MODELING UNCERTAINTY AND HETEROGENEITY IN MARK-RECAPTURE AND OCCUPANCY EXPERIMENTS

by

Wendell O. Challenger
M.Sc., Queen's University, 2004

a Thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
in the Department
of
Statistics and Actuarial Science

SIMON FRASER UNIVERSITY
Fall 2010

APPROVAL

Name:	Wendell O. Challenger
Degree:	Doctor of Philosophy
Title of Thesis:	Modeling Uncertainty and Heterogeneity in Mark-Recapture and Occupancy Experiments

Examining Committee: Dr. Derek Bingham
Chair

Dr. Carl J. Schwarz, Senior Supervisor

Dr. Brad McNeney, Supervisor

Dr. Rick Routledge, Supervisor

Dr. Tim Swartz, SFU Examiner

Dr. Roger Pradel, External Examiner,
Research Director at CNRS,
Center for Evolutionary and Functional Ecology

Date Approved: October 7, 2010

SFU SIMON FRASER UNIVERSITY
LIBRARY

Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website <www.lib.sfu.ca> at: <http://ir.lib.sfu.ca/handle/1892/112>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

Last revision: Spring 09

Abstract

Ecological data sets often present problems such as detection, heterogeneity and uncertainty. Capture-recapture and occupancy frameworks deal with issues of detection for individuals and species respectively. Although developed separately, both share commonalities that are highlighted in this thesis. We present two developments in capture-recapture and one in occupancy.

In capture-recapture experiments parameter heterogeneity is often dealt with by stratifying the population in advance (e.g. by sex), but strata assignment may not always be possible. For example in the MilleLacs fishery, catachability is known to differ by sex, but sex could not be determined on all sampling occasions leaving some individuals with an "unknown" sex designation. This heterogeneity can severely bias estimates of abundance if the data are simply pooled and treated as coming from a single large population. In the first development the super-population approach is extended to handle uncertainty in strata assignments.

Heterogeneity in catchability can also be result when animals do not behave in a completely independent fashion. The assumption of independence is a long-standing assumption of capture-recapture models and is known almost never to hold. We take the first steps to relaxing this assumption by modeling the dependency in capture outcomes between pair-bonded individuals. Animals in a pair-bond remain in close proximity to one another, and if one member of the pair is observed, we are also more likely to observe the other member. We estimate the "correlation" (termed rho) in capture events using simulated data, as the motivating example (a harlequin duck study) was more complicated than originally expected.

In the final development we highlight the similarities between capture-recapture and occupancy models by rewriting occupancy models to use the capture-recapture multievent approach. In this development we also provide design protocols for permanent and temporary site monitoring and provide a power analysis comparing the efficiency of both protocols to detect changes in relative abundance. Finally, we provide a complicated multi-year analysis of the anuran frog data set to demonstrate the ability of the proposed framework to handle complex biological problems.

Keywords: Abundance estimation; Capture-mark-recapture; Jolly-Seber; Cormack-Jolly-Seber; Occupancy; Latent states; Heterogeneity; Uncertainty; State-space; Statistical power

I would like to dedicate this thesis to my father, Barry Challenger (1942-2009).

“Statistical thinking will one day be as necessary a qualification for efficient citizenship as the ability to read and write.”

—H.G. WELLS

Acknowledgments

I am indebted to many people and organizations over the course of my Ph.D. First I would like to thank my friends and family, who put up with many late nights solving problems, my forgetfulness (which at times reached dizzying heights) and my occasional bouts of grumpiness when handling stubborn problems. This especially extends to my wife Karen who is a rock and encouraged me regularly to get out and on my bike, keeping me sane in the final stages.

On a slightly more serious note I am also indebted to my supervisor Dr. Carl J. Schwarz who unquestioningly took a student with a strong background in biology, but with understanding of theoretical statistics. If we are to proceed into the new age of interdisciplinary science, we need more people with Carl's perspective.

I also would to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Mathematics of Information Technology and Complex Systems (MITACS) Accelerate program for providing funding in terms of scholarships. I would not have been able to complete my Ph.D. without it. More specifically, the super-population was funded by an NSERC PGS A scholarship and the ρ-CJS and multi-state occupancy work were funded by separate MITACS internships with BigHorn WildLife Technologies Ltd. and the Vancouver Aquarium respectively. I learned a lot during my internships with Beth MacCallum and Jeff Marliave.

Finally, I would like to thank everyone in the Department of Statistics and Actuarial Science at Simon Fraser University. This especially includes all my fellow graduate students who were always friendly and more than willing to lend a hand. It is an amazing department and I would encourage anyone considering a career in statistics to consider spending time here.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Modeling ecological data presents a number of interesting statistical challenges. It has been argued that ecology is a science of mechanisms (Krebs, 1991). Statistical models have been viewed as a tool for exploring and understanding the processes affecting wildlife populations (Nichols, 2001) and carry the meaning of science (Giere, 2004). Models that incorporate plausible ecological mechanism are generally preferred over models that adequately fit the data, but are not ecological interpretable (Taper et al., 2008). Yet, ecological data often violates assumptions of most standard statistical techniques. Data tends not to be normally distributed and can be over-dispersed relative to a number of standard statistical approaches (e.g. generalized linear models; GLM).

One of the most pervasive problems with ecological data is our general inability to observe animals in the field, a problem commonly termed "detectability." Animals are oftentimes sparsely distributed across a landscape making the probability of encounter low. Further adding to this difficulty, they often possess natural camouflage and exhibit behavioural patterns (e.g. nocturnal foraging) that can further add to their cryptic nature. In fact our ability to reliably detect the simplest of ecological measures, the presence or absence of a species, is quite limited (Craig and Roberts, 2001; Lindenmayer et al., 2001; MacKenzie et al., 2002; Tyre et al., 2003).

If not handled correctly, issues of detection will lead to bias in estimates in even the simplest of experiments. Suppose a researcher is investigating the relationship between brood size (number of offspring) and survivorship for adult females of a particular species of bird. The researcher may opt for a natural experiment, looking at relationship between survival and naturally occurring brood sizes. Alternatively, the researcher may take a more experimental approach by manipulating brood sizes directly to control for parental quality. While experimental manipulations are preferable in terms of statistical inference, both experimental approaches can suffer from the issue of detection.

The researcher, naïve to issues of detection, would most likely determine survival by counting the number of adult females present at the beginning and end of the breeding season. The resulting difference in counts could then be modeled as a binomial experiment, with survivorship as the probability of a "success," allowing for standard GLM techniques to be employed. Unfortunately, in doing so, the researcher has assumed birds that were not observed at the end of the breed season did not survive, when in fact they may have survived and were simply missed, or they may have survived and emigrated from the area. Any of these possibilities will bias estimates of the ecological process of interest (survival).

In response to these difficulties a number of specialized techniques have evolved that directly model the detection process along with ecological processes of interest. The two methods in this dissertation (mark-recapture and occupancy), use repeated sampling schemes in order to estimate the degree that "detectability" affects the data. In both cases a series of indicator variables (i.e. '1' and '0') are used to describe observations on each of the sampling occasions.

Mark-recapture, as its name implies, involves marking animals with unique tags, releasing the captured animal back into the population, then attempting to recapture the same marked animals at later sampling occasions. This produces a series of 1's and 0's, termed a capture history, indicating if an individual was caught on a specific sampling occasion. For example, a capture history of `101` indicates an experiment with three sampling occasions, where the individual was captured on the first occasion, tagged and released, missed on the second and captured again on the third. A capture history of `100` indicates the individual was captured and released on the first occasion, but not recaptured in any further sampling occasions. For more complex experiments, the indicator variable '1' can be replaced by codes to indicate various states or groupings (i.e. breeding status, sex, etc).

The occupancy framework operates on a similar premise, but rather than uniquely tracking individuals, the framework follows predefined geographical areas, termed sites, over successive sampling occasions. Here a '1' indicates that the species of interest was observed on the site, while a '0' would indicate the species was not observed. Depending on assumptions, the '0' can indicate that the site was either unoccupied, or occupied but the species was missed. For an encounter history of 101, if we can assume the site was closed to changes in occupancy (e.g. the sampling periods were close enough together in time), then the ‘0’ observation has a similar meaning to the mark-recapture interpretation, in that the species was there but was missed.

The unique capture and encounter histories can then be modeled as categorical outcomes in a multinomial distribution. The power of these techniques emerges as we build ecologically relevant models to describe each of the possible unique encounter or capture outcomes in the multinomial distribution. Commonly, a shared set of parameters and covariates across the population, or groups within, is used to represent various sampling and ecological processes hypothesized to occur. For example, in modeling a ‘0’ occurrence a mark-recapture experiment could mean the animal could have died (survivorship), emigrated or was available for capture but simply missed on a given occasion (capture) depending on the sampling history before and after that zero. By using different parameters to represent each of these possibilities, we are able to define complex probability statements to describe these processes.

The primary focus of this dissertation has been to provide additional practical tools for researchers to answer ecological questions. To date there has been an extensive development of methods based on the direct sampling of individuals such as mark-recapture, while other frameworks such as occupancy are still newly emerging. This thesis will focus on extending the occupancy framework to handle a variety of ecological states, other than the presence or absence of a particular species or set of species. This includes relative measures of abundance, breeding status, and multi-species occupancy all within the framework of a multi-year study. For mark-recapture, we focus on dealing with additional classification uncertainty in abundance estimates and relaxing assumptions of independent fates among individuals. The latter can especially be important for a number of avian species, such as migratory birds, that exhibit strong pair-bonding and frequently have shared fates.

Part I

Mark-Recapture Models

Chapter 2

Introduction to Part I of the Thesis

In this first part of the dissertation, I present extensions made to existing models in the mark-recapture framework. Within the mark-recapture framework my focus has been on extending open population models to handle issues of parameter heterogeneity, incorporating additional types of uncertainty that are known to occur in the data, and taking the first steps to relax a long standing assumption of independence among individuals.

Generally, open population models are the most commonly employed models to analyze mark-recapture experiments. Unlike closed models the population size is allowed to change during sampling (e.g. by processes such as birth/immigration and death/emigration). These models are an important method for estimating demographic parameters which provides biologists in-depth knowledge about the driving forces behind changes in population size.

Within open populations models, the Jolly-Seber (JS) and the Cormack-Jolly-Seber (CJS) are the two predominant modeling frameworks. The JS models consider both births and deaths (survival), while the CJS framework only models survival by conditioning on first capture. By making assumptions on the underlying catchability of unobserved individuals the JS framework also allows for estimates of abundance. While abundance estimates are of value, earlier formulations estimated population size only as derived parameters and did not explicitly include parameters for abundance in the model likelihood.

Within both frameworks I have been interested in relaxing modeling assumptions so that we may better model the true nature of the collected data. In Chapter 3 extend the super-population formulation (JS framework) of Schwarz and Arnason (1996) to allow for uncertainties in group assignments. This type of work has already been accomplished in the CJS framework, but the JS framework presents unique challenges as we need to account for group membership upon entry into the population, which the CJS framework does not consider due to conditioning of first capture. The work was published as Challenger and Schwarz (2009) in the peer-reviewed conference proceedings *Modeling demographic processes in marked populations* which was compiled from papers of the 2007 European Bird Ringing Technical Meeting and is the third volume in the *Environmental and Ecological Statistics* series.

Within the CJS open population framework in Chapter 4 I begin the first work to relax the longstanding assumption that all individuals in the population behave in an independent manner. Specifically, I consider the dependency between mating pairs in the form of pairwise dependent capture outcomes. I term this the ρ-CJS framework.

Chapter 3

Mark-recapture Jolly-Seber Abundance Estimation with Classification Uncertainty[1]

3.1 Introduction

The capture-recapture (CR) experimental protocol provides a flexible approach for making inferences on animal populations that may be hard to observe otherwise. Inferences can range from individual level processes, such as survivorship, to population level processes such as abundance and recruitment. The Jolly-Seber (JS) model is used for open populations, where both births (or immigration) and deaths (or emigration) are explicitly modeled. Current applications of the JS abundance model include estimating overall and group specific abundances.

The JS class of models need to make key assumptions on the marked and unmarked animals in the population of interest. Specifically, both segments of the population are assumed to be homogeneous in catchability and demographic parameters such as survivorship. This assumption of homogeneity is not only required at the population level but also at the individual level, where parameters are assumed to be the same among animals. However, populations often have underlying structures that can violate the latter assumption of animal homogeneity. For example, animals with differing age and/or sex may exhibit

---

[1] Published as Challenger and Schwarz (2009).

differences in catchabilities and/or different apparent survivorship.

If left unmodeled the heterogeneity may result in over-dispersion, model misfit, and potentially biased estimates. Heterogeneity can be handled by modeling affected parameters in terms of individual covariates or, at a coarser level, by modeling the heterogeneity through discrete partitions. When, over the course of the experiment, membership to a partition is fixed (e.g. sex), the term "group" is often used, while the term "state" often implies the potential for membership can change. For systems where membership to a partition is fixed (i.e. groups), classical parameter grouping methods (e.g. Jolly, 1965; Lebreton et al., 1992) are often used, while for systems where membership can change (non-fixed states) are handled by multi-state (e.g. Lebreton and Pradel, 2002) and multievent models (Pradel, 2005). These approaches not only assume that heterogeneity may be effectively partitioned, but that the state/group can also be accurately assessed on capture. Therefore, the ability to assign state/group represents an additional source of variability, which up until recently has not been included in many models.

State/group assignment uncertainty can be divided into two distinct types: identification uncertainty and mis-assignment uncertainty. Identification uncertainty represents a probability of identifying a state/group when observed, while mis-assignment is the probability of assigning the incorrect state/group. Depending on the system, one or both of these types of error may occur separately or in synchrony. Models have been developed to deal with uncertainties in identification (e.g. Nichols et al., 2004), state assignment (e.g. Fujiwara and Caswell, 2002; Kendall et al., 2003; Pradel, 2005), as well as both types simultaneously (Pradel, 2005). However, these advances are restricted to the Cormack-Jolly-Seber (CJS) framework and as such do not allow for abundance estimates. Since abundance estimates are often of interest to wildlife managers, there is a need to extend these approaches to the JS framework.

In extending the JS framework to deal with state uncertainty, we deal with states that are fixed over the duration of the experiment. Fixed characteristics, such as sex, are often of interest to wildlife managers and partition heterogeneity in the population. However, due to factors, such as lack-of-training, identification of these states on captures is not always known. To this end we consider the situation where sex can only be accurately assigned on some of the capture occasions or not at all.

When dealing with group uncertainty for fixed characteristics, it can be appropriate to look at identification uncertainty rather than mis-assignment (e.g. see Nichols et al., 2004). Primarily, each sampling occasion gives an independent opportunity to determine the true underlying group. As such, group assignments are not required on each sampling occasion, since future and previous occasions provide useful information. In this manner positive group assignments may be restricted to situations where the group is known with certainty, with all remaining scenarios handled by assignment to an unknown group. This is in contrast to non-fixed characteristics (e.g. breeding state) that can transition between sampling occasions and where there is an emphasis on making a state assignment each every capture occasion (but see Pradel, 2005).

In the present paper we extend the approach used by Nichols et al. (2004) to deal with sex uncertainty, to the JS framework by modifying the super-population formulation of Schwarz and Arnason (1996). Positive assignments were restricted to situations where sex was known with certainty and all other situations were dealt with as unknown designations. This model was then applied to walleye (*Stizostedion vitreum*) data from Mille Lacs, Minnesota, where sexing designations were not always possible due to staffing restrictions. The model was then used to determine sex-specific abundances and sex-specific recruitment over multiple years.

3.2 Survey Protocol

The survey protocol follows the standard JS protocol with groups except that on each capture occasion, 3 possible events may occur: unobserved, group unknown, or a positive group identification. We consider the case of 2 groups, males and females. We assume that any positive assignment is definitive and as a consequence strictly forbid the assignment of more than one sex to a tag history. For situations where group uncertainty exists the *unknown* designation should be used.

This is intended for situations where there are a mixture of experts and non-experts conducting the survey and each type of observer has different capabilities to assign group membership (i.e. sex) on capture. Both types of observers are expected to record tag numbers correctly, while only expert observers are allowed to assign group membership. Upon capture experts may choose to assign group membership (i.e. male or female), or designate group membership as *unknown*, while captures by non-experts are given the *unknown* group membership. In doing so we give expert observers the option of forgoing group membership assignment, should the need arise, and allow non-expert observations to be incorporated.

This treatment of groups differs slightly from the more classical grouping methods (i.e. Jolly, 1965; Lebreton et al., 1992; Schwarz and Arnason, 1996) that assume all animals can be definitively assigned to their respective groups at some point during the experiment. We relaxed this assumptions by modeling the probability of encountering individuals with unresolved sex as being a composite of both sexes. In this manner we take an approach that is similar both Nichols et al. (2004) and the post-stratification method developed by Conroy et al. (1999), which used covariates to assign sex post capture to juvenile animals. However, both approaches were developed for the CJS framework (excluding abundance estimation) and unlike the post-stratification we make no requirements on the availability of covariates that can be used to predict sex. Our approach also differs from *ad hoc* techniques that use a unique category with its own demographic parameters to deal with the unknown sex designations. If handled in such a manner group specific demographic parameters, such as apparent survivorship, will be positively biased due to the expected increase in encounters for histories where sex has been resolved (Nichols et al., 2004).

Finally, in the most general case of our model, it is assumed group assignments will be attempted on each and every sampling occasion. While useful for estimating assignment probabilities, it is not an absolute requirement. If proper model restrictions are employed, experimental protocols can be designed where group assignment is only attempted on certain capture occasions (see the example section).

3.3 Notation

Classical mark-recapture experiments use a capture history vector of '1' or '0' to represent whether an animal was caught (1) or not caught (0) during k sampling occasions. In our case we distinguish between the unobserved state (0) and three possible observed "groups": unknown (1), female (2) and male (3). Here the *female* and *male* represent a positive identification, while *unknown* represents uncertainty in assignment. As such the capture history will still be a vector of length k, but will consisting of 0's, 1's, 2's and 3's instead of the traditional 0's and 1's. For example, the history `313` would be a male that was caught on all three sampling occasions but was positively identified as male on the 1st and 3rd occasions only.

Lastly, because positive identifications are considered definitive, a single tag history cannot contain both male and female identifications. As a result histories such as `312` are not allowed and possible observed tag histories will not be a permutation of all states, but instead a restricted subset.

3.3.1 Statistics and Indicator Variables

k - number of sampling occasions in the experiment.

m - number of uniquely observable tag histories. Note that this does not include the unobserved tag history ($\boldsymbol{\omega}_0$).

i - the index for tag histories, where $i = 0, 1, 2, \ldots, m$; $i = 0$ is used to denote the unobserved tag history.

j - the index for sampling occasion, where $j = 1, 2, \ldots, k$.

n_j - total number of animals caught at sampling occasion j.

v_j - total number of animals lost on capture at sampling occasion j.

$\boldsymbol{\omega}_i$ - capture and identification history vector, where $\boldsymbol{\omega}_i = [\omega_{i1}, \omega_{i2}, \ldots, \omega_{ik}]'$ and

$$\omega_{ij} = \begin{cases} 3 & \text{for animals captured at sampling occasion } j \text{ and identified as male} \\ 2 & \text{for animals captured at sampling occasion } j \text{ and identified as female} \\ 1 & \text{for animals captured at sampling occasion } j \text{ and identified as unknown sex} \\ 0 & \text{for animals not captured at sampling occasion } j \end{cases}$$

ζ_i - the sex indicator for history $\boldsymbol{\omega}_i$, where

$$\zeta_i = \begin{cases} \{M\} & \text{for } \max(\omega_{i1}, \omega_{i1}, \ldots, \omega_{ik}) = 3 \\ \{F\} & \text{for } \max(\omega_{i1}, \omega_{i1}, \ldots, \omega_{ik}) = 2 \\ \{M, F\} & \text{for } \max(\omega_{i1}, \omega_{i1}, \ldots, \omega_{ik}) \leq 1 \end{cases}$$

Animals with an unknown sex designation are modeled as a composite of both sexes.

$\boldsymbol{\omega}_i^c$ - a capture history vector, where $\boldsymbol{\omega}_i^c = [\omega_{i1}^c, \omega_{i2}^c, \ldots, \omega_{ik}^c]'$ and

$$\omega_{ij}^c = \begin{cases} 1 & \text{if} \quad \omega_{ij} \geq 1 \quad \text{- a captured animal} \\ 0 & \text{otherwise} \end{cases}$$

$\boldsymbol{\omega}_i^s$ - sex identification vector, where $\boldsymbol{\omega}_i^s = [\omega_{i1}^s, \omega_{i2}^s, \ldots, \omega_{ik}^s]'$ and

$$\omega_{ij}^s = \begin{cases} 1 & \text{if} \quad \omega_{ij} > 1 \quad \text{- a sexed animal} \\ 0 & \text{otherwise} \end{cases}$$

$n_{\boldsymbol{\omega}_i}$ - number of animals with tag history $\boldsymbol{\omega}_i$.

f_i - first occasion when animals with tag history $\boldsymbol{\omega}_i$ were captured.

l_i - last occasion when animals with tag history $\boldsymbol{\omega}_i$ were captured.

κ_i - loss-on-capture indicator for tag history $\boldsymbol{\omega}_i$, where

$$\kappa_i = \begin{cases} 1 & \text{lost on sampling occasion } l_i \\ 0 & \text{not lost on any sampling occasions} \end{cases}$$

3.3.2 Fundamental Model Parameters

N - the *super-population* is the total of animals that were present at the start of the study or entered the system between any pair of sampling occasion and survived to the next sampling occasion.

β_j - the expected fraction of the super-population that enters the population at sampling occasion j and survived to sampling occasion $j+1$ for $\{j : j = 0, 1, \ldots, k-1\}$. β_0 is the expected fraction of animals alive just prior to the first sampling occasion and $\sum_{j=0}^{k-1} \beta_j = 1$.

π_j^s - the probability that an animal entering the population between j and $j+1$ is of sex s, for $s \in \{M, F\}$ and $\{j : j = 0, 1, \ldots, k-1\}$. The parameter π_0^s is the proportion of animals of sex s alive in the population prior to the first sampling occasion and $\sum_s \pi_j^s = 1$, for all j.

p_j^s - the sex-specific probability an animal will be caught at occasion j, given that the animal is alive at sampling occasion j, for $s \in \{M, F\}$ and $\{j : j = 1, 2, \ldots, k\}$.

δ_j^s - the probability, given an animal has been caught, that the sex will be positively identified at sampling occasion j, for $s \in \{M, F\}$ and $\{j : j = 1, 2, \ldots, k\}$. The probability the animal's sex cannot be identified is $(1-\delta_j^s)$. Positive identifications are considered to be definitive.

ϕ_j^s - The sex-specific probability that an animal survives and remains in the population from sampling occasion j to sampling occasion $j+1$, given it was alive and in the population at sampling occasion j, for $s \in \{M, F\}$ and $\{j : j = 1, 2, \ldots, k\}$.

ν_j - the probability of losing an animal when it is captured at sampling occasion j. Equal probability is assumed for both sexes. A loss is any event whereby the animal is not released into the catchable population after capture.

3.3.3 Functions of Parameters

β_j^* - the expected fraction of the super-population remaining to enter the population between sample occasion j and $j+1$, for $\{j : j = 0, 1, \ldots, k-1\}$. Note that $\beta_0^* = \beta_0$, $\beta_j^* = \frac{\beta_j}{\sum_{i=j}^{k-1} \beta_i}$, and $\beta_{k-1}^* = 1$. The β_j^*'s were used when fitting the likelihood, because they are unconstrained, taking on any value in $[0,1]$, yet maintain the $\sum_{j=0}^{k-1} \beta_j = 1$ constraint.

χ_j^s - The probability an animal of sex s is not observed again after period j.

$$\chi_j^s = \begin{cases} 1 - \phi_j^s + \phi_j^s \left(1 - p_{j+1}^s\right) \chi_{j+1}^s & j < k \\ 1 & j = k \end{cases} \quad \text{for } s \in \{M, F\}.$$

ψ_j^s - The probability of an animal of sex s entering the population and remaining unobserved prior to sampling occasion j.

$$\psi_{j+1}^s = \begin{cases} \beta_0 \pi_0^s & j+1 = 1 \\ \psi_j^s \left(1 - p_j^s\right) \phi_j^s + \beta_j \pi_j^s & j+1 > 1 \end{cases} \quad \text{for } s \in \{M, F\},$$

where π_0^s is the proportion of animals of sex s alive just prior to the first sampling period, while π_j^s, for $j > 0$, is the sex-specific proportion of the entrants.

B_j^s - The total number of animals of sex s that enter the system between sampling occasion j and $j+1$, for $\{j : j = 0, 1, \ldots, k-1\}$ and $s \in \{M, F\}$. B_0^s represents the number of animals alive (male or female) just prior to the first sampling occasion. The remaining B_j^s's are referred to as the net births and represent the process of live births, immigration and/or recruitment into the population of interest. Also note $E\left[B_j^s\right] = N\beta_j\pi_j^s$ and $N = \sum_s \sum_{j=0}^{k-1} B_j^s$.

N_j^s - The sex-specific population size at sampling occasion j.

$$E\left[N_{j+1}^s\right] = \begin{cases} N\beta_0\pi_0^s & j+1 = 1 \\ E\left[N_j^s\right] \phi_j^s + N\beta_j\pi_j^s & j+1 > 1 \end{cases} \quad \text{for } s \in \{M, F\}.$$

When losses on capture occur, a different formulation is required (see Appendix A.1).

3.4 Model Development

3.4.1 Assumptions

The standard mark-recapture assumptions are made. As well, we assume that both marked and unmarked animals of the same sex exhibit the same sex-specific catchability $\{p_i^s\}$ and survivorship $\{\phi_i^s\}$ (homogeneity). Also important is the assumption that newly captured unmarked animals are a random sample of all unmarked animals in the population. Additional assumptions include a homogeneous probability of being sexed, once captured, that is shared within each sex and homogeneity of demographic parameters for animals regardless of whether or not the animal was successfully sexed. This last assumption could be violated for example if juveniles are harder to sex and exhibit different survivorship. It is also assumed that the underlying characteristic (sex) is fixed throughout the experiment and the underlying population can be dichotomously divided into either males or females. Finally, as with most capture-recapture experiments it is assumed that tags are not lost; are not misread; and are unique to each animal; that sampling is instantaneous; and that the study area is constant throughout the experiment.

3.4.2 Likelihood

The model we propose is a direct extension of the super-population model proposed by Schwarz and Arnason (1996). While Schwarz and Arnason did allow for group specific demographic parameters (i.e. $\{\beta_{jg}\}, \{p_{jg}\}, \{\phi_{jg}\}$), we go a step further by modeling the probability of a positive group assignment (the identification probability), as well as modeling tag histories in which assignment did not occur.

With the JS capture-recapture models, it is common practice to break the likelihood into distinct portions. Following the notation of Schwarz and Arnason (1996) the likelihood of the super-population model can be written as

$$L = L_1^A(N, \{\beta_i\}, \{p_i\}, \{\phi_i\}) \times L_1^B(\{\beta_i\}, \{p_i\}, \{\phi_i\}) \times L_2(\{\nu_i\}) \times L_3(\{p_i\}, \{\phi_i\}).$$

These components make up $L_1^A \times L_1^B = P(\text{first capture} | \{\beta_i\}, \{p_i\}, \{\phi_i\})$, $L_2 = P(\text{loss-on-capture} | \{\nu_i\})$ and $L_3 = P(\text{recapture} | \{p_i\}, \{\phi_i\})$ respectively. Schwarz and Arnason divided the probability of first capture into two components (L_1^A and L_1^B), by first conditioning the complete likelihood on the total number of unmarked animals observed in the experiment (we use n_{obs} in place of $u.$). This formulation used was to develop a conditional estimate of $\hat{N}$ by first maximizing $L_1^B \times L_2 \times L_3$ and then using the estimates to derive the conditional MLE of $\hat{N}$. Asymptotically, the final result will be equivalent to maximizing the entire likelihood (Sanathanan, 1972; Schwarz and Arnason, 1996).

We formulate our model in a similar manner except we combine the L_1^B, L_2 and L_3 terms ($L_1^{B*} = L_1^B \times L_2 \times L_3$) so that the probability expression for each observed tag history can be modeled directly. Also similar to the group specific model proposed by Schwarz and Arnason (1996), we have male and female specific demographic parameters. However, we also model the probability an entrant will be of a particular sex (π_j^s) and the probability the sex can be identified, conditional on capture (δ_j^s).

We start by modeling the total number of observed tag histories ($n_{\text{obs}} = \sum_{i=1}^m n_{\boldsymbol{\omega}_i}$ and $n_{\boldsymbol{\omega}_0} = N - n_{\text{obs}}$), where $L_1^A = [n_{\text{obs}}|N] \sim \text{Binomial}\,(N, 1 - P(\boldsymbol{\omega}_0))$. Here $\boldsymbol{\omega_0}$ is the unobserved capture history ($\boldsymbol{\omega}_0 = [0, 0, \ldots, 0]'$) and $P(\boldsymbol{\omega}_0)$ is the probability of a male or female entering the population at some time prior to, or during, the study and remained unobserved for the remainder of the sampling occasions.

$$P(\boldsymbol{\omega}_0) = \sum_{s \in \{M,F\}} \sum_{j=0}^{k-1} \beta_j \pi_j^s (1 - p_{j+1}^s) \chi_{j+1}^s \tag{3.1}$$

Next, we model the distribution conditional upon being seen at least once, $L_1^{B*} = \left[\{n_{\boldsymbol{\omega}_i}\} \,|\, n_{\text{obs}}\right] \sim \text{Multinomial}\left(n_{\text{obs}}, \{\lambda_{\boldsymbol{\omega}_i}\}\right)$, where $\lambda_{\boldsymbol{\omega}_i} = \frac{P(\boldsymbol{\omega}_i)}{(1 - P(\boldsymbol{\omega}_0))}$. The probability of the observed tag history $P(\boldsymbol{\omega}_i)$, is

$$\begin{aligned} P(\boldsymbol{\omega}_i) \quad = \quad & \sum_{s \in \zeta_i} \psi_{f_i}^s \left\{ \prod_{j=f_i}^{l_i} (p_j^s)^{\omega_{ij}^c} \left(1 - p_j^s\right)^{(1-\omega_{ij}^c)} \right\} \\ & \times \left\{ \prod_{j=f_i}^{l_i} \left(\delta_j^s\right)^{\omega_{ij}^s \times \omega_{ij}^c} \left(1 - \delta_j^s\right)^{(1-\omega_{ij}^s) \times \omega_{ij}^c} \right\} \\ & \times \left\{ \prod_{j=f_i}^{l_i - 1} \phi_j^s \right\} \left(\chi_{l_i}^s\right)^{(1-\kappa_i)} \\ & \times \left\{ \prod_{j=f_i}^{l_i - 1} (1 - \nu_j)^{\omega_{ij}^c} \right\} \times (1 - \nu_{l_i})^{(1-\kappa_i)} (\nu_{l_i})^{\kappa_i}. \end{aligned} \tag{3.2}$$

Note that for the case where an animal is seen only once ($f_i = l_i$) the terms $\prod_{j=f_i}^{l_i-1} \phi_j^s$ and $\prod_{j=f_i}^{l_i-1} \nu_j$ are both treated as evaluating to 1.

The probability of determining the correct sex is represented by the δ_j^s parameter, which is conditional upon capture (i.e. $\omega_{ij}^s \times \omega_{ij}^c$). For animals where the sex identification has been positively determined at some point during the experiment $\zeta_i \in \{M\}$ or $\zeta_i \in \{F\}$, otherwise $\zeta_i \in \{M, F\}$. As a result, for tag histories where sex has been confirmed, only demographic parameters associated with that sex (or group) are used. For example the tag history `0133` (or `OUMM`) is a male that was captured on the second, third and fourth sampling occasions. On the first capture, its sex could not be identified, but on subsequent occasions a definitive sex assignment was made. The probability of this history (excluding the loss-on-capture component) will be

$$P(\texttt{0133}) = \psi_2^M p_2^M (1 - \delta_2^M) \phi_2^M p_3^M \delta_3^M \phi_3^M p_4^M \delta_4^M.$$

However, for tag histories where a definitive group identification is not possible, the histories are modeled as belonging to both groups, but with an unidentified status (i.e. $\omega_{ij}^s = 0$ for all j where $\omega_{ij}^c = 1$). As a result the $(1 - \delta_j^s)$ term appears in association with each capture occasion. For example the tag history `0111` (or `OUUU`) has the same capture history as the previous example, but without a definitive sex identification. In this case the probability of this tag history will be

$$\begin{aligned} P(\texttt{0111}) &= \psi_2^M p_2^M (1 - \delta_2^M) \phi_2^M p_3^M (1 - \delta_3^M) \phi_3^M p_4^M (1 - \delta_4^M) \\ &\quad + \psi_2^F p_2^F (1 - \delta_2^F) \phi_2^F p_3^F (1 - \delta_3^F) \phi_3^F p_4^F (1 - \delta_4^F). \end{aligned}$$

This form of grouping differs from more classical grouping methods (i.e. Jolly, 1965; Lebreton et al., 1992; Schwarz and Arnason, 1996) that assume all animals can be definitively assigned to their respective groups during at least one of the encounters. By doing so the fully parameterized model (no model constraints, either group or time) will be equivalent to running multiple independent capture recapture experiments in parallel on each of the possible groups. In contrast, for the approach used here the *unknown* designation will be a composite of the possible positive state assignments, even in the most parameterized model. As a result, the complete likelihood for this model will be

$$L_{\text{complete}} = L_1^A \times L_1^{B*}$$

$$= \left\{ \binom{N}{n_{\text{obs}}} (1 - P(\boldsymbol{\omega}_0))^{n_{\text{obs}}} P(\boldsymbol{\omega}_0)^{N - n_{\text{obs}}} \right\} \times \left\{ \frac{n_{\text{obs}}!}{n_{\boldsymbol{\omega}_1}! n_{\boldsymbol{\omega}_2}! \cdots n_{\boldsymbol{\omega}_m}!} \prod_{i=1}^{m} \left(\frac{P(\boldsymbol{\omega}_i)}{(1 - P(\boldsymbol{\omega}_0))} \right)^{n_{\boldsymbol{\omega}_i}} \right\}. \quad (3.3)$$

3.4.3 Model Constraints, Link Functions and Covariates

We used parameter index matrices (PIM) as implemented in MARK (White and Burnham, 1999) to provide a flexible modeling environment, as suggested by Lebreton et al. (1992). This allows a general way to specify parameter restrictions on the fully time dependent model. Four separate PIM's were used for the $\{\pi_j^M\}$, $\{p_j^s\}$, $\{\delta_j^s\}$ and $\{\phi_j^s\}$ model parameters respectively. Covariates may be included through the use of design matrices. Similar to MARK, parameters may also be fixed at known values. Fixing parameter values can be useful for creating testable hypothesis such as a 50:50 incoming sex ratio (i.e. $\pi_j^M = 0.5$ for all $j > 0$).

A fifth, limited PIM, was implemented for $\{\beta_j\}$ parameters, where only basic constraints may be imposed. Some examples include death only ($\beta_0 = 1, \beta_{1:k-1} = 0$), no recruitment for certain periods (e.g. $\beta_1 = 0$), and constrained entrance (e.g. $\beta_1 = \beta_2$ or $\beta_1 = \beta_2 = 0.1$). Covariates cannot however be supplied for the $\{\beta_j\}$ parameters (see Schwarz and Arnason, 1996, for an explanation).

Finally, common link functions were used to restrict parameter estimates to remain between 0 and 1 (see Lebreton et al., 1992). Available link functions follow the MARK implementation and include the sin, logit, log and identity links.

3.4.4 Parameter Redundancy

As mark-recapture models become more complex, the large parameter sets needed to describe the modeled processes may lead to parameter redundancy (Catchpole and Morgan, 1997).

Recently, Catchpole and Morgan (1997), Catchpole et al. (1998) and Catchpole and Morgan (2001) developed a technique for assessing parameter redundancy using existing computer algebra packages that are capable of performing symbolic math. The advantage over numeric techniques is that numerical criteria are not needed to identify uniquely estimable parameters and in some cases estimable parameter combinations can also be determined.

Following the implementation by Gimenez et al. (2003), we used the software package

Maple (Version 10) to determine the model deficiency as well as the uniquely identifiable parameters. In our case the most general, time dependent, model was found to have a parameter deficiency of 4. The full set of the time dependent classification parameters $\{\delta_j^s\}$ were identifiable. As in the simple JS super-population model, $\{\phi_j^s : j = 1, 2, \ldots, k-2\}$ and $\{p_j^s : j = 2, 3, \ldots, k-1\}$ were identifiable. Finally, as the sampling occasions increase the middle proportion of entrants $\{\beta_j : j = 2, 3, \ldots, k-2\}$ and the probability of a male entrant $\{\pi_j : j = 2, 3, \ldots, k-2\}$ also become uniquely estimable. The remaining parameters formed complex estimable combinations. It should be noted that sparse data sets may further increase the parameter deficiency, as is the case with the example study.

3.4.5 Parameter Estimation

Parameter estimates were derived in a manner similar to the procedure described by Schwarz and Arnason (1996). Estimates of $\{\hat{\nu}_j\}$ were found by maximizing the L_2 loss-on-capture component. The L_1^{B*} was then maximized with respect to the remaining parameters given $\{\hat{\nu}_j\}$. Finally, the estimates $\{\hat{\beta}_j\}$, $\{\hat{\pi}_j\}$, $\{\hat{p}_j^s\}$, $\{\hat{\delta}_j^s\}$ and $\{\hat{\phi}_j^s\}$ were used to estimate $\hat{N}$ using L_1^A, where $\hat{N} = \left[n_{\text{obs}}/(1 - \widehat{P(\boldsymbol{\omega}_0)})\right] \left(\text{greatest integer} \leq n_{\text{obs}}/(1 - \widehat{P(\boldsymbol{\omega}_0)})\right)$ as per Sanathanan (1972).

At this point it should be noted that in this formulation of the JS model, beyond the assumption on unmarked animals (which is untestable), there is no information about N or β_0 in the capture histories (see Link and Barker, 2005). That said, the capture histories do contain information on the remaining β_j parameters and given reasonable survival estimates, the estimates of N and β_0 may still be fair.

Finally, the L_1^{B*} was maximized numerically using a quasi-Newton (variable metric) procedure in R. The Hessian matrix was then numerically determined using the full likelihood (3.3) and point estimates from the conditional fit. The delta method was then used to determine the variance-covariance matrix for all derived parameters, such as $\{\hat{B}_i^s\}$ and $\{\hat{N}_i^s\}$. It should be noted that the variance calculations for $\{\hat{N}_i^s\}$ excluded the additional variance associated with estimating the sex ratio of the loss-on-captures that were not successfully sexed (see Appendix A.1). This represents a rather rare scenario as well as a small source of variation in most applications.

3.4.6 Goodness of Fit

The goodness-of-fit tests for a JS model involves the L_1^{B*} component, where the likelihood has been conditioned on the observed histories. Both model deviance and the Pearson chi-squared statistic were used. The model deviance is computed by comparing the likelihood of the fitted model to the saturated model, where every $\boldsymbol{\omega}_i$ history has the probability $\frac{n_{\boldsymbol{\omega}_i}}{n_{\text{obs}}}$. The difference between the saturated and fitted model (3.4) will be $D \sim \chi^2_{(m-p-1)}$, where the number of degrees of freedom equal is to the difference in the number of parameters between the two models $(m-p-1)$. The number of parameters in the saturated model is the number of unique tag histories m minus 1, while p is the number of parameters estimated in the L_1^{B*} component.

$$D = 2\left[\sum_{i=1}^{m} n_{\boldsymbol{\omega}_i}\left(\log\left(\frac{n_{\boldsymbol{\omega}_i}}{n_{\text{obs}}}\right) - \log(\lambda_{\boldsymbol{\omega}_i})\right)\right] \tag{3.4}$$

In addition to deviance, the Pearson chi-squared goodness-of-fit (GOF) statistic was also used

$$\chi^2 = \sum_{i=1}^{m} \frac{(o_i - e_i)^2}{e_i} \tag{3.5}$$

where $o_i = n_{\boldsymbol{\omega}_i}$ is the observed tag history frequency and $e_i = N \times P(\boldsymbol{\omega}_i) = n_{\text{obs}} \times \frac{P(\boldsymbol{\omega}_i)}{1-P(\boldsymbol{\omega}_0)}$ is the expected frequency. Under the hypothesis that the model is correct $\chi^2 \sim \chi^2_{(m-p-1)}$. Asymptotically, both tests should be equivalent, however differences can occur depending on the adequacy of the $\chi^2_{(m-p-1)}$ approximation. Specifically, for small frequencies there is some evidence to suggest that χ^2 may preform better that D, since D can be unduly influenced by very small frequencies (see Cressie and Read, 1989).

3.4.7 Model Selection

By using the PIM's and design matrices, many different models can be specified following the notation of Lebreton et al. (1992). Of particular interest will be the time-varying parameterization (i.e. $\{\pi_j\}$) , time independent (e.g. $\pi.$) and group independent parameterization (e.g. $p_j^M = p_j^F$ for $j = 1, 2, \ldots, k$). Also of interest may be entrance restrictions such as a death only model ($\beta_0 = 1; \beta_j = 0$ for $j = 1, 2, \ldots, k-1$), or death only for specific time periods (e.g. $\beta_2 = 0$). Finally, also of interest may be situations where sex assignment does not occur in every sampling occasion (i.e. $\delta_j^s = 0$ for some j). These models can be specified using the respective parameter PIMs and specifying constants for the fixed parameters.

A variety of techniques can be used to choose between competing sets of models and model averaging for models that have equal support (see Burnham and Anderson, 2002). Techniques used for model selection included the Akaike information criterion (AIC_c) and QAIC_c. Both balance overall model fit to the data with number of parameters needed, but QAIC_c also corrects for lack of fit and effective sample size. It should be noted the lack of fit used for the QAIC_c calculations was based on the Pearson GOF statistic (3.5). Finally, a candidate model set was chosen *a prior*; and the best fit was selected from competing models.

3.5 Example

Our example is concerned with estimating the number of walleye in Mille Lacs Lake, Minnesota, for which the study design is outlined in Schwarz (2004). Briefly, a three-year mark-recapture study was initiated in 2002. In all years tags were applied in two phases, first via trap netting on the spawning grounds and second during the angling season from launch-boats. Recoveries came from angling, trap netting, tribal harvests and a gillnet assessment at the end of the season. Nearly complete sexing occurred during the spawning ground releases and gillnet assessment at the beginning and end of each season respectively. However, due to manpower restrictions, only partial sexing was possible during the angling portion of the season, resulting in capture histories with unknown sex designations.

For the purpose of the example, all 3 years are considered (2002 to 2004), but only two sampling occasions per year are included, the spawning ground releases and harvest at the start of the season and the angling season that occurs over the summer, but prior to the end of the season gill net assessment, which was not included. This results in a total of six sampling occasions, with two sampling occasions occurring every year (Table 3.1). The data set is provided in Appendix A.2 (Table A.1).

3.5.1 Model Constraints

The first sampling event within each year (sampling occasions $j = 1, 3, 5$), was limited to fish greater than 14 inches in length. For the second event within each year (sampling occasions $j = 2, 4, 6$), the minimum size requirement was increased 16 inches so as to preclude recruitment. Therefore, in the model recruitment was not allowed between occasions 1 and 2,

Table 3.1: Summary of Mille Lacs study design.

	2002		2002		2002	
	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
Survey Type	Spring Tagging + Tribal Harvest	Launch Boats + Angling	Spring Tagging + Tribal Harvest	Launch Boats + Angling	Spring Tagging + Tribal Harvest	Launch Boats + Angling

occasions 3 and 4, and occasions 5 and 6 (i.e. the following values were fixed: $\beta_1, \beta_3, \beta_5 = 0$). New recruits were however allowed between occasions 2 and 3, and between occasions 5 and 6 (i.e. $\beta_2, \beta_4 \geq 0$). These restrictions reflect allowable catch restrictions and as a result the $\{\beta_j\}$ parameters represent recruitment into the size restricted fishery, rather than births.

The walleye data set also suffered from issues of tag loss, mainly between 2002 and 2003 (see Cowen and Schwarz, 2006). As a result, the estimates of ϕ_2^s (apparent survivorship between periods 2 and 3) are expected to be lower than the estimates of ϕ_4^s as as such these values were never restricted to be equal.

3.5.2 Model Selection and Best Model Estimates

Due to the known issue of tag loss in this population (see Cowen and Schwarz, 2006) model misfit was expected. Since the example is for illustration purposes, the best model presented was judged so by a combination of model goodness of fit criteria (model deviance and the over-dispersion estimate $\hat{c}$), derived parameter estimates (population sizes), and AIC model ranking. The fundamental parameter estimates from the best fitting model are displayed in Table 3.2, with derived parameters, population size (N_j^s) and recruitment (Bj^s), displayed in Table 3.3.

The model fit is poor ($\hat{c} = 128$) for two reasons. First, the very large sample sizes imply that small discrepancies in fit can be detected quite easily. Second, as noted in Cowen and Schwarz (2006), tag loss, particularly on the first winter, is a serious problem in this study. The latter would require extensive modeling to incorporate properly. Nevertheless, despite the apparent lack-of-fit, the estimates are reasonable and match quite well to those from other work (Schwarz, 2004; Cowen and Schwarz, 2006).

In particular, the model picks up the known higher catchability of male fish versus female fish on the spawning ground (occasions $j = 1, 3, 5$; p_j^M and p_j^F estimates, Table 3.2) and the lower catchability during the angling season (occasions $j = 2, 4, 6$; p_j^M and p_j^F estimates, Table 3.2). When compared to estimates from an equivalent simple JS, that does not include sex, the naive catchability estimates occurred somewhere between the sex-specific estimates (Table 3.4).

The initial male population estimate was low ($\hat{N}_1^M \approx 252,000$ versus $\hat{N}_1^F \approx 490,000$, Table 3.3), but is not unexpected. Tribal harvest (during the spring) consists mostly of male fish. Over the years, this will tend to lower the sex ratio of the standing population, as seen in the female/male ratio column (3.3). We are puzzled though, by the apparent high male proportion of new recruits prior to 2003. We tried several different initial values, but the resulting estimates were always high.

Yearly survival rates for 2002 are likely biased low because of the excessive tag loss observed between 2002 and 2003. The estimates appear more reasonable for later periods when tag loss was much less prevalent. Arnason and Mills (1981) showed that homogeneous tag loss resulted in unbiased estimates of population size, but biased estimates of recruitment.

Finally, our estimates of population size are comparable to those in Cowen and Schwarz (2006) which accounted for tag loss, but not a lack of sex identification. The estimate of total population size are also comparable to those from a simple JS model pooled over both sexes (Table 3.3).

3.6 Discussion

Abundances, both general and sex or group-specific, are of often of interest to biologists and managers alike, but obtaining such estimates can be difficult in cases where the group status was not always be measurable. If the population sex ratio is known and constant then such assignments are not needed (Zhang et al., 2005). However, situations where the sex ratio is known and stable is rare. The method presented here provides a general approach to deal with unknown designations that avoids biased demographic estimates that can arise if classification uncertainty is otherwise dealt with in an *ad hoc* fashion.

The JS extension described here is an intermediate between the classical grouping approach, where the state is fixed and always measurable and the multi-state extension of the JS model (Schwarz and Dupuis, 2007). If sex is known all the time then the experiment could

Table 3.2: Point estimates (SE) from the best model fit. Values have not been adjusted for over-dispersion ($\hat{c} = 128$). Fixed parameters are shown without a SE.

	2002		2003		2004	
	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
β_{j-1}	0.846 (0.029)	0	0.052 (0.026)	0	0.102 (0.014)	0
π^M_{j-1}	0.339 (0.024)	0	0.489 (0.050)	0	0.489 (0.050)	0
p^M_j	0.131 (0.013)	0.013 (0.002)	0.283 (0.008)	0.010 (0.001)	0.191 (0.017)	0.005 (0.001)
p^F_j	0.011 (0.001)	0.025 (0.002)	0.024 (0.002)	0.019 (0.002)	0.029 (0.005)	0.008 (0.002)
δ^M_j	0.981 (0.001)	0.183 (0.011)	0.981 (0.001)	0.183 (0.011)	0.981 (0.001)	0.183 (0.011)
δ^F_j	0.968 (0.005)	0.072 (0.002)	0.968 (0.005)	0.072 (0.002)	0.968 (0.005)	0.072 (0.002)
ϕ^M_j	0.750 (0.067)	0.624 (0.055)	0.829 (0.071)	0.903 (0.068)	0.829 (0.071)	–
ϕ^F_j	0.903 (0.068)	0.650 (0.054)	0.795 (0.093)	0.829 (0.151)	0.795 (0.093)	–

Table 3.3: Estimated male and female specific birth, abundances and sex ratio by sampling occasion. Note that population size (N_j^s) are in 1000's. Also note that B_0^s are not shown as they are the same as N_1^s values. SE have not been adjusted for over-dispersion. The final column represents estimates from a simple JS model pooled over sex.

	Births (B_{j-1}^s)		Population Sizes (N_j^s)			Population Total (N_j)	
Year	Females	Males	Females	Males	F/M Ratio	Sex-Specific	Simple JS
2002 ($j = 1$)	–	–	490 (34)	252 (25)	1.9	743 (46)	695 (10)
($j = 2$)	0 (−)	0 (−)	441 (27)	176 (31)	2.5	617 (45)	677 (10)
2003 ($j = 3$)	22 (11)	23 (11)	302 (23)	125 (3)	2.4	427 (23)	342 (5)
($j = 4$)	0 (−)	0 (−)	238 (27)	92 (10)	2.6	330 (28)	327 (5)
2004 ($j = 5$)	45 (10)	44 (5)	231 (39)	118 (13)	2.0	351 (42)	326 (5)
($j = 6$)	0 (−)	0 (−)	182 (41)	89 (20)	2.0	271 (45)	312 (5)

Table 3.4: Recapture probabilities estimates (SE) from the sex-specific JS and a simple JS model. SE have not been adjusted for over-dispersion.

	Sex-Specific		Simple JS
Year	p_j^F	p_j^M	p_j
2002 ($j = 1$)	0.011 (0.001)	0.142 (0.013)	0.055
($j = 2$)	0.024 (0.001)	0.014 (0.002)	0.020
2003 ($j = 3$)	0.025 (0.002)	0.282 (0.008)	0.131
($j = 4$)	0.020 (0.002)	0.010 (0.001)	0.017
2004 ($j = 5$)	0.028 (0.005)	0.193 (0.018)	0.101
($j = 6$)	0.008 (0.002)	0.005 (0.001)	0.007

be broken down into a simple stratification and results in males and females abundances being estimated separately. As a result biologically relevant restrictions, such as a 50:50 sex ratio of incoming entrants, cannot be imposed. One of the nice features of the proposed model is that the probability of an entrant's sex is directly modeled in the likelihood. This makes it a fairly trivial task to impose constraints on the sex ratio of entrants, which in turn could be useful for testing a variety of biological hypotheses that may be specified in a multi-model framework (see Burnham and Anderson, 2002). The model presented is also flexible in regards to identification probabilities. If definitive sex assignments can always be made (over either a portion or the entire experiment), then this formulation can still be used by fixing the the appropriate δ_j^s parameter to 1.

While formulation presented here dealt explicitly with sex as the underlying group of interest, the model notation can also handle multiple fixed groups. In this case the superscript s would be used to represent the various groups, rather than sex. Also an additional restriction is needed so that probability of entrant belonging to a specific group sums to one for all sampling occasions (i.e. $\sum_s \pi_j^s = 1$ for all j). Lastly, for the unknown group tag history i the assigned group ζ_i would be the set of all possible fixed groups in the system.

This extension is not however without limitations. In dealing with classification uncertainty, explicit assumptions were made on how group assignments occurred. Specifically, any positive assignments were considered to be definitive. This was intended for sampling situations where there is a mixture of observers (e.g. experts and non-experts) and only a subset is capable of making positive assignments. This does not however protect against mis-assignments. If the proportion of mis-assignment is large, then group-specific estimates may become biased. While it is possible to adapt the modeling framework to deal with mis-assignment in addition to group identification uncertainty, mis-assignment can be largely controlled if field practitioners ensure that any uncertainty in an animal's group assignment is treated as an inability to assign a group.

Finally, it is also assumed that there is no unmodeled heterogeneity in the identification probabilities. This may also not be the case if animals that have less physical development or are diseased have identification probabilities that differ from the rest of the population. Such conditions will not only affect the identification probabilities, but such animals may also exhibit different demographic parameters which would violate homogeneity assumptions. In this case further groupings or the addition of covariates may be needed to model the potential heterogeneity in identification probabilities and demographic parameters.

Chapter 4

Modeling Pairwise Dependent Capture Outcomes in Mark-Recapture Experiments

4.1 Introduction

Mark-recapture models make assumptions of independence among individuals in the experiment in order to formulate the likelihood as a product of individual likelihood components. This approach has arisen due both to convenience and practicality. While we may suspect that various dependency structures exist in a population (e.g. preferential selection of habitat, as well as social behaviour such as herding and pair-bonds), generally there is little available information on which to act. The assumption of independence then becomes necessary to make mark-recapture problems mathematically tractable.

While the topic of dependency has been considered in other relevant techniques (e.g. distance sampling; Buckland et al., 2010) it has received little to no consideration in mark-recapture methodology. Few however would suggest that associations within a population, such as mating pairs, can be expected to act in a completely independence manner. Clearly, there there is a need to handle these types of situations, yet to date we have no modeling framework available that can even adequately handle information on mating pairs.

Of the potential forms of dependency, pair-bonds (also termed lifelong social monogamy) provides the best first step. There are a number of instances of known social monogamy in wild populations and include perennially monogamous species such as various water fowl (e.g. ducks, geese, and swans), large seabirds (e.g. albatrosses), many birds of prey (e.g. eagles and ospreys), cranes, some carnivores (e.g. foxes and wolves), as well annually monogamous species such as number of passerines such as songbirds. While far from exhaustive, the list demonstrates the potential availability of such information. In addition, this type of dependency is the most likely to be well documented. Most wildlife researchers are interested in the various aspects associated with reproduction. Information on mating pairs would certainly qualify under this umbrella.

Our goal then is to the take the first steps towards building such a modeling framework. As a first effort we introduction a method for handling pairwise capture dependency in the Cormack-Jolly-Seber (CJS) framework (Cormack, 1964; Jolly, 1965; Seber, 1965). Dependency is assumed to occur between individuals in a mating pair, which we can expect to be caught together with a probability that differs from other individuals not in the mating pair. We introduce a sampling correlation parameter ρ to represent this process and herein referring to our proposed model as belonging to the ρ-CJS modeling framework.

In developing the approach we have also adopted a state-space implementation, based on the Gimenez et al. (2007) individual state-space formulation. State-space models (SSM) provide a very general framework that view the complete realization of the data to be a combination of a state (e.g. survival) and observation (e.g. capture) process. The strength of this approach is the ability to specify models with highly complex interactions in relatively simple terms. This contrasts with more traditional approaches that rely on complex probability statements which make future model modifications more difficult. Our chosen approach should also allow practitioners an easier path for modifying the proposed models to handle idiosyncrasies of their data sets.

4.1.1 Motivation

The main motivation for our work comes from a long-term studies of harlequin ducks in the McLeod River region located Alberta, Canada (see Bond et al., 2009). The primary objective of the study is to monitor wild harlequin populations by obtaining estimates of demographic processes such as survival. Due to the high capture rates and relatively small population size, the study has created and maintained an in-depth pairing catalog of known individuals and known pairings. Since mating pairs are often observed together on the same capture events, there was interest in being able to account for this non-independence when generating demographic estimates.

Harlequins also make a good first candidate for modeling dependence as they typically exhibit long term pair-bonds, changing mates only when a particular partner dies. For the initial formulation we will assume individuals strictly pair for life, disallowing mate switching. Generally, pair-bonds are initiated on the wintering ground (West Coast) prior to migration to the McLeod River region in the spring for reproduction. As such, we can expect the majority of individuals entering the experimental area will be either single or already in a pair bond allowing us, for the time being, to ignore the possibility of new mating pairs being formed once individuals have entered the experiment. In future models we hope to relax these restrictions.

The data set also has unique characteristics that make a state-space implementation preferable. For example, it is suspected that females exhibit fidelity for a specific breeding site, while males exhibit fidelity only for females. As such, we can expect the male to permanently emigrate from the site if the female mate dies. These types of complexities can be handled in relatively straight manner when employing in a state-space formulation.

Our motivation for first extending the CJS framework to account for pair dependency is that the CJS framework is the predominant form of open population mark-recapture models used in the literature today. The framework obtains estimates of survival, while dealing with the nuisance issues such as detectability in the form of encounter or capture probabilities. Issues of detection occur when we observe 0's in the capture history even though the individual was alive and available for capture. The CJS also also provides a simpler framework to work with than other open population models such as Jolly-Seber models, that also estimate abundance.

4.2 Model Development

We will be considering a mark-recapture experiment with T capture occasion with all mating pairings formed prior to the entrance of either individual into experiment. Further we assume dependence may occur between capture outcomes within mating pairs, but assume independence between pairs. We also assume that survival is a completely independent process.

Typically, mark-recapture experiments are represented by a capture history for each particular animal as a sequence of the form `00110`, where '1' represents a capture and '0' no capture. The underlying state of the animal is its "alive status" with '1' for alive and a '0' for dead. We know the "alive status" for a captured individual will be a '1', but typically we do not know the "alive status" after last capture. The observations for an experiment where n individuals that were caught at least once can then be represented in a matrix of dimensions $n \times T$.

For the ρ-CJS model we need to consider a total $m < n$ independent entities, that consist of either single individuals or mating pairs. For entity j there are four possible female/male pair states: alive/alive, alive/dead, dead/alive and dead/dead. (Note we use dead and absent interchangeably, as is commonly practiced in the CJS literature.) Let $\mathbf{Z}_{j,t}$ be a random state vector talking the values $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, $(0, 0, 1, 0)$ and $(0, 0, 0, 1)$ to represent these states for entity j on sampling occasion t. Individuals without mates are restricted to two of the possible four states depending on their sex.

Similarly, let $\mathbf{Y}_{j,t}$ be a random observation vector taking the values $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, $(0, 0, 1, 0)$ and $(0, 0, 0, 1)$, but representing whether both members of the pair were captured; only the female member of the pair was captured or a lone female was captured; only the male member of the pair or a lone male was captured; or no capture of an animal.

The fundamental model parameters are as follows:

$\phi^s_{j,t}$ - Probability of survival from occasion t to $t+1$ for an individual of sex $s \in \{M, F\}$ in entity j given that it was alive at t.

$p^s_{j,t}$ - Marginal probability of capture for an individual sex $s \in \{M, F\}$ on occasion t belonging to entity j given that it was alive at t.

ρ - The correlation coefficient for pairwise capture outcomes of individuals in a mating pair.

We have only introduced one new parameter ρ that doesn't exist in the standard CJS framework. It should also be noted that capture probability $p^s_{j,t}$ is now defined as a marginal probability of capture only for individuals belonging to a pair, for single individuals it retains its standard definition.

4.2.1 Modeling Pairwise Captures

For a given mating pair we assume there will be dependence in the Bernoulli capture outcomes for the pair. We can produce a 4×4 table of the possible pairwise capture outcomes in terms of our marginal capture probabilities and ρ (Table 4.1). The expression for $p^{fm}_{j,t}$ is derived from the definition of covariance (see Appendix B.1) with the rest of the entries following from the constraints placed on the margins. It is also worth noting that we use superscripts that are lowercase (e.g. p^{m0}) for parameters that are functions of fundamental parameters (e.g. p^M), which use capital superscripts.

By modeling pairwise capture outcomes as a function of ρ we are making strong assumptions about how individuals in a mating pair associate with one another. For example, arguments could be made that once mated, the pair may act in a completely different fashion (e.g. the pair could act as a super or meta-individual). For these types of cases we will want to make the fewest possible assumptions on the probabilities of Table 4.1. Let $p^{fm}_{j,t}$, $p^{f0}_{j,t}$ and $p^{m0}_{j,t}$ be the pairwise, female only and male only capture probabilities conditional on that both individuals in mating pair j are alive. The complement $p^{00}_{j,t} = 1-p^{fm}_{j,t}-p^{f0}_{j,t}-p^{m0}_{j,t}$ is the probability of no capture for pair j. The only constraint required is that $p^{fm}_{j,t}+p^{f0}_{j,t}+p^{m0}_{j,t} \leq 1$ for all j and all t.

Instead of estimating ρ we will have to estimate $p^{fm}_{j,t}$, $p^{f0}_{j,t}$ and $p^{m0}_{j,t}$ separately from the capture probabilities of single individuals ($p^s_{j,t}$ for $\{s \in M, F\}$). This will likely result in an over-parameterized model requiring further constraints (beyond the constraint on the sum) to be added in order to make the model estimable. Finally, also note that the definition of $p^s_{j,t}$ changes slightly, becoming probability of capture for single individuals only.

4.2.2 Observation and State Equations

State-space models describe the evolution of two time-series processes; one is referred to as the state process and the other the observation process (Buckland et al., 2004). The state process is not observable (e.g. the alive/dead state of birds after last capture) and the observation process is assumed to be a function of the state process. Typically, observations for a given point in time are modeled as being conditionally dependent on the state process at that same time point. Furthermore, the state process is assumed to be a first-order Markov process, that is the current state is only depend on the state in the previous time step.

To fully specify a state-space model we need to specify the initial state distribution along with the distributions for the state and observation processes. Statistical inference is then accomplished by integrating the joint distribution of the observation and state processes over the state process. If a Bayesian approach is employed, numerical techniques for deriving

Table 4.1: Pairwise capture probabilities for the ρ-CJS model formulation for female (F) and male (M) of mating pair. Note pair and time dependent subscripts have been omitted for clarity and $\sigma^F = \sqrt{p^F(1-p^F)}$ and $\sigma^M = \sqrt{p^M(1-p^M)}$.

	Not captured (M)	Captured (M)	Marginal (F)
Not captured (F)	$p^{00} = 1 - p^F - p^M + \rho\sigma^F\sigma^M + p^F p^M$	$p^{m0} = p^M - \rho\sigma^F\sigma^M - p^F p^M$	$1 - p^F$
Captured (F)	$p^{f0} = p^F - \rho\sigma^F\sigma^M - p^F p^M$	$p^{fm} = \rho\sigma^F\sigma^M + p^F p^M$	p^F
Marginal (M)	$1 - p^M$	p^M	

the posterior distribution, such as Markov-chain Monte Carlo (MCMC), will automatically integrate over the possible values of the state-space.

To gain an intuitive understanding of how this might work, consider the capture history `110` in a regular CJS experiment. We know the individual was alive for the first two occasions, but may be alive or dead on the last occasion. Normally, these possibilities are accounted for when formulating the marginal capture probability statement. For example in a standard CJS model the marginal probability statement $P(110) = \phi_{i,1}p_{i,2}(\phi_{i,2}(1-p_{i,3}) + 1-\phi_{i,2})$ would be directly used in the likelihood. The term $\phi_{i,2}(1-p_{i,3})$ accounts for the possibility individual i survived from the second to third occasion, but was not captured in the third. The term $1-\phi_{i,2}$ accounts for the alternate possibility where individual i died between the second and third occasions and was therefore not available for capture.

In a state-space formulation the integration of this final state is performed by directly considering each possible outcome. If we take a Bayesian approach at each iteration of the MCMC chain one of the two states on the third occasion ($t = 3$) would be realized. When the chain reaches the stationary distribution, and we are sampling from the posterior, these two state-space possibilities would be then sampled with the correct proportion. In ϕ_2 of the cases a '1' will be selected (corresponding to the animal being alive at $t = 3$) and in the other $1-\phi_2$ cases a '0' will be selected (corresponding to the animal being dead at $t = 3$). In each case the "complete data" likelihood would be evaluated using one of these two states. The resulting posterior will have the same properties as if we had integrated the likelihood directly.

A key advantage of the state-space formulation is that the "complete-data" likelihood is based on knowing the latent state, making the model simpler to write down. For example (Gimenez et al., 2007) provided an individual state-space CJS formulation (which models the states of each individual), where the state and observation processes are modeled as

$$X_{i,t+1}|X_{i,t} \sim \text{Bernoulli}(X_{i,t}\phi_{i,t}), \tag{4.1}$$

$$Y_{i,t}|X_{i,t} \sim \text{Bernoulli}(X_{i,t}p_{i,t}), \tag{4.2}$$

respectively, for $t \geq e_i$ and $p_{i,e_i} = 1$; e_i indicates the first occasion individual i was captured. Here $X_{i,t}$ is a random variable designating the state of individual i on occasion t and $Y_{i,t}$ is the observation. Note that CJS mark-recapture models condition on first capture, removing the first capture probability from the probability statement for an individual capture history.

In the state-space formulation this is accomplished by setting $p_{i,e_i} = 1$. The likelihood (not shown) is the integral of the observation/state joint distribution over the possible state-space. The initial state distribution, used to form the likelihood, is a simple point mass as the animal is known to be alive.

This reduction of the problem into observation and state components also follows approaches such as the matrix population models of Caswell (2001), the multievent model of Pradel (2005) and the multistate occupancy model discussed later in this thesis. However, while these approaches focus on specifying complex capture history probabilities in a simpler manner, the state-space approach is more generalizable and can handle any situation where the distributions for the initial state, the state process, and the observation processes are known.

ρ-CJS State Equations

For the proposed ρ-CJS model the state equation for $\mathbf{Z}_{j,t}$ can be formulated as either a multinomial process or a combination of Bernoulli trials. The multinomial state equation is defined as

$$\mathbf{Z}_{j,t+1}|\mathbf{Z}_{j,t} \sim \text{Multinomial}\left(1, \mathbf{Z}_{j,t}\begin{bmatrix} \phi^F_{j,t}\phi^M_{j,t} & \phi^F_{j,t}(1-\phi^M_{j,t}) & (1-\phi^F_{j,t})\phi^M_{j,t} & (1-\phi^F_{j,t})(1-\phi^M_{j,t}) \\ 0 & \phi^F_{j,t} & 0 & 1-\phi^F_{j,t} \\ 0 & 0 & \phi^M_{j,t} & 1-\phi^M_{j,t} \\ 0 & 0 & 0 & 1 \end{bmatrix}\right). \tag{4.3}$$

Here pairs may may transition from both being alive to one or both being dead. We assume independent survival of the male or female, and that matings occur prior to entry into the experiment so that once single, an individual does not form another mating pair. These are somewhat arbitrary restrictions that may be relaxed in future work.

Rather than modeling state transitions as being a multinomial process, we can alternatively define the state equation in terms of n total individuals. In this case the state equation will be defined as

$$X_{i,t+1}|X_{i,t} \sim \text{Bernoulli}\left(X_{i,t}\phi_{i,t}\right) \tag{4.4}$$

for $t \leq f_i$, where f_i is the first known instant where individual i has entered the experiment (and known to be alive), and $X_{i,t}$ is the individual's alive/dead state on occasion t. Next, the mating-pair state vector for pair j on occasion t will be defined as

$$\mathbf{Z}_{j,t}|X_{l,t}, X_{k,t} = \left[\begin{array}{cccc} X_{k,t}X_{l,t}, & X_{k,t}(1-X_{l,t}), & (1-X_{k,t})X_{l,t}, & (1-X_{k,t})(1-X_{l,t}) \end{array} \right], \tag{4.5}$$

for female k and male l of pair j. For the alternative formulation care must be taken when defining f_i, the first occasion in which individual i has entered the experiment. For example if the known mate of individual i enters the experiment prior to the first capture of individual i, the first occasion of the mate will be used.

Generally, we prefer formulation (4.4, 4.5) as it allows more flexibility for future modifications, such as the eventual inclusion of mate switching.

4.2.3 ρ-CJS Observation Process

All CJS models condition on first capture, restricting inference to individuals that have been observed, removing the first capture probability from the likelihood. For the ρ-CJS framework first captures differs somewhat, because entities may consist of either single individuals or pairs of individuals. For single individuals, the first capture conditioning does not change from standard CJS models. However, for entities consisting of pairs we approach first capture differently as there are three possible capture outcomes: female only (F0), male only (M0) and both (FM).

Let e_j be the occasion where entity j is encountered for the first time. The observation model for the first capture occasion will be

$$\mathbf{Y}_{j,t}|\mathbf{Z}_{j,t} \sim \text{Multinomial}\left(1, \mathbf{Z}_{j,t} \begin{bmatrix} \frac{p^{fm}_{j,t}}{1-p^{00}_{j,t}} & \frac{p^{f0}_{j,t}}{1-p^{00}_{j,t}} & \frac{p^{m0}_{j,t}}{1-p^{00}_{j,t}} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\right) \text{ for } t = e_j. \tag{4.6}$$

The first row of (4.6) contains pairwise capture probabilities for the three possible capture outcomes, conditioned on the overall probability of capture $1-p^{00}_{j,t}$. The the second and third rows correspond to states where only the female or male are alive on first capture. Here we use a value of 1 which is equivalent to removing the first capture probability. Entities were both individuals are dead cannot be captured also receive a value of 1 which corresponds to neither individual individual being seen.

Subsequent captures are modeled in a manner that is consistent with multistate mark-recapture models, with the following observation equation

$$\mathbf{Y}_{j,t}|\mathbf{Z}_{j,t} \sim \text{Multinomial}\left(1, \mathbf{Z}_{j,t}\begin{bmatrix} p^{fm}_{j,t} & p^{f0}_{j,t} & p^{m0}_{j,t} & 1-p^{fm}_{j,t}-p^{f0}_{j,t}-p^{m0}_{j,t} \\ 0 & p^{F}_{j,t} & 0 & 1-p^{F}_{j,t} \\ 0 & 0 & p^{M}_{j,t} & 1-p^{M}_{j,t} \\ 0 & 0 & 0 & 1 \end{bmatrix}\right) \text{ for } e_j < t \leq T. \tag{4.7}$$

If both individuals are alive (the first row) we used the derived pairwise capture probabilities (e.g. $p^{fm}_{j,t}$), otherwise we use the appropriate marginal capture probability $p^{s}_{j,t}$. For the state where both individuals are dead, neither are captured with probability 1.

Special Case: Permanent Male Emigration

One of the interesting characteristics with the McLeod harlequin study is possibility that males emigrate from the site when their mate dies. In this case we would expect single males to only produce '0' observations, resulting in the following alternate observation equation

$$\mathbf{Y}_{j,t}|\mathbf{Z}_{j,t} \sim \text{Multinomial}\left(1, \mathbf{Z}_{j,t}\begin{bmatrix} p^{fm}_{j,t} & p^{f0}_{j,t} & p^{m0}_{j,t} & 1-p^{fm}_{j,t}-p^{f0}_{j,t}-p^{m0}_{j,t} \\ 0 & p^{F}_{j,t} & 0 & 1-p^{F}_{j,t} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}\right) \text{ for } e_j < t \leq T \tag{4.8}$$

used in place of the original observation equation (4.7).

4.2.4 Likelihood Formulation

Statistical inference (Bayesian or frequentist) requires a fully formed likelihood to be specified. Assuming that mating pairs are independent of one another, the likelihood can be constructed as the product of the individual mating-pair likelihood components. As such, the likelihood component of mating-pair j is the probability of obtaining observation matrix $\mathbf{Y}_j = (\mathbf{Y}_{j,e_j}, \ldots, \mathbf{Y}_{j,T})$ determined by integrating over all the possible states. For pair j the likelihood component will determined as

$$\int_{\mathbf{Z}_{j,e_j}} \cdots \int_{\mathbf{Z}_{j,T}} \left[\mathbf{Y}_{j,e_j}|\mathbf{Z}_{j,e_j}\right] \left[\mathbf{Z}_{j,e_j}\right] \left\{ \prod_{t=e_j+1}^{T} \left[\mathbf{Y}_{j,t}|\mathbf{Z}_{j,t}\right] \left[\mathbf{Z}_{j,t}|\mathbf{Z}_{j,t-1}\right] \right\} d\mathbf{Z}_{j,e_j} \dots d\mathbf{Z}_{j,T} \tag{4.9}$$

where $\left[\mathbf{Y}_{j,e_j}|\mathbf{Z}_{j,e_j}\right]$ is the distribution describing the first captures (4.6) and $\mathbf{Z}_{j,e_j}$ is one of three possible initial alive states, which is assumed to be known with certainty. The distribution $[\mathbf{Y}_{j,t}|\mathbf{Z}_{j,t}]$ is used for subsequent captures (4.7). Finally, $[\mathbf{Z}]$ is the distribution used to describe our latent alive/dead status matrix $\mathbf{Z}$ and may be specified by either (4.3) or (4.4 and 4.5) state formulations. The full likelihood is then formed by taking the product of the m (pairs) independent components.

4.2.5 Model Assumptions

Standard assumptions under the mark-recapture framework are that tags are not lost, overlooked and are recorded correctly, sampling is instantaneous, emigration from the sampling area is permanent and the fate of each marked individual (capture and survival) is independent to that of any other individual (e.g. Pollock et al., 1990; Williams et al., 2002). We relax the last standard assumption so that the capture of each entity (the mating pair or singleton) is independent, as well as making some new assumptions.

It is assume that mating pairings are known with certainty and can be discerned at some point during the experiment. For pairs where only one mate is caught, we assume it is possible discern the captured individual has a mate. Second, we assume that all pairwise matings occur prior to a pair's entrance into the experiment. Furthermore, individuals that enter the experiment without a mate, will not become mated during the course of the experiment. While these last assumptions may not be realistic for all system the provide a starting point for modeling pairwise dependence in mark-recapture experiments.

4.2.6 Example Probability Statement of a Capture History

It is often helpful to consider the probability expression for sample capture histories. For example, we will use the capture history `301` to indicate a mating pair that was captured together on the 1st occasion, both missed on the 2nd occasion and only the female was caught on the 3rd occasion. In this history we know the female survived the duration of the experiment, but the male was only observed on the first occasion.

To develop a probability statement for history `301` we need to consider the possible fates of the female along with the male simultaneously. The resulting marginal probability statement will be

$$P(301) = \frac{p_{j,1}^{fm}}{1 - p_{j,1}^{00}} \phi_{j,1}^{F} \left(\phi_{j,1}^{M} p_{j,2}^{00} \left(\phi_{j,2}^{F} \left(\phi_{j,2}^{M} p_{j,3}^{f0} + (1 - \phi_{j,2}^{M}) p_{j,3}^{F} \right) \right) + (1 - \phi_{j,1}^{M})(1 - p_{j,2}^{F}) \phi_{j,2}^{F} p_{j,3}^{F} \right). \tag{4.10}$$

where the subscript j refers to the j^{th} pair of animals. At first glance the probability statement may seem quite complex, but it can be broken down into smaller pieces by consider the possible each unique combination of fates separately (Figure 4.1).

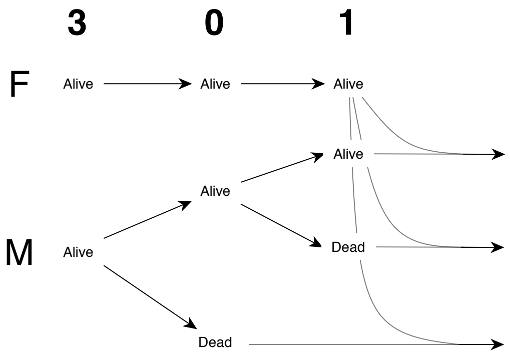
Common to all three unique combination outcomes is the leading term $\frac{p_{j,1}^{fm}}{1-p_{j,1}^{00}}$ which represents the probability of observing both mates on the first capture occasion, given a capture event occurred and both mates were alive. For the situation that only one of the two mates were alive we would exclude the first capture probability as is the case in standard CJS models.

After first capture each of the possible unique paths are modeled in a similar fashion to standard CJS models except at each time interval we need to include the correct sex-specific survival probabilities (e.g. $\phi_{j,t}^{F}$) for both the male and female of a given pair. The only real departure from standard CJS type models is the change in capture probability terms used which depends on whether or not the male is alive on a given occasion. For example in both fates (a) and (b) in Figure 4.1 the probability of observing a '`0`' on the second occasions is $p_{j,2}^{00}$ because both the female and male are alive on the second occasion, but for (c) the capture probability is $(1 - p_{j,2}^{F})$ because only the female is alive. Summing over these individual fate outcomes gives us back our original marginal probability statement (4.10).

4.2.7 Contrasting ρ-CJS and Standard CJS Models

While the ρ-CJS considers the history of mating pairs, it should be noted that without knowledge of the mate pairing structure, there is no way to distinguish the observations from a ρ-CJS experiment with that of a standard CJS models. For example, if we consider again the capture history `301`, we can view it as be constructed of a male `100` history and a female `101` history.

If we deconstruct the the combined female/male histories into individual histories and



$$P(301)$$

$$=$$

$$\frac{p^{fm}_{j,1}}{1-p^{00}_{j,1}}\phi^{F}_{j,1}\phi^{M}_{j,1}p^{00}_{j,2}\phi^{F}_{j,2}\phi^{M}_{j,2}p^{f0}_{j,3} \qquad \text{(a)}$$

$$+$$

$$\frac{p^{fm}_{j,1}}{1-p^{00}_{j,1}}\phi^{F}_{j,1}\phi^{M}_{j,1}p^{00}_{j,2}\phi^{F}_{j,2}(1-\phi^{M}_{j,2})p^{F}_{j,3} \qquad \text{(b)}$$

$$+$$

$$\frac{p^{fm}_{j,1}}{1-p^{00}_{j,1}}\phi^{F}_{j,1}(1-\phi^{M}_{j,1})(1-p^{F}_{j,2})\phi^{F}_{j,2}p^{F}_{j,3} \qquad \text{(c)}$$

Figure 4.1: The capture history `301` will consist of a female that is known to be alive for all three sampling occasions and a male that has three potential fates over the experiment. The overall probability of this history is the combination of the three possible female/male fates.

determine the expected frequencies, it can be shown (rather surprisingly) that these frequencies do not differ between the ρ-CJS and standard CJS models if we exclude the `00..0` history which can only occur in the ρ-CJS setting. This is the result of restricting the marginal capture probabilities when paired to be the same as the capture probabilities of single individuals.

4.3 Numerical Example

We compare a standard CJS type approach against the proposed ρ-CJS framework on a simulated data set with a known amount of pairwise capture dependency. Data was simulated the statistical programming language R (R Development Core Team, 2010) with parameter values that appear to be reasonably based on the harlequin study. Fates (alive/dead states) were simulated for each individual using a constant survivorship across all individuals, occasions and sexes (i.e. $\phi^{s}_{j,t} = \phi = 0.8$ for $s \in \{F, M\}$). A subset of individuals were then placed in pairs and multinomial draws for pairwise capture outcomes were made for individuals in mating pairs using the pairwise probabilities of Table 4.1 with $\rho = 0.25$. Captures for single individuals were generated using binomial draws. In both cases a sex specific capture probabilities was used that was constant across occasions and individual/pairs (i.e. $p^{F}_{j,t} = p^{F}$ and

$p^M_{j,t} = p^M$). The capture outcomes and pairing information was then used to construct the observation matrix.

In order to fit the standard CJS model we had to exclude individuals with a `0...00` individual capture history as these histories cannot be observed in a regular CJS type framework. In contrast, the ρ-CJS framework can produce this type of individual history if we know the mate exists but the mate was never captured. In this case the `0...00` type history may be hidden in summary histories such as `101` or `202`.

In total we simulated histories for 1500 individuals over $T = 8$ occasions. Not all individuals entered the experiment, resulting total of 1463 individuals included in the ρ-CJS model and 1407 individuals included in the standard CJS model.

The parameter estimates for the CJS data set were obtained through maximum likelihood using Program MARK (White and Burnham, 1999) and the parameter estimates for the ρ-CJS model were obtained using a Bayesian approach and implemented in WinBUGS (Spiegelhalter et al., 2003). The WinBUGS code is provided in Appendix B.2. Multiple chains were initiated, the first 10,000 samples from the posterior were discarded for burn-in, then an additional 100,000 samples of the posterior were taken with a thinning of 10, to reduce autocorrelation between some of the parameters. This resulted in a total 10,000 usable samples from the posterior distribution for each parameter.

As would be expected both approaches yield point estimates that appear to be unbiased (Table 4.2). However the ρ-CJS model was able to successfully estimate the degree of correlation between capture outcomes, whereas the standard CJS model is naive to such processes. Furthermore the ρ-CJS also produced slightly lower measure of uncertainty, due to the extra information used by ρ-CJS framework (i.e. individual `0...00` histories). Goodness-of-fit tests available in Program MARK did not indicate a problem except for $\hat{c}$, the estimate of model over-dispersion. The observed $\hat{c}$ was 1.68, with a bootstrapped $\hat{c}$ value of 1.88 (se 0.15). Both show an indication of over-dispersion in the data relative to the model. A $\hat{c}$ value of 1.88 was used to adjust the standard error estimates of naïve CJS model, which increased the standard error estimates slightly (Table 4.2).

4.4 Discussion

We present an important first step to laying the framework needed for the consideration of dependency structures in mark-recapture experiments. Of the standard assumptions made

Table 4.2: Parameter estimates from simulated data for both the standard CJS approach, which is naive to any dependency in pairwise capture outcomes and the ρ-CJS approach which models this type of dependency. Standard errors for the naïve CJS model were not adjusted for over-dispersion.

Parameter	True Value[1]	Naïve CJS (MARK)			ρ-CJS (WinBUGS)			
		Est.	SE	SE[2]	Mean	SD	95 % CI	
ϕ	0.800	0.794	0.006	0.008	0.794	0.006	0.784	0.805
ρ	0.250	-	-	-	0.258	0.031	0.196	0.320
p^F	0.731	0.746	0.011	0.015	0.740	0.009	0.722	0.758
p^M	0.818	0.826	0.010	0.013	0.822	0.009	0.805	0.839
p^{00}	0.092	-	-	-	0.090	0.008	0.075	0.105
p^{f0}	0.091	-	-	-	0.088	0.007	0.076	0.101
p^{m0}	0.177	-	-	-	0.170	0.008	0.154	0.187
p^{fm}	0.641	-	-	-	0.652	0.010	0.632	0.671

[1] Parameter values used in that data simulation. The concept of a "true" parameter value does not exist in a Bayesian framework.

[2] Standard errors adjusted for over-dispersion using a $\hat{c}$ value of 1.88.

in the CJS modeling framework (see Pollock et al., 1990; Williams et al., 2002), this is the first time the long standing assumption of independent fates among individuals has been approached. While violations of independence are often cited as an issue in analysis of model misfit, we are the first to provide a potential solution.

In this formulation we focused on modeling the pairwise dependency that may occur in capture outcomes only, assuming independence in the survivorship process. While dependency in survivorship may exist, we feel this restriction is justified. Mark-recapture sampling occasions are considered instantaneous and tend to occur during periods of reproduction when the majority of individuals in the population are available for sampling. Under these conditions we may expect individuals in a mating pair to be associated together more often with each other than with individuals who are not in the mating pair. However, this association may be temporarily seen during sampling periods only, after which the mating pair may or may not share fates. This makes it unclear whether or not survivorship dependency needs to be handled in addition to capture dependency.

We suspect dependency in survivorship will be highly dependent on the system under study. If individuals in a mating pair are always be associated with one another, then we see no reason why survivorship dependency cannot be handled in a manner similar to the method we have outlined for captures. However, if individuals in a mating pair are only associated together temporarily, then considerations of dependency may not be necessary. Note however that these considerations of survivorship dependency are based on individuals sharing fates.

Dependency in survivorship may also exist if individuals in a mating pair select other individuals of similar quality. In this case, we may expect individuals of lower or higher quality to be associated with other individuals of similar quality. Because quality and survivorship are often related, individuals with higher or lower survivorship may also be associated together, due to mate selection. We suspect this possibility could be handled with the careful use of random effects between individuals in a pair and represents a simpler modeling case than dependency in captures.

Any of these future considerations will be highly dependent on a flexible modeling framework. To these ends we were successfully able to implement a state-space model in a Bayesian framework using the common Bayesian software package WinBUGS. state-space implementation as this type of model formulation is very promising for modeling ecological data sets (see Buckland et al., 2004; Gimenez et al., 2007).

In terms of the model presented, we prefer the individual state equation using Bernoulli trials as opposed to provided multinomial state equation. This type of formulation will be easier to carry forward as additional model features are added and also allows for the incorporation of additional information about the alive/dead states if there exists knowledge outside the current study. For example, an individual may be observed as being alive in another study, but was not captured recently in the current study. This type of state-space formulation makes including this type of additional information fairly trivial. Furthermore, if we are to accommodate more complex forms of interactions such as mate switching, an individual state equation will be essential. Lastly it should be noted that the current model may also be formulated in the multievent framework (Pradel, 2005).

While we have taken the important first steps towards solving this problem, we are still not in a position to approach the motivational data set due to various intricacies such as the occurrence mate switching, which the proposed model currently does not handle. Mate switching can occur even in socially monogamous species when one of the mates dies, or in the context of our framework, mate switching may occur if an individual enters the experiment single then finds a mate at a later occasion. Switching mates adds additional problems as we can no longer represent the experiment in terms of a simple multi-state capture history, because the definition of an entity (in our case either a single individual or a mating pair) changes as the mate changes.

There are a couple of potential approaches that could be taken to handle the problem. First the standard capture history data format will have to be changed to an single entity (single or mating pair) format, with information about the individuals making up the entity and the capture occasions. Currently our WinBUGS model takes this approach to the data.

Secondly, mate switching requires that we either know the exact occasion, or range of occasions that switching takes place, otherwise another process will need to be included in the model that accommodates for the process of pairing, breakups and repairing. For any of these scenarios we must also carefully consider the entrance indicator for each individual, as this is used in part to correctly set the period we know an individual is alive. Incorrectly specifying the entrance will lead to biased parameter estimates. Specifically, entrance indicator must accommodate for the situation where one of the mates is captured earlier in the experiment, but at the time is not paired to the individual of interest. In this case the mate that was captured earlier provides no additional information about when the individual of interest entered the experiment.

Setting aside the topic of entrances, handling switching is still a non-trivial exercise for all situations except when the exact timing of the switch is known. Since this is an unlikely situation we will consider the situations where switching takes place during some interval or we need to introduce a new process to the model.

If we only know that switching took place within an interval we may opt to condition out those capture observations from the observation equation, since we do not know with observations came from a single individual and which capture outcomes came from a paired individual. Alternatively, we can introduce another process into the model representing the pairing process, for example a pairing probability. Introducing this would then allow one to integrate over the possible states in the known switching interval.

While this may seem like a straight forward solution, introducing a pairing probability also introduces additional problems. For starters there is the implicit assumption that there is no restriction in the availability of mates, and that all single individuals wish to form mating pairs. While the latter assumption is not particularly problematic, the former could be in smaller populations. Incidentally, we would also expect smaller populations to be the prime candidate for this type of modeling framework due to the work required to acquire a detailed mating record. This problem could be overcome if we knew the exact number of available individuals, however this would also require moving away from CJS framework towards a Jolly-Seber (JS) framework (see Schwarz and Seber, 1999; Williams et al., 2002), where individuals available for capture are explicitly modeled. Interesting, if an adequate solution is determined, modeling the pairing process would also one to relax our initial assumption of knowing the initial entity state. Similar to switching, one could also integrate over the possible initial states, rather than setting them as a knowing quantity.

Generally, this direction in modeling approach is also consistent with the general approach of incorporating more information on the alive/dead state of an individual than a regular CJS type framework allows. Typically, in a CJS model, we only know about the interval from first capture to last capture and the birth/entrance to the study is left censored and the death/exit from the study is right censored. The CJS framework conditions on first capture to remove the issue of left censoring and integration to handle the second. In contrast the JS framework integrates out both types, thereby making stronger assumptions. Our proposed model is somewhere in between as we are able to include some information outside the regular CJS interval.

In particular, we can include more information about the alive/dead state when a mate is known to be alive due to its association with its mate (e.g. the male is not observed but the female is still producing offspring). Depending on the situation we can widen the interval we know a given animal is alive, providing more information and more precise estimates. In addition, unlike the regular CJS framework the proposed ρ-CJS framework is explicitly able to include `0...00` type individual histories, for example a `101` history could be a mating-pair where only the female was caught (a `101` female history and a `000` male history) or a single female. The difference being that we knew the initial state upon capture. Again this starts to blur the line between a JS and a CJS type model.

Another interesting feature of our proposed modeling framework is that if we break the joint capture histories into individual capture histories (e.g. a `303` history consists of both a female `101` history and a male `101` history) we find that the expected frequency of observed individual capture histories (excluding the `0...00` type individual histories) is identical to a CJS type framework.

Regular CJS type experiments contain no information about pairwise capture dependencies and the mating information is an auxiliary component to the regular mark-recapture type experiment. The equivalence in observed frequencies is due to our assumptions that pairwise captures are distributed as dependent Bernoulli trials. In a ρ-CJS expiment we restricted the marginal capture probabilities of paired individuals to be the same as the capture history for single individuals, thereby producing similar observed frequences. While the expected frequencies are equivalent other distributional characteristics will differ between a ρ-CJS experiment and a regular CJS type experiment. Furthermore this outcome is only valid if pairwise captures follow our assumptions, to which we also provided an alternative capture model that will likely result in a divergence in observed frequencies between regular CJS and ρ-CJS experiments.

Often mark-recapture experiments produce data that is over-dispersed relative to the models. Typically this is quantified with the $\hat{c}$ statistic, the model dispersion divided by the degrees of freedom, where values greater than one show potential problems of over-dispersion. Model over-dispersion is often attributed to problems such as non-independence without citing definitive proof. Now that the initial workings it would be worth investigate the effect varying levels of pairwise dependency has on the estimated uncertainty of model parameters.

Part II

The Multiple Season Multi-State Occupancy Framework

Chapter 5

Modeling Ecology Processes Using Ecological States

5.1 Introduction

Understanding driving forces behind changes in wildlife populations is a central scientific goal for ecologists, conservationists, and wildlife managers, yet after a century of research many topics still remain hotly debated. For example, the mechanisms behind density dependence, a well accepted and often cited population process, is still being called into question (e.g. the role of individual variation as a mechanism; Benton et al., 2006). In general, a variety of factors have been argued to be behind the difficulty. These have ranged from the type of mathematics historically used (Nieminen, 2008), to the difficulties involved in obtain high quality data sets (e.g. Wolda and Dennis, 1993; Hanski and Henttonen, 1994). We side with the latter concern over the availability of quality data sets and further observe that most studies have chosen to focus on sampling individuals which can be notoriously difficult.

While individuals may provide the ultimate level of inference, sampling can be notoriously difficult. Species are often under study because they are rare or endangered, with populations that are sparsely distributed across the landscape. This naturally leads to sampling inefficiencies. Sampling individuals often requires physical capture, presenting managers and conservationists with a potential trade-off between gaining information and potential disturbance. In particular, when studying rare species a disproportionately larger percentage of the population must be sampled in order to obtain estimates of acceptable precision. This in turn represents a larger impact due to sampling than would be the case for a more plentiful species. A balance then between impact on the study species (through sampling) and knowledge must be forged. Muddying this situation further is the fact that rare and endangered species are often regulated by stronger governmental restrictions, adding administrative overhead and further costs to study deployments.

5.1.1 Ecological States

These types of shortcomings give rise to other approaches such as using ecological states. An ecological state is defined simply as any unique configuration of ecological characteristics considered important by expert opinion (e.g. Stringham et al., 2003). State-and-transition type models have gained widespread use in rangeland ecology as an effective way to model plant communities (Herrick et al., 2006; Bestelmeyer et al., 2009; Knapp and Fernandez-Gimenez, 2009) and has even gained governmental agency adoption (Stringham et al., 2003). Within wildlife ecology, a state-based approach has also been gaining traction, with examples including relative measures of abundance (e.g. occupancy MacKenzie et al., 2002; Tyre et al., 2003; Royle and Link, 2005), species richness (Wickham et al., 1997) as well as other indicators of population health (e.g. active breeding Nichols et al., 2007). The use of ecological states is also compatible with seminal ideas in ecology such as patch dynamics (Wu and Loucks, 1995; Wu and Levin, 1997)

Ecological states shift the focus to geographical units, often termed sites, as the statistical "population". Observations are formed by visiting the site and recording an ecological state. For example, in presence/absence data the observer records whether or not a particular species is present. Statistical inference may then be focused on estimating the proportion of landscape exhibiting a given ecological state or modeling the probability that a given state is exhibited as a function of covariates. For presence/absence data (usually denoted as a '1' or '0') the proportion of the landscape occupied is often estimated, and can be used as a surrogate measure for abundance. Monitoring studies would then track changes in the underlying population size, by tracking the increase or decreases in the proportion of the landscape occupied.

Compared to methodologies focused on individuals, ecological state variables also can be more efficient in terms of sampling effort as well as providing a way to model available knowledge. For example, the presence of a species can be confirmed by only observing one member of a species. In contrast, techniques such as mark-recapture need to repeatedly observe the *same* individual. Also it can be easier to incorporate available information (e.g. traditional knowledge) with scientific knowledge using ecological states (Knapp and Fernandez-Gimenez, 2009).

While there are potential benefits to using ecological states, issues of detection have been largely overlooked. Detection issues have long been considered to be a common problem in wildlife abundance data (see Williams et al., 2002), yet it has been largely ignored for ecological state data (MacKenzie et al., 2002; Tyre et al., 2003; Royle and Link, 2005; MacKenzie et al., 2006). For example, even our ability to detect even the most basic of ecological states, the presence or absence of a species, is known to be quite limited (Craig and Roberts, 2001; Lindenmayer et al., 2001; MacKenzie et al., 2002; Tyre et al., 2003). Ignoring these issues will bias estimates of ecological processes, even when there are only small levels of detection error (MacKenzie et al., 2002; Martin et al., 2005).

In this portion of the thesis we build an analysis framework for handling ecological state data, with special consideration given to handling detection issues. Building on the work of MacKenzie et al. (2002) and Royle and Link (2005), who proposed the site-occupancy framework and the single-season mixture multinomial model, we derive a multiple season, multi-state occupancy framework based on the multievent model of Pradel (2005).

5.2 Motivating Example: The North American Amphibian Monitoring Program

Our intent is to develop and refine techniques that have real-world applicability. In doing so it can be instructive to consider a motivating example when constructing the analysis framework. In our case we have chosen the North American Amphibian Monitoring Program (NAAMP), which uses call indices (a measure of call frequency) as a surrogate for abundance.

Amphibians have been experiencing world-wide declines (even in relatively pristine environments) and are considered “bioindicators” for environmental health due their high sensitivity to contaminants (Blaustein, 1994; Gardner, 2001). Consequently, amphibian monitoring programs have been initiated across Canada and the United States to gather baseline data and track changes in anuran populations (frogs and toads). These programs rely on volunteers to collect abundance data and as a result have focused on the use of surrogate measures of abundance which are easier to obtain and require less training than more direct techniques such as mark-recapture. Generally, ponds with higher anuran abundances have higher vocalization rates (Nelson and Graves, 2004). A call index is simply a standardized categorization of the call frequency (e.g. 0-none, 1-low, 2-medium, 3-high) that was formalized and standardized with the creation of the NAAMP. Changes in the underlying abundances is then monitored by tracking changes in the call index. For example, ponds showing consistently showing lower call indices over time are assumed to be undergoing a decrease in the underlying population abundance.

For simple presence/absence data, MacKenzie et al. (2002) proposed the occupancy modeling framework as a way to handle detectability, but the occupied/unoccupied dichotomy for the ecological state is too coarse. Site-occupancy models take an areas of interest are divided into sites (in the case of the NAAMP, ponds are sites), which are repeatedly sampled for the observed presence or absence of a species. Observations are then summarized as an encounter history, for example '`101`' is a site that was surveyed three times and a species was observed on the first and third occasions. Site-occupancy models treat the underlying state of the site as latent (occupied with probability Ψ) with the observation being dependent on the underlying state. For occupied sites a '1' is observed with probability p_t on survey t. If surveys occur close together in time we can assume the underlying state has not changed and the probability of observing that encounter history becomes

$$P(\texttt{101}) = \Psi p_1(1 - p_2)p_3. \tag{5.1}$$

The outcome can then summarized in terms of encounter histories, which is modeled as multinomial with the unique encounter making up the possible outcomes and the encounter history probabilities as the cell probabilities.

By modelling states (occupied/unoccupied) and observations (detected/undetected) separately, finely nuanced models can be created to suit a variety of unique biological situations. For example the probability of observing the species may be defined on a site- and survey-specific level (e.g. p_{it}), which can then be modeled as a function of site-specific covariates. In a similar manner a site specific probability of occupancy Ψ_i can also be modeled as a function of covariates.

While a flexible framework for presence/absence data, the occupied/unoccupied dichotomy for state variable is too coarse for many ecological situations. In the context of the NAAMP data set, ponds that produce a call index of '3', '2' or '1' would all be treated as 'present' (e.g. Weir et al., 2005). Areas that consistently produced a call index of '3' would be treated as possessing the same ecological state (occupied) as those producing an index of '1', even though higher index values are known to be associated with higher abundances. An *ad hoc* approach to deal with this multi-state data in the site-occupancy framework would be to model detection probability p and/or occupancy probability Ψ as a function of the call index value (e.g. we may expect sites with higher abundance will also have a higher site-specific value of p). This of course is arguably a backwards approach to modeling the true nature of the data.

For data sets such as the NAAMP, an explicit multi-state occupancy framework is the next logical step. Most ecological state variables can be thought of in terms of multiple states of species occupancy. Examples include relative measures of abundance (ordinal abundance levels), successful reproduction (occupied and occupied with reproduction) or even in terms of co-occurrence of one or more species (species A, species A and B, etc). Similar to the classical site-occupancy models, such a framework should also give special consideration to issues of detectability. Here the issue of "detectability" becomes more complex as there is general failure to observed the species in any state (detection) or a site may be observed as occupied, but is mistakenly classified as exhibiting an ecological state other than the true state (misclassification). Finally, a multi-state occupancy framework must also accommodate multiple years or seasons of data, so that expected changes in the underlying ecological state may be estimated. This will allow trends in the ecological state (and factors affecting those trends) to be modeled and estimated, giving ecologists and managers important tools to track and interpret important changes in wildlife populations.

5.3 Hierarchical Models and the Multievent Framework

5.3.1 Hierarchical Models

Hierarchical statistical models are recognized as one of the most promising approaches for specifying complex ecological models (e.g. Royle and Young, 2008; Bolker et al., 2009; Ponciano et al., 2009). These types of statistical models allow researches a great amount of flexibility. For example, extra variability may be included in parameters that may otherwise be treated as fixed. This can correctly address issues such as individual heterogeneity, survey designs (such as blocking) or to acknowledge the often stochastic nature of environmental processes. However, not all hierarchal models are equally appropriate for ecological data.

Central to this issue are the model components and the handling of over-dispersion. Hierarchical approaches such as generalized linear mixed models (GLMM) have been suggested as a practical approach for creating ecological models (e.g. Bolker et al., 2009), but oftentimes neglect important considerations such as detectability. GLMM type approaches do not clearly separate the ecological mechanism from the observation process, instead representing observations as being generated from a single process, potentially with random effects included as an additive effect.

This presents two drawbacks, because we know detection to be an important consideration in ecological data (Craig and Roberts, 2001; Lindenmayer et al., 2001; MacKenzie et al., 2002; Tyre et al., 2003). Without a clear modeling paradigm that incorporates detection, researchers may under appreciate the relevance of the issue. It makes it easier to neglect model components representing how observations are obtained and may lead to models which erroneously view over-dispersion as an outcome of variability in the ecological mechanism, rather than as a consequence of the ecological mechanism and the inability to fully observe the system.

Royle and Young (2008) argued for a type of hierarchical modeling approach that has long been employed in mark-recapture and more recently occupancy models. These types of models define separate components for ecological mechanism or state and the observation process, where the expected value (conditional on the underlying state) is used for building models around the observation process.

For example, in site-occupancy models, the underlying state will either be occupied or unoccupied (denoted as a 1 and 0 respectively). We let Z represent the latent occupancy state. When visiting a site we can observe the site as either being occupied or empty (also denoted as 1 or 0 respectively. We will let Y represent our observed occupancy status. If a site is occupied, that is $z = 1$, then we can expect to observe a '1' with probability p (the detection probability), which we can summarize as $[y|z = 1] \sim \text{Bernoulli}(p)$. More complex models for the positive detections can then be constructed, by modeling the expected value, p, as a function of covariates. Generally, if $z = 0$ only a '0' will be observed, which we can similarly state as $[y|z = 0] \sim \text{Bernoulli}(0)$. Traditionally, this latter distinction is not directly stated, since mark-recapture and occupancy models have typically worked directly with the marginal distribution $[y|Z][Z]$, since their first introduction (Jolly, 1965; Seber, 1965). We will be looking at this latter distinction in more detail shortly.

Experiments involve multiple sampling occasions, and as such the entire summary of observations, termed histories, is modeled as using multinomial distributions. Each unique history capture (mark-recapture) or encounter (site-occupancy) history makes up one of the multinomial outcomes. Returning to our earlier example, a history of '`101`' for a site-occupancy model indicates an experiment where a site was visited on three occasions and the species observed on the 1st and 3rd sampling occasions only. The biological relevance of the model is incorporated when we define the multinomial cell probabilities. In this case we must first define Ψ, the probability that a site is occupied, and p_t as the probability of a positive detection in survey t. Assuming the underlying state of the site has not changed between the three surveys we get the encounter probability written in (5.1).

Because the species was observed in at least one of the occasions, we know the site was occupied, simplifying our probability statement. For cases of the all zero history (i.e. '`000`'), we can't be certain of the underlying state so we must consider both possibilities. In this case

$$P(\texttt{000}) = 1 - \Psi + \Psi(1 - p_1)(1 - p_2)(1 - p_3).$$

which accounts for the possibilities the site was not occupied, or was occupied, but not detected on any surveys.

In both cases we are directly specifying the marginal probability and have implicitly accounted for all possible states. While on the surface this may seem straight forward, but as the number of possible states increase, specifying the marginal probability becomes more problematic because we need to consider all combinations simultaneously. As a result it can be easy to overlook implicit assumptions or completely overlook certain outcomes.

Keeping with our first example, the probability statement for $P(\texttt{101})$ assumes that no false positives can occur, that is an unoccupied site will only produce observations of '0'. We can make this assumption clearer by defining q_t to be the probability that an unoccupied site produces a '0', thus producing the following,

$$P(\texttt{101}) = \Psi p_1(1 - p_2)p_3 + (1 - \Psi)(1 - q_1)q_2(1 - q_3).$$

This new statement accounts for the possibility that the encounter history '`101`' was produced by either an occupied or unoccupied site. By making the explicit constraint $q_t = 1$ (i.e. unoccupied site produces a '0' with certainty) we can retrieve our original probability statement.

The problem with the first probability statement (5.1) is that at a glance it is not entirely self-evident that we have integrated across all states. While such assumption may seem trivial for simple models like the site-occupancy model, multiple states of occupancy will creates many more types of misclassification errors. A clear modeling paradigm is needed to facilitate complex multi-state occupancy models, so that can be readily understood by practitioners of varying mathematical abilities.

Ideally, focus should be placed on defining the components of the system, rather than marginal probability statements. For example, what are the latent ecological states of interest, how are these states related to each other and over time (i.e. can states change between sampling occasions; can one state transition directly to the next?) and what types of observations can be produced by a particular state. Once specified, these components should naturally define the marginal likelihood, without having to define complex probability statements.

5.3.2 Incorporating the Multievent Framework

Pradel (2005) provided a significant step forward in terms of model formulation with the multievent framework for the multi-state Cormack-Jolly-Seber (CJS) class of mark-recapture models. The frameworks attempts to model the true nature of capture histories, by looking at the observed data in terms of latent states and observations, which are termed events. On a given sampling occasion an animal exhibits one of a set of possible states, for example being alive is one obvious state. The state however is not directly observable, instead an event is observed from a set of possible events that may be associated with a given state.

The use of the term “event” is used to clarify the concept that observations are separate but related to underlying state of interest. For example, consider an experiment where the states of interest are alive, alive and breeding and dead. Determining breeding status directly can be difficult, instead events such as “sitting on an egg” or “standing near a nest” may give differing insights into the underlying state. For the first event it is clear that the animal is alive and breeding, but for the second it is only clear the animal is alive, whether or not it is breeding cannot be determined. As such, the second event should be modeled as coming from either alive states with appropriate event probabilities, while the former event would from the alive and breeding state, which would have different event probabilities.

This alone is not a radical departure from multi-state CJS models already in use, the true strength comes in simplified mathematics used to define the model. More specifically, the multievent framework dispenses with the need to explicitly provide overly complicated marginal probabilities. Instead, model formulation is broken down into component pieces which include an initial state vector, as well as transition and events probability matrices. Because both mark-recapture and occupancy models share a number essential characteristics in their basic formulations, we can easily adapt approach to the occupancy framework.

Returning to our previous occupancy example, where we consider both missed detection and false positives, we can take a multievent type approach by first defining state and event components. Our states would be a row vector containing the probability of being either unoccupied or occupied,

$$\begin{bmatrix} 1-\Psi & \Psi \end{bmatrix}.$$

Next a 2 × 2 event probability matrix needs to be defined to account for all possible observed events, conditional on the underlying state. In this case we have two events (0 - not observed, 1 - observed as present) which may occur for each of the two states (0 - unoccupied and 1 - occupied). Using the same notation as before, the resulting event matrix becomes

$$\begin{array}{ccc} & \multicolumn{2}{c}{\text{event}} \\ \text{state} & 0 & 1 \\ 0 & q_t & 1-q_t \\ 1 & 1-p_t & p_t \end{array}.$$

Here each row must sum to one as events are mutually exclusive. By arranging the events in a matrix it is readily apparent that false positives (i.e. observing an ‘1’ when the site is unoccupied) are possible.

With the state and event components defined, the last task is adapting the multievent framework is to determine the marginal probability of an encounter history using matrix algebra. To do so we employ a simple mathematical trick where for each observation we take the corresponding column from the event matrix and place it on the main diagonal before taking the product allowing the mathematical statement to resolve correctly. For the example encounter history of ‘`101`’ again, the marginal probability is then defined as

$$\begin{aligned} P(\texttt{101}) &= \begin{bmatrix} 1-\Psi & \Psi \end{bmatrix} \begin{bmatrix} 1-q_1 & 0 \\ 0 & p_1 \end{bmatrix} \begin{bmatrix} q_2 & 0 \\ 0 & 1-p_2 \end{bmatrix} \begin{bmatrix} 1-q_3 & 0 \\ 0 & p_3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad (5.2) \\ &= \Psi p_1(1-p_2)p_3 + (1-\Psi)(1-q_1)q_2(1-q_3). \end{aligned}$$

The final column vector $\mathbf{1}$ sums the separate probability statements for each state, making the overall equation a marginal probability.

This is very similar to the probability statement used in the multievent model, where the probability of capture history h is determined as,

$$P(h) = \pi_e D\left(B_e^0(o_e, \cdot)\right)\left(\prod_{i=e+1}^{T} \Phi_{i-1} D(B_i(o_i, \cdot))\right)\mathbf{1}_N. \tag{5.3}$$

Here π_e is the initial state probabilities the multievent model (similar in meaning as the vector containing Ψ in (5.2)), B_e^0 is a matrix containing the event probabilities on the first capture (necessary for Cormack-Jolly-Seber mark-recapture models) and B_i is a matrix containing the event probabilities for subsequent captures. The notation $D(B(o_e, \cdot))$ is used to denote taking a row vector of matrix B corresponding to event o and placing it on a diagonal of a matrix with dimensions the same as B. Note the only difference is that we are using columns from the event matrix rather than rows, and we have multiple observations for a given state, which is similar to the robust design in mark-recapture (Kendall et al., 1997). Finally, Φ_i is the state transition matrix which includes the probability of survival between occasions. We will consider state transitions when we formally introduce the multi-state occupancy model in the next chapter.

While matrix formula in (5.2) may at first glance appear somewhat cumbersome, there is utility in this approach. Practitioners can focus on defining only the state and event matrices. The marginal probability is then defined as a byproduct of the state probability vector and event probability matrix. For example, if to make the assumption that false positives cannot occur (i.e. an unoccupied site can only produce a ‘0’) we need only redefine our event probability matrix as

$$\begin{array}{cc} & \text{event} \\ \text{state} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \left[\begin{array}{cc} 1 & 0 \\ 1-p_t & p_t \end{array}\right], \end{array}$$

and make the appropriate substitutions in the marginal probability statement above.

The strength of a multievent like approach is that it simplifies the model formulation for researchers, allows for a wide range of models to be defined, while all the while making assumptions clearer. As was shown, false positives can be handled by simply redefining the event matrix, rather than defining a whole new framework as has been the case in the past (e.g. Royle and Link, 2006). This allows greater ability of practitioners to redefine models to their specific research needs as well being very scalability in terms of software deployment.

A drawback however is that without careful consideration it may be fairly easy to create models that produce parameter combinations that are not estimable, leaving practitioners to assign appropriate restrictions in order for the model to be estimable. That said, there are a variety of techniques available to identify parameter redundancy (e.g. symbolic algebra and analytic-numerical, see: Catchpole et al., 2002; Gimenez et al., 2003).

Given the clear advantages of the multievent framework, we generalize this approach to the problem of multi-state occupancy over multiple seasons in the next section. Models are built by defining three components, an initial state probability vector, along with transition and event matrices. These separate components are then combined to define the marginal likelihood, which can be shown to produced a number of popular occupancy formulations with the correct model parameterization.

Chapter 6

The Multi-Season, Multi-State Occupancy Framework

6.1 Introduction

Several authors have considered single-season multi-state occupancy models. Royle and Link (2005) modeled a single-season of the NAAMP call index data using a mixture multinomial model which can be viewed as a multi-state extension to occupancy modeling. Each site was viewed as possessing an unobservable abundance index, capable of producing a maximum observed call index of the same level (e.g. a site that has an abundance index of '3' can produce observed indices of '3', '2', '1' and '0'). Sites that produced an observed value of '3' were assumed to be a direct observation of the state, while all other observed indices were assumed to potentially come from a mixture of misclassification and true detection. Observations were modeled as a mixture of multinomials, with each multinomial being dependent on the underlying latent state. In addition to relative measures of abundance, Nichols et al. (2007) expanded the definition of occupancy to include two ordinal states, with and without reproduction in order to model populations of the Northern spotted owl. Sites were first model as either being occupied or not, then conditional on being occupied, further divided depending on whether or not successful reproduction had occurred. Similar to the mixture multinomial model, observations of occupied and reproduction were assumed to be situations were the true underlying state had occurred while all other observations (occupied without reproduction and not occupied) were assumed to come from a mixture of correct and missed detection.

While these models represent important progress in developing multi-state occupancy model, these models may be too specific for a general adoption to a wide variety of problems and can only be applied to a single-season of data. Most ecological studies with state variables encompass multiple seasons of data, where the underlying state can be assumed to change between seasons, and the models are still too specific to be applied to a wider range of ecological state data. We present a generalized multi-state occupancy model that can be used on both single season and multi-season data sets. The generalization is based on an over-parameterized model as the starting point. This model is then reformulated through a combination of parameter pooling, and a re-parameterization of pooled parameters in terms of biological relevant processes. This requires great care and consideration on the part of practitioners, but makes model assumptions very clear.

Also important to this framework is considerations of goodness-of-fit (GOF). Often the goal is to build models that describe some element of reality, in order to test ecological hypotheses and predict future observations. Assessing how well these models fits is a cornerstone to such an approach. If the model does not fit the data, testing competing ecological hypotheses becomes a murky practice. While GOF statistics exist for occupancy models, and can potentially be extended for multi-state occupancy, these tend to be omnibus measures and of questionable value because they only test whether or not the entire model fits and does not allow finer considerations. Further more, these omnibus GOF statistics often rely on asymptotic results and distributional assumptions that, are not appropriate for many data sets. As result, computationally intensive methods, such as the bootstrap, are required in order for these tests to be useable. For larger monitoring data sets such as the NAAMP this may be all together infeasible due to computational time requirements. Instead, we present a case specific GOF test for the NAAMP data set, which is less computationally intensive and provides more insight into where lack of fit may exist.

Finally, we also consider topics of optimal design. At the heart of these approaches is the notion that sites are repeatedly visited over time in order to estimate detection rates. What is however unclear, is the optimal number of repeat visits in terms of effort versus precision. Also in designing long term studies there may be questions whether it is better to repeatedly sample the same sites over the entire duration or to attempt to sample new sites each season.

6.2 Sampling Protocol

Before we delve into the multi-state multiple season model formulation, it is important to clearly define the sampling protocol associated with occupancy models. Occupancy models use geographical units, termed "sites," as the observational unit, where sites may be contiguously connected (e.g. Figure 6.1) or separated by distance. The goal generally is to estimate proportion of sites that exhibit a certain trait (e.g. being occupied by the study species) as well as transition rates between traits for multi-season models (e.g. for presence/absence occupancy models, transitions would simply be colonization and extinction rates). Here a season is simply an arbitrarily defined unit of time, where we may expect the state to remain closured to changes. The state/trait may however change between seasons.

For multi-state occupancy, there are K possible occupancy states, resulting in a total set of $K + 1$ possible ecological states, denoted as $\mathcal{Z} = \{0, z_1, z_2, \ldots z_K\}$. Note that we explicitly include '0' in the set of possible occupancy states, because an empty site is always possible. For multiple seasons, we consider a scenario where individual sites independently change between a set $\mathcal{Z}$ states over S seasons, where transitions occur only between seasons and sites are closed to changes within a season.

Sampling protocol for multiple season studies occurs on a primary and secondary level, where randomly selected sites are repeatedly sampled T_s times during state closure of a given season (Figure 6.1). However, on any given survey the state of the site is assumed to be unobservable (or latent) and instead one of $E+1$ possible events (set $\mathcal{U} = \{0, u_1, u_2, \ldots, u_E\}$) are observed in its place. Again '0' is explicitly included in the set of events is interpreted as "not observed." While the number of possible events may differ from the possible latent states, events are assumed to be conditionally dependent on the underlying state. The relationship between events and states will be more clearly defined when we introduce the event probabilities (π) shortly.

Sites also do not have to be surveyed for the entire duration of the experiment resulting in two general types of experimental protocol set ups: temporary and permanent monitoring sites. With the temporary sampling protocol practitioners would randomly select a new sample of R_s sites each season, while with permanent monitoring the same set of R sites selected at the start of the experiment are followed through time (see Figure 6.1). Sampling protocols in which sites are only sampled for a subset of seasons can be handled by the permanent monitoring framework if certain assumptions hold (expanded upon later).

Temporary monitoring protocol will allow practitioners to estimate state proportions but not transition rates and may be useful situations where the sampling frame may change between seasons (e.g. ephemeral sites such as ponds). For comparison purposes, we compare the power of both protocols to detect changes in state proportions.

A summary of the survey data collected for a site summarized as row vector of events $\mathbf{h}_i$, termed the encounter history. For example the encounter history $\mathbf{h}_i = 021\ 000\ 12\ 11$ would be a four season design with three surveys per season during the first two seasons and two surveys per season for the remaining seasons. During season two, all surveys resulted in no observations, resulting in three 0's. Missing data can be specified with a '-', for example $\mathbf{h}_i$ = 021 `---` 12 11 has a similar meaning as before except no surveys were performed in the second season. Individual events are denoted with y, or more specifically $y_{i,st}$ for the observed event observed during survey t of season s on site i.

It should be noted that while primary sampling occasions are termed "seasonal," the definition is strictly up to the investigator. Multiple primary sampling occasions could be defined as occurring within a year, or across multiple years, the context depends on study objectives and the type of occupancy states investigated. The only major assumption is that there is population closure during the secondary sampling occasions (i.e. the state does not change). Between seasons however the population is open to the usual demographic processes (birth, death, immigration, emigration).

6.3 Model Formulation

We consider a scenario where individual sites move independently through a set of $K+1$ finite population states $\mathcal{Z} = \{0, z_1, z_2, \ldots z_K\}$, and through a finite number of seasons S. States may be be both nominal, ordinal and/or a combination therein. For example, a site-occupancy model may be implemented by using the nominal categories of 0="none" and z_1="present," while an ordinal approach could be implemented by specifying the states as 0 ="none," z_1="low," z_2="medium" and z_3="high." Combination of both ordinal and nominal approaches are also possible, for example "breeding and at low abundance." The implication in using nominal as ordinal states comes in how the event and transition distributions is parameterized. Surveys (secondary sampling occasions) occur during state closure and produce an event from $\mathcal{U}$ possible events, which is assumed to be conditionally dependent on the underlying state.

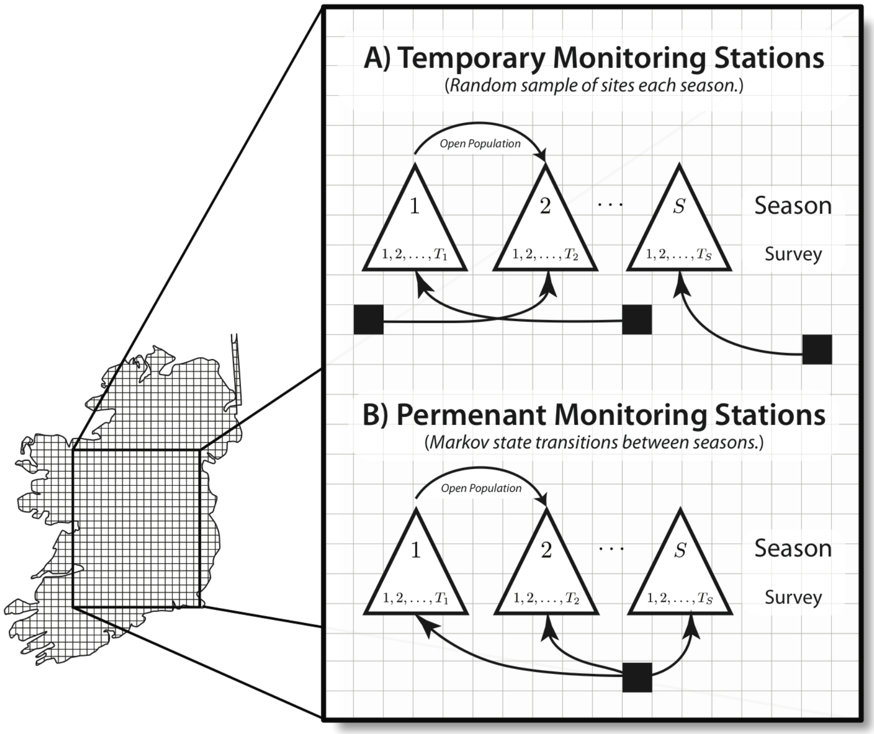


Figure 6.1: Sampling scheme for temporary and permanent monitoring protocols. Each season a new set of sites are chosen and repeatedly survey for temporary monitoring sites, while for permanent monitoring sites sites are repeatedly sampled over multiple seasons. Sites in this example are contiguously arranged, but this is not required.

6.3.1 Event Probabilities

In formulating the event probabilities, each state may produce any of the possible events from $\mathcal{U}$, but that the individual event probabilities may differ by state. For example, events that not expected to occur can be modeled as occurring with probability zero. This probabilities may be modeled in an $K+1$ by $E+1$ matrix, where each row represents a state, and each column an event. Furthermore, these probabilities may be expected to vary by site and survey occasion.

As such, let $\pi_{i,st}^{e|k}$ be the probability of observing event u_e (for $e = 0, 1, \ldots, E$) on survey t of season s, we assume the probability also dependent on the underlying state z_k. As such the detection process can be described as a set of conditional probabilities, which we deliberately start with the most general case. Let $\pi^{e|k}_{i,st}$ be the conditional probability of observing event u_e at site i in season s on occasion t, given that site i is of state z_k, that is $P(Y_{i,st} = u_e|Z_{i,s} = z_k) = \pi^{e|k}_{i,st}$. We can represent the event probabilities for site i on occasion t of season s as a $K + 1 \times E + 1$ matrix,

$$\boldsymbol{\pi}_{i,st} = \begin{array}{c|c} \begin{array}{cccc} \multicolumn{4}{c}{\text{Events}} \\ u_0 & u_1 & \cdots & u_E \\ \hline \end{array} & \text{States} \\ \begin{bmatrix} \pi^{0|0}_{i,st} & \pi^{1|0}_{i,st} & \cdots & \pi^{E|0}_{i,st} \\ \pi^{0|1}_{i,st} & \pi^{1|1}_{i,st} & \cdots & \pi^{E|1}_{i,st} \\ \vdots & \vdots & \ddots & \vdots \\ \pi^{0|K}_{i,st} & \pi^{1|K}_{i,st} & \cdots & \pi^{E|K}_{i,st} \end{bmatrix} & \begin{array}{c} z_0 \\ z_1 \\ \vdots \\ z_K \end{array} \end{array}$$

where each row represents a state and each column an event. Because events are treated as being mutually exclusive, $\boldsymbol{\pi}_{i,st}$ will be row stochastic (i.e. $\pi^{0|k}_{i,st} = 1 - \sum_{e=1}^{E} \pi^{e|k}_{i,st}$ for all i, s, t and k). When the number of events equals the number of states, the diagonal represents correct classification and the off diagonal represents the various forms of misclassification. We assume that the event distribution will be a multinomial $[y_i|Z_i] \sim \text{multinomial}(\{\pi^{e|k}_i\})$, where $\{\pi^{e|k}_i\}$ are the multinomial cell probabilities corresponding to the appropriate row in $\boldsymbol{\pi}_{i,st}$ corresponding to Z.

For most practical applications the site, survey and season specific event probabilities $\boldsymbol{\pi}_{i,st}$ will produce an over-parameterized model that contains more parameters that can be estimated. This is intentional as the $\pi^{e|k}_{i,st}$'s are intended to be a flexible place holder which will be redefined, through a combination of parameter pooling, re-parameterization and covariate models. By deliberately starting with the most general case practitioners are forced to consider how the event probabilities may change (or remain the same) across sites, seasons and occasions. For example, event probabilities may be assumed to be relatively constant across sites, suggestion a practitioner may want to pool across sites (e.g. $\pi^{e|k}_{i,st} = \pi^{e|k}_{st}$ for all i). Similarly a practitioner may want to pool across sites and surveys within a season (e.g. $\pi^{e|k}_{i,st} = \pi^{e|k}_{s}$ for all i and t).

In other situations event probabilities can be expressed as a function of site-specific covariates. Site-specific models could be created on either on a per survey basis (i.e. the full $\pi^{e|k}_{i,st}$ parameterization would be used) or even on a per season basis by first pooling over surveys then modeling (i.e. modeling $\pi^{e|k}_{i,s}$ instead of $\pi^{e|k}_{i,st}$). However, covariate models will be more complex to implement than for simple site-occupancy models given the row constraint in $\boldsymbol{\pi}_{i,st}$.

Initial over-parameterization may be solved by a re-parameterization of the event probability matrix. Often event probabilities are too general to possess biologically relevant meaning. By introducing parameters and parameter combinations with biological meaning, the event matrix can be re-parameterized in terms of biologically relevant processes. Once re-parameterized these new parameters can then be modeled as functions of site-specific covariates tools such as logistic or multinomial regression models, depending on the form of the re-parameterization. We provide examples of both approaches in the example section.

6.3.2 Initial State Distribution and Transition Probabilities

Permanent Monitoring Protocol

For the permanent site monitoring protocol, we need to consider the initial latent state distribution and state transition probabilities. Here we are assuming that sites may transition between states only between seasons and process obeys a Markov chain (that is future occupancy states only depends current occupancy states and not past occupancy states).

For the initial state distribution, if $Z_{i,s}$ is the latent state of site i at season s, then let

$$\boldsymbol{\Phi}_{i,0} = \left[\begin{array}{cccc} \Pr(Z_{i,0} = z_0) & \Pr(Z_{i,0} = z_1) & \cdots & \Pr(Z_{i,0} = z_K) \end{array} \right] \tag{6.1}$$

be the initial state probability vector and $Z_{i,1} \sim \text{Multinomial}(1, \boldsymbol{\Psi}_{i,0})$. Note that $\boldsymbol{\Phi}_{i,0}$ is a stochastic vector and sums to one. The elements of $\boldsymbol{\Phi}_{i,0}$ have been specified in a general manner to accommodate a variety of formulations. For example, in relatively simple case, if Ψ_i^k is the probability that site i has the initial state z_k prior to the start of the experiment, then we may choose to specify the initial state vector as

$$\boldsymbol{\Phi}_{i,0} = \left[\begin{array}{cccc} 1 - \sum_{k=1}^{K} \Psi_i^k & \Psi_i^1 & \ldots & \Psi_i^K \end{array} \right].$$

More complex formulations are presented in the example section.

For the permanent monitoring protocol state transitions are assumed to be memory-less (and hence Markovian). Let $\Phi_{i,s}^{jk}$ be probability that site i transitions from abundance state z_j in season s to abundance state z_k in season $s+1$. The transition probabilities for a particular site and season can be expressed as a $K+1 \times K+1$ matrix,

$$\mathbf{\Phi}_{i,s} = \begin{bmatrix} \Phi_{i,s}^{00} & \Phi_{i,s}^{01} & \cdots & \Phi_{i,s}^{0K} \\ \Phi_{i,s}^{10} & \Phi_{i,s}^{11} & \cdots & \Phi_{i,s}^{1K} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{i,s}^{K0} & \Phi_{i,s}^{K1} & \cdots & \Phi_{i,s}^{KK} \end{bmatrix}_{K+1\times K+1} , \tag{6.2}$$

where $\Phi_{i,s}^{j0} = 1 - \sum_{k=1}^{K} \Phi_{i,s}^{jk}$ for all states, seasons and sites.

Temporary Monitoring Protocol

For the temporary monitoring protocol we have made no assumptions about state transitions because each site is only observed during one of the seasons, during state closure. As such, we must estimate the latent state distribution for each season and each site, that is for site i in season s, the latent occupancy state distribution will be:

$$\mathbf{\Phi}_{i,s} = \begin{bmatrix} \Pr(Z_{i,s} = z_0) & \Pr_i(Z_{i,s} = z_1) & \cdots & \Pr(Z_{i,s} = z_K) \end{bmatrix}. \tag{6.3}$$

Similar to the conditional detection probabilities both the latent state distribution and the transition probabilities are over-parameterized.

6.3.3 Likelihood Formulation

Permanent Monitoring Protocol (Markov Transitions)

For permanent monitoring sites the initial latent occupancy state distribution and the transitions between states must be explicitly modeled. Because between seasons changes in states obey a Markov chain, the probability of any multi-season encounter history can then be expressed as

$$P(\mathbf{h}_i) = \mathbf{\Phi}_i^0 \left[\prod_{s=1}^{S-1} \mathbf{P}_{i,s} \mathbf{\Phi}_{i,s} \right] \mathbf{P}_{i,S} \cdot \mathbf{1}. \tag{6.4}$$

In this expression $\mathbf{P}_{i,s}$ is a diagonal matrix with the elements equal the product of within-season event probabilities for each state. The final component $\mathbf{1}$ is a column vector with $K{+}1$ elements, ensures the matrix algebra resolves correctly producing a marginal probability.

The matrix $\mathbf{P}_{i,s}$ can be viewed as containing the within season event probabilities for each possible state. Formally it is defined as

$$\mathbf{P}_{i,s} = \prod_{t}^{T_s} D\left(\boldsymbol{\pi}_{i,st}(o_{i,st})\right), \tag{6.5}$$

where $D\left(\boldsymbol{\pi}_{i,st}(o_{i,st})\right)$ is the column vector of $\boldsymbol{\pi}_{i,st}$, corresponding to event o, placed on the diagonal.

For example consider a system with $K = 3$ states and $E = 3$ events, with three secondary sampling occasions performed in season s. If the within-season encounter history for site i was '`313`' then $\mathbf{P}_{i,s}$ would be structured as

$$\begin{aligned}
\mathbf{P}_{i,s} &= \begin{bmatrix} \boldsymbol{\pi}_{i,s1}^{3|0} & 0 & 0 & 0 \\ 0 & \boldsymbol{\pi}_{i,s1}^{3|1} & 0 & 0 \\ 0 & 0 & \boldsymbol{\pi}_{i,s1}^{3|2} & 0 \\ 0 & 0 & 0 & \boldsymbol{\pi}_{i,s1}^{3|3} \end{bmatrix} \begin{bmatrix} \boldsymbol{\pi}_{i,s2}^{1|0} & 0 & 0 & 0 \\ 0 & \boldsymbol{\pi}_{i,s2}^{1|1} & 0 & 0 \\ 0 & 0 & \boldsymbol{\pi}_{i,s2}^{1|2} & 0 \\ 0 & 0 & 0 & \boldsymbol{\pi}_{i,s2}^{1|3} \end{bmatrix} \begin{bmatrix} \boldsymbol{\pi}_{i,s3}^{3|0} & 0 & 0 & 0 \\ 0 & \boldsymbol{\pi}_{i,s3}^{3|1} & 0 & 0 \\ 0 & 0 & \boldsymbol{\pi}_{i,s3}^{3|2} & 0 \\ 0 & 0 & 0 & \boldsymbol{\pi}_{i,s3}^{3|3} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{\pi}_{i,s1}^{3|0}\boldsymbol{\pi}_{i,s2}^{1|0}\boldsymbol{\pi}_{i,s3}^{3|0} & 0 & 0 & 0 \\ 0 & \boldsymbol{\pi}_{i,s1}^{3|1}\boldsymbol{\pi}_{i,s2}^{1|1}\boldsymbol{\pi}_{i,s3}^{3|1} & 0 & 0 \\ 0 & 0 & \boldsymbol{\pi}_{i,s1}^{3|2}\boldsymbol{\pi}_{i,s2}^{1|2}\boldsymbol{\pi}_{i,s3}^{3|2} & 0 \\ 0 & 0 & 0 & \boldsymbol{\pi}_{i,s1}^{3|3}\boldsymbol{\pi}_{i,s2}^{1|3}\boldsymbol{\pi}_{i,s3}^{3|3} \end{bmatrix}.
\end{aligned}$$

Depending on how the event probabilities are formulated the expression produced by (6.5) may simplify greatly. For example, suppose it is impossible for a site with state z_0, z_1 or z_2 to produce a '3'. In that case there would only be one non-zero element, $\boldsymbol{\pi}_{i,s1}^{3|3}\boldsymbol{\pi}_{i,s2}^{1|3}\boldsymbol{\pi}_{i,s3}^{3|3}$. We explore this further in our example section.

The marginal likelihood of $\mathbf{h}$, the matrix of all encounter histories, can be specified if all sites are assumed to have independent fates, that is

$$L\left(\{\boldsymbol{\Psi}_i\}, \{\boldsymbol{\Phi}_{i,s}\}, \{\boldsymbol{\pi}_{i,s}\}, \boldsymbol{\theta} | \mathbf{h}_1, \ldots, \mathbf{h}_R\right) = \prod_{i=1}^{R} \Pr(\mathbf{h}_i). \tag{6.6}$$

Temporary Monitoring Protocol and the Single-Season Likelihood

In the temporary monitoring protocol a new set of sites is sampled on each primary sampling occasion, and there is no information available for individual state transitions, and we model the data as series of independent single year likelihoods. Such an approach may be useful if the goal to simply understand factors affecting distribution of a states across a landscape, rather than to monitor or model transitions in occupancy states over time.

Because we do not know the true underlying state of the site we can compute the marginal probability of an encounter history for a given site and season as

$$\Pr(\mathbf{h}_{i,s}) = \mathbf{\Phi}_{i,s}\mathbf{P}_{i,s} \cdot \mathbf{1}, \tag{6.7}$$

where $\mathbf{\Phi}_{i,s}$ is the latent state distribution (6.3) and $\mathbf{P}_{i,s}$ is a diagonal matrix defined by (6.5). Again $\mathbf{1}$ is a column vector with $K+1$ elements that allows the matrix algebra to resolve correctly. Assuming the all sites are independent of one another, a single-season likelihood can be formed as a product of the individual encounter histories across sampled sites,

$$L_s\left(\{\Psi_{k,s}\}, \{\boldsymbol{\pi}_s\} | \mathbf{h}_{1,s}, \ldots, \mathbf{h}_{R_s,s}\right) = \prod_{i=1}^{R_s} \Pr(\mathbf{h}_{i,s}), \tag{6.8}$$

where R_s is the number of sites sampled in season s. Finally, multiple seasons are handled by taking the product of the individual seasonal likelihoods,

$$L\left(\{\mathbf{\Psi}_s\}, \{\boldsymbol{\pi}_s\} | \mathbf{h}_1, \ldots, \mathbf{h}_R\right) = \prod_{s=1}^{S} L_s, \tag{6.9}$$

we have now assumed seasons observations from different are also independent of one another.

6.3.4 Dealing with Missing Data

Missing observations can be handled fairly easily under either multi-season formulation if the missingness of the data is not related to either the underlying state or the parameter values. Then, current formulations can be used by making a simple substitution in (6.5). For any missing observation we replace $D\left(\boldsymbol{\pi}_{i,st}(o_{i,st})\right)$ with $D\left(\mathbf{1}\right)$, where $\mathbf{1}$ is a vector of 1's of length $K+1$.

For example, returning to our previous example of a system with $E = K = 3$ and a history of '313' in season s, suppose the second history is missing (i.e. $\mathbf{h}_{i,s} =$ 3-3) then the four entries of $\mathbf{P}_{i,s}$ would be

$$\mathbf{P}_{i,s} = \begin{bmatrix} \boldsymbol{\pi}_{i,s1}^{3|0} & 0 & 0 & 0 \\ 0 & \boldsymbol{\pi}_{i,s1}^{3|1} & 0 & 0 \\ 0 & 0 & \boldsymbol{\pi}_{i,s1}^{3|2} & 0 \\ 0 & 0 & 0 & \boldsymbol{\pi}_{i,s1}^{3|3} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{\pi}_{i,s3}^{3|0} & 0 & 0 & 0 \\ 0 & \boldsymbol{\pi}_{i,s3}^{3|1} & 0 & 0 \\ 0 & 0 & \boldsymbol{\pi}_{i,s3}^{3|2} & 0 \\ 0 & 0 & 0 & \boldsymbol{\pi}_{i,s3}^{3|3} \end{bmatrix}.$$

This approach can be extended to handle complete season's worth of observations that are missing on a site (for example data loss), or if sites are not sampled every season. The end result is that the missing value is essentially ignored when computing the likelihood value. The only caveat with the latter approach (missing primary or secondary surveys) is that missingness is assumed not to relate to the underlying occupancy state. For example, of a site is not sampled in the second season due to the fact occupancy was not confirmed in the first season one, then the missing data cannot be handled in this manner as sites with state z_0 in the first season are more likely not to be surveyed in season two than other sites. A special likelihood would have to be formulated in this circumstance.

6.4 Model Parameterization Philosophy

Thus far we have deliberately presented an over-parameterized model that allows for separate event probabilities for each and every event and underlying state combination, across all sites, surveys and seasons. Such a model will not be estimable even with the richest of data sets. Instead, the parameterization is intended to provide place holders that may be restricted or re-paramaterized to suit a particular situation. Before estimation is attempted, we envision a three step process to redefine the general model to one more specific for the situation at hand:

1. parameter pooling;
2. model re-parameterization and constraints; and
3. the inclusion of linear functions of covariates.

Starting with an overly general model forces practitioners to carefully consider their model before proceeding with data fitting. In the first step, parameters are pooled together to reflect only the key features of heterogeneity in the system. This has the additional benefit of providing an easy framework to define and test important features of the system.

For example, if the conditional event probabilities are believed to be roughly similar across sites, then a practitioner may want to drop the i subscript from $\pi_{i,st}^{e|k}$ thereby using the same values of $\{\pi\}$ for all sites but different values depending on the season, the survey and the underlying state. We would state this restriction as $\pi_{i,st}^{e|k} = \pi_{st}^{e|k}$. A further restriction of $\{\pi\}$ could be to constrain the event probabilities across surveys within a season (e.g. situations where secondary samples are close together in time), but different values between seasons. We could state the restriction as $\pi_{i,st}^{e|k} = \pi_{s}^{e|k}$. Assuming the models are estimable we could then examine which these two purposed parameter pooling better describes the data by either likelihood ratio tests or information theoretic approaches such as Akaike Information Criterion (AIC) ranking or evidence ratios (see Burnham and Anderson, 2002).

Once parameters have been pooled, the next step is to re-parameterize the event and state probabilities in terms of biological processes as well as setting any appropriate parameter constraints. The $\{\pi\}$ and $\{\Phi\}$ parameters should seen as place holders that can be re-expressed in terms of biologically meaningful parameters. As we will see in the example section, observed events are often the result of one or more biological processes occurring on the site. By re-parameterizing π and ϕ in terms of these processes, we can build very specific models for each component separately. Furthermore, the re-parameterization allows each biological process to be modeled further using linear functions of covariates, the final step in the process.

We outline all of these procedures in our examples section.

6.4.1 Goodness-of-fit

By building models that emphasize plausible ecological mechanisms first and foremost, the topic of goodness-of-fit presents an interesting challenge. Ideally, the models in this context represent scientific hypotheses (Royle and Young, 2008) and favouring approaches such as multiple hypothesis testing to assess support for various ecological hypotheses (Burnham and Anderson, 2002). While arguably a better approach to understanding complex systems than simple hypothesis testing (although see Stephens et al., 2005), assessing model fit becomes a somewhat overlooked component to building ecologically relevant models. If none of the models in the model set adequately fit the data, then their appropriateness in critiquing underlying scientific hypotheses is unclear. If models are to carry the meaning of scientific hypotheses, the issues of fit need to be properly considered.

To date only omnibus goodness-of-fit tests are available for occupancy models (e.g. MacKenzie and Bailey, 2004). Based on a Pearson chi-squared statistic, the proposed technique compares the observed and expected frequencies of the unique encounter histories and can accomodate site-specific covariate models. Typically, covariate models can create problems due to the potential of continuous covariate values to create a large number of outcome classes. However, by basing the goodness-of-fit classes on unique encounter histories (rather than a combination of covariate values and encounter histories), we limit the number of possible classes and we can calculate the expected frequency of each unique history by pooling over sites all sites in the experiment. Because the test statistic will no longer follow a chi-square distribution a parametric bootstrap procedure is now required to determine whether an observed observed encounter history frequency is too large.

Adapting this approach to the proposed multi-state occupancy model is fairly straight forward. The Pearson chi-squared statistic can be computed as

$$\chi^2 = \sum_{h=1}^{H} \frac{(O_h - E_h)^2}{E_h}$$

where H is the number of unique encounter histories and O_h and E_h are the observed and expected frequencies of encounter history h. The expected number of sites exhibiting encounter history h can be determined by pooling over all sites, that is

$$E_h = \sum_{i=1}^{R} P(\mathbf{X}_i = h)$$

where R is the number of sites sampled and $P(\mathbf{X}_i = h)$ is determined by (6.4) for permanent monitoring designs and by (6.7) for the temporary monitoring protocol. With the temporary monitoring protocol it would likely be advantageous to consider each season independently and then combine the χ^2 values from each season.

To assess the adequacy of the fitted model, we determine the probability of the observed χ^2 statistic using a parametric bootstrap (Buckland and Garthwaite, 1991). Following from MacKenzie and Bailey (2004), the parametric bootstrap proceeds as follows:

1. Fit the model to observed data and derive parameter estimates for the event probabilities $\widehat{\pi_{i,st}^{e|k}}$, the initial state probabilities $\widehat{\mathbf{\Phi}}_{i,0}$ and the transition probabilities $\widehat{\mathbf{\Phi}}_{i,s}$.

2. Using the parameter estimates to compute the test statistic χ^2_{obs}.

3. Then for each site, denoted as i, do the following:

   (a) Assign an initial state from $\mathcal{Z}$ to site i using a multinomial draw with cell probabilities $\widehat{\mathbf{\Phi}}_{i,0}$. We will denote this state as $z_{i,1}$.

   (b) Next for $s = 1 \ldots S$ each seasons do the following

      i. Make repeated multinomial draws from $\mathcal{U}$, using the row probabilities from $\widehat{\pi^{e|k}_{i,st}}$ associated with state $z_{i,s}$ and assign T_s events to site i in season s. Record these values as the observations for season s and ensure it is of the same structure as the observed data (e.g. same number of surveys, missing values, etc).

      ii. Next, if $s < S$, assign a new state to i for season $s+1$, by making a multinomial draw using the row probabilities from $\widehat{\mathbf{\Phi}}_{i,s}$ associated with being in state k in season s. Denote this new state as $z_{i,s+1}$.

   The observed events from Steps 3a-3b will make up the encounter history of site i.

4. Fit the same structural model as in Step 1, except using the data from Step 3. Calculate χ^2_{b} and store the results.

5. Repeat Steps 3-4 a large number of times, B, to approximate the distribution of the sampling statistic, assuming the fitted model is correct.

6. Compare χ^2_{obs} to the approximate distribution (χ^2_{B}) and determine the probability of observing χ^2_{Obs} or a more extreme value. This will be the p-value for the goodness-of-fit test.

As suggested by MacKenzie and Bailey (2004) an estimate of over-dispersion $\hat{c}$ may be obtained by following formula suggested by White et al. (2002),

$$\hat{c} = \frac{\chi^2_{\text{obs}}}{1/B \sum_b^B \chi^2_{\text{b}}}.$$

The over-dispersion estimate $\hat{c}$ can then be used adjust our model selection procedures and standard errors (Burnham and Anderson, 2002). If the model adequately fits then $\hat{c}$ should be close to one. Values of $\hat{c}$ greater than one suggest there is more variability in the data than expected by the model, while values less than one suggest less variability.

While this approach does allow for a goodness-of-fit, shortcomings are present. Calculating GOF in this manner can be problematic for large data sets because it is computationally infeasible (we explore this problem further in our example section). Furthermore omnibus tests provide little information about what portion of the model does not adequately fit; and only indicates that the whole model does not adequately describe the data - we will not broach the topic of what defines adequacy.

We agree with other authors (e.g. Gelman et al., 2004) that graphical methods for assessing fit should also be considered. As such we have also developed a formulation-specific GOF procedure that is less computationally intensive than the bootstrap procedure outlined above. It is outlined in our numerical examples section.

Chapter 7

Special Cases

7.1 Site-occupancy Modeling, Single Species ($K = 2$)

MacKenzie et al. (2002) introduced site-occupancy modeling as a way to model species occurrences across a landscape, where there are two underlying states z_0=“absent” and z_1=“present” and two observed events are u_0=“not observed,” and u_1=“observed as occupied”. This represents a special case where $K = 1$ and $E = 1$.

In the original formulation it is a assumed that only an occupied site may produce a sighting, as such we can expect the following events to be associated with the underlying states,

$$\begin{aligned} z_0 &\rightarrow u_0 \\ z_1 &\rightarrow u_0 \text{ or } u_1. \end{aligned}$$

Let p_{it} representing the probability of observing the species on site i on occasion t, resulting in the following event probability matrix:

$$\boldsymbol{\pi}_{i,st} = \boldsymbol{\pi}_{i,t} = \begin{array}{c|c} \multicolumn{1}{c}{\text{Events}} & \\ \begin{array}{cc} u_0 & u_1 \end{array} & \text{States} \\ \hline \left[\begin{array}{cc} 1 & 0 \\ 1 - p^1_{i,t} & p^1_{i,t} \end{array}\right] & \begin{array}{c} z_0 \\ z_1 \end{array} \end{array}.$$

Note that the first row of $\boldsymbol{\pi}_{i,st}$ contains a zero, this restriction is imposed because it is assumed only the event u_0 may be observed when the site is unoccupied z_0.

Because this is a single-season model we only need to define the initial latent state distribution. Let Ψ_i representing the probability that site i was occupied, which gives the following initial state distribution

$$\mathbf{\Phi}_{i,0} = \begin{bmatrix} 1-\Psi_i & \Psi_i \end{bmatrix}.$$

To estimate parameters the temporary monitoring protocol likelihood (6.9) can be used with $S = 1$. The p_{it} and Ψ_i parameters can further modeled with covariates.

For a multiple seasons occupancy model we need to adjust the detection probabilities to include seasonality and include transition probabilities. Let $p_{i,st}$ be the probability of observing the species if the site is occupied, for site i in season s on survey t. The event probabilities will then be defined as

$$\boldsymbol{\pi}_{i,st} = \begin{bmatrix} 1 & 0 \\ 1-p_{i,st} & p_{i,st} \end{bmatrix}.$$

Following the multiple season site-occupancy formulation of MacKenzie et al. (2006), γ_s and ϵ_s are used to represent the probability of colonization (a site transitioning from unoccupied to occupied) and extinction (the reverse) in season s. Here we will assume all sites have the same probability of extinction or colonization. The state transition probabilities for season s will be defined as

$$\mathbf{\Phi}_{i,s} = \mathbf{\Phi}_s = \begin{bmatrix} 1-\gamma_s & \gamma_s \\ \epsilon_s & 1-\epsilon_s \end{bmatrix}.$$

The probability of an encounter history is then computed with (6.4) with the likelihood for all the observed histories specified by (6.6). If the population of sites were located in a region that experiences roughly the same environmental conditions, it could be possible model γ_s or ϵ_s in terms of regional factors to gain a better understanding of the extinction and colonization processes.

7.2 Site-occupancy with Multiple Species, Including Biological Interactions

In addition to abundance and distributional information about a single species, ecologists and managers often wish to understand the how a variety of species interact with each other. A variety of biological interactions exist from neutralism, where species interact but do not affect each other, to more severe forms such as predation and parasitism (one gains while the other loses) or commensalism where the interaction benefits both. Because all types of interactions, beside neutralism, has the potential to affect both the distribution of both species, multi-state occupancy can be used to gain insight into these processes.In the simplest scenario we can use multi-state occupancy to model the presence or absence of two species (denoted as A and B), producing four possible states and events (Table 7.1).

Table 7.1: The underlying states and events for a two species co-occurrence model.

State			**Event**		
z_0	-	Neither species is present	u_0	-	Neither species is observed
z_A	-	Only species A is present	u_A	-	Only species A is observed
z_B	-	Only species B is present	u_B	-	Only species B is observed
z_{AB}	-	Both species are present	u_{AB}	-	Both species are observed.

If we are willing to assume that false positives are not possible, that is trained observers may miss a species but will not falsely observe species, then each state can only produce a subset of events:

$$\begin{aligned} z_0 &\rightarrow u_0 \\ z_A &\rightarrow u_0 \text{ or } u_A \\ z_B &\rightarrow u_0 \text{ or } u_B \\ z_{AB} &\rightarrow u_0, u_A, u_B \text{ or } u_{AB}. \end{aligned}$$

7.2.1 Event Probabilities

This application differs somewhat from previous multi-state examples as the underlying state of sites is the result of two separate biological processes, the presence and/or absence of the individual species. As such, not all events can be produced by all states. For example, state z_A can only produce the events u_0 and u_A, since species B is not available, while the state z_{AB} can produce all possible events. This will produce the following general pattern for the event probabilities:

$$
\begin{array}{ccc}
 & \text{Events} & \\
 & \begin{array}{cccc} u_0 & u_A & u_B & u_{AB} \end{array} & \text{States} \\
\boldsymbol{\pi}_{i,st} = & \left[\begin{array}{cccc}
1 & 0 & 0 & 0 \\
\pi_{i,st}^{0|A} & \pi_{i,st}^{A|A} & 0 & 0 \\
\pi_{i,st}^{0|B} & 0 & \pi_{i,st}^{B|B} & 0 \\
\pi_{i,st}^{0|AB} & \pi_{i,st}^{A|AB} & \pi_{i,st}^{B|AB} & \pi_{i,st}^{AB|AB}
\end{array}\right] & \begin{array}{c} z_0 \\ z_A \\ z_B \\ z_{AB} \end{array}
\end{array} .
$$

Next we may re-parameterized the $\{\pi\}$ parameters in terms of appropriate biological processes. MacKenzie et al. (2006) proposed a single-season multiple species formulation, where the conditional event probabilities (termed "detection probabilities") were modeled separately depending on whether a single both species occupied the site. For example, if vocalization is used to detect the presence of species A and the occurrence of species B causes species A to vocalize less (and therefore be more elusive to observers) then we may expect $\pi_{i,st}^{A|A} > \pi_{i,st}^{A|AB}$ and $\pi_{i,st}^{AB|AB} < \pi_{i,st}^{A|AB} \times \pi_{i,st}^{B|AB}$. The proposed parameterization (assuming all sites share the same event probabilities and extending for S season) is as follows:

p_{st}^{A} - the probability of detecting A, given it is present and species B is absent in survey occasion t of season s,

p_{st}^{B} - the probability of detecting B, given it is present and species A is absent in survey occasion t of season s,

r_{st}^{AB} - the probability of detecting both species simultaneously, given that both are present in survey occasion t of season s,

r_{st}^{Ab} - the probability of detecting only detecting species A, given that both are present in survey occasion t of season s,

r_{st}^{aB} - the probability of detecting only detecting species B, given that both are present in survey occasion t of season s.

The conditional event probability matrix then take the form

$$\boldsymbol{\pi}_{i,st} = \boldsymbol{\pi}_{st} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1-p^A_{st} & p^A_{st} & 0 & 0 \\ 1-p^B_{st} & 0 & p^B_{st} & 0 \\ 1-r^{AB}_{st}-r^{Ab}_{st}-r^{aB}_{st} & r^{Ab}_{st} & r^{aB}_{st} & r^{AB}_{st} \end{bmatrix}.$$

By using different event probabilities for each of the possible underlying states, interactions between species detection and species presence may be estimated. However, there may also be little or no effect such effect present, in this case we would expect $\pi^{A|A}_{i,st} = \pi^{A|AB}_{i,st}$ and $\pi^{AB|AB}_{i,st} = \pi^{A|AB}_{i,st} \times \pi^{B|AB}_{i,st}$. As such, we may want to reformulate the conditional event probabilities as

$$\boldsymbol{\pi}_{i,st} = \boldsymbol{\pi}^*_{st} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1-p^A_{st} & p^A_{st} & 0 & 0 \\ 1-p^B_{st} & 0 & p^B_{st} & 0 \\ (1-p^A_{st})(1-p^B_{st}) & p^A_{st}(1-p^B_{st}) & (1-p^A_{st})p^B_{st} & p^A_{st}p^B_{st} \end{bmatrix}.$$

Note that use the * notation to indicate an alternate, species independent form of $\boldsymbol{\pi}_{st}$. We may even consider going a step further and use the same conditional event probabilities for all surveys in a season,

$$\boldsymbol{\pi}_{i,st} = \boldsymbol{\pi}^*_{s} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1-p^A_{s} & p^A_{s} & 0 & 0 \\ 1-p^B_{s} & 0 & p^B_{s} & 0 \\ (1-p^A_{s})(1-p^B_{s}) & p^A_{s}(1-p^B_{s}) & (1-p^A_{s})p^B_{s} & p^A_{s}p^B_{s} \end{bmatrix}.$$

Investigators may then select between the potential conditional event models (e.g. $\boldsymbol{\pi}_{st}$, $\boldsymbol{\pi}^*_{st}$ and $\boldsymbol{\pi}^*_{s}$) by using standard tools such as likelihood ratio tests and AIC rankings.

7.2.2 Initial States

States are defined as in terms of the presence or absence of species A and B (Table 7.1). Let Ψ^A be the probability that species A occurs on a site, regardless of species B, and Ψ^B the probability that species B occurs regardless of species A, which can be interpreted as the marginal probabilities of observing each of the respective species. If the occurrences of species A and B occurs independently, then the probability of co-occurrence will be the product of the marginal probabilities $\Psi^A \times \Psi^B$. Species however often do not occur independently, instead they may interact with each other through a variety of mechanisms (e.g. competition, mutualism; see Table 7.2). As a result, we can expect we must estimate the probability of co-occurrence separately. Let Ψ^{AB} be the probability that species A and B co-occur, which may differ from $\Psi^A \times \Psi^B$ depending on the type of interaction (Table 7.2).

Because Ψ^A represents the marginal we can compute the probability that site i only contains species A at the start of the experiment as $P(z_A) = \Psi^A - \Psi^{AB}$ (the marginal minus the intersection). This same arguments can be used for species B resulting in the following latent state probabilities,

$$\begin{aligned} P(z_0) &= 1 - \Psi^A - \Psi^B + \Psi^{AB} \\ P(z_A) &= \Psi^A - \Psi^{AB} \\ P(z_B) &= \Psi^B - \Psi^{AB} \\ P(z_{AB}) &= \Psi^{AB} \end{aligned}$$

and the following initial state vector,

$$\mathbf{\Phi}_{i,0} = \mathbf{\Phi}_0 = \left[\begin{array}{cccc} \Pr(z_0) & \Pr(z_A) & \Pr(z_B) & \Pr(z_{AB}) \end{array} \right]$$

which has been generalized to all sites. In this case we are willing to make the assumption that all sites posses the same probability of occurrence so that we may estimate the proportion of sites across the landscape that are occupied by either or both species.

7.2.3 State Transitions

There has been little work exploring, a multi-season formulation for the multi-species site-occupancy mode. However, it is fairly straightforward to specify transition probabilities for the Markov model assuming either independent or dependent colonization and extinction process. Similar to the initial state and event probabilities, we need to consider the possibility that species A may have an association with species B which may be antagonistic, neutral, commensal or mutualistic relative to species B.

Table 7.2: Potential biological interactions between two species (A and B) and the potential effect on the probability of co-occurrence. Note that '0' represents no effect, '+' a positive effect and '−' a negative effect.

Type	**Effect on A**	**Effect on B**	**Expected**Ψ^{AB}
Neutralism	0	0	$\Psi^{AB} = \Psi^A \times \Psi^B$
Amensalism	$-$	0	$\Psi^{AB} < \Psi^A \times \Psi^B$
Commensalism	$+$	0	$\Psi^{AB} > \Psi^A \times \Psi^B$
Competition	$-$	$-$	$\Psi^{AB} < \Psi^A \times \Psi^B$
Mutualism	$+$	$+$	$\Psi^{AB} > \Psi^A \times \Psi^B$
Predation or Parasitism	$+$	$-$	$\Psi^{AB} \leq \Psi^A \times \Psi^B$

In terms of site-occupancy modeling this potential for competition can be expected to affect extinction and colonization probabilities, depending on the nature of the competition. Assuming a Markovian process, the probability of a species colonizing or going extinct in season $s + 1$, will depend on whether one of the competing species occupies the site in season s. If colonization and extinction processes can be viewed as occurring instantaneously between seasons, the following parameterization can be used

γ_s^A - the probability that a site is colonized by species A in season $s+1$, given neither species are present in season s

γ_s^B - the probability that a site is colonized by species B in season $s+1$, given neither species are present in season s

γ_s^{Ab} - the probability that a site is colonized by species A in season $s+1$, given species B is present in season s.

γ_s^{aB} - the probability that a site is colonized by species B in season $s+1$, given species A is present in season s.

ϵ_s^A - the probability that a site goes extinct for species A in season $s + 1$, given that species B is not present in season s

ϵ_s^B - the probability that a site goes extinct for species B in season $s + 1$, given that species A is not present in season s

ϵ_s^{Ab} - the probability that a site goes extinct for A in season $s+1$, when when species B is present in season s.

ϵ_s^{aB} - the probability that a site goes extinct for B in season $s+1$, when when species A is present in season s.

Because the colonization and extinction process is viewed as instantaneous all transitions from the state z_0 will be identical to the independent formulation (first row of $\mathbf{\Phi}_s^*$). All other transitions depend on the presence or absence of the competing species. The resulting Markov transition probability matrix will be

$$\mathbf{\Phi}_{i,s} = \mathbf{\Phi}_s = \begin{bmatrix} (1-\gamma_s^A)(1-\gamma_s^B) & \gamma_s^A(1-\gamma_s^B) & (1-\gamma_s^A)\gamma_s^B & \gamma_s^A\gamma_s^B \\ \epsilon_s^A(1-\gamma_s^{aB}) & (1-\epsilon_s^A)(1-\gamma_s^{aB}) & \epsilon_s^A\gamma_s^{aB} & (1-\epsilon_s^A)\gamma_s^{aB} \\ (1-\gamma_s^{Ab})\epsilon_s^B & \gamma_s^{Ab}\epsilon_s^B & (1-\gamma_s^{Ab})(1-\epsilon_s^B) & \gamma_s^{Ab}(1-\epsilon_s^B) \\ \epsilon_s^{Ab}\epsilon_s^{aB} & (1-\epsilon_s^{Ab})\epsilon_s^{aB} & \epsilon_s^{Ab}(1-\epsilon_s^{aB}) & (1-\epsilon_s^{Ab})(1-\epsilon_s^{aB}) \end{bmatrix}.$$

For a neutral relationship, we can expect species A and species B to experience colonization and extinction rates independently of one another, which can be modeled using the following parameters,

γ_s^A - the probability that a site is colonized by species A in season $s+1$, given it is not currently occupied by species A in season s, regardless of species B.
γ_s^B - the probability that a site is colonized by species B in season $s+1$, given it is not currently occupied by species A in season s, regardless of species B.
ϵ_s^A - the probability that a site goes extinct for species A in season $s+1$, given species A was present in season s, regardless of species B.
ϵ_s^B - the probability that a site goes extinct for species B in season $s+1$, given species B was present in season s, regardless of species A.

Any transition between states will simply be the product of the appropriate processes. For example, the transition $z_0 \to z_{AB}$ will occur with probability $\gamma_s^A\gamma_s^B$, because both species must colonize simultaneously. For transition $z_0 \to z_B$ to occur between season s and $s+1$, species A must not colonize, but species B does, resulting in a transition probability of $(1-\gamma_s^A)\gamma_s^B$. The transition matrix may be specified as

$$\mathbf{\Phi}_s^* = \begin{bmatrix} (1-\gamma_s^A)(1-\gamma_s^B) & \gamma_s^A(1-\gamma_s^B) & (1-\gamma_s^A)\gamma_s^B & \gamma_s^A\gamma_s^B \\ \epsilon_s^A(1-\gamma_s^B) & (1-\epsilon_s^A)(1-\gamma_s^B) & \epsilon_s^A\gamma_s^B & (1-\epsilon_s^A)\gamma_s^B \\ (1-\gamma_s^A)\epsilon_s^B & \gamma_s^A\epsilon_s^B & (1-\gamma_s^A)(1-\epsilon_s^B) & \gamma_s^A(1-\epsilon_s^B) \\ \epsilon_s^A\epsilon_s^B & (1-\epsilon_s^A)\epsilon_s^B & \epsilon_s^A(1-\epsilon_s^B) & (1-\epsilon_s^A)(1-\epsilon_s^B) \end{bmatrix}.$$

We use the notation $\mathbf{\Phi}_s^*$ as independent colonization and extinction processes represent a special case that is not likely to occur in multi-species studies.

Similar to the conditional event probabilities, tests for species interaction in the colonization and extinction processes may be assessed by comparing the models using either $\mathbf{\Phi}_s^*$ or $\mathbf{\Phi}_s$ formulations by employing standard tools such as a likelihood ratio tests or AIC ranking.

Finally, it should be noted that by clearly separating the observational process from the underlying state transition process, allows for great flexibility in hypothesis testing. For example, it is possible to model independent species detection probabilities, but dependent colonization and extinction processes.

7.3 Site-occupancy with Successful Reproduction

The Northern Spotted Owl (NSO) has been steadily declining for the last four decades, due to a combination of habitat loss and competition with barred owl. Despite decades of mark-recapture studies, there have been difficulties relating individual measurements to survivorship and to overall population declines (Anthony et al., 2006).

Nichols et al. (2007) took a different approach to the problem by providing a specific multi-state extension to the single-season site-occupancy model that accounted for successful reproduction. The formulation uses to three possible states along with three possible events (Table 7.3).

Table 7.3: The underlying states and events of the site-occupancy with reproduction model.

State			**Event**		
z_0	-	The site is unoccupied	u_0	-	Species not observed
z_1	-	Occupied with no production of young	u_A	-	Species observed
z_2	-	Occupied with successful reproduction	u_B	-	Species observed with young

Similar to the site-occupancy formulation false positives were considered not to occur. As such we get the following association between states and events,

$$\begin{aligned} z_0 &\rightarrow u_0 \\ z_1 &\rightarrow u_0 \text{ or } u_A \\ z_2 &\rightarrow u_0, u_A \text{ or } u_B. \end{aligned}$$

In modeling the events it was assumed that species detection was the result of two different processes (dependent on whether or not young were present) and as well as a conditional process of determining whether or not reproduction had occurred. As such events were modeled using the following parameters,

$p^1_{i,t}$ - the probability of detecting the species, given it is present and without young,
$p^2_{i,t}$ - the probability of detecting the species, given it is present and with young,
$\delta^1_{i,t}$ - the conditional probability of detecting evidence of reproduction, given the site is occupied with young.

producing the following event probability matrix,

$$\boldsymbol{\pi}_{i,st} = \boldsymbol{\pi}_{i,t} = \begin{array}{c|c} \begin{array}{ccc} & \text{Events} & \\ u_0 & u_A & u_B \end{array} & \text{States} \\ \hline \left[\begin{array}{ccc} 1 & 0 & 0 \\ 1-p^1_{i,t} & p^1_{i,t} & 0 \\ 1-p^2_{i,t} & p^2_{i,t}(1-\delta_{i,t}) & p^2_{it}\delta_{it} \end{array}\right] & \begin{array}{c} z_0 \\ z_1 \\ z_2 \end{array} \end{array} .$$

Notice that the event of observing reproductive evidence (e.g. young) can only occur if the site is occupied with successful reproduction (i.e. state z_2). This assumes that experienced field observers are collecting the data (a justified assumption in this case), and other studies may wish to consider these types misclassifications more thoroughly. Also note we have excluded the s subscript for clarity because this was a single-season model.

Similar to the multispecies model states can be viewed as the result of two separate processes. The parameter Ψ^1_i denotes the probability site i was occupied regardless of reproductive state, while Ψ^2_i represents the probability of successful reproduction, given an occupied site. The initial state distribution for site i will then be

$$\begin{aligned} P(z_0) &= 1 - \Psi^1_i \\ P(z_1) &= \Psi^1_i(1 - \Psi^2_i) \\ P(z_2) &= \Psi^1_i \Psi^2_i, \end{aligned}$$

producing the following initial state vector,

$$\boldsymbol{\Phi}_{i,0} = \left[\begin{array}{ccc} 1 - \Psi^1_i, & \Psi^1_i(1 - \Psi^2_i), & \Psi^1_i \Psi^2_i \end{array}\right].$$

7.3.1 Extending to Multiple Seasons

It is relatively straight forward to extend the formulation to multiple seasons following the multi-season site-occupancy example. Since a primary concern of researchers is to understand the factors affecting successful reproduction, it is important to consider the role site-specific factors may play on site colonization and the ability of sites to support reproduction. To do so will a require site-specific parameterization.

Following the logic of the initial state distribution we can model transitions as the independent process of colonization or extinction and followed by reproduction if the site is colonized. This leads to the following site and season-specific parameters,

$\gamma_{i,s}$ - the probability that site i is colonized in season s

$\epsilon_{i,s}$ - the probability that site i goes extinct in season s

$\alpha_{i,s}$ - the probability that site i is suitable for reproduction in season s, given that a site is occupied,

which in turn will produce the following state transition matrix,

$$\mathbf{\Phi}_{i,s} = \begin{bmatrix} 1-\gamma_{i,s} & \gamma_{i,s}(1-\alpha_{i,s}) & \gamma_{i,s}\alpha_{i,s} \\ \epsilon_{i,s} & (1-\epsilon_{i,s})(1-\alpha_{i,s}) & (1-\epsilon_{i,s})\alpha_{i,s} \\ \epsilon_{i,s} & (1-\epsilon_{i,s})(1-\alpha_{i,s}) & (1-\epsilon_{i,s})\alpha_{i,s} \end{bmatrix}.$$

Logistic regression can then be used to model any of the transition parameters. For example suitability could be modeled as a function of site-specific covariates thought to be important to facilitating reproduction.

7.4 Relative Measures of Abundance (Ordinal Abundance Classes)

Among ecologists and managers there is a growing interest in the use of indirect abundance measurements to track changes in population size over time. Leaving aside debates on the appropriateness of different surrogate measure, we will focus on how to incorporate detection, a topic which there is still much debate (e.g. MacKenzie and Kendall, 2002; Royle and Link, 2005). The North American Amphibian Monitoring Program (NAAMP) provides a very useful example of this type of approach.

Since 2001, the NAAMP uses as standardized ranked categorical calling index to estimate anuran abundance. Currently, the protocol is deployed across 17 states and provinces in the United States and Canada and monitors about 22 anuran species. The ranked categorical call index is based on male anuran vocalization frequencies, both of which have been shown to be correlated with true underlying abundance of a site, including both males and females (Nelson and Graves, 2004).

The ranked call index would then be considered the observed event in this system and takes on four possible levels: no calls (0), discrete non-overlapping calls (u_1 or 1), discrete overlapping calls (u_2 or 2) and full chorus (u_3 or 3; Table 7.4, for a full description see Weir and Mossman, 2005). (Table 7.4). The difficulty comes in relating the events to an underlying state. Higher call indices are generally associated with higher abundance, but the association is clouded by issues of detection; while on average a site may produce a high call index, on any given occasion a lower call index may be observed.

Royle and Link (2005) proposed a model where the underlying states are formulated in terms of the maximum possible call index a site can produce (termed the abundance class; Table 7.4). On any given survey a site will produce a call index the same value as or lower than its abundance class. For example a site that possesses an abundance class of ‘2’ could produce call a call index of ‘0’, ‘1’ or ‘2’. The possible events produced by each state can be summarized as:

$$\begin{aligned} z_0 &\rightarrow u_0 \\ z_1 &\rightarrow u_0 \text{ or } u_1 \\ z_2 &\rightarrow u_0, u_1 \text{ or } u_2 \\ z_3 &\rightarrow u_0, u_1, u_2 \text{ or } u_3. \end{aligned}$$

The abundance class formulation can be viewed as a multi-state occupancy problem. Royle and Link (2005) viewed the problem in terms of mixture of multinomial distributions of differing sizes depending on the underlying state. An abundance class of ‘1’ had a binomial outcome (i.e. a call index of either ‘0’ or ‘1’), while an abundance class of ‘2’ or ‘3’ had multinomial outcomes with three and four outcomes respectively. A marginal likelihood, integrating over states, was then used to generate parameter estimates.

Table 7.4: The underlying states and events for the NAAMP abundance class model.

States			Events		
z_0	-	The site is unoccupied	u_0	-	no calls
z_1	-	An abundance level capable of producing a call index of 1 or lower	u_1	-	discrete non-overlapping calls
z_2	-	An abundance level capable of producing a call index of 2 or lower	u_2	-	discrete overlapping calls
z_3	-	The site possesses an abundance capable of producing a call index of 3 or lower	u_3	-	full chorus

7.4.1 Event Probabilities

Although technically each abundance class can produces number of possible outcomes, we can be generalize all outcomes as coming from the same set of possible categorical outcomes, with unobservable categories having a cell probability of probability zero. As such each of these of the multinomial cell probabilities will make up the rows in the event probability matrix. Using a notation akin to that of Royle and Link (2005), the general event matrix will be formulated as

$$\boldsymbol{\pi}_{i,st} = \boldsymbol{\pi} = \begin{array}{c|c} \begin{array}{cccc} \multicolumn{4}{c}{\text{Events}} \\ u_0 & u_1 & u_2 & u_3 \end{array} & \text{States} \\ \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ \pi^{0|1} & \pi^{1|1} & 0 & 0 \\ \pi^{0|2} & \pi^{1|2} & \pi^{2|2} & 0 \\ \pi^{0|3} & \pi^{1|3} & \pi^{2|3} & \pi^{3|3} \end{array}\right] & \begin{array}{c} z_0 \\ z_1 \\ z_2 \\ z_3 \end{array} \end{array}.$$

where the diagonal probabilities ($\pi^{k|k}$ for $k = 1, 2, 3$) represent a correct classification, while $\pi^{j|k}$ for $j < k$ represents the possible types of misclassification, that is an observed call index lower than the true abundance class. Note that we have excluded the site and time specific subscripts for clarity and the upper diagonal is zero because observing call indices greater than the underlying abundance class is assumed not to occur.

The general parameterization doesn't convey the ordinal nature of the data, for example the probability of observing a call index matching the true abundance class will likely be higher than a lower call index. In this case we may also want to investigate factors that effect these probabilities. While it is possible to estimate each of the individual event probabilities, it is more conducive to view events as a processes of correct detection and misclassification, where factors affecting each can be independently investigated.

Royle and Link (2005) provided a re-parameterization that viewed events in terms separate process of correct classification (i.e. $p^k = \pi^{k|k}$) and conditional misclassification (β^{kj} for $j < k$). The conditional misclassification β^{kj} represents the probability of observing call index j, or lower, given that the correct classification of call index k was not made first (e.g. $\pi^{2|3} = (1 - p^3_{it})\beta^{32}$). The original analysis also considered single-season with a site and survey specific correct classification probability $p^k_{i,t}$ but common misclassification probability shared across all sites. As such, the single-season event matrix is redefined as

$$\boldsymbol{\pi}_{i,st} = \boldsymbol{\pi}_{it} = \begin{array}{cccc|c}
\multicolumn{4}{c}{\text{Events}} & \\
u_0 & u_1 & u_2 & u_3 & \text{States} \\ \hline
1 & 0 & 0 & 0 & z_0 \\
q^1_{it} & p^1_{it} & 0 & 0 & z_1 \\
(1-\beta^{21})q^2_{it} & \beta^{21}q^2_{it} & p^2_{it} & 0 & z_2 \\
(1-\beta^{31})(1-\beta^{32})q^3_{it} & \beta^{31}(1-\beta^{32})q^3_{it} & \beta^{32}q^3_{it} & p^3_{it} & z_3
\end{array}$$

where $q^k_{it} = 1 - p^k_{it}$ is the complement. Note that the correct classification p^k_{it} and misclassification β^{kj} probabilities are used as a type of continuation ratio. Using a continuation ratio also ensures that each row in $\boldsymbol{\pi}_{it}$ automatically sums to one as long as p^k_{it} and β^{kj} are probabilities.

Direct estimates p_{it} will not be possible, because the model will be over-parameterized; instead an additive model of the form

$$\operatorname{logit}(p^k_{it}) = \alpha_k + \sum_{l=1}^{L} b_l x_{l,it} \quad \text{for } k = 1, 2, 3$$

was used to model correct detection. A separate α_k intercept is used for each of the k correct classification parameters, but with the same additive effect shared across classes. This assumes that although the base line detection of each class of correct classification probability may differ, covariates have the same additive effect (on the logit scale).

The NAAMP protocol recognizes three well defined sampling windows, within which different species of anuran are expected to have their peak breeding activity and highest vocalization rate. Each survey t was conducted in a separate survey window producing the following additive model

$$\begin{aligned} \text{logit}(p_{i1}^{k}) &= \alpha_k + b_1 \\ \text{logit}(p_{i2}^{k}) &= \alpha_k \\ \text{logit}(p_{i3}^{k}) &= \alpha_k + b_3 \end{aligned} \tag{7.1}$$

where a separate intercept α_k is used for each of the three possible classes (i.e. $k = 1, 2, 3$). The second survey window is treated as the baseline, while all other survey windows have the same additive effect for all three classes. Other potential covariates such as survey window were also investigated by Royle and Link (2005), but have not been included here. We will be re-examining the NAAMP data set in our numerical examples section.

Chapter 8

Numerical Example: The NAAMP Data Set

8.1 Introduction

The NAAMP data set represent a complex challenge and a good test case for the flexibility of the proposed framework. While the survey protocol attempts to standardized on a number of parameters, anuran detection rates can vary substantial from one sampling occasion to the next, across years, sites and geographically (e.g. by region) and may be affected large number of factors (Grant et al., 2005; Royle and Link, 2005; Weir et al., 2005). Of three studies analyzing the NAAMP call data specifically, two gave careful consideration to potential factors affecting detection (Table 8.1). All studies to date have been in one region, Maryland, and used observations from trained staff.

Table 8.1: Summary of the characteristics of previous NAAMP call index analyses.

Study	Survey Years	Framework	Event Covariates	Study Region
Weir et al. (2005)	2002	Site-occupancy	Yes[a,b]	Maryland
Royle and Link (2005)	2001	Abundance Class	Yes[a,c]	Maryland
Mackenzie et al. (2009)	2001-2005	Abundance Class	No	Maryland

[a] The effects of survey temperature on call events were investigated.

[b] Call event seasonality was modeled as a sinusoidal day-of-year function.

[c] Call event seasonality was modeled via survey window blocking.

Using a site-occupancy framework, Weir et al. (2005) presented the most exhaustive investigation into the environmental factors effecting calling events of ten anuran species. Factors included seasonality (i.e. day-of-year), ambient temperature, time after sunset, wind, cloud cover, moon illumination, rain fall, type of habitat and anthropomorphic sources such as observer experience and proximity to human activity. Of these seasonality and temperature appeared to be among the strongest effects. However, the site-occupancy approach treats all observable call indices (i.e. '1', '2' or '3') as a single response category (i.e. occupied or a '1'). As such, additional abundance related information is discarded and potential heterogeneity that may exist between detection rates for different ecological states is left unmodeled. In addition, a site-occupancy approach provides no consideration of how environmental factors may affect the misclassification rates (e.g. observing a '1' at a site that has an latent occupancy state of '3').

The analyses by Royle and Link (2005) and Mackenzie et al. (2009) took a multi-state occupancy approach, but examined the effects of environmental factors in a simpler fashion. Royle and Link (2005) provided the foundation of the multi-state formulation for the NAAMP call indices and considered the effects of seasonality and temperature on the probability of correct classification. Misclassification probabilities were however modeled as being constant. The multi-year analysis of Mackenzie et al. (2009) was strictly focused on estimating the latent occupancy state proportions (i.e. the proportion of the landscape that had an abundance class of 0, 1, 2 and 3) and did not consider any covariate factors. Instead, a general event, initial state, and transition probability structure was used.

Of the two studies that investigated the effect of covariates on detection, the manner in which seasonality was handled differed substantially. Weir et al. (2005) used a fairly complicated approach, modeling baseline detection probabilities with a sinusoidal day-of-year function on the logit scale with additive effects of other factors. By contrast Royle and Link (2005) took a simpler approach modeling seasonality in terms of survey window block effect. The NAAMP protocol defines three possible survey windows to cover the calling phenology (the annual timing of when a species calls) of monitored species. While we may expect detection to be changing on per day basis, peak detection for any given species is likely to occur in one of the three survey windows and depending on how rapidly detection changes across and within survey windows, blocking adequately describe changes in detection due to seasonality.

Our goal in analyzing the NAAMP data set will be to provide a further investigation the factors effecting event probabilities by comparing and contrast the two previous approaches used to model seasonality. In our analysis we will also include additional regional areas (ten states in total as opposed to one) and focusing on volunteer data, rather than observations obtained from trained staff. Our preliminary analysis of the NAAMP data set includes observations from a six year period (2001 - 2006) for two species, *Hyla versicolor* (Gray Treefrog) and *Bufo fowleri* (Fowler's Toad), which have traditionally demonstrated lower detection probabilities (Weir et al., 2005). We also be focusing our attention on how seasonality and temperature effect both correct classification as well as misclassification probabilities. In doing so we will be employing an information theoretic approach whereby we rank and compare a set of *a priori* models based on our objectives (see Burnham and Anderson, 2002).

8.2 Model Formulation

In designing the models we have focused on modeling event probabilities, with minimal structure on the initial occupancy states and intermediate structure on the state transitions (Figure 8.1). Our goal is simply to demonstrate how potential biological processes may be modeled under our multi-state occupancy framework.

8.2.1 Event Probabilities

Although standardization exists we may still expect event probabilities may change on a per-survey basis, and we will need a site, season, and survey specific event formulation. Starting with the alternate event probability formulation provided by Royle and Link (2005) and modeling all parameters as being survey-specific, the following event probability matrix is obtained:

$$\boldsymbol{\pi}_{i,st} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ q^1_{i,st} & p^1_{i,st} & 0 & 0 \\ (1-\beta^{21}_{i,st})q^2_{i,st} & \beta^{21}_{i,st}q^2_{i,st} & p^2_{i,st} & 0 \\ (1-\beta^{31}_{i,st})(1-\beta^{32}_{i,st})q^3_{i,st} & \beta^{31}_{i,st}(1-\beta^{32}_{i,st})q^3_{i,st} & \beta^{32}_{i,st}q^3_{i,st} & p^3_{i,st} \end{bmatrix}, \tag{8.1}$$

where $q^k_{i,st} = 1 - p^k_{i,st}$ is the complement of the correct classification probability $p^k_{i,st}$ and $\beta^{kj}_{i,st}$ is the conditional misclassification probability as defined by Royle and Link (2005). Note

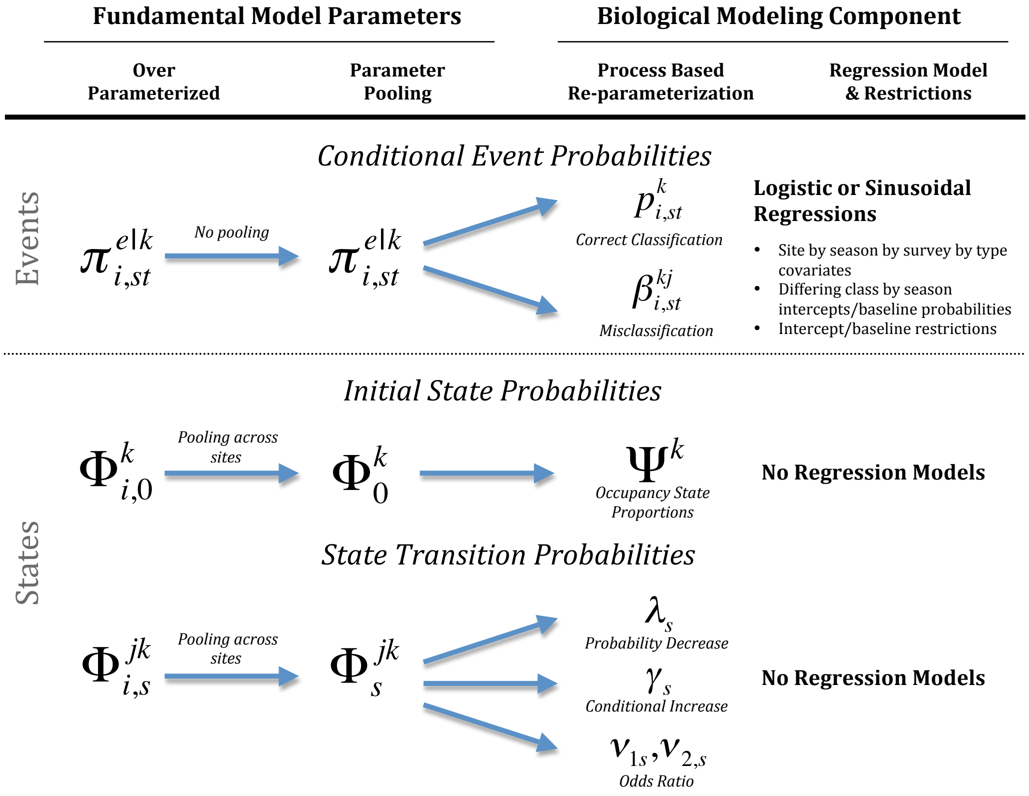


Figure 8.1: Summary of the NAAMP example analysis model formulation, re-parameterization and parameter regressions. The over-parameterized fundamental parameterized were first pooled according to end modeling goals. Event probabilities were site-, season- and survey-specific, while initial state probabilities were pooled across all sites and transitions were pooled across sites but left season-specific. After parameter pooling, fundamental parameters were then re-parameterized in terms of relevant biological processes at which point further regression type modeling was used on event probabilities.

that unlike previous formulations which modeled the conditional misclassification as being constant, we will be treating it similarly to correct classification by making it site, season and survey specific.

Primarily we are interested in further understanding how seasonality and temperature may effect event probabilities. To do so we will be have to develop two sets of general regression equations, one for modeling seasonality in terms of survey windows and the second for modeling seasonality as a day-of-year function.

Modeling Seasonality With Survey Windows

Modeling seasonality in terms of survey windows is a straightforward implementation of standard logistic model approaches, with the exception that we must accommodate differences in baseline detection rates that may occur between between parameter classes and between seasons (i.e. years in the context of the NAAMP). We follow the approach taken by Royle and Link (2005) whereby we use different intercept terms to specify different different baseline probabilities for distinctions we view as important, such as the parameter class and seasons. The covariate effect is then modeled as providing consistently the same effect across these various baseline probabilities (e.g. see equation 7.1 in the example formulations). As such we are assuming that the covariate effect acts identically across classes or season, but that we must accomodate basic differences between these classes and seasons.

In our case we wish to accommodate different baseline probabilities for the different classes within a fundamental parameter (e.g. the different correct classification probabilities) and study year (i.e. season). Because we will be believe covariates effect will most likely differ between fundamental parameter groupings we will separately consider regression equations for correct classification and misclassification fundamental parameter groupings. This results in two general regression equations:

$$\operatorname{logit}\left(p^k_{i,st}\right) = \mu^p_{s,k} + \sum_{l=1}^{L_p} b^p_l x^p_{l,i,st} \qquad \text{for } k = 1,2,3;\; s = 1,2,\ldots,6 \tag{8.2}$$

$$\operatorname{logit}\left(\beta^{kj}_{i,st}\right) = \mu^\beta_{s,kj} + \sum_{l=1}^{L_p} b^\beta_l x^\beta_{l,i,st} \qquad \text{for } k = 1,2,3,\ j < k \text{ and } s = 1,2,\ldots,6, \tag{8.3}$$

where $\mu^p_{s,k}$ and $\mu^\beta_{s,kj}$ are class and season specific intercepts for correct and conditional misclassification probabilities respectively. These allow different baseline probabilities to be assigned based on fundamental parameter class and by season (in this context year). Onto these baseline probabilities the additive covariate effect by fundamental parameter are added. Covariates values for the correct classification and misclassification regressions are designated by $x^p_{l,i,st}$ and $x^\beta_{l,i,st}$, with b^p_l and b^β_l representing the respective regression coefficients.

Exploring seasonality in terms of survey windows with the added effect of survey temperature produces following regression models,

$$\begin{aligned} \text{logit}\left(p^k_{i,st}\right) &= \mu^p_{s,k} + b_1 Win1_{i,st} + b_3 Win3_{i,st} + c_1 Temp_{i,st} + c_2 Temp^2_{i,st} && (8.4) \\ \text{logit}\left(\beta^{kj}_{i,st}\right) &= \mu^\beta_{s,kj} + d_1 Win1_{i,st} + d_3 Win3_{i,st} + e_1 Temp_{i,st} + e_2 Temp^2_{i,st}, && (8.5) \end{aligned}$$

where $Win1_{i,st} = 1$ and $Win3_{i,st} = 1$ if the Julian date of survey t that occurred on site i in season s is within the range $[45, 110)$ and $[150, 200)$ respectively, otherwise both are zero. $Temp_{i,st}$ and $Temp^2_{i,st}$ represent the linear and quadratic temperature values for a given survey. Note that, the date ranges are slightly larger than those specified in the NAAMP protocol.

Modeling Seasonality Using A Day-Of-Year Function

While survey windows provide a straightforward approach, it may not provide a sufficiently realistic approximation of the changes that may occur in detection rates over time. Instead, we may expect event probabilities to change steadily and predictably on a per day basis. Weir et al. (2005) modeled these types of changes using a sinusoidal regression (also known as “harmonic regression” see Bloomfield, 1976), as a way to model these potential changes.

Similar to the survey windows approach, we have designed the regression model parameterization to allow different baseline probabilities based on the fundamental parameter class and the season (i.e. year). However, we take a slightly different approach by defining a separate day-of-year detection curve rather than different constant baseline probability. Again we wish to consider correct and misclassification probabilities separately producing the following set of general regression curves,

$$\text{logit}\left(p^k_{i,st}\right) = \mu^p_k + \alpha^p_k \cdot \cos\left(\frac{2\pi d_{i,st}}{365} + \delta^p_s\right) + \sum_{l=1}^{L_p} b^p_l x^p_{l,i,st} \quad \text{for } k = 1,2,3 \text{ and } s = 1,2,\ldots,6 \quad (8.6)$$

$$\text{logit}\left(\beta^{kj}_{i,st}\right) = \mu^{\beta}_{kj} + \alpha^{\beta}_{kj}\cdot\cos\left(\frac{2\pi d_{i,st}}{365} + \delta^{\beta}_{s}\right) + \sum_{l=1}^{L_p} b^{\beta}_{l} x^{\beta}_{l,i,st} \text{ for } k = 2, 3,\ j < k \text{ and } s = 1, 2, \ldots, 6 \quad (8.7)$$

where μ^p_k and μ^β_{kj} are the class specific intercepts, α^p_k and α^β_{kj} characterize the class specific amplitudes and δ^p_s and δ^β_s are the season specific phases within a 365 day cycle. The Julian day a specific survey occurred on is specified by $d_{i,st}$ and the additional covariate values.

Directly interpretation of the individual parameters (e.g. μ^p_k, α^p_k and δ^p_s) can be awkward as the μ^p_k represents the baseline detection rate at one quarter of a cycle from either the min or max detection and the cosine wave form occurs on the logit scale. Instead, it is best to view this component as providing a baseline detection rate on to which other environmental effects are added. It should also be noted that we have made the simplifying assumption that for a fundamental parameter class (e.g. p^3) all seasonal curves share the same curve profile with the exception of phase. The intensity of the detection probability can change between parameter classes. Finally, we have also allowed Julian day corresponding to peak detection to change between seasons.

In the context of our analysis, seasonality and temperature is modeled using the following regression equations,

$$\text{logit}\left(p^{k}_{i,st}\right) = \mu^{p}_{k} + \alpha^{p}_{k}\cdot\cos\left(\frac{2\pi d_{i,st}}{365} + \delta^{p}_{s}\right) + c_1\text{Temp}_{i,st} + c_2\text{Temp}^2_{i,st} \quad (8.8)$$

$$\text{logit}\left(\beta^{kj}_{i,st}\right) = \mu^{\beta}_{kj} + \alpha^{\beta}_{kj}\cdot\cos\left(\frac{2\pi d_{i,st}}{365} + \delta^{\beta}_{s}\right) + e_1\text{Temp}_{i,st} + e_2\text{Temp}^2_{i,st}. \quad (8.9)$$

8.2.2 States and Transitions

Initial State Distribution Ψ^k_i

Goals of a specific analysis will dictate whether the initial latent state distribution should be modeled as site specific or pooled across sites. The research objective requires estimates of the proportion of the landscape exhibiting a given latent abundance state, the pooled parameterization is required. Otherwise a covariate model can be used to specify a site-specific initial state distribution.

Depending on the goals of a specific analysis it may or may not be of interest to model the initial state distribution as being site specific or pooled across sites. If the goal is estimation of the proportion of sites across the landscape that exhibit a given occupancy state then pooling sites will be necessary. For predictive goals (i.e. understanding what factors are associated with a given occupancy state) site-specific initial state probabilities will be a prerequisite. Note than an exception exists if a site specific model and the covariate distribution across the landscape is known then the latent state distribution across the landscape may also be estimated.

In our case an argument can be made for both approaches. As a monitoring study the NAAMP is primarily interested in understanding changes in the proportion of the landscape used (i.e. occupied) and changes in the level of occupancy. However it would also be naïve to suggests all sites share the same state distribution. The NAAMP data set spans 17 states (of which we analyzed 10), it is unlikely that all states have the same latent abundance class distribution nor the same transition probabilities.

That said, the presented analysis only is focused on modeling the event probabilities to simplify the analysis as much as possible. We are therefore making the simplifying assuming that $\Psi_i^k = \Psi^k$ for all i and k.

Abundance Class Seasonal Transitions Φ_s

Given the monitoring objective of the NAAMP likely there is interest in understanding transitions in terms of aggregate statistics. We will be modeling all sites using common transition probabilities (i.e. $\mathbf{\Phi}_{i,s} = \mathbf{\Phi}_s$), and we will re-parameterizing the $\mathbf{\Phi}_s$ matrix in terms of processes that are believed to be biologically meaningful. As the abundance class model is a form of multi-level occupancy, we turn to site-occupancy framework for inspiration.

The multi-season site-occupancy model proposes that changes in occupancy results from two separate processes that occur between season. These processes are colonization (probability γ_s) and extinction (probability ϵ_s; MacKenzie et al., 2006). While highly applicable to a strict site-occupancy framework, these processes are somewhat broad from a multi-state occupancy context. Sites may experience decreases in abundance without local extinction, and likewise increases in abundance without colonization. In a monitoring context, both of these changes would be of interest to researchers, with priority placed on detecting declines. (Most monitoring programs are put in place because of concerns over the continued survival of a particular species.)

For multi-state occupancy we can break down changes in the occupancy state into the process of decline (roughly akin to the extinction process in site-occupancy, but defined as a decrease in the abundance class), and then given a decline has not occurred the conditional process of an increase. The probability of no change, becomes the complement of these two processes. Formally, we use the following parameters,

λ_s - The probability a site decreases by one or more abundance levels between seasons s and $s+1$, given a decrease is possible (i.e. $Z \neq 0$).

γ_s - The conditional probability of an increase in the latent abundance class between seasons s to $s+1$, given an decrease has not occurred and an increase is possible (i.e. $Z \neq K$).

$\nu_{1,s}$ - Given an decrease in abundance has occurred, $\nu_{1,s}$ is the recursive continuation ratio, that is the probability of selecting the next lowest abundance class.

$\nu_{2,s}$ - Given an increase in abundance has occurred, $\nu_{2,s}$ is the recursive continuation ratio, that is the probability of selecting the next highest abundance class.

Given that we are assuming all sites are subject to the same transition probabilities between season $\mathbf{\Phi}_s$ is now re-expresses as,

$$\mathbf{\Phi}_{i,s} = \mathbf{\Phi}_s = \begin{bmatrix} (1-\gamma_s) & \gamma_s\nu_{2,s} & \gamma_s(1-\nu_{2,s})\nu_{2,s} & \gamma_s(1-\nu_{2,s})(1-\nu_{2,s}) \\ \lambda_s & (1-\lambda_s)(1-\gamma_s) & (1-\lambda_s)\gamma_s\nu_{2,s} & (1-\lambda_s)\gamma_s(1-\nu_{2,s}) \\ \lambda_s(1-\nu_{1,s}) & \lambda_s\nu_{1,s} & (1-\lambda_s)(1-\gamma_s) & (1-\lambda_s)\gamma_s \\ \lambda_s(1-\nu_{1,s})(1-\nu_{1,s}) & \lambda_s(1-\nu_{1,s})\nu_{1,s} & \lambda_s\nu_{1,s} & (1-\lambda_s) \end{bmatrix}.$$

8.2.3 Multi-model Inference and Compact Model Syntax

In ecology, it is becoming more common place to investigate multiple competing hypotheses when interpreting natural phenomenal (Burnham and Anderson, 2002; Johnson and Omland, 2004). Within the context of our example analysis we have multiple competing hypotheses about the way event probabilities can be affected by day-of-year and temperature effect.

For example, does a more simple approach such as survey window blocking adequately describe seasonality compared to a more complex approach such as a sinusoidal regression? Furthermore, do we need to include further restrictions within any one of these approaches (e.g. across season)? In the most general form, both seasonality formulations allow for differences in the baseline probabilities by functional class (i.e. correct and misclassifications probabilities) as well as subclass (e.g. $p^3_{i,st}$ versus $p^2_{i,st}$) and by year. We may wish to explore scenarios were these baseline probabilities are restricted across subclass (e.g. misclassification: $\beta^{31}_{i,st} = \beta^{21}_{i,st}$) or across years (e.g. $\beta^{31}_{i,1t} = \beta^{31}_{i,2t} = \ldots = \beta^{31}_{i,6t}$).

With these goals in mind we have developed a compact model syntax to represent the various models and constraints. The syntax uses a general structure with an indicator for the fundamental model parameter (e.g. 'p' for the correct detection class $p^k_{i,st}$) followed by further indicators (optional) then class and seasonal constraints. For example, the model syntax $p\{cos\}\{1,2,3\}\{.\}$ would indicate the correct detection regression that uses the sinusoidal day-of-year function for seasonality and has a differing baseline probability based on class, but is shared across seasons (i.e. $p^k_{i,1t} = p^k_{i,2t} = \cdots = p^k_{i,6t}$ for all values of i, t and k). In this case $p^k_{i,st}$ is defined by a sinusoidal day-of-year function (8.8) and the restriction is accomplished by the constraint $\delta^p_1 = \delta^p_2 = \cdots = \delta^p_6$. For the survey window regression (8.4) the equivalent model (denoted as $p\{win\}\{1,2,3\}\{.\}$) is accomplished by constraining $\mu^p_{1,k} = \mu^p_{2,k} = \cdots = \mu^p_{6,k}$ for all k.

While the manner in which constraints are enacted differ, both result in a roughly equivalent outcome. For the survey windows regression a constant baseline probability is defined across seasons for each class, while for the sinusoidal regression each class has a separate curve which is shared across seasons. In both cases the compact model syntax relates these concepts without detailing how the constraints were executed.

Finally, noting that the model syntax deals with class and season constraints only. We do not consider constraints across either the i or t indices because covariate values (i.e. day of observation and temperature) will change by site and survey occasion. For the analysis we used a fairly flexible state transition model so that we could focus finding the best event model formulation.

8.3 Goodness-of-fit for the NAAMP Ordinal Abundance Class Model

Using the omnibus goodness-of-fit measure provided in Section 6.4.1 may be computationally infeasible for large data, due to the time taken to numerically maximize the models after each bootstrap sample is generated. Model adequacy however can still be assessed in this scenario by employing a similar approach to Bayesian posterior predictive checking.

If the model fits, then replicate data generated under the model (using MLE's from the model as parameter values) should be similar to the observed data. Lack-of-fit tests can be formulated to test various underlying assumptions of the model by comparing the arrangement of the observed data to the relevant empirical distribution generated from the bootstrap sample. For example, the probability of observing a given encounter history, or a site-specific encounter history (for site-specific covariate models) could be compared. These potentially, "test statistic" would be akin, in concept to the Bayesian posterior p-value (Meng, 1994) and may allow for the incorporation of other techniques such as a model adequacy distance statistical as proposed by Lindsay (2004).

While this approach offers promise, computational short comings may also exist for complex data sets such as the NAAMP. Long encounter histories (either through large numbers of within season surveys and/or many seasons) present a large number of possibilities that bootstrap samples may not adequately cover without being prohibitively large. This problem can however be reduced for multi-state occupancy models that use the latent abundance class formulation, an ordinal occupancy states formulation.

In the latent abundance class formulation (and potentially with other similar ordinal occupancy formulations) the length of the encounter histories can be reduced by considering the maximum observed index within a season. During this period the information about the underlying state has not changed, and multiple surveys can be condensed into a univariate summary. For example, the true underlying state could be theoretically determined by infinitely resampling the site during a state of closure (i.e. within a season) and taking the maximum, that is

$$\lim_{T_s \to \infty} \max\{h_{i,st} : t = 1, 2, \ldots, T_s\} = z_{i,s} \quad \text{for } s = 1, 2, \ldots, S, \tag{8.10}$$

if the probability of the correct classification event is non-zero. That is, the maximum observed index converges in probability to the true underlying state, if and only if the probability of a correct classification event is non-zero. Since this assumption can be expected to be true in most cases, the maximum index within a season effectively summarizes all the available information on the underlying state and will be used as the basis of our goodness-of-fit statistic.

Building off of this, a number of lack-of-fit criteria can be created by following a general procedure:

1. Fit the model to observed data and derive parameter estimates for the event probabilities $\widehat{\pi^{e|k}_{i,st}}$, the initial state probabilities $\widehat{\mathbf{\Phi}_{i,0}}$ and the transition probabilities $\widehat{\mathbf{\Phi}_{i,s}}$.

2. Using the parameter estimates compute for each site, denoted as i, do the following:
   (a) Assign an initial state from $\mathcal{Z}$ to site i using a multinomial draw with cell probabilities $\widehat{\mathbf{\Phi}_{i,0}}$. We will denote this state as $z_{i,1}$.
   (b) Next for $s = 1 \ldots S$ each seasons do the following
      i. Make repeated multinomial draws from $\mathcal{U}$, using the row probabilities from $\widehat{\pi_{i,st}^{e|k}}$ associated with state $z_{i,s}$ and assign T_s events to site i in season s. Once T_s events have been assigned record the maximum index as the observation for season s.
      ii. Next, if $s < S$, assign a new state to i for season $s + 1$, by making a multinomial draw using the row probabilities from $\widehat{\mathbf{\Phi}_{i,s}}$ associated with being in state k in season s. Denote this new state as $z_{i,s+1}$.

   The observed events from Steps 2a-2b will make up the maximum index encounter history of site i. Record the row vector as row i in $\boldsymbol{X}^b$.

3. Repeat Step 2 a large number of times, B, this is the bootstrap sample which will be used to approximate the distribution of lack-of-fit metrics based on the maximum index.

4. Using the observed data for each site, denoted as i, do the following:
   (a) For $s = 1, 2, \ldots, S$ determine the maximum observed occupancy index in season s and record this as the observation for season s.
   (b) Record these observed maximum indices for site i as row i in $\boldsymbol{X}^{\text{obs}}$.

   $\boldsymbol{X}^{\text{obs}}$ will be used in the calculation of any lack-of-fit metric.

5. Generate a lack-of-fit metric using $\boldsymbol{X}^{\text{obs}}$ and compare to the lack-of-fit empirical distribution derived from the $\boldsymbol{X}^b$ sample.

Within the context of our NAAMP example, we use the bootstrap distribution to assess whether the top selected model adequately predicts the transitions of states between seasons. While we cannot fully assess this due to issues of detectability, we can compare observed and expected changes between maximum indices between seasons. Note that we have deliberately substituted the word “change” for transition, since we cannot assess the true rate of state transitions.

8.4 Results

Call chorus data for *Hyla versicolor* and *Bufo fowleri* from 2001 to 2006 in 10 states was obtained from the online from the United States Geological Survey website[1]. For *H. versicolor* there were a total of 28,613 surveys recorded at 2,867 sites over 6 years, while for 22,282 total surveys for *B. fowleri* were recorded at 2,257 sites.

The analysis was approached in stages, where we considered eight event model scenarios each with additional possible multiple parameter restrictions based on parameter class and season. Each scenarios contained used either a sinusoidal or survey window approach to define baseline probabilities. Models for a temperature effect included, effects on correct detection only, or effects on both correct and misclassification. The remaining scenarios use either a survey windows or sinusoidal regression approach for the correct classification probabilities, with a constant misclassification probability. These scenarios are intended to provided a frame of reference to the Royle and Link (2005) analysis which modeled a constant misclassification probability.

Within each event scenario we also consider an equivalent set of parameter restrictions. In its most general form all regressions have differing baseline probabilities based on fundamental parameter classes and season. Such a high degree of parameterization could potentially represent an over-parameterized model. We also developed an *a priori* a set of biologically relevant class and season restrictions for the correct and misclassification probabilities (Table 8.2). These restrictions can be applied within any of the event scenarios and represent different ways to define the baseline correct or misclassification probabilities in terms of either class or season.

In total we consider two season restrictions (either $\{s\}$ or $\{.\}$) and three class restrictions for correct classification and four class restrictions for misclassification. There are 48 possible event model restrictions within any given event models scenario. Within each scenario all models were ranked by AIC, and we present only the top three models from each scenario for each of the two species (see Tables 8.3 and 8.4 for the top *Bufo fowleri* sinusoidal and survey windows models and Tables 8.5 and 8.6 for the top *Hyla versicolor* sinusoidal and survey windows models).

Support for the different event model scenarios was evaluated by comparing the AIC values of the top baseline restriction model within each of the event scenarios (Table 8.7).

---

[1] `http://www.pwrc.usgs.gov/naamp/index.cfm?fuseaction=app.dataDownload`

Table 8.2: Summary of the various tested restrictions placed on the correct and misclassification probabilities baseline probabilities. All combinations of class and season restrictions where used the different event scenarios tested. for example $p\{k\}\{s\}$ would be a model that uses different baseline probabilities for each class and season.

Fundamental	Class		Season	
Parameter	Restriction	Syntax	Restriction	Syntax
Correct Classification	–	$p\{1,2,3\}$	–	$\{s\}$
	$p^2_{i,st} = p^3_{i,st}$	$p\{1,23\}$	$p^k_{i,1t} = p^k_{i,2t} = \cdots = p^k_{i,6t}$	$\{.\}$
	$p^1_{i,st} = p^2_{i,st} = p^3_{i,st}$	$p\{.\}$		
Misclassification	–	$\beta\{21,31,32\}$	–	$\{s\}$
	$\beta^{21}_{i,st} = \beta^{31}_{i,st}$	$\beta\{21\|31,32\}$	$\beta^{kj}_{i,1t} = \beta^{kj}_{i,2t} = \cdots = \beta^{kj}_{i,6t}$	$\{.\}$
	$\beta^{31}_{i,st} = \beta^{32}_{i,st}$	$\beta\{21,31\|32\}$		
	$\beta^{21}_{i,st} = \beta^{31}_{i,st} = \beta^{32}_{i,st}$	$\beta\{.\}$		

While the top baseline restrictions within each event scenario were generally close in terms of AIC support (e.g. 0-10 AIC units, see top ranked models in each scenario in Tables 8.3 and 8.4 for *Bufo fowleri*, and Tables 8.5 and 8.6 for *Hyla versicolor*) the differences between event scenarios was startlingly large, on the order of hundreds of AIC units (Table 8.7). As a rule of thumb models within 10 AIC units of one another are generally considered plausible (Burnham and Anderson, 2002).

The top scenario for both *Bufo fowleri* and *Hyla versicolor* used the day-of-year (sinusoidal) regression approach, with separate baseline probabilities for each season for both correct and misclassification probabilities. Both top models also included survey specific temperature effects for both correct and misclassification probabilities.

For *Bufo fowleri* the next closest scenario was 112 AIC units away and used survey windows to model within-season changes, had separate baseline probabilities for each season, and included temperature effects for both correct and misclassification probabilities. In fact of top five scenarios, three used the sinusoidal regression to model within season changes, while two used survey windows and most included some form of temperature effects (Table 8.7). Finally, the top models in each scenario had used differing baseline probabilities for each season.

For *Hyla versicolor* the sinusoidal regression approach was used in all of the top three

Table 8.3: Top AIC ranked day-of-year regression event models for *Bufo fowleri*. Models have been ranked within each set, and sets ranked overall by the top model.

Model Constraints	np	AIC	ΔAIC	$\hat{\bar{N}}_1$	$\hat{\bar{N}}_2$	$\hat{\bar{N}}_3$	$\hat{\bar{N}}_4$	$\hat{\bar{N}}_5$	$\hat{\bar{N}}_6$	$\hat{b}_1$	$\hat{b}_2$	$\hat{c}_1$	$\hat{c}_2$	$\hat{d}_1$	$\hat{d}_2$	$\hat{e}_1$	$\hat{e}_2$
Day-of-year regression $\left(p^k_{i,st} \;\&\; \beta^{kj}_{i,st}\right)$ **with temperature** $\left(p^k_{i,st} \;\&\; \beta^{kj}_{i,st}\right)$																	
$p\{cos\}\{1,23\}\{s\}\{temp\}$ $\beta\{cos\}\{21\|31,32\}\{s\}\{temp\}$	39	10458.6	0.0	0.76	0.78	0.89	0.81	0.83	0.69	–	–	0.34	-0.02	–	–	0.22	-0.01
$p\{cos\}\{1,2,3\}\{s\}\{temp\}$ $\beta\{cos\}\{21\|31\|32\}\{s\}\{temp\}$	39	10462.5	3.9	0.77	0.79	0.91	0.82	0.85	0.71	–	–	0.34	-0.02	–	–	0.22	-0.01
$p\{cos\}\{1,2,3\}\{.\}\{temp\}$ $\beta\{cos\}\{21,31,32\}\{s\}\{temp\}$	38	10463.3	4.7	0.95	1.00	1.29	1.18	1.26	1.11	–	–	0.34	-0.02	–	–	0.25	-0.01
Day-of-year regression $\left(p^k_{i,st} \;\&\; \beta^{kj}_{i,st}\right)$ **with temperature** $\left(p^k_{i,st}\right)$																	
$p\{cos\}\{1,2,3\}\{s\}\{temp\}$ $\beta\{cos\}\{21,31,32\}\{s\}$	41	10592.8	0.0	0.61	0.64	0.78	0.69	0.68	0.59	–	–	0.37	-0.02	–	–	–	–
$p\{cos\}\{1,23\}\{s\}\{temp\}$ $\beta\{cos\}\{21,31,32\}\{s\}$	39	10596.4	3.5	0.62	0.64	0.79	0.69	0.71	0.60	–	–	0.37	-0.02	–	–	–	–
$p\{cos\}\{1,23\}\{s\}\{temp\}$ $\beta\{cos\}\{21\|31\|32\}\{s\}$	35	10597.3	4.4	0.61	0.63	0.78	0.69	0.69	0.61	–	–	0.38	-0.02	–	–	–	–
Day-of-year regression $\left(p^k_{i,st} \;\&\; \beta^{kj}_{i,st}\right)$																	
$p\{cos\}\{1,2,3\}\{s\}$ $\beta\{cos\}\{21,31,32\}\{s\}$	39	10896.4	0.0	0.57	0.59	0.77	0.68	0.69	0.57	–	–	–	–	–	–	–	–
$p\{cos\}\{1,23\}\{.\}$ $\beta\{cos\}\{21\|31,32\}\{s\}$	30	10904.2	8.0	0.62	0.64	0.80	0.74	0.69	0.60	–	–	–	–	–	–	–	–
$p\{cos\}\{1,2,3\}\{.\}$ $\beta\{cos\}\{21\|31,32\}\{s\}$	32	10907.2	11.4	0.62	0.65	0.81	0.74	0.69	0.61	–	–	–	–	–	–	–	–

Table 8.4: Top AIC ranked survey window event models for *Bufo fowleri*. Models have been ranked within each set, and sets ranked overall by the top model.

Model Constraints	np	AIC	ΔAIC	$\hat{\tilde{N}}_1$	$\hat{\tilde{N}}_2$	$\hat{\tilde{N}}_3$	$\hat{\tilde{N}}_4$	$\hat{\tilde{N}}_5$	$\hat{\tilde{N}}_6$	$\hat{b}_1$	$\hat{b}_2$	$\hat{c}_1$	$\hat{c}_2$	$\hat{d}_1$	$\hat{d}_2$	$\hat{e}_1$	$\hat{e}_2$
Survey Windows $\left(p^k_{i,st}\ \&\ \beta^{kj}_{i,st}\right)$ **with temperature** $\left(p^k_{i,st}\ \&\ \beta^{kj}_{i,st}\right)$																	
$p\{1,2,3\}\{s\}\{temp\}$ $\beta\{21,31\vert 32\}\{s\}\{temp\}$	53	10570.5	0.0	0.82	0.82	1.20	1.32	1.22	1.10	-0.17	-0.94	0.41	-0.02	-1.62	-0.48	0.26	-0.01
$p\{1,2,3\}\{s\}\{temp\}$ $\beta\{21,31,32\}\{s\}\{temp\}$	59	10572.7	2.3	0.85	0.85	1.21	1.32	1.24	1.13	-0.18	-0.94	0.41	-0.02	-1.64	-0.48	0.26	-0.01
$p\{1,2,3\}\{.\}\{temp\}$ $\beta\{21,31,32\}\{s\}\{temp\}$	44	10582.5	11.7	0.76	0.90	1.22	1.20	1.17	0.93	-0.22	-0.88	0.40	-0.03	-1.58	-0.51	0.27	-0.01
Survey Windows $\left(p^k_{i,st}\ \&\ \beta^{kj}_{i,st}\right)$ **with temperature** $\left(p^k_{i,st}\right)$																	
$p\{1,23\}\{s\}\{temp\}$ $\beta\{21\vert 31,32\}\{s\}$	45	10754.5	0.0	0.58	0.67	0.68	0.71	0.64	0.64	-0.73	-1.19	0.43	-0.03	-2.15	0.55	–	–
$p\{1,23\}\{s\}\{temp\}$ $\beta\{21\vert 31\vert 32\}\{s\}$	39	10755.4	0.9	0.58	0.67	0.68	0.72	0.64	0.64	-0.71	-1.19	0.43	-0.03	-2.16	0.54	–	–
$p\{1,23\}\{s\}\{temp\}$ $\beta\{21,31\vert 32\}\{s\}$	45	10756.1	1.6	0.58	0.68	0.68	0.71	0.63	0.60	-0.73	-1.18	0.43	-0.03	-2.15	0.54	–	–
Survey Windows $\left(p^k_{i,st}\ \&\ \beta^{kj}_{i,st}\right)$																	
$p\{1,23\}\{s\}$ $\beta\{21\vert 31,32\}\{s\}$	43	11150.0	0.0	0.58	0.64	0.68	0.75	0.60	0.55	-1.32	-0.25	–	–	-2.40	0.32	–	–
$p\{1,23\}\{s\}$ $\beta\{21,31,32\}\{s\}$	49	11151.7	1.7	0.57	0.64	0.71	0.77	0.63	0.58	-1.27	-0.26	–	–	-2.41	0.31	–	–
$p\{1,2,3\}\{s\}$ $\beta\{21,31,32\}\{s\}$	55	11152.9	2.9	0.67	0.75	0.90	1.00	0.90	0.79	-1.18	-0.21	–	–	-2.38	0.24	–	–
Survey Windows $\left(p^k_{i,st}\right)$ **with temperature** $\left(p^k_{i,st}\right)$																	
$p\{1,23\}\{s\}\{temp\}$ $\beta\{21,31\vert 32\}\{s\}$	43	11227.1	0.0	0.50	0.66	0.62	0.64	0.64	0.48	-1.07	-1.04	0.43	-0.03	–	–	–	–
$p\{1,23\}\{s\}\{temp\}$ $\beta\{21\vert 31\vert 32\}\{s\}$	37	11229.0	1.9	0.50	0.66	0.61	0.65	0.63	0.47	-1.07	-1.04	0.43	-0.03	–	–	–	–
$p\{1,2,3\}\{s\}\{temp\}$ $\beta\{21,31\vert 32\}\{s\}$	49	11234.6	7.5	0.51	0.66	0.61	0.63	0.64	0.49	-1.07	-1.03	0.43	-0.03	–	–	–	–
Survey Windows $\left(p^k_{i,st}\right)$																	
$p\{1,23\}\{s\}$ $\beta\{21\vert 31\vert 32\}\{s\}$	35	11679.7	0.0	0.57	0.64	0.61	0.63	0.55	0.43	-1.85	-0.11	–	–	–	–	–	–
$p\{1,23\}\{s\}$ $\beta\{21\vert 31,32\}\{s\}$	41	11683.7	4.0	0.61	0.61	0.59	0.63	0.55	0.42	-1.85	-0.10	–	–	–	–	–	–
$p\{1,23\}\{s\}$ $\beta\{21,31\vert 32\}\{s\}$	41	11683.7	4.0	0.61	0.61	0.59	0.63	0.54	0.43	-1.85	-0.10	–	–	–	–	–	–

Table 8.5: Top AIC ranked day-of-year regression event models for *Hyla versicolor*. Models have been ranked within each set, and sets ranked overall by the top model.

Model Constraints	np	AIC	ΔAIC	$\hat{\hat{N}}_1$	$\hat{\hat{N}}_2$	$\hat{\hat{N}}_3$	$\hat{\hat{N}}_4$	$\hat{\hat{N}}_5$	$\hat{\hat{N}}_6$	$\hat{b}_1$	$\hat{b}_2$	c_1	c_2	d_1	d_2	e_1	e_2
Day-of-year regression $\left(p^k_{i,st} \;\&\; \beta^{kj}_{i,st}\right)$ **with temperature** $\left(p^k_{i,st} \;\&\; \beta^{kj}_{i,st}\right)$																	
$p\{cos\}\{1,2,3\}\{.\}\{temp\}$ $\beta\{21,31,32\}\{s\}\{temp\}$	38	21762.1	0.0	1.02	1.28	1.34	1.25	1.26	1.20	–	–	0.19	-0.02	–	–	0.06	-0.01
$p\{cos\}\{1,2,3\}\{s\}\{temp\}$ $\beta\{21,31,32\}\{s\}\{temp\}$	43	21768.3	6.2	1.09	1.30	1.35	1.26	1.28	1.22	–	–	0.17	-0.02	–	–	0.07	-0.01
$p\{cos\}\{1,2,3\}\{.\}\{temp\}$ $\beta\{21\|31,32\}\{s\}\{temp\}$	36	21771.0	8.9	1.17	1.36	1.41	1.31	1.33	1.27	–	–	0.17	-0.02	–	–	0.07	-0.01
Day-of-year regression $\left(p^k_{i,st} \;\&\; \beta^{kj}_{i,st}\right)$ **with temperature** $\left(p^k_{i,st}\right)$																	
$p\{cos\}\{1,2,3\}\{.\}\{temp\}$ $\beta\{21,31,32\}\{s\}$	36	21834.3	0.0	0.99	1.26	1.33	1.25	1.26	1.21	–	–	0.20	-0.02	–	–	–	–
$p\{cos\}\{1,2,3\}\{s\}\{temp\}$ $\beta\{21,31,32\}\{s\}$	41	21836.4	2.1	1.00	1.26	1.33	1.24	1.26	1.22	–	–	0.20	-0.02	–	–	–	–
$p\{cos\}\{1,2,3\}\{.\}\{temp\}$ $\beta\{21\|31,32\}\{s\}$	34	21847.2	12.9	1.04	1.31	1.38	1.28	1.30	1.26	–	–	0.20	-0.02	–	–	–	–
Day-of-year regression $\left(p^k_{i,st} \;\&\; \beta^{kj}_{i,st}\right)$																	
$p\{cos\}\{1,2,3\}\{.\}$ $\beta\{21,31,32\}\{s\}$	34	22049.9	0.0	1.18	1.33	1.40	1.33	1.31	1.31	–	–	–	–	–	–	–	–
$p\{cos\}\{1,2,3\}\{s\}$ $\beta\{21,31,32\}\{s\}$	39	22050.0	0.1	1.18	1.33	1.40	1.33	1.31	1.29	–	–	–	–	–	–	–	–
$p\{cos\}\{1,2,3\}\{.\}$ $\beta\{21\|31,32\}\{s\}$	32	22054.1	4.2	1.24	1.39	1.46	1.38	1.35	1.35	–	–	–	–	–	–	–	–

Table 8.6: Top AIC ranked survey window event models for *Hyla versicolor*. Models have been ranked within each set, and sets ranked overall by the top model.

Model Constraints	np	AIC	ΔAIC	$\hat{\hat{N}}_1$	$\hat{\hat{N}}_2$	$\hat{\hat{N}}_3$	$\hat{\hat{N}}_4$	$\hat{\hat{N}}_5$	$\hat{\hat{N}}_6$	$\hat{b}_1$	$\hat{b}_2$	c_1	c_2	d_1	d_2	e_1	e_2
Survey Windows $\left(p^k_{i,st}\ \&\ \beta^{kj}_{i,st}\right)$ **with temperature** $\left(p^k_{i,st}\ \&\ \beta^{kj}_{i,st}\right)$																	
$p\{1,2,3\}\{s\}\{temp\}$ $\beta\{21\|31,32\}\{.\}\{temp\}$	43	22692.0	0.0	1.12	1.27	1.48	1.44	1.31	1.31	-1.44	1.76	0.14	-0.02	-2.56	1.18	0.07	-0.01
$p\{1,2,3\}\{s\}\{temp\}$ $\beta\{21,31\|32\}\{.\}\{temp\}$	43	22693.3	1.3	1.40	1.35	1.49	1.44	1.31	1.31	-1.42	1.74	0.14	-0.02	-2.51	1.16	0.07	-0.01
$p\{1,2,3\}\{s\}\{temp\}$ $\beta\{21\|31\|32\}\{.\}\{temp\}$	42	22701.0	9.0	1.14	1.29	1.50	1.46	1.33	1.33	-1.43	1.77	0.14	-0.02	-2.54	1.18	0.07	-0.01
Survey Windows $\left(p^k_{i,st}\ \&\ \beta^{kj}_{i,st}\right)$ **with temperature** $\left(p^k_{i,st}\right)$																	
$p\{1,2,3\}\{s\}\{temp\}$ $\beta\{21,31\|32\}\{.\}$	41	22791.7	0.0	1.36	1.26	1.41	1.42	1.31	1.31	-1.50	1.68	0.15	-0.02	-2.68	1.40	–	–
$p\{1,2,3\}\{s\}\{temp\}$ $\beta\{21,31\|32\}\{s\}$	51	22800.0	8.3	1.39	1.26	1.44	1.38	1.33	1.33	-1.49	1.68	0.15	-0.02	-2.70	1.41	–	–
$p\{1,2,3\}\{s\}\{temp\}$ $\beta\{21\|31,32\}\{.\}$	41	22800.8	9.1	1.12	1.23	1.45	1.45	1.31	1.31	-1.50	1.71	0.15	-0.02	-2.73	1.42	–	–
Survey Windows $\left(p^k_{i,st}\ \&\ \beta^{kj}_{i,st}\right)$																	
$p\{1,2,3\}\{.\}$ $\beta\{21,31,32\}\{.\}$	25	22960.33	0.00	1.15	1.32	1.39	1.33	1.35	1.35	-1.63	2.16	–	–	-2.76	1.39	–	–
$p\{1,2,3\}\{.\}$ $\beta\{21,31\|32\}\{.\}$	24	22961.35	1.03	1.15	1.32	1.39	1.33	1.35	1.35	-1.62	2.17	–	–	-2.77	1.38	–	–
$p\{1,23\}\{s\}$ $\beta\{21,31\|32\}\{.\}$	33	22963.81	3.48	1.34	1.30	1.41	1.43	1.26	1.26	-1.68	2.06	–	–	-2.81	1.42	–	–
Survey Windows $\left(p^k_{i,st}\right)$ **with temperature** $\left(p^k_{i,st}\right)$																	
$p\{1,23\}\{s\}\{temp\}$ $\beta\{21\|31,32\}\{.\}$	33	23768.73	0.00	1.00	1.11	1.08	1.03	1.09	1.03	-2.48	1.87	0.13	-0.02	–	–	–	–
$p\{1,23\}\{s\}\{temp\}$ $\beta\{21,31,32\}\{.\}$	34	23769.79	1.06	1.00	1.10	1.07	1.03	1.09	1.03	-2.47	1.87	0.13	-0.02	–	–	–	–
$p\{1,23\}\{s\}\{temp\}$ $\beta\{21\|31,32\}\{s\}$	43	23770.43	1.70	0.98	1.09	1.06	1.04	1.09	1.04	-2.48	1.88	0.14	-0.02	–	–	–	–
Survey Windows $\left(p^k_{i,st}\right)$																	
$p\{1,23\}\{s\}$ $\beta\{21\|31,32\}\{.\}$	34	23958.7	0.0	0.96	1.10	1.07	1.05	1.07	1.10	-2.78	2.27	–	–	–	–	–	–
$p\{1,23\}\{s\}$ $\beta\{21,31,32\}\{.\}$	35	23959.9	1.3	0.96	1.09	1.07	1.05	1.07	1.10	-2.77	2.27	–	–	–	–	–	–
$p\{1,23\}\{s\}$ $\beta\{21\|31,32\}\{s\}$	44	23960.1	1.4	0.94	1.08	1.06	1.05	1.08	1.11	-2.78	2.28	–	–	–	–	–	–

scenarios. For correct detection the baseline probabilities were constant across season, while for misclassification the top models included seasonality effects. The remaining scenarios that used survey windows also favoured the inclusion of temperature effects on both correct and misclassification probabilities. Finally, the top models in these remaining scenarios also had a mixed support for the inclusion of seasonality in baseline probabilities.

Table 8.7: *Bufo fowleri* and *Hyla versicolor* event scenarios rankings by AIC. Within each event scenario various parameter restrictions were compared via AIC ranking and only the top restriction model from each event scenario are displayed below. The column np is the number of model parameters.

A) *Bufo fowleri*

Event Model				
Correct Classification	Misclassification	np	AIC	ΔAIC
$p\{cos\}\{1,23\}\{s\}\{temp\}$	$\beta\{cos\}\{21\vert 31,32\}\{s\}\{temp\}$	39	10458.6	0
$p\{win\}\{1,2,3\}\{s\}\{temp\}$	$\beta\{win\}\{21,31\vert 32\}\{s\}\{temp\}$	53	10570.5	111.9
$p\{cos\}\{1,2,3\}\{s\}\{temp\}$	$\beta\{cos\}\{21,31,32\}\{s\}$	41	10592.8	134.2
$p\{win\}\{1,2,3\}\{s\}\{temp\}$	$\beta\{win\}\{21\vert 31,32\}\{s\}$	45	10754.5	295.9
$p\{cos\}\{1,23\}\{s\}$	$\beta\{cos\}\{21\vert 31,32\}\{s\}$	39	10896.1	437.5
$p\{win\}\{1,23\}\{s\}$	$\beta\{win\}\{21\vert 31,32\}\{s\}$	43	11150.0	691.4
$p\{win\}\{1,2,3\}\{s\}\{temp\}$	$\beta\{21,31\vert 32\}\{s\}$	43	11227.1	768.5
$p\{win\}\{1,2,3\}\{s\}$	$\beta\{21,31\vert 32\}\{s\}$	35	11679.7	1221.1

B) *Hyla versicolor*

Event Model				
Correct Classification	Misclassification	np	AIC	ΔAIC
$p\{cos\}\{1,2,3\}\{.\}\{temp\}$	$\beta\{cos\}\{21,31,32\}\{s\}\{temp\}$	38	21762.1	0
$p\{cos\}\{1,2,3\}\{.\}\{temp\}$	$\beta\{cos\}\{21,31,32\}\{s\}$	36	21834.3	72.2
$p\{cos\}\{1,2,3\}\{.\}$	$\beta\{cos\}\{21,31,32\}\{s\}$	34	22049.9	287.8
$p\{win\}\{1,2,3\}\{s\}\{temp\}$	$\beta\{win\}\{21\vert 31,32\}\{.\}\{temp\}$	43	22692.0	929.9
$p\{win\}\{1,2,3\}\{s\}\{temp\}$	$\beta\{win\}\{21,31\vert 32\}\{.\}$	41	22791.7	1029.6
$p\{win\}\{1,2,3\}\{.\}$	$\beta\{win\}\{21,31,32\}\{.\}$	25	22960.3	1198.2
$p\{win\}\{1,23\}\{s\}\{temp\}$	$\beta\{21\vert 31,32\}\{.\}$	33	23768.7	2006.6
$p\{win\}\{1,23\}\{s\}$	$\beta\{21\vert 31,32\}\{.\}$	34	23958.7	2196.6

Figure 8.2 plots the day-to-day changes in the estimated correct classification and misclassification probabilities over the final study year (2006) for both species based on the model with the greatest support. For both species correct detection and misclassification probabilities are virtually zero in the first survey window. For *Hyla versicolor* both correct

and misclassification probabilities peak in the third survey window (Figure 8.2a and 8.2b). For *Bufo fowleri* correct detection probabilities peaked between the second and third survey windows (Figure 8.2c), while misclassification peaked in third survey window (Figure 8.2d). For both species misclassification peaked later in the year than correct classification probabilities.

The timing of the peak of correct classification curves also matched estimates by Weir et al. (2005), however the probability of correctly identifying higher latent abundance indices (i.e. $p^2_{i,st}$ and $p^3_{i,st}$) was higher noticeably higher that the overall detection probabilities in their site-occupancy approach.

The effects of changes in ambient temperature, on the correct and misclassification probabilities can also be seen (Figure 8.3). Because the top models for both *Hyla versicolor* and *Bufo fowleri* include a quadratic temperature effect we can look at the optimal survey temperature for each species by Julian date (Figure 8.3). For the 10 regions analyzed the optimal survey temperature for observing *Hyla versicolor* is very close to the expected temperature, while for *Bufo fowleri* the optimal temperature appears to be much higher.

Finally, as a demonstration of our proposed goodness-of-fit procedure (Section 8.3) we assessed the support for the top *Hyla versicolor* using a graphical approach to look at changes in the underlying abundance states between seasons. In total there are $\binom{S}{2}(K+1)$ possible pairwise comparisons that can be investigated. We selected the 16 possible changes in maximum index frequency between the first to second season (Figure 8.4). The change frequency observed in the data is indicated with an arrow against the empirical distribution provided by the bootstrap sample.

Of the 16 possible transitions between seasons one and two, three potential transitions appear to be of concern. The frequency of the observed $0 \rightarrow 0$ change is much higher than predicted by the model. Also of potential concern are the $0 \rightarrow 3$ and $3 \rightarrow 0$ changes, where the observed frequency appears to be much lower than expected.

8.5 Discussion

While event probabilities can be viewed as nuisance parameters, a better understanding the factors affecting calling events may help in the understanding of calling phenologies, an important component of conservation and management efforts (Hocking et al., 2008).

Knowledge on the optimal environmental conditions for maximizing correct detection probabilities can help tailor survey efforts for particular species of interest or help to best direct survey efforts for multi-species monitoring. A better understanding of calling event timing can also lend potentially lend insight into calling and reproductive phenologies, but caution is needed as there is not always a direct association between calling and reproduction (Donnelly and Guyer, 1994).

While we found strong support for modeling seasonality using a sinusoidal day-of-year approach in both *Hyla versicolor* and *Bufo fowleri*, this may not be appropriate for all anuran species. Some species (e.g. tropical) may have more prolonged reproductive periods with periods of calling bouts followed by periods of relative quiescence (Green, 1990; Donnelly and Guyer, 1994; McCauley et al., 2000). If quiescence is not associated with known covariates (e.g. temperature), then the sinusoidal structure will be inappropriate. Furthermore, the sinusoidal day-of-year approach suggests that calling events smoothly increase and decrease over time. For species that exhibit explosive calling or single sustained calling bouts, it is possible that true calling pattern may be abrupt and may not as smooth as suggested by the sinusoidal function.

For both species, models that included a temperature covariate in both correct and misclassification probabilities were also strongly supported. This is not surprising as temperature been found to be an important factor in the timing of both anuran breeding (Beebee, 1995; Reading, 1998; Blaustein et al., 2001; Gibbs and Breisch, 2001) and anuran calling (Tryjanowski et al., 2003; Blaustein et al., 2001; Weir et al., 2005; Royle and Link, 2005). While we noted this effect on event probabilities, temperature has only been found to be important to the timing of *Hyla versicolor* breeding (Gibbs and Breisch, 2001), but not *Bufo fowleri* (Blaustein et al., 2001).

Excluding tropical species (which tend to have prolonged breeding periods that may not be as tightly associated with temperature) there is conflicting evidence over the importance of temperature on anuran breeding timing within the temperate-zone anuran literature, as well as general trends towards earlier breeding due to global warming (e.g. Beebee, 1995; Reading, 1998; Blaustein et al., 2001). Interestingly, a direct association between breeding date and temperature was not found in *Bufo fowleri* (Blaustein et al., 2001), but an indirect association was found for *Hyla versicolor* (Gibbs and Breisch, 2001).

It is hard not to re-evaluate this conflict within the context of detectability. To date most, of these studies, use a “first day of breeding” metric to measure the commencement of reproduction, but none have considered the incorporation of missed classification. It is completely plausible for short term environmental conditions to affect when the first day of breeding is observed, without actually affecting the true biological breeding period. In this case the noise added to the observation process may make it more difficult to detect subtle shifts of global climate change. Instead, approaches that attempt to isolate the true biological breeding period from other sources of noise may clear up these types of conflicts.

Thus, it may be better to measure peak reproductive timing for episodic temperate-zone breeders rather than less clearly defined concepts such as "commencement." What defines commencement is open to interpretation, it could be defined by the first observed call or breeder, or once a certain (presumably low) number of calls or breeders has been surpassed (i.e. to avoid observer mistakes). Both are set at an an arbitrary level by an investigator. In contrast by defining the timing of breeding in terms of a peak activity, there is a clear defined metric that is also biologically important, the only requirement is that activity can be modeled using a smooth change point function, as we have done in the current study.

For temperate-zone breeders if we are willing to concede that calling rates and breeding activity are related, then derived parameters such as $t_{\max}$, from the sinusoidal regression framework, may provide a plausible metric. If we define peak breeding by peak calling periods, then peak detection rates (which are affected by calling) will also indicate peak breeding. In addition, the estimate will also incorporate other factors (e.g. temperature) as affecting detection rates and should therefore be less sensitive to short term environmental changes.

While there is clear support for the inclusion of temperature effects in all event probabilities, future studies may want re-examining the manner in which temperature is modeled. Generally, the NAAMP protocol allows for sampling across a five month period, a time interval in which the average temperature can be expected to change in temperate-zone areas such as the ones in the study. As such, the day-of-year effect (defined by sinusoidal function) may be confounded with the temperature effect if the timing of anuran breeding is triggered by changes in temperature or a temperature threshold, as has been suggested for some anuran species (e.g. Reading, 1998).

If this is the case temperature could be instrumental in the timing of breeding and by extension the timing of calling and calling event probabilities. While temperature may effect the phenology of breeding and calling, temperature could also play an additional role in the intensity (and calling event) heard on any given sampling occasion. In order to get at this latter temperature effect, while decoupling the temperature and timing effect, future studies may wish to consider centering the temperature covariate on the expected temperature for a given Julian day. This new temperature covariate would become a way to define unseasonably higher or lower temperatures.

Finally, in specifying the model we assumed all 10 states from which the experiment was conducted possessed the same initial occupancy proportion $\mathbf{\Phi}_{i,0} = \mathbf{\Phi}_0$. However, if we look at the distribution of the maximum observed indices by region (Figures C and C for *Hyla versicolor* and *Bufo fowleri* respectively) it is evident that different regions likely posses different initial occupancy state distributions. For example, for *Hyla versicolor* states such as Delaware Virginia, West Virginia appear to have a higher proportion of zero sites than states such as Maine, Massachusetts and New Hampshire. Potentially, this represents a model misspecification which may be produce parameter estimates that under estimate the proportion of zero sites in these areas. If so, then the observed data would have a higher $0 \rightarrow 0$ transition, because there are more zero sites than predicted.

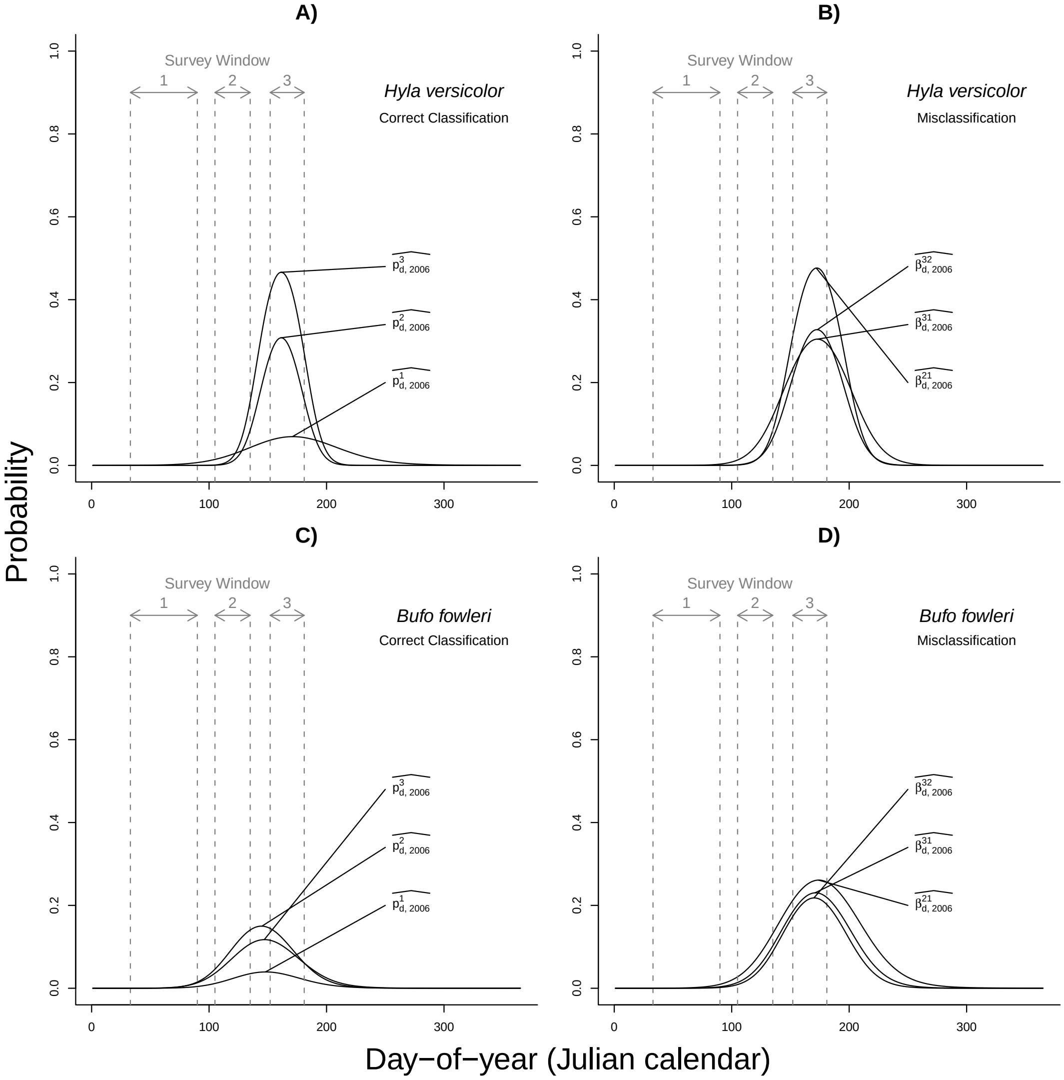


Figure 8.2: Estimated seasonal correct classification and misclassification probability curves for *Hyla versicolor* and *Bufo fowleri* in 2006, using the highest ranked day-of-year sinusoidal regression model (Table 8.7). Correct and misclassification probabilities has been also been adjusted for temperature, by using the expected mean temperature for a given Julian date. The current NAAMP suggested survey windows are indicated with dashed lines. In all cases the correct detection misclassification probabilities has been also been adjusted for temperature, by using the expected mean temperature for a given Julian date and NAAMP survey windows are indicated with dashed lines.

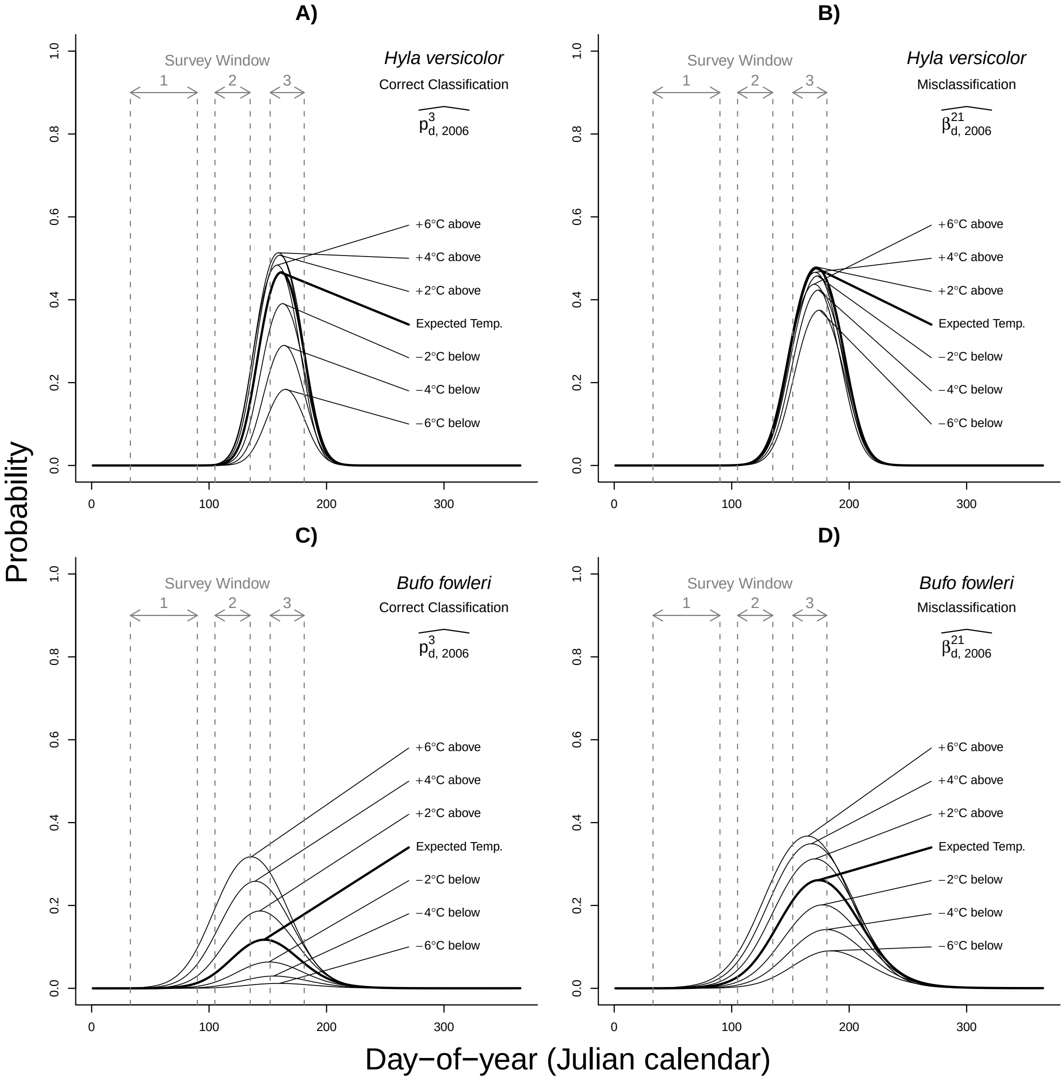


Figure 8.3: The effect of ambient temperatures, above and below the expected mean temperature by Julian date, on the estimated probability of correct classification $\widehat{p^k_{d,2006}}$ conditional misclassification $\widehat{\beta^{kj}_{d,2006}}$ for *Hyla versicolor* and *Bufo fowleri* in 2006, using the highest ranked day-of-year sinusoidal regression model (Table 8.7). Estimated probability for the expected ambient temperature by Julian date is displayed with a thicker line weight, while ambient temperatures above and below are displayed with a thinner line weight. NAAMP survey windows are indicated with dashed lines.

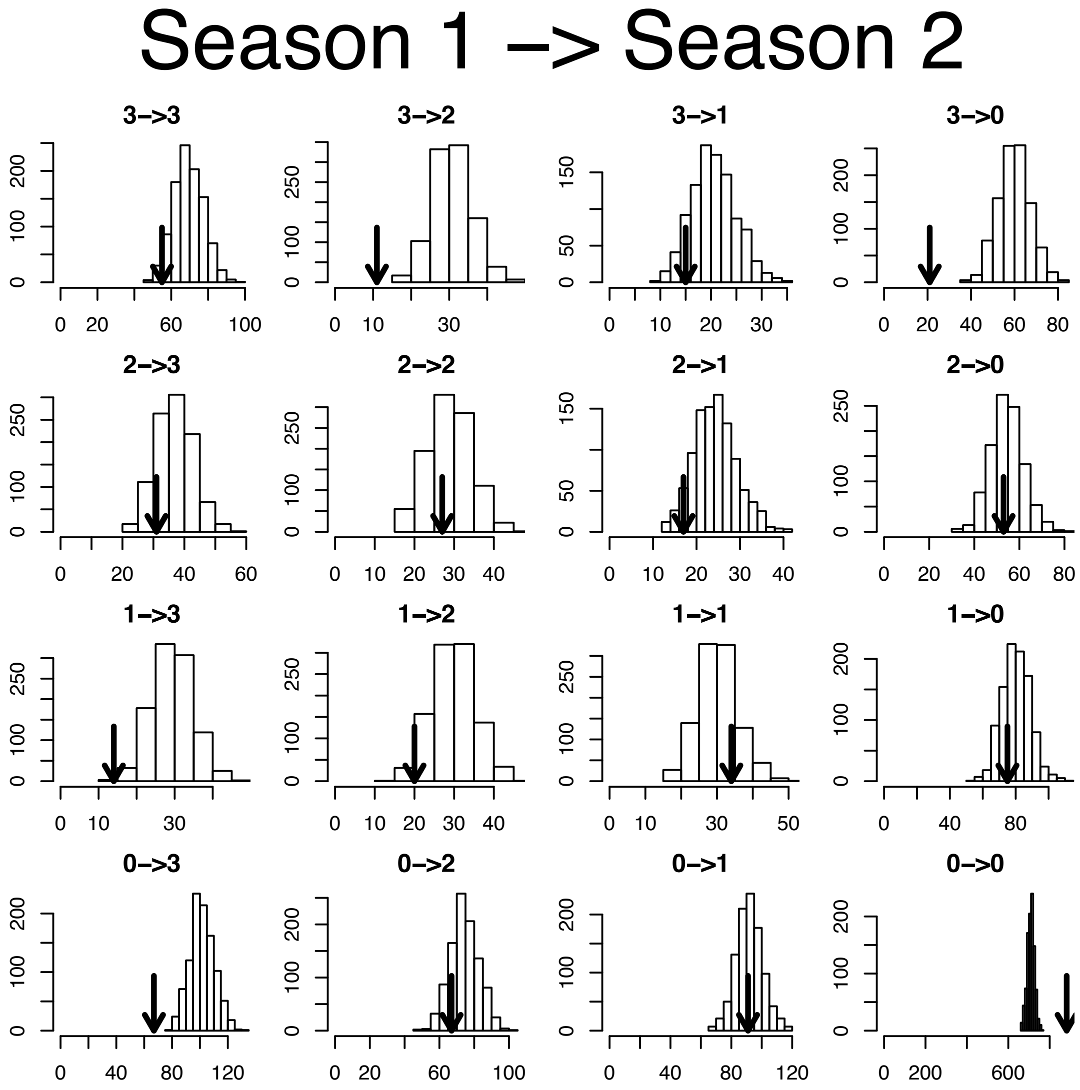


Figure 8.4: Example lack-of-fit transition graph for the top *Hyla versicolor* model for possible changes maximum indices observed from season one to season two. Arrows indicate the frequency observed in the data, while the histogram shows the empirical distribution generated from the parametric bootstrap.

Chapter 9

Power Analysis: Contrasting Permanent and Temporary Monitoring Protocols

Our proposed multi-state occupancy formulation provides for two types of sampling protocols: permanent and temporary sampling protocols. Both are intended to handle different types of monitoring situations, but will also differ in their power to detect changes in underlying occupancy state proportions. The differences in power between these two sampling protocols is unknown. Furthermore, there are other design considerations, such as whether to partition total survey effort into sampling additional sites or number of repeat samples. As a guide, we have conducted a power analysis to investigate the trade off between number of repeat samples and total sampling effort using the abundance class formulation proposed by (Royle and Link, 2005) for both sampling protocols using real world event probabilities.

The abundance class formulation (see Section 7.4) represents a potentially useful formulation for long term monitoring studies, making it a prime candidate for further investigation. As the basis of our power analysis we wish to consider designs that minimize effort for potential experimenters. For the sake of the analysis, we will assume that each survey, whether a repeat sample of an existing site, or the inclusion of a new site as having equal cost. We look to minimize the total survey effort, that is number of sites multiplied by number of repeat samples that still provides the same power to detect changes.

For the purposes of this analysis we will be using estimated event probabilities from both studies on both anuran (Royle and Link, 2005) and rockfish (Marliave and Challenger, 2009). While the two studies have a different interpretations of what constitute a latent abundance class represents, they do provide a useful starting point for the types real-world parameter values that may be encountered. Both are single year studies, and as a possible extension to these studies we consider the ability to detect changes in the underlying abundance class distribution between the first and second season in a two season study (here we equate years to seasons).

Power is interpreted as the ability to detect changes in univariate summaries statistics suggested by Royle and Link (2005); the average abundance class $\bar{N}$, and probability of occupancy $1 - \Psi_0$.

9.1 Deriving Univariate Summary Statistics for Multiple Seasons

For the latent abundance class formulation Royle and Link (2005) purposed two univariate summary statistics, the average latent abundance distribution, $\bar{N}$ and the probability of occupancy $1 - \Psi_0$. Within a multi-season context the average latent abundance class for season s is defined as,

$$\bar{N}_s = E[Z] = \sum_{k=0}^{K} k\Psi_{k,s},$$

where $\Psi_{k,s}$ is the expected proportion of sites with latent abundance class k in season s.

Because we are interested in detecting changes in the latent abundance class distribution between years we can define

$$\Delta\bar{N}_{s\to v} = E[\bar{N}_v - \bar{N}_s] = \bar{N}_v - \bar{N}_s \tag{9.1}$$

as the expected difference between the average latent abundance class of seasons s and seasons v, where $s < v$. Similarly, managers may be interested in tracking changes in the proportion of sites occupied between any given season. Since the expected proportion of sites occupied for season s is $1 - \Psi_{0,s}$, we can define

$$\Delta(1 - \Psi_0)_{s\to v} = \Psi_{0,s} - \Psi_{0,v} \tag{9.2}$$

as the change in expected proportion of sites occupied between season s and season v for $s < v$.

In both cases the functional invariance property of maximum likelihood estimators allows estimates of the summary measure be obtained by by substituting in the appropriate MLE's. For the temporary sampling protocol the $\Psi_{0,s}$ and $\Psi_{0,v}$ will be estimated directly, however for the permanent sampling protocol this will be a derived parameter. Let be $\boldsymbol{\Psi}_s$ be a row vector such that $\boldsymbol{\Psi}_s = [\Psi_{0,s}, \Psi_{1,s}, \ldots, \Psi_{K,s}]$, then

$$E\left[\boldsymbol{\Psi}_s\right] = \boldsymbol{\Phi}_0 \prod_{j=1}^{s-1} \boldsymbol{\Phi}_j \qquad \text{for} \quad s = 2, 3, \ldots, S, \tag{9.3}$$

where $\boldsymbol{\Phi}_0$ is the initial latent abundance state distribution and $\boldsymbol{\Phi}_j$ are the transition probabilities from season j to season $j+1$. Estimates are again obtained by substituting the appropriate MLE's in (9.3). All sites are also assumed to share the same transition probabilities.

9.2 Methods

As a test scenario we consider in a two season study where the study area experiences 30% drop in the average latent abundance $\bar{N}$ between seasons (Figure 9.1). In the first season the initial latent abundance distribution was $\boldsymbol{\Phi}_0 = [0.1, 0.2, 0.3, 0.4]$, $\bar{N}_1 = 2$, decreased to $\bar{N}_2 = 1.399846$ in the second season with a latent abundance distribution of $\boldsymbol{\Psi}_2 = [0.3569, 0.159575, 0.2103037, 0.2732213]$.

This scenario was based on the λ-γ-ν transition structure (see Section 8.2.2), where the probability of a decrease in latent abundance was $\lambda_1 = 0.35$, with a conditional probability of increase being $\gamma_1 = 0.035$, and a continuation ratio of $\nu_{1,1} = \nu_{2,1} = 0.15$. Note that for the temporary monitoring design the state distribution $\boldsymbol{\Psi}_s$ will be estimated directly each year. Overall, this scenario results in a $\Delta\bar{N}_{1\rightarrow 2} = -0.6001537$ and an occupancy change of $1 - \Psi_{0,1} = 0.9$ in the first season to $1 - \Psi_{0,2} = 0.6431$ in the second season, for an overall change of $\Delta(1 - \Psi_0)_{1\rightarrow 2} = -0.2569$, a 28.5% decrease in the probability of occupancy.

We considered designs that used $T_s = 2, 3, 4, 5, 6$ surveys each season, but with a fixed number of total surveys each season. The number of sites surveyed is a then function of the protocol, the number of repeat surveys, and the total survey effort. For the permanent sampling protocol, the total number of sites surveyed is determined by taking the number

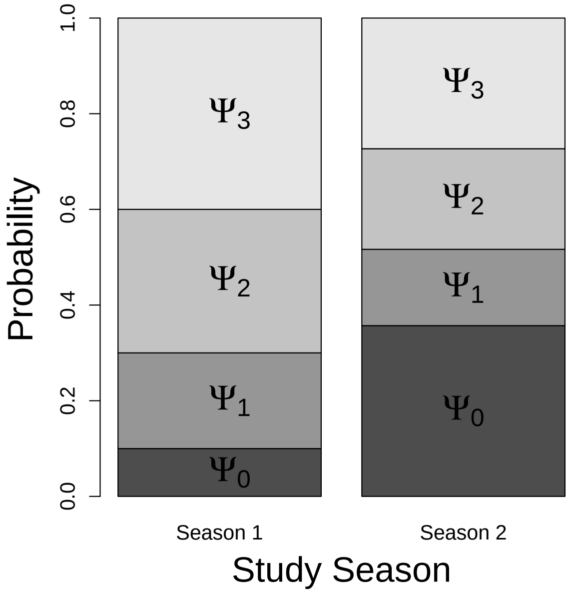


Figure 9.1: Change in the latent abundance distribution from the first to second season for both test scenarios involving anuran and rockfish event probabilities. Overall the change results in a 30% decrease in the average latent abundance and a 28.5% decrease in occupancy.

of total surveys performed in any season dividing it by the number T_s. For the temporary protocol this was the number of sites surveyed each season. In the case where this produced a non-integer result, the number of sites was rounded to the nearest integer and the total sampling effort was adjusted accordingly.

For each experimental setting 2500 runs were completed based on either anuran (Table 9.1a) or rockfish event probabilities (Table 9.1b) and the corresponding statistical hypotheses tested were,

$$H_0 : \Delta \bar{N}_{1\to 2} = 0$$
$$H_1 : \Delta \bar{N}_{1\to 2} \neq 0,$$

for the change in average latent abundance class and

$$H_0 : \Delta(1 - \Psi_0)_{1\to 2} = 0$$
$$H_1 : \Delta(1 - \Psi_0)_{1\to 2} \neq 0,$$

and for change in occupancy. The number of times the null hypothesis was rejected at the

$\alpha = 0.05$ level was used to calculate power. For computational efficiency, hypotheses were tested using Wald-type test statistic with variance estimates derived from the Delta method.

Table 9.1: Estimated event probabilities from the best supported model for A) North American Amphibian Monitoring Program (Royle and Link, 2005) and B) the Vancouver Aquarium rockfish dive surveys (Marliave and Challenger, 2009).

A) NAAMP Call Data (2001)

	$P(U=0\|Z)$	$P(U=1\|Z)$	$P(U=2\|Z)$	$P(U=3\|Z)$
$Z=0$	1.000	0	0	0
$Z=1$	0.758	0.242	0	0
$Z=2$	0.772	0.149	0.078	0
$Z=3$	0.647	0.125	0.149	0.078

B) Aquarium Rockfish Dive Surveys (2006)

	$P(U=0\|Z)$	$P(U=1\|Z)$	$P(U=2\|Z)$	$P(U=3\|Z)$
$Z=0$	1.000	0	0	0
$Z=1$	0.690	0.310	0	0
$Z=2$	0.310	0.380	0.310	0
$Z=3$	0.018	0.093	0.579	0.310

9.3 Results

Generally, the permanent sampling protocol produced more powerful, in some cases requiring less than half the total effort to detect changes in the average latent abundance or occupancy. This difference in efficiency was similar for both univariate summaries. However, the optimal design differed greatly depending on the event probabilities used.

For anuran event probabilities the most powerful design involved $T_s = 6$ repeat samples per season, which held for either sampling protocol and for both univariate statistics (Figure 9.2). For $\Delta\bar{N}_{1\to2}$ 80% power was achieved with 390 and 690 total surveys per season for permanent and temporary sampling protocols, while for $\Delta(1-\Psi_0)_{1\to2}$ 350 and 540 total surveys were required respectively. In this case $\Delta(1-\Psi_0)_{1\to2}$ produced slightly more power full designs and the temporary protocol requires approximately 1.5-1.7 times as much effort as the permanent sampling protocol. It should be noted that higher $T_s > 6$ designs may be more powerful, but were not tested.

For rockfish event probabilities however, $T_s = 2$ tended to produce more powerful designs for both sampling protocols and univariate summary statistics. The only exception was for detecting changes in $\Delta(1 - \Psi_0)_{1\to2}$ with the temporary sampling protocol, where $T_s = 3$ appears to produce equivalent or slightly more powerful designs that $T_s = 2$ repeated samples per season. For $\Delta\bar{N}_{1\to2}$ 80% power was achieved with 75 and 150 total surveys per season for the permanent and temporary protocols respectively, while for $\Delta(1-\Psi_0)_{1\to2}$ 80% power was achieved with 75 and 210 total surveys per season respectively. For the permanent sampling protocol both univariate sampling statistics had comparable power, while for the temporary design $\Delta\bar{N}_{1\to2}$ was slightly more efficient.

In terms of sampling efficiency in all cases the permanent sampling protocol produced designs that were 1.5 to 2.8 times more efficient at the 80% power level in terms of total surveys performed (Figure 9.4). For all the tested scenarios the univariate summary statistics performed fairly comparably, with the notable exception that for the anuran event probabilities combined with the temporary sampling protocol, the $\Delta\bar{N}_{1\to2}$ univariate summary required about 1.3 times as much effort for 80% power as $\Delta(1 - \Psi_0)_{1\to2}$.

9.4 Discussion and Conclusions

While the optimal number of repeat surveys design differed depending on whether rockfish or anuran event probabilities were used, the permanent sampling design showed a strong advantage over the temporary sampling designing in all the scenarios investigated. Within the scenarios tested a permanent sampling approach showed anywhere from a 1.5 to 2.8 times higher efficiency in terms of total sampling effort required to reach a power of 80% to detect change. As such, practitioners should focus on implementing a permanent monitoring scheme whenever possible. However, for situations where this is not possible, practitioners should be prepared to adjust their total survey effort accordingly.

Interestingly, anuran and rockfish event probabilities produced a fairly large divergence in the optimal repeated sampling effort. For a given total effort (surveys per season) the optimal design for both permanent and temporary protocols used $T_s = 6$ and $T_s = 2$ repeated samples per season for anuran and rockfish event probabilities respectively. This equates to a three times difference in the number of sites sampled for the same given total effort. Note that, $T_s = 6$ was the highest number of repeat samples tested and in the case of the anuran analysis a higher value of T_s may have produced a slightly more powerful design.

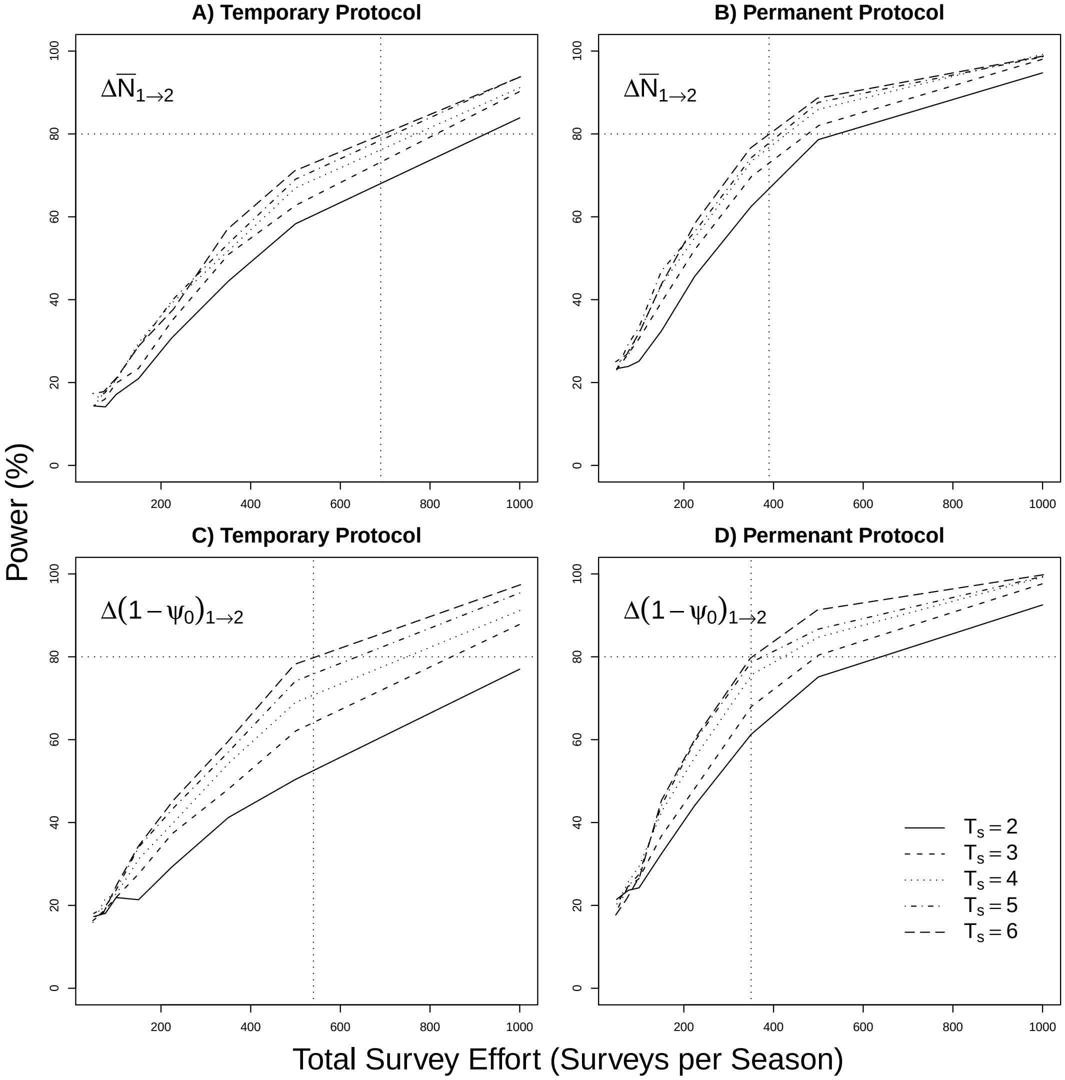


Figure 9.2: Power curves for the anuran event probabilities (Table 9.1) for designs using $T_s = 2, 3, 4, 5, 6$ repeat surveys per season in a two season experiment. Power was judged as the ability at the $\alpha = 0.05$ level to detect changes in average latent abundance $\Delta\bar{N}_s$ using A) temporary and B) permanent sampling protocols as well as to detect changes in occupancy $\Delta(1 - \Psi_0)_s$ using C) temporary and D) permanent sampling protocols. Horizontal dotted line indicates 80% power to detect a change and vertical dotted line indicates the total survey effort for the most powerful design tested. Total survey effort is the number of sites multiplied by the number of repeat surveys per season.

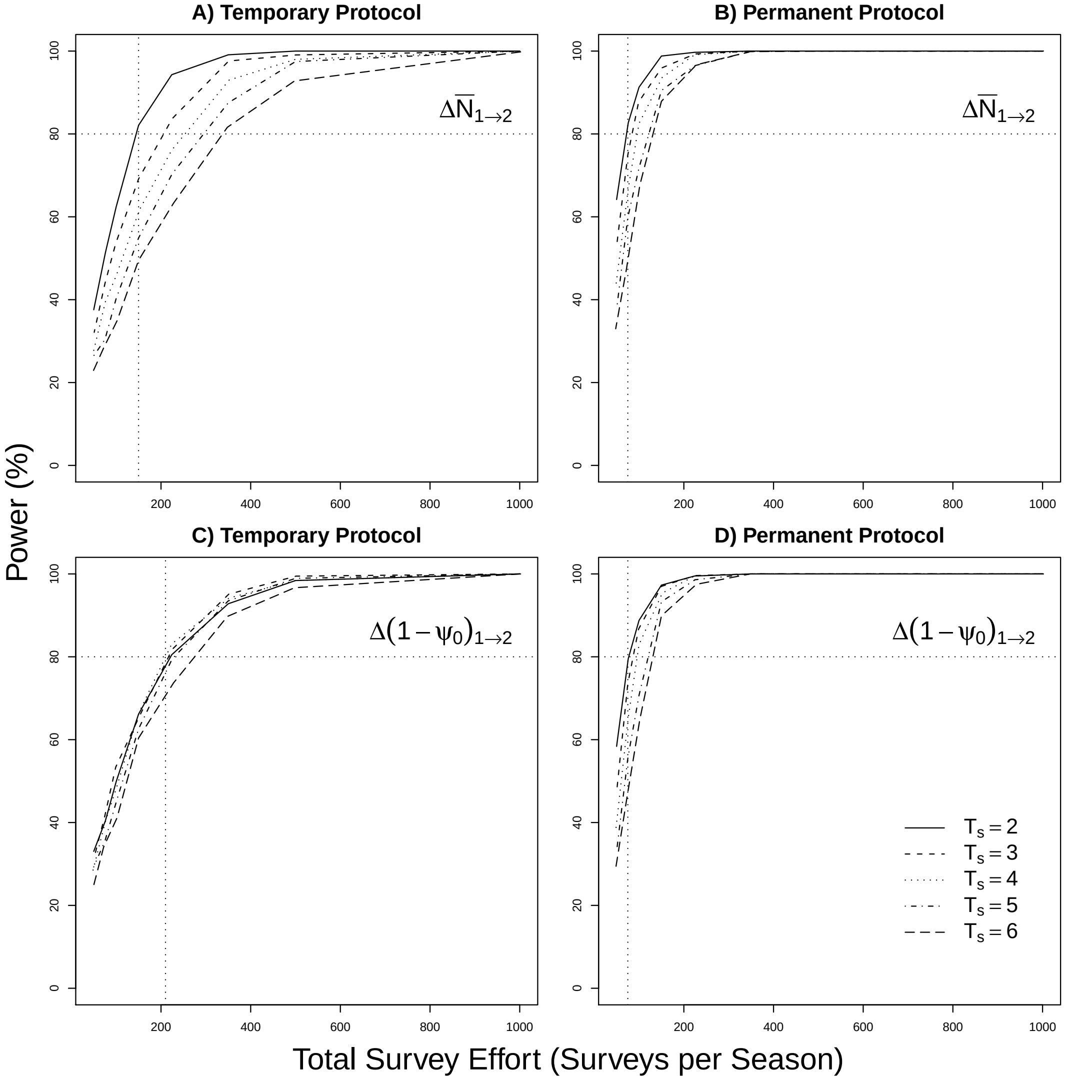


Figure 9.3: Power curves for the rockfish event probabilities (Table 9.1) for designs using $T_s = 2, 3, 4, 5, 6$ repeat surveys per season in a two season experiment. Power was judged as the ability at the $\alpha = 0.05$ level to detect changes in average latent abundance $\Delta\bar{N}_s$ using A) temporary and B) permanent sampling protocols as well as to detect changes in occupancy $\Delta(1 - \Psi_0)_s$ using C) temporary and D) permanent sampling protocols. Horizontal dotted line indicates 80% power to detect a change and vertical dotted line indicates the total survey effort for the most powerful design tested. Total survey effort is the number of sites multiplied by the number of repeat surveys per season.

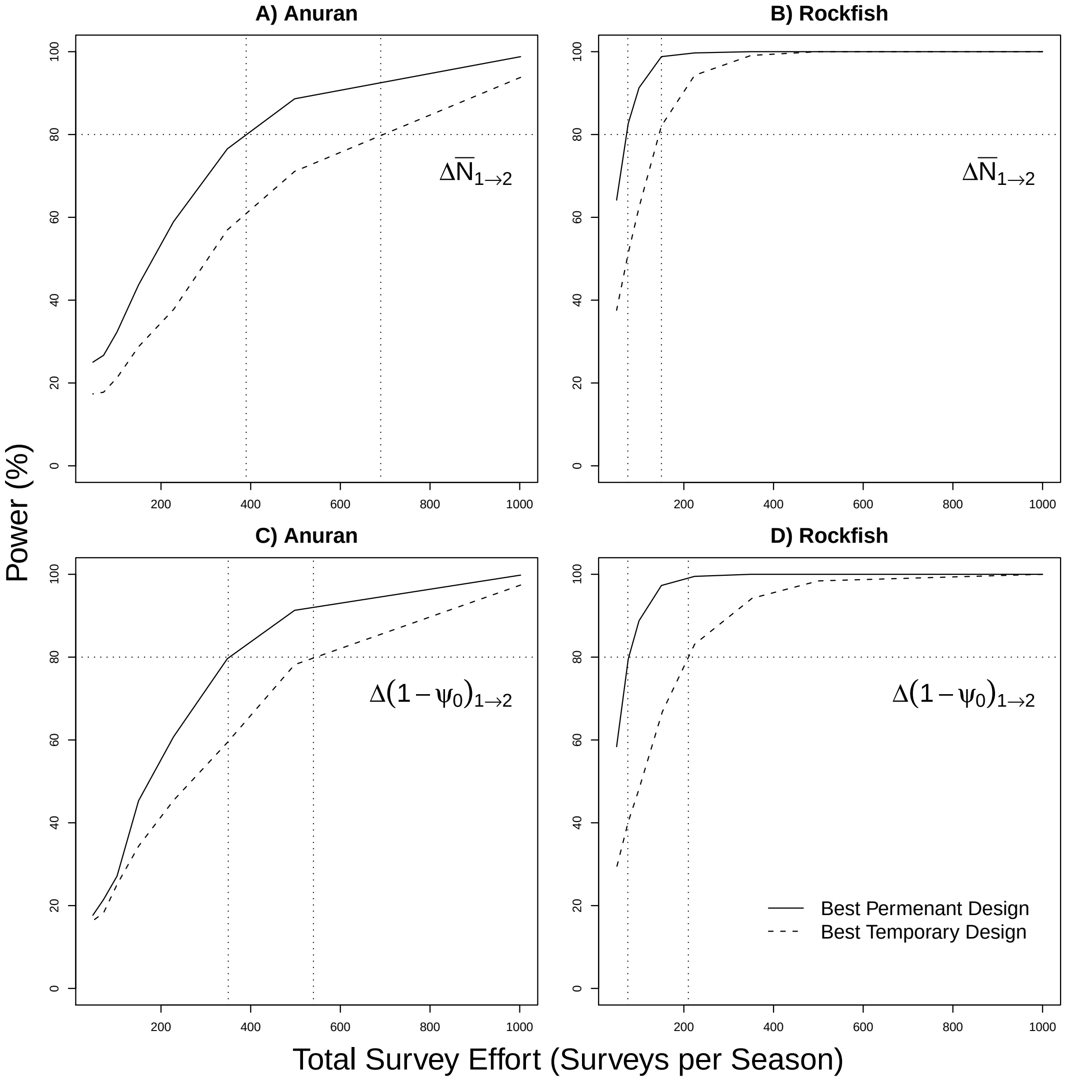


Figure 9.4: Power curves for the ability of permanent (solid line) and temporary (dashed line) sampling protocols to detect differences in $\Delta\bar{N}_{1\rightarrow 2}$ using A) anuran and B) rockfish event probabilities as well for differences in $\Delta(1-\Psi_0)_s$ using C) anuran and D) rockfish event probabilities. The most powerful design for each protocol is displayed. Horizontal dotted line indicates a power of 80% to detect a change, while horizontal lines indicate the total survey effort required.

Also note, that designs of $T_s = 1$ would require some sort of additional modeling structure, such as covariates, in order to make the model estimable.

For the site-occupancy formulation detection probability is known to be an important factor for determining the optimal number of repeat samples (MacKenzie and Royle, 2005). In this case the relationship is fairly straightforward due to the structural simplicity of the model. Generally, low levels of detection result in higher repeat samples being optimal. The abundance class multi-state occupancy formulation likely also follows this rule of thumb, however it is harder to critically assess due to the complexity of event probabilities. The anuran event probabilities generally showed lower correct detection probabilities, and a higher probability of observing a zero compared to the rockfish event probabilities. It was not surprising that the optimal number of repeat samples was also higher. However, it would be advisable for practitioners to conduct their own power analysis before implementing designs, as optimal design may also depend on univariate summary being used, the type of changes in the latent state distribution and the overall event probability structure. We only considered a very narrow range of possible configurations.

Of particular interest may be further comparisons between univariate statistics and simpler models such as the site-occupancy formulation. In our analysis we chose a scenario where the percent changes in average latent abundance $\bar{N}$ and occupancy $1 - \Psi_0$ where roughly equivalent (30 and 28.5% declines respectively) and it is not a surprise that the power of these two univariate summaries were comparable. However, it is unclear how both of these univariate summaries would perform under different changes in the latent state distribution. For example, a general decline in the average latent abundance, without changes in the underlying proportion of sites occupied would likely favour the use of $\Delta\bar{N}$. As such, it would be highly recommend that practitioners wishing to monitor changes in wildlife population using the abundance class formulation should consider using both univariate summaries simultaneously.

The power analysis also considered a very specific transition structure for changes in occupancy states over time. The transition formulation used for the permanent monitoring protocol was relatively sparse compared to a more general transition formulation (i.e. 3 versus 12 free parameters). This was done purposefully to ensure the power results for the permanent and temporary formulations were as directly comparable as possible. For each formulation there were three free parameters used to model either the transition probabilities between season one and two or the latent abundance class distribution in the second season.

Differing numbers of free parameters could affect the power calculations due to the extra variance introduced by having to estimate additional parameters and we may expect the performance difference between protocols to lessen. In addition, it is also unclear what the potential effect of misspecifying transition probabilities will have on the ability to detect changes in the latent state distribution.

A potential benefit of the temporary sampling protocol is that it makes no assertions on the underlying transition structure between seasons. As such, it makes fewer assumptions and may be preferable for situations where there is little scientific significance in estimating transition probabilities. If this is true, but permanent sampling protocol is more cost efficient (e.g. adding new sites is more costly than sampling existing sites), then practitioners should use a completely free state transition structure accepting a potential loss of power. Further investigation into the power differences between a the permanent protocol with a free transition matrix and the temporary protocol is needed.

Investigators may also may wish to explore the differences in power between between a simpler site-occupancy model and the more complex multi-state models such as the abundance class models. Generally, it is believed that more complex models, such as the abundance class model, make better approximate of reality as they discard less information. For example, the abundance class model will better describe possible heterogeneity in detection probabilities, caused by larger populations being more easier to detect (Royle and Nichols, 2003). The logic then carries that these models are therefore more applicable to monitoring applications. However, to date these types of assertions have not been tested and it is possible that more complex models may be no better at detecting basic changes in ecological states, than simpler models that make fewer overall assumptions.

Such assumptions may be investigated using a technique called Model Structural Adequacy (MSA) analysis (Taper et al., 2008). MSA analysis is a simulation-based methodology than assesses the ability of statistical models to answer scientific questions. Candidate models are repeatedly fit to data simulated by a complex process model (CPM; a proxy for reality) which contains more structure than can be described by any of the candidate models and is based on available expert opinions. The ability of candidate models to answer scientific questions (e.g. changes in the underlying abundance state distribution) are then assessed allowing for the affects of structural errors on scientific outcomes to be approached. While highly dependent on the formulation of the CPM, this methodology would allow for the investigation of suspected factors. For example, detection heterogeneity is a well documented problem in both mark-recapture and occupancy models (Huggins, 1989; Coull and Agresti, 1999; Pledger, 2000; Royle and Nichols, 2003). Various forms of detection heterogeneity could be included (e.g. population dependent) in the CPM as well other potential violations such as different functional relationships between true population size and the underlying ecology state. Under these scenarios, the ability of site-occupancy and abundance class models to detect changes in the true population could be assessed. This would allow practitioners to choose the most powerful and robust framework, for a given scientific goal.

Chapter 10

Monitoring and Evaluating Rockfish Conservation Areas in British Columbia[1]

The use of abundance indices can have broad application within ecological studies. Off the coast of British Columbia, Canada there are about 35 species of rockfish, all of which have been experiencing declines in recent years, especially within the Howe Sound. Due to the tendency of rockfish to be long live with a relatively long 22-year generation time (Yamanaka et al., 2004), Fisheries and Oceans Canada initiated a Rockfish Conservation Areas (RCA) as part of the Rockfish Conservation Strategy to address concerns the decline observed in the Howe Sound. To assess the efficacy of the RCA designation the Vancouver Aquarium performed rockfish point counts at various locations within and outside the designated RCAs. However it was not clear how to best analyze the data.

The choice of point counts provided implementation benefits, but also brought analysis drawbacks. Because of the conservational status of rockfish, surveying can be difficult. Point counts were ideal because sampling is not invasive and can be performed easily at a low cost. However from a statistical stand point counts also suffer from a number of issues, such as detectability and double counts (individuals are not uniquely identified).

As part of their natural behaviour most rockfish are cavity hiders and may take shelter or hide in deep rock piles depending on the conditions and time of year. Although the

---

[1] Published as Marliave and Challenger (2009).

underlying abundance of rockfish in an area may not change substantially between surveys (the species is long lived) there can be substaintial changes in the number observed or counted due to detectability. Furthermore, under certain circumstances the available density of rockfish may be higher that can be enumerated by divers, leading to either under counts or potentially double counts, depending how individual divers react to the situation.

That said, point counts do provide some useful information about the underlying latent abundance level. In barren areas without fish, point counts will always produce zero. As abundance increases we can expect the average count to also increase. This relationship should hold assuming factors affecting detection can largely be treated as random; we generally believe this to be the case. Furthermore, we may choose group point counts into classes. This will help reduce observation repeatability, yet should still retain most of the biological meaning. For example, from a biological perspective it may not be important to distinguish between 100 or 150 individuals, but instead it may be more crucial to distinguish the difference between 100, 1000 or 10000 individuals. As such we developed, in the accompanying paper, a relative abundance index that uses orders of magnitude differences in point counts.

Finally, this paper also provides a good example of how to use a conditional odds to model the latent abundance distribution. We modeled the underlying abundance class distribution as a linear combination of site specific factors such as boulder coverage, whether or not the site belonged to an RCA or control area, and accounted for potential regional area effects within the Howe Sound. Below is a copy of the manuscript published in the Canadian Journal of Fisheries and Aquatic Sciences, Volume 66 in 2009 (pages 995-1006).

10.1 Introduction

Howe Sound, one of North America's southernmost fjords, is located northwest of Vancouver and extends from West Vancouver north to Squamish, adjoining the Strait of Georgia at its southern entrances (Fig. 10.1), and the area incorporates many islands and clusters of islands. Most of the shoreline in Howe Sound is steep rock or mud slopes.

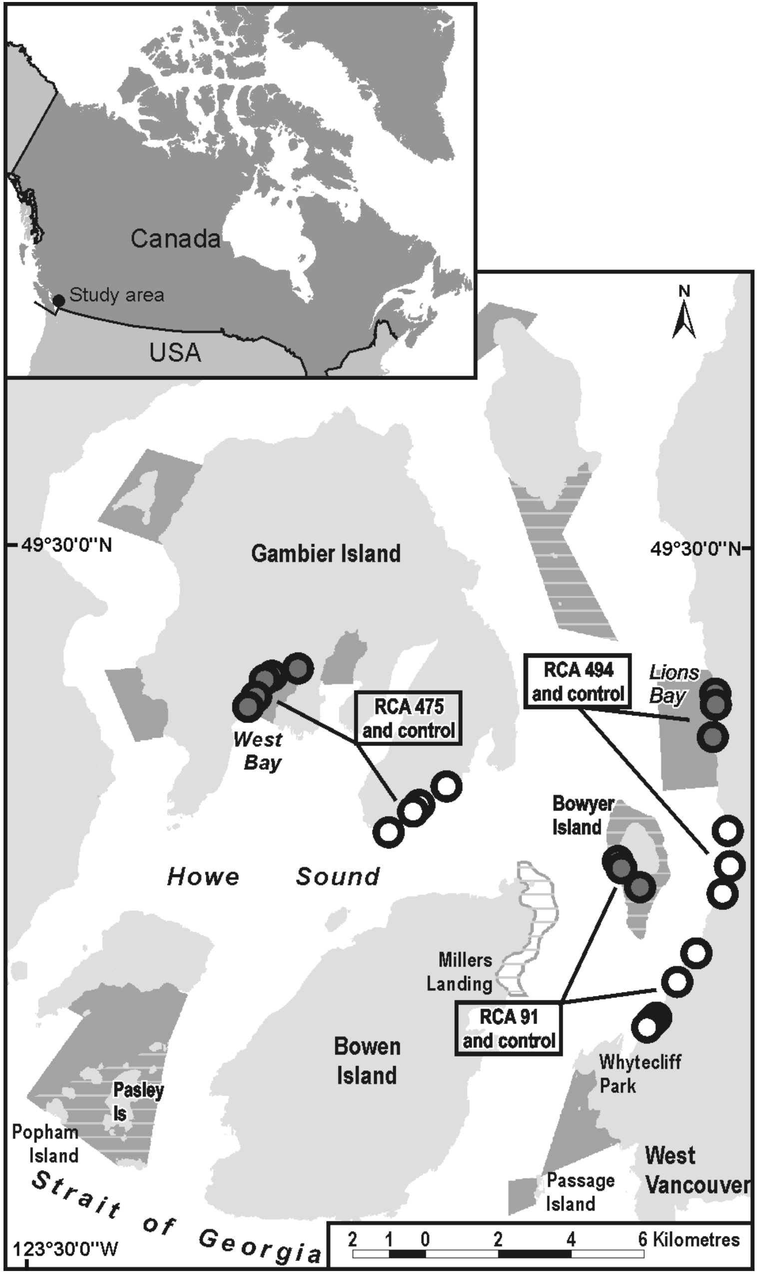
About 35 species of rockfish live off the coast of British Columbia (BC) (Hart, 1973; Love and Thorsteinson, 2002). Howe Sound was one of the first regions in BC to exhibit serial depletion of rockfishes on nearshore reefs after the live fishery for rockfish and lingcod was initiated in the late 1970s (Love and Thorsteinson, 2002). Research and monitoring indicate that the populations of five inshore rockfish species (copper – *Sebastes caurinus*, quillback – *S. maliger*, China – *S. nebulosus*, tiger – *S. nigrocinctus* and yelloweye – *S. ruberrimus*) are now at low levels, and a sixth, *S. melanops* (black rockfish), is depleted in the vicinity of Vancouver. Although scientific estimates of inshore rockfish abundance are not available for Howe Sound, recent research estimates that lingcod biomass (*Ophiodon elongatus*, a species of groundfish that is ecologically similar to rockfish) in Howe Sound is less than 1% of the historic biomass (Logan et al., 2005). This is in contrast to other parts of the Strait of Georgia where lingcod biomass is generally 7-20% of historic levels. Similarly, inshore rockfish abundances in Howe Sound are likely to be relatively more depleted than elsewhere along the BC coast. Monitoring of quillback rockfish suggests that the population in the Strait of Georgia region (that encompasses Howe Sound) has been reduced by over 75% in less than one 22-year generation (Yamanaka et al., 2004).

In 2004, Fisheries and Oceans Canada (DFO) initiated a Rockfish Conservation Strategy to address concerns about inshore rockfish/lingcod fisheries along the coast of British Columbia. One component of the strategy was the designation in 2004 of Rockfish Conservation Areas (RCAs), areas closed to hook-and-line fishing and designed to halt rockfish population decline (4 RCAs were in Howe Sound). Additional RCAs were established in 2007, with one rescinded and eight added for a total of 11 RCAs in Howe Sound (Fig. 10.1). Since establishment of RCAs, Vancouver Aquarium research divers have done scuba dive census work on inshore rockfishes both inside and outside RCAs in Howe Sound.

In 2004/2005 the Vancouver Aquarium conducted a pilot study to compare adult rockfish densities inside and outside the original four Howe Sound RCAs, partly in order to enable data submission to the public process toward final RCA selections. Because of the ongoing RCA selection process, these 2004/2005 dives were focused on known sites of high rockfish abundance outside the original RCAs, and were thus biased. Some of the highest observed densities of adult rockfish remained outside those four RCAs (e.g. Passage Island reefs exhibited 110 adult quillback rockfish per hour effort), with one RCA exhibiting the lowest densities (Millers Landing, on Bowen Island, exhibited 0.3 adult quillback rockfish per hour; unpublished data). In 2007 the Millers Landing RCA was deleted, and Passage Island RCA established. Despite the RCA modifications, the pilot study suggested further evaluation of RCAs in Howe Sound was needed in order to establish a clear basis for future monitoring throughout BC of this conservation management regime.

The present study compares the adult rockfish abundances between RCA and non-RCA

Figure 10.1: Randomly selected dive sites for RCAs (gray circles) and for adjacent control areas (white circles) that are outside the DFO rockfish habitat model predicted areas. Horizontal lines indicate preliminary 2004 RCAs; final RCAs are gray (Millers Landing RCA was rescinded in 2007).



areas in Howe Sound and relates these differences to habitat substrate. Through a sequence of activities and analyses, we verified our methods, documented the substrate associations of the rockfish, and evaluated the utility of the government predictive model and public process for RCA designations (cf. Ardron and Wallace, 2005), in terms of the rocky fjord environment of Howe Sound.

We also introduce and expand on a new type of model used to quantify non-invasive measures (e.g. point counts) of population abundances (see MacKenzie et al., 2002; Royle and Link, 2005). The technique uses a relative measure of abundance (i.e. none, low, medium and high), allows for misclassifications and requires no direct interaction with the study species. We expand the model to allow for direct comparisons of abundance between areas, which we use to compare abundances between RCA and control areas as well as between areas with differing habitat substrate. As such, the model serves as an independent evaluation of the rockfish habitat model currently used by DFO and by other fisheries researchers (Ardron, 2003). Our work also represents a novel application of the technique of occupancy modeling in fisheries science. Finally, a combination of this model with simple telemetry and dive surveys will enable unbiased, non-invasive monitoring of rockfish abundance inside and outside RCAs in BC.

10.2 Materials and Methods

10.2.1 Dive Technique Overview

Dives were conducted in 2005-2006, including paired point count dives at known sites (2005), simultaneous visual/video point count dives (2005) and a series of paired point count dives at randomly selected dive sites (2006). In 2005 the paired point count dives and simultaneous visual/video point count dives were employed to compare different counting techniques and to make preliminary observations about the relationship between rockfish abundance and substrate, including sites of known high rockfish abundance. In 2006, paired point count dives at random sites were used to test the DFO predictive model in terms of rockfish occurrence. In 2007 and 2008, sidescan sonar was used to predict microscale occurrence of piled boulders and associated rockfish abundance within RCAs, and tested by means of dive surveys.

The rockfish species monitored in this study include the quillback (*Sebastes maliger*), copper (*S. caurinus*), yelloweye (*S. ruberrimus*), tiger (*S. nigrocinctus*), black (*S. melanops*), yellowtail (*S. flavidus*), vermilion (*S. miniatus*) and Puget Sound (*S. emphaeus*). The first five species are managed together in Canada as inshore rockfishes and are important components of the live fishery aimed at Asian markets, a fishery that originated in Vancouver, BC (Love and Thorsteinson, 2002). The other rockfish that is managed as an inshore species in BC, the China rockfish (*Sebastes nebulosus*), does not occur in protected, inland marine waters like Howe Sound (Lamb and Edgell, 1986), so was not encountered in this study.

Divers used the roaming diver technique (Martell et al., 2000), investigating any suitable habitat within the depth range for that dive. Divers would swim as a pair, each directing attention in the opposite direction, up-slope versus down-slope; hand signals were used to avoid any double-counts. During these dives, rockfish were identified to species and life stage (adult, juvenile < 20cm, "prejuvenile" or early juvenile with color pattern distinct from that of adult). The prejuvenile color pattern was usually seen on fish up to 14-17 months age. To judge juvenile size, divers would landmark a fish against substrate features (line fish up against background) and then judge their hand against those features, the gloved hand having been measured against a 20cm scale. Dives encompassed all seasons in 2005 and summer/fall of 2006. At certain sites the suitable rockfish habitat did not extend deeper than 10-15m, whereas at other sites (notably Passage Island reefs), all rockfish habitat was at 15-35m depth. In 2006 dives were limited to < 20m.

10.2.2 Technique Verification by Video Survey

The comparison of videotape versus direct diver counts in 2005 consisted of one diver videotaping along a depth transect (roaming diver technique) while the dive partner conducted a direct visual count. The visual diver had to remain on the offshore side and slightly behind the videographer, using only lighting from the camera, in order to provide advantage to the videotaping in terms of any flight responses by fish.

A total of 24 video/visual dives was conducted in four areas, six dives per area. The areas were Popham Island (49° 21.45N, 123° 28.99W), Bowyer Island (49° 25.78N, 123° 16.48W), Whytecliff Park MPA (49° 22.62N, 123° 17.17W), and Passage Island (49° 20.44N, 123° 18.88W). For 17 such video dives, a second dive was conducted on the same date, at the same location, with both divers teaming on a direct visual count, using dive lights to investigate crevices.

In all these dives, rockfishes were identified to species and life stage, as for regular paired dives, with the video tapes scored for total counts of different species and life stages after the dive. Tukey's Honestly Significant Difference (HSD) was used to control experimental-wise error in the multiple pairwise comparisons of log transformed count per unit effort (CPUE) data, with CPUE equal to number of rockfish recorded per hour diving time.

10.2.3 Video Surveys by Substrate Category

Habitat covariates and count rates within specific dive tracks from the 2005 dives were also scored. The videos were broken down into one-minute segments in which rockfish occurrence was recorded together with habitat type by the methods of Krieger and Ito (1999). Krieger and Ito (1999) evaluated submersible videos of deepwater rockfishes, scoring minute-by-minute segments according to slope, substrate and rockfish abundance. For this study, the substrate of crushed shell was added to Krieger and Ito's mud, sand, pebble, cobble (60-250 mm), boulders and bedrock. Both primary (50-100%) and secondary (10-50%) substrates were recorded. Boulders were recorded by abundance category (B1 – scarce, B2 – scattered, B3 – always in view, B4 – piled). Slope categories were 1 = 0-5°, 2 = 6-20°, 3 = 21-45°, and 4 = 46-90°.

10.2.4 Abundance Class Study Design

Since the 2004/2005 dive sites were not randomly selected, a second study was conducted in 2006. Three regional areas containing one RCA were chosen (Lions Bay, Bowyer Island and West Bay; Fig. 10.1). Within each regional area, a control area was chosen to be as close as possible to the respective RCA, but outside of the DFO predicted rockfish abundance areas designated by DFO's rockfish habitat model (pers. comm., L. Lacko, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, BC V9R 5K6 Canada). Within each RCA and control area a 5 km stretch of shoreline was divided into 200m sectors, for a total of 25 potential sites per area. Within each RCA and control area three sites were chosen at random and each site was visited three times over several months. If the first dive of a site yielded a zero-count of rockfish, an additional randomly selected dive site was added to that area. In total 69 dives were conducted at 11 RCA sites and 12 control sites.

On each dive the rockfish point counts, bottom substrate types and active search time (bottom time) were recorded using timed segments. Once divers reached depth, dive time was noted along with substrate type and rockfish counts commenced. When a new substrate type was encountered, the time was recorded and the next timed segment began. This process continued for the duration of the dive. Within each timed segment all rockfish species were counted according to life history stage (prejuvenile, juvenile and adult) by the roaming divers. In addition to the point counts, divers also recorded the presence/absence of macroscopic plant and animal species, however these data were not included in the analysis. The predominant species in this region are copper and quillback rockfish; since other species were observed in trivial numbers, the analysis was restricted to these two species. The relatively shallow dive depths were expected to yield relatively more copper rockfish, since quillbacks prefer greater depths than coppers (Love and Thorsteinson, 2002)

For the analysis, point counts were first standardized to counts per hour and the rate was used to designate the observed abundance index ($y = 0, 1, 2, 3$) of a site on a specific sampling occasion. Abundance indices were separated roughly by orders of magnitude (Table 10.1). Orders of magnitude were chosen since observed count rates spanned approximately three orders of magnitude and it was felt that distinguishing between orders of magnitude would result in a more repeatable index.

Finally, the first dives were on June 27, 2006 and the final dives on October 24, 2006, all within the period of the seasonal thermocline and abundant food when copper rockfish tend not to hide (Patten 1973). None of the sites had been surveyed previously, although for the Bowyer Island RCA (one of the original 2004 RCAs) other parts of that RCA had been investigated extensively. It took about a month to complete initial dives at all 23 sites, and the final dives at the different sites were completed over a period of nearly two months. Dives were at < 20m depth.

Table 10.1: Observed abundance index (y) in relation to count rate.

Observed Count Rate	Abundance index	Interpetation
no fish observed	$y = 0$	None
$0 < \text{fish} \cdot \text{h}^{-1} \leq 10$	$y = 1$	Low
$10 < \text{fish} \cdot \text{h}^{-1} \leq 100$	$y = 2$	Medium
$100 < \text{fish} \cdot \text{h}^{-1} \leq 1000$	$y = 3$	High

Note: Rockfish count per unit effort (CPUE) are standardized to a per-hour count.

10.2.5 Use of Side-scan Sonar to Predict Rockfish Occurrence

Following the various surveys described above, a brief series of dives was guided in 2007/2008 by use of side-scan sonar for identifying isolated, small patches of piled boulders. The initial side-scan survey was conducted at NW Bowyer Island, outside the random dive sites for 2006, described above, where results in 2004/2005 had yielded isolated patches of high abundance of copper rockfish along the shoreline. Six potential patches were located in a 1 km stretch of shoreline and then divers explored these sites for verification of substrate type and rockfish occurrence. The same method was used in 2008 on the central reefs in the West Bay RCA (again, outside the 2006 shoreline random dive sites) and at another RCA in Howe Sound, Dommett Point RCA. Iterative dive verifications enabled perfection of interpreting side-scan images for identifying piled boulder habitat at the shoreline. The shoreline throughout the area for random site selection at Bowyer Island was surveyed by side-scan sonar in 2008 as well.

10.2.6 Multilevel Occupancy Model

Modeling point counts directly poses difficulties due to the detectability issues often associated with ecological measures (see MacKenzie et al., 2006). As such we treat point counts as general index of abundance (none, low medium and high) on a site. Each site is assumed to possess an underlying and unobservable latent abundance class, which gives rise to the observed abundance index. In doing so we are able to separate the observation process from the underlying abundance state, thereby making allowances for detectability. The model used for analyzing the data was based on the mixture multinomial model proposed by Royle and Link (2005). Here we modify the approach by using the log-odds to model the effect of covariates on the latent abundance class distribution.

Latent Abundance Class Distribution

Each site is assumed to have an unobservable (i.e. latent) abundance class $N_i = 0, 1, 2, 3$ that is fixed for the duration of sampling and is treated as being ordinal (i.e. $0 < 1 < 2 < 3$). Let $\Psi_{i,k}$ be the probability that site i is of abundance class k for $k = 0, 1, 2, 3$. The latent abundance class distribution will be multinomial $[N_i] \sim \text{Multinomial}\,(1, \{\Psi_{ik}\})$.

To investigate potential ability of factors such as belonging to an RCA, regional area within Howe Sound and the average boulder coverage to affect latent abundance class distribution, a conditional odds model was formulated, where the log-odds is of the form

$$\begin{aligned}\ln\left(\frac{P(N_i = k)}{P(N_i < k)}\right) &= \theta_k + a_1 R_i + b_1 \bar{B}_i + c_1 A1_i + c_2 A2_i + d_1 R_i \bar{B}_i \\ &\quad + e_1 R_i A1_i + e_2 R_i A2_i + f_1 \bar{B}_i A1_i + f_2 \bar{B}_i A2_i \quad \text{for } k = 3, 2, 1.\end{aligned} \tag{10.1}$$

Here R_i indicates whether site i is located in an RCA, $\bar{B}_i$ is the average percent boulder coverage of the site and $A1_i, A2_i$ indicates whether the site belongs to either the West Bay, Bowyer Island regions respectively. The Lions Bay effect will be picked up in the intercept and the regional area effects (c_1, c_2) are then interpreted relative to Lions Bay. Finally, in addition to the main effects, all two-way interactions were also considered.

The conditional odds model considers the odds that a site belongs to the highest abundance class ($N = 3$) versus a lower class to be a function of covariates. For lower classes (i.e. $N = 2, 1$), the odds are conditional on the site not belonging to a higher abundance class. A property of this formulation is that by constraining to a single intercept for all abundance classes (i.e. $\theta_3 = \theta_2 = \theta_1 = \theta$) the constraint $\Psi_3 > \Psi_2 > \Psi_1$ will be imposed. This could represent a scenario where habitat may be considered as either suitable or not suitable, and depending on the scale of sampling (i.e. the size of a site), the abundance classes $N = 1, 2$ could represent a situation where only part of the site is suitable rockfish habitat (the alternative being that the site has suboptimal substrate like bedrock).

Conditional Detection Probabilities

Upon each sampling occasion, the observed abundance index (the event) $\{y : y = 0, 1, 2, 3\}$ may occur where the observed index y will be either equal to or lower than the true latent abundance class N. As such it is implied that the observation $y > N$ will never occur. For example, a site with the latent abundance class $N = 3$, may produce observations such as $y = 0, 1, 2, 3$, but for sites where $N = 1$ we may only observe $y = 0, 1$.

Let $\pi_{it,ke}$ be the conditional probability of observing event $y_{it} = e$ for site i on occasion t, given that $N_i = k$. Therefore, for site i on occasion t the correct detection occurs with probability $\pi_{it,kk}$, while the misclassifications ($y_{it} < N_i$) occur with probability $\pi_{it,ke}$, since the misclassification $y_{it} > N_i$ occurs with probability 0, $\pi_{it,ke} = 0$ for $e > k$. Furthermore, for sites where $N_i = 0$, the event $y_{it} = 0$ occurs with probability 1, that is $\pi_{it,00} = 1$ for all t. Finally, note that the constraint $\sum_{e=0}^{3} \pi_{it,ke} = 1$ for all $k = 0, 1, 2, 3$ for all sites and occasions exists, and that the distribution of the observed indices is multinomial, conditional on N_i, $[Y_{it}|N_i = k] \sim \text{Multinomial}\left(1, \{\pi_{it,ke}\}\right)$, where $\{\pi_{it,ke}\}$ are i^{th} site's multinomial cell probabilities.

Royle and Link (2005) suggested an alternate formulation for conditional detection probabilities, wherein correct classification and misclassification are considered separate processes. We adapt this approach using a correct detection probability that is both site and occasion-specific ($\pi_{it,kk} = p_{it,k}$), but with a conditional misclassification probability that is shared across all sites and all occasions (i.e. $\pi_{it,ke} = \beta_{ke}$ for $e < k$; see Table 10.2). Note, the β_{ke} parameters represent a conditional misclassification probability. For example, if $N = 3$, β_{32} is the probability of observing $y = 2$, given that $y \neq 3$, while β_{31} is the probability of of observing $y = 1$, given $y \neq 2, 3$.

Lastly, there may be concern that an association between rockfish and boulders could affect the probability of a correct detection ($p_{it,k}$). As such, the following correct detection model was used:

$$\text{logit}\left(p_{it,k}\right) = \alpha_k + \gamma B_{it} \quad \text{for } k = 1, 2, 3, \tag{10.2}$$

where B_{it} is the percent boulder coverage encountered on site i during sampling occasion t. This allows for the probability of a correct classification to vary by site and occasion in response to the amount of boulder coverage encountered during a dive.

Marginal Likelihood

The joint likelihood of $\{\pi_{it,ke}\}$ and N_i produces a hierarchical model in the form $[y|N][N]$, where $[y|N]$ is the condition detection multinomial and $[N]$ could be interpreted as a random effect. As indicated by Royle and Link (2005) a marginal likelihood of y may be obtained by integrating out the effect of the latent parameter N producing the following likelihood,

$$L\left(\boldsymbol{\theta}, \boldsymbol{\pi}, \boldsymbol{\Psi}|\mathbf{Y}\right) = \prod_{i=1}^{R} \left\{ \sum_{k=0}^{3} \left[\prod_{t=1}^{3} P\left(Y_{it} = y_{it}|N_i = k; \{\pi_{it,ke}\}\right) \right] \cdot \Psi_{ik} \right\}. \tag{10.3}$$

Parameter estimates may then be obtained by maximum likelihood. Note that R is the total number of sites.

Table 10.2: Multinomial cell probabilities (general and alternate formulation) for the conditional sampling distribution of the observed abundance index for the site i (adapted from Royle and Link (2005)).

	$P\left(Y_{it}=0 \mid N_i\right)$	$P\left(Y_{it}=1 \mid N_i\right)$	$P\left(Y_{it}=2 \mid N_i\right)$	$P\left(Y_{it}=3 \mid N_i\right)$
$N_i=0$	$\pi_{it,00}=1$	$\pi_{it,01}=0$	$\pi_{it,02}=0$	$\pi_{it,03}=0$
$N_i=1$	$\pi_{it,10}=q_{it,1}$	$\pi_{it,11}=p_{it,1}$	$\pi_{it,12}=0$	$\pi_{it,13}=0$
$N_i=2$	$\pi_{it,20}=(1-\beta_{21})q_{it,2}$	$\pi_{it,21}=\beta_{21}q_{it,2}$	$\pi_{it,22}=p_{it,2}$	$\pi_{it,23}=0$
$N_i=3$	$\pi_{it,30}=(1-\beta_{31})(1-\beta_{32})q_{it,3}$	$\pi_{it,31}=\beta_{31}(1-\beta_{32})q_{ti,3}$	$\pi_{it,32}=\beta_{32}q_{it,3}$	$\pi_{it,33}=p_{it,3}$

Note: Rows correspond to possible latent abundance classes;s $q_{it,k}=1-p_{it,k}$ is used in order to keep the notation compact and the same misclassification probabilities are used for all sites and occasions.

Model Selection and Submodel Parameterization

The fully parameterized model described in (10.3) could potentially represent an over-fit of the data. Thus we may wish to consider sensible *a priori* parameter constraints on the conditional detection probabilities as well as the factors affecting the latent abundance class distribution. Four models were considered for both correct classification and misclassification giving a total of 16 possible conditional detection models (Table 10.3). In addition to these models, we also consider a possible 12 models describing the effect of factors on the latent abundance class distribution (Table 10.4). Finally, if we also consider restricting the intercept of the log-odds model to be equal for all abundance classes, this brings the total to 384 possible models.

Table 10.3: *A priori models* used for the analysis the conditional correct classification probability.

Model	Description of $p_{it,k}$ and β_{kj} structure.
Correct Classification	
$p(.)$	No boulder effect (i.e. $\gamma = 0$) with the a single intercept for all classes (i.e. $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$)
$p(k)$	No boulder effect, but separate intercept for each class.
$p(. + B)$	Occasion specific boulder effect on $p_{i,k}$, but a single intercept for all classes.
$p(k + B)$	Occasion specific boulder effect on $p_{i,k}$, with separate intercept for each class.
Misclassification	
$\beta(.)$	A single probability for all misclassification probabilities (i.e. $\beta_{21} = \beta_{31} = \beta_{32} = \beta$).
$\beta(31 = 21)$	Separate probability for β_{32} and the constraint $\beta_{31} = \beta_{21} = \beta$.
$\beta(32 = 31)$	Separate probability for β_{21} and the constraint $\beta_{32} = \beta_{31} = \beta$.
$\beta(k)$	Separate probabilities for each type of misclassification.

Rather than consider all models, we have employed a two stage model selection process using a small sample corrected Akaike Information Criterion (AICc) to rank and select models (see Burnham and Anderson, 2002). First the conditional detection model was chosen by fitting all possible conditional detection models with the $\Psi(R, B, A, RA)$ model both with and without the intercept constraint. Afterwards the top conditional detection model was used to fit and rank the various effect models.

Table 10.4: *A priori* models used to assess factor effects on $\Psi_{i,k}$.

Model	Description of $\Psi_{i,k}$ structure
$\Psi(R, B, A, I)$	Full model with RCA (R), boulder (B), region (A) and all two-way interactions.
$\Psi(R, B, A)$	RCA, boulder and regional main effects only.
$\Psi(R, B)$	RCA and boulder main effects.
$\Psi(R, B, RB)$	RCA and boulder main effects with an RCA*boulder interaction.
$\Psi(R, A)$	RCA and regional main effects.
$\Psi(R, A, RA)$	RCA and regional main effects, with an RCA*region interaction.
$\Psi(B, A)$	Boulder and regional main effects.
$\Psi(B, A, BA)$	Boulder and area main effects, with a boulder*region interaction.
$\Psi(R)$	RCA main effect.
$\Psi(B)$	Boulder main effect.
$\Psi(A)$	Regional main effects.
$\Psi(.)$	Null model with no main effects.

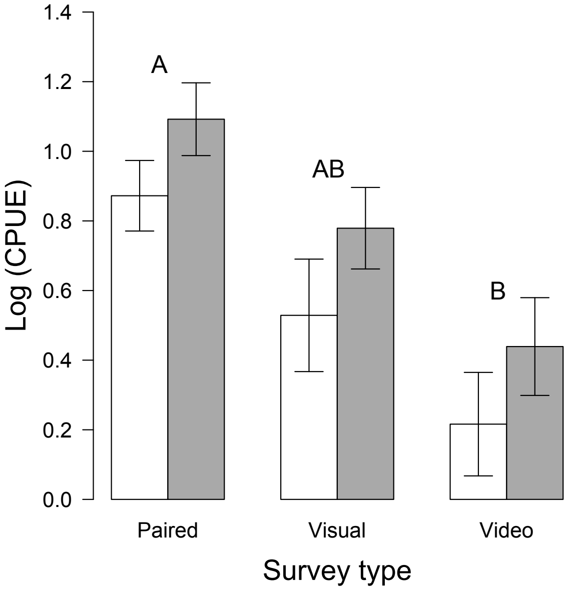
Support for model i in the model set $\mathcal{R}$ was assessed using the Δ_i value (read Delta $\mathrm{AIC_C}$), where $\Delta_i = \mathrm{AIC_C}_i - \min_{\mathcal{R}}(\mathrm{AIC_C}_i)$. In general, models with a value of $\Delta_i \leq 2$ have strong support, while models with $2 < \Delta_i < 7, 8, 9$ are considered plausible. However, models with $\Delta_i > 10$ are generally considered to have little to no support. Finally, model weights were determined using the Δ_i values, which allowed for model averaged effect estimates and standard errors to be calculated (Burnham and Anderson, 2002). This attempts to reduce potential spurious effects that a single model-based approach can produce.

10.3 Results

10.3.1 Video Survey/Method Verification

Log transformed CPUE for adults and all age classes of rockfish were compared between video/visual and paired visual surveys for 17 dates in 2005 to see if detection rates differed for the different survey techniques (Fig. 10.2). A global comparison showed all survey methods to be significantly different ($p < 0.05$) for comparisons between adults or all age classes combined. For those 17 pairs of dives, the overall average count (number of rockfish per hour) was 102 from videotape, 142 for direct visual count (alongside videographer), and 199 for paired diver counts. The method of paired diver counts therefore achieves a more complete census than can be obtained from videotaping.

Figure 10.2: Log transformed CPUE (count per hour) for adults and all age classes of rockfish between survey types (Paired = paired visual, Visual = visual alongside videographer, and Video =counts from videotape) for 17 dive dates. Different letters indicate pairwise significant difference ($p < 0.05$) for comparisons between adults or all age classes combined. Tukey's Honestly Significant Difference was used to control experimental-wise error. Unshaded bars for adults and shaded bars for all ages. Bars indicate standard error.

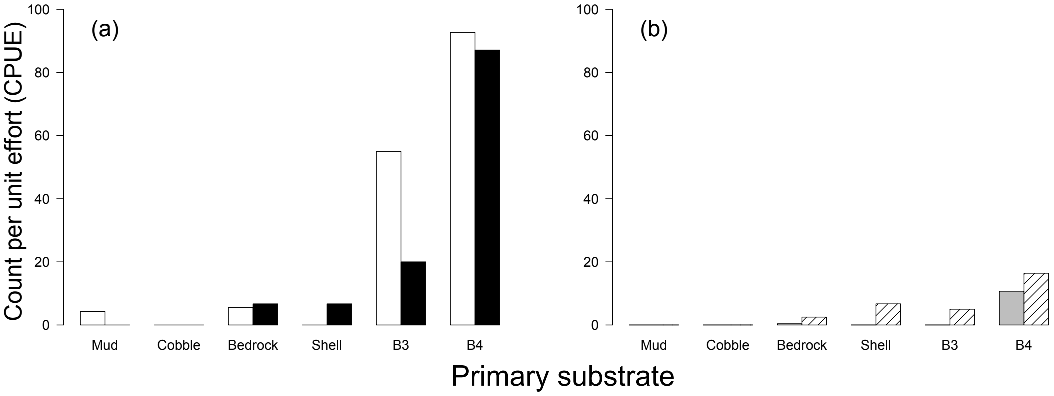


10.3.2 Video Surveys by Substrate Category

The substrate counts for the minute-by-minute scoring of the 2005 videotapes revealed no trend in rockfish abundance according to bottom slope, whereas adult rockfish appeared to strongly prefer piled boulders (> 80 adults per hour for copper and quillback) compared to other substrates (Fig. 10.3a). Adult copper rockfish are more likely than adult quillbacks to be found around scattered boulders or mud (rarely on mud, associated with hard objects like logs). Juvenile coppers and quillbacks also appeared to prefer piled boulders (Fig. 10.3b). In addition, prejuvenile coppers had newly recruited to both piled boulders and bedrock, whereas prejuvenile quillbacks occurred (at high densities averaging counts over 220 per hour) almost exclusively on cobble (a few in adjacent piled boulders). It should be noted that, aside from the video surveys, equally patchy and abundant prejuvenile coppers were seen at a number of boulder and bedrock reefs covered with the seaweed *Agarum fimbriatum*.

Examining deeper video dives where quillback rockfish dominated, a trend was evident for quillbacks to cluster at high density in larger areas of piled boulders (ten 4-6 min. segments averaged 19.4 (SD 14.3) adult quillbacks per segment), whereas 1 min. boulder segments ($n = 17$) averaged 1.8 (SD 2.7) quillbacks. Since the 2-3 min. segments ($n = 20$) averaged 7.8 (SD 6.2) adult quillbacks, the quillbacks were roughly twice as dense (3.1-3.9 fish per min.) in all longer segments than fish in single-minute segments, where zero counts were five times more frequent. For adult copper rockfish at shallower dive sites, however, no such trend was evident, with slightly higher densities of coppers in single-minute segments ($n = 9$, mean 4.1 (SD 4.5) fish), versus ca. 2 per minute in segments of 2-19 minutes ($n = 31$). All of these video segments were classified as piled boulders, though, whereas the boulder substrate classification in the 2006 random dives included scattered boulders as well as piled boulders.

Figure 10.3: CPUE (count per unit effort - 1hr) according to primary substrate type in minute-counts from 2005 videotapes for (*a*) adult and (*b*) juvenile copper versus quillback rockfish abundance.White bars for adult copper, black for adult quillback, gray for juvenile copper, wide hatching for juvenile quillback. Note that B3 is monolayer boulders, and B4 is piled boulders.



10.3.3 Abundance Class Model

In the 69 dives conducted at randomized sites in 2006 a total of 2,440 rockfish of all life stages was counted, including 1,881 copper (in 74 dive segments by substrate) and 527 quillback (in 47 dive segments). The next most abundant species was the Puget Sound rockfish, for which 26 fish were counted in two dive segments. Only two juvenile yelloweye rockfish were counted (in two dive segments) and single individuals of tiger, vermilion, yellowtail and black rockfish. All of these dives were shallow dives at < 20m depth (for a summary of counts by site see Table 10.5). In terms of habitat 1,821 fish were counted over piled boulders (CPUE $= 197.22$ fish $\cdot$ h^{-1}), 554 on bedrock (92.59 fish $\cdot$ h^{-1}), 49 on cobble (41.41 fish $\cdot$ h^{-1}), 16 on mud (4.75 fish $\cdot$ h^{-1}) and 0 on sand. The fish sighted over mud were hovering under sunken logs or moorage buoy anchor blocks.

Prior to the main analysis, the 16 potential detection models (Table 10.3) were assessed and ranked using the $\Psi(R, B, A, RA)$ effects model. The top detection model was $p(.)\beta(.)$ model for both the constrained and separate intercepts log-odds model sets. In both cases the next closest was the $p(.)\beta(31 = 21)$ model with a Δ_i of 1.8 and 4.3. The highest ranked detection model with a boulder effect on correct classification was the $p(B)\beta(.)$ model, which only had a Δ_i of 4.9 and 7.1 respectively indicating very little support. Therefore the $p(.)\beta(.)$ model was chosen as it had the highest rank with strong overall support (Akaike model weights of 52 and 82% for the constrained and separate intercept models respectively; note that the top model always has $\Delta_i = 0$), in addition to having the simplest structure. This detection model was used for the remainder of the analysis.

Next, to test whether being located in an RCA has any effect on the abundance class distribution, a subset of models was fit. Only the $\Psi(.)$, $\Psi(R)$, $\Psi(A)$, $\Psi(R, A)$ and $\Psi(R, A, RA)$ effect models were considered, since habitat differences between RCA and control sites may exist, potentially confounding an RCA effect. Fitting both the constrained and free intercept models produced nearly identical model rankings with the $\Psi(.)$ model as having the highest support, followed by the $\Psi(A)$ model (Table 10.6). The top model with an RCA effect had Δ_i values of 2.55 and 3.07 respectively, suggesting very weak support for an RCA effect, considering the top model was the null model. Furthermore, the model averaged estimate was -0.42 (SE 0.85) and -0.65 (SE 0.82) respectively, which is in the opposite direction of a beneficial RCA effect, but the estimate has a large standard error. The 95% confidence interval includes zero, also indicating no effect. Finally, there was some support

Table 10.5: Summary of the dive surveys within RCA and control sites.

		West Bay			**Lions Bay**			**Bowyer**		
		Sites	Count	Time	Sites	Count	Time	Sites	Count	Time
RCA	Boulder	3	135	162	2	789	193	1	30	66
	No boulder	2	146	104	1	15	63	2	26	142
Control	Boulder	1	156	85	3	946	201	3	136	143
	No boulder	3	132	203	0	0	0	2	35	107

Note: Sites designated as “boulder” have more that 20% boulder pile coverage. Counts are for juvenile and adult copper and quillback rockfish and time represents bottom time duration where active searching occurred by divers. The surveys are summarized by general area within the Howe Sound.

for a regional area effect ($\Delta_i = 1.41$ and 1.52 respectively) with highest log-odds estimated for the Lions Bay region, although effect estimates were again indistinguishable from zero. As such there is little evidence that being located in an RCA or a regional area in Howe Sound has any effect on the underlying abundance class.

The effect of changes in the average percent boulder coverage was investigated by fitting and ranking all possible effect models (Table 10.7). Here a strong effect was found. Both the constrained and free intercept log-odds models produced nearly identical model rankings, with strong support for the boulder only model $\Psi(B)$. In addition, the next closest model $\Psi(R, B)$ was over 3 AICc units away, suggesting little support for the predictive ability of any other factors. In fact each of the top five models included a boulder effect, suggesting support for these models is primarily due to the boulder term. As such it appears that, of the factors investigated, the only real factor predicting the latent abundance class distribution was boulder coverage.

Boulder coverage also had a large and positive effect on the latent abundance class distribution. The model averaged estimate for the boulder effect was 0.039 (SE 0.016) and 0.06 (SE 0.023) on the log-odds scale for the constrained and free intercept models respectively. Using the model averaged estimates from the fixed intercept model, if the boulder coverage of a site increased by 10 percentage units, the odds of the site belonging to the highest abundance class ($N = 3$) increase by a factor of 1.5 (95% CI 1.1, 1.9). For an increase of 50 percentage units, the odds would increase by a factor of 7 (95% CI 2.0, 24.8). Keep in mind this is a multiplicative increase of the odds, not the odds itself. As such this demonstrates a markedly positive effect of boulder coverage. By comparison, being located in an RCA or one of the regional areas had no such discernible effect.

It is worth noting that the constrained intercept log-odds model consistently had an AIC_Cvalue 2.5 to 12 units lower than the comparable free intercept model (Table 10.7). Since the same data set was fit for all models, the AIC_Cvalue can be used to select between the constrained and freed intercept log-odds models; the AIC_Cvalues suggest using the constrained intercept log-odds model. The premise behind the constrained intercept model was that habitat was either suitable for rockfish or not. As such the lower abundance classes (i.e. $N = 2, 1$) could be the result of sampling artifacts, where only a portion of the site is suitable (for example a mud bottom site that includes the edge of an adjacent boulder pile).

Finally, model averaged estimates of correct classification and misclassification probabilities were also investigated. Using the constrained intercept model, the model averaged

Table 10.6: Model rankings for the constrained and free intercept models for the RCA and regional area effects model set.

Model^a	Parm.^{b}	$\log(L)^c$	AICc	Δ_i	w_i	$\hat{a}_1$	$\hat{c}_1$	$\hat{c}_2$
Constrained intercept log-odds model								
$\Psi(.)$	3	-79.40	165.90	0.00	0.53			
$\Psi(A)$	5	-77.10	167.30	1.41	0.26		-2.19 (1.27)	-1.90 (1.28)
$\Psi(R)$	4	-79.20	168.50	2.55	0.15	-0.36 (0.65)		
$\Psi(R, A)$	6	-76.90	170.50	4.56	0.05	-0.41 (0.68)	-2.16 (1.22)	-1.90 (1.24)
$\Psi(R, A, RA)$	8	-76.20	177.50	11.55	0.00	-6.72 (23.93)	-8 (23.91)	-7.58 (23.91)
Free intercept log-odds model								
$\Psi(.)$	5	-78.80	170.80	0.00	0.56			
$\Psi(A)$	7	-75.90	172.40	1.52	0.26		-3.23 (1.63)	-2.80 (1.57)
$\Psi(R)$	6	-78.60	173.90	3.07	0.12	-0.47 (0.71)		
$\Psi(R, A)$	8	-75.20	175.30	4.50	0.06	-1.03 (0.93)	-3.48 (1.60)	-3.08 (1.56)
$\Psi(R, A, RA)$	10	-74.80	185.30	14.47	0.00	-7.997 (NA)	-10.15 (NA)	-9.47 (NA)

Note: Ranking was based on a small sample sized corrected AIC value ($\mathrm{AIC_C}$). The Δ_i column indicates the model $\mathrm{AIC_C}$ value relative to the smallest $\mathrm{AIC_C}$value in the constrained for free intercept model sets. The w_i indicates the corresponding Akaike model weight. Estimates for the RCA ($\hat{a}_1$) and regional area ($\hat{c}_1$, $\hat{c}_2$) effects are provided, parentheses indicate standard error. Note that both model sets were fit on the same data set and can be combined for an overall $\mathrm{AIC_C}$ ranking. This however was not done.

[a] The $p(.)\beta(.)$ detection model was used in all instances.

[b] Number of parameters.

[c] Log-likelihood.

Table 10.7: The top five effect models for both the constrained and free intercept model sets.

Model^a	Parm.^{b}	$\log(L)^c$	AIC$_\text{C}$	Δ_i	w_i	$\hat{a}_1$	$\hat{b}_1$	$\hat{c}_1$	$\hat{c}_2$
Constrained intercept log-odds model									
$\Psi(B)$	4	-72.9	155.7	0.00	0.75		0.039 (0.014)		
$\Psi(R,B)$	5	-72.8	158.8	3.12	0.16	-0.14 (0.77)	0.039 (0.014)		
$\Psi(B,A)$	6	-72.3	161.3	5.58	0.05		0.038 (0.015)	-1.03 (1.26)	-1.30 (1.29)
$\Psi(R,B,RB)$	6	-72.6	161.8	6.11	0.04	0.34 (1.03)	0.049 (0.022)	-1.12 (1.28)	-1.41 (1.34)
$\Psi(R,B,A)$	7	-72.2	165.1	9.36	0.01	-0.30 (0.79)	0.037 (0.015)		
Free intercept log-odds model									
$\Psi(B)$	6	-70.9	158.5	0.00	0.80		0.059 (0.022)		
$\Psi(R,B)$	7	-70.7	162.1	3.61	0.13	-0.49 (0.88)	0.059 (0.023)		
$\Psi(B,A)$	8	-69.9	164.9	6.41	0.03		0.053 (0.020)	-1.62 (1.52)	-1.94 (1.54)
$\Psi(R,B,RB)$	8	-70.1	165.3	6.78	0.03	0.30 (1.18)	0.087 (0.047)		
$\Psi(R,B,A)$	9	-69.6	169.1	10.66	0.00	-0.79 (0.93)	0.052 (0.020)	-1.91 (1.67)	-2.30 (1.72)

Ranking was based on a small sample sized corrected AIC value (AIC$_\text{C}$). The Δ_i column indicates the model AIC$_\text{C}$ value relative to the smallest AIC$_\text{C}$value in the constrained for free intercept model sets. The w_i indicates the corresponding Akaike model weight. Estimates for the RCA ($\hat{a}_1$), average percent boulder coverage ($\hat{b}_1$) and regional area ($\hat{c}_1$, $\hat{c}_2$) effects are provided, parentheses indicate standard error. Note that both model sets were fit on the same data set and can be combined for an overall AIC$_\text{C}$ ranking. This however was not done.

[1] The $p(.)\beta(.)$ detection model was used in all instances.

[b] Number of parameters.

[c] Log-likelihood.

95% confidence interval for the correct classification probability $p_.$ was (0.21, 0.6), while for $\beta_.$ it was (0.52, 0.86). Since this is markedly different from one, it was appropriate to account for detectability during the analysis. Lastly, it should be noted that the model set containing all possible factor models was used in the model averaged calculations.

10.3.4 Side-scan Sonar for Locating Piled Boulders

The first side-scan exploration was at NW Bowyer Island in October 2007, where six apparent rock piles were detected, five of which were verified by diving to consist of piled boulders inhabited by copper rockfish. The sixth, however, proved to be large boulders resting on bedrock or embedded in sand, with no rockfish associated (Fig. 10.4). These sites where the boulder piles were detected corresponded to the general area where high counts of rockfish had occurred in 2004.

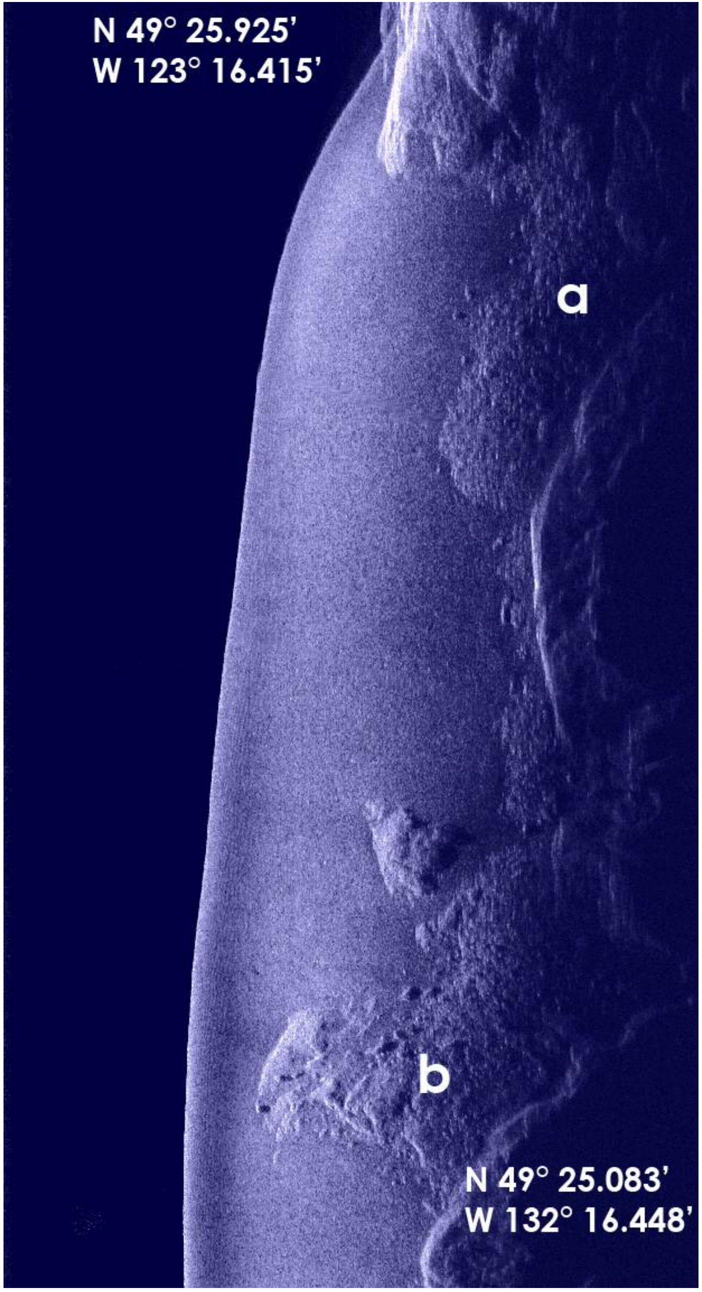
The center of West Bay RCA was explored with side-scan sonar, and only one rock pile was detected on extensive, high relief bedrock reef. That rock pile yielded the only rockfish sighted on that entire reef in verification diving, only six copper rockfish being seen, all at that small, isolated, cluster of boulders. Similarly, at the Dommett Point RCA, one small area was detected by side-scan with evidence of rock piles. A high density of quillback rockfish was found at the deeper end of that strip, then numerous copper rockfish in the shallower boulders. It should be noted that a high complexity bedrock area, lacking boulders, in that same RCA also had abundant copper rockfish.

The area from which random sites were selected at the Bowyer Island RCA was surveyed in its entirety by side-scan, and no boulder piles were detected. The boulder area which had been seen by divers was a very shallow rock slope with relatively sparse boulders in very shallow depths.

10.4 Discussion

The predictive model used by DFO did not isolate the areas in Howe Sound where higher abundances of rockfish occurred. The DFO model uses complexity of bathymetry patterns to identify potential rockfish habitat, as defined as high level of changing slope of substrate (Ardron, 2003), which tends to identify bedrock reef with ledges and ridges, but not piled rocks, which require a finer scale of telemetry, as with side-scan sonar . Similarly, the catch records used in the DFO model could not be summarized on as small a spatial scale as exists

Figure 10.4: Sidescan image at northwest Bowyer Island for which two areas of satisfactory rockfish habitat were predicted: Divers verified that (a) the upper cluster of rocks consisted of multilayered, piled boulders inhabited by copper rockfish, whereas (b) the area at the bottom of the image consisted of bedrock with large boulders embedded in sand or resting on bedrock, without any deep crevices for rockfish to hide in.



for isolated rock piles. Thus, the degree of fine-scale habitat affinity by rockfish for piled boulders, and the spatial scale at which that boulder habitat occurs, is on a smaller order of spatial magnitude than the size scale of RCAs, which themselves are at the lower limit of management practicality. It must therefore be anticipated that in many areas there will be patchy abundance of rockfish both inside and outside of designated RCAs, as documented in this study. The real challenge will be to determine whether significant differences in average fish size and density emerge through time within the confines of RCAs, in comparison with fish in areas outside the RCAs.

Nearly all of the shoreline in Howe Sound is rocky, and rockfish were observed in all but one of the randomly selected sites, both RCA and control (one control site in the West Bay area was mud bottom, with no sunken logs or moorage blocks, and thus with no rockfish). Other soft bottom sites had hard structures that harbored isolated rockfish.

We used the roaming diver technique, which is a search technique. We knew piled boulders were preferred habitat so we looked for boulders. The explanation for the contradictory results from sites classed as “boulder” at two of the areas (West Bay and Bowyer) results from concentration of dive effort on the limited, shallow boulder habitat patches that we were able to locate (for example a mud bottom site that includes the edge of an adjacent boulder pile). The most striking difference between major areas is that the boulders we found at Lions Bay were extensive and deeply layered, and ranged from shallow water to deeper than our dive limit of 20m. At the other sites we found boulders primarily at the shoreline, in depths as shallow as 5m, where fewer rockfish (only coppers occur shallow) were found, and the boulders were usually scattered rather than piled.

For the random site study we anticipated that different RCAs and control areas would vary in quality, as they had in our 2004/2005 surveys. Irvine (2008) discusses the necessity of accounting (via public process) for legitimate social and economic factors that need to be traded off in assessing risk management, as with RCA selection. Ardron and Wallace (2005) had shown that final RCA selections for the initial designation series in 2004 reflected the public process in addition to scientific modeling, and therefore were not as equally representative of optimal rockfish habitat and abundance as the original proposed RCAs based on the model only. We furthermore knew from 2004/2005 results that the Bowyer RCA, for example, had particular spots of high abundance that had been excluded from the random dive site selections in 2006 in order to avoid bias. In 2008 we used side-scan sonar to determine the extent to which site selection within an area affected CPUE. The very patchy occurrence of boulders affects results within any small sampling site more than larger area effects. Nevertheless, whereas discrete areas of very high abundance of piled boulders are known to exist at Bowyer Island, our side-scan results failed to reveal any such hot spots for the West Bay RCA or within the shoreline strip used for random selections at Bowyer Island RCA. Without using telemetry to examine all the shoreline of Howe Sound, we can only assume that the areas we used for random site selections represented the diversity of habitat quality across Howe Sound. Also note that both video and sidescan results yielded low to intermediate rockfish abundance on bedrock away from piled boulders.

It is not a surprise then that we were not able to demonstrate any differences between RCA and control sites. While we cannot rule out power issues when interpreting the lack of an RCA effect, the approach used was able to detect other factors, such as the strong association between boulder piles and rockfish. Furthermore, the positive association is also in agreement with the data from the 2005 minute-counts from videotapes and divers who observed a close association of rockfish with piled boulders. Although the affinity of rockfish for piled boulders had been known, and has been described in the literature (Love et al. 2002), the abundance class model verified this habitat association without the necessity of any risky assumptions about true detection rate being one.

Interestingly, the support for a fixed intercept log-odds model is also consistent with the observation that piled boulders are an important habitat type for rockfish. By fixing the intercept we imposed the following constraint $\Psi_3 > \Psi_2 > \Psi_1$. That is, the estimated frequency of $N = 3$ sites is restricted to be greater than the frequency of $N = 2$ sites, which again will be greater than the frequency of $N = 1$ sites. If boulder piles are the defining habitat for high density rockfish, then we may expect that most boulder piles will be capable of supporting roughly the same density of fish. As such the classes $N = 2$ and $N = 1$ may represent an artifact of sampling, where a site only contains an edge of a boulder pile. Alternatively, $N = 2$ and $N = 1$ could also represent suboptimal substrate like bedrock. Note that video analyses showed differing tendencies between copper and quillback rockfish in terms of relative extent of rock piles and rockfish densities, quillbacks tending less to aggregate at the smallest boulder patches. We believe our approach could be extended further to deal with such intricacies.

The model used also represents a departure for fishery science as it directly approaches the issue of detectability, a phenomenon inherent to most ecological measures (MacKenzie et al., 2006). While there may be concern that converting point counts to abundance indices (rather than using a general linear model with Poisson errors) discards valuable information, we do not feel this is the case. Raw point counts become uninterpretable if detectability of individuals changes between sampling occasions. Even though our indices were separated by orders of magnitude in the CPUE (ensuring a high degree of repeatability compared to straight point counts), the probability of a correct classification was still estimated to be substantially lower than 1; in fact it was estimated to be between 0.21 and 0.60. From our perspective the issue of detectability needs to be addressed more by fisheries scientists.

While with classical count models, such as the Poisson model, it is possible to account for factors like bottom time (i.e. how long you are counting), the best approach for handling detectability is not clear. It is not hard to imagine situations (like hiding behavior) where animal availability changes between sampling occasions. If the factors causing the detectability are known then they can be handled in a way similar to bottom time; however, if the factors are not well known, or even known at all, then inferences about the count rate will become inseparably confounded with issues of detectability. As such, it would only take small changes in detectability to introduce biases into such analyses. In these types of situations, mixture models, such as the abundance class model, tend to represent the best approach (Martin et al., 2005).

The issue of detectability aside, the approach taken also has other benefits that may make it an ideal candidate for long term monitoring studies. By using indices of abundance, rather than the point counts, the error between observers of differing skill set will be reduced as the measure is much easier to repeat. Since long term monitoring studies can have personnel changes, this would also help to ensure consistency across years. Furthermore, the method is relatively easy to employ, represents a non-invasive approach to monitoring and does not have to be restricted to measures of relative abundance. Other relevant biological states (e.g. reproduction, multiples species occurrences, risk effects) could potentially be modeled.

A key assumption of the model, however, is that the abundance class of sites is closed to changes during sampling. While such assumptions are likely to be violated by many species of fish, homing and site fidelity have been demonstrated in a number of rockfish species (Carlson and Haigh, 1972; Markevich, 1988; Matthews, 1990b). The home ranges for both copper and quillback rockfishes have been shown to occur on a scale of $< 30\text{m}^2$ (Matthews, 1990b,a), which is smaller than the 200m sampling scale used. Finally, small changes in site specific population levels may occur as long as the underlying latent abundance class does not change. Since orders of magnitude were used to differentiate between observed index levels, a rather substantial change in the underlying population level would be required before a site's underlying latent abundance class could be expected to change. Lastly, it should be noted that a type of adaptive sampling was used in the field to include 5 additional sites into the study design. An accommodation was not made for this sampling structure in the analysis due to the already small sample size of the study.

Howe Sound represents fjord geography, and does not have the extensive shallow soft bottom areas typical of the continental shelf, so the DFO model should be of greater utility for coastal areas without steep, glaciated shorelines. Note, however, that the DFO model was only used to reveal high complexity (rocky) bottom, at spatial scales that cannot be of utility for determining presence of boulders, much less boulders in multilayered deposits. In addition to high complexity, deposits of boulders from glacial retreat or from debris torrents or landslides are needed to create optimum habitat for these inshore rockfishes. As sonar telemetry methods increasingly permit remote detection of rock piles, direct confirmation of rockfish demography by diving or remote video will be required for determining the finer scale of rockfish distribution within a region. Yoklavich et al. (2007) have successfully used remote seabed habitat classification, based on side-scan, multibeam and seismic mapping of georeferenced seafloor substrata, together with manned submersible line-transect methods, to estimate rockfish abundance in the Cowcod Conservation Areas of southern California. Similar manned submersible observations had been used for determining habitat associations of rockfishes on Hecata Bank in Oregon (Stein et al., 1992) and in Monterey Bay in California (Yoklavich et al., 2000).

In the present study, diver counts yielded superior results compared to counts from videotape. It is also likely that the lack of maneuverability of a submersible limits the ability of observers to peer under ledges or behind rocks in the manner that divers can accomplish. When manned submersible observations are being conducted over habitat including rockfalls at the base of canyon walls (Yoklavich et al., 2000), the results of this study suggest that it may be reasonable to calculate up to twice the abundance observable from videotapes in that type of complex seabed habitat. It may, however, remain feasible only for scuba divers in shallow depths <40m to detect winter hiding behavior in rock piles (Patten, 1973; Carlson and Barr, 1977).

All of the 2007 RCAs in Howe Sound that we have studied encompass rockfish populations that may prosper under protection from fishing pressure. The present data will afford a baseline for comparison with the future that will enable determination of how well enforcement succeeds in protecting rockfish survival and abundance within these closed areas. Similarly, these baseline data will enable determination of whether rockfish abundance trends diverge between the protected and the unprotected portions of the Howe Sound shoreline. The success of Rockfish Conservation Areas as a management and conservation tool will depend on whether changes of demography can be documented inside the RCAs, relative to demography in outside areas. A higher future abundance of mature adults of inshore rockfish species can only be predicted for the RCAs if enforcement results in significant reduction of fishing mortality (especially poaching) in those areas. Side-scan telemetry for locating suitable rock piles within randomly selected study areas, inside and outside RCAs, can be used as a basis for documentary census work by divers, ROV or manned submersible that, analyzed with the abundance class model, will enable unbiased evaluation of the conservation impact of RCAs in British Columbia or of fishery closures elsewhere.

Chapter 11

Framework Discussion

The introduced framework is intended to provide flexible and robust statistical method for modeling ecological states. One of the key aspects is the formal handling of detection and classification errors. We accomplish this by following Pradel (2005) and explicitly separating the ecological state from the observation process, termed events. Researcher are then required to explicitly define both the relationship between the underlying ecological states and the observed events as well as how states change over time using a straightforward matrix formulation. This approach differs from the common approach of assuming the state of interest can be observed without any type of classification error (e.g. Petersen et al., 2009). Violations of this assumption can bias estimates of ecological processes, limiting its usefulness. In developing this aproach we also provided a comprehensive framework to help define these relationships, estimate parameters and assess overall model fit.

The underlying premise of our approach is that ecological processes can be abstracted in terms of states and transitions. We defined ecological states as any unique configuration of ecological characteristics, which is consistent with other approaches (e.g. Stringham et al., 2003). This type of reductionist approach has a number of advantageous when describe complex ecological systems. For example, multiple ecological characteristics may be considered simultaneously without the need to make direct assumptions about the structure of the association and can be useful for acknowledging the limits in available knowledge.

Simplifications are often needed in order to make hypothesis tests for key ecological processes tractable and to account for limits in available knowledge. For example, there would be little utility in formulating hypotheses about absolute changes in abundance for a species, if abundance estimates are known to be problematic. That said, there may still be enough information available to reformulate a hypothesis in terms of a relative increase or decrease in abundance. This requires a pragmatic approach to knowledge, instead of blind drive towards the goal of building models that closely matches the "truth" as possible.

The motivation behind abundance class formulation provides a good example of this type of thinking. The difficulties in obtaining direct estimates of abundance (e.g. due to issues of scarcity, financial constraints, lack of manpower, etc.) necessitated the use of proxy measures (i.e. the anuran call index), which can be interpreted as ordinal states of abundance. Depending on the management and/or research objectives, this less detailed proxy measure may still provide sufficient information for management decisions. For example, if the management goal is to monitor the effect habitat policies have on anuran abundance absolute measures of abundance may not provide much more practical information than is already attainable with realtive measures of abundance.

The use of ecological states to represent more complex underlying structure can also provide other benefits such as reducing observed variability and providing an easier way to amalgamate differing types of information. If we return to the example of abundance, the cost of obtaining abundance estimates may result in small sample sizes and high degree of error in the subsequent estimates. However, the easier nature of obtain state measures may provide a more precise estimates of change due to sheer differences in sample size. In addition, ecological state measures may also be more reproducible in the field, providing another mechanism to reduce variability. Finally, ecological states provide a straightforward way to incorporate normally incompatible types of information (e.g. local or traditional knowledge with scientific knowledge; Knapp and Fernandez-Gimenez, 2009).

Our main contribution to this modeling approach was to formalized issues of state detection via a hierarchical modeling structure that should not be confused with other hierarchical approaches used in ecology such as GLMMs. GLMMs attribute the often over-dispersed nature of in ecological data sets to variability in the ecological process itself and not the observation process, such as we have modeled. While it is possible to incorporate additional variation attributable to the observation process using a type GLM approach, it is not straightforward to relate this variability directly to the underlying ecological state such as we have done. And with our approach, it is possible to have some states produce more observational variation than others. For example, a high abundance sites may produce less variability in an observed index than a low or marginal abundance sites, due to issues of scarcity.

Furthermore, it is fairly straightforward to add additional hierarchal structure such as random effects to our modeling framework. The primary difficulty will be parameter estimation, but can be accomplished using data cloning (Ponciano et al., 2009) or by employing a Bayesian methodology (see Gelman et al., 2004). That said, we believe the outlined framework provides a lot of utility for describing biological processes. Practitioners should prioritize investigating the interplay between events and states before adding additional hierarchies to the overall model.

Our adherence to state-and-transition models also brings drawbacks in terms of model assumptions. Our framework has two distinct levels of abstraction, both of which may be difficult for model practitioners to verify. On the basic level is the assumption that the assemblages of characteristics we define as states carry relevant biological meaning and that changes in these assemblages can be modeled as transitions. This shortcoming is true of all state-and-transition models. The second level of abstraction is the relationship between observed events and the underlying states. Potentially verifying either of these relationships will not be possible or at least very difficult. The abstraction also places additional emphasis on the model build strategy and the available tools to assess goodness-of-fit.

We agree with Burnham and Anderson (2002) that careful consideration should always be given to the *a priori* set of questions and goals before specifying models. Ideally, the state, transition and event structures should be based on the best available expert opinion and past knowledge. Afterwards, all available tools for assessing goodness-of-fit should be used to verify any shortcomings of the specified models. This can be accomplished in a variety of methods including using the general omnibus goodness-of-fit test we provided, as well as the *ad hoc* visualization technique we derived for the abundance class formulation. Other model fit assessments may also be built off our general approach depending on the type of predictions the model can make. This puts a strong emphasis on the practitioner to think about ways the appropriateness of model structure may be tested, which is an often overlooked component of statistical ecology.

Another promising route for assessment is the Model Structural Adequacy (MSA) analysis framework (Taper et al., 2008). The simulation-based methodology assesses the ability of statistical models to answer scientific questions by fitting candidate models to data simulated by a complex process model (CPM; a proxy for reality). The CPM is based on the best available expert opinion and is designed to produced features either thought or suspected to occur in reality. Variations to the CPM may be produced to test the ability of candidate models to answer scientific questions. For example if we return to anuran abundance class formulation, a CPM could track real changes in abundance, while indices could be produced based on various density assumptions. The ability for the abundance class framework to track changes over time under various assumptions could then be tested.

The MSA analysis does an end-run around the issues of goodness-of-fit by directly testing the model robustness to answer scientific questions given a set of plausible circumstances. Arguably, most important aspect of ecological models is their ability to adequately answer purposed scientific questions. This perspective differs from many other statistical applications where the ability to make predictions tends to be preferred. In this way ecological models may also be better suited to an MSA analysis and may be a more appropriate way to assess complex ecology models than more traditional goodness-of-fit tests, which are primarily concerned with assess whether the observed data originates from a particular distribution, a fact we know to be false. The caveat with this type of approach is its heavy dependence on the quality of the CPM and the ability to ask the right kinds of questions. However, by combining a number of approaches, for example graphical, traditional goodness-of-fit measures and MSA analyses, it should be possible to verify many of the underlying model assumptions in states, transitions and the state to event structure.

Finally, in terms of employing the multi-state occupancy framework, the temporary and permanent sampling protocols were intended handle the two most common sampling situations. The main purpose of the temporary protocol is to handle situations where the sampling frame is expected to change from season to season, for example ephemeral ponds. In this situation the permanent sampling protocol is nonsensical. For situations where sites may be sampled in all possible seasons can be handled by the permanent sampling protocol, seasons that are missed are assumed to be missing completely at random. Ideally, this means that related to either the underlying state, event or underlying model parameter. We currently do not provide for other types of "missingness."

Part III

Summary

Chapter 12

Discussion

The topic of heterogeneity has been a defining theme of this thesis. In the super-population extension in order to account for capture heterogeneity caused by sexes, the problem of how to handle uncertainty in strata assignments needed to be solved. In the ρ-CJS work, capture heterogeneity caused by mating pairs sharing fates was handled by directly modeling the mating pair association (the first such effort in mark-recapture methodology). Finally, the driving force behind the multi-state occupancy work was the need to account for detection heterogeneity caused by differences in underlying site abundances.

Another other unifying theme was the compartmentalization of statistical models in terms of an underlying an ecological process of interest (which is not always observable) and a process generating the observation that is a function of the underlying state. This perspective was embodied in the use of the multievent framework to define the multi-state multi-season occupancy framework and the state-space formulation used in the ρ-CJS work. This type of explicit separation is relatively recent development in ecological statistics and will likely become the *de facto* method for modeling ecological experiments in general as it more closely approaches the true nature of the observed data.

Together these themes underly the general difficulty in analyzing ecological data. The underlying biological process of interest is difficult to observe and observations tend to be highly variable due to problems of detection. In addition to this problem, statistical models are also tasked with job of carrying the meaning of science. The combination of these factors makes the task of adequately assessing model fit crucial, yet the complexities of model makes this task difficult. Currently, our methods for assessing fit have not kept pace with the rise in complexity.

The Model Structure Adequacy (MSA) approach, mentioned in Chapters 9 and 10, presents an interesting perspective for future work whereby we worry more about testing the impact assumptions have on the ability to answer scientific questions rather than a particular statistical definition of fit. In terms of the multi-state occupancy more work is needed understanding the impact of using ecological states as proxies for not being able to observe numerical abundances. Using the MSA approach as a guiding philosophy, I would like to reconsider the power analysis using a known abundance distributions with different choices in the cut-points used generate the observed abundance indices. The idea would be get a better understanding of the impact differing field protocols may have on the ability to detect changes in the true numerical abundance. While not a direct application of the MSA type approach it is in keeping with it's pragmatic view that the ultimate worth of a statistical model is the ability to correctly answer scientific questions.

Other interesting extensions to the use of relative abundance data would be the inclusion of observer error in the observation itself. For example it would be interesting to consider abundance observations in terms of a ranges. A field observer on a given sampling occasion could record that somewhere between 50 and 100 individuals were observed rather than attempt a "precise" count. If we assume an underlying true abundance distribution exists, then it should be possible to assign a probability to the observation.

By presenting count data in a range the data would also more closely represent the types of data collected in the field. Often field staff are forced to create the illusion of precision counts. For example, it is fairly unlikely a technician could accurately count 225 birds on a given sampling occasion, but it should be possible of a technician to accurately produce a count range. By allowing abundance to be recorded in terms of a range, field technicians would be encouraged to communicate uncertainties in their assignments. Similar to the ρ-CJS work a state-space type approach would likely be used, due to its flexibility and ability to simplify model specification

Lastly, we have just scratched the surface with the innovative ρ-CJS model. The next step will be to make an allowance for mate switching, which will require a number of considerations from re-imagining the data structure of a mark-recapture experiment to including a process in the model that represents how individuals form mating pairs. Once these additions are complete, the model will be able to handle the original motivating data set and should be ready for peer-reviewed publication.

Afterwards, the next step will be extend to the overall approach to Jolly-Seber models.

While the Cormack-Jolly-Seber framework was a useful place to start, Jolly-Seber models will allow for the consideration all uncaptured individuals. This could provide great utility, especially if we continue using a state-space approach. For example, one potential benefit will be an improved ability to handle partial knowledge of mating pairs. By using state-space approach the entire population will be "known" at each occasion opening up the possibility to assign potential matches based on suitability. If, for example, we know all the single individuals that are alive on a given occasion, we can assume that these are the available matches. If a Bayesian approach is taken it would be trivial to assign posterior probabilities to a selection of potential mates. I can see such a feature being of great utility to behavioural ecologists.

Bibliography

Anthony, R. G., Forsman, E. D., Franklin, A. B., Anderson, D. R., Burnham, K. P., White, G. C., Schwarz, C. J., Nichols, J. D., Hines, J. E., Olson, G. S., Ackers, S. H., Andrews, L. S., Biswell, B. L., Carlson, P. C., Diller, L. V., Dugger, K. M., Fehring, K. E., Fleming, T. L., Gerhardt, R. P., Gremel, S. A., Gutierrez, R. J., Happe, P. J., Herter, D. R., Higley, J. M., Horn, R. B., Irwin, L. L., Loschl, P. J., Reid, J. A., Sovern, S. G., 2006. Status and trends in demography of northern spotted owls, 1985-2003. Wildlife Monographs 163, 1–48.

Ardron, J., 2003. GIS recipe for determining benthic complexity. In: Bremen, J. (Ed.), Marine geography: GIS for the oceans and seas. Esri Press, Redlands CA, pp. 169–175.

Ardron, J., Wallace, S., 2005. Modelling inshore rockfish habitat in British Columbia: a pilot study. In: Wright, D., Scholtz, A. (Eds.), Place matters: geospatial tools for marine science. University of Oregon Press, Eugene OR, pp. 47–68.

Arnason, A. N., Mills, K. H., 1981. Bias and loss of precision due to tag loss in jolly-seber estimates for mark-recapture experiments. Canadian Journal of Fisheries and Aquatic Sciences 38 (9), 1077–1095.

Beebee, T. J. C., 1995. Amphibian breeding and climate. Nature 374 (6519), 219–220.

Benton, T. G., Plaistow, S. J., Coulson, T. N., 2006. Complex population dynamics and complex causation: devils, details and demography. Proceedings of the Royal Society B: Biological Sciences 273 (1591), 1173–1181.

Bestelmeyer, B. T., Tugel, A. J., Peacock, G. L., Robinett, D. G., Sbaver, P. L., Brown, J. R., Herrick, J. E., Sanchez, H., Havstad, K. M., 2009. State-and-transition models for heterogeneous landscapes: A strategy for development and application. Rangeland Ecology & Management 62 (1), 1–15.

Blaustein, A. R., 1994. Chicken little or nero fiddle - a perspective on declining amphibian populations. Herpetologica 50 (1), 85–97.

Blaustein, A. R., Belden, L., Olson, D., Green, D., Root, T., Kiesecker, J., 2001. Amphibian breeding and climate change. Conservation Biology 15 (6), 1804–1809.

Bloomfield, P., 1976. Fourier analysis of time series: an introduction. Whiley, New York.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., White, J.-S. S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends In Ecology & Evolution 24 (3), 127–135.

Bond, J. C., Iverson, S. A., Maccallum, N. B., Smith, C. M., Bruner, H. J., Esler, D., 2009. Variation in breeding season survival of female harlequin ducks. Journal of Wildlife Management 73 (6), 965–972.

Buckland, S. T., Garthwaite, P. H., 1991. Quantifying precision of mark-recapture estimates using the bootstrap and related methods. Biometrics 47 (1), 255–268.

Buckland, S. T., Laake, J. L., Borchers, D. L., 2010. Double-observer line transect methods: levels of independence. Biometrics 66 (1), 169–177.

Buckland, S. T., Newman, K. B., Thomas, L., Koesters, N. B., 2004. State-space models for the dynamics of wild animal populations. Ecological Modelling 171 (1-2), 157–175.

Burnham, K. P., Anderson, D. R., 2002. Model Selection and multi-model inference: a practical information-theoretic approach, 2nd Edition. Springer-Verlag, New York.

Carlson, H., Barr, L., 1977. Seasonal changes in spatial distribution and activity of two species of Pacific rockfishes, *Sebastes flavidus* and *S. ciliatus*, in Lynn Canal, southeastern Alaska. Marine Fisheries Review 39 (3), 23–24.

Carlson, H., Haigh, R., 1972. Evidence for a home site and homing of adult yellowtail rockfish, *Sebastes flavidus*. Journal of the Fisheries Research Board of Canada 29 (7), 1011–1014.

Caswell, H., 2001. Matrix population models. Sinauer Associates, Sunderland, MA.

Catchpole, E. A., Morgan, B. J., 1997. Detecting parameter redundancy. Biometrika 84 (1), 187–196.

Catchpole, E. A., Morgan, B. J., 2001. Deficiency of parameter-redundant models. Biometrika 88 (2), 593–598.

Catchpole, E. A., Morgan, B. J., Freeman, S. N., 1998. Estimation in parameter-redundant models. Biometrika 85 (2), 462–468.

Catchpole, E. A., Morgan, B. J., Viallefont, A., 2002. Solving problems in parameter redundancy using computer algebra. Journal of Applied Statistics 29 (1-4), 625–636.

Challenger, W. O., Schwarz, C. J., 2009. Mark-recapture Jolly-Seber abundance estimation with classification uncertainty. In: Patil, G. P., Thomson, D. L., Cooch, E. G., Conroy, M. J. (Eds.), Modeling Demographic Processes In Marked Populations. Vol. 3 of Environmental and Ecological Statistics. Springer US, pp. 827–844.

Conroy, M. J., Senar, J. C., Hines, J. E., Domenech, J., 1999. Development and application of a mark-recapture model incorporating predicted sex and transitory behaviour. Bird Study 46 Supplement, S62–73.

Cormack, R., 1964. Estimates of survival from the sighting of marked animals. Biometrika 51 (3-4), 429–438.

Coull, B., Agresti, A., 1999. The use of mixed logit models to reflect heterogeneity in capture-recapture studies. Biometrics 55 (1), 294–301.

Cowen, L., Schwarz, C. J., 2006. The Jolly-Seber model with tag lossolly-seber model with tag loss. Biometrics 62 (3), 699–705.

Craig, M. D., Roberts, J. D., 2001. Evaluation of the impact of time of day, weather, vegetation density and bird movements on outcomes of area searches for birds in eucalypt forests of south-western Australia. Wildlife Research 28 (1), 33–39.

Cressie, N., Read, T. R., 1989. Pearson's χ^2 statistic and the loglikelihood ratio statistic G^2: a comparative review. International Statistical Review 57, 19–43.

Donnelly, M. A., Guyer, C., 1994. Patterns of reproduction and habitat use in an assemblage of neotropical hylid frogs. Oecologia 98 (3-4), 291–302.

Fujiwara, M., Caswell, H., 2002. Estimating population projection matrices from multi-stage mark-recapture data. Ecology 83, 3257–3265.

Gardner, T., 2001. Declining amphibian populations: a global phenomenon in conservation biology. Biodiversity and Conservation 24 (2), 25–44.

Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B., 2004. Bayesian data analysis, 2nd Edition. Chapman and Hall, London.

Gibbs, J., Breisch, A., 2001. Climate warming and calling phenology of frogs near Ithaca, New York, 1900-1999. Conservation Biology 15 (4), 1175–1178.

Giere, R. N., 2004. How models are used to represent reality. Philosophy Of Science 71 (5), 742–752.

Gimenez, O., Choquet, R., Lebreton, J.-D., 2003. Parameter redundancy in multistate capture-recapture models. Biometrical Journal 45 (6), 704–722.

Gimenez, O., Rossi, V., Choquet, R., Dehais, C., Doris, B., Varella, H., Vila, J.-P., Pradel, R., 2007. State-space modelling of data on marked individuals. Ecological Modelling 206 (3-4), 431–438.

Grant, E. C., Jung, R. E., Nichols, J. D., Hines, J. E., 2005. Double-observer approach to estimating egg mass abundance of pool-breeding amphibians. Wetlands Ecology and Management 13 (3), 305–320.

Green, A. J., 1990. Determinants of chorus participation and the effects of size, weight and competition on advertisement calling in the tungara frog, *Physalaemus pustulosus* (Leptodactylidae). Animal Behaviour 39, 620–638.

Hanski, I., Henttonen, H., 1994. Traces of population dynamics past. Trends In Ecology & Evolution 9 (1), 2–4.

Hart, J., 1973. Pacific fishes of Canada. Bulletin / Fisheries Research Board of Canada 180, 740.

Herrick, J., Bestelmeyer, B. T., Archer, S., Tugel, A., Brown, J. R., 2006. An integrated framework for science-based arid land management. Journal of Arid Environments 65 (2), 319–335.

Hocking, D. J., Rittenhouse, T. A. G., Rothermel, B. B., Johnson, J. R., Conner, C. A., Harper, E. B., Semlitsch, R. D., 2008. Breeding and recruitment phenology of amphibians in Missouri oak-hickory forests. American Midland Naturalist 160 (1), 41–60.

Huggins, R. M., 1989. On the statistical-analysis of capture experiments. Biometrika 76 (1), 133–140.

Irvine, J. R., 2008. The successful completion of scientific public policy: lessons learned while developing Canada's Wild Salmon Policy. Environmental Science & Policy 12 (2), 140–148.

Johnson, J., Omland, K., 2004. Model selection in ecology and evolution. Trends in Ecology & Evolution 19 (2), 101–108.

Jolly, G. M., 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. Biometrika 52 (1/2), 225–247.

Kendall, W. L., Hines, J. E., Nichols, J. D., 2003. Adjusting multistate capture-recapture models for misclassification bias: manatee breeding proportions. Ecology 84 (4), 1058–1066.

Kendall, W. L., Nichols, J. D., Hines, J. E., 1997. Estimating temporary emigration using capture-recapture data with Pollock's robust design. Ecology 78, 563–578.

Knapp, C. N., Fernandez-Gimenez, M. E., 2009. Understanding change: integrating rancher knowledge into state-and-transition models. Rangeland Ecology & Management 62 (6), 510–521.

Krebs, C. J., 1991. The experimental paradigm and long-term population studies. Ibis 133, 3–8.

Krieger, K. J., Ito, D. H., 1999. Distribution and abundance of shortraker rockfish, *Sebastes borealis* and rougheye rockfish, *S. aleutianus*, determined from a manned submersible. Fishery Bulletin 97 (2), 264–272.

Lamb, A., Edgell, P., 1986. Coastal fishes of the Pacific Northwest. Harbour Publishing, Madeira Park, BC.

Lebreton, J. D., Burnham, K. P., Clobert, J., Anderson, D. R., 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case-studies. Ecological Monographs 62 (1), 67–118.

Lebreton, J. D., Pradel, R., 2002. Multistate recapture models: modelling incomplete individual histories. Journal of Applied Statistics 29 (1-4), 353–369.

Lindenmayer, D. B., Cunningham, R. B., Donnelly, C. F., Incoll, R. D., Pope, M. L., Tribolet, C. R., Viggers, K. L., Welsh, A. H., 2001. How effective is spotlighting for detecting the greater glider (*Petauroides volans*)? Wildlife Research 28 (1), 105–109.

Lindsay, B. G., 2004. Statistical distances as loss functions in assessing model adequacy. In: Taper, M. L., Lele, S. R. (Eds.), The nature of scientific evidence: statistical, philosophical and empirical considerations. The University of Chicago Press, Chicago, pp. 439—488.

Link, W. A., Barker, R. J., 2005. Modeling association among demographic parameters in analysis of open population capture-recapture data. Biometrics 61 (1), 46–54.

Logan, G., de la Mare, W., King, J., Haggarty, D., 2005. Management framework for Strait of Georgia lingcod management framework for Strait of Georgia lingcod. Pacific Scientific Advice Review Committee Research Document 2005/048, Fisheries and Oceans Canada.

Love, M. M. Y., Thorsteinson, L., 2002. The rockfishes of the Northeast The rockfishes of the Northeast Pacific. University of California Press, Berkeley CA.

MacKenzie, D. I., Bailey, L. L., 2004. Assessing the fit of site-occupancy models. Journal of Agricultural, Biological, and Environmental Statistics 9 (3), 300–318.

MacKenzie, D. I., Kendall, W. L., 2002. How should detection probability be incorporated into estimates of relative abundance? Ecology 83 (9), 2287–2293.

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., Langtimm, C. A., 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83 (8), 2248–2255.

MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., Hines, J. E., 2006. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier, San Diego CA.

Mackenzie, D. I., Nichols, J. D., Seamans, M. E., Gutierrez, R. J., 2009. Modeling species occurrence dynamics with multiple states and imperfect detection. Ecology 90 (3), 823–835.

MacKenzie, D. I., Royle, J. A., 2005. Designing occupancy studies: general advice and allocating survey effort. Ecology 42, 1105–1114.

Markevich, A., 1988. Nature of territories and homing in the eastern Sea-Perch, *Sebastes taczanowski*. Journal of Ichthyology 28 (1), 161–163.

Marliave, J., Challenger, W. O., 2009. Monitoring and evaluating rockfish conservation areas in british columbia. Canadian Journal of Fisheries and Aquatic Sciences 66 (6), 995–1006.

Martell, S., Walters, C., Wallace, S., 2000. The use of marine protected areas for conservation of lingcod (*Ophiodon elongatus*). Bulletin of Marine Science 66 (3), 729–743.

Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., Tyre, A. J., Possingham, H. P., 2005. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecology Letters 8 (11), 1235–1246.

Matthews, K. R., 1990a. An experimental-study of the habitat preferences and movement patters of copper, quillback, and brown rockfishes (*Sebastes* spp). Environmental Biology of Fishes 29 (3), 161–178.

Matthews, K. R., 1990b. A telemetric study of the home ranges and homing routes of copper and quillback rockfishes on shallow rocky reefs. Canadian Journal of Zoology 68 (11), 2243–2250.

McCauley, S. J., Bouchard, S. S., Farina, B. J., Isvaran, K., Quader, S., Wood, D. M., Mary, C. M. S., 2000. Energetic dynamics and anuran breeding phenology: insights from a dynamic game. Behavioral Ecology 11 (4), 429–436.

Meng, X. L., 1994. Posterior predictive p-values. Annals of Statistics 22 (3), 1142–1160.

Nelson, G. L., Graves, B. M., 2004. Anuran population monitoring: comparison of the North American Amphibian Monitoring program's calling index with mark-recapture estimates for *Rana clamitans*. Journal of Herpetology 38 (3), 355–359.

Nichols, J. D., 2001. Using models in the conduct of science and management of natural resources. In: Shenk, T., Franklin, A. (Eds.), Modeling in natural resource management: development, interpretation, and application. Island Press, Washington, DC, pp. 11–34.

Nichols, J. D., Hines, J. E., MacKenzie, D. I., Seamans, M. E., Gutierrez, R. J., 2007. Occupancy estimation and modeling with multiple states and state uncertainty. Ecology 88 (6), 1395–1400.

Nichols, J. D., Kendall, W. L., Hines, J. E., Spendelow, J. A., 2004. Estimation of sex-specific survival from capture-recapture data when sex is not always known. Ecology 85 (12), 3192–3201.

Nieminen, J. K., 2008. Is bad mathematics useless? Ecological Modelling 213 (3-4), 468–470.

Patten, B. G., 1973. Biological information on copper rockfish in Puget Sound, Washington. Transactions of the American Fisheries Society 102 (2), 412–416.

Petersen, S. L., Stringham, T. K., Roundy, B. A., 2009. A process-based application of state-and-transition models: a case study of western juniper (*Juniperus occidentalis*) encroachment. Rangeland Ecology & Management 62 (2), 186–192.

Pledger, S., 2000. Unified maximum likelihood estimates for closed capture-recapture models using mixtures. Biometrics 56 (2), 434–442.

Pollock, K., Nichols, J. D., Brownie, C., Hines, J. E., 1990. Statistical-inference for capture-recapture experiments. Wildlife Monographs 107, 1–97.

Ponciano, J. M., Taper, M. L., Dennis, B., Lele, S. R., 2009. Hierarchical models in ecology: confidence intervals, hypothesis testing, and model selection using data cloning. Ecology 90 (2), 356–362.

Pradel, R., 2005. Multievent: An extension of multistate capture-recapture models to uncertain states. Biometrics 61 (2), 442–447.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Reading, C. J., 1998. The effect of winter temperatures on the timing of breeding activity in the common toad *Bufo bufo*. Oecologia 117 (4), 469–475.

Royle, J. A., Link, W. A., 2005. A general class of multinomial mixture models for anuran calling survey data. Ecology 86 (9), 2505–2512.

Royle, J. A., Link, W. A., 2006. Generalized site occupancy models allowing for false positive and false negative errors. Ecology 87 (4), 835–841.

Royle, J. A., Nichols, J. D., 2003. Estimating abundance from repeated presence-absence data or point counts. Ecology 84 (3), 777–790.

Royle, J. A., Young, K. V., 2008. A hierarchical model for spatial capture-recapture data. Ecology 89 (8), 2281–2289.

Sanathanan, L., 1972. Estimating size of a multinomial population. Annals of Mathematical Statistics 43 (1), 142–152.

Schwarz, C. J., 2004. Estimating the number of walleye in Mille Lacs, Minnesota–Final Report. Tech. rep., Minnesota Department of Natural Resources.

Schwarz, C. J., Arnason, A. N., 1996. A general methodology for the analysis of capture-recapture experiments in open populations. Biometrics 52 (3), 860–873.

Schwarz, C. J., Dupuis, J. A., 2007. A Bayesian approach to the multistate Jolly-Seber capture-recapture model. Biometrics 63, 1015–1022.

Schwarz, C. J., Seber, G. A. F., 1999. Estimating animal abundance: Review III. Statistical Science 14 (4), 427–456.

Seber, G. A. F., 1965. A note on the multiple-recapture census. Biometrika 52 (1/2), 249–259.

Spiegelhalter, D., Thomas, A., Best, N., Lunn, D., 2003. Winbugs user manual. version 1.4 winbugs user manual. version 1.4 (http://www.mrc-bsu.cam.ac.uk/bugs.). Technical report, Medical Research Council Biostatistics Unit.

Stein, D., Tissot, B., Hixon, M., Barss, W., 1992. Fish-habitat associations on a deep reef at the edge of the Oregon continental shelf. Fishery Bulletin 90 (3), 540–551.

Stephens, P. A., Buskirk, S. W., Hayward, G. D., Rio, C. M. D., 2005. Information theory and hypothesis testing: a call for pluralism. Journal of Applied Ecology 42 (1), 4–12.

Stringham, T., Krueger, W., Shaver, P., 2003. State and transition modeling: an ecological process approach. Journal of Range Management 56 (2), 106–113.

Taper, M. L., Staples, D. F., Shepard, B. B., 2008. Model structure adequacy analysis: selecting models on the basis of their ability to answer scientific questions. Synthese 163 (3), 357–370.

Tryjanowski, P., Rybacki, M., Sparks, T., 2003. Changes in the first spawning dates of common frogs and common toads in western Poland in 1978-2002. Annales Zoologici Fennici 40 (5), 459–464.

Tyre, A. J., Tenhumberg, B., Field, S. A., Niejalke, D., Parris, K., Possingham, H. P., 2003. Improving precision and reducing bias in biological surveys: Estimating false-negative error rates. Ecological Applications 13 (6), 1790–1801.

Weir, L. A., Mossman, M. J., 2005. North American Amphibian Monitoring Program (NAAMP). In: Lannoo, M. J. (Ed.), Amphibian Declines: Conservation Status of United States Species. University of California, Berkeley, pp. 307–313.

Weir, L. A., Royle, J. A., Nanjappa, P., Jung, R. E., 2005. Modeling anuran detection and site occupancy on North American Amphibian Monitoring Program (NAAMP) routes in Maryland. Journal of Herpetology 39 (4), 627–639.

White, G. C., Burnham, K. P., 1999. Program mark: survival estimation from populations of marked animals. Bird Study 46 Supplement, 120–139.

White, G. C., Burnham, K. P., Anderson, D. R., 2002. Advanced features of Program MARK. In: Fields, R., Warren, R. J., Okarma, H., Sievert, P. R. (Eds.), Wildlife, land, and people: priorities for the 21st century. Proceedings of the Second International Wildlife Management Congress, 2 July 1999, Godollo, Hungary, pp. 368–377.

Wickham, J. D., Wu, J., Bradford, D. F., 1997. A conceptual framework for selecting and analyzing stressor data to study species richness at large spatial scales. Environmental Management 21 (2), 247–257.

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

Wolda, H., Dennis, B., 1993. Density dependence tests, are they? Oecologia 95 (4), 581–591.

Wu, J., Levin, S. A., 1997. A patch-based spatial modeling approach: conceptual framework and simulation scheme. Ecological Modelling 101 (2-3), 325–346.

Wu, J., Loucks, O. L., 1995. From balance of nature to hierarchical patch dynamics: A paradigm shift in ecology. Quarterly Review of Biology 70 (4), 439–466.

Yamanaka, K., Lacko, L., Lochead, J., Martin, J., Haigh, R., Grandin, C., West, K., 2004. Stock assessment framework for inshore rockfish. Canadian Science Advisory Secretariat Research Document 2004/068, Fisheries and Oceans Canada.

Yoklavich, M., Greene, H., Cailliet, G., Sullivan, D., Lea, R., Love, M., 2000. Habitat associations of deep-water rockfishes in a submarine canyon: an example of a natural refuge. Fishery Bulletin 98 (3), 625–641.

Yoklavich, M. M., Love, M. S., Forney, K. A., 2007. A fishery-independent assessment of an overfished rockfish stock, cowcod (*Sebastes levis*), using direct observations from an occupied submersible. Canadian Journal of Fisheries and Aquatic Sciences 64 (12), 1795–1804.

Zhang, L., Liu, L., You, N., 2005. Estimating population size in logistic capture-recapture models with a known sex ratio. Communications in Statistics -Theory and Methods 34 (1), 37–44.

Appendix A

Appendices for Jolly-Seber with Uncertainty

A.1 Computing N_i^s when Removals are Present

If removals or losses occur, then the $\{N_j^s\}$ estimates need to be adjusted using

$$E\left[N_{j+1}^s\right] = \begin{cases} N\beta_0\pi_0^s & j=0 \\ E\left[N_j^s\right]\phi_j^s + N\beta_j\pi_j^s - L_j^{s*}\phi_j^s & j>0 \end{cases} \quad \text{for } s \in \{M,F\} \tag{A.1}$$

where L_j^{s*} are the sex-specific removals/losses at time j and the term $-L_j^{s*}\phi_j^s$ represent the number of losses expected to have have survived to time $j+1$ if they were not removed from the population at time j. The value of L_j^{s*} may be known if all losses can be sexed, otherwise the sex ratio of removals and/or losses must be estimated.

Let α_j^s be the expected proportion of unknown captures that are of sex s, for $s \in \{M,F\}$. An estimate of α_j^s can be found as

$$E\left[\alpha_i^s\right] \approx \frac{E\left[N_j^s\right]p_j^s(1-\delta_j^s)}{E\left[N_j^s\right]p_j^s(1-\delta_j^s) + E\left[N_j^{s^c}\right]p_j^{s^c}(1-\delta_i^{s^c})} \quad \text{for } s \in \{M,F\} \tag{A.2}$$

where s^c is the complement sex and $N_j^s p_j^s(1-\delta_j^s)$ represents the expected number of animals, at time j, of sex s, caught, but whose sex remains unidentifiable. The expected number of sex-specific losses for time j can then be determined as

$$E\left[L_j^{s*}\right] = L_j^s + L_i^u E\left[\alpha_j^s\right] \quad \text{for}\, s \in \{M, F\} \tag{A.3}$$

where L_j^u is the number of losses with unknown sex and L_j^s is the number of losses whose sex is known. The estimated value L_j^{s*} is then used in (A.1).

A.2 Walleye Data Set

Table A.1: Mille Lacs Lake walleye data set analyzed in the example section. Data represents six sampling occasions over three consecutive years (2002-2004). Sampling occurred twice a year, once in the spring and once in the summer (see Table 3.1). Counts are provided for unique tag histories with counts with negative counts reflecting a loss on capture during the last capture. For a given tag history: `0` - no capture; `U` - capture, with unknown sex designation; `M` - capture, with a male sex designation; and `F` - capture, with a female sex designation.

Tag History	Count	**Tag History**	Count	**Tag History**	Count	**Tag History**	Count
00000F	3	00M00U	-16	0U0UM0	1	M000M0	-190
00000F	-365	00M0F0	13	0UF000	70	M000U0	6
00000M	1	00M0F0	-13	0UF000	-49	M000U0	-11
00000M	-133	00M0M0	1814	0UF00U	-1	M00U00	17
00000U	752	00M0M0	-1155	0UF0F0	4	M00U0U	1
00000U	-648	00M0MM	-1	0UF0F0	-2	M00UM0	3
0000F0	4646	00M0MU	2	0UFU00	4	M00UM0	-3
0000F0	-1949	00M0MU	-2	0UM000	106	M00UMU	-1
0000FF	-9	00M0U0	12	0UM000	-145	M0F000	7
0000FU	15	00M0U0	-53	0UM0M0	12	M0F000	-7
0000FU	-5	00MM00	-6	0UM0M0	-8	M0F0M0	2
0000M0	11637	00MU00	87	0UMU00	4	M0M000	1213
0000M0	-9052	00MU00	-31	0UMUM0	1	M0M000	-526
0000MF	-2	00MUM0	12	0UU000	1	M0M00M	-1
0000MM	-1	00MUM0	-5	0UU000	-6	M0M00U	1
0000MU	14	00MUU0	-1	F00000	3347	M0M00U	-2
0000MU	-22	00U000	164	F00000	-1751	M0M0F0	1
0000U0	120	00U000	-633	F0000F	-1	M0M0F0	-1
0000U0	-497	00U0F0	-1	F0000U	4	M0M0M0	274
000F00	-333	00U0M0	7	F000F0	28	M0M0M0	-64
000M00	-193	00U0M0	-6	F000F0	-10	M0M0MU	2
000U00	3710	00U0U0	-1	F000M0	7	M0M0MU	-2
000U00	-808	00UU00	1	F00F00	-2	M0M0U0	1
000U0F	-2	0F0000	71	F00U00	12	M0M0U0	-1
000U0M	-1	0F0000	-462	F00UF0	-1	M0MM00	-1
000U0U	13	0F00F0	-1	F0F000	56	M0MU00	9
000U0U	-5	0FM000	1	F0F000	-12	M0MU00	-2
000UF0	34	0FM000	-1	F0F00U	1	M0MUM0	1
000UF0	-38	0M0000	48	F0F0F0	3	M0U000	14
000UFU	1	0M0000	-277	F0F0F0	-2	M0U000	-16
000UM0	38	0M00M0	1	F0FU00	2	M0U0M0	2
000UM0	-49	0MM000	1	F0FUF0	1	M0U0M0	-1
000UU0	1	0MM000	-2	F0M000	9	MM0000	-5
000UU0	-4	0U0000	10796	F0M000	-3	MU0000	110
00F000	5042	0U0000	-865	F0M0M0	1	MU0000	-12
00F000	-1631	0U000F	-4	F0M0M0	-1	MU00M0	5
00F00U	7	0U000M	-2	F0U000	2	MU00M0	-2
00F00U	-1	0U000U	15	FF0000	2	MUM000	15
00F0F0	83	0U000U	-3	FF0000	-3	MUM000	-9
00F0F0	-29	0U00F0	41	FF00M0	-1	MUM0M0	5
00F0FU	1	0U00F0	-61	FM0000	1	U00000	129
00F0M0	9	0U00M0	43	FU0000	53	U00000	-728
00F0M0	-5	0U00M0	-84	FU0000	-1	U0000U	1
00F0U0	-3	0U00MF	-1	FUF000	-1	U000F0	1
00FF00	-2	0U00U0	1	FUM000	1	U000M0	1
00FU00	59	0U00U0	-7	M00000	12944	U0M000	8
00FUF0	2	0U0F00	-5	M00000	-16310	U0M000	-1
00FUF0	-1	0U0M00	-1	M0000M	-2	U0M0M0	4
00FUM0	1	0U0U00	135	M0000U	2	U0U000	1
00M000	18928	0U0U00	-1	M0000U	-2	UF0000	-1
00M000	-12299	0U0U0U	1	M000F0	4	UU0000	-1
00M00M	-5	0U0UF0	5	M000F0	-1		
00M00U	10	0U0UF0	-1	M000M0	487		

Appendix B

Chapter 3 Appendices

B.1 Derivation of Joint Capture Probability

Let Y^F and Y^M be two random variables representing the Bernoulli capture outcomes of a male and female in a mating pair on a given capture occasion. Y^F and Y^M will have expected values of p_F and p_M, and standard deviations of σ_F and σ_M respectively. By definition, the correlation coefficient $\rho^{F,M}$ between Y^F and Y^M will be

$$\rho^{F,M} = \frac{\operatorname{cov}(Y^F, Y^M)}{\sigma^F \sigma^M} = \frac{E\left[(Y^F - p^F)(Y^M - p^M)\right]}{\sigma^F \sigma^M}.$$

Using the definition of variance for a Bernoulli experiment, if we expand and take the expectation we can show that,

$$\begin{aligned}\rho^{F,M} &= \frac{E\left[Y^F Y^M - Y^F p^M - Y^M p^F + p^F p^M\right]}{\sqrt{p^F(1-p^F)}\sqrt{p^M(1-p^M)}} \\ &= \frac{E\left[Y^F Y^M\right] - p^F p^M}{\sqrt{p^F(1-p^F)}\sqrt{p^M(1-p^M)}} \\ &= \frac{p^{fm} - p^F p^M}{\sqrt{p^F(1-p^F)}\sqrt{p^M(1-p^M)}}. \qquad \text{(B.1)}\end{aligned}$$

Note we have simply defined p^{fm} as being $E\left[Y^F Y^M\right]$, the joint capture probability. Next we can use (B.1) to define the joint capture probability directly in terms of the individual marginal capture probabilities and the correlation coefficient, giving

$$p^{fm} = \rho^{F,M}\sqrt{p^F(1-p^F)}\sqrt{p^M(1-p^M)} + p^F p^M. \tag{B.2}$$

This allows us to then define the possible joint capture probabilities (Table 4.1), since we know the marginal probability of capture must be p^F and p^M for females and males respectively.

B.2 WinBUGS Code for Fitting $\rho - CJS$ Models using the Individual State-space Formulation

```
########################################################################
# MODEL                                                                #
# State-space formulation of the rho Cormack-Jolly-Seber (rho-CJS)     #
# observations = 0 (non-encountered),     1 (female encounter only),   #
#               2 (male encoutner only), 3 (both encounted)            #
# states = 0 (dead) and 1 (alive)                                      #
########################################################################

model {
  #-----------------
  # SPECIFY PRIORS
  #-----------------

  #--- Survivorship ---
  cphi ~ dbeta(1,1)
  for(t in 2:T){
    phi[t-1] <- cphi
  }

  #--- CAPTURE COMPONENT ---
  # correlation coefficient [-1.0, 1.0]
  ws ~ dbeta(1,1)
  rho <- 2*ws  - 1.0

  # Additive Male Effect
  b1 ~ dnorm(0, 0.5)

  # Intercept
  ctheta ~ dnorm(0, 0.5)

  # Compute Occasion specific capture probability
  for(t in 1:T){
    theta[t]  <- ctheta # Capture intercept on logit scale
```

```
  # Sex and occasion specific capture probabilities (without non-independence)
  logit(pf[t]) <- theta[t]
  logit(pm[t]) <- theta[t] + b1

  #  Joint Capture probabilities for paired individuals
   p00[t] <- 1 - pf[t] - pm[t] + pfm[t]
   pf0[t] <- pf[t] - pfm[t]
   pm0[t] <- pm[t] - pfm[t]
   pfm[t] <- rho*sqrt(pf[t]*(1-pf[t]))*sqrt(pm[t]*(1-pm[t])) + pf[t]*pm[t]
}

#-------------------------
#  STATE PROCESS
#-------------------------
# Model Assumes independent survival
for (i in 1:n){
  # After First Encounter
  for(t in (e[i] +1):T){
      PrX[i,t] <- phi[t-1] * X[i,t-1]
      X[i,t] ~ dbern(PrX[i,t])
  }
}
#-------------------------
#  OBSERVATION PROCESS
#-------------------------
# Observation model assumes a linear pairwise dependence
# occurs in capture outcomes between paired animals.
# Capture outcomes are modeled as multinomial for all types
# (i.e. single only and both alive) using the appropriate
# cell probabilities.

# For each pair/individual by occasion
for (j in 1:nx) {

 # First encounter for a pair. This represents a bit of a work
 # around (at the moment) we the entrances for a pair will be the same.
 epair[j] <- min(e[D[j,2]], e[D[j,3]])

 # Indicator for whether current occasion is first encounter
 f[j] <- equals(epair[j], D[j,1])

 # Overall Conditioning Indicator (occasion <= epair[j])
 # Necessary in case we have obs (i.e. 0's) before entrance.
 c[j] <- step(epair[j] - D[j,1])

 # Indicators for the capture outcome
 obs00[j] <- equals(D[j,4], 1)
```

```
   obsF0[j] <- equals(D[j,4], 2)
   obsM0[j] <- equals(D[j,4], 3)
   obsFM[j] <- equals(D[j,4], 4)

   # Prob obs 00 - 1
   Pr0[j,1] <- f[j]*obs00[j] +
               (1-f[j])*(1-sum(Pr0[j,2:4]))

   # Prob obs F0 - 2
   Pr0[j,2] <- c[j]*(1-f[j])*obsF0[j] +
               c[j]*f[j]*X[D[j,2],D[j,1]]*X[D[j,3],D[j,1]]*pf0[D[j,1]]/(1-p00[D[j,1]]) +
               c[j]*f[j]*X[D[j,2],D[j,1]]*(1 - X[D[j,3],D[j,1]]) +
               (1-c[j])* X[D[j,2],D[j,1]]*X[D[j,3],D[j,1]]*pf0[D[j,1]] +
               (1-c[j])* X[D[j,2],D[j,1]]*(1 - X[D[j,3],D[j,1]])*pf[D[j,1]]

  # Prob obs M0 - 3
   Pr0[j,3] <- c[j]*(1-f[j])*obsM0[j] +
               c[j]*f[j]*X[D[j,2],D[j,1]]*X[D[j,3],D[j,1]]*pm0[D[j,1]]/(1-p00[D[j,1]]) +
               c[j]*f[j]*(1-X[D[j,2],D[j,1]])*X[D[j,3],D[j,1]] +
               (1-c[j])*X[D[j,2],D[j,1]]*X[D[j,3],D[j,1]]*pm0[D[j,1]] +
               (1-c[j])*(1 - X[D[j,2],D[j,1]])* X[D[j,3],D[j,1]]*pm[D[j,1]]
   # Prob obs FM - 4
   Pr0[j,4] <- c[j]*(1-f[j])*obsFM[j] +
               c[j]*f[j]*X[D[j,2],D[j,1]]*X[D[j,3],D[j,1]]*pfm[D[j,1]]/(1-p00[D[j,1]]) +
               (1-c[j])*X[D[j,2],D[j,1]]*X[D[j,3],D[j,1]]*pfm[D[j,1]]

   # Draw Observed as Multinomial Outcome
   D[j,4] ~ dcat(Pr0[j,1:4])
  }
}
############################################################
# DATA
# ‘T’  is the number of encouter occassions
# ‘n’  is the number of entities (individuals and/or mating pairs)
# ‘nx’ total number of observations in ‘D’, typically ‘T’ times ‘n’
# ‘e’  is the vector of first encounters (‘n’ components). Entry ‘n’+1
#      is set to T+1 as a workaround for no mates.
# ‘X’  is a matrix with dimensions ‘n’+1 times ‘T’ containing known
#      alive/dead states for all ‘n’ individuals in the experiment with
#      ‘0’ values for elements before entrance (i.e. e[i]) and ‘NA’ values
#      for alive/dead states after last capture/last known alive status.
#      The final row ‘n’+1 is a workaround for no mates and is set to all 0’s
# ‘D’  is a matrix with dimensions ‘nx’ times 4 where
#       - the first column specifies the current encounter occasion (1,...,‘T’),
#       - the second column specifies the row index in ‘X’ for the female
#         (alone or paired) if present, otherwise the row index ‘n’+1 is
#         specified to indicating no mate.
#       - the third column specifies the row index in ‘X’ for the male
```

```
#         (alone or paired) if present, otherwise the row index ‘n’+1 is
#         specified to indicating no mate.
#       - the fourth column provides the observation as a categorical response
#         1 - No detection,
#         2 - Female only capture
#         3 - Male only capture
#         4 - Both female and male were captured
############################################################
list(
 T = 8,
 n = 1470,
 nx = 5876,
 e = c(
  1, 2, 1, ..., 9
 ),
 X = structure(
  .Data = c(
        1,    1,    1,    1,    1,    1,    1,   NA,
        0,    1,   NA,   NA,   NA,   NA,   NA,   NA,
        1,   NA,   NA,   NA,   NA,   NA,   NA,   NA,
        .
        .
        .
        0,    0,    0,    0,    0,    0,    0,    0
  ),
  .Dim = c(1471, 8),
 ),
 D = structure(
  .Data = c(
        1,    1,  702,    4,
        1,    2,  703,    4,
        1,    3,  704,    4,
        .
        .
        .
        8, 1471, 1345,    1
  ),
  .Dim = c(5876, 4)
 )
)
```

Appendix C

NAAMP Observed Maximum Call Index Frequencies

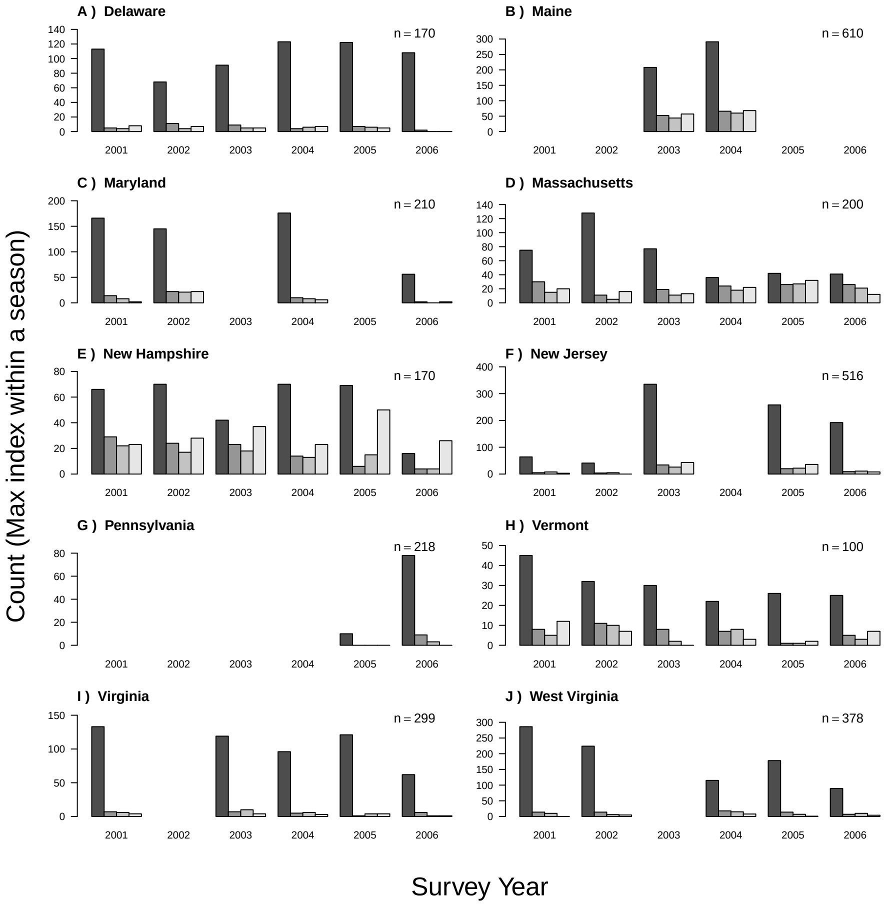


Figure C.1: Empirical summary of maximum observed call indices for *Hyla versicolor* by season and region. Within each season the maximum observed call index was determined and summarized for one of the 10 surveyed regions (states). Ordinal call indices are ranked from lowest (0 - dark grey) to highest (3 - white).

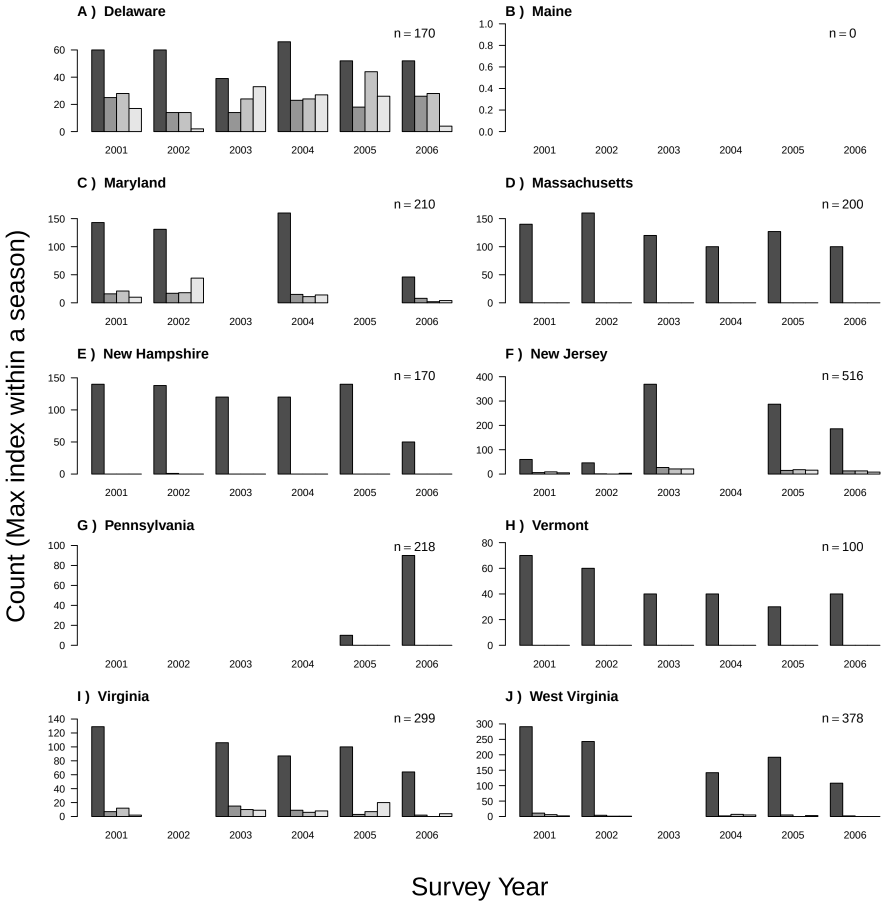


Figure C.2: Empirical summary of maximum observed call indices for *Hyla versicolor* by season and region. Within each season the maximum observed call index was determined and summarized for one of the 10 surveyed regions (states). Ordinal call indices are ranked from lowest (0 - dark grey) to highest (3 - white).