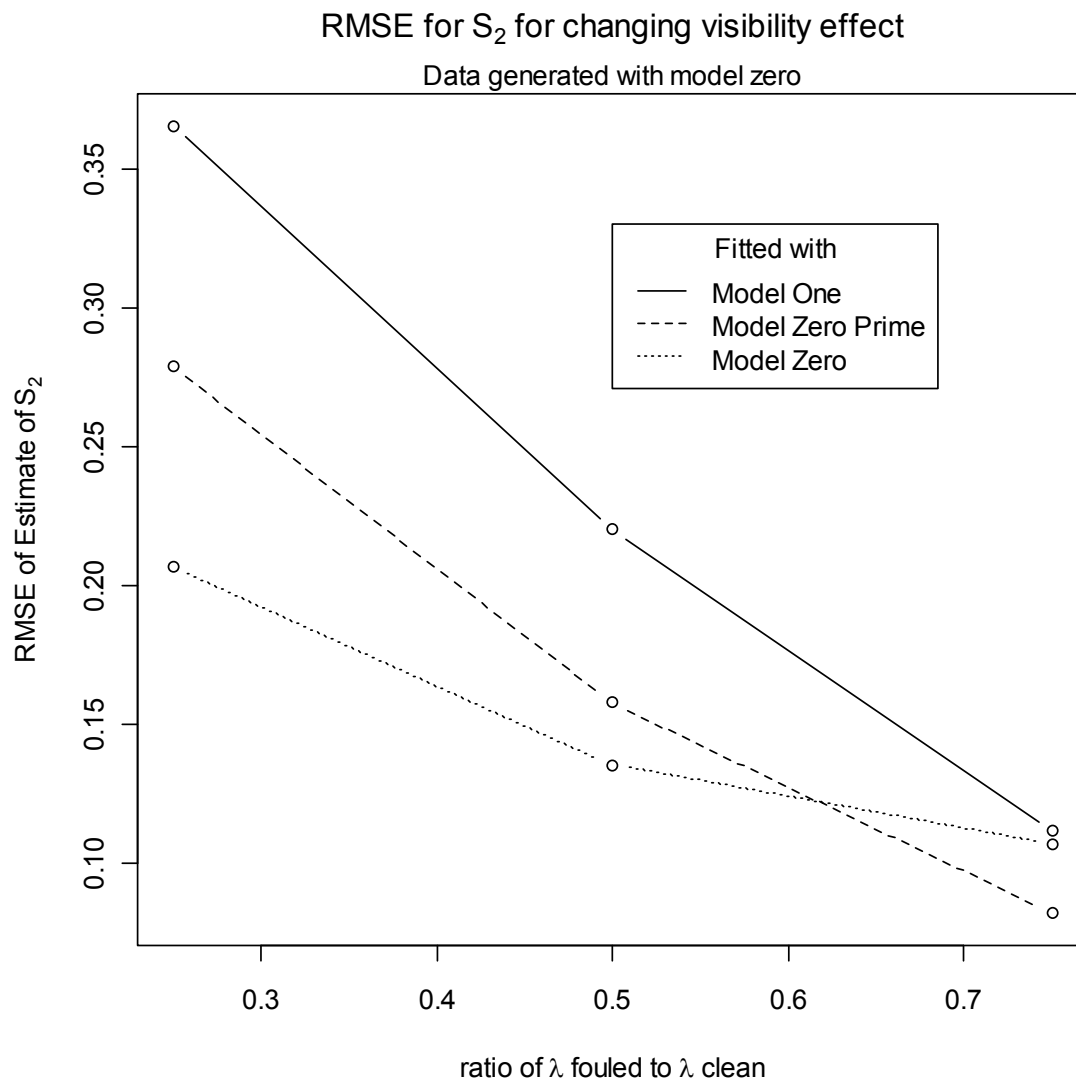
Chapter 3

Instantaneous rates tagging models allowing for delayed mixing of newly tagged cohorts:

partial year tabulation of recaptures

Abstract

Instantaneous rates tagging models can be used to estimate natural mortality and fishing mortality rates from multi-year tagging studies in which cohorts of tagged animals are released at the start of each year. The models can include additional parameters to account for a delay in mixing of newly tagged animals with previously tagged animals. One such model allows for the newly tagged population to become fully mixed in less than one year. Here a new model, referred to as the partial-year tabulation model for delayed mixing (delayed pyt model), is proposed in which the year is divided into parts and tag returns are tabulated by parts of the year rather than a full year. This is beneficial when there is delayed mixing because it achieves greater precision and provides estimates of the instantaneous rate of fishing mortality in the first year, which cannot be estimated when tag returns are tabulated by full year. The new model can be used at little or no extra cost. The superiority of the delayed pyt model is demonstrated through Monte Carlo simulation.

Introduction

Multi-year tagging studies are often used to obtain estimates of survival rates and instantaneous rates of fishing and natural mortality. Modern, multiyear tagging models were first described by Seber (1970) and Robson and Youngs (1971). Brownie et al. (1978, 1985) popularized the approach by describing a suite of models that enables the user to estimate age- and year-specific survival rates and the models are now commonly referred to as Brownie models.

The biggest difficulty in conducting tagging studies is to ensure that the tagged population is thoroughly mixed with the population at large so that recaptures of tagged animals reflect what is happening to the untagged population. Unfortunately, the Brownie models cannot accommodate a lack of mixing of newly tagged animals into the population (Youngs and Robson 1975). Hoenig et al. (1998a,b) used additional information about the timing of the fishery within the year to reparameterize the Brownie models in terms of instantaneous rates of fishing and natural mortality. This enables one to estimate the components of mortality. Significantly, this approach can accommodate delayed mixing of newly tagged animals into the untagged population (Hoenig et al. 1998b).

At the start of each of several successive time periods a sample of the population, termed a cohort, is captured, tagged, and released. Recaptures of tagged animals are tabulated by period, which is typically a year. In this paper annual tagging periodicity is assumed. The data consist of counts r_{ij}, which are the realizations of (i.e., observations on) the random variables R_{ij}, i.e., the r_{ij} are the observed number of recaptures of animals

tagged in year i ($i = 1, 2, \ldots, I$) and recaptured in year j ($j = i, \ldots, J$, with $J \geq I$). The estimates of rates of natural and fishing mortality are calculated using full years of data.

For models parameterized in terms of instantaneous rates, parameters can be included to allow for a delay in the mixing of the newly tagged animals with the untagged and previously tagged population (Hoenig et al. 1998b). One of the Hoenig et al. (1998b) models allows for the newly tagged animals to be mixed after a full year at liberty and another formulation assumes mixing is complete after a partial year. For both models, an estimate of the normal instantaneous rate of fishing mortality during the first year of the study is not possible due to confounding of parameters.

If newly tagged animals fully mix with the population in less than a year, the tabulation of recoveries by year results in a loss of efficiency (in terms of larger standard errors for every parameter). In this paper, we present a new model, the partial year tabulation model for delayed mixing (delayed pyt), in which recoveries are tabulated by portion of the year. The year is split in parts such that in the first part of the year, part (a), the newly tagged animals are not fully mixed with the population at large and in part (b), the second part of the year, the newly tagged animals have fully mixed with the population at large. The data consist of an upper trapezoidal array where the cells contain counts r_{ijk}, which are the realizations (i.e., observations) of the random variables R_{ijk}, i.e., the observed numbers of recaptures of animals tagged in year i ($i = 1, 2, \ldots, I$) and recaptured in part k, $k \in \{a, b\}$, of year j, $j = i, \ldots, J$ with $J \geq I$.

The performance of the delayed pyt model will be evaluated by Monte Carlo simulation. The new model will be compared to the models of Hoenig et al. (1998 a,b). I will show that, by making better use of the available information, an estimate of the

61

instantaneous rate of fishing mortality in the first year is possible and there is an increase

in the overall efficiency of estimation.

Instantaneous Rates Tagging Models

Hoenig et al. (1998a) formulated the Brownie model in terms of instantaneous

rates of fishing and natural mortality given information on the timing of the fishery. The

Brownie models are parameterized in terms of the survival rate in year j, S_j, and the tag

recovery rate in year j, f_j. The survival rate, denoted S_j, is defined as the expected

fraction of the population alive at the start of year j that is still alive at the end of the year.

The survival rate in year j can be expressed in terms of the instantaneous rates of fishing

(F_j, yr $^{-1}$) and natural mortality (M, yr $^{-1}$) by,

(1) $S_j = \exp(- F_j - M)$.

This holds regardless of the relative timing of the fishing and natural mortality. The tag

recovery rate in year j, f_j, can be expressed as a product of its components (Pollock et al.

1991; Hoenig et al. 1998a)

(2) $f_j = \phi \, \lambda \, u_j$,

where ϕ is a combination of short term survival rate from tagging and short term

probability of tag retention, u_j is the exploitation rate, or the expected fraction of the

population alive at the start of the year that dies due to harvesting during the year, and λ is

the tag reporting rate, or the probability that a tag will be reported if the fish is

recaptured.

It is necessary to know the timing of the fishery during the year in order to relate

the exploitation rate to the instantaneous rates of fishing and natural mortality as will be

shown below. Hoenig et al. (1998a) provided a general model formulation which allows

for an arbitrary fishing pattern throughout the year. When there is constant fishing

intensity throughout the year, exploitation in year j can be expressed:

$$(3) \quad u_j = u_j(F_j, M) = \frac{F_j}{F_j + M}(1 - \exp(-F_j - M)).$$

The notation of a u followed by parentheses is used for brevity to indicate u is a function

of the parameters in parentheses. It is possible to fit a model where M varies over time

but here it is assumed to be constant both within and over years. If the year is split into

two parts, part (a) and part (b), then the exploitation rates in part (a) and (b) for year j, u_{ja}

and u_{jb}, are given by,

$$(4a) \quad u_{ja} = \frac{p_j F_j}{p_j F_j + \gamma M}(1 - \exp(-p_j F_j - \gamma M))$$

$$(4b) \quad u_{jb} = \frac{q_j F_j}{q_j F_j + (1-\gamma)M}(1 - \exp(-q_j F_j - (1-\gamma)M))\exp(-p_j F_j - \gamma M)$$

where p_j is the proportion of the fishing mortality occurring in part (a) of the year, q_j is

the proportion of the fishing mortality occurring in part (b) of the year (note $q_j = 1 - p_j$), γ

is the fraction of the year that elapses in part (a), and F and M are as defined previously.

The value of p is determined under the assumption that the fraction of fishing mortality

occurring is part (a) is proportional to the fraction of fishing effort that occurs in part (a)

(if determining relative effort in the parts of a year is difficult one can use relative catch,

as suggested by Hoenig et al. 1998a). This assumption is generally not critical unless the

natural mortality rate, M, is high (Hoenig et al. 1998a). The exploitation rate in year j is

the sum of the exploitation rates in parts (a) and (b), i.e.,

$$(5) \quad u_j = u_{ja} + u_{jb}.$$

Note that under this formulation, the fishing mortality is assumed to be constant with time within a time period but may vary among periods.

The recapture data (r_{ij}) can be explained by modeling the expected cell counts in terms of survival rates and exploitation rates which are themselves functions of the fishing and natural mortality rates. It is then a straightforward matter to construct the likelihood function for the data in order to estimate the parameters.

Fully Mixed Case

When the tagged cohort is fully mixed with the population at large prior to the start of fishing, the expected recovery from the cohort tagged in year i and recovered in year j, $E(R_{ij})$, can be expressed:

$$(6) \quad E(R_{ij}) = \begin{cases} N_i \phi \lambda u_j & , \ i = j \leq I \\ N_i \phi \lambda u_j \prod_{h=i}^{j-1} S_h & , \ i < j \leq J \end{cases}$$

where N_i is the number tagged and released in year i and the other parameters are as defined previously. There is an implicit category for all animals of a cohort that are never seen again, Y_i. This can be expressed as:

$$(7) \quad Y_i = N_i - \sum_{h=i}^{J} r_{ih} \ .$$

Maximum likelihood estimation (MLE) is generally used to obtain estimates for the model parameters based on the observed data. MLE methods consist of writing an expression, called the likelihood function, for the probability of having obtained the observed data; the likelihood is written in terms of the unknown parameters of interest.

The values of the parameters which maximize the likelihood function are called the maximum likelihood estimates. Generally, numerical methods are used to solve for the maxmimum likelihood estimates. Maximum likelihood estimators have the following desirable asymptotic properties (under the regularity conditions): unbiased, unique, consistent, minimum variance, and normally distributed (e.g., Hogg et al. 2005; Wackerly et al. 2002).

The likelihood function for instantaneous rates models can be expressed as the product of independent multinomial distributions of tag returns over time, with each tagged cohort giving rise to a multinomial distribution. The general form of the likelihood function Λ for product multinomial models can be expressed as:

$$(8) \quad \Lambda \propto \prod_{i=1}^{I} \left(\prod_{j=i}^{J} P_{ij}^{r_{ij}} \right) \left(1 - \sum_{j=i}^{J} P_{ij} \right)^{Y_i} ,$$

where the symbol $\propto$ means "is proportional to," P_{ij} is the cell probability of recovering a tagged animal in year j given that it was tagged in year i,

$$(9) \quad P_{ij} = \frac{E(R_{ij})}{N_i} ,$$

and the other quantities are as defined previously. The likelihood function is proportional to the product of the cell probabilities raised to the power of observed recaptures. The right hand portion of the equation involving one minus a sum is the expression for the animals in a cohort that are never seen again after tagging, that is, the cell probability of never being seen again raised to the power of the number of tagged animals which are never seen again, given by Y_i (Equation 7).

Delayed Mixing Lasting a Full Year

Hoenig et al. (1998b) extended the instantaneous rates models to allow for the newly tagged animals to not fully mix with the previously tagged population until either a full year or partial year has elapsed. As a result of this delay, the newly tagged animals in year j experience an abnormal instantaneous rate of fishing mortality, $F_j{}^*$, for the first year at liberty ($F_j{}^*$ is considered abnormal because it does not reflect the mortality which previously tagged animals experience as the newly tagged cohort has not yet fully mixed). Values of $F_j{}^*$ are generally not of interest but are included in the model to avoid bias in the estimates of the F_j's. For a fishery with constant fishing intensity over the year the exploitation rate, $u_i{}^*$, for the first year a cohort is at liberty is now expressed as,

$$(10) \quad u_i{}^* = u_i{}^* \left(F_i{}^*, M \right) = \frac{F_i{}^*}{F_i{}^* + M} \left(1 - \exp\left(-F_i{}^* - M \right) \right) .$$

The asterisk is used here to denote that the rate of fishing mortality is abnormal which causes the exploitation rate to be different from that of previously tagged animals. The expected recoveries from the cohort tagged in year i and recovered in year j can be expressed,

$$(11) \quad E(R_{ij}) = \begin{cases} N_i \phi \lambda u_j{}^* (F_j{}^*, M) & , \ j = i \\ N_i \phi \lambda u_j (F, M) S_i{}^* & , \ j = i+1 \\ N_i \phi \lambda u_j (F, M) S_i{}^* \prod_{h=i+1}^{j-1} S_h & , \ j > i+1 \end{cases}$$

where $S_i{}^*$ is the abnormal survival rate in year i for the cohort tagged in year i that has not fully mixed with the previously tagged population,

66

(12) $S_i^* = \exp\left(-M - F_i^*\right)$

and the other parameters are the same as described previously.

Delayed Mixing Lasting Part of the Year

As shown by Hoenig et al. (1998b), if the newly tagged animals fully mix in less than a year, the year can be split into two parts, part (a) and part (b), such that animals are fully mixed into the population at large by the end of part (a). The exploitation rates in part (a) and (b) for the first year each cohort is at liberty are given by u_{ia}^* and u_{ib}^*

(13a) $u_{ia}^* = u_{ia}^*\left(F_i^*, M\right) = \dfrac{p_i F_i^*}{p_i F_i^* + \gamma M}\left(1 - \exp\left(-p_i F_i^* - \gamma M\right)\right)$

(13b) $u_{ib}^* = u_{ib}^*\left(F_i^*, F_i, M\right) = \dfrac{q_i F_i}{q_i F_i + (1 - \gamma)M}\left(1 - \exp\left(q_i F_i - (1 - \gamma)M\right)\right)\exp\left(-p_i F_i^* - \gamma M\right).$

The asterisk on u_{ib}^* occurs because the exploitation rate in part (b) of year i depends on F_i^* (the abnormal fishing mortality rate in part (a) of year i).

If newly tagged animals are assumed to be fully mixed after part of the year, rather than the full year, the expected recaptures change to:

(14) $E(R_{ij}) = \begin{cases} N_i \phi \lambda \left(u_{ia}^*\left(F_i^*, M\right) + u_{ib}^*\left(F_i^*, F_i, M\right)\right) & , i = j \\ N_i \phi \lambda \left(u_{ja}\left(F_j, M\right) + u_{jb}\left(F_j, M\right)\right)S_i^* & , j = i + 1 \\ N_i \phi \lambda \left(u_{ja}\left(F_j, M\right) + u_{jb}\left(F_j, M\right)\right)S_i^* \displaystyle\prod_{h=i+1}^{j-1} S_h & , j > i + 1 \end{cases}$

where S_i^* is the abnormal survival rate for the first year each cohort is at liberty given by

(15) $S_i^* = \exp\left(-p_i F_i^* - q_i F_i - M\right).$

Note that the abnormal survival rate in the year of tagging, S_i^*, is specified differently in equations 12 and 15. For equation (12) delayed mixing lasts a full year and thus animals must survive the abnormal instantaneous rate of fishing mortality, F_i^*, for the full first year at liberty, year i. For equation (15) animals experience an abnormal instantaneous rate of fishing mortality, F_i^*, for only part of year i, the first year the cohort is at liberty, and experience the normal rate of fishing mortality, F_i, for the remainder of the year.

Note that under the delayed mixing model, where the delay in mixing lasts part of the year, for year $i = 1$ there is no separate F_1 and F_1^* ; instead there is only F_1^* because the parameters are not separably estimable (Hoenig et al. 1998b). Logically, in year one of a tagging study, all of the animals are newly tagged so there are no previously tagged animals from which to estimate the normal rate of fishing mortality.

As before, maximum likelihood estimation methods can be used to estimate the parameters because the likelihood under delayed mixing is of the same form as found in Equation 8.

New Model: partial year tabulation for delayed mixing lasting part of year

A new model, called the delayed pyt model, is now developed for the case where delayed mixing occurs but the newly tagged cohort fully mixes with the untagged and previously tagged population in less than a year. As before, the year is divided into parts such that in part (a) of the year, the newly tagged cohort is not fully mixed with the population at large, and by part (b) of the year the cohort is fully mixed with the population at large and experiencing the same rate of fishing mortality as previously tagged cohorts. The key feature of the delayed pyt model is that tag recaptures are

tabulated separately for parts (a) and (b) of every year. The expected recaptures, from the

cohort tagged in year i and recaptured in part (a) of year j, can be expressed as:

$$
(16a) \quad E(R_{ija}) = \begin{cases} N_i \phi \lambda u_{ia}{}^* \left(F_i^*, M \right) & , i = j \\ N_i \phi \lambda u_{ja} \left(F_j, M \right) S_i^* & , j = i+1 \\ N_i \phi \lambda u_{ja} \left(F_j, M \right) S_i^* \prod_{h=i+1}^{j-1} S_h & , j > i+1 \end{cases}
$$

and for part (b) of year j,

$$
(16b) \quad E(R_{ijb}) = \begin{cases} N_i \phi \lambda u_{ib}{}^* \left(F_i^*, F_i, M \right) & , i = j \\ N_i \phi \lambda u_{jb} \left(F_j, M \right) S_i^* & , j = i+1 \\ N_i \phi \lambda u_{jb} \left(F_j, M \right) S_i^* \prod_{h=i+1}^{j-1} S_h & , j > i+1 \end{cases}
$$

where $u_{ia}{}^*$ and $u_{ib}{}^*$ are given by equations (13a) and (13b), respectively, u_{ja} and u_{jb} are

given by equations (4a) and (4b), respectively, and the remaining parameters are as

defined previously. Note that equations (16a) and (16b) differ from equation (14) in that

the tabulations are done by parts of the year for the delayed pyt model and by full year in

equation (14).

The likelihood is constructed similarly to equation (8), but with the year broken

into parts (a) and (b) and the tag recaptures tabulated by parts. The recapture cell

representing the tagged animals which are never seen again is given by

$$
(17) \quad W_i = N_i - \sum_{j=i}^{J} \sum_{k \in \{a,b\}} r_{ijk}
$$

Thus the likelihood can be expressed

$$(18) \quad \Lambda \propto \prod_{i=1}^{I} \left(\prod_{j=i}^{J} \prod_{k\in\{a,b\}} P_{ijk}^{r_{ijk}} \right) \left(1 - \sum_{j=i}^{J} \sum_{k\in\{a,b\}} P_{ijk} \right)^{W_i}$$

where P_{ijk} is the cell probability of recovering a tagged animal in part k, $k \in \{a,b\}$, of

year j given that it was tagged in year i,

$$(19) \quad P_{ijk} = \frac{E(R_{ijk})}{N_i},$$

and the remaining parameters are as defined previously.

Simulations

Methods

Monte Carlo simulations were performed to evaluate the performance of the

delayed pyt model and to compare the results with those from the partial-year delayed-

mixing model described by Hoenig et al. (1998b) using simulated data reflecting delayed

mixing lasting part of the year. The simulations consisted of three years of tagging (years

1, 2, and 3) and four full years of recaptures, that is, recaptures for periods 1a, 1b, 2a, 2b,

3a, 3b, 4a, and 4b for the new model and years 1, 2, 3, and 4 for the Hoenig et al. (1998b)

delayed mixing model. One thousand animals were tagged each year and two thousand

product multinomial datasets were simulated for each scenario. The scenarios simulated

delayed mixing lasting for half a year. That is, after half of a year has passed the newly

tagged animals were fully mixed with the untagged population as well as with the

previously tagged animals. Note that ϕ and λ are confounded and cannot be estimated

separately. For the purpose of the Monte Carlo simulations the product of $\phi\lambda$ was set to 1.0, indicating the tag reporting rate was one-hundred percent, no tags were shed immediately after tagging, and there was no short-term tag-induced mortality.

The parameters used to generate the datasets under two scenarios are given in Tables 1 and 2. For each dataset the fraction of fishing effort and thus, by assumption, fishing mortality in part (a) is equal to that in part (b) ($p = q$). One dataset (Table 2) had values of the rates of fishing mortality close to each other (0.22 to 0.28) and the other (Table 1) had a wider range of rates of fishing mortality (0.20 to 0.45).

The following two models were fitted to each of the datasets:

1) Delayed mixing model, as described by Hoenig et al. (1998b), in which mixing is assumed after half a year; data are tabulated by year.

2) Delayed pyt model (new model), in which mixing occurs after half a year; data are tabulated by half year.

Computations were done using the statistical language R (R Development Core Team 2008) as described below.

A second group of simulations was conducted in order to evaluate how well a model selection criterion would perform in choosing the best fitting model. Two models were compared: the delayed pyt model and a fully mixed model in which the tag recaptures were tabulated by parts of the year (referred to as the fully mixed pyt model). The formulation for the fully mixed pyt model is similar to that of the delayed pyt model, except anywhere an $F_j{}^*$ appears it is replaced by F_j. Tabulating recaptures by parts of the year does not increase the precision of the fully mixed pyt model relative to the fully

mixed model with tag recaptures by year, but it does allow for model comparison between models with immediate and delayed mixing of newly tagged animals.

A commonly used model selection criterion is Akaike's Information Criterion (AIC) (Anderson 2008). AIC is defined as:

(20) $AIC = -2\log_e(\Lambda) + 2k,$

where Λ is the value of the maximum for the likelihood function for the model and k is the number of parameters in the model. The model which has the lower AIC value is considered the better fitting model. Additionally, one can define ΔAIC, which is the difference in AIC values between an alternative model and the best fitting model. A general rule of thumb is that a model with a ΔAIC of less than two or three is considered a plausible alternative to the best fitting model (Anderson 2008).

For the second group of simulations, the values of fishing mortality and natural mortality were those in Table 1. The values of the abnormal rates of fishing mortality, F_{ja}^*'s, relative to the normal rates of fishing mortality F_{ja}'s (note $F_{ja}=0.5F_j$) were varied in order to see what effect the magnitude of this difference had on model performance and on AIC for model selection (Tables 3, 4, 5 and Figure 1). Parameter estimates and standard errors were calculated for an 'AIC-selected model', that is, for each simulation the estimates for the parameters from the model (fully mixed pyt or delayed pyt models) which had the lower AIC value were used, and then these 10,000 model-selected estimates were averaged (Tables 3, 4, and 5).

The function 'rmultinomial' (Chasalow 2005) in the R package 'combinat' was used to generate multinomial datasets with specified sample sizes and cell probabilities. The function 'nlm' was used to minimize the negative log likelihood function and obtain

the Hessian (R Development Core Team 2008). Standard errors were estimated by

inverting the Hessian matrix using the R function 'solve'. The true standard error was

determined from the variability (standard deviation) of the two thousand estimates for

each parameter. The output from the Monte Carlo simulations includes: estimates of the

parameters for each simulated dataset and the bias, standard error, and bias of the

estimated standard error. Additionally the root mean square error (RMSE) for each

parameter was calculated using

$$(21)\ RMSE = \sqrt{bias^2 + variance} = \sqrt{\frac{\sum_{i=1}^{T}(\hat{\theta}_i - \theta)^2}{T}},$$

where T is the number of simulated datasets (2,000) and $\hat{\theta}_i$ is the ith estimate of the

parameter whose true value is Θ. The parameters estimated for the new model are: F_1,

F_2, F_3, F_4, $F_{1a}{}^*$, $F_{2a}{}^*$, and $F_{3a}{}^*$. The parameters estimated for the Hoenig et al. (1998b)

model are F_2, F_3, F_4, $F_1{}^*$, $F_{2a}{}^*$, and $F_{3a}{}^*$. (Hoenig et al. (1998b) parameterized the

model in terms of $F_i{}^*$'s, but the simulations here were done with $F_{ia}{}^*$'s and the

relationship is $pF_i{}^* = F_{ia}$). There is no $F_{4a}{}^*$ since tagging occurred for just three years.

Simulation Results

    In the first set of simulations, when the F_j's vary greatly from year to year (range

0.20 to 0.45), both the delayed pyt model and the delayed mixing model of Hoenig et al.

(1998b) yield essentially unbiased estimates of the F_j's ($j = 2,…,4$) and $F_{ja}{}^*$'s ($j = 2, 3$)

(Table 1). For all the F_j's and F_{ja}'s, the delayed pyt model produces estimates with

smaller standard deviations and smaller RMSE's. Only the delayed pyt model yields an

estimate of F_1: mean $\hat{F}_1 = 0.200$ ($F_1 = 0.20$) with a RMSE of 0.023. Both models yield

essentially unbiased estimates of the standard errors.

When the F_j's are close in value to each other, both the delayed pyt model and the

delayed mixing model of Hoenig et al. (1998b) yield essentially unbiased estimates of the

F_j's ($j = 2,\ldots,4$) and F_{ja}*'s ($j = 2, 3$) (Table 2). The delayed pyt model produces an

unbiased estimate of M while the delayed mixing model of Hoenig et al. (1998b) is

negatively biased (% bias for $\hat{M} = -9\%$). Only the delayed pyt model produces an

estimate of F_1. However it is worth noting that the delayed pyt model produces a mean

estimate of $\hat{F}_1 = 0.220$ ($F_1 = 0.22$) and thus is unbiased, with a RMSE of 0.033. Both

models yield essentially unbiased estimates of the standard errors.

When the data are generated with the values of the F_j's being closer together than

the previous simulation (Table 2), the performance of the delayed pyt model improves

relative to the Hoenig et al. model by having lower RMSE's for the estimates of M, F_j's

($j = 1, 2, 3$), and F_{ja}*'s ($j = 1, 2, 3$).

For the second set of simulations, in which the data was such that the tag

recaptures were tabulated by parts of the year, the percentage of the 10,000 simulations in

which the delayed pyt model had a lower AIC value than the fully mixed pyt model

varied depending on the multiplier between F_j* and F_j (Figure 1). In fact, when the data

come from a case of full mixing (Table 3) and one uses AIC as the model selection

criterion, the wrong model is only selected 10.66% of the time. Hence, uncertainty about

model selection is low. In fact, bias of the estimates under model selection is essentially

zero (as with the fully mixed and delayed pyt models). The RMSE is slightly higher

when model selection is used than when the fully mixed pyt model is used. This is

because for 10.66% of the 10,000 simulations the less precise model (delayed pyt model) was selected.

When the data come from a case of slight delayed mixing, i.e, $F_j^* = 1.1F_j$, the delayed pyt model is selected 27.21% of the time (Figure 1). Thus, the incorrect model, the fully mixed pyt model, is selected 72.79% of the time, yielding parameter estimates with larger biases than the delayed pyt model (Table 5). Interestingly, for four of the five parameters (F_2, F_3, F_4, and M) the AIC-selected model has the lowest RMSE. This is because 72.79% of the time a more precise model (the fully mixed pyt model) is selected and 27.21% of time the essentially unbiased model (the delayed pyt model) is selected, resulting in estimates that have lower RMSE's than using either model alone.

If the data come from a stronger case of delayed mixing, i.e., $F_j^* = 0.75F_j$, the delayed pyt model is selected 93.1% of the time (Figure 1). Now, the AIC-selected model is tied for the lowest RMSE for just two of the parameters (F_1 and F_3) and the delayed pyt model has the lowest RMSE for all parameters (including the two ties) (Table 4).

Discussion

In order to determine if it is appropriate to apply the delayed pyt model it must first be determined if delayed mixing is a problem.   If a model for delayed mixing is applied when the newly tagged cohorts fully mix with the previously tagged, the delayed pyt model will likely still outperform the delayed mixing model of Hoenig et al. (1998b) by yielding estimates closer to those found when fitting the fully mixed model and the delayed pyt model will have smaller standard errors for each parameter (Tables 1, 2).

However, the delayed pyt model will likely be less precise than the fully mixed model since the additional parameters reduce precision.

When delayed mixing does occur for part of the year, the delayed pyt model provides more precise estimates of the rates of fishing mortality and natural mortality compared to the delayed mixing model of Hoenig et al. (1998b). Furthermore, if the fully mixed model is applied when delayed mixing occurs, the estimates may be biased for all parameters. In fact, whether there is full mixing or delayed mixing for part of a year, the delayed pyt model will generally outperform the delayed mixing model of Hoenig et al. (1998b) by providing estimates with smaller estimated standard errors and provide an estimate of the rate of fishing mortality in the first year (which the Hoenig et al. (1998b) model does not provide).

In order to use the delayed pyt model, it is necessary to determine the value of p, the proportion of fishing mortality occurring in part (a) of the year. The assumption of p is not critical for the estimation of any value except for F_1 (Hoenig et al. 1998b). It is critical for F_1 because the value of p determines the multiplier used to relate fishing mortality estimated for part (b) of the year to the fishing mortality over the whole year.

Additionally, the degree to which the rate of fishing mortality varies across the years will affect the precision of the estimates of the normal and abnormal rate of fishing mortality as shown by the simulations (c.f. Tables 1, 2). The greater the variability in the rate of fishing mortality the more precise the estimates will be.

Tabulating the recaptures on a finer scale than the periodicity of tagging should not be costly. Pollock and Raveling (1982) showed that when conducting a tagging study it is important to determine the date of tag recapture in order to avoid biased estimates of

survival. Thus, assuming the advice of Pollock and Raveling is followed, knowing whether the tag return is from part (a) or part (b) of the year should add little or no additional cost to the study.

For these simulations the product of $\varphi\lambda$ was assumed to be one. In practice one would either use the data to estimate $\varphi\lambda$ or use additional information to obtain an outside estimate. With a large dataset, i.e., high numbers of recaptures and many years of tagging and recaptures, it is possible to estimate $\varphi\lambda$ from the observed information. It is also possible to obtain an outside estimate of $\varphi\lambda$ using methods such as high-reward tagging studies or planted tag studies (Hoenig et al. 1998a). Thus, the delayed pyt model can be used in a much more general setting in which there is an outside estimate of $\varphi\lambda$ or when $\varphi\lambda$ is estimated from the observed information.

This is only the second study to show that tabulating tag recoveries on a finer scale than the periodicity of tagging can be advantageous. The other study, by Waterhouse (2010), described a generalization of Brownie model zero (Brownie et al. 1978, 1985) to address tag fouling occurring in the first part of the first year a cohort is at liberty with the result that newly tagged animals are more visible than previously tagged ones. It is yet to be seen what further benefits can accrue from tabulating recapture data on a finer scale than the tagging periodicity.

Acknowledgments

References

Anderson, D. R. 2008. Model based inference in the life sciences: a primer on evidence.
Springer. New York, NY.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson (1978). Statistical
inference from band recovery data: a handbook. 2nd Ed. U.S. Fish and Wildlife
Service Resource Publication 131. Washington, D.C.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson (1985). Statistical
inference from band recovery data: a handbook. U.S. Fish and Wildlife Service
Resource Publication 156. Washington, D.C.

Chasalow, S. 2005. combinat: combinatorics utilities. R package version 0.0-6.
Available: http://cran.r-project.org/web/packages/combinat/index.html. (July
2010).

Hoenig, J. M. , N. J. Barrowman, W.S. Hearn, and K.H. Pollock. 1998a. Multiyear
Tagging Studies Incorporating Fishing Effort Data. Canadian Journal of Fisheries
and Aquatic Sciences. 55:1466-1476.

Hoenig, J. M. , N. J. Barrowman, K.H. Pollock, E.N. Brooks, W.S. Hearn and T.
Polacheck. 1998b. Models for Tagging Data That Allow for Incomplete Mixing
of Newly Tagged Animals. Canadian Journal of Fisheries and Aquatic Sciences.
55:1477-1483.

Hogg, R. V., McKean, J. W. , and A. T. Craig. 2005. Introduction to Mathematical
Statistics. 6th Ed. Pearson Prentice Hall. Upper Saddle River, New Jersey.

Pollock, K. H. and D. G. Raveling. 1982. Assumptions of Modern Band-Recovery
 Models, with Emphasis on Heterogeneous Survival Rates. The Journal of
 Wildlife Management. 46(1):88-98.

Pollock, K.H., J.M. Hoenig, and C.M. Jones. 1991. Estimation of Fishing and Natural
 Mortality Rates when a Tagging Study is Combined with a Creel Survey or Port
 Sampling. American Fisheries Society Symposium 12:423-434.

R Development Core Team. 2008. 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 Youngs, W.D. 1971. Statistical analysis of reported tag recaptures in
 the harvest from an exploited population. Biometrics Unit Mimeographic Series
 Paper No. BU-369-M. Cornell University, Ithaca, N.Y.

Seber, G.A.F. 1970. Estimating time-specific survival and reporting rates for adult birds
 from band returns. Biometrika, 57: 313-318.

Wackerly, D.D., W. Mendenhall, R.L. Scheaffer. 2002. Mathematical Statistics with
 Applications. 6th Ed. Duxbuy of Thomson Learning. Pacific Grove, CA.

Waterhouse, L. 2010. Partial Year Tagging Models: accounting for tag fouling and
 delayed mixing, Master's thesis. Virginia Institute of Marine Science, The
 College of William and Mary.

Youngs, W.D. and D.S. Robson. 1975. Estimating survival rates from tag returns: model
 test and sample size determination. Journal of the Fisheries Research Board of
 Canada. 32: 2365-2371.

Table 1. Simulation results for estimating the instantaneous rate of fishing mortality in years one through four (F_1, F_2, F_3, F_4), the abnormal instantaneous rate of fishing mortality in part (a) of years two and three (F_{2a}^*, F_{3a}^*), and the instantaneous rate of natural mortality (M). Data were generated with contrasting F_j's across years and with $p = q$, that is with 50% of F_j occurring in part (a) of year j. RMSE is root mean squared error. Mean $\hat{SE}$ refers to the mean of the 2,000 standard errors, calculated by inverting the Hessian. SE of estimates refers to the true standard error, i.e., the standard deviation, of the 2,000 estimates of each parameter. Hoenig et al. refers to the delayed mixing model of Hoenig et al. (1998b) with recaptures tabulated by year and delayed pyt refers to the new model in which recaptures are tabulated by parts of the year. The smaller value for each performance measure for each parameter is in bold.

Model Fitted	Parameter	True Value	Mean estimate	Bias	Mean $\hat{SE}$	SE of estimates	RMSE
Delayed pyt	F_1	0.20	0.200	0.0002	0.023	0.023	0.023
Delayed pyt	F_2	0.30	0.300	**0.0002**	**0.023**	**0.023**	**0.023**
Delayed pyt	F_3	0.25	0.251	**0.0006**	0.021	**0.021**	**0.021**
Delayed pyt	F_4	0.45	0.453	**0.0035**	**0.048**	**0.046**	**0.046**
Delayed pyt	F_{1a}^*	0.05	0.050	-0.0001	0.007	0.007	0.007
Delayed pyt	F_{2a}^*	0.60	0.600	**0.0004**	**0.030**	**0.029**	**0.029**
Delayed pyt	F_{3a}^*	0.15	0.150	**-0.0003**	0.013	**0.013**	**0.013**
Delayed pyt	M	0.20	0.200	**0.0002**	0.028	**0.027**	**0.027**
Hoenig et al.	F_1	0.20	--	--	--	--	--
Hoenig et al.	F_2	0.3	0.300	-0.0004	0.029	0.029	0.029
Hoenig et al.	F_3	0.25	0.252	0.0017	0.030	0.030	0.030
Hoenig et al.	F_4	0.45	0.456	0.0061	0.057	0.055	0.055
Hoenig et al.	$F_{1a}^* = pF_1^*$	0.05	--	--	--	--	--
Hoenig et al.	$F_{2a}^* = pF_2^*$	0.6	0.602	0.0019	0.036	0.035	0.035
Hoenig et al.	$F_{3a}^* = pF_3^*$	0.15	0.150	-0.0004	0.021	0.021	0.021
Hoenig et al.	M	0.2	0.202	0.0016	0.033	0.032	0.032

Table 2. Simulation results for estimating the instantaneous rate of fishing mortality in years one through four (F_1, F_2, F_3, F_4), the abnormal instantaneous rate of fishing mortality in part (a) of years two and three (F_{2a}^*, F_{3a}^*), and the instantaneous rate of natural mortality (M). Data were generated with similar F_j's across years and with $p = q$, that is the proportion of F_j occurring in part (a) of year j equals 50%. RMSE is root mean squared error. Mean $\hat{SE}$ refers to the mean of the 2,000 standard errors, calculated by inverting the Hessian. SE of estimates refers to the true standard error, i.e., the standard deviation, of the 2,000 estimates of each parameter. Hoenig et al. refers to the delayed mixing model of Hoenig et al (1998b) with recaptures tabulated by year and delayed pyt refers to the new model in which recaptures are tabulated by parts of the year. The smaller value for each performance measure for each parameter is in bold.

Model Fitted	Parameter	True Value	Mean estimate	Bias	Mean $\hat{SE}$	SE of estimates	RMSE
Delayed pyt	F_1	0.22	0.220	-0.0002	0.033	0.033	0.033
Delayed pyt	F_2	0.28	0.281	**0.001**	**0.030**	**0.030**	**0.030**
Delayed pyt	F_3	0.25	0.251	**0.001**	**0.030**	0.031	0.031
Delayed pyt	F_4	0.23	0.233	**0.003**	**0.039**	0.040	0.040
Delayed pyt	F_{1a}^*	0.62	0.621	0.0012	0.031	0.031	0.032
Delayed pyt	F_{2a}^*	0.6	0.600	**0.000**	0.031	0.030	0.030
Delayed pyt	F_{3a}^*	0.68	0.680	**0.000**	0.034	0.034	0.034
Delayed pyt	M	0.2	0.199	**-0.001**	0.045	0.044	0.044
Hoenig et al.	F_1	0.22	--	--	--	--	--
Hoenig et al.	F_2	0.28	0.282	0.002	0.043	0.044	0.044
Hoenig et al.	F_3	0.25	0.246	-0.004	0.040	0.040	0.040
Hoenig et al.	F_4	0.23	0.223	-0.007	0.042	0.043	0.043
Hoenig et al.	$F_{1a}^* = pF_1^*$	0.62	--	--	--	--	--
Hoenig et al.	$F_{2a}^* = pF_2^*$	0.6	0.596	-0.004	0.037	0.035	0.035
Hoenig et al.	$F_{3a}^* = pF_3^*$	0.68	0.676	-0.004	0.037	0.037	0.038
Hoenig et al.	M	0.2	0.182	-0.018	0.050	0.050	0.054

Table 3. Simulation results for estimating the instantaneous rate of fishing mortality in years one through four (F_1, F_2, F_3, F_4) and natural mortality (M). Data were generated with contrasting F_j's across years and with $p = q$, that is with 50% of F_j occurring in part (a) of year j. Here $F_j^* = 1.0F_j$, i.e., full mixing. RMSE is root mean squared error. SE of estimates refers to the true standard error, i.e., the standard deviation, of the 10,000 estimates of each parameter. AIC-selected refers to the mean estimate calculated using parameter estimates for the model having the smaller AIC for each simulated dataset. For each parameter, the smallest value for SE of estimates and RMSE is in bold.

Model fitted	Parameter	True Value	Mean Estimate	Bias	SE of estimates	RMSE
Delayed pyt	F1	0.2	0.200	0.0001	0.024	0.024
Delayed pyt	F2	0.3	0.300	-0.0001	0.022	0.022
Delayed pyt	F3	0.25	0.250	0.0000	0.020	0.020
Delayed pyt	F4	0.45	0.452	0.0016	0.047	0.047
Delayed pyt	M	0.2	0.200	-0.0004	0.028	0.028
Fully mixed pyt	F1	0.2	0.200	-0.0001	**0.016**	**0.016**
Fully mixed pyt	F2	0.3	0.300	0.0000	**0.017**	**0.017**
Fully mixed pyt	F3	0.25	0.250	0.0000	**0.015**	**0.015**
Fully mixed pyt	F4	0.45	0.452	0.0016	**0.043**	**0.043**
Fully mixed pyt	M	0.2	0.200	-0.0001	**0.025**	**0.025**
AIC-selected	F1	0.2	0.200	0.0000	0.019	0.019
AIC-selected	F2	0.3	0.300	0.0000	0.018	0.018
AIC-selected	F3	0.25	0.250	0.0001	0.017	0.017
AIC-selected	F4	0.45	0.452	0.0017	0.044	0.044
AIC-selected	M	0.2	0.200	-0.0002	0.026	0.026

Table 4. Simulation results for estimating the instantaneous rate of fishing mortality in years one through four (F_1, F_2, F_3, F_4) and the instantaneous rate of natural mortality (M). Data were generated with contrasting F_j's across years and with $p = q$, that is with 50% of F_j occurring in part (a) of year j. Here $F_j^* = .75F_j$. RMSE is root mean squared error. SE of estimates refers to the true standard error, i.e., the standard deviation, of the 10,000 estimates of each parameter. AIC-selected refers to the mean estimate calculated using the parameter estimates for the model having the smaller AIC for each simulation. For each parameter, the smallest value for SE of estimates and RMSE is in bold.

| Model fitted	Parameter	True Value	Mean Estimate	Bias	SE of estimates	RMSE
Delayed pyt	F1	0.2	0.200	0.0001	0.024	**0.024**
Delayed pyt	F2	0.3	0.300	-0.0001	0.022	**0.022**
Delayed pyt	F3	0.25	0.250	0.0000	0.020	**0.020**
Delayed pyt	F4	0.45	0.452	0.0016	0.047	**0.047**
Delayed pyt	M	0.2	0.200	-0.0004	0.028	**0.028**
Fully mixed pyt	F1	0.2	0.170	-0.0305	**0.014**	0.034
Fully mixed pyt	F2	0.3	0.265	-0.0354	**0.015**	0.038
Fully mixed pyt	F3	0.25	0.220	-0.0297	**0.014**	0.033
Fully mixed pyt	F4	0.45	0.400	-0.0505	**0.039**	0.064
Fully mixed pyt	M	0.2	0.163	-0.0373	**0.026**	0.045
AIC-selected	F1	0.2	0.199	-0.0012	0.024	**0.024**
AIC-selected	F2	0.3	0.299	-0.0012	0.023	0.023
AIC-selected	F3	0.25	0.249	-0.0008	0.020	**0.020**
AIC-selected	F4	0.45	0.450	0.0002	0.048	0.048
AIC-selected	M	0.2	0.199	-0.0014	0.029	0.029

Table 5. Simulation results for estimating the instantaneous rate of fishing mortality in years one through four (F_1, F_2, F_3, F_4) and the instantaneous rate of natural mortality (M). Data were generated with contrasting F_j's across years and with $p = q$, that is with 50% of F_j occurring in part (a) of year j. Here $F_j^* = 1.1 F_j$. RMSE is root mean squared error. SE of estimates refers to the true standard error, i.e., the standard deviation, of the 10,000 estimates of each parameter. AIC-selected refers to the mean estimate calculated using the parameter estimates for the model having the smaller AIC for each simulation. For each parameter, the smallest value for SE of estimates and RMSE is in bold.

Model fitted	Parameter	True Value	Mean Estimate	Bias	SE of estimates	RMSE
Delayed pyt	F1	0.2	0.200	0.0002	0.024	0.024
Delayed pyt	F2	0.3	0.300	-0.0001	0.022	0.022
Delayed pyt	F3	0.25	0.250	0.0001	0.020	0.020
Delayed pyt	F4	0.45	0.452	0.0021	0.047	0.047
Delayed pyt	M	0.2	0.200	-0.0003	0.028	0.028
Fully mixed pyt	F1	0.2	0.212	0.0125	**0.017**	**0.021**
Fully mixed pyt	F2	0.3	0.315	0.0146	**0.017**	0.023
Fully mixed pyt	F3	0.25	0.262	0.0123	**0.016**	0.020
Fully mixed pyt	F4	0.45	0.473	0.0234	**0.045**	0.051
Fully mixed pyt	M	0.2	0.214	0.0137	**0.025**	0.029
AIC-selected	F1	0.2	0.200	-0.0002	0.023	0.023
AIC-selected	F2	0.3	0.300	-0.0002	0.021	**0.021**
AIC-selected	F3	0.25	0.250	0.0004	0.019	**0.019**
AIC-selected	F4	0.45	0.452	0.0018	**0.045**	**0.045**
AIC-selected	M	0.2	0.200	-0.0002	0.027	**0.027**

Figure 1. The percentage of the 10,000 simulations for which the delayed pyt model is selected as a result of having the lower AIC value (solid line) or when the AIC value of the delayed pyt model is at least two smaller than the fully mixed model (dashed line). Both models used data with tag recaptures tabulated by parts of the year. The values of the F_j's can be found in Table 3. The value of F_j^* was specified as a multiple of F_j in order to see how this affected model performance.
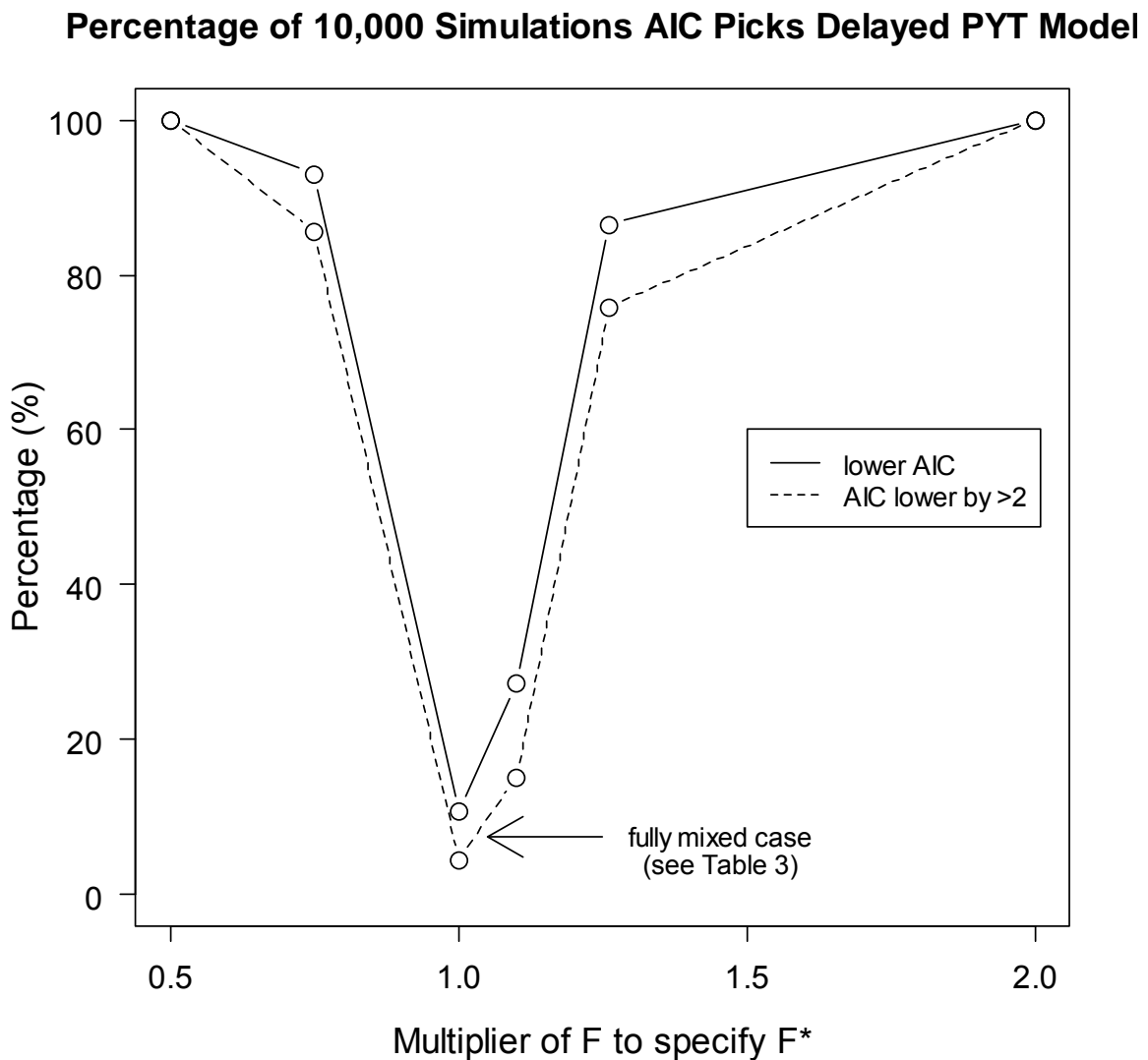
Percentage of 10,000 Simulations AIC Picks Delayed PYT Model

APPENDIX A

Choosing between models 0, 0', and 1: Lake trout example

In practice, when choosing which model best fits the data it is common to use a measure of goodness-of-fit such as Akaike's Information Criterion (AIC) (Anderson 2008). AIC is defined as:

(1) $\quad AIC = -2\log_e(\Lambda) + 2k,$

where Λ is the likelihood function for the model (here, as defined in Thesis Chapter 2) and k is the number of parameters in the model. The model which has the lowest AIC value is considered the best fitting model. Additionally, one can define ΔAIC which is the difference in AIC values between the alternative model and the best fitting model. A general rule of thumb is that a model with a ΔAIC of less than two or three is considered a plausible alternative to the best fitting model (Anderson 2008).

With multinomial count data it is not uncommon to find more variation than can be explained by the model, that is overdispersion, and one can calculate an overdispersion parameter (Anderson 2008). If the overdispersion parameter is large (>1), one can use quasilikelihood AIC (QAIC) to select among alternative models and adjust standard errors.

Since AIC is based on maximum likelihood techniques, an important rule when comparing models is that they must all use the same data. When the data are tabulated by parts of the year, the data are no longer the same as when they are tabulated by full year in that the value of the maximum of the likelihood function is changed. Therefore, when using AIC to select models, the models must all have tag recaptures tabulated on the same time scale.

Here AIC will be used to choose between model 1 (where tag visibility is assumed constant across cohorts in a given year), model 0' (new model, where tag visibility is different from previously tagged cohorts for the first half of the first year a cohort is at liberty), and model 0 (where tag visibility is different from previously tagged cohorts for the entire first year). All model specifications can be found in Thesis Chapter 2. Models 1 and 0 will be reparameterized to reflect tag recaptures by parts of the year rather than full year.

Model 1 tag recaptures by part of the year

The Brownie formulation of model 1 has tag recaptures tabulated on the same periodicity as that of tagging, which is typically a year. If recaptures are tabulated separately for each of two parts of the year, the expectations of model 1 can be expressed:

$$E\left(R_{ijk}\right) = \begin{cases} N_i\, f_{jk} & k = a,\ j = i \\ N_i \displaystyle\prod_{h=i}^{j-1} S_h f_{jk} & k = a,\ j > i \\ N_i S_{jk} f_{jk} & k = b,\ j = i \\ N_i \displaystyle\prod_{h=i}^{j-1} S_h S_{jk} f_{jk} & k = b,\ j > i \end{cases}$$

where k represents the part of the year, $k \in \{a, b\}$, and the other parameters are as defined in Chapter 2. Note there is an implicit category of those tagged animals which are never seen again. The likelihood is constructed from the expected values in the same was as described in chapter 2.

Model 0 tag recaptures by part of the year

Similarly, if recaptures are tabulated separately for each of two parts of the year, the expectations of model 0, in which the tag visibility is different for newly tagged cohorts from previously tagged cohorts for the entire first year, can be expressed:

$$
E(R_{ijk}) = \begin{cases} N_i f_{jk}^{*} & k = a,\ j = i \\[2mm] N_i \prod_{h=i}^{j-1} S_h f_{jk} & k = a,\ j > i \\[2mm] N_i S_{jk} f_{jk}^{*} & k = b,\ j = i \\[2mm] N_i \prod_{h=i}^{j-1} S_h S_{jk} f_{jk} & k = b,\ j > i \end{cases},
$$

where k represents the part of the year, $k \in \{a, b\}$, and the other parameters are as defined in Chapter 2. Note there is an implicit category of those tagged animals which are never seen again. The likelihood is constructed from the expected values in the same was as described in chapter 2.

Lake trout example

Youngs and Robson (1975) described a tagging study of lake trout (*Salvelinus namaycush*) in Cayuga Lake, New York, that included 10 years of tagging and 10 years of recaptures. Following the methods of Hoenig et al. (1998a), I use the first five years of tagging and recaptures in the following examples and assume the product of $\phi\lambda$ is known to be 0.18. These data are believed to represent the case where the tag reporting rate for the first year a cohort is at liberty is equal to the tag reporting rate of previously tagged cohorts within the same year. The dataset was modified in the following ways to create two examples.

The first example consisted of modifying the original dataset (Table 1) from Youngs and Robson (1975) to give recaptures by half year (Table 2). The exploitation rate for the first half of the year using:

$$(2) \; u_{ja} = \frac{F_j}{F_j + M}(1 - \exp(-F_j/2 - M/2)).$$

M and F are divided by two to account for the fact the year was divided into parts (each part being half a year). The values of F_j and M used here (Table 4) were those from fitting the instantaneous rates model that assumes the animals are fully mixed to the original data (tabulation by year) in Table 1. The recaptures in part (a) of the year, r_{ija}, were calculated by taking the yearly recaptures (Table 1) and multiplying by the fraction of the exploitation rate occurring in part (a), i.e., u_{ja} / u_j, and then rounding to the nearest whole number. The recaptures in part (b) of the year, r_{ijb}, were calculated by subtracting the recaptures in part (a) from the total recaptures, r_{ij}, that is,

$$(3) \; r_{ijb} = r_{ij} - r_{ija}.$$

The recaptures tabulated by parts of the year are in Table 2.

For the second example, the data were modified to reflect higher tag visibility (thus higher tag recovery rates) for the first part of the first year each cohort was at liberty. This was done by modifying the tag recaptures by parts of the year (Table 2). For each cohort, the first cell (that is, part (a) recaptures for the first year each cohort is at liberty) was multiplied by 2.5 to reflect a tag visibility that is 2.5 times that of tags that have been at liberty for more than half a year. The cells to the right of the first cell were left alone and the tag recaptures are shown in Table 3.

Three different models were applied to each dataset as follows:

1) Model 1 by parts (Brownie et al. 1978, 1985), in which the tag reporting rate is assumed to be equal for all cohorts within a given year, data are tabulated by parts of the year.

2) Model 0' (Thesis Chapter 2), in which the tag reporting rate for the first half of the first year a cohort is at liberty is different from previously tagged cohorts, data are tabulated by parts of the year

3) Model 0 by parts (Brownie et al. 1978, 1985), in which the tag reporting for the entire first year a cohort is at liberty is assumed to be different from previously tagged cohorts, data are tabulated by parts of the year.

The likelihoods were maximized using the same methods described for the Monte Carlo simulations and the models were evaluated using the same metrics as described for the Monte Carlo simulations (Thesis Chapter 2). The fit of models 0, 0', and 1 were compared using AIC and ΔAIC since they all have recaptures tabulated by parts of the year.

Results

All three models should provide essentially unbiased estimates when data arise from a process where tag visibility is equal across cohorts within a given year. Using AIC as the model selection criteria we see that model 1 has the smallest AIC (Table 5). The overdispersion parameter was calculated for the full model, model 0 by parts, and found to be less than 1, so QAIC did not need to be used. The ΔAIC values for models 0' and 0 are more than 2 AIC units larger, implying that these models are not good

candidates for this observed information. All three models produce similar estimates for the survival rates. Furthermore, model 1 has the smallest standard errors. Model 0' appears to outperform model 0 in that it has a smaller AIC value and has smaller standard errors for each estimate of survival. Additionally, model 0' provides each estimate of survival one year sooner than model 0 (thus, no estimate of S_4 using model 0 with only five years of tagging).

For the second lake trout dataset, which was modified to reflect a higher tag visibility for newly tagged cohorts compared to the previously tagged animals for the first part of the first year each cohort is at liberty, it is expected that both models 0' and 0 will provide unbiased estimates of survival (see Monte Carlo simulation results, Thesis Chapter 2). In fact, model 0' is the most appropriate model as it has the smallest AIC value (Table 6). Again the overdispersion parameter was calculated for the full model, model 0 by parts, and found to be less than one so QAIC was not necessary. The ΔAIC values for models 0 and 1 are both above 2. The estimates of survival rates from model 0 are closer to those from model 0' than the estimates of survival from model 1, which is expected since model 1 does not allow for a change in tag recovery rate as a result of different tag visibilities. Model 0' also has smaller standard errors than model 0. Furthermore, model 0' provides estimates of survival rates one year sooner than with model 0.

References

Anderson, D. R. 2008. Model based inference in the life sciences: a primer on evidence. Springer. New York, NY.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson (1978). Statistical inference from band recovery data: a handbook. 2nd Ed. U.S. Fish and Wildlife Service Resource Publication 131. Washington, D.C.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson (1985). Statistical inference from band recovery data: a handbook. U.S. Fish and Wildlife Service Resource Publication 156. Washington, D.C.

Hoenig, J. M. , N. J. Barrowman, W.S. Hearn, and K.H. Pollock. 1998a. Multiyear Tagging Studies Incorporating Fishing Effort Data. Canadian Journal of Fisheries and Aquatic Sciences. 55:1466-1476.

Youngs, W.D. and D.S. Robson. 1975. Estimating survival rates from tag returns: model test and sample size determination. Journal of the Fisheries Research Board of Canada. 32: 2365-2371.

Table 1. Lake trout (*Salvelinus namaycush*) recapture data published by Youngs and Robson (1975) from Cayuga Lake, New York.

Year tagged	Number Tagged	Year recaptured				
		1	2	3	4	5
1	1048	72	44	8	9	4
2	844		74	30	20	7
3	989			54	48	13
4	971				74	24
5	863					48

Table 2. Modified lake trout (*Salvelinus namaycush*) recapture data from Table 1. Recaptures for each year have been split into two parts of the year, with the recaptures being allocated by the exploitation rate for the portion of the year. It was assumed that half of the instantaneous fishing and natural mortality takes place in each half of the year. The estimate of exploitation rate was calculated using estimates of instantaneous rates of fishing and natural mortality estimated from the first five years of tagging and recaptures (Table 4, column 2).

Year Tagged	Number Tagged	Year of Recapture									
		1a	1b	2a	2b	3a	3b	4a	4b	5a	5b
1	1048	42	30	26	18	5	3	5	4	2	2
2	844			44	30	17	13	12	8	4	3
3	989					31	23	29	19	7	6
4	971							44	30	14	10
5	863									27	21

Table 3. Lake trout (*Salvelinus namaycush*) recapture data originally published by

Youngs and Robson (1975) from Cayuga Lake, New York, modified to reflect a tag

visibility for the first part of the first year each cohort is at liberty that is 2.5 times greater

than the tag visibility of tagged cohorts that have been at liberty for a longer period of

time.

Year Tagged	Number Tagged	Year of Recapture									
		1a	1b	2a	2b	3a	3b	4a	4b	5a	5b
1	1048	105	30	26	18	5	3	5	4	2	2
2	844			110	30	17	13	12	8	4	3
3	989					78	23	29	19	7	6
4	971							110	30	14	10
5	863									68	21

Table 4. Estimates of instantaneous rates of fishing and natural mortality from the fully mixed model for lake trout data published by Youngs and Robson (1975). See Table 1 for cell recapture values.

<u>Estimate</u>

Parameter	Fully Mixed Model
F1	0.568
F2	0.670
F3	0.403
F4	0.628
F5	0.350
M	0.114

Table 5. Estimates and their standard errors, for survival rates from the three models, for lake trout data (Youngs and Robson 1975) modified to have tag recaptures tabulated by parts of year. See Table 2 for cell recapture values. Note that there is no estimate of S_4 for model 0 by parts. This is because only five years of recaptures were used. The bold values indicate the estimates for the model that is presumed to be most appropriate based on lowest AIC. Model 1 p denotes model 1 by parts and model 0 p denotes model 0 by parts.

	Estimate				Estimated Standard Error		
Parameter	Model 1 p	Model 0'	Model 0 p		Model 1 p	Model 0'	Model 0 p
S_1	**.40**	.36	.30		.06	.07	.08
S_2	**.53**	.54	.56		.08	.10	.12
S_3	**.60**	.60	.59		.09	.12	.18
S_4	**.44**	.44	--		.09	.14	--
AIC	**4944**	4951	4956				
ΔAIC	0	7	12				

Table 6. Estimates and their standard errors, for survival rates from the three models, for lake trout data (Youngs and Robson 1975) modified to reflect a tag visibility for the first part of the first year a cohort is at liberty that is 2.5 times that of the tag visibility for cohorts that have been at liberty longer, with tag recaptures tabulated by parts of the year. See Table 3 for cell recapture values. Note that there is no estimate of S_4 for model 0 by parts. This is because only five years of recaptures were used. The bold values indicate the estimates for the model that is presumed to be most appropriate based on lowest AIC. Model 1 p denotes model 1 by parts and model 0 p denotes model 0 by parts.

Parameter	Estimate			Estimated Standard Error		
	Model 1 p	Model 0'	Model 0 p	Model 1 p	Model 0'	Model 0 p
S_1	.27	**.36**	.30	.04	.07	.08
S_2	.42	**.54**	.56	.06	.10	.12
S_3	.41	**.60**	.59	.05	.12	.18
S_4	.30	**.44**	--	.05	.14	--
AIC	6398	**6375**	6379			
ΔAIC	23	0	4			

APPENDIX B

Choosing between an instantaneous rates model with full mixing and delayed

mixing: Lake trout example

In practice, when choosing which model best fits the data it is common to use a

measure of goodness-of-fit such as Akaike's Information Criterion (AIC) (Anderson

2008). AIC is defined as:

(1) $\quad AIC = -2\log_e(\Lambda) + 2k,$

where Λ is the likelihood function for the model (here, as defined in Thesis Chapter 3)

and k is the number of parameters in the model. The model which has the lowest AIC

value is considered the best fitting model. Additionally, one can define ΔAIC which is

the difference in AIC values between the alternative model and the best fitting model. A

general rule of thumb is that a model with a ΔAIC of less than two or three is considered

a plausible alternative to the best fitting model (Anderson 2008).

With multinomial count data it is not uncommon to find more variation than can

be explained by the model, that is overdispersion, and one can calculate an

overdispersion parameter (Anderson 2008). If the overdispersion parameter is large (>1),

one can use quasilikelihood AIC (QAIC) to select among alternative models and adjust

standard errors.

Since AIC is based on maximum likelihood techniques, an important rule when

comparing models is that they must all use the same data. When the data are tabulated by

parts of the year, the data are no longer the same as when they are tabulated by full year

in that the value of the maximum of the likelihood function is changed. Therefore, when

using AIC to select models, the models must all have tag recaptures tabulated on the same time scale.

Here AIC will be used to choose between the fully mixed model (described in Thesis Chapter 3) and the model with delayed mixing lasting part of the year of Hoenig et al. (1998a, b) (described in Thesis Chapter 3). Additionally, the partial year tabulation for delayed mixing lasting part of year model (delayed pyt model) will be fit to the data, but this model will not be compared in terms of AIC since the tag recaptures are tabulated by parts of the year rather than full year. All model specifications can be found in Thesis Chapter 3.

Lake trout data example

Youngs and Robson (1975) described a tagging study of lake trout (*Salvelinus namaycush*) in Cayuga Lake, New York, that included 10 years of tagging and 10 years of recaptures. Following the methods of Hoenig et al. (1998b), I use the first five years of tagging and recaptures in the following examples and assume the product of $\phi\lambda$ is known to be 0.18. The dataset was modified in the following ways to create two examples.

The first example consisted of using the original dataset (Table 1) from Youngs and Robson (1975) as well as modifying the data to give recaptures by half year (Table 2). These data are believed to represent full mixing of newly tagged animals into the population at large (Hoenig et al. 1998b). The data were split into parts of the year by calculating the exploitation rate for the first half of the year using:

$$(2) \quad u_{ja} = \frac{F_j}{F_j + M}(1 - \exp(-F_j/2 - M/2)).$$

M and *F* are divided by two to account for the fact the year was divided into parts (each part being half a year). The values of F_j and *M* used above (Table 4, column 2) were those from fitting the instantaneous rates model that assumes the animals are fully mixed to the original data (tabulation by year) in Table 1. The recaptures in part (a) of the year, r_{ija}, were calculated by taking the yearly recaptures (Table 1) and multiplying by the fraction of the exploitation rate occurring in part (a), i.e., u_{ja} / u_j, and then rounding to the nearest whole number. The recaptures in part (b) of the year, r_{ijb}, were calculated by subtracting the recaptures in part (a) from the total recaptures, r_{ij}, that is,

(3) $r_{ijb} = r_{ij} - r_{ija}$.

The recaptures tabulated by parts of the year are in Table 2.

For the second example, the data were modified to reflect delayed mixing lasting half of a year, following methods similar to those of Hoenig et al. (1998b). Using the modified lake trout dataset, which is assumed to represent full mixing, with the year split in parts (Table 2), the recaptures for the first part of the first year each cohort is at liberty (the r_{iia} recapture cells) were adjusted to two thirds of the original value to reflect a lower exploitation rate as a result of delayed mixing. The cells to the right of the first cell of recaptures for each cohort had to be adjusted to reflect this increase in survival rate (a decrease in exploitation is equivalent to an increase in survival). By adjusting for this increase in survival rate the estimates of the F_j's, the normal rates of fishing mortality, should remain roughly the same as those obtained with the fully mixed model fitted to the original data. The estimates of F_j and *M* (Table 4), from fitting the fully mixed model, were used to calculate the exploitation rate for the first part of the first year each cohort was at liberty (Equation 2). The abnormal instantaneous rate of fishing mortality for part

102

(a) of the year, $0.5F_i*$, which results in a two-thirds reduction in the calculated exploitation rate was solved iteratively. That is, the left hand side of equation (2) was multiplied by two-thirds and then the value of F_j* that made the right hand side equal to the left was solved for iteratively (with M held constant). The cells to the right were adjusted, to reflect the increase in the survival rate, by multiplying each recapture cell (the cells to the right of the r_{iia} recapture cell), r_{ijk}, by $exp(0.5F_i - 0.5Fi*)$ where F_i and F_i* are from the fully mixed model and the iterative search, respectively. The recaptures were rounded to the nearest whole number and are shown in Table 3. Parts of the year were added together to give recaptures over the whole year.

Three different models were applied to each dataset as follows:

4) Fully mixed model, as described by Hoenig et al. (1998a), data are tabulated by year. This model is an instantaneous rates model which assumes the newly tagged animals fully mix with the untagged and previously tagged population before the fishing takes place.

5) Hoenig et al. model, as described by Hoenig et al. (1998b), in which mixing is assumed after half a year; data are tabulated by year.

6) Delayed pyt model, as described in Thesis Chapter 3, in which mixing is assumed after half a year; data are tabulated by half year.

The likelihoods were maximized using the same methods described for the Monte Carlo simulations (Thesis Chapter 3). The models were evaluated using the same metrics as described for the Monte Carlo simulations (Thesis Chapter 3), including estimates of the parameters for each dataset and estimated standard errors. The fit of the fully mixed model and the Hoenig et al. model were compared using AIC and ΔAIC since they both

have recaptures tabulated by full year. No AIC or ΔAIC analysis was done for the delayed pyt model since tag recaptures are by part of the year.

Results

All three models should provide essentially unbiased estimates when data arise from a process where all cohorts are fully mixed into the population at large immediately after tagging. Under complete mixing, the fully mixed model should be the most efficient. In fact, when the models are fitted to data presumed to represent full mixing (Tables 1, 2), all three models produce similar estimates with the fully mixed model producing the smallest estimated standard errors for all parameters (Table 4). The fully mixed model has the lowest AIC and the Hoenig et al. model is five ΔAIC units away, meaning the only model that should really be considered of the two is the fully mixed model. Interestingly, the delayed pyt model has smaller estimated standard errors than the Hoenig et al. model (1998b) for each parameter. As expected, the estimates of F_j* are approximately equal to the estimates of the corresponding F_j for the new model and the delayed mixing model of Hoenig et al. (1998b) since the data presumably represent full mixing. For eight of the ten parameters, the new model provides estimates closer to the estimates from the fully mixed model which are presumed to be closest to the truth. The new model outperforms the delayed mixing model of Hoenig et al. (1998b) by better reproducing the results of the fully mixed model, yielding smaller estimated standard errors for all parameters, and providing an estimate for F_1 (which the delayed mixing model does not provide).

For the second modified lake trout dataset (Table 3), where delayed mixing occurs for the first part of the first year each cohort is at liberty, both the new model and the delayed mixing model of Hoenig et al. (1998b) should yield unbiased parameter estimates. As expected, the fully mixed model's estimates of F_1, F_2, F_3, F_4, F_5, and M (Table 5) are different from those of the other two models, as the model incorrectly assumes for the first year at liberty a newly tagged cohort experiences the same instantaneous rate of fishing mortality (and thus survival) as previously tagged cohorts. The delayed pyt model's and Hoenig et al. model's estimates of F_1, F_2, F_3, F_4, F_5, and M are approximately equal to the values expected (bold values in Table 5 correspond to bold values in Table 4). These values were expected to be equal because the data were modified (decrease in exploitation for first half of first year and cells to the right increased to reflect higher survival) in such a way that the normal rates of fishing mortality and natural mortality (F's and M) should be equal to their values from the fully mixed case (Table 4). Note that the fully mixed model yields estimates with the smallest standard errors, so if the estimates were believed to be accurate this would result in one assuming high precision for incorrect estimates. The Hoenig et al. model has the lower AIC value, and the ΔAIC value for the fully mixed model is four units, which means of the two models only the Hoenig et al. model should be considered. As with the previous lake trout example, the fully mixed model provides the smallest estimates of standard errors for all parameters, followed by the delayed pyt model, and the Hoenig et al. model yields the largest estimated standard errors for all parameters. The delayed pyt model outperforms the Hoenig et al. model by better reproducing the expected parameter estimates (bold values in Table 4).

References

Anderson, D. R. 2008. Model based inference in the life sciences: a primer on evidence. Springer. New York, NY.

Hoenig, J. M. , N. J. Barrowman, W.S. Hearn, and K.H. Pollock. 1998a. Multiyear Tagging Studies Incorporating Fishing Effort Data. Canadian Journal of Fisheries and Aquatic Sciences. 55:1466-1476.

Hoenig, J. M. , N. J. Barrowman, K.H. Pollock, E.N. Brooks, W.S. Hearn and T. Polacheck. 1998b. Models for Tagging Data That Allow for Incomplete Mixing of Newly Tagged Animals. Canadian Journal of Fisheries and Aquatic Sciences. 55:1477-1483.

Youngs, W.D. and D.S. Robson. 1975. Estimating survival rates from tag returns: model test and sample size determination. Journal of the Fisheries Research Board of Canada. 32: 2365-2371.

Table 1. Lake trout (*Salvelinus namaycush*) recapture data published by Youngs and Robson (1975) from Cayuga Lake, New York.

Year tagged	Number Tagged	Year recaptured				
		1	2	3	4	5
1	1048	72	44	8	9	4
2	844		74	30	20	7
3	989			54	48	13
4	971				74	24
5	863					48

Table 2. Modified lake trout (*Salvelinus namaycush*) recapture data from Table 1. Recaptures for each year have been split into two parts of the year, with the recaptures being allocated by the exploitation rate for the portion of the year. It was assumed that half of the instantaneous fishing and natural mortality takes place in each half of the year. The estimate of exploitation rate was calculated using estimates of instantaneous rates of fishing and natural mortality estimated from the first five years of tagging and recaptures (Table 4, column 2).

Year Tagged	Number Tagged	1a	1b	2a	2b	3a	3b	4a	4b	5a	5b
1	1048	42	30	26	18	5	3	5	4	2	2
2	844			44	30	17	13	12	8	4	3
3	989					31	23	29	19	7	6
4	971							44	30	14	10
5	863									27	21

Table 3. Lake trout (*Salvelinus namaycush*) recapture data originally published by Youngs and Robson (1975) from Cayuga Lake, New York, modified using methods similar to those of Hoenig et al. (1998b) to show delayed mixing of each cohort for the first part of the first year at liberty. Recaptures from parts of the year, (a) and (b), can be added together for yearly recaptures.

Year Tagged	Number Tagged	1a	1b	2a	2b	3a	3b	4a	4b	5a	5b
						Year of Recapture					
1	1048	28	33	30	20	5	4	1	4	2	2
2	844			30	34	18	14	13	9	4	3
3	989					20	25	32	22	8	6
4	971							30	34	14	11
5	863									18	22

Table 4. Estimates and their standard errors, for instantaneous rates of fishing and natural mortality from the three models, for lake trout data published by Youngs and Robson (1975). See Tables 1 and 2 for cell recapture values. Note that the delayed mixing model of Hoenig et al. (1998b) for recaptures tabulated by year does not allow for an estimate of F_1. Also note that there are no estimates of the $F_j*'s$ from the fully mixed model because these parameters do not exist in the model. The bold values indicate the estimates for the model that are presumed to be most appropriate.

	Estimate			Standard error		
Parameter	Fully Mixed Model	Hoenig et al. Model	Delayed pyt Model	Fully Mixed Model	Hoenig et al. Model	Delayed pyt Model
F1	**0.568**	-----	0.569	0.076	-----	0.110
F2	**0.670**	0.698	0.696	0.071	0.121	0.091
F3	**0.403**	0.449	0.414	0.048	0.094	0.060
F4	**0.628**	0.693	0.641	0.078	0.172	0.093
F5	**0.350**	0.414	0.391	0.050	0.146	0.076
M	**0.114**	0.135	0.132	0.040	0.055	0.043
F1*	-----	0.568	0.565	-----	0.077	0.075
F2*	-----	0.667	0.663	-----	0.099	0.070
F3*	-----	0.392	0.407	-----	0.064	0.050
F4*	-----	0.602	0.629	-----	0.095	0.076
F5*	-----	0.399	0.393	-----	0.069	0.060
AIC	4216	4221	ΔAIC	0	5	

Table 5. Estimates and their standard errors, for instantaneous rates of fishing and natural mortality from the three models, for lake trout data modified to represent delayed mixing during the first part of the first year each cohort is at liberty. See Table 3 for cell recapture values. Note that the delayed mixing model of Hoenig et al. (1998b) does not allow for an estimate of F_1. Also note that there are no estimates of the F_j*'s from the fully mixed model because these parameters do not exist in the model. The bold values indicate the values that are presumed to be the best estimates since they are approximately equal to the estimates from the unmodified data (see Table 4).

	Estimate			Standard error		
Parameter	Fully Mixed Model	Hoenig et al. Model	Delayed pyt Model	Fully Mixed Model	Hoenig et al. Model	Delayed pyt Model
F1	0.462	-----	**0.588**	0.066	-----	0.108
F2	0.610	**0.798**	**0.761**	0.066	0.130	0.094
F3	0.345	**0.505**	**0.429**	0.042	0.098	0.060
F4	0.544	**0.796**	**0.651**	0.067	0.187	0.089
F5	0.298	**0.495**	**0.390**	0.042	0.175	0.072
M	0.112	**0.158**	**0.149**	0.039	0.052	0.042
F1*	-----	0.480	0.476	-----	0.069	0.067
F2*	-----	0.550	0.587	-----	0.084	0.063
F3*	-----	0.315	0.337	-----	0.054	0.042
F4*	-----	0.508	0.526	-----	0.082	0.064
F5*	-----	0.325	0.323	-----	0.060	0.051
AIC	4054	4050	ΔAIC	4	0	

Thesis References

Anderson, D. R. 2008. Model based inference in the life sciences: a primer on evidence. Springer. New York, NY.

Beverton, J. H. and S. J. Holt. 1957. On the dynamics of exploited fish populations. Chapman and Hall. London, Great Britain.

Brownie, C. , Anderson, D. R. , Burnham, K. P. , and D. S. Robson. 1978. Statistical inference from band recovery data: a handbook. 2nd Ed. U.S. Fish and Wildlife Service Resource Publication 131. Washington, D.C.

Brownie, C. , Anderson, D. R. , Burnham, K. P. , and D. S. Robson. 1985. Statistical inference from band recovery data: a handbook. U.S. Fish and Wildlife Service Resource Publication 156. Washington, D.C.

Burnham, K. P. and D. R. Anderson. 2002. Model Selection and Mutimodel inference. A practical Information-Theoretic Approach. 2nd Ed. Springer Science and Business Media, LLC. USA.

Chasalow, S. 2005. combinat: combinatorics utilities. R package version 0.0-6. Available: http://cran.r-project.org/web/packages/combinat/index.html. (July 2010).

Curtis, J.M.R. 2006. Visible Implant Elastomer Color Determination, Tag Visibility, and Tag Loss: Potential Sources of Error for Mark–Recapture Studies. North American Journal of Fisheries Management. 26: 327-337.

Dicken, M.L., A.K. Booth, and M.J. Smale. 2006. Preliminary observations of tag shedding, tag reporting, tag wounds, and tag biofouling for raggedtooth sharks (*Carcharias taurus*) tagged off the east coast of South Africa. ICES Journal of

Marine Science. 63: 1640-1648.

Hoenig, J. M. , N. J. Barrowman, W.S. Hearn, and K.H. Pollock. 1998a. Multiyear Tagging Studies Incorporating Fishing Effort Data. Canadian Journal of Fisheries and Aquatic Sciences. 55:1466-1476.

Hoenig, J. M. , N. J. Barrowman, K.H. Pollock, E.N. Brooks, W.S. Hearn and T. Polacheck. 1998b. Models for Tagging Data That Allow for Incomplete Mixing of Newly Tagged Animals. Canadian Journal of Fisheries and Aquatic Sciences. 55:1477-1483.

Hogg, R. V., McKean, J. W. , and A. T. Craig. 2005. Introduction to Mathematical Statistics. 6th Ed. Pearson Prentice Hall. Upper Saddle River, New Jersey.

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

Lowry, M. B. and I. M. Suthers. 1998. Home range, activity and distribution patterns of a temperate rocky-reef fish, *Cheilodactylus fuscus*. Marine Biology. 132(4): 569-578.

Medley, P. A. H. and C. H. Ninnes. 1999. A stock assessment for the conch (*Strombus gigas* L.) fishery in the Turks and Caicos Islands. Bulletin of Marine Science. 64(3): 399-406.

Pollock, K. H. and D. G. Raveling. 1982. Assumptions of Modern Band-Recovery Models, with Emphasis on Heterogeneous Survival Rates. The Journal of Wildlife Management. 46(1):88-98.

Pollock, K.H., J.M. Hoenig, and C.M. Jones. 1991. Estimation of Fishing and Natural Mortality Rates when a Tagging Study is Combined with a Creel Survey or Port Sampling. American Fisheries Society Symposium 12:423-434.

R Development Core Team. 2008. 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.

Reeves, K. S. and D. L. Buckmeier. 2009. Mortality, Predation, and Tag Visibility of Fish Marked with Visible Implant Elastomer Tags. North American Journal of Fisheries Management. 29: 323-329.

Ricker, W. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin of the fisheries research board of Canada. 191. Department of environment fisheries and marine service, Ottawa, Canada.

Robson, D.S., and Youngs, W.D. 1971. Statistical analysis of reported tag recaptures in the harvest from an exploited population. Biometrics Unit Mimeographic Series Paper No. BU-369-M. Cornell University, Ithaca, N.Y.

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

Seber, G.A.F. 1970. Estimating time-specific survival and reporting rates for adult birds from band returns. Biometrika, 57: 313-318.

Seber, G. A. F. 2002. The Estimation of Animal Abundance and related parameters. 2nd Ed. The Blackburn Press. Caldwell, NJ.

SEDAR (Southeast Data, Assessment, and Review). 2007. Caribbean Queen Conch. SEDAR 14, Stock Assessment Report 3. North Charleston, South Carolina.

Tarbath, D. B. 1999. Estimates of growth and natural mortality of the blacklip abalone

    (*Haliotis rubra*) in Tasmania. Marine Research Laboratories, Tasmanian

    Aquaculture and Fisheries Institute. Technical Report No. 3, University of

    Tasmania, Tasmania. Available:

    http://www.tafi.org.au/publications/TechnicalReport_3.pdf. (July 2010).

Verweij, M. C. and I. Nagelkerken. 2007. Short and long term movement and site

    fidelity of juvenile Haemulidae in back-reef habitats of a Caribbean embayment.

    Hydrobiologia. 592(1):257-270.

Wackerly, D.D., W. Mendenhall, R.L. Scheaffer. 2002. Mathematical Statistics with

    Applications. 6th Ed. Duxbuy of Thomson Learning. Pacific Grove, CA.

Waterhouse, L. 2010. Partial Year Tagging Models: accounting for tag fouling and

    delayed mixing, Master's thesis. Virginia Institute of Marine Science, The

    College of William and Mary.

White, G. C. 1983. Numerical Estimation of Survival Rates from Band-Recovery and

    Biotelemetry Data. The Journal of Wildlife Management. 47(3):716-728.

White, G. C. 1992. PC SURVIV User's Manual. Dept. of Fishery and Wildlife Biology,

    Colorado State University, Fort Collins, CO. Available: http://www.mbr-

    pwrc.usgs.gov/software/survive.html. (July 2010).

Youngs, W.D. and D.S. Robson. 1975. Estimating survival rates from tag returns: model

    test and sample size determination. Journal of the Fisheries Research Board of

    Canada. 32: 2365-2371.

VITA

Lynn Waterhouse

Born in Nashville, Tennessee on 5 January 1985. Graduated from The Miami Valley High School in 2003. Earned presidential scholarship and membership in John W. *Berry* Sr. *Scholars* Programs at University of Dayton. Graduated with a B.S. in biology with minor in economics from the University of Dayton in 2007. Entered the Master of Science program at the School of Marine Science, Virginia Institute of Marine Science, College of William & Mary in 2007. Received a VIMS fellowship.