DISSERTATION

METHODS FOR INCORPORATING POPULATION DYNAMICS AND DECISION THEORY IN

CACKLING GOOSE MANAGEMENT

Submitted by

Perry J. Williams

Department of Fish, Wildlife, and Conservation Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2016

Doctoral Committee:

Advisor: William L. Kendall

Mevin B. Hooten
Joel A. Schmutz
Yu Wei

ABSTRACT

METHODS FOR INCORPORATING POPULATION DYNAMICS AND DECISION THEORY IN

CACKLING GOOSE MANAGEMENT

Natural resource management is often challenging due to competing stakeholder values, limited resources, or uncertainty in ecosystem processes. Cackling goose (*Branta hutchinsii minima*) management faces all of these challenges. Many Native Alaskan subsistence hunters living in rural communities rely on the seasonal influx of cackling geese each spring as a source of food security. In addition to food, subsistence hunting provides cultural and spiritual benefits directly tied to the values of many Native Alaskans. When cackling geese migrate out of Alaska, most birds fly to the Willamette Valley in Oregon, and the lower Columbia River Basin in Washington. Here, the cackling goose narrative is much different. In their wintering area, cackling geese browse on private agricultural fields, destroying crops. The destruction of crops causes loss of agricultural yield, and therefore, loss of income. Hazing geese from fields has become a daily requirement for many farmers, consuming time, energy, and financial resources. In previous years, various programs (including the Oregon Goose Control Task Force) provided services to ameliorate cackling goose depredation on private land. These programs were, at least in part, successful at reducing depredation. Funding for these programs was limited, and when financial resources ended, the programs disbanded leaving farmers to address the cackling goose depredation problem once again. The competing interests between farmers and subsistence hunters has resulted in contentious debate related to the management objectives associated with the cackling goose population as the Pacific Flyway Council prepares to revise their cackling goose management plan.

In addition to competing stakeholder values, population dynamics of cackling geese are largely unknown. Uncertainty associated with the population dynamics of cackling geese limits the ability

of decision makers to predict the population response of potential management actions. Thus, even after management objectives have been identified, the question remains as to how to choose management actions that best meet the decision makers' objectives. Harvest regulations, consisting primarily of recreational hunting season length and daily bag limits, are the primary management tool used to control cackling goose abundance. Information on how harvest regulations affect population dynamics will provide a template for choosing management actions that optimize the objectives associated with cackling goose management.

My objective was to address challenges associated with cackling goose management by 1) developing and implementing methods for estimating cackling goose population dynamics based on empirical data, 2) describe and implement quantitative methods for addressing multi-objective problems, and 3) develop methods for incorporating information from population dynamics models with quantitative decision-theoretic methods to identify recreational harvest regulations that optimize a decision maker's objectives.

In Chapter 1, I develop a Bayesian hierarchical integrated population model (IPM) for assessing population dynamics of cackling geese at two temporal scales, inter-annually and intra-annually. I fit the model using two data sources, including 26 years of mark-resight data from neck collared cackling geese and 33 years of relative abundance estimates obtained from aerial surveys. I estimated age-, sex-, season-, and year-specific survival probabilities, annual fecundity, detection probability, relative population size, and age structure of cackling geese during years ranging from 1982–2014. Although the model developed in Chapter 1 was specific to cackling geese, the model can be generalized for other species by reparameterizing the link between vital rates and abundance.

In Chapter 2, I provide a guide to quantitative methods for solving multi-objective problems (i.e., multi-objective optimization). I apply two of the methods described in this chapter to address

the problem of selecting a management goal for cackling geese, given the different competing stakeholder objectives related to cackling goose management. A requirement of multi-objective optimization is the ability to quantify each of the decision maker's objectives into objective functions. I provide examples of objective functions related to cackling goose management. I also developed a web-based tool to allow cackling goose management decision makers to alter inputs for these objective functions.

Chapter 3 is an extension of Chapter 2 in which I demonstrate the flexibility of the multi-objective optimization framework by applying it to a different multi-objective problem; the problem of model selection in which an investigator balances competing objectives of model fit and model complexity. I reconcile several model selection procedures within the multi-objective optimization framework. Additionally, I propose a method for model selection based on post- optimization specification of preferences in which a decision maker first identifies the set of Pareto optimal solutions, and then chooses a model for which to base inference/predictions.

In Chapter 4, I combine the concepts and methods presented in Chapters 1 and Chapter 2 by incorporating population dynamics and multi-objective optimization into a decision theoretic framework to find optimal state-dependent daily bag limits for cackling goose management. Specifically, I describe a Markov decision process (MDP) in which I used a model of population dynamics to inform transition probabilities of moving from abundance n_t to abundance n_{t+1}, for potential management actions including no hunting, restrictive daily bag limits, and liberal daily bag limits. I used multi-objective optimization to inform action-specific reward functions. Finally, I identified an optimal policy using stochastic dynamic programming.

Complex ecosystem processes often necessitate complex models for inferring ecosystem dynamics and quantifying uncertainty. This dissertation provides a framework for 1) using novel

techniques to model population dynamics, and 2) methods to incorporate statistical models explicitly into a decision-theoretic framework to inform wildlife management. As advanced statistical methods for ecological inference continue to proliferate, so should research focused on the development of rigorous methods for using statistical analyses to inform decisions. This dissertation provides flexible, quantitative tools to inform cackling goose management. The utility of this dissertation for cackling goose management is most likely not in the specific output reported within these pages, but the flexibility of the framework to incorporate alternative input from decision makers and stakeholders. One of the primary challenges of any decision problem is quantifying objectives or objective functions. Given these objective functions can be quantified, the methods described here can be used to facilitate management decisions.

In addition to addressing specific issues associated with cackling goose management, this work contributes more generally to the science of decision theory/structured decision making (SDM; Fig. 0.1). Necessary for the SDM process are models that predict the consequences of potential actions a decision maker can take. Bayesian models of population dynamics, such as those developed in Chapter 1 and 4, are important contributions to the modeling toolkit required for SDM (Fig. 0.1). Likewise, Chapter 2 contributes to the SDM analysis toolkit, providing an overview of methods and examples for incorporating preferences among objectives in a multi-objective problem, a necessity in most, if not all SDM processes (Fig. 0.1). Finally, Chapter 4 describes methods for linking the modeling described in Chapter 1, and multi-objective optimization described in Chapter 2, with the other components of SDM. Chapter 4 provides an example of how complex ecological models, objective functions, and optimization routines can be incorporated into wildlife management and/or conservation decisions.

Future work related to this dissertation are suggested in each chapter. I highlight several of those suggestions here. First, in Chapter 1 I note that bias in survival estimates propagates to bias

in fecundity estimates in IPMs. This suggests that unbiased information on fecundity estimates might help account for bias in survival estimates. Simulation studies could help clarify the relationship between bias in survival and fecundity estimates in IPMs. Second, in Chapter 1, I used two data sources to estimate survival of cackling geese. However, a third data source exists for cackling geese and many other harvested animals; band recovery data from harvested birds. An extension of the IPM fit in Chapters 1 and 4 could include band recovery data to estimate survival. Third, computation time potentially limits the flexibility of IPMs; the computation time for the model fit in Chapter 1 exceeded 900 hours. A computationally efficient alternative to the IPM fit in Chapters 1 and 4 would be to conduct a piece-wise analysis of each data source. That is, first fit a model containing one data set; then use the information from the first analysis to facilitate fitting a model to the second data set (e.g., using the results of the first analysis as a prior distribution in the second analysis). A comparison of the results from each approach would help future data analysts decide if the information that resulted from fitting the IPM is worth the additional computational cost. Fourth, in Chapter 4 I identified an optimal management policy of the MDP using an objective function identified using *a priori* specification of stakeholder preferences. An extension of the work in Chapter 4 is to identify an optimal management policy of the MDP based on *a posteriori* selection of Pareto optimal solutions. This would require using stochastic dynamic programming to optimize the MDP for many different objective functions, each objective function implying a different set of stakeholder weights (preferences). The benefits of this multi-objective MDP with *a posteriori* specification of preferences is that decision makers could select an optimal policy without explicitly selecting stakeholder weights. Each Pareto policy could be presented at a stakeholder meeting, and stakeholders could potentially agree on a policy to implement. Finally, one of the assumptions of MDP is that the current state (e.g., population size) is observable without error. Partially Observable MDPs (POMDPs) relax this assumption. Comparing optimal

management policies calculated using MDPs and POMDPs would provide information on the robustness of the optimal solution with respect to the assumption of correctly observing the current state. The observation model in hierarchical models could provide a framework for estimating the observation error required in POMDPs, and therefore, POMDPs are potentially a straightforward extension of Chapter 4.
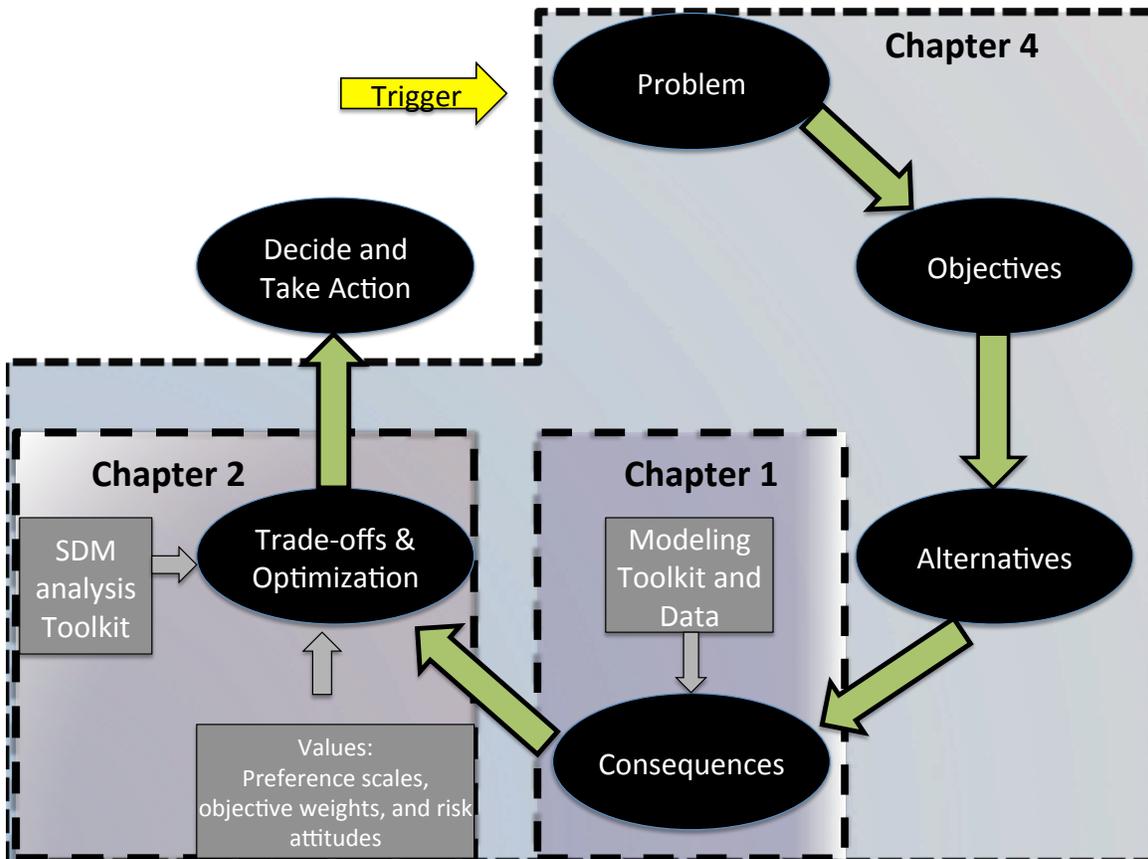
FIGURE 0.1. Diagram of the structured decision making process modified from www.fws.gov. Chapter 1 is concerned with developing and implementing methods for modeling population dynamics. Chapter 2 is concerned with quantitative decision-making methods that incorporate stakeholder preferences into the decision making process. Not shown is Chapter 3, which is an extension of Chapter 2 for model selection, a common multi-objective problem in wildlife research and management. Finally, Chapter 4 combines methods developed in Chapter 1 and 2 and describes methods for incorporating other remaining components of the structured decision making process.

TABLE OF CONTENTS

LIST OF TABLES

A Bayesian Integrated Population Model for Assessing Cackling Geese Demography at Two Temporal Scales

1.1. Summary

Seasonal and annual population dynamics for a species are critical mechanisms of population change. Integrated population models (IPMs) permit the simultaneous analysis of data on population vital rates and abundance. Most of the guidance for development of IPMs have occurred at the inter-annual scale. I developed a Bayesian hierarchical IPM that allows for simultaneous inference on both intra- and inter-annual population dynamics. I applied the model to data collected on cackling geese (*Branta hutchinsii minima*). Cackling geese are an important source of food security for Native Alaskan subsistence hunters during spring and fall migration. They winter in the Lower Columbia River Basin, Washington, and the Willamette Valley, Oregon where they depredate agricultural crops causing yield loss. Management of cackling geese is concerned with balancing abundance such that the population provides adequate subsistence hunting opportunities, but limits crop depredation. To better understand the mechanisms driving population change to facilitate efficient management, I estimated age-, sex-, season-, and year-specific survival probabilities. Additionally, I estimated annual fecundity, detection probability, relative population size, and age structure of cackling geese from 1982–2014 using 33 years of abundance data and 10,428 collar-marked geese. Mean seasonal survival of two-month periods was lowest for juveniles (0.90), then sub-adults (0.94), then adults (0.95). Mean annual survival was 0.54, 0.69, and 0.75 for juveniles, sub-adults, and adults, respectively. Sex did not appear to affect additive annual survival.

Mean fecundity of cackling geese was 1.67 birds per breeding pair for the years 1982–2004, and 2.16 for the years 1982–2004 and 2011–2013. The mean percentage of pre-breeding adults in the population was 71%. Long-term trends appeared to be driven by the number of adults in the population while short-term (annual) variation appeared to be driven by fecundity. The IPM developed in this paper can be easily generalized to other life-cycles by re-parameterizing the link between vital rates and abundance.

1.2. INTRODUCTION

Population dynamics are affected by temporal variation in vital rates that result from changes in environmental, physiological, and/or anthropogenic conditions. Vital rates are often assessed at two temporal scales: intra-annually (e.g., seasonally) or inter-annually (Ims and Andreassen 2000, Sillett and Holmes 2002, Folk et al. 2007). Identifying differences in vital rates at each scale is useful for elucidating mechanisms that limit and regulate populations (Fretwell 1972, Caughley 1977, Caswell 2001). Assessing intra-annual variation in vital rates helps identify life stages, seasons, and geographical regions within an annual cycle that have the largest effect on population dynamics (Flockhart et al. 2015, Hostetler et al. 2015). Assessing inter-annual variation in vital rates allows estimates of demographic stochasticity and population change due to larger scale temporal processes (e.g., policy change, climate change; Lande et al. 2003). Examination of each of the two temporal scales separately provides valuable information, but only a partial picture of population dynamics compared to an evaluation of both scales simultaneously.

Hierarchical models are often fit with multiple data sources (Hanks et al. 2011). Integrated population models (IPMs) are a specific class of hierarchical models for which data on vital rates and abundance are simultaneously analyzed (Besbeas et al. 2002, Schaub and Abadi 2011). Abundance data are realizations of a demographic process defined by specific vital rates and therefore

contain information on vital rates. Incorporating abundance data with other vital rate data permits precise parameter estimation and estimation of some unobservable vital rates (Schaub and Abadi 2011). IPMs are a relatively new modeling framework for wildlife populations and most guidance on their development and implementation have occurred at the inter-annual temporal scale (Besbeas et al. 2002, Brooks et al. 2004, Schaub et al. 2007, Abadi et al. 2010, Cave et al. 2010, Schaub and Abadi 2011, Kéry and Schaub 2012). IPMs for the intra-annual scale are much less common. In a review of models used to examine intra-annual population dynamics (i.e., full-annual-cycle population models), Hostetler et al. (2015) were unable to find any published examples of IPMs at the intra-annual scale. This observation was not surprising because one or more data sets required to fit IPMs often occur at the annual time scale. It is common, however, that one data set occurs at the annual temporal scale and one occurs at a finer temporal scale. In cases with disparate temporal resolutions of data, one method to fit an IPM is to align the data by dissolving the resolution of the finer-scale data to match the scale of the coarser-scale data (e.g., Davis et al. 2014). This is advantageous in that the benefits of IPMs for parameter estimation can be realized and standard models, software, and code developed for IPMs at the inter-annual temporal scale can be implemented. It is disadvantageous in that it results in loss of information from the data collected at finer resolution. Appropriately scaling different temporal data sources for inference is important because mechanisms driving the temporal dynamics of a process at one scale might not be important or are inoperative at another scale (Gotway and Young 2002). Identifying methods for using disparate data or information to make inference is a general problem in ecological, statistical, and scientific investigation (Cressie 1996, Jelinski and Wu 1996, King 1997, Gotway and Young 2002). Methods to fit disparate temporal scales of data within an IPM framework will permit the advantages of IPMs for estimating parameters and allow the simultaneous evaluation of multiple temporal scales. I demonstrate the use of disparate temporal data in an IPM framework for cackling geese (*Branta*

hutchinsii minima) for which survival data were collected intra-annually and abundance data were collected inter-annually.

Cackling geese are migratory birds that nest on the coastal plain of the Yukon-Kuskokwim (YK) Delta, Alaska and primarily winter in the Willamette Valley and Lower Columbia River Valley in Oregon and Washington (Fig. 1.1; Sedinger and Bollinger 1987, Pacific Flyway Council 1999). The cackling goose population has experienced substantial variation in size and distribution since the mid-1960s. Autumn counts in northern California ranged from $> 350,000$ birds in the mid-1960s to $< 30,000$ birds in 1983 (Fig. 1.2; O'Neill 1979, Raveling 1984, King and Derksen 1986). More recently, the U.S. Fish and Wildlife Service (USFWS) estimated abundance at 281,000 birds in 2014 (Fig. 1.2). The cackling goose winter distribution shifted in the mid-1980s. Formerly, most birds wintered in the Central Valley in California (Nelson and Hansen 1959, Sedinger and Bollinger 1987). After 1993, the majority of the population wintered in the Willamette Valley and Lower Columbia River Valley (Pacific Flyway Council 1999). The extreme changes in abundance were thought to be linked to harvest levels at both ends of their migratory route (Raveling 1984, Raveling et al. 1992). Mechanisms driving changes in distribution are unknown, but hypothesized to be related to drought in California between 1986–1992 or expansion of the commercial grass industry in Oregon (Pacific Flyway Council 1999).

Cackling geese represent an important food source for subsistence hunters in Alaska. They are regularly the second-most harvested avian species during subsistence hunts (Naves 2010; 2011; 2012). The cackling goose subsistence hunt has historically occurred from when birds arrive in the spring (Apr–May) through the end of August, with a 30-day legal cessation implemented after nest initiation. In addition to subsistence hunting, recreational hunting of cackling geese typically occurs from mid-September in Alaska to mid-March in parts of Oregon and Washington. In their winter range, cackling geese graze on private agricultural land, causing yield loss for farmers. The

Pacific Flyway Council, in conjunction with the USFWS, is currently reviewing its cackling goose management plan to identify solutions to balance the differing stakeholder interests associated with cackling geese. Due to these competing interests of stakeholders and observed changes in cackling goose abundance and distribution, understanding the intra- and inter-annual temporal variability in population dynamics is imperative for developing an efficient management plan that addresses the concerns of Native Alaskan subsistence hunters, Oregon and Washington farmers, recreational hunters, and other non-consumptive benefactors. Examining life history traits at two temporal scales helps provide a more precise estimate of the anthropogenic effect on population size of cackling geese.

My objectives were to 1) develop an IPM framework that permits inference at two temporal scales using a simple reconciliation of seasonal survival estimates to annual survival estimates and link these estimates to annual abundance data, 2) apply the model to cackling geese data with abundance collected annually and mark-resight data collected seasonally, and 3) use the results of the model to examine the temporal variation in vital rates of cackling geese. Although the IPM I developed was specific to the life cycle of cackling geese, the framework applies to similar types of data for any species or life cycle.

1.3. Methods

1.3.1. An IPM for inference at two temporal scales. Kéry and Schaub (2012) described a three step process to build an integrated population model. First, develop a population model that links vital rates to changes in population size. Second, describe the likelihood for each set of data. Third, construct the joint likelihood. I follow this framework to develop a Bayesian hierarchical integrated population model to assess temporal dynamics at both intra- and inter-annual scales. I develop the model for a species with 3 age classes (juveniles, sub-adults, adults), annual

abundance data (y_t) collected pre-breeding, and breeding occurring by adults. Later, I relax the assumption that breeding only occurs as adults, but start with this assumption for simplicity in demonstration. I use mark-resight data in which individual animals have been captured, marked, and released and data exist on whether marked animals have been resighted in seasons $s = 1, ..., S$ in year $t = 1, ..., T$. The data on seasonal survival for each individual i ($\boldsymbol{m}_i$) is a vector of length $J = S \times T$ of 1s and 0s, depending on if individual i was seen or not seen, respectively, in the season-year combination $j = s + (t - 1) \times S$. I used this life cycle and sampling design as an example because of its applicability to the cackling goose population to which I subsequently fit the model. This life cycle can be generalized by appropriately reparameterizing the link between abundance and demographic vital rates.

1.3.1.1. *Step 1: Linking population size to vital rates.* The deterministic link between annual abundance and annual vital rates can be described generally as:

$$(1.1) \qquad\qquad \boldsymbol{n}_t = \boldsymbol{A}\boldsymbol{n}_{t-1} \,, t = 2, ..., T,$$

where $\boldsymbol{n}_t$ is a vector with elements consisting of the abundance of each age/stage class at time t, and $\boldsymbol{A}$ is a matrix linking $\boldsymbol{n}_{t-1}$ to $\boldsymbol{n}_t$ through survival at time $t - 1$ ($\boldsymbol{\phi}_{t-1}$), fecundity at time $t - 1$ (f_{t-1}), and potentially other vital rates (e.g., breeding propensity, clutch size; Caswell 2001). For a population with three age classes (juveniles, sub-adults, adults), pre-birth data on $\boldsymbol{n}_t$, and breeding occurring as adults, (1.1) can be described as

$$\begin{bmatrix} n_{sub,t} \\ n_{ad,t} \end{bmatrix} = \begin{bmatrix} 0 & \frac{f_{t-1}\phi_{juv,t-1}}{2} \\ \phi_{sub,t-1} & \phi_{ad,t-1} \end{bmatrix} \times \begin{bmatrix} n_{sub,t-1} \\ n_{ad,t-1} \end{bmatrix},$$

where $n_{sub,t}$ and $n_{ad,t}$ are the number of sub-adults and adults, respectively, in the pre-birth popu-lation at time t, f_t is the year-specific fecundity of a pair of adult animals, and $\phi_{a,t}$ are age-specific annual survival probabilities $a \in \{$juv, sub, ad$\}$. $n_{sub,1}$ and $n_{ad,1}$ can be either fixed or assigned prior distributions based on available information. Suppose now seasonal data exist to estimate $\phi_{s,a,t}$, where $s = 1, ..., S$ seasons or partitions of a year. Annual survival is a function of seasonal survival. That is: $\phi_{a,t} = f(\phi_{s,a,t})$. Assuming survival is independent among seasons, and $\phi_{s,a,t}$ is estimated for all $s \in t$, f is simply the product function with respect to s:

$$(1.2) \qquad \phi_{a,t} = \prod_{s=1}^{S} \phi_{s,a,t}.$$

If $\phi_{s,a,t}$ are estimated using Markov chain Monte Carlo (MCMC) methods, $\phi_{a,t}$ is a derived param-eter. The equivariance property of MCMC derived parameters (Hobbs and Hooten 2015) permits estimation of the posterior distribution of $\phi_{a,t}$ (i.e., $[\phi_{a,t}|\boldsymbol{y}, \boldsymbol{M}]$, where $\boldsymbol{y}$ is the vector of abundance data and $\boldsymbol{M}$ is a matrix of mark-resight data with elements $m_{i,j}$). Point estimates and credible in-tervals can then be obtained from the posterior distribution. Now suppose $\phi_{s,a,t}$ are estimated from data using a generalized linear regression model with logit link function: $\text{logit}(\boldsymbol{\phi}) = \boldsymbol{X\beta}$, where $\boldsymbol{X}$ is a design matrix of indicator variables that identify the season, age, year and potentially other individual or population-level covariates of each resighting for each individual, and the parameters $\boldsymbol{\beta}$ describe the effect of the covariates on survival. I chose the logit function to map the support of $\boldsymbol{X\beta}$ (i.e., $\mathbb{R}^1$) to (0–1); any link function meeting this criteria could be used. Annual survival for each time and age class can be derived from the function: $\phi_{a,t} = \prod_{s=1}^{S} \text{expit}\{\beta_0 + \beta_{1,s} + \beta_{2,a} + \beta_{3,t}\}$, where $\beta_{1,1} = 0$, and is accounted for in the intercept. It follows that $\boldsymbol{A}$ can be expressed as:

$$\boldsymbol{A} = \begin{bmatrix} 0 & \frac{1}{2} f_{t-1} \prod_{s=2}^{S} \text{expit}\{\beta_0 + \beta_{1,s} + \beta_{2,juv} + \beta_{3,t}\} \\ \prod_{s=1}^{S} \text{expit}\{\beta_0 + \beta_{1,s} + \beta_{2,sub} + \beta_{3,t}\} & \prod_{s=1}^{S} \text{expit}\{\beta_0 + \beta_{1,s} + \beta_{2,ad} + \beta_{3,t}\} \end{bmatrix},$$

providing a simple link between seasonal survival, annual survival, and population size for the IPM framework. Alternatively, the life cycle could be represented using $\mathbf{A} = \mathbf{B}_S \cdots \mathbf{B}_2 \mathbf{B}_1$, where S are the number of seasons, and:

(1.3)

$$\mathbf{B}_1 = \begin{bmatrix} 0 & 0 & \frac{f_t}{2} \\ 0 & \phi_{sub,1,t} & 0 \\ 0 & 0 & \phi_{ad,1,t} \end{bmatrix}, \mathbf{B}_i = \begin{bmatrix} \phi_{juv,i,t} & 0 & 0 \\ 0 & \phi_{sub,i,t} & 0 \\ 0 & 0 & \phi_{ad,i,t} \end{bmatrix}, \mathbf{B}_S = \begin{bmatrix} 0 & 0 & 0 \\ \phi_{juv,S,t} & 0 & 0 \\ 0 & \phi_{sub,S,t} & \phi_{ad,S,t} \end{bmatrix},$$

where $i = 2, ..., S-1$, and $\phi_{a,s,t} = \text{expit}\{\beta_0 + \beta_{1,s} + \beta_{2,a} + \beta_{3,t}\}$, for all age classes a. (Caswell and Fujiwara 2004). Annual survival is a derived parameter from the estimated posterior distributions of $\boldsymbol{\beta}$.

1.3.1.2. *Step 2: Defining the likelihood of each data set.* Eq. (1.1) is deterministic and thus far not linked to random realizations of observed data. To describe the stochastic nature of the data relative to the observation error and process error that occur, probability distributions are assumed. For this example, I assume abundance data collected pre-breeding are normally distributed with mean μ_t, and variance σ^2. I assumed $\mu_t = \boldsymbol{n}_t'\mathbf{1}$, where $\mathbf{1}$ represents a vector of 1s with length equal to the number of age classes in $\boldsymbol{n}_t$ (μ_t is the sum of adults and sub-adults at time t). Thus: $y_t \sim \text{N}(\boldsymbol{n}_t'\mathbf{1}, \sigma^2)$, where σ^2 represents the observation variance. Other probability models could be used instead of the normal distribution using $\boldsymbol{n}_t'\mathbf{1}$ as the mean structure and appropriately describing the variance (e.g., $y_t \sim \text{Poisson}(\lambda_t)$, mean and variance = $\lambda_t = \boldsymbol{n}_t\mathbf{1}'$). Note that these distributions assume abundance data can be over-counted as well as under-counted. If the true population size cannot be over-counted and data exist to estimate detection probability (false negatives), the binomial distribution could be used. If an investigator is unsure about the exact form of $\boldsymbol{A}$, and wants to include "demographic stochasticity" (*sensu* Kéry and Schaub 2012), a probability distribution could be assigned to $\boldsymbol{n}_t$. For example: $\log(\boldsymbol{n}_t) \sim \text{N}(\log(\boldsymbol{An}_{t-1}), \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a

covariance matrix for different age classes representing variability or covariance not captured by the functions comprising $\boldsymbol{A}$.

A common method to model the random nature of seasonal survival probability and detection probability is using a zero-inflated Bernoulli model (i.e., Cormack-Jolly-Seber model). Assume:

$$
(1.4) \qquad m_{i,j} \sim \begin{cases} 0 & , z_{i,j} = 0 \\ \\ \text{Bern}(p_j) & , z_{i,j} = 1 \end{cases},
$$

where $z_{i,j}$ is the latent, true state (alive $= 1$, dead $= 0$) of individual i during resight occasion j. p_j is the resight probability of a marked individual during resight event j. Assume p_j is estimated using a generalized linear model with a logit link: $\text{logit}(\boldsymbol{p}) = \boldsymbol{W}\boldsymbol{\alpha}$, where $\boldsymbol{W}$ is a design matrix for the detection probability parameters $\boldsymbol{\alpha}$. We want to estimate the posterior distributions of $\boldsymbol{\alpha}$. The true latent state of each animal can be modeled as:

$$
(1.5) \qquad z_{i,j} \sim \begin{cases} 0 & , z_{i,j-1} = 0 \\ \\ \text{Bern}(\phi_{i,j}) & , z_{i,j-1} = 1 \end{cases},
$$

where the first case ensures that if an individual was dead during the previous resight event it remained dead during the present resight event. The second case assumes that if an individual was alive during the previous resight event it was alive with probability equal to the season and time specific survival probability $\phi_{i,j}$. Given these models for the random nature of the data, we can specify the likelihood for each process. The likelihood of the abundance data, assuming normal distributions for observation data and process variance is

$$
L(\boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{f} | \boldsymbol{y}) = \prod_{t=1}^{T} [y_t | \mu_t][\mu_t | \boldsymbol{\beta}, \boldsymbol{f}],
$$

9

where $\boldsymbol{\beta}$ are the parameters used to estimate the seasonal survival rates in (1.3). The likelihood of the mark-resight data, assuming zero-inflated Bernoulli models, is

$$L(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{Z}|\boldsymbol{M}) = \prod_{i=1}^{n}\prod_{j=2}^{J}[m_{i,j}|\boldsymbol{\alpha}]^{z_{i,j}}I_{m_{i,j}=0}^{1-z_{i,j}}[z_{i,j}|\boldsymbol{\beta}]^{z_{i,j-1}}I_{z_{i,j}=0}^{1-z_{i,j-1}},$$

where $\boldsymbol{Z}$ is a matrix (with elements $z_{i,j}$) of the true, unknown latent state of survival, $\boldsymbol{M}$ is a matrix of the mark-resight data, and I is an indicator variable that equals 1 when its subscript is true (e.g., $I_{m_{i,j}=0} = 1$ for all cases when $m_{i,j} = 0$). It is evident that both likelihoods depend on $\boldsymbol{\beta}$, and thus $\boldsymbol{\beta}$ will be estimated from each set of data, despite data occurring at disparate temporal scales.

1.3.1.3. *Step 3: Formulate the joint likelihood and posterior distribution.* Given the likelihood functions described in step 2, the joint likelihood of the parameters given the data is:

$$L(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{f}, \sigma^2|\boldsymbol{y}, \boldsymbol{M}) =$$

$$\prod_{t=1}^{T}[y_t|\mu_t, \sigma^2][\mu_t|\boldsymbol{\beta}, \boldsymbol{f}]\prod_{i=1}^{n}\prod_{j=2}^{J}[m_{i,j}|\boldsymbol{\alpha}]^{z_{i,j}}I_{m_{i,j}=0}^{1-z_{i,j}}[z_{i,j}|\boldsymbol{\beta}]^{z_{i,j-1}}I_{z_{i,j}=0}^{1-z_{i,j-1}}.$$

Assuming $n_{ad,1}$ and $n_{sub,1}$ are fixed and known, defining priors for $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, and $\boldsymbol{f}$ completes the information required to specify the joint posterior distribution:

$$[\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{f}, \sigma^2|\boldsymbol{y}, \boldsymbol{M}] \propto$$

(1.6)
$$\prod_{t=1}^{T}[y_t|\mu_t, \sigma^2][\mu_t|\boldsymbol{\beta}, f]\prod_{i=1}^{n}\prod_{j=2}^{J}[m_{i,j}|\boldsymbol{\alpha}]^{z_{i,j}}I_{m_{i,j}=0}^{1-z_{i,j}}[z_{i,j}|\boldsymbol{\beta}]^{z_{i,j-1}}I_{z_{i,j}=0}^{1-z_{i,j-1}}[\boldsymbol{\beta}][\boldsymbol{\alpha}][\boldsymbol{f}][\sigma^2].$$

From the joint posterior distribution, the full-conditional distributions of $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, $\boldsymbol{Z}$, $\boldsymbol{f}$, and σ^2 can be identified (see Appendix A for full-conditional distributions).

1.3.2. APPLICATION TO CACKLING GEESE DATA. In this section, I apply a specific formulation of the *IPM for inference at two temporal scales* to two data sources on cackling geese:

abundance data and mark-resight data. Both data sources have been used independently to inform management decisions for cackling goose harvest regulations (Pacific Flyway Council 1999) and a subset of the mark-resight data was used to obtain inference for seasonal and annual survival for years 1982–1988 (Raveling et al. 1992). However, the data have not been combined in a simultaneous analysis of data sources or temporal scales within an IPM framework. I describe the current application of the data for management and the benefits gained from combining the data in a single analysis.

1.3.2.1. *Abundance data.* The USFWS, Division of Migratory Birds, Alaska have conducted aerial surveys of geese nesting in the coastal plain of the YK Delta each spring since 1985. For a detailed description of the study design and results see U.S. Fish and Wildlife Service (2014). Briefly, the study design consisted of pilots flying an amphibious aircraft (either Cessna 206 or Quest Kodiak) along pre-determined transects at 145–170 km/h at a height of 30–45 m. All surveys occurred between 29 May–24 June. The timing of the surveys was designed to coincide with goose egg-laying and incubation. The transects extended from the coastline of the YK Delta, to approx. 50 km inland to the edge of the upland tundra. The distance between transects was either 1.6, 3.2, 6.4, or 12.9 km, depending on the strata. Strata were delineated across the coastal plain of the YK Delta based on perceived goose abundance; strata with higher perceived abundance received more sampling effort (i.e., they had less distance between aerial transects). The total survey area encompassed $> 12,000$ km^2 of tundra wetlands, with approx. 2,350 km of flight distance flown on 235 transects. An observer counted and recorded the number of cackling geese within transects during surveys. The observer also identified group size (i.e., single, pair, or ≥ 5 birds). Observers' counts were calibrated to account for unobservable female geese (presumably on nests) by multiplying the number of observed single birds by two. This assumed single birds

were paired with unobservable females. The calibrated count (i.e., indicated total birds; ITB) was

(1.7) $$y_t = 2(\text{single birds}_{i,t} + \text{pairs}_{i,t}) + \text{birds in flocks}_{i,t},$$

where $t = 1, ..., T$ (corresponding to: 1985,...,2014) were the years. Note that it was possible to over-estimate and under-estimate abundance during the surveys, and support of y_t includes positive real numbers. The USFWS, Pacific Flyway has used the equation $\tilde{n}_t = 3.35y_t$ to estimate annual population size and the three-year average: $\delta_t = \frac{1}{3}(\tilde{n}_t + \tilde{n}_{t-1} + \tilde{n}_{t-2})$ for establishing harvest regulations for cackling geese in year t (Pacific Flyway Council 2015). The value 3.35 was obtained from an additional data source described in the *scaling relative abundance to true abundance* section, below. The cackling goose hunting season is closed when δ_t decreases below 80,000 birds and subsequently re-opened when δ_t increases above 110,000 birds. The current population objective is 250,000 birds (Pacific Flyway Council 1999).

Prior to 1985, cackling goose abundance was estimated using peak fall counts in the Klamath Basin and the Sacramento Valley (Fig. 1.2; O'Neill 1979). This included first counting Taverner's geese (*Branta hutchinsii taverneri*) and cackling geese in an aircraft in their winter areas. Abundance of cackling geese and Taverner's geese was then partitioned using ground-based estimates of the proportion of cackling geese relative to Taverner's geese. I used the peak fall count estimates divided by 3.35 for cackling geese data on y_t from 1982–1984, because the current surveys to calculate y_t began in 1985.

1.3.2.2. *Mark-resight data.* The U.S. Geological Survey (USGS), the USFWS, the University of California, Davis, Alaska Department of Fish and Game, California Department of Fish and Game, Oregon Department of Fish and Wildlife, and Washington Department of Fish and Wildlife captured, marked and resighted cackling geese from 1982–2004 and from 2011–2014. From 1982–2004, USGS captured birds on the YK Delta by driving them into corrals while they were flightless

in late-July–early-August. They marked each of the captured birds with a USFWS/USGS aluminum leg band and a yellow plastic neck band. The neck bands had a three-digit alpha-numeric code that allowed observers to uniquely identify each marked bird from a distance. Personnel from the USGS, the California Department of Fish and Game, the University of California, Davis, and the USFWS resighted birds in 1982–1983, 1985–2004 in the following time/location combinations: late summer on the YK Delta (Aug–Sep), fall (Oct–Nov) in Oregon, Washington, and California, winter (Dec–Feb) in Oregon, Washington, and California, spring (Mar–Apr) in Oregon, Washington, and California, early summer on the YK Delta, and during subsequent capture occasions on the YK Delta (late-Jul–Aug; Fig. 1.3). A total of 9,085 birds marked with neck collars between 1982–2003 were used for this analysis. Of these birds, 4985, were juveniles (2,348 F, 2,637 M), and 4,100 were either sub-adults or adults but indistinguishable (2,104 F, 1,996 M). Of the 4,100 sub-adults and adults I omitted the first year of resight data from analysis. Doing so insured that each bird could be classified as an adult. From 2011–2013 cackling geese were captured on both the YK Delta, Alaska and in Oregon. Rocket nets were used in Oregon. A total of 1,343 captured birds from this time-period were used for analyses (672 F, 671 M; Appendix B). Of these, 50 were juveniles (18 F, 32 M) and 1,293 were either sub-adults or adults (654 F, 639 M). These 1,343 birds were fitted with neck collars and aluminum leg bands. Newly captured adults were not omitted from the first year of analysis from these data because preliminary analyses on the mark-resight data from 1982–2004 suggested little difference in survival between sub-adults and adults. Further, omitting the data during the shorter time frame between 2011–2014 would have limited the inference on adult survival in 2011. Several (98) birds captured during 2011–2013 were survivors from the 1982–2004 banding period. These surviving birds could be classified immediately as adults. During 2011–2014 birds were observed in their wintering areas in California, Oregon, and Washington.

1.3.2.3. *Data Analysis.* To analyze the cackling geese data at two temporal scales, I developed and applied a Bayesian hierarchical IPM following the methods described in *An IPM for inference at two temporal scales*. I assumed the aerial abundance data from $t = 1, ..., 33$ (corresponding to 1982–2014) calculated using (1.7) was normally distributed, with mean $\mu = n_t$, and variance σ^2. The latent variable n_t represented pre-breeding abundance. I assumed an inverse-gamma distribution for the variance parameter σ^2. I parameterized the inverse-gamma distribution such that the mean and variance equaled $5,000^2$. I chose $5,000^2$ because I assumed the value was large enough to realistically reflect the uncertainty of y_t, but small enough to inform other components of the model (e.g., fecundity). The maximum annual standard error estimated for y_t was 5,202 (mean = 2,976; T. Sanders, personal communication). Although it was possible for y_t to be larger than true abundance based on (1.7), it is more likely that true abundance was underestimated due to imperfect detection during aerial surveys. Thus, I assumed n_t was an estimate of relative abundance. I defined n_t by the process:

$$n_t = n_{sub,t} + n_{ad,t},$$

(1.8) $$n_{sub,t} = \frac{n_{sub,t-1}}{2}\rho_{sub,t-1}f_{sub,t-1}\phi_{juv,t-1} + \frac{n_{ad,t-1}}{2}\rho_{ad,t-1}f_{ad,t-1}\phi_{juv,t-1},$$

$$n_{ad,t} = n_{sub,t-1}\phi_{sub,t-1} + n_{ad,t-1}\phi_{ad,t-1},$$

where $n_{sub,t}$ were the number of sub-adult birds in year $t = 1, ..., 33$, $n_{ad,t}$ were the number of adult birds, $\rho_{sub,t}$ and $\rho_{ad,t}$ were breeding propensies for each age, $f_{sub,t}$ and $f_{ad,t}$ were fecundity estimates for each age (a composite reproductive statistic defined as the mean number of birds per pair that survived until summer banding), and $\phi_{juv,t}$, $\phi_{sub,t}$, and $\phi_{ad,t}$ were annual survival probabilities. I prescribed 1-year-old birds to be sub-adults and birds $\geq$ 2-years-old to be adults. The age-specific variables for breeding propensity and fecundity in (4.4) were not individually identifiable from the data or strength of the prior information. Therefore, I reparameterized $n_{sub,t}$

14

such that:

$$n_{sub,t} = \frac{n_{ad,t-1}}{2} f_{t-1} \phi_{juv,t-1},$$

$$f_{t-1} = \frac{n_{sub,t-1} \rho_{sub,t-1} f_{sub,t-1}}{n_{ad,t-1}} + \rho_{ad,t-1} f_{ad,t-1}.$$

The value f_{t-1} was identifiable in the IPM framework (assuming age structure at time 1 is known) and $f_{t-1} \times \frac{n_{ad,t-1}}{2}$ is interpretable as the number of juveniles produced in year $t-1$ by sub-adults and adults that survived to summer banding. As $\rho_{sub,t-1} f_{sub,t-1} \to 0$, $f_{t-1} \to f_{ad,t-1} \rho_{ad,t-1}$. That is, if either sub-adult breeding propensity or sub-adult fecundity approach zero, the estimate of f_{t-1} approaches the average number of juveniles produced that survive to the banding period, per adult pair. I set $n_{ad,1} = y_1 - n_{sub,1} = y_1 - 0.2y_1$ for year 1. The value 0.2 was calculated by estimating the average age distribution for adults and sub-adults for years 1983–2014 for three runs of the MCMC algorithm, each with different starting values of $n_{ad,1}$ and $n_{sub,1}$ ranging from 1) equal age distribution to 2) 90% adults to 3) 90% sub-adults. I averaged over different starting values to improve robustness of the choice of $n_{ad,1}$ and $n_{sub,1}$. I assessed the sensitivity of the choice of $n_{ad,1}$ and $n_{sub,1}$ by choosing different proportions of y_1 for sub-adults and adults. I examined the following proportions: (0,1), (0.2, 0.8),(0.4, 0.6),...,(1,0), and examined age distributions through time, starting at 1982. The choice of $n_{ad,1}$ and $n_{sub,1}$ had only a small effect on the first two years, and no visual effect after 1985. I assumed a uniform prior for fecundity, $f_t \sim \text{Unif}(0, 15)$. I assumed the demographic stochasticity of the process was accounted for in the uncertainty in each parameter and therefore did not assume an additional probability model for the process in (4.4).

To model the mark-resight data $m_{i,j}$, for $i = 1, ..., 10,428$ individuals, $t = 1, ..., 33$, years, $s = 1, ..., 6$ seasons, and $j = 1, ..., 192$ resight occasions ($j = s + (t - 1) \times S$), I used the state-space formulation of the Cormack-Jolly-Seber model described in (1.4), where $z_{i,j}$ was the

latent, true state (alive $= 1$, dead $= 0$) of cackling goose i during resight occasion j. p_j was the resight probability of a marked goose during resight event j. I modeled the true latent state of each goose using (1.5). Survival was estimated for each season, age, sex, and year. The parameters p_j and $\phi_{a,sex,j}$ were modeled as $\text{logit}(\boldsymbol{p}) = \boldsymbol{W}\boldsymbol{\alpha}$ and $\text{logit}(\boldsymbol{\phi}) = \boldsymbol{X}\boldsymbol{\beta}$. The matrix $\boldsymbol{W}$ consisted of indicator variables identifying the season and year of each potential resighting, allowing detection probability to vary additively by season and year. The matrix $\boldsymbol{X}$ distinguished birds in the following categories: season, age of the bird during resight events, sex, and year it was resighted (1982–2005, 2011–2014). Thus, survival was allowed to vary additively for each season, age, sex, and year. I assumed vague priors for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ defined by $\text{N}(\boldsymbol{0}, 1.5^2\boldsymbol{I})$, where $\boldsymbol{I}$ is the identity matrix. I derived annual survival ($\phi_{a,sex,t}$) for each age class, sex, and year using (1.2). I marginalized sex-specific annual survival rates by iterating the conditional expectations of survival given sex. A complete model statement is given in Appendix C.

1.3.2.4. *Scaling relative abundance to true abundance.* The Pacific Flyway currently uses the aerial survey data described in (1.7) to make annual harvest regulations for cackling geese and other species (Pacific Flyway Council 1999). The aerial survey data underestimate fall cackling goose abundance because many birds are undetected and because they occur pre-breeding, omitting juvenile birds from the counts. The Pacific Flyway has accounted for this by using the equation $\tilde{n}_t = 3.35 y_t$, where $3.35 = \frac{\sum_{t \in \mathscr{T}} \eta_t}{\sum_{t \in \mathscr{T}} y_t}$, $\mathscr{T} = \{1989 - 2003, 2011 - 2013\}$, and η_t were population estimates derived from an additional data source described in Pacific Flyway Council (1999). The additional data source consisted of the neck-collared geese and resight information used in this study, combined with the number of observed unmarked geese in the periods $\mathscr{T}$. Specifically, the population estimates η_t were calculated using: $\eta_t = \frac{\text{unique marks}_t}{\tilde{p}_t} \times \frac{\text{unmarked}_t}{\text{marked}_t}$, where unique marks$_t$ were the number of unique neckbands observed, $\tilde{p}_t$ was the detection probability, and $\frac{\text{unmarked}_t}{\text{marked}_t}$ was the ratio of unmarked-to-marked birds observed (Pacific Flyway Council 1999). The detection

probability $\tilde{p}_t$ were estimated independently of this study and reported in Pacific Flyway Council (1999). The estimates for η_t have been considered the closest estimate of true abundance available which is why they were used to estimate a scaler for y_t to approximate true abundance. Note that η_t is an estimate of the number of unmarked birds in the population. To scale my estimates of relative abundance, using the corrected mark-resight abundance estimates $\boldsymbol{\eta} = \frac{\text{unique marks}_t}{\tilde{p}_t} \times \frac{\text{unmarked}_t + 1}{\text{marked}_t}$, I calculated a multiplier c_{IPM} that minimized the sum of squared-error distance between the posterior mean of pre-breeding abundance and the mark-resight estimates of abundance, $\boldsymbol{\eta}$. The multiplier c_{IPM} from pre-breeding abundance to the estimated true abundance assumes c_{IPM} is constant among years. The resulting estimates of abundance ($n_{tot,t} = c_{IPM} E(n_t | \boldsymbol{y}, \boldsymbol{M})$) incorporated three data sources: aerial survey data, mark-resight survival survival data, and mark-resight population estimation data.

I calculated posterior distributions using a Markov Chain Monte Carlo algorithm in R version 3.0.2 and C++ (R Core Team 2013). The full algorithm is provided in Appendix D. I calculated marginal posterior distributions for the parameters $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, $\boldsymbol{f}$, and σ^2 and the derived parameters: $\boldsymbol{n}_{sub}$, $\boldsymbol{n}_{ad}$, $\boldsymbol{n}$, $\boldsymbol{n}_{tot}$, $\phi_{s,a,sex,t}$, $\phi_{a,sex,t}$, $p_{s,t}$ and annual population growth rate $\boldsymbol{\lambda}$. Preliminary analyses indicated that the β and α parameters associated with 1984 and 2004–2010 did not converge due to no mark-resight data occurring during those periods. Therefore, I prescribed strong priors for these parameters (i.e., I set $\beta = 0$ and $\alpha = -10$) to ensure convergence. This assumes that there was no year effect on survival during those years, and detection probability was approximately zero. I obtained 3 parallel chains of 200,000 iterations using the algorithm. I removed the first 25,000 burn-in iterations. I examined convergence using trace plots and Gelman-Rubin diagnostics.

1.4. RESULTS

A total of 8,193 of 10,428 neck-collared geese were resighted at least once. The oldest recorded bird was a female captured in Alaska as an adult on 27 Jul 1985 and recaptured in Oregon on 5 Dec 2013 (> 29 years old). Resighting between 2011–2014 occurred only in California, Oregon, and Washington, and therefore estimates of survival and detection probability during these years from May–Aug were estimated from the 1982–2004 resight data. All other posterior distributions reported had Gelman Rubin diagnostic statistics < 1.01 from the 25,000–200,000 MCMC iterations of three chains. It required 933.5 hours to obtain 200,000 iterations on a workstation with 3.0 GHz Intel Xeon Eight-Core and 64 GB of memory.

Mean seasonal detection probability, $E(E(p_j|\boldsymbol{M}))$, was 0.26 (range = 0.01–0.63; Fig. 1.4). Detection probability was higher in California, Oregon, and Washington than Alaska and usually the highest from Dec–Feb (Fig. 1.5). Detection probability decreased from 1993 to 2004 corresponding to a decreased banding and survey effort (Appendix B, Fig. 1.4). Most of the 95 % credible intervals of $\boldsymbol{\alpha}$ did not overlap 0, with the exception of the parameters associated with detection probability in 2001–2002 and 2003–2004 indicating these years had similar detection probability to 1982-1983 (Fig. 1.6).

Mean seasonal survival probability, $E(E_{j|a}(\phi_{s,a}|\boldsymbol{y}, \boldsymbol{M})), = 0.90$ (range: 0.78–0.94) for juveniles, 0.94 (range: 0.86–0.96) for sub-adults, and 0.95 (range: 0.89–0.97) for adults. Survival was highest when geese were in their breeding areas in Alaska from May–Jun (Fig. 1.7). Survival was lowest in Oct–Nov when birds were in their winter areas. The lowest and highest mean seasonal survival probability occurred in Oct–Nov, 2013 and May–Jun, 1987, respectively. Mean annual survival probability for juveniles, sub-adults, and adults ($E(E(\phi_{a,sex,t}|\boldsymbol{y}, \boldsymbol{M}))$) was 0.54 (range: 0.26–0.66), 0.69 (range: 0.45–0.79), and 0.75 (range: 0.53–0.83), respectively (Fig. 1.8). Annual survival probability was highest during 1987–1988 and lowest during 2013–2014. Sex did

18

not appear to additively affect survival probability: $E(\beta_{sex=female}) = 0.02$ (Fig. 1.9). Survival estimates for all age classes were consistent during years with no hunting (1984-1984–1993-1994) or restricted hunting (i.e., one or two daily bird bag limit; 1994-1995–1997-1998,2011-2012) and were lower during years when liberal hunting occurred (i.e., three or four daily bird bag limit; 1998-1999–2004-2005,2012-2013; Fig. 1.8).

Mean fecundity for all years $E(E(f_t|\boldsymbol{y}, \boldsymbol{M})) = 2.37$ birds per pair (range: 0.25–7.33; Fig. 1.10). Mean fecundity for the years 1982–2004 was 1.67 birds per pair. Mean fecundity was lowest in 1997 and highest in 2013. The five years with the smallest mean fecundity were 1982, 1983, 1990, 1997, and 2001. The largest mean fecundity occurred in 1991, 1993, 2002, 2012, and 2013 (Fig. 1.10). Fecundity estimates were large from 2011–2013 relative to other years in which fecundity estimates were available (1982–1983, 1985–2003; Fig. 1.10). The large fecundity estimates corresponded to reduced apparent survival during these years (Fig. 1.8) and relatively stable population counts in 2012–2014 (Fig. 1.2).

The mean proportion of adults just prior to breeding $E(E\frac{n_{ad,t}}{n_t}|\boldsymbol{y}, \boldsymbol{M})) = 0.71$ (range: 0.39–0.97; Fig. 1.11). The population grew between 1985–1997, except in 1991 (Figs. 1.12, 1.13). Low fecundity in 1997 and reduced survival after 1998, corresponding with liberalization of hunting regulations, appeared to stabilize population growth (Figs. 1.8, 1.10, 1.11, 1.13, 1.12).

The value of c_{IPM} that minimized the sum of squared error between mark-resight abundance estimates η_t and posterior means of pre-breeding relative population size was 3.36.

1.5. DISCUSSION

1.5.1. INTRA-ANNUAL CACKLING GOOSE POPULATION DYNAMICS. Seasonal survival estimates were similar among seasons with the exception of Oct–Nov (Fig. 1.7). Reduced survival in Oct–Nov correspond to both fall migration and recreational hunting. Most ecological theory

related to migration assumes migration is costly and therefore reduces survival (Greenberg 2005). However, seasonal survival rates of adult female greater snow geese (*Chen caerulescens atlantica*) were not lower during migration compared to other seasons, after the effects of hunting were removed, and seasonal survival of Pacific brant (*Branta bernicla nigricans*) were higher during migration compared to other seasons (Ward et al. 1997, Gauthier et al. 2001). I did not explicitly estimate hunting mortality but the reduced survival rate during fall migration occurred during periods of recreational hunting. Thus, it is possible that either recreational hunting or the effects from migration (or both) could have been responsible for reduced survival during these periods. Survival probability during spring migration (May–Jun) was higher than any other season. Further, high survival probability during years with either no recreational hunting or restricted recreational hunting suggest that migration alone was not responsible for low survival rates and that recreational hunting during fall migration reduced survival. Seasonal survival estimates were highest during May–Jun corresponding to spring migration, and spring subsistence hunting.

I detected little difference (0.02) in additive annual survival between males and females. These findings are consistent with many other studies of geese (Samuel et al. 1990, Francis et al. 1992*b*, Rexstad 1992, Ward et al. 1997, Menu et al. 2002). However, I did not examine if seasonal survival was a function of sex. Schmutz and Ely (1999) showed that seasonal survival differed among seasons by sex, but that annual survival was similar between sexes. Raveling et al. (1992), using some of the same data analyzed in this study, found evidence of higher survival of females during summer in the years 1982–1988. The difference in survival was small (i.e., 0.061). There was evidence of difference in age-specific survival, with adults having the largest survival followed by sub-adults, then juveniles (Figs. 1.8, 1.9).

1.5.2. INTER-ANNUAL CACKLING GOOSE POPULATION DYNAMICS. Abundance estimates of cackling geese increased rapidly between 1985–1997 (Figs. 1.12, 1.13). This period corresponded with population growth in other goose species in North America including, Canada geese (*Branta canadensis*), snow geese (*Chen caerulescens*), and white-fronted geese (*Anser albifrons*; Ankney 1996, Menu et al. 2002). Both recreational and subsistence hunting was closed for cackling geese during most of the period of population growth, with the exception of 1994–1997, in which subsistence hunting was open and recreational hunting was restricted (i.e., 1–2 bird daily bag limit). Population growth is most affected by adult survival in long-lived species such as cackling geese (Lebreton and Clobert 1991). Additionally, studies on other goose species suggest that hunting mortality is additive and reduces goose survival (Francis et al. 1992*b*, Rexstad 1992, Hestbeck 1994, Schmutz et al. 1997, Gauthier et al. 2001). Thus, the additional mortality due to liberalization of hunting regulations after 1997 could have accounted for population stabilization. Menu et al. (2002) found similar trends in stabilization of population growth following re-opening of hunting of greater snow geese in the late 1970s, but did not have survival estimates to confirm whether hunting reduced survival. I found that survival estimates were not lower during restrictive recreational hunting regulations and open subsistence hunting, but were lower during periods with liberal recreational hunting regulations. Mean annual adult survival between 1985–1997 was 0.76, compared to 0.71 during all other times of this study, and 0.66 during periods of liberal bag limits. The restrictive recreational hunting regulation of a 1-bird daily bag limit from the period of 1994–1997, combined with legalization of subsistence hunting did not appear to affect annual survival probability compared to years when hunting was illegal because abundance continued to increase and survival estimates appeared unchanged (Figs. 1.8, 1.13). The subsistence harvest season includes most of the time cackling geese are in Alaska (minus a 30-day period during nesting), but most subsistence harvest occurs when geese first arrive in the spring (Schmutz et al. 1997, Hupp

et al. 2008, Naves 2010; 2011; 2012). Relatively high spring survival combined with high seasonal and annual survival rates from 1994–1997 suggest that differences in regulations on subsistence hunting had a small effect on survival rates compared to factors during other times of the year (Fig. 1.7). These results are in contrast to survival rates for Pacific brant, which were lowest in late spring (15 Apr–1 Jun) on the YK Delta, corresponding to periods of greatest subsistence harvest (Ward et al. 1997). Although survival of cackling geese was relatively high during spring, it should be noted that changes in pre-breeding survival have a disproportionately large effect on population size relative to changes in survival during other parts of the year and small decreases in survival will result in larger *per capita* take (Doubleday 1975, Schmutz et al. 1997).

In addition to reduced survival following 1997, the fecundity rate in 1997 was low and could have been partially responsible for the stagnation of population growth. Low fecundity rates in 1997 are consistent with an independent data set on nest success of cackling geese in 1997 (C. Ely, personal communication). Nest survival and productivity of arctic nesting birds are largely driven by two ecological processes: predation and gosling growing conditions (Lepage et al. 1998, Lake et al. 2008, Rizzolo et al. 2014). Arctic fox (*Vulpes lagopus*) are thought to affect cackling goose nest success. Additionally, cackling goose density and the abundance of food, drive gosling growth after fledging (Lake et al. 2008, Rizzolo et al. 2014). I did not explicitly link these covariates to estimated fecundity due to limited data. However, the USFWS (unpublished data) have collected presence-absence data on the proportion of surveyed plots with active arctic fox sign during nesting surveys of geese on the YK Delta. The five years with the highest proportion of plots with active fox sign included 2001, 2002, 2003, 2004, and 2008. Of these years with estimated fecundity in this study, only the low fecundity estimated in 2001 corresponded well with large proportion of active fox sign. This suggests either that 1) the proportion of sampled plots that contain active fox sign is not a good indicator of population level abundance of arctic fox or that 2) mechanisms

other than arctic fox (i.e., cackling goose density and food availability) drive differences in annual fecundity.

Fecundity estimates were high and contained large credible intervals during the period 2012–2014 relative to other periods of the study (Fig. 1.10). Although there was considerable variation in the 95% credible intervals, the mean of the posterior distribution of $f_{2012} = 6$, which seems biologically unlikely. Fecundity was not estimated using explicit data and was estimated by interfacing abundance estimates and survival data with the life cycle of cackling geese. Thus, the large values of fecundity could be the result of over-estimating y_t or underestimating survival. One potential explanation is that survival was underestimated due to differing survival rates between neck-collared birds and non-collared birds. Samuel et al. (1990) found that survival rates of neck-banded Canada geese were not different than those of leg-banded geese and (Menu et al. 2000) found similar results for greater snow geese. Negative effects of collars have been observed in small Canada geese (Alisauskas and Lindberg 2002). Because collars were used in all years of this study, this does not explain the reduced survival estimated for 2012–2014, unless different banding techniques were used. Another explanation is that hunters were targeting neck-collared birds, and this occurred at a relatively higher rate during 2012–2014. In a comparison between the ratio of collared to non-collared cackling geese in the field during resighting events in 2011–2014 and the ratio of harvested collared geese to harvested non-collared geese, the ratio of harvested birds was higher than that observed in the wild (B. Reischus, unpublished data). This suggests that hunters might have been targeting neck-collared cackling geese, reducing the apparent survival. It is unknown whether this targeting of neck-collared geese increased following the 2011–2013 banding efforts of cackling geese. Collecting additional data on fecundity could help calibrate apparent survival to true survival in an IPM framework. It is also possible the y_t was largely over-estimated in 2013–2014 which would increase fecundity estimates in 2012–2013. Finally, it is possible that

because these estimates of survival occurred on the temporal boundary of observed data, the estimates were not "stabilized" by subsequent data, as occurred in previous years.

The IPM provided insight on age structure of the cackling goose population from 1982–1984, and 1985–2004 (conditional on an assumed age structure in 1982; Fig. 1.11). The low-frequency (long-term) variation in relative abundance appeared to be most correlated with adult abundance and the higher frequency (short-term) variation appeared to be correlated with juvenile abundance. This is not surprising as fecundity is closely related to annual weather patterns and expected to be highly variable compared to adult survival rates that are more consistent among years (Raveling 1978, Francis et al. 1992a, Skinner et al. 1998, Menu et al. 2002, Souchay et al. 2013).

Mean annual survival rates estimated in Raveling et al. (1992), using some of the same data as this study, were about 25% lower than my estimates of annual survival. Raveling et al. (1992) estimated a 25% neck-band loss rate. The estimated 25% band-loss rate in Raveling et al. (1992) was calculated by comparing survival estimates of mark-resight data to survival estimates based on abundance data. As an alternative and more direct approach for estimating neck-band loss, I examined the retention of neck-bands from geese that were captured during retrap events when leg bands could be uniquely identified. Of > 400 geese retrapped that were marked with a neck band and leg band, only three lost their neck band and therefore I considered neck-band loss to be negligible ($< 1\%$). However, this does not explain the difference in survival estimates in Raveling et al. (1992) and this study. One possible explanation for the difference in survival estimates is the abundance data used in the IPM framework could have had a large effect on the survival parameters. Only a small proportion (0.13) of the total number of resighted birds occurred from 1982–1988. Thus, abundance data might have had a disproportionately large effect on survival estimates for early years, compared to later years when more mark-resight data was available for parameter estimation.

The multiplier c_{IPM} that minimized the squared-error loss between y_t and $E(n_t|\boldsymbol{y}, \boldsymbol{M})$ for years $\mathscr{T}$ was 3.36. This value is similar to the current scaler used to estimate true abundance using y_t, 3.35. Although the results are similar, c_{IPM} has useful statistical properties. It is optimal with respect to the squared-error loss function and it incorporates survival estimates in its calculation through the posterior distribution of n_t.

1.5.3. OTHER INSIGHTS ON INTEGRATED POPULATION MODELS. In the integrated population model described above, the prior distribution σ^2 served as a tuning parameter for the strength of each data set (i.e., abundance data, mark-resight data). Small values of σ^2 would force abundance estimates to be close to y_t and estimates of $\boldsymbol{\beta}$ to change accordingly. Large values of σ^2 would permit flexibility in abundance estimates allowing $\boldsymbol{\beta}$ to be estimated from the mark-resight data. I chose mean and variance parameters for σ^2 large enough to allow realistic uncertainty in abundance estimates, but still allow y_t to inform vital rate parameters. If an investigator is confident in abundance estimates, then smaller values of σ^2 can be chosen (i.e., either a smaller mean or by fixing the value of σ^2. This would permit more precise estimates of fecundity because the range of fecundity values that result in values close to y_t are restricted. This information can inform sampling design for both abundance and mark-resight studies, as well as help to understand the value of information that results from each data source.

I also examined sensitivity to more informative priors for fecundity. The priors were $f_t \sim$ Gamma$(1, 1)$. In combination with large values of σ^2, the informative priors drove abundance estimates to zero because the fecundity regulated by the prior was not large enough to compensate for the estimated survival parameters. A smaller mean of σ^2 or vague priors for fecundity provided better model fit. Smaller values of σ^2 forced the estimates of $\boldsymbol{\beta}$ to change from those estimated from the data. Vague priors for fecundity allowed $\boldsymbol{\beta}$ to be estimated from the data. Thus, certain specifications of priors in an integrated population model might result in bias in other estimated

parameters and there is a trade-off among the sources of information in survival and fecundity. If the preponderance of data support survival with limited or potentially erroneous prior information on fecundity, I recommend vague priors for fecundity.

IPMs are a powerful tool for the combined analysis of vital rate data and abundance data. They provide precise parameter estimation and the ability to estimate parameters without explicit data. Most of the published literature for methods and application of IPMs are directed towards the inter-annual temporal scale. To realize the benefits of IPMs for parameter estimation, investigators have dissolved data occurring at a finer scale to match the inter-annual temporal scale. This results in loss of information and the inability to use abundance data to estimate seasonal vital rates. For many populations, intra-annual temporal population dynamics are important for guiding management decisions. For the cackling goose population, I found that seasonal survival was highest when birds arrive in the spring and lowest in Oct–Nov. This information could potentially help guide management decisions for a species that is harvested on both ends of its migratory range during different parts of the year. I developed and implemented an IPM that estimated vital rate parameters based on disparate temporal data. The model benefited from the IPM framework in that seasonal and annual survival rates were estimated from two data sources, and fecundity and population structure, parameters with no explicit data, were estimable. The specific model I fit assumed known population structure in at least one year suggesting age-structure data is an important source of information for future IPM models. The model provided new insights on cackling goose population dynamics not obtainable through individual analyses of each data set alone.
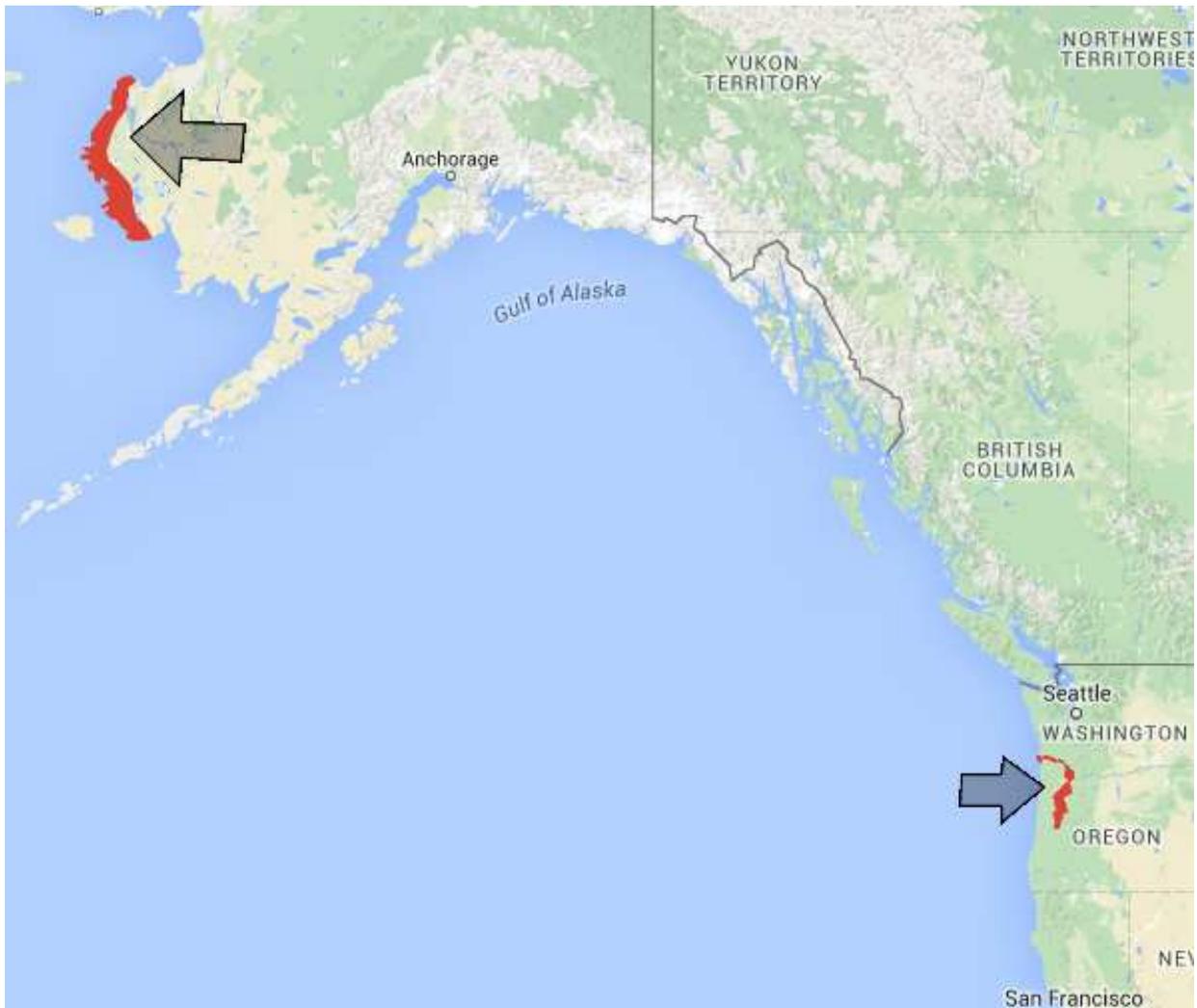
F<small>IGURE</small> 1.1. Cackling geese (*Branta hutchinsii minima*) nest in the Yukon-Kuskokwim Delta, AK and primarily winter in the Lower Columbia River Valley, WA and Willamette Valley, OR.
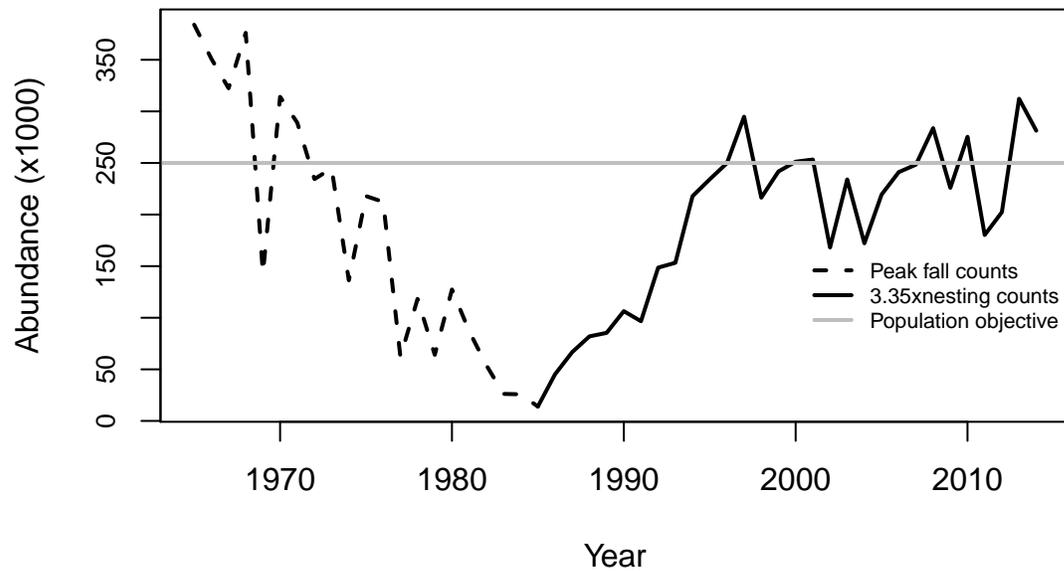
FIGURE 1.2. Estimated cackling goose abundance from 1965–2014. Prior to 1985 abundance was estimated using peak fall counts in the Klamath Basin and Sacramento Valley, CA. Since 1985 abundance ($\tilde{n}_t = 3.35y_t$), has been estimated during nesting using aerial surveys on the Yukon-Kuskokwim Delta, AK.
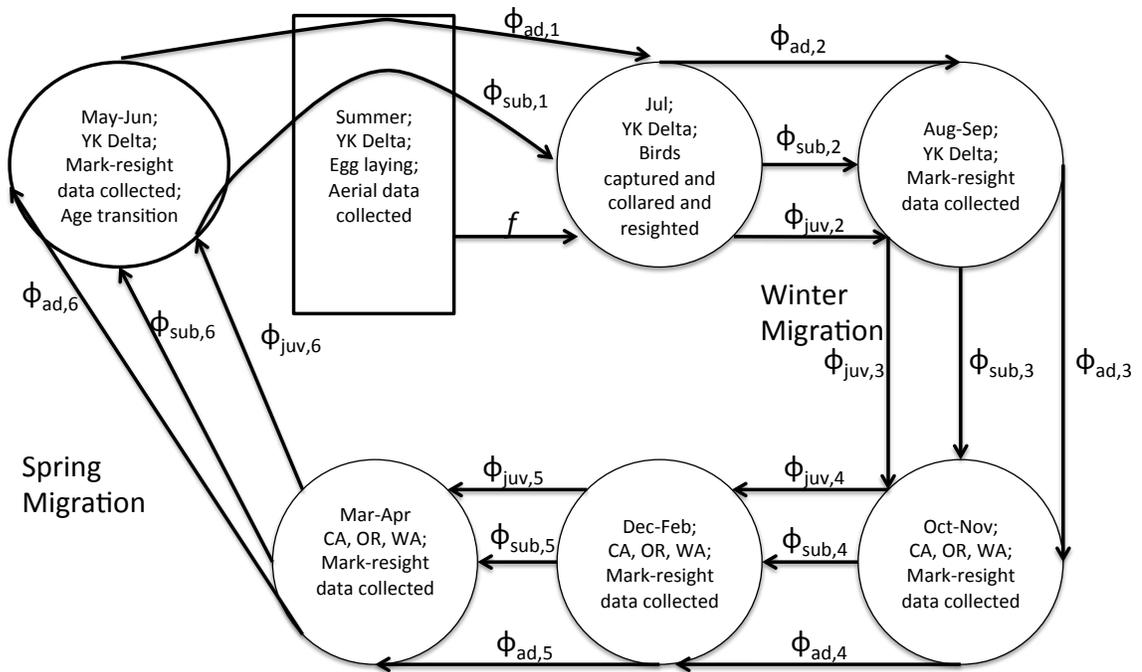
FIGURE 1.3. Life-cycle and data collection periods for cackling geese. The value *f* represents a composite reproductive parameter for the mean per pair number of juveniles born that survived to banding in July.
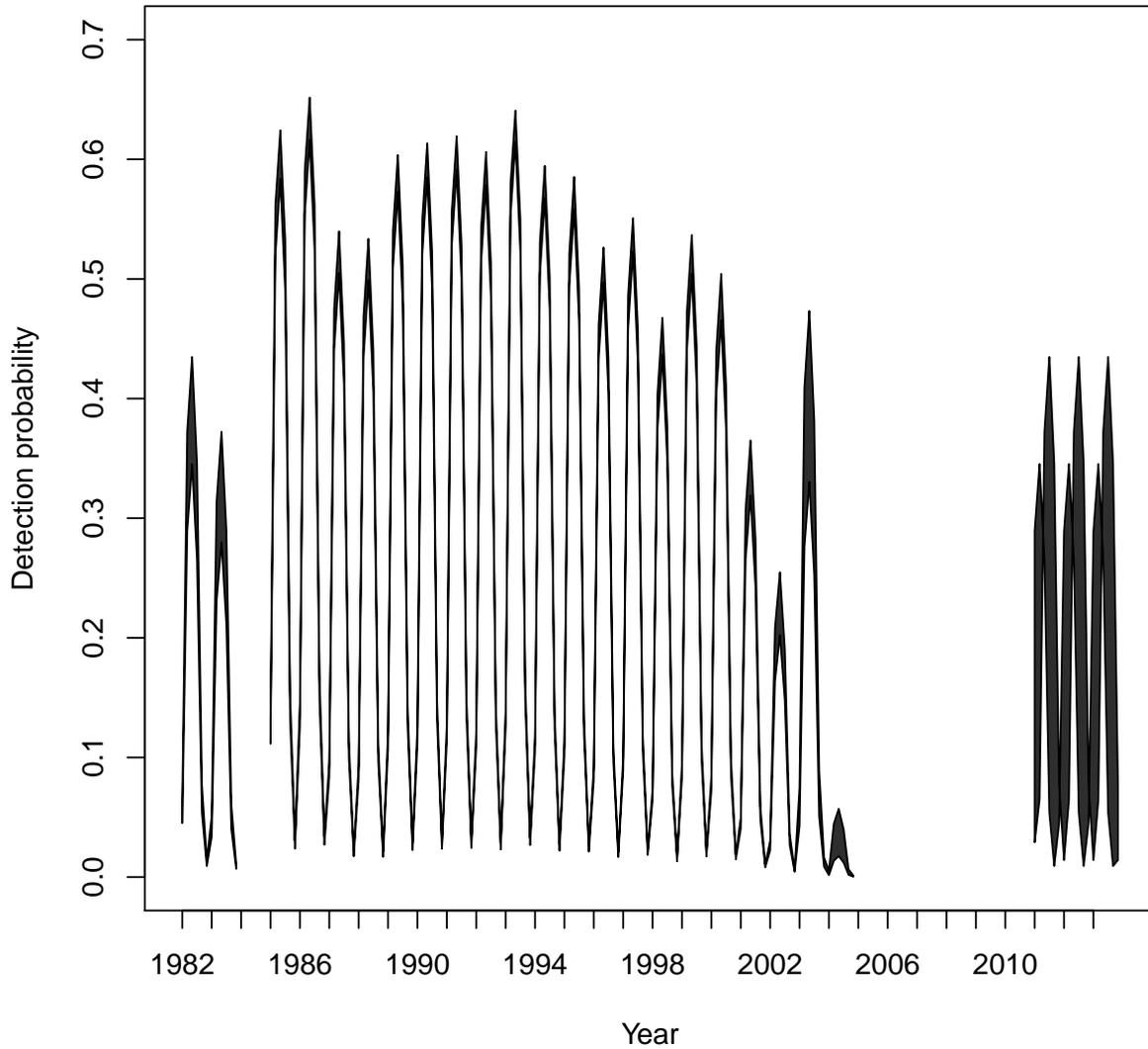
FIGURE 1.4. Seasonal detection probability for resighting neck-collared cackling geese from 1982–2014. The seasons correspond to late summer (mid-Jul–mid-Aug) in AK, Oct–Nov in CA, OR, and WA, Dec–Feb in CA, OR, WA, Mar–Apr in CA, OR, and WA, spring (Apr–Jun) in AK, and during banding drives in AK (mid-Jul–mid-Aug. No geese were resighted during 1984–1985 and 2004–2010.
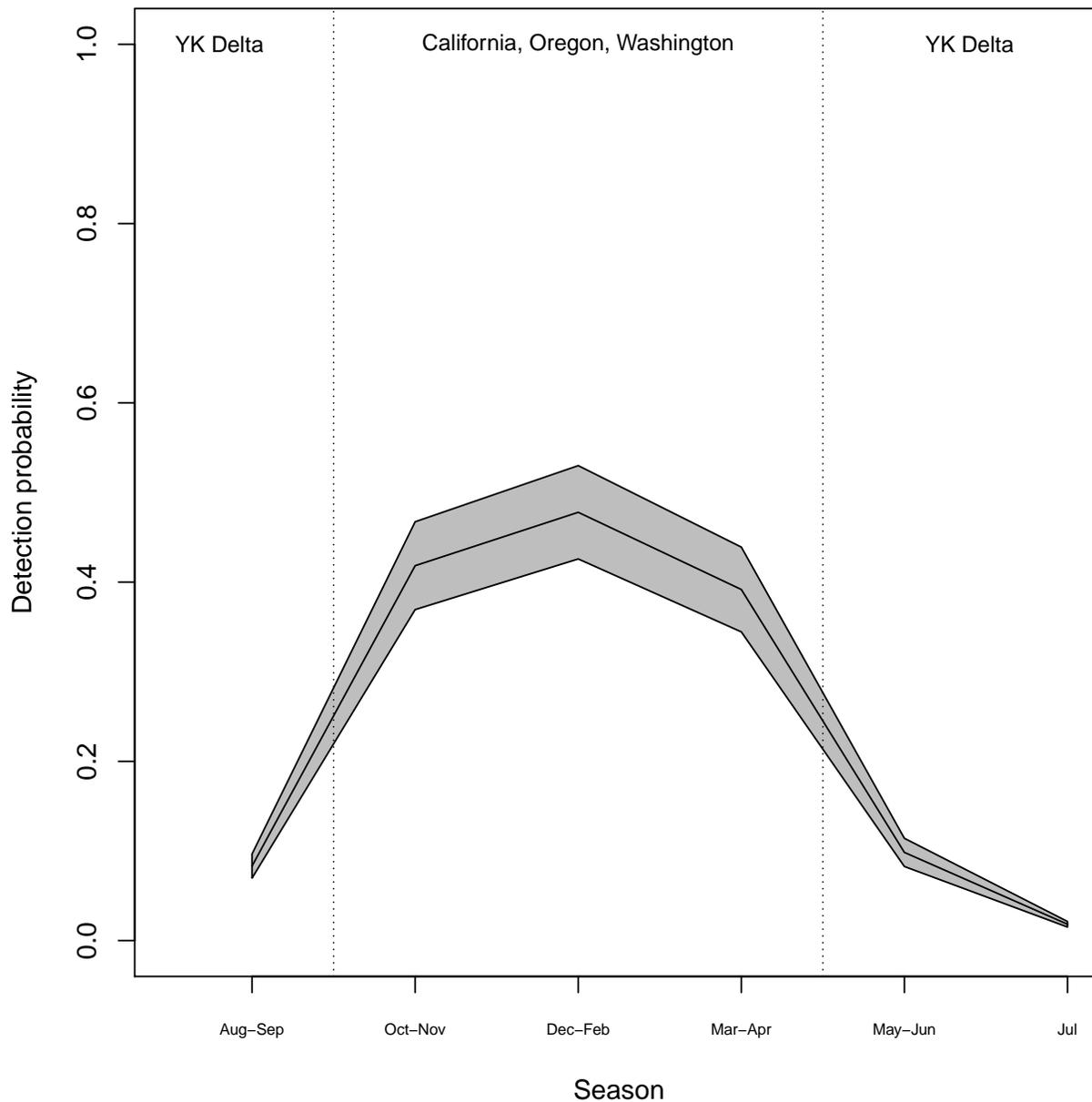
FIGURE 1.5. Mean of posterior means (solid line) of seasonal detection probability for resighting neck-collared cackling geese for 6 periods with mean taken over years in which observers were looking for geese (1982-1984, 1985-2004, 2011-2014). The shaded region is the mean $\pm 2 \times SE$.
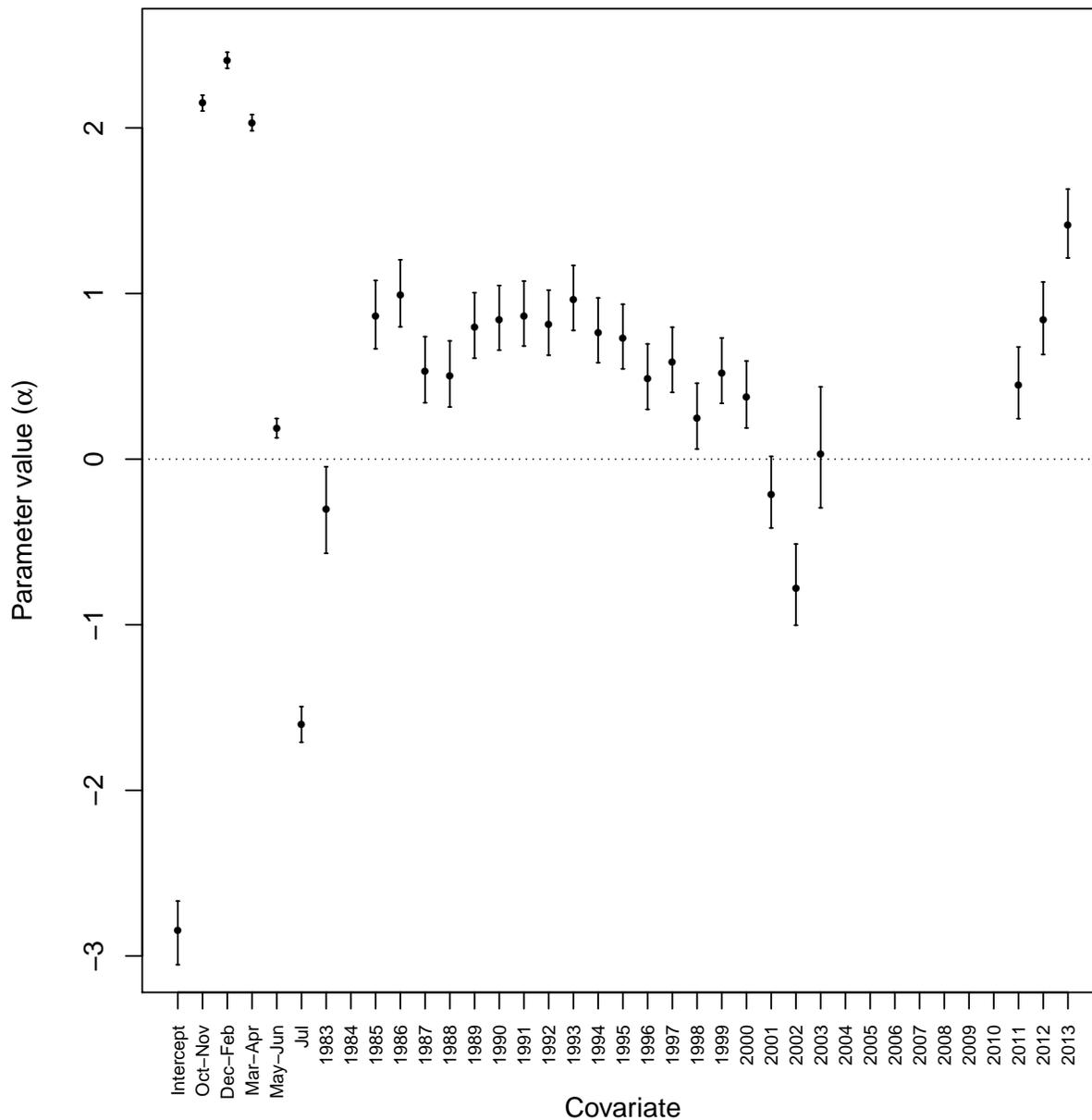
FIGURE 1.6. Posterior mean and 95% credible intervals for each of 37 detection parameters α of the integrated population model used to examine population dynamics of cackling geese from 1982–2014. The parameters associated with detection probability during the sampling periods in 1984–1985, and 2005–2010 were set to -10 because no data were collected during these periods.
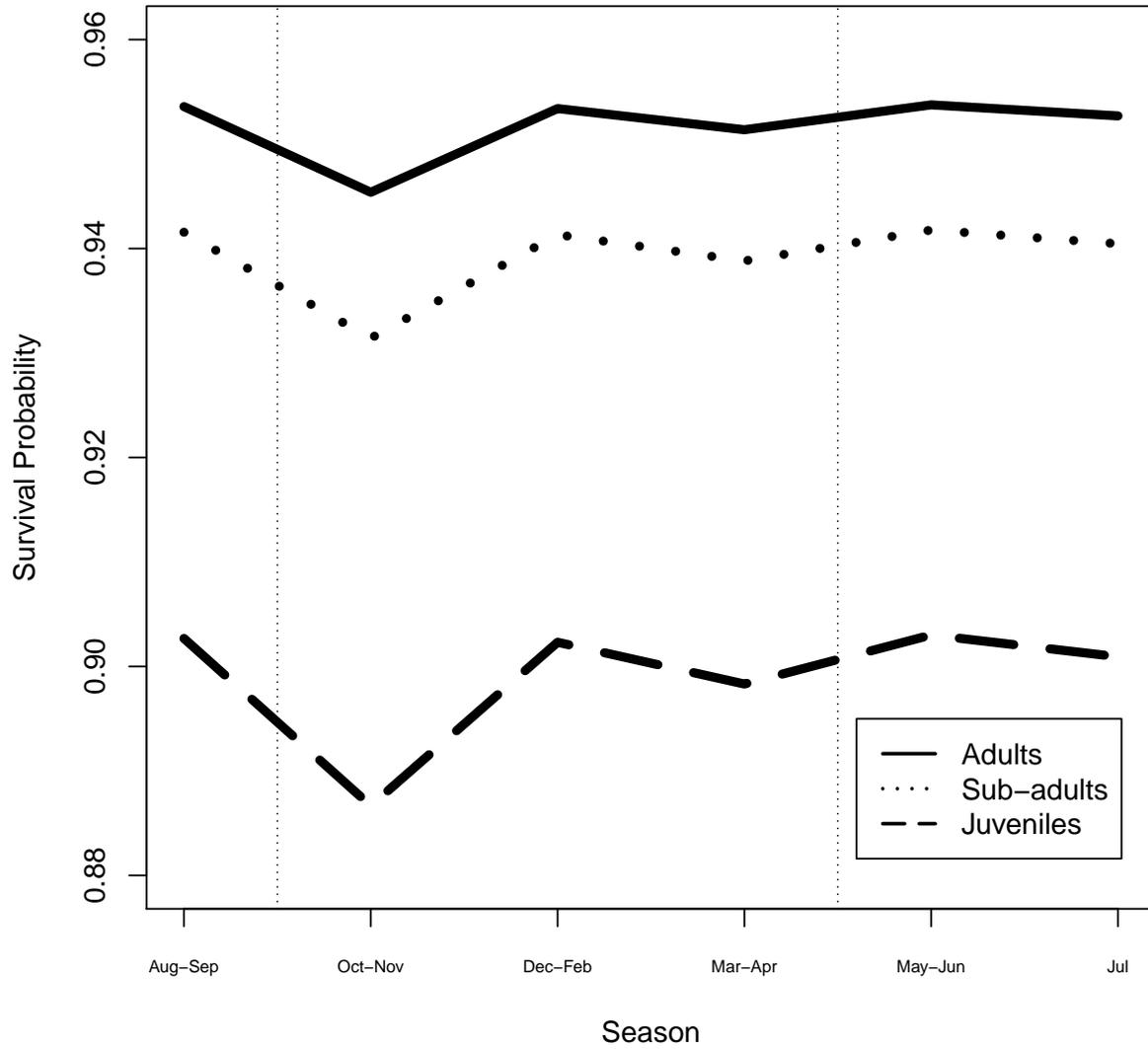
FIGURE 1.7. Mean seasonal survival probability for juvenile, sub-adult, and adult neck-collared cackling geese estimated by taking the mean of the posterior means of seasonal survival probability for each year with mark-resight data on cackling geese (1982–1983, 1985–2003, 2011–2013).
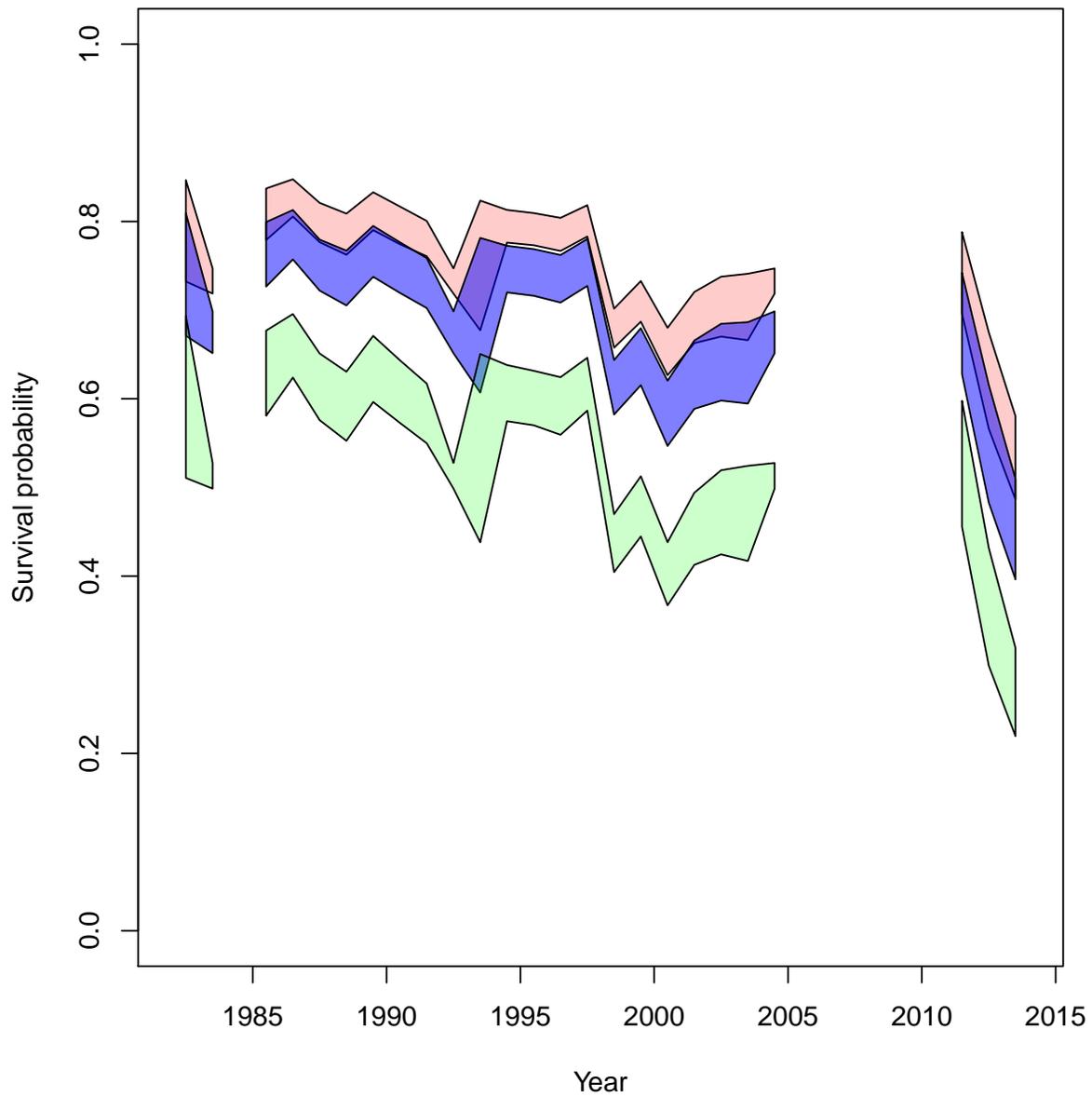
FIGURE 1.8. Posterior mean and 95% credible intervals of annual survival probability of three age classes of neck-collared cackling geese from 1982–2014. Red = adults, blue = sub-adults, and green = juveniles.
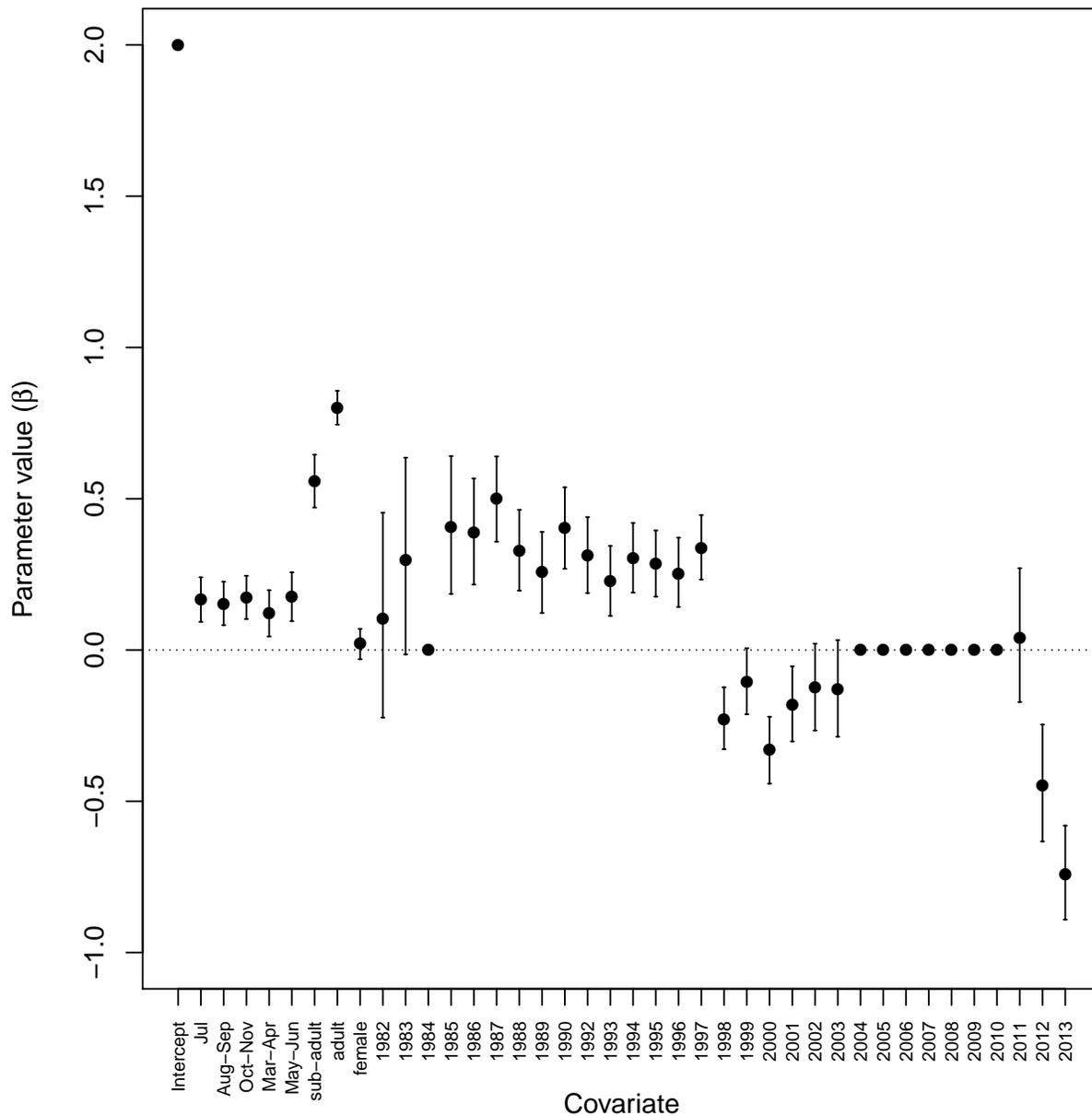
FIGURE 1.9. Posterior mean and 95% credible intervals for each of 40 survival parameters β of the integrated population model used to examine population dynamics of cackling geese from 1982–2014. The parameter associated with survival probabilities during the sampling periods in 1984–1985, and 2005–2010 were set to zero because no data were collected during these periods.
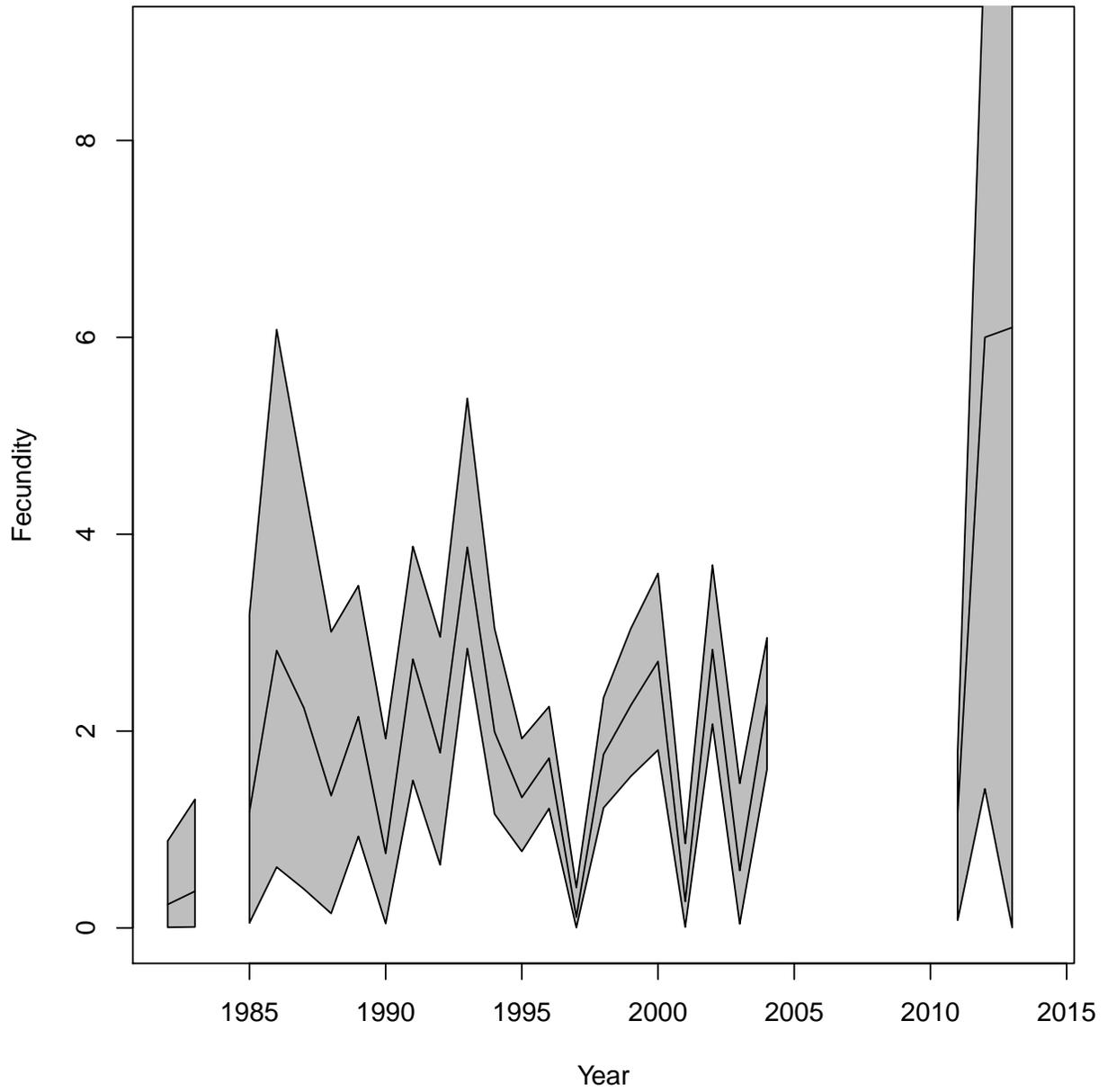
FIGURE 1.10. Posterior mean and 95% credible intervals for fecundity of cackling geese between 1982–2013.
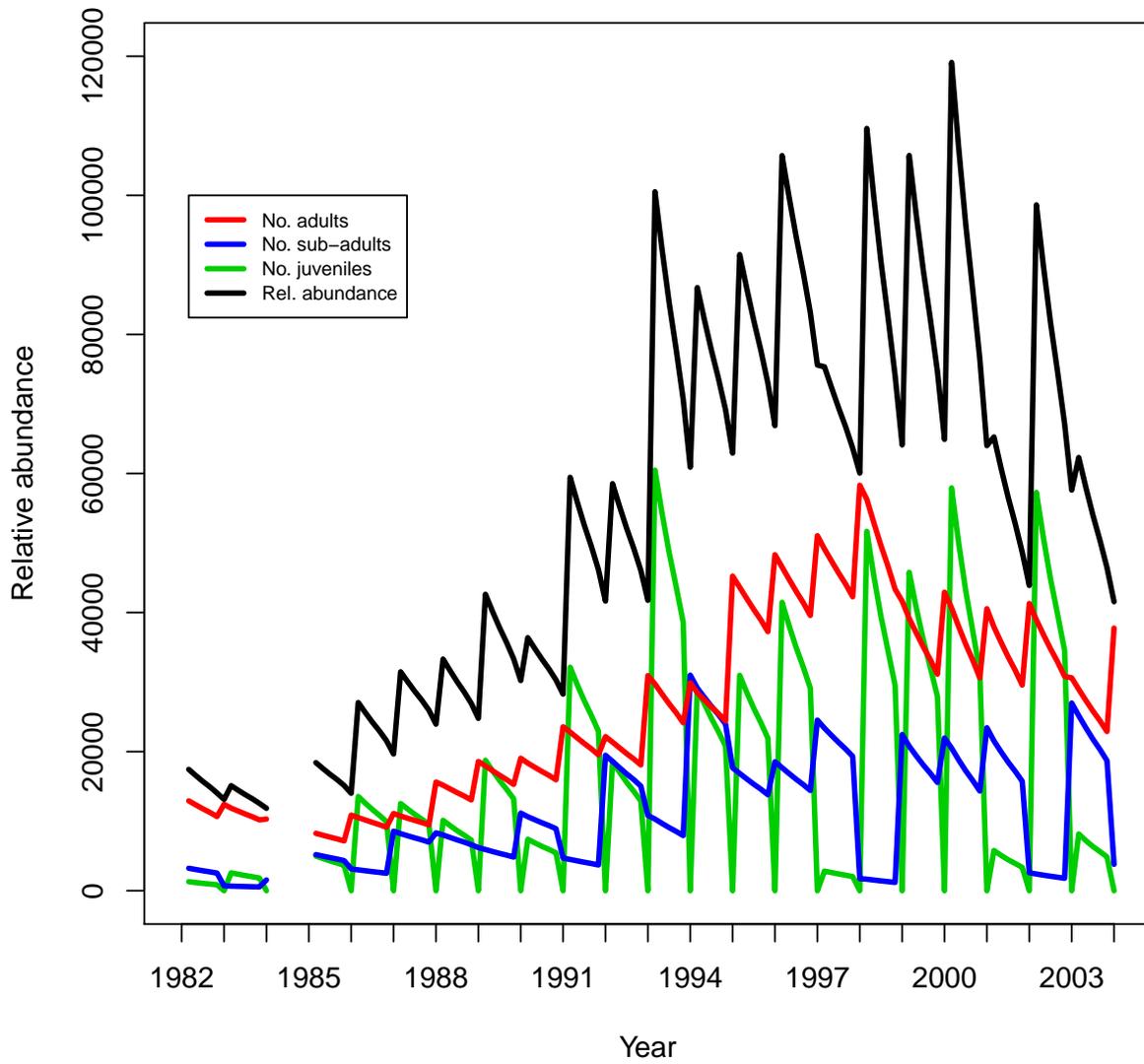
FIGURE 1.11. Posterior means of relative abundance of three age classes of cackling geese and their sum (relative abundance) from 1982–2004.
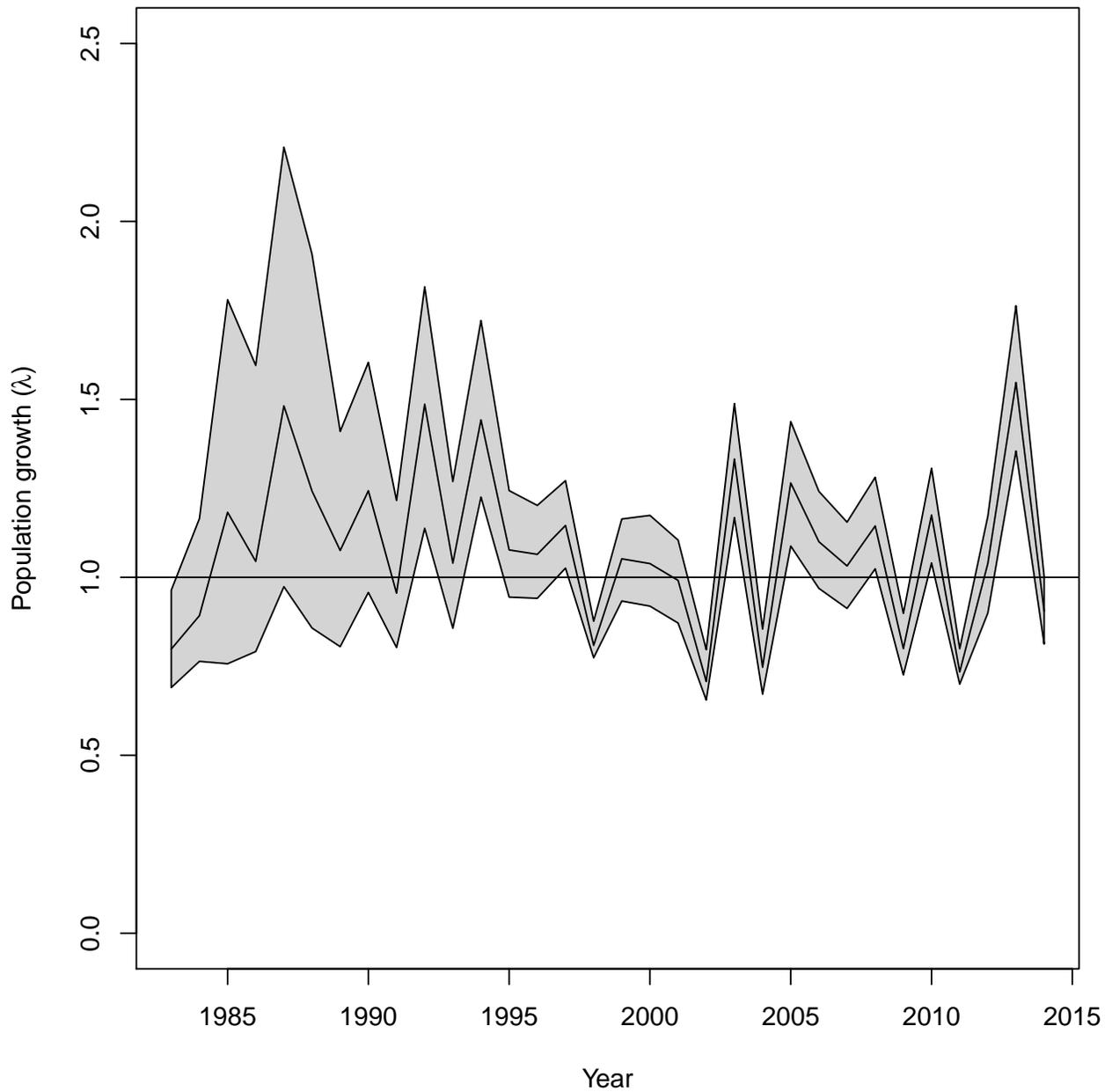
FIGURE 1.12. Posterior mean and 95% credible intervals of annual population growth rate ($\lambda_t = n_t/n_{t-1}$) of cackling geese from 1982–2014.
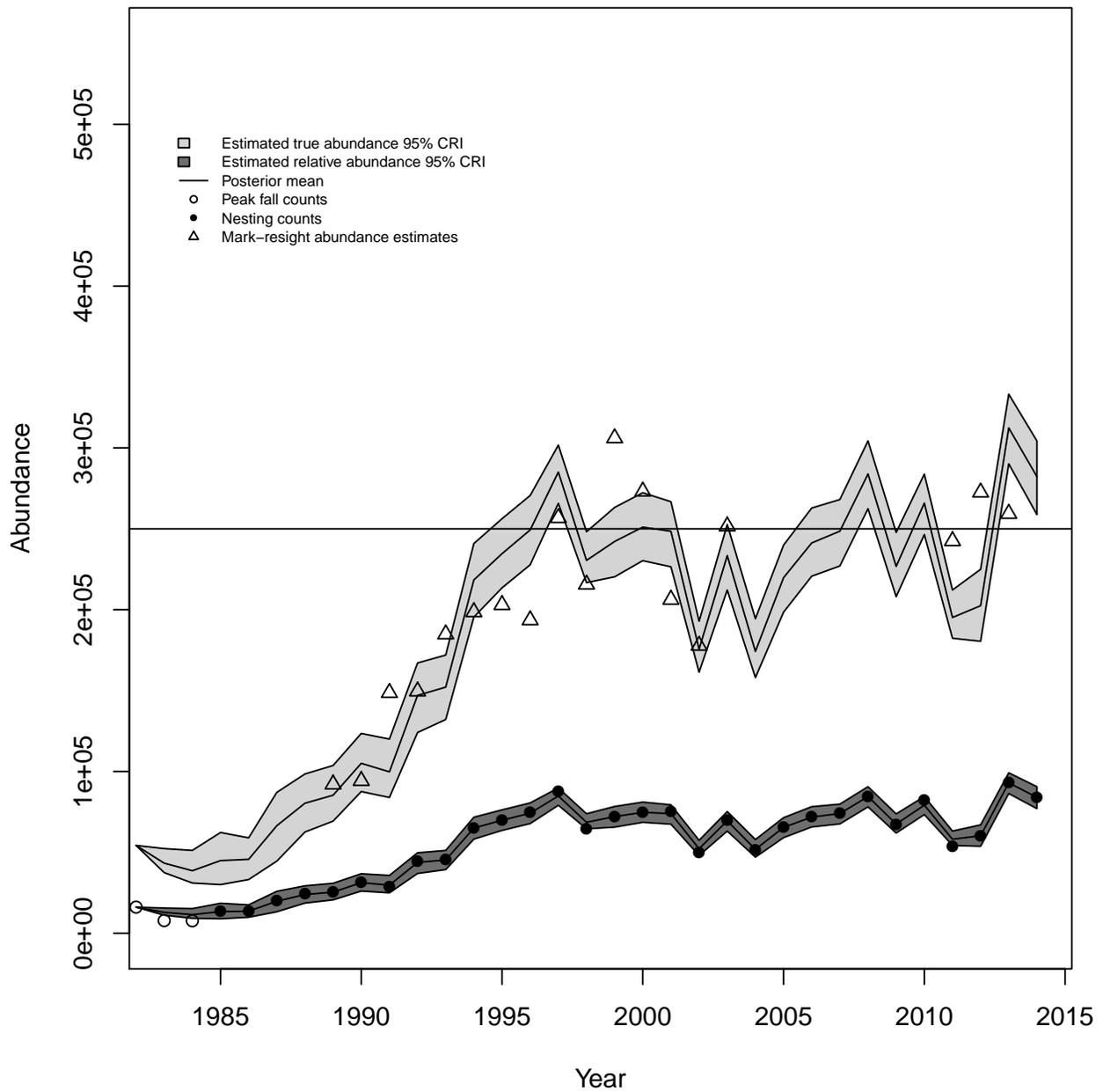
FIGURE 1.13. Estimated pre-breeding relative abundance and scaled estimates of true abundance using the scaler $c_{IPM} = 3.36$ of cackling geese from 1982–2014 fit using an integrated population model.

CHAPTER 2

A GUIDE TO MULTI-OBJECTIVE OPTIMIZATION FOR ECOLOGICAL PROBLEMS WITH AN APPLICATION TO CACKLING GOOSE MANAGEMENT

2.1. SUMMARY

Choices in ecological research and management are the result of balancing multiple, often competing, objectives. Familiar examples include natural resource management and developing models of ecological processes. *Multiple-objective optimization* (MOO) is a formal decision-theoretic framework for solving multiple objective problems in terms of Pareto optimal solutions. I discuss Pareto optimality and its relationship to MOO problems. A large class of methods for solving MOO problems can be separated into two strategies: modeling preferences pre-optimization (the *a priori* strategy), or modeling preferences post-optimization (the *a posteriori* strategy). The *a priori* strategy requires describing preferences among objectives without knowledge of how preferences affect the resulting decision. In the *a posteriori* strategy, the decision maker simultaneously considers a set of solutions (the Pareto optimal set) and makes a choice based on the trade-offs observed in the set. I describe several methods for modeling preferences pre-optimization, including: the bounded objective function method, the lexicographic method, and the weighted-sum method. I discuss modeling preferences post-optimization through examination of the Pareto optimal set. I applied each MOO strategy to the natural resource management problem of selecting a management goal for cackling goose (*Branta hutchinsii minima*) abundance. Cackling geese provide food security to Native Alaskan subsistence hunters in the goose's nesting area, but depredate crops on private agricultural fields in wintering areas. I developed objective functions to represent the

competing objectives related to cackling goose management and identified an optimal solution first using the *a priori* strategy, and then by examining trade-offs in the Pareto set using the *a posteriori* strategy. I used four methods for selecting a final solution within the *a posteriori* strategy; the most common optimal solution (i.e., the mode), the most robust optimal solution (i.e., the solution that minimized the standard deviation), and two solutions based on maximizing a restricted portion of the Pareto set. Three of four methods in the *a posteriori* strategy resulted in a similar decision to the *a priori* strategy. I discuss MOO with respect to natural resource management, but it is sufficiently general to cover any ecological problem that contains multiple competing objectives that can be quantified using objective functions.

Key words decision analysis, decision theory, dominated choice, efficiency frontier, Pareto frontier, Pareto optimality

2.2. INTRODUCTION

Ecological decisions that require balancing multiple objectives are pervasive. Examples include endangered species management (e.g., maximizing species persistence while minimizing cost; Maguire et al. 1987), managing harvested species (e.g., maximizing cumulative harvest while maintaining population objectives; Johnson et al. 1997), and choosing statistical models to infer ecological processes (i.e., maximizing model fit while minimizing model complexity). When a decision maker has multiple competing objectives, a solution that simultaneously optimizes each objective does not exist; improving one objective results in a trade-off from another. Solving multi-objective decision problems requires incorporating decision-maker preferences among objectives into the decision problem (either explicitly or implicitly) to reach a compromise solution. Formal methods for evaluating multiple-objective decision problems exist for problems too complex to informally evaluate them (Keeney 1982).

Methods used to address multiple objective problems have been developed in many fields (Keeney and Raiffa 1976, Belton and Stewart 2002, Marler and Arora 2004, Mendoza and Martins 2006, Diaz-Balteiro and Romero 2008) under many names (e.g., multi-attribute optimization, multi-criteria decision analysis, multi-objective optimization). The process of solving a multi-objective problem consists of identifying or soliciting the objectives of the decision maker, identifying potential actions, weighing the potential actions (or the predicted outcome of the actions) with respect to each objective, and making a choice. Scientific investigation can be used to predict the result of potential actions, but science alone is insufficient to address competing objectives because incorporating preferences among objectives requires value-based judgement (Holland-Bartels and Pierce 2011). Consider an early application of decision analysis for managing wildlife populations, the case of the Sumatran rhino (*Dicerorhinus sumatrensis;* Maguire et al. 1987). Sumatran rhinos are critically endangered (van Strien et al. 2008) and now thought to be extinct in Malaysia (Havmøller et al. 2015). Maguire et al. (1987) identified two management objectives for Sumatran rhinos: maximizing species persistence and minimizing cost. The list of potential management actions included maintaining the status quo, controlling poaching, developing a new reserve, expanding a current reserve, building fencing, translocation of rhinos, and developing a captive breeding program (Maguire et al. 1987) . For each action, the authors used available information to estimate the cost (ranging from $0 to $3.69 million dollars) and expected value of the probability of extinction (ranging from 0 to 0.98), quantifying the trade-offs that were expected in terms of objectives. The quantification of objectives in Maguire et al. (1987) did not reveal an unequivocal action, as actions that maximized species persistence were more costly. Ultimately, a decision maker can implement one management strategy (which might consist of combinations of management actions). Thus, how can we formally combine a quantification of objectives (as in Maguire et al. 1987) with objective preferences to select a final, optimal decision? *Multi-objective*

optimization (MOO) is a formal mathematical framework for explicitly incorporating objectives and objective preferences to evaluate decisions.

I outline the MOO framework and describe two strategies for solving MOO problems. Each strategy incorporates objective preferences into the decision problem. The strategies differ in the order in which preferences are incorporated; the first strategy (the *a priori* strategy) incorporates preferences pre-optimization and the second (the *a posteriori* strategy) incorporates them post-optimization. To demonstrate MOO, I apply both strategies to a common natural resource management problem: selecting an abundance objective for a population that affects multiple stakeholders differently. Cackling geese (*Branta hutchinsii minima*) nest on the coastal plain of the Yukon Kuskokwim (YK) Delta. They are an important food source for Native Alaskan subsistence hunters. Ecosystem stability, the satisfaction of recreational hunters, and other non-consumptive users also depend on them. In their wintering area in Oregon and Washington (primarily in the Willamette Valley), cackling geese congregate on private agricultural fields and eat crops, resulting in loss of agricultural yield for landowners. Selecting a population objective that balances the different stakeholder opinions is challenging. Multi-objective optimization is general, spanning many disciplines, and strategies used to solve MOO problems provide a framework for making defensible, transparent choices for natural resource management and ecological decisions in general.

2.3. THE MULTI-OBJECTIVE OPTIMIZATION PROBLEM

Multi-objective optimization assumes a decision maker can quantify the value of a decision with respect to the decision maker's objectives. Examples of natural resource management problems that explicitly quantified objectives include Maguire et al. (1987) who quantified objectives for the Sumatran rhino using the expected probability of extinction and cost. Johnson et al. (1997)

described an objective function that quantified the objectives of maximizing the expected cumulative mallard (*Anas platyrhynchos*) harvest through time and sustaining a population size >8.1 million mallards. Converse et al. (2013) quantified the probability of successful population establishment, cost, public relations, and others for whooping crane (*Grus americana*) reintroductions. A function that quantifies the value of the potential actions θ from a set of possible choices of actions Θ relative to an objective is termed an objective function (Keeney and Raiffa 1976, Williams et al. 2002). Objective functions are inherently subjective because they are used to quantify the aim or interest of a decision maker (Hennig and Kutlukaya 2007). For consistency with MOO literature, I denote the objective function using $f(\theta)$ (notation definitions are also reported in Table 2.1 for reference). Objective functions are synonymous with loss functions, utility functions, or reward functions described in other fields (Williams et al. 2002, Berger 2013). The actions a decision maker can consider can be either a discrete set (as was the case with the Sumatran rhino) or continuous (e.g., choosing a target population). Without loss of generality, I assume that the set of actions to consider is continuous. The set of actions from which a decision maker chooses is represented by Θ. A specific action in the set of Θ is represented by θ. The value of the objective function (or utility) for a specific action is represented by $f(\theta)$. When a decision maker has one objective to maximize, and the objective function is unimodal, the decision maker can simply choose the value for θ that optimizes the objective function $f(\theta)$ (Fig. 2.1**A**). Decisions become difficult when decision makers must consider more than one objective. A single optimal solution for multiple competing objective functions does not exist without compromise. There are many (possibly infinite) solutions that represent trade-offs among competing objectives. MOO is concerned with methods for choosing among these trade-off solutions.

The MOO problem is defined as:

$$f(\boldsymbol{\theta}^*) = \text{optimum}_\theta f(\boldsymbol{\theta}),$$

$$\text{where } f(\boldsymbol{\theta}) = (f_1(\boldsymbol{\theta}), f_2(\boldsymbol{\theta}), ..., f_k(\boldsymbol{\theta})),$$

(2.1)

$$\text{such that } g_j(\boldsymbol{\theta}) \leq c_j, \ j = 1, 2, ..., J,$$

$$\text{and } h_l(\boldsymbol{\theta}) = d_l, \ l = 1, 2, ..., L,$$

where $f_i(\boldsymbol{\theta})$ represent the k different, potentially competing objective functions, $f(\boldsymbol{\theta})$ is a set of the different objective functions, and g_j and h_l represent J inequality constraints and L equality constraints, respectively (Hwang and Masud 1979). "Optimum" refers to either the minimum or maximum depending on the nature of each objective function (e.g., minimize cost, maximize abundance). I use only the minimum in several subsequent equations to avoid redundancy. The elements of the vector of variables $\boldsymbol{\theta}$ are known formally as *design variables*. Design variables are the combination of choices the decision maker can choose for inputs into the decision problem. For example, if a decision maker with limited personnel is concerned with eradicating invasive species and can use both chemical and mechanical treatments, then: $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3]'$, where $\theta_1 =$ amount of chemical treatment, $\theta_2 =$ amount of mechanical treatment, and $\theta_3 =$ the amount of personnel hours. A decision maker must simultaneously consider the effect from the combined choices θ_1, θ_2, and θ_3. The set of possible design variables from which a decision maker can choose ($\boldsymbol{\Theta}$) is termed the *feasible design space* and is defined by all potential combinations of choices of $\boldsymbol{\theta}$ that meet the constraints (i.e., $\{\boldsymbol{\theta}|g_j(\boldsymbol{\theta}) \leq c_j, \ j = 1, 2, ..., J, \ h_l(\boldsymbol{\theta}) = d_l, \ l = 1, 2, ..., L\}$). The constraints limit the potential combinations of choices by formally considering items such as budgetary constraints, legal mandates, etc. The *feasible criterion space* includes the values of $f(\boldsymbol{\theta})$ for each $\boldsymbol{\theta}$ in the feasible design space (i.e., $f(\boldsymbol{\theta})|\boldsymbol{\theta} \in \boldsymbol{\Theta}\}$). Consider an example to clarify notation and concepts.

Assume a decision maker wants to maximize $f_1(\theta)$ and $f_2(\theta)$. The functions of $f_1(\theta)$ and $f_2(\theta)$ are general but might represent various natural resource management problems (e.g., satisfaction of two opposing stakeholders for with respect to possible densities of wolves in a management unit). Attainment of each objective depends on only one design variable $\theta \in \Theta$ (e.g., the possible densities of wolves that could be considered; Fig. 2.1**B**). Assume Θ includes the potential choices $[0 - 20]$. Also assume the inequality constraints: $g_1(\theta) = \theta \leq 17$ and $g_2(\theta) = -\theta \geq 3$. These constraints limit the choices of θ such that $3 \leq \theta \leq 17$ and the feasible design space are values θ in the domain $[3 - 17]$. That is, although it might be desirable to one stakeholder to have densities of wolves < 3 wolves per unit, managers might not want densities this low because the population might become extirpated (Fig. 2.1**B**). Smaller values of θ imply higher attainment of $f_1(\theta)$ and lower attainment of $f_2(\theta)$. The feasible criterion space are the 2×1 vectors $[f_1(\theta), f_2(\theta)]'$ for all $\theta \in [3 - 17]$. The feasible criterion space can be viewed graphically (when $k \leq 3$ and θ includes one variable) by plotting the paired values of $f_1(\theta)$ and $f_2(\theta)$ against each other on opposing axes (Fig. 2.1**C**). Optimal attainment of $f_1(\theta)$ and $f_2(\theta)$ (i.e., $\theta = \theta_1^*$ and $\theta = \theta_2^*$) occur at 4 and 16, respectively (Fig. 2.1**B**). No choice of θ simultaneously optimizes $f_1(\theta)$ and $f_2(\theta)$ (i.e., $\theta_1^* \neq \theta_2^*$); a single solution to the MOO problem does not exist. The decision maker must include additional information to reduce the set of potential solutions to a single solution; the additional information required is a decision maker's preferences among objectives.

2.4. PARETO OPTIMAL SOLUTIONS AND SPECIFICATION OF PREFERENCES

Pareto optimality is a concept in MOO in which optimality is defined with respect to trade-offs that are required to improve an objective (Deb 2001). A Pareto optimal solution is any solution in which there are no other candidate solutions that improve achievement of at least one objective without hindering the achievement of another objective. For example, if we assume an increase

in the numbers of cackling geese corresponds to an increase in the objective function associated with subsistence hunting, and a decrease in the objective function associated with agriculture, then any management goal is a Pareto optimal solution because it requires a trade-off in objectives. More rigorously, a Pareto optimal solution is an action $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}$ in which there is no other action $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ such that both $f(\boldsymbol{\theta}) \leq f(\boldsymbol{\theta}^*)$, and $f_i(\boldsymbol{\theta}) < f_i(\boldsymbol{\theta}^*)$ for at least one function i (note: I use the minimum, without loss of generality Deb 2001, Marler and Arora 2004). A MOO problem with competing objectives has a large (potentially infinite) set of Pareto optimal solutions. The Pareto set (or Pareto frontier or efficiency frontier) is the set of Pareto optimal solutions. The Pareto set excludes all choices that are dominated by at least one other solution (Deb 2001). A dominated solution is a solution in which there exists another solution that is as good or better for all objectives, and better for at least one objective. Consider the example in Fig. 2.1. Any value of θ between 4 and 16 is a Pareto optimal solution. For all $4 < \theta < 16$, to improve $f_2(\theta)$ requires a trade-off from $f_1(\theta)$. Likewise, to improve $f_1(\theta)$ requires a trade-off from $f_2(\theta)$. Note that decisions for $\theta < 4$ and $\theta > 16$ are dominated solutions; both objectives can be improved simultaneously by increasing or decreasing θ, respectively.

Methods for solving (2.1) can be separated into two strategies: specification of preferences pre-optimization (the *a priori* approach) or post-optimization (the *a posteriori* approach; Deb 2001, Marler and Arora 2004). A Pareto set exists in both strategies.

2.4.1. THE *a priori* STRATEGY. Every choice in a Pareto set corresponds to a specific ranking (or preference) of objectives. When decision makers specify preferences pre-optimization they are attempting to identify the choice in the Pareto set that most closely aligns with their *a priori* perceptions of the importance of each objective. The information used to assign preferences can include qualitative or quantitative values obtained from a variety of methods ranging from personal opinion to formal theoretical development (e.g., Akaike 1973, Saaty 1988, Mustajoki et al. 2005).

After preferences are assigned pre-optimization, the Pareto optimal solution associated with those preferences is identified. The identification of the Pareto solution associated with a set of preferences occurs using various mathematical functions (Table 2.2). These functions typically take one of two forms; using a composition of objective functions or constraining the feasible criterion space. A composition of objective functions combines the multiple objective functions into a single objective function (e.g., Converse et al. 2013). This typically requires specifying weights that represent the relative importance of each of the multiple objective functions. The optimum of the resulting composite function reflects the Pareto optimal solution associated with the selected weights/preferences. Constraining the feasible criterion space results in an optimization of (2.1) that occurs over a single objective function (e.g., Johnson et al. 1997). The constraints are set based on the preferences of the decision maker. The optimum of the single objective function represents the Pareto optimal solution associated with the constrained space. As an example of the *a priori* approach, I consider three different functions for assigning preferences pre-optimization to identify an optimal solution: the bounded objective function method, the lexicographic method, and the weighted sum method. The bounded objective function method and the lexicographic method constrain the feasible criterion space. The weighted-sum method is a composition of functions. These three methods, or combinations of them, are sufficiently general to cover a diverse array of ecological decision problems. I have also included several other methods in Table 2.2 for reference.

The bounded objective function method assigns preferences for k objective functions by constraining $k-1$ objective functions to preferred ranges of values, then optimizes the final objective function within the constrained space (Marler and Arora 2004). This effectively reduces the number of objective functions to one with a constrained feasible criterion space that meets the preferences (or are in some tolerable range) of all other objectives. The MOO problem in (2.1) for the

bounded objective function method is defined by

$$f(\boldsymbol{\theta^*}) = \min_\theta f_1(\boldsymbol{\theta}),$$

such that: $c_{l,i} \leq f_i(\boldsymbol{\theta}) \leq c_{u,i}, i = 2, ..., k$, where $c_{l,i}$ and $c_{u,i}$ represent the lower and upper bounds, respectively, for each objective function $i = 2, ..., k$. As an example, suppose θ represents the amount of land purchasable for conservation of an endangered species and ranges between 0 and 2,000 ha. Suppose we have two objective functions ($k = 2$). Let $f_1(\theta)$ be a function that returns a value of conservation to the endangered species from the amount of purchased land. For example, the function $f_1(\theta) = \frac{1.1 \times 10^6}{1 + e^{-0.005(\theta - 700)}}$ is a non-decreasing function that has diminishing marginal returns for the amount of land purchased and is on a similar scale to to $f_2(\theta)$ (Fig. 2.2). Let $f_2(\theta)$ represent the cost of land. Assume the cost of land is linearly related to the size of land and 1 ha of land costs \$1,000 (Fig. 2.2). The decision maker cannot simultaneously maximize $f_1(\theta)$ and minimize $f_2(\theta)$ because both are increasing functions of θ. Suppose the decision maker prefers cost ($f_2(\theta)$) be bounded to ≤ 1 million dollars, with 0 as a minimum bound. The optimal solution can be found using the bounded objective function method by minimizing $f_1(\theta)$ such that $0 \leq f_2(\theta) \leq 1,000,000$. The optimal solution is the maximum of $f_1(\theta)$ in the constrained feasible criterion space determined by $0 \leq f_2(\theta) \leq 1,000,000$ and equals 1,000 ha (Fig. 2.2). Constraints are a natural choice for monetary objectives (or other objectives with explicit boundaries in the feasible criterion space) because they represent realistic budgetary conditions.

The lexicographic method involves ordering objectives by importance and subsequent iterative optimization. The intuition behind the lexicographic method is simple. First identify the order of preference of each objective function; the most important function first (Fig. 2.3). Next identify the optimal solutions for the most important objective function (Fig. 2.3**B**). Given multiple optimal solutions to the most important objective function (i.e., multiple global optima), choose

the solutions that optimizes the second objective function (Fig. 2.3**C**). Continue through the re-maining objective functions until one solution is identified (Fig. 2.3**D**). The MOO problem for the lexicographic method is defined by

(2.2) $$\min_\theta f_i(\boldsymbol{\theta}),$$

subject to $f_j(\boldsymbol{\theta}) \leq f_j(\boldsymbol{\theta}_j^*)$, $j = 1, 2, ...i - 1$, if $i > 1; i = 1, 2, ...k$, where $f_j(\boldsymbol{\theta}_j^*)$ is the optimal value of the jth objective function. The objective functions are ranked in order of importance from $i = 1, 2, ..., k$ with $i = 1$ being the most important. Note that $f_j(\boldsymbol{\theta}_j^*)$ is unique but $\boldsymbol{\theta}_j^*$ is not necessarily unique. Therefore, the decision maker chooses the value of $\boldsymbol{\theta}_j^*$ that optimizes $f_i(\boldsymbol{\theta})$. The optimal solution is sensitive to the ordering of objective functions (Fig. 2.3). The lexico-graphical method is useful when a primary objective must be met. For example, when legislative mandates (e.g., endangered species act, migratory bird treaty act) require a decision maker meet one objective; given the first objective is met, optimize with respect to subsequent objectives. The lexicographic method has limited use with an increasing or decreasing primary objective function because the optimum of the primary objective function will occur at a boundary. The hierarchical method (Table 2.2) is a generalization of the lexicographic method that relaxes the constraints in (2.2) to be within some tolerance level (δ) of the optimal solutions of preceding objective functions. That is, $f_j(\boldsymbol{\theta}) \leq (1 + \frac{\delta_j}{100})f_j(\boldsymbol{\theta}_j^*)$.

The weighted-sum method and its variants are the most common methods for solving MOO problems across disciplines (see Williams 1998, Converse et al. 2013, for ecological applications). Common variants of the weighted-sum method include the analytic hierarchy process (Saaty 1990), the simple multi-attribute weighting technique (SMART; Edwards 1977), the preference ranking organization method for enrichment evaluations (PROMETHEE), the elimination and choice ex-pressing reality (ELECTRE), and several model selection criteria (which I demonstrate in Chapter

3). A weighted sum of multiple functions is described by:

$$(2.3) \qquad f(\boldsymbol{\theta}) = \sum_{i=1}^{k} w_i f_i(\boldsymbol{\theta}).$$

The weights w_i are chosen by the decision maker to reflect the importance of each objective. Methods for selecting the weights have been the focus of considerable research and discussion in fields ranging from statistics to social science (e.g., Akaike 1973, Wierzbicki 1986, Saaty 1988, Rao and Roy 1989, Goodwin and Wright 2004). Weights are often constrained such that $w_i \geq 0$ and $\sum_{i=1}^{k} w_i = 1$ to aid in interpreting weights as relative importance of objectives. The MOO problem using a weighted sum is:

$$(2.4) \qquad f(\boldsymbol{\theta}^*) = \min_{\boldsymbol{\theta}} \sum_{i=1}^{k} w_i f_i(\boldsymbol{\theta}),$$

such that $g_j(\boldsymbol{\theta}) \leq c_j$, $j = 1, 2, ..., J$, and $h_l(\boldsymbol{\theta}) = d_l$, $l = 1, 2, ..., L$. Consider the objective functions used in the lexicographic example (Fig. 2.3). Instead of the lexicographic method, suppose we apply the weighted-sum method to incorporate objective preferences into the MOO problem. Suppose each objective is equally important such that weights $w_i = 1/3$ for $i = 1, 2, 3$. The resulting objective function is: $f(\theta) = 1/3 f_1(\theta) + 1/3 f_2(\theta) + 1/3 f_3(\theta)$ (Fig. 2.4**A**). The optimal solution is $\theta^* = 1$ (*c.f.*, the lexicographic method in which $\theta^* = 0.8$). Had the weights been, for example, $w_1 = 1/2, w_2 = 1/2, w_3 = 0$, the optimal solution would have been $\theta^* = 0.73$ (Fig. 2.4**B**). The weighted-sum method has several advantages that contribute to its widespread use. The intuition is simple, it provides flexibility sufficient to cover a diverse array of problems, and it provides a method for examining the second strategy to MOO, selecting preferences post-optimization (through multiple optimizations over many different combinations of weights). Often, the real challenge with the weighted-sum method is agreeing on a specific set of weights.

The preceding three examples of specifying preferences pre-optimization are intuitively simple, yet cover a diverse array of applications. For problems in which none of the above methods are sufficient to reduce the problem to a solution, combinations of the methods could be applied. Consider a problem with four objectives: minimize cost, maximize species persistence, maximize non-hunting recreation, and maximize hunting recreation. Cost could be constrained at the onset of a problem by identifying the available budget. If species persistence must be obtained the lexicographic method could be used to constrain the decision space. Finally, non-hunting and hunting recreation could be reduced to a single objective using the weighted-sum method. The decision maker could optimize the combined recreation objective within the constrained decision space determined by cost and meeting the species persistence objective. The resulting choice would keep cost under budget, meet the species persistence objective, and maximize the combined non-hunting/hunting recreation objective, given the constraints. Alternative functions for specification of preferences pre-optimization are reported in Table 2.2. Each of these methods must meet certain criteria to ensure the resulting solution is part of the Pareto optimal set. Marler and Arora (2004) provide a thorough discussion of the requirements for Pareto optimality for the three methods described here and additional methods in Table 2.2.

Solutions to MOO problems using the *a priori* method are sensitive to the decision maker's choice of preferences and constraints. The level of sensitivity is unknown when preferences are set. It is recommended, that after a solution is identified, a sensitivity analysis be conducted to examine the robustness of the final solution to small changes in weights or preferences (Barron and Schmidt 1988, Rios Insua 1990).

2.4.1.1. *The* a posteriori *strategy.* Selecting Pareto solutions using the *a posteriori* strategy is similar to an exhaustive sensitivity analysis. To select a Pareto optimal solution using the *a posteriori* method, the decision maker first identifies as many Pareto optimal solutions as possible. Pareto

optimal solutions are identified using various algorithms including: evolutionary algorithms (Deb 2001), physical programming (Messac et al. 2001, Messac and Mattson 2002), normal boundary intersection methods (Das and Dennis 1997; 1999), normal constraint methods (Messac et al. 2003), graphical methods (e.g., Fig. 2.1**D**), and an iterated weighted-sum approach. Graphical methods include plotting the values of the objective functions for each action θ on opposing axes, and are applicable when $k \leq 3$ and $\boldsymbol{\theta}$ is one-dimensional. The iterated weighted-sum approach is the most common optimization technique, and is applicable for many ecological applications (but see Athan and Papalambros 1996, Das and Dennis 1997, for limitations). The iterated weighted-sum approach involves solving the optimization in (2.4) repeatedly, for different values of weights $\boldsymbol{w}$ (usually in equal increments from 0–1 for each w_i). Evolutionary algorithms have also increased in popularity due to general applicability and fast computation time. There is now a field of research known as evolutionary multiobjective optimization which refers to solving MOO problems using evolutionary algorithms (Deb 2001, Coello et al. 2007).

When a decision maker identifies the Pareto set they obtain information on the trade-offs required for each potential choice $\boldsymbol{\theta}$. The decision maker uses the knowledge of the resulting trade-offs among Pareto optimal solutions (and other information relevant to the problem) to choose a solution from the Pareto set. The choice implies a decision maker's preferences among objectives because each Pareto optimal solution is associated with a set of preferences. Thus, the main difference between pre- and post-specification of preferences is the former requires describing the relative preferences among objectives without any knowledge of the consequences. The latter uses consequences to aid the decision.

Identifying the Pareto set is not always trivial when optimizing a large number of multivariate objective functions. Development of computational methods and algorithms to identify a Pareto set is a primary focus in optimization research (e.g., Balachandran and Gero 1984, Deb 2001,

Shan and Wang 2005, Zitzler et al. 2008). Algorithms are assessed relative to their *necessary* and *sufficient* conditions required for Pareto optimality (Marler and Arora 2004). If an algorithm meets necessary conditions, the only solutions that are Pareto optimal are solutions to the algorithm. That is, all Pareto optimal solutions can be found using the algorithm. If an algorithm meets sufficient conditions, any solution to the algorithm is Pareto optimal. If an algorithm is sufficient but not necessary, any solution to the algorithm is Pareto optimal, but there are other Pareto optimal solutions than cannot be found with the algorithm. Any solution to (2.4) is sufficient for Pareto optimality (Athan and Papalambros 1996). However, whether a solution to (2.4) is necessary to be in the Pareto set depends on the shape of the feasible criterion space. A solution to (2.4) is necessary if the feasible criterion space is convex (Athan and Papalambros 1996). When the feasible criterion space is not convex, there will be Pareto optimal solutions that are not solutions to (2.4). Alternative algorithms (e.g., evolutionary algorithms, compromise programming) might be preferred in these cases.

When a Pareto set is identified in a MOO problem the decision maker is left with the task of choosing among the Pareto set to obtain a final decision. This is similar to the example of the Sumatran rhino described above, with the exception that five of the seven actions presented for the Sumatran Rhino were dominated by two actions (in terms of the objectives of minimizing extinction probability and cost) and would not appear in the Pareto set (Maguire et al. 1987). The advantage of Pareto optimization is that the decision maker first reduces the potential solutions to Pareto optimal solutions. Then the decision maker identifies the trade-offs inherent in the Pareto set. After the trade-offs are identified, any number of methods can be used to make a choice. For example, applying constraints (as is done for the bounded objective function method), identifying solutions that are the most robust to changes in objective functions, identifying "elbows"

in the Pareto frontier (i.e., where a small gain in one objective results in a large trade-off in another objective). The final step of choosing among the Pareto solutions requires problem-specific considerations that are subjective and depend on each decision.

2.5. APPLICATION OF MULTI-OBJECTIVE OPTIMIZATION TO SELECT A MANAGEMENT GOAL FOR CACKLING GEESE

I consider the resource management problem of choosing a management goal for cackling geese. The Pacific Flyway, in collaboration with the U.S. Fish and Wildlife Service, state and Canadian wildlife agencies, and representatives of Oregon and Washington farmers and Native Alaskan subsistence hunters, are in the process of revising the cackling goose management plan (Pacific Flyway Council 1999). An important part of the plan is the management goal. The management goal is currently 250,000 birds. Cackling goose abundance is estimated annually using aerial survey data (see Chapter 1). Annual management decisions related to hunting regulations are made based on the three-year average of abundance estimates. The hunting regulations are intended to keep the population size at or near the management goal. The Pacific Flyway must balance the competing interests of subsistence hunters in Alaska and private agricultural farmers in Oregon and Washington. Additional considerations the Pacific Flyway face include the satisfaction of recreational hunters, the ecological integrity of the community on the YK Delta, and other non-consumptive benefactors of the cackling goose population. I focus on the objectives related to subsistence hunters and farmers for this example. Based on these objectives, I first develop objective functions related to the management objectives. I then apply the weighted-sum method (with equal weights) to develop a composite objective function to maximize. I then examine post-optimization-specification of preferences by optimizing a large number of composite functions obtained from the weighted-sum method with varying stakeholder weights. My objective was to

demonstrate the applicability of MOO as a tool for selecting a management goal and not to provide a solution to the problem. Thus I use variables in place of values specific to the problem.

2.5.1. OBJECTIVE FUNCTIONS. MOO requires specification of objective functions. An objective function $f_i(\theta)$ associated with the management goal θ of cackling geese quantifies the value of each potential $\theta \in \Theta$ relative to the objective it represents. For pre-optimization specification of preferences, I consider three objective functions. Two objective functions, $f_s(\theta)$ and $f_a(\theta)$, represent the Pacific Flyway's competing objectives related to Native Alaskan subsistence hunters and agriculture in Oregon and Washington, respectively. The third objective function $f_b(\theta)$ is concerned with balancing the importance of the other two objective functions. All three objective functions depend on one variable θ, that represents the number of geese for the management goal.

The task of formulating objective functions is about the translation of an interest or aim to the formal language of mathematics (Hennig and Kutlukaya 2007). There are an infinite number of possible objective functions that might be chosen. For pragmatic reasons, it is necessary to consider only a small number of possible choices (Hennig and Kutlukaya 2007). I developed the objective functions based on simple axioms and hypotheses. These objective functions do not necessarily represent the views of the associated stakeholders and are only used for demonstrating the MOO procedure. In applied settings, objective functions should be developed in close collaboration with resource managers and/or stakeholders. Collaborating with these groups to develop objective functions helps ensure the objective functions accurately represent their beliefs. It also promotes acceptance of the decision analysis process. Without collaborative development of objective functions it is unlikely that managers/stakeholders will be satisfied with the solution, regardless of how closely the objective functions reflect their objectives. Formal techniques for soliciting information to develop objective functions is beyond the scope of this Chapter and I refer readers to other literature (e.g., Keeney and Raiffa 1976, Gregory et al. 2012).

I developed the objective function associated with subsistence harvest based on three axioms. The first axiom was that if hunting regulations did not permit subsistence hunting, the value of $f_s(\theta)$ would be zero. The current management plan closes hunting season when $\theta < 80,000$ birds. I denote the value at which hunting is closed as χ for generality. The second axiom was that an increase in the management objective θ was associated with larger values of $f_s(\theta)$. That is, more birds implied higher utility. The third axiom was diminishing marginal returns. When considering a very large population of geese (i.e., a population size θ in which a hunter providing reasonable effort could harvest as many geese as needed; e.g., 1 million birds) it is hard to distinguish between the utility of θ and $\theta + \tilde{\theta}$ (e.g., 1.01 millions birds). That is, the marginal utility of additional birds beyond θ is too small to be recognized by subsistence hunters. By the second axiom, $\theta + \tilde{\theta}$ is better than θ, but a large value of θ would already provide enough birds to maximize storage and consumption potential. The additional $\tilde{\theta}$ birds would provide little benefit, given the large population size of θ; $f_s(\theta)$ is bounded by some upper limit. The upper limit reflects maximum storage capacity and/or maximum consumptive capacity of subsistence hunters. The objective function should therefore be concave, increasing faster at low values of θ, then reaching an asymptote. I used Bernoulli's utility function for wealth to describe the value or utility of each management goal θ (eq. (2.5); Bernoulli 1954). I selected an upper limit beyond the cackling goose carrying capacity, K so the function was close to linear in the population range of $\chi, ..., K$ birds. A linear utility function would result in the largest management goal of all concave functions, and thus is an optimal choice for subsistence hunters (from a game-theoretic perspective). Any other concave function would result in a sub-optimal choice of management goal with respect to subsistence hunters. I scaled the resulting objective function between 0 and 1 for values of θ between 0 and K (Fig. 2.5**A**).

The impact of cackling geese on agriculture depends on both the number of birds in the population and the distribution of the wintering population relative to public and private land. The mechanisms that determine habitat use by cackling geese relative to public and private land and relative to abundance are unknown. I developed two hypotheses describing the extremes of the potential distribution of geese relative to public and private land with respect to abundance. The hypotheses were: 1) geese feed exclusively on private land and depredate crops proportionately to population size, and 2) geese first use public land until the number of geese approaches the public-land carrying capacity. The remaining geese use private land after the public-land carrying capacity is met and depredate crops proportionately to population size. I represented the public land carrying capacity with the symbol K_{pl}. I developed a negative linear function scaled between 0 and 1 to represent the first hypothesis ($f_{a,1}$; Fig. 2.5**B**). I developed a piece-wise linear function ($f_{a,2}$) that equaled 1 for $\theta = 0, 1, ..., K_{pl}$ and decreased linearly for $\theta > K_{pl}$ with slope equal to the slope of $f_{a,1}$ for the second hypothesis. Assuming each hypothesis was equally likely, I averaged the two hypotheses to obtain the agricultural objective function ($f_a(\theta)$; Fig 2.5**B**).

Finally, I developed an objective function ($f_b(\theta)$) that represented balancing competing stakeholder interests. The objective function $f_b(\theta)$ had small values when $f_s(\theta)$ and $f_a(\theta)$ were far apart and large values (i.e., 1) when $f_s(\theta) = f_a(\theta)$. This objective function assumed equal attainment of each objectives ($f_a(\theta)$ and $f_s(\theta)$) is preferred to unequal attainment of each objective. I used the negative squared-error loss function, scaled between 0 and 1 for $f_b(\theta)$ (Fig. 2.5**C**). The

mathematical form of each objective function was:

$$
f_s(\theta) = \begin{cases} 0 \,, \theta \leq \chi \\[2ex] \log(b + c\theta) \,, \theta > \chi \end{cases},
$$

$$
f_a(\theta) = 1/2 f_{a,1}(\theta) + 1/2 f_{a,2}(\theta),
$$

(2.5)

$$
f_{a,1}(\theta) = 1 - \frac{\theta}{\max(\theta)},
$$

$$
f_{a,2}(\theta) = \begin{cases} 1 \,, \theta \leq K_{pl} \\[2ex] 1 - \frac{\theta - K_{pl}}{\max(\theta)} \,, \theta > K_{pl} \end{cases},
$$

$$
f_b(\theta) = -\left(f_s(\theta) - f_a(\theta)\right)^2,
$$

for $\theta = 0, ..., K$. I scaled the functions $f_s(\theta)$, $f_a(\theta)$, and $f_b(\theta)$ between 0 and 1 using the equation $\frac{f(\theta) - \min(f(\theta))}{\max(f(\theta)) - \min(f(\theta))}$.

Given the assignment of objective functions, any of the methods described in the previous section or in Table 2.2 could be used to incorporate preferences among objectives to identify an optimal solution. The non-decreasing and non-increasing nature of $f_s(\theta)$ and $f_a(\theta)$, respectively, make several of the MOO techniques trivial. The bounded objective function will result in an objective at the minimum or maximum bound conditional on which objective is ranked highest. The lexicographic method will result in an objective at either 0, $\max(f_b(\theta))$, or K, conditional on what objective is ranked highest. The Pareto set includes all $\theta = 0, ..., K$. I considered the weighted-sum method to identify an optimal management goal.

Specifying weights for each objective is difficult when objectives represent stakeholder beliefs or values. Perhaps the most politically palatable set of weights for stakeholders with competing

objectives are those in which $w_1 = w_2 = ... = w_k$ for k stakeholders. I assumed $w_1 = w_2 = w_3 = 1/3$ representing equal weight among stakeholders and equal weight to the penalty for differences in $f_s(\theta)$ and $f_a(\theta)$. The resulting objective function was:

$$(2.6) \qquad f(\theta) = 1/3 f_s(\theta) + 1/3 f_a(\theta) + 1/3 f_b(\theta).$$

The weighted-sum method combined the multiple objective functions described in (2.5) to a single objective function that could be optimized using standard techniques (e.g., visually inspecting a plot). The resulting function $f(\theta)$ and the optimal solution θ^* are shown in Fig. 2.5**D**. The solution is sensitive to the choices of χ, K, K_{pl}, and the weights $\mathbf{w}$. The value χ is determined by hunting regulations and the easiest to identify (80,000 for cackling geese). The carrying capacity, K, could potentially be estimated using historical abundance data when available. The private land carrying capacity, K_{pl}, could be estimated using bioenergetic requirements of cackling geese and the amount of food produced on available public land (e.g., McWilliams and Raveling 2004). I developed a web-based application to explore the optimal management goal in which the decision makers can identify values of χ, K, and K_{pl} relevant to the decision problem (https://perrywilliams.shinyapps.io/popobjective).

2.5.2. POST-OPTIMIZATION SPECIFICATION OF PREFERENCES. The management goal θ^* shown in Fig. 2.5**D** is the Pareto optimal solution associated with the choice of weights $\mathbf{w} \equiv 1/3$. A decision maker might wish to view the sensitivity of the optimal management goal relative to the choice of weights, or alternatively, view the implied weights of a given management objective. To examine these trade-offs, I calculated the Pareto optimal set and plotted each Pareto optimal solution with respect to the implied weights of the solution (Fig. 2.6). To calculate the Pareto

optimal set I calculated optimal solutions of objective functions described by (2.6) with differing values of w_i, such that $w_i > 0$ and $\sum_{i=1}^{k} w_i = 1$ (i.e., values in the weight space; Fig. 2.6).

I plotted the optimal solution for each combination of weights in the weight space to examine how changes in preferences affect changes in the management goal, and which management goals were relatively robust to changes in preferences (Fig. 2.6). Examination of Fig. 2.6 reveals several important points related to the information obtained using an *a posteriori* approach that was not available using the *a priori* approach. First, the objective weights of 1/3 applied to each objective function appears robust to small changes in **w** because it is on a relatively flat surface in Fig. 2.6. Thus, the choice of θ^* as a management goal was relatively insensitive to the choice of weights. Second, as w_b approaches zero the small changes in the remaining weight assigned to w_a and w_s result in large differences in the resulting optimal management goal. Third, as w_b approaches one, differences in weights have a small effect on the optimal management goal. Thus, the most robust management goal is one that has a large values of w_b. The second and third points illustrate the importance of an objective function that balances the competing objective functions; without it, the management goal can reflect a large disparity between the implied weights given to each stakeholder; with it, the implied weights are relatively insensitive to the management goal.

Given the trade-off solutions shown in Fig. 2.6, I considered four methods for selecting a final management goal. The first method was to choose the Pareto optimal solution that was the most common solution to $\theta^* = \max\{f(\theta)\}$ for all combinations of $w_i, i = 1, 2, 3$ such that $\sum_{i=1}^{3} w_i = 1$ (i.e., the weight space), excluding management goals that were equal to 0 and K. The second method was the management goal that was most robust to change in stakeholder weights estimated using the standard deviation of 400 neighbors found using a 21×21 matrix centered at each point (for matrices that extended beyond the boundary, the boundary was used). The third method was bounding $w_a < w_s$ and maximizing $f_a(\theta)$. Finally, the fourth method was bounding $w_s < w_a$

and maximizing $f_s(\theta)$. The resulting optimal solutions are shown in Figs. 2.7, 2.8, and 2.9. The first method resulted in a solution similar to the solution obtained using the *a priori* method with equal stakeholder weights (Fig. 2.7). The second method also resulted in a solution similar to the solution obtained using the *a priori* method with equal stakeholder weights, but implied different stakeholder weights. Specifically, $w_a = 0$, $w_s = 0$, and $w_b = 1$ (Fig. 2.8). The third method resulted in a solution of $\theta^* = K$ (Fig. 2.8). This was due to Pareto optimal solutions occurring at K even when w_a was forced to less than w_s. This solution is not likely a viable option for cackling goose management. The fourth method resulted in a solution at the same location in the weight space as the second option (i.e., $w_a = 0$, $w_s = 0$, and $w_b = 1$; Fig. 2.8). Thus, three of four methods resulted in similar management goals for the cackling goose population. Thus, a decision maker could select the resulting management goal with the justification that it was robust to stakeholder weights and maximizes the combined objective functions.

2.6. DISCUSSION

Researchers and practitioners across disciplines face multi-objective decision problems on a regular basis. There are challenges at all levels of multi-objective problems. Stakeholder objectives can be difficult to quantify, alternative actions are not always clear, and how a decision relates to objectives is often uncertain. I discussed one component of the multi-objective problem: quantitative tools and concepts to identify a solution given well defined objective functions, a set of possible actions, and preferences among objectives. The MOO problem concerns selecting a Pareto optimal solution from a set of trade-off solutions. Selection occurs by specifying the preferences explicitly or implicitly among the multiple objectives. Selecting preferences occurs either pre-optimization in which the decision maker articulates their preferences among objectives without knowledge of the resulting outcome, or post-optimization when the trade-offs from each

Pareto optimal solution can be seen. The main difference between pre- and post-specification of preferences is the former requires describing the relative preferences among objectives without any knowledge of the consequences; the latter uses consequences to aid the decision. The value of this difference depends on the context of the decision problem.

I described three methods for optimizing a multi-objective problem using pre-optimization specification of preferences (the bounded objective function method, the lexicographic method, and the weighted sum method). The combination of these methods or alternative methods described in Table 2.2 are applicable to a diverse array of problems. The optimal solution is sensitive to the choice of method. However, in a review of studies that used different optimization procedures in parallel for the same application, Huang et al. (2011) found the recommended course of action did not vary significantly with the method applied. I found evidence that different optimization procedures resulted in similar outcomes in the cackling goose problem of selecting a management goal. The *a priori* strategy using the weighted-sum method with equal weights resulted in similar management goals to three of the four solutions obtained using the *a posteriori* strategy. However, one of the solutions from the *a posteriori* strategy resulted in a different solution; one that would not likely be acceptable from a management perspective, as the optimal solution was equal to the carrying capacity.

The weighted-sum method is a useful tool for both the *a priori* and *a posteriori* strategy. Variants of the weighted-sum method (e.g., SMART) have been used in many ecological applications (e.g., Reynolds 2001, Reynolds and Hessburg 2005, Converse et al. 2013). The specific weighted-objective function I used for cackling geese is analogous to the SMART technique (Edwards 1977). However, the weighted-sum method is more general and is used in applications ranging from engineering (Marler and Arora 2004) to statistical and mathematical tools including mixture models,

Fourier transforms, and model selection (see Chapter 3). Using the weighted-sum method to identify the Pareto set is computationally straightforward.

The *a priori* strategy is more common in ecological applications (e.g., Johnson et al. 1997, Brown et al. 2001, Reynolds 2001, Duke and Aull-Hyde 2002, Ananda and Herath 2003, Herath 2004, Redpath et al. 2004, Snyder et al. 2004, Converse et al. 2013), than the *a posteriori* strategy (e.g., Groot and Rossing 2011). The *a priori* strategy is prescriptive; given preferences, a solution is prescribed. The *a posteriori* strategy does not require explicitly describing preferences; all preferences can be examined, and a choice implies preferences, but is based on the trade-offs observed. Deb (2001) argues that an *a posteriori* strategy is more methodical, more practical, and less subjective. He also concedes that if preferences can be reliably articulated, there is no reason to identify trade-off solutions. The *A priori* strategy is useful in decision problems consisting of stakeholders and/or decision makers that do not agree on a solution, but might agree on the inputs of a decision problem. For example, if the stakeholders associated with the cackling goose management problem agreed on values of χ, K, K_{pl}, $\mathbf{w}$ and their respective objective functions, and adhered to the decision analysis framework, an optimal management goal could be prescribed. If they cannot agree on inputs into the decision problem, decision making methods other than MOO might be preferred (e.g., conflict resolution). Knowledge of consequences *via* an *a posteriori* strategy might bias a decision maker who desires a decision based on pre-optimization specification of preferences. These situations might result in a game-theoretic decision problem (Von Neumann and Morgenstern 2007) instead of a MOO decision problem. That is, when competing stakeholders select objective functions or other inputs into the decision problem to optimize θ with respect to their interests, and not to reflect the true relationship between $f(\theta)$ and their beliefs or values. Thus, an *a posteriori* strategy is likely more useful for a decision maker who has competing objectives but with non-competing interests. For example, decisions regarding optimal allocation of limited

resources as in the case of the Sumatran rhino. In these cases, knowledge of the trade-offs among

Pareto optimal solutions might facilitate the decision process.

I focused on examples related to natural resource management. Several other applications of multi-objective optimization relevant to ecology include model selection, life history evolution, and behavior ecology. Model selection is concerned with balancing the competing objectives of model fit and model complexity (Burnham and Anderson 2002). Numerous methods have been developed to address this problem, including the development of several information criteria (e.g., Akaike's information criterion, Schwartz' information criterion Akaike 1973, Schwarz et al. 1978). Information criteria are weighted averages of functions related to model fit (usually assessed using model likelihoods) and model complexity (Chapter 3). Life history evolution concerns balancing vital rates to maximize fitness (e.g., clutch size vs. survival probability; Lack 1947). Similarly, behavioral ecology concerns balancing competing interests such as finding food and avoiding predators (Mangel and Clark 1986). MOO provides a basis for developing models to investigate hypotheses related to life history and/or behavioral ecology.

Finally, although I limited the discussion of objective functions in this paper, appropriately modeling a decision maker's objectives is important for MOO to be useful, regardless of which strategy or optimization formula is used. Any optimal solution from a MOO problem is only good as the objective functions describing the aims or interests of the decision maker. Analogous to statistical models of ecological processes, objective functions are models of decision maker preferences. As such, objective functions do not perfectly reflect reality. However, parsimonious objective functions can be useful tools for facilitating complex decisions (Kendall 2001, Williams 2015, Williams and Hooten In Review). Further, in the case of objective uncertainty (e.g., Fig. 2.5**B**), multiple objective functions/models could be considered with the goal of reducing uncertainty among objective functions through time. Methods developed for reducing ecological model

uncertainty through time (i.e., adaptive resource management) could be employed for reducing

objective function uncertainty through time.

Table 2.1: Notation and definitions of components in multi-objective optimization.

Notation	Description
θ	Design variables. An action or choice a decision maker can choose. $\boldsymbol{\theta}$ implies decisions for >1 variable.
θ^*	An optimal solution.
Θ	The set or list of actions from which the decision maker can choose. Θ is the potential combinations of potential actions when a decision maker must make choices for >1 variable.
$f_i(\theta)$	Individual objective function that describes the value of each choice $\theta \in \Theta$.
$f(\boldsymbol{\theta})$	A set of multiple objective functions that depend on (potentially many) choices $\boldsymbol{\theta}$.
$g(\boldsymbol{\theta})$	Inequality constraints.
$h(\boldsymbol{\theta})$	Equality constraints.
Feasible design space	The choices of $\boldsymbol{\theta}$ that meet the constraints.
Feasible criterion space	The values of $\boldsymbol{f}(\boldsymbol{\theta})$ for the feasible design space.

Table 2.2: Standard functions or methods for incorporating decision maker preferences (in the form of objective weights w_i or goals b_j) into multi-objective optimization problems. For more details see included references or Marler and Arora (2004).

Name of function or method	Function ($f =$)	See for details
Weighted global criterion 1	$\sum_{i=1}^{k} w_i(f_i(\boldsymbol{\theta}))^p, \quad f_i(\boldsymbol{\theta}) > 0 \; \forall i$	Yu and Leitmann (1974), Zeleny (1982), Chankong and Haimes (1983)
Weighted global criterion 2	$\sum_{i=1}^{k} (w_i f_i(\boldsymbol{\theta}))^p, \quad f_i(\boldsymbol{\theta}) > 0 \; \forall i$	Yu and Leitmann (1974), Zeleny (1982), Chankong and Haimes (1983)
Weighted global criterion 3	$(\sum_{i=1}^{k} w_i(f_i(\boldsymbol{\theta}) - f_i^o)^p)^{1/p}, \quad f_i(\boldsymbol{\theta}) > 0 \; \forall i$	Yu and Leitmann (1974), Zeleny (1982), Chankong and Haimes (1983)
Weighted global criterion 4	$(\sum_{i=1}^{k} (w_i^p(f_i(\boldsymbol{\theta}) - f_i^o)^p)^{1/p}, \quad f_i(\boldsymbol{\theta}) > 0 \; \forall i$	Yu and Leitmann (1974), Zeleny (1982), Chankong and Haimes (1983)
Hierarchical	$f_j(\boldsymbol{\theta}) \leq (1 + \frac{\delta_j}{100})f_j(\boldsymbol{\theta}_j^*),$ $j = 1, 2, ... i - 1, i > 1, \qquad i = 1, 2, ... k$	Osyczka (1984)

Weighted Tchebycheff	$w_i(f_i(\boldsymbol{\theta}) - f_i^o)$	Marler and Arora (2004)
Augmented weighted Tchebycheff	$w_i(f_i(\boldsymbol{\theta}) - f_i^o) + \rho \sum_{j=1}^{k}(f_j(\boldsymbol{\theta}) - f_j^o)$	Steuer and Choo (1983)
Modified weighted Tchebycheff	$w_i(f_i(\boldsymbol{\theta}) - f_i^o + \rho \sum_{j=1}^{k}(f_j(\boldsymbol{\theta}) - f_j^o))$	Kaliszewski (1987)
Exponential weighted-criterion	$\sum_{i=1}^{k}(e^{pw_i} - 1)e^{pf_i(\boldsymbol{\theta})}$	Athan and Papalambros (1996)
Weighted-product	$\prod_{i=i}^{k}(f_i\boldsymbol{\theta})^{w_i}$	Bridgman (1922)
Goal programming	$\sum_{j=1}^{k}|d_j|, d_j = ||b_j - f_j(\boldsymbol{\theta})||$	Charnes and Cooper (1977)
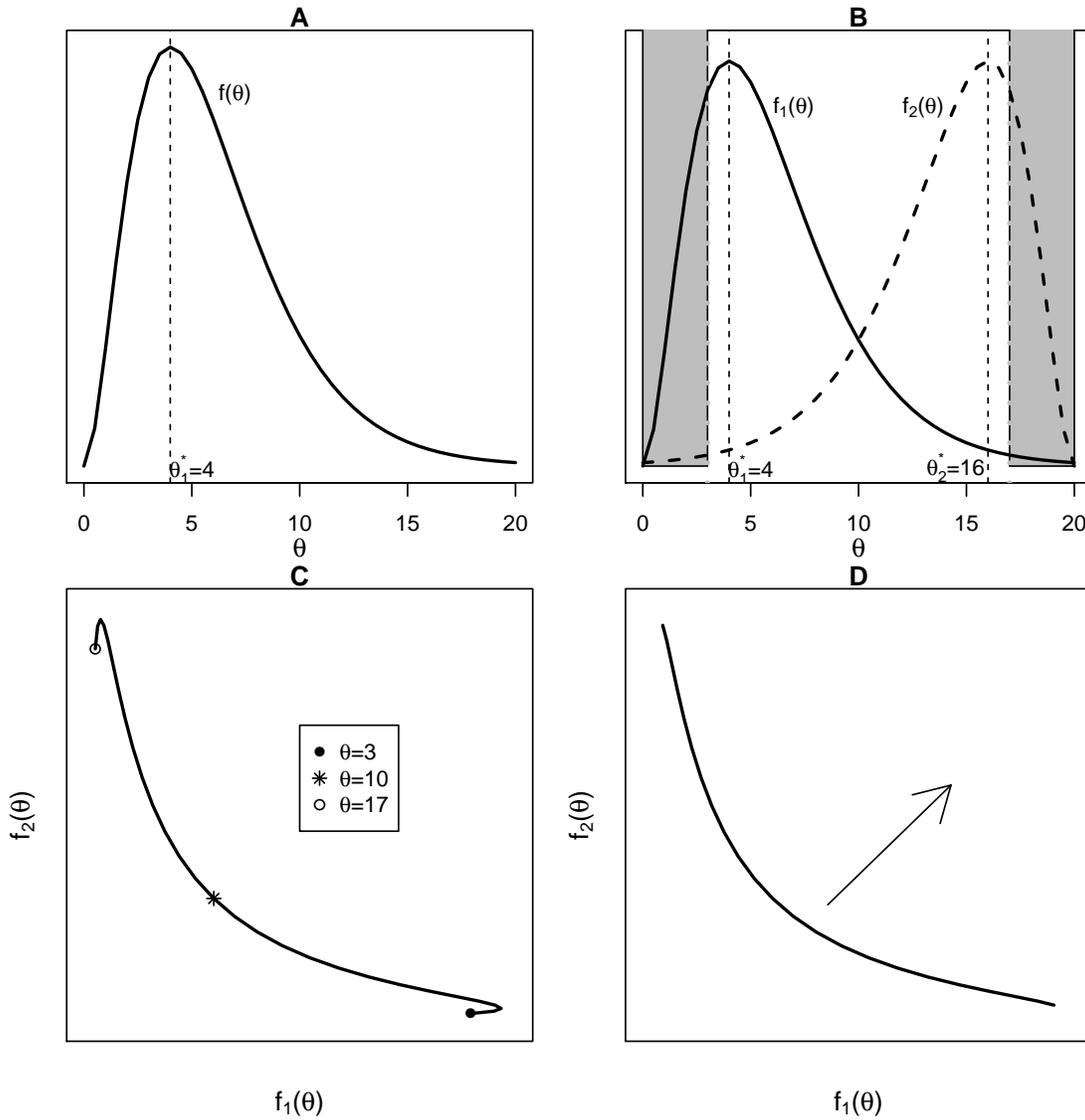
FIGURE 2.1. **A:** Single objective function $f_1(\theta)$ with optimal solution $\max_\theta(f_1(\theta)) = 4$. **B:** Multiple objective functions and feasible design space (i.e., possible choices for θ in the domain [3–17]; white area). The optimal values for each objective are: $\max_\theta(f_1(\theta)) = 4$ and $\max_\theta(f_2(\theta)) = 16$. Since $\theta_1^o \neq \theta_2^o$ no value of θ simultaneously optimizes $\boldsymbol{f}(\theta)$. **C:** Feasible criterion space for multi-objective optimization problem (i.e., $\{\boldsymbol{f}(\theta)|3 \leq \theta \leq 17\}$). **D:** Plot of the Pareto set (i.e., efficiency frontier or Pareto frontier), in which the dominated solutions in **C** were removed. Values towards the top right of the graph are preferred. Each solution on the Pareto frontier is Pareto optimal and represents a trade-off between objective 1 and objective 2.

cost > 1 million

$f_2(\theta)$ $f_1(\theta)$

$\theta^*=1000$
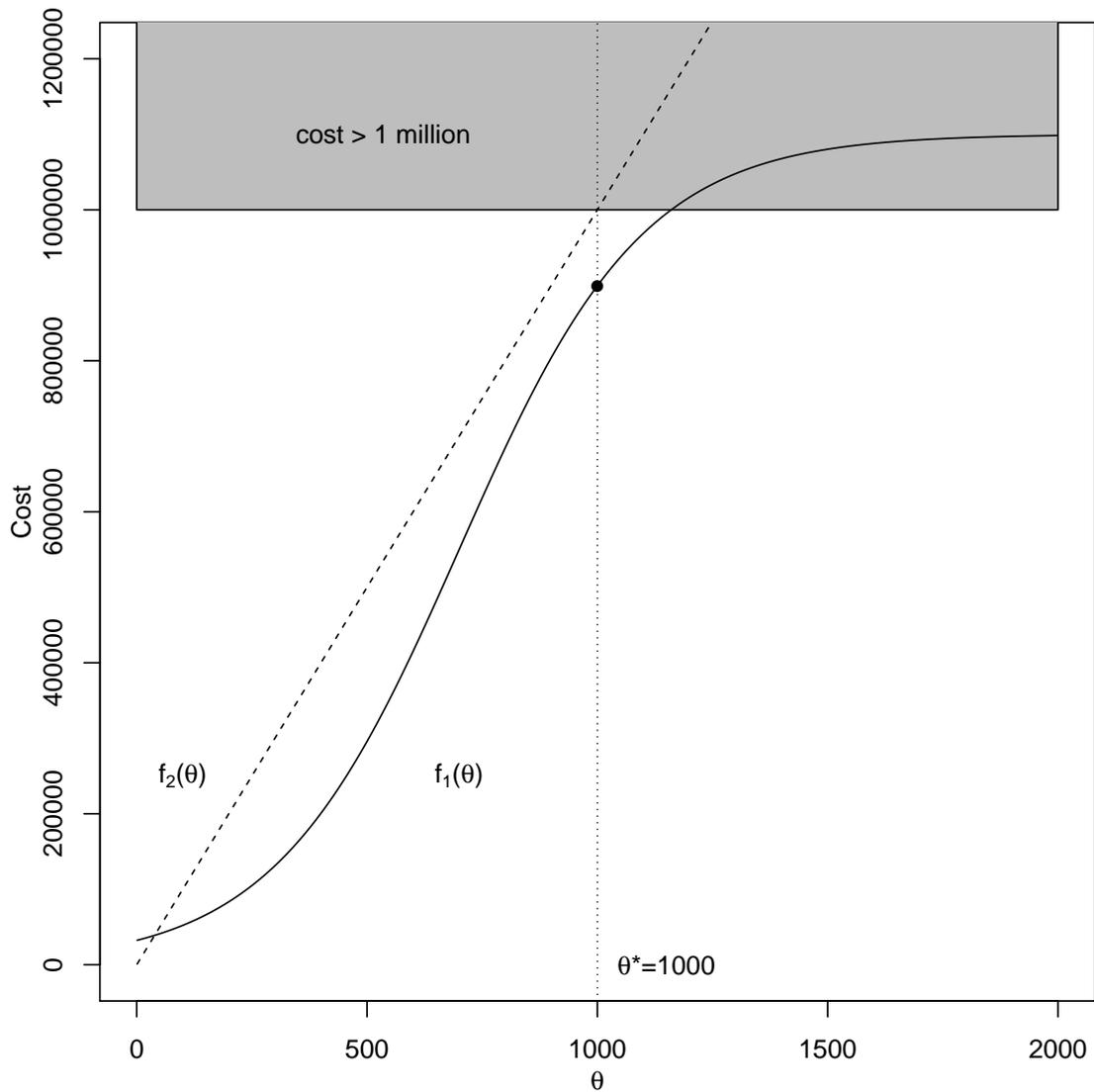
FIGURE 2.2. Example of *Bounded objective function method* with two objective functions representing the value of conservation to an endangered species (solid line) and price (dotted line) for different values of the amount of land that could be purchased (θ). Price was bounded to be between 0 and 1,000 ha. The optimal solution is the value that maximized the value of conservation within the bounds, and occurred at 1,000 ha.
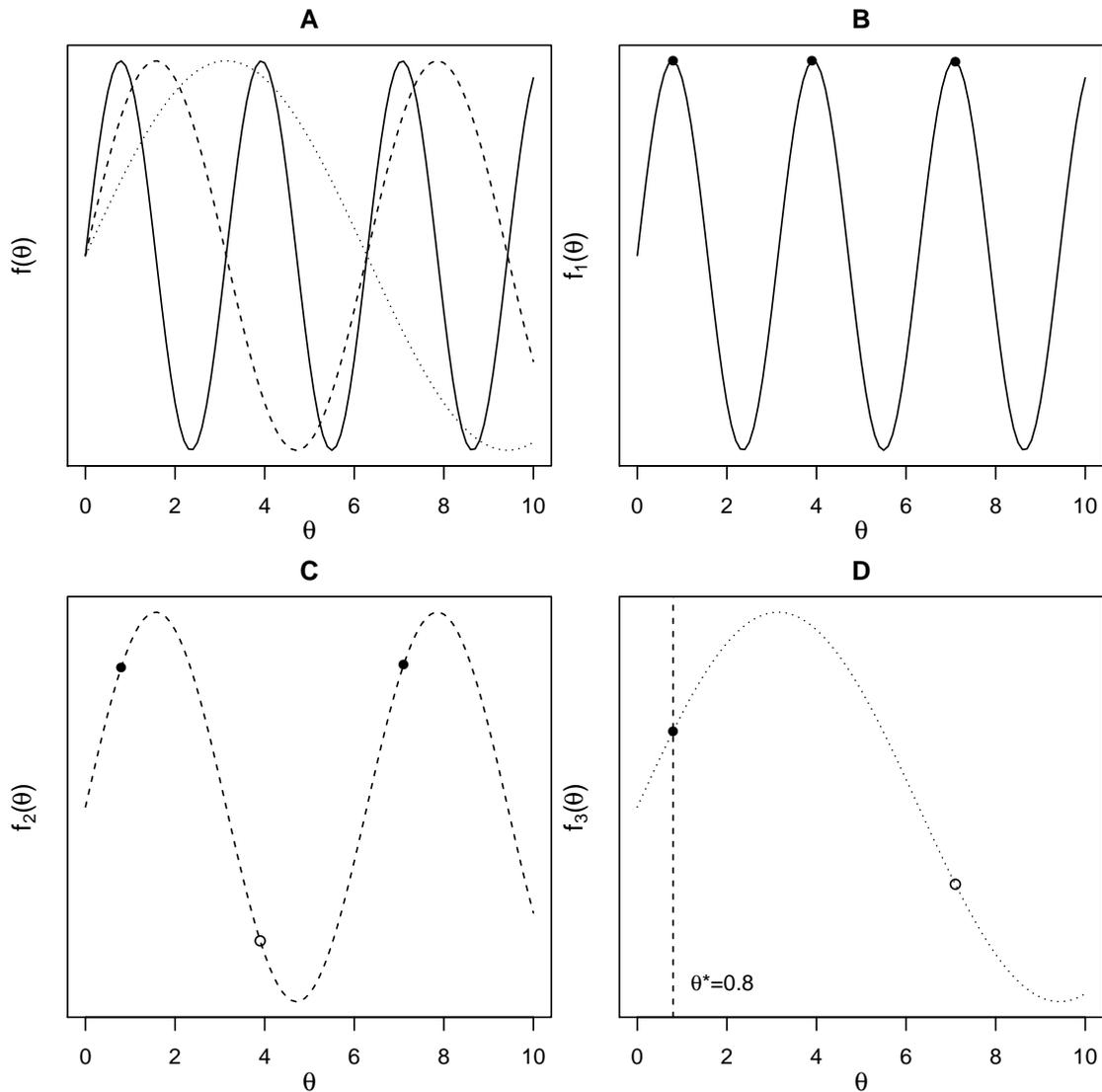
FIGURE 2.3. *Lexicographical method* for solving multi-objective optimization problems. **A:** Three objective functions with no common solution. **B:** Highest priority objective function and three optimal solutions. **C:** Second highest priority objective function with the three optimal solutions of the highest priority solution shown. The open dot represents a sub-optimal solution with respect to the second objective function. **D:** Least important objective function with the two optimal solutions of the second highest priority objective function. The final solution is the closed dot.
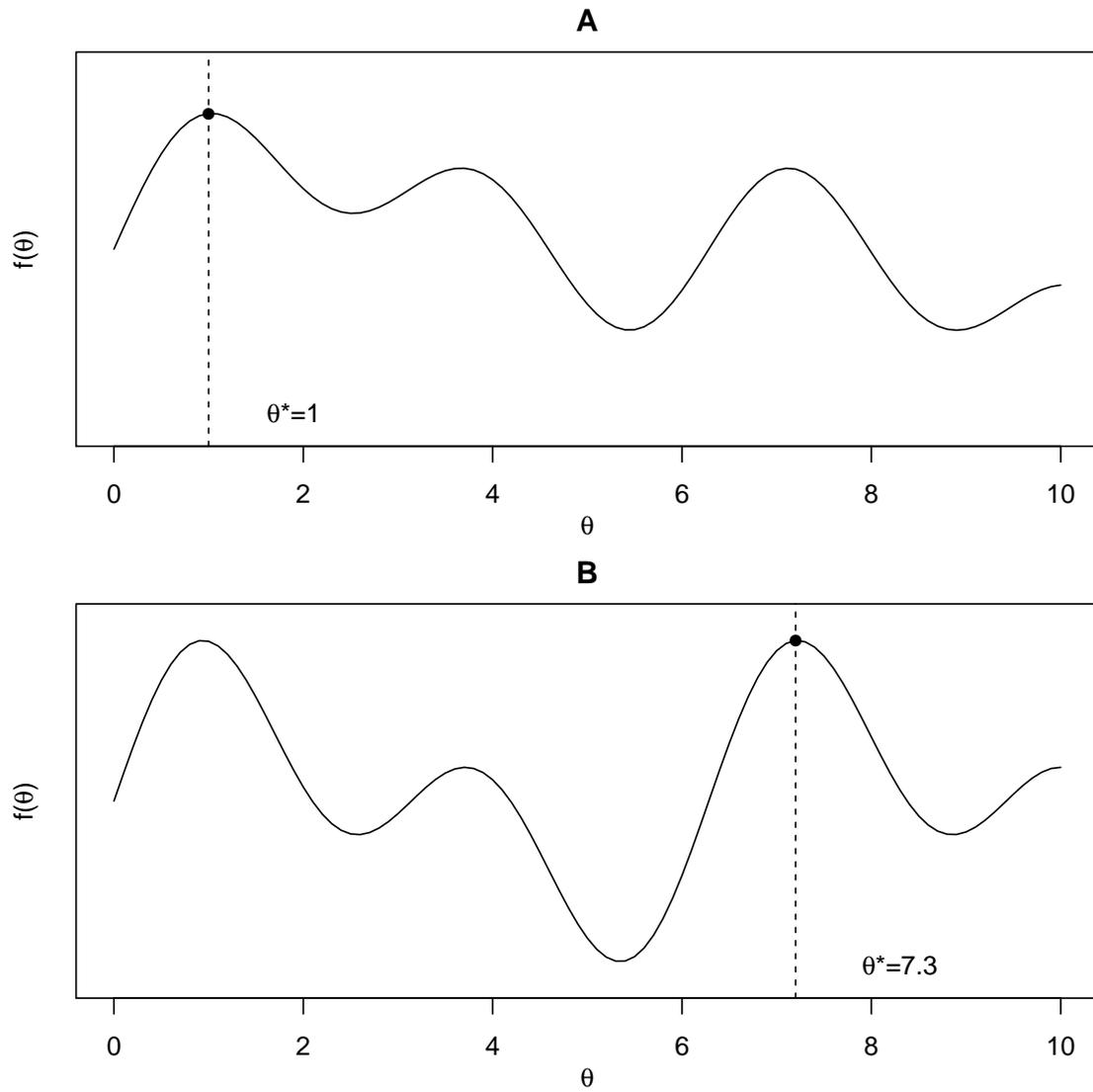
FIGURE 2.4. *The weighted-sum method* applied to the objective functions shown in Fig. 3**A**. The resulting function to optimize was **A**: $\theta^* = \max(1/3 f_1(\theta) + 1/3 f_2(\theta) + 1/3 f_3(\theta))$ and **B**: $\theta^* = \max(1/2 f_1(\theta) + 1/2 f_2(\theta) + 0 f_3(\theta))$.
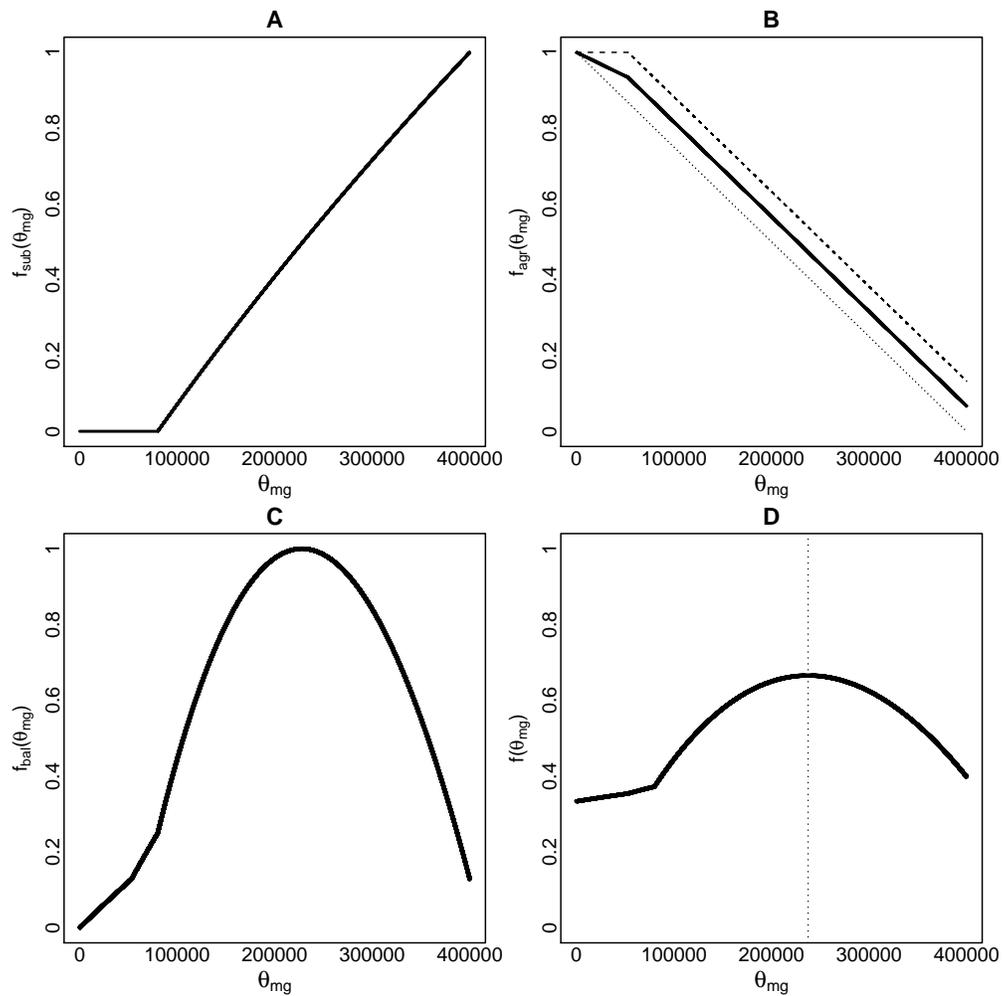
FIGURE 2.5. Objective functions for selecting a cackling goose management goal. **A**: Objective function representing the objective of maintaining a large population size to improve subsistence hunting opportunities. **B**: Objective function representing the objective of minimizing crop depredation in cackling goose wintering areas. The dashed line represents the hypothesis that cackling geese first use public land until the carrying capacity is reached (assumed to be 54,000 in this example), then decreasing proportionately to abundance. The dotted line assumes geese forage exclusively on private land. The solid line is the average of the dotted and dashed lines. **C**: Negative squared error loss between $f_a(\theta)$ and $f_s(\theta)$, scaled between 0 and 1. **D**: Weighted sum of $f_a(\theta)$, $f_s(\theta)$, and $f_b(\theta)$, with weights $\mathbf{w} \equiv 1/3$.

Weight space (front view) — **Mgmt. goals (front view)** — **Mgmt. goals (back view)** — **Mgmt. goals (side view)**
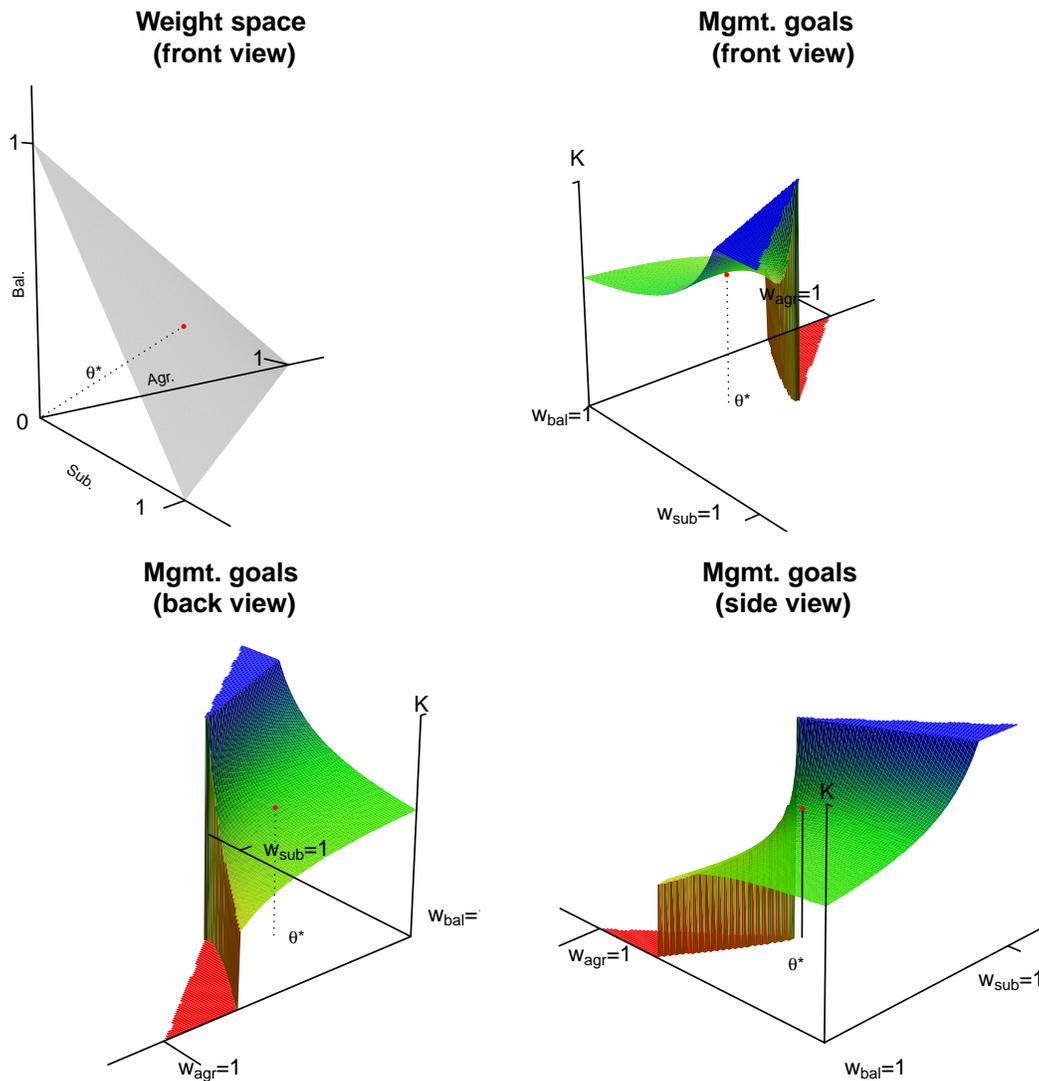
FIGURE 2.6. **Weight space:** Plane representing the possible combinations of $w_i, i = 1, 2, 3$ such that $\sum_{i=1}^{3} w_i = 1$ and $w_i \geq 0$ used in the equation $f(\theta) = \sum_{i=1}^{n} f_i(\theta) w_i$. The red dot indicates equal weights among the three objectives of the management goal of balancing the competing objectives of subsistence harvest, agriculture, and balancing the first two objective functions. **Mgmt. goals:** three angles of a surface of Pareto optimal solutions for a cackling goose management goal obtained by optimizing the weighted objective function with varying values of weights $w_i, i = 1, 2, 3$ such that $\sum_{i=1}^{3} w_i = 1$. The X-Y plane of the three colored figures is the weight space. The Z-axis is the value of the optimal management goal for each combination in the weight space. The optimal solution obtained from equal weights = 1/3 is shown with the red point.
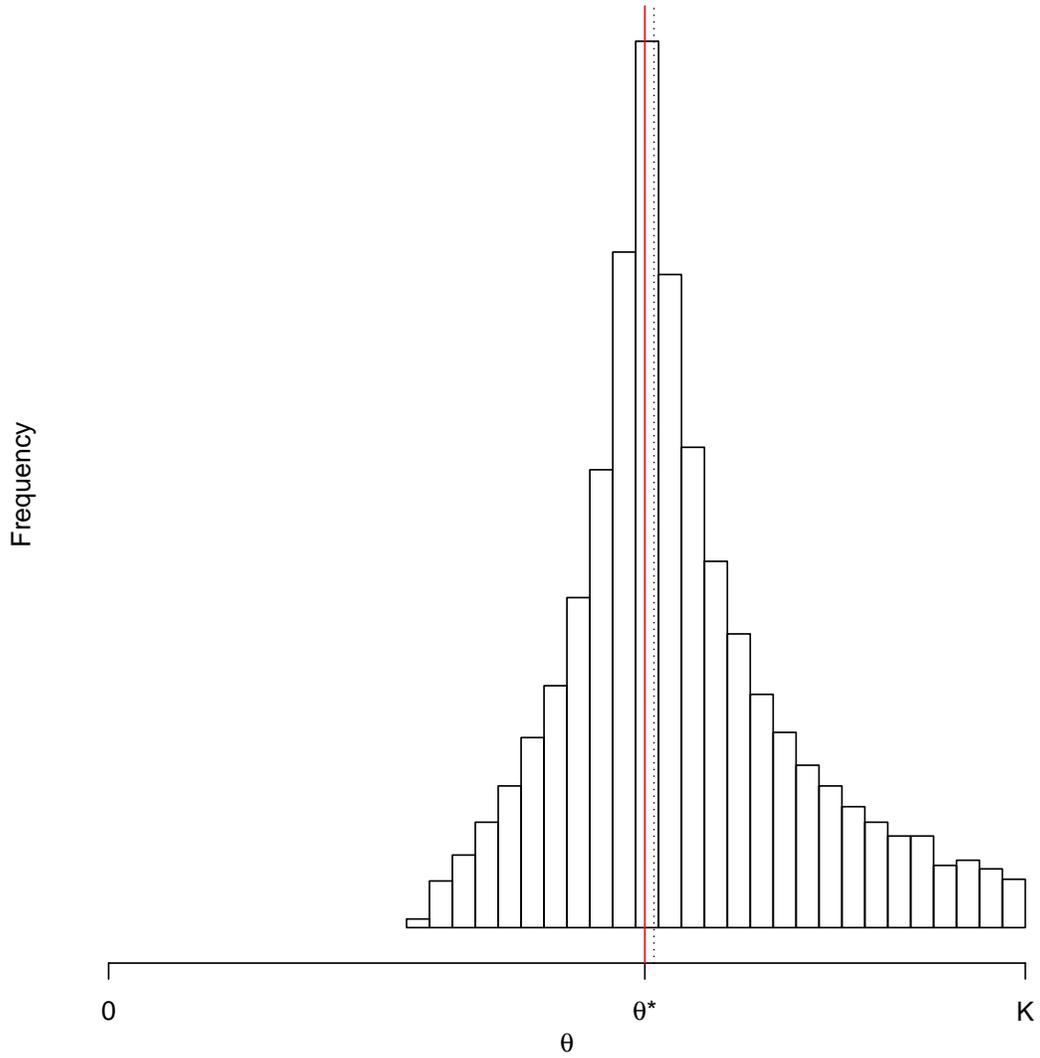
FIGURE 2.7. Distribution of optimal management goals found using the weighted-sum method with varying weight combinations of $w_i, i = 1, 2, 3$ such that $\sum_{i=1}^{3} w_i = 1$ and $w_i \geq 0$ used in the equation $f(\theta) = \sum_{i=1}^{n} f_i(\theta) w_i$. The red vertical line represents the mode; the management goal that was most often selected as optimal. Also shown (black vertical dotted line) is the Pareto optimal solution obtained using equal weights among objectives.
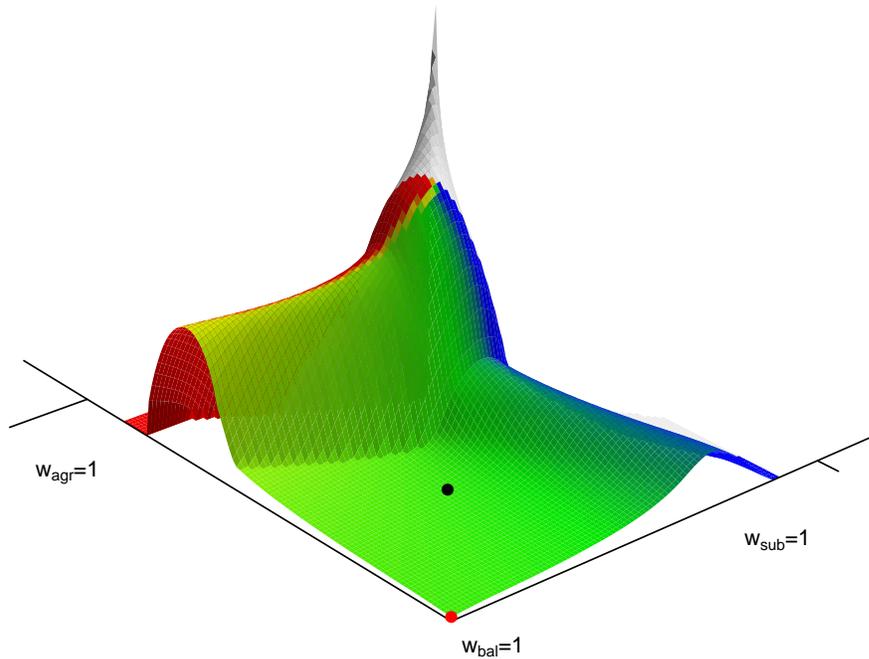
FIGURE 2.8. Surface of the standard deviation (SD) of Pareto optimal management goals shown in Fig. 2.6 used as an indicator of robustness of the management decision to changes in objective weights. The SD was taken with respect to a 21×21 matrix of neighboring values centered at each point with boundaries repeated ten times at each boundary. The optimal management goal with with the smallest standard deviation (red dot) occurs at $w_a = 0$, $w_s = 0$, and $w_b = 1$. Also shown (black dot) is the Pareto optimal solution obtained using equal weights. The colors used in Fig. 2.6 were maintained for reference.
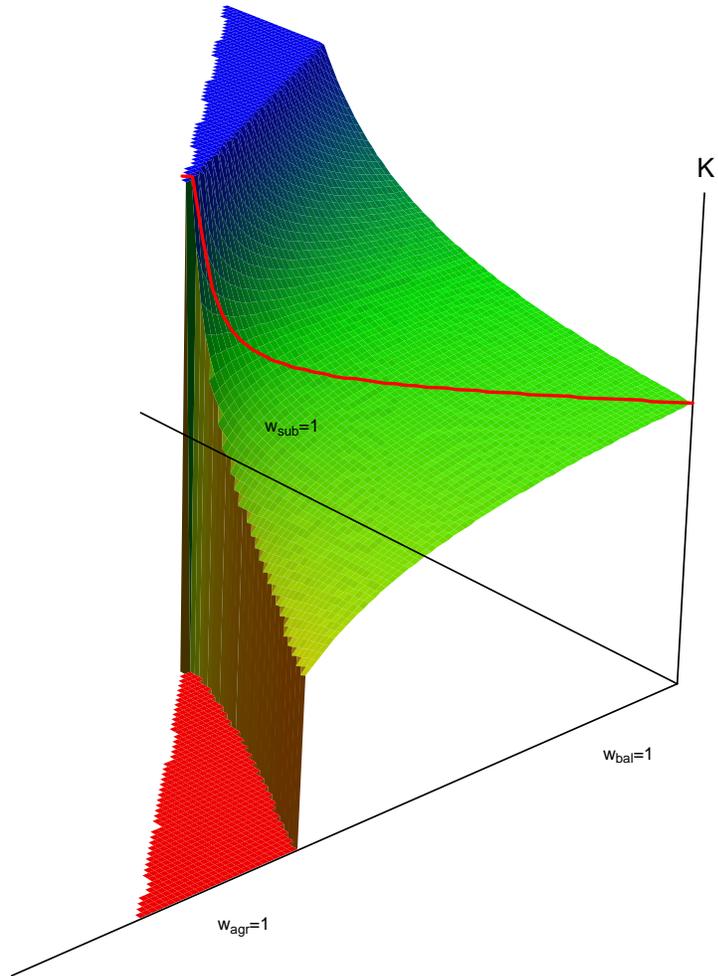
FIGURE 2.9. Surface of Pareto optimal solutions for a cackling goose management goal obtained by optimizing the weighted objective function with varying values of weights $w_i, i = 1, 2, 3$ such that $\sum_{i=1}^{3} w_i = 1$. The red line represents equal weights for the subsistence objective and the agriculture objective. Thus, for the decision rule that bounds $w_a < w_s$ and chooses the max with respect to $f_a(\theta)$ occurs at $\theta^* = K$. The decision rule that bounds $w_s < w_a$ and chooses the max with respect to $f_s(\theta)$ occurs at $w_a = 0$, $w_s = 0$, and $w_b = 1$; the same result that occurred using the standard deviation method (Fig. 2.8).

MODEL SELECTION AS A MULTI-OBJECTIVE OPTIMIZATION PROBLEM

3.1. SUMMARY

Multiple-objective optimization (MOO) is a formal decision-theoretic framework for solving multiple objective problems. Model selection is a MOO problem concerned, heuristically, with balancing the competing objectives of model fit and model complexity. I describe the model-selection problem within the MOO framework. I discuss two strategies of solving the MOO problem; modeling preferences pre-optimization and post-optimization. Most model selection methods are consistent with solving MOO problems *via* specification of preferences pre-optimization. I reconcile these methods within the MOO framework. I also consider model selection using post-optimization specification of preferences. That is, by first identifying Pareto optimal solutions, and then selecting among Pareto optimal solutions. Two concepts commonly used in model selection, competing models (i.e., considering models with ΔAIC< 2) and model averaging (*sensu* Burnham and Anderson 2002, p. 150), do not necessarily result in Pareto optimal solutions to the MOO problem of maximizing model fit and minimizing model complexity. Competing models and model averaging represent solutions to an additional objective function associated with minimizing model uncertainty. I demonstrate concepts with an application to variable selection in multiple linear regression.

Key words competing models, decision theory, multimodel inference, Pareto frontier, optimal solution

3.2. INTRODUCTION

Multi-objective optimization (MOO) is a formal decision-theoretic framework for optimizing problems with more than one objective (Chapter 2). MOO is commonly used in engineering, economics, and other fields in which decision makers must balance trade-offs between ≥ 2 competing

objectives (Marler and Arora 2004). For example, maximizing species persistence, while minimizing cost (Maguire et al. 1987). Often, a single solution that optimizes multiple objectives does not exist. When a decision maker has competing objectives, a solution that is optimal for one objective might not be optimal for the other objective. With competing objectives there exists many (possibly infinite) solutions that might be considered "optimal" (i.e., Pareto optimal; Chapter 2). However, in most decision contexts, a decision maker can only make one choice. To choose among potential optimal solutions, a decision maker must include their preferences among objectives to identify a final solution. MOO provides a mathematical framework for quantifying preferences for examining multi-objective problems. The MOO framework is described generally as

$$f(\boldsymbol{\theta}^*) = \text{optimum}_\theta f(\boldsymbol{\theta}),$$

$$\text{where } f(\boldsymbol{\theta}) = (f_1(\boldsymbol{\theta}), ..., f_k(\boldsymbol{\theta})),$$

(3.1)
$$\text{such that } g_j(\boldsymbol{\theta}) \leq c_j, \ j = 1, 2, ..., J,$$

$$\text{and } h_l(\boldsymbol{\theta}) = d_l, \ l = 1, 2, ..., L,$$

where $f(\boldsymbol{\theta})$ is a set of k different, potentially competing objective functions $f_i(\boldsymbol{\theta})$, $i = 1, ..., k$, and g_j and h_l represent J inequality constraints and L equality constraints, respectively (Marler and Arora 2004, Cohon 2013). Pareto optimality is a concept of optimality used for (3.1) when no value of $\boldsymbol{\theta}$ simultaneously optimizes each function f_i. A Pareto optimal solution for a minimization problem is a solution $\boldsymbol{\theta}^* \in \Theta$ in which there is no other solution $\boldsymbol{\theta} \in \Theta$ such that both $f(\boldsymbol{\theta}) \leq f(\boldsymbol{\theta}^*)$, and $f_i(\boldsymbol{\theta}) < f_i(\boldsymbol{\theta}^*)$ for at least one function i (Deb 2001, Marler and Arora 2004). For decision problems with competing objectives, there are many, (potentially infinite) Pareto optimal solutions. The set of solutions that are Pareto optimal are known as the Pareto set (or Pareto frontier or efficiency frontier). Each solution in a Pareto set has an implied set of preferences for the objective functions f_i (Deb 2001). Thus, choosing among a set of Pareto optimal solutions requires assuming (either implicitly or explicitly) preferences for the objective functions f_i. Preferences among objective functions can be specified pre- or post-optimization, representing two separate strategies

to solving (3.1) (see Chapter 2). When specifying preferences pre-optimization, decision makers explicitly describe preferences of objective functions and select the Pareto optimal solution associated with their choice of preferences. When specifying preferences post-optimization, decision makers first examine the set of Pareto optimal solutions. Then the decision maker chooses the final Pareto optimal solution based on the trade-offs observed among the set. The choice implies decision-maker preferences.

One of the most common methods for incorporating preferences for f_i into a decision problem pre-optimization is the weighted-sum method (Athan and Papalambros 1996, Das and Dennis 1997, Cohon 2013). The weighted-sum method is described by

$$(3.2) \qquad f(\boldsymbol{\theta}) = \sum_{i=1}^{k} w_i f_i(\boldsymbol{\theta}),$$

in which the optimal solution is

$$(3.3) \qquad f(\boldsymbol{\theta}^*) = \text{optimize}_{\boldsymbol{\theta}} \sum_{i=1}^{k} w_i f_i(\boldsymbol{\theta}).$$

The weights w_i are chosen by the decision maker to reflect the importance of each objective function f_i. The weighted-sum method is a composition of functions that results in a single objective function over which to optimize. When optimizing one objective function, an unequivocal optimal choice can be made.

Model selection is a multiple-objective problem that can be considered in terms of balancing the competing objectives of model fit and model complexity. Model selection is one of the most common multiple-objective problems across disciplines and numerous model-selection methods are available (e.g., Akaike 1973, Mallows 1973, Schwarz et al. 1978, Gelfand and Ghosh 1998, Burnham and Anderson 2002, Hooten and Hobbs 2015). There is no consensus among statisticians on best methods for model selection (Hooten and Hobbs 2015).

My objective was to examine model selection within the MOO framework and to demonstrate that several methods commonly used for model selection are specific cases of the MOO problem solved using the weighted-sum method with *a priori* specification of preferences. I examine concepts of the MOO framework, specifically Pareto optimality, as it relates to several common model

selection methods. Finally, I examine the second strategy of MOO, post-specification of preferences, and its application to the model selection problem. I demonstrate the concepts presented using a simple example involving variable selection in a multiple linear regression model used to model a simulated data set.

3.3. MODEL SELECTION AS A MOO PROBLEM

Methods for model selection typically consist of minimizing a weighted sum of two functions. Heuristically, these functions represent 1) model fit and 2) model complexity. That is,

$$(3.4) \qquad f(\boldsymbol{\theta}^*) = \min_{\boldsymbol{\theta}} \sum_{i=1}^{2} w_i f_i(\boldsymbol{\theta}).$$

where $\boldsymbol{\theta}^*$ represents the optimal solutions from the set of design variables $\boldsymbol{\theta}$, describing fit and complexity of any model, w_i are weights given to the importance of the objectives associated with model fit and complexity, and f_i are functions that quantify the value of model fit and complexity. Clearly, (3.4) is a specific form of the MOO problem defined in (3.3). A large body of theoretical justification within the field of statistics has been developed for choices of objective functions $f_i(\boldsymbol{\theta})$ and their corresponding weights w_i (Akaike 1973, Mallows 1973, Schwarz et al. 1978, Gelfand and Ghosh 1998, Burnham and Anderson 2002, Link and Barker 2006, Hooten and Hobbs 2015). Although there is no consensus among statisticians on specific model selection methods, most of the theoretical development related to model selection can be described by two general functions for f_i. Differences in model selection criteria are often the result of different choices in weights. The most common objective function for model fit is the negative log-likelihood of the data, given parameters. That is, if f_1 is the objective function associated with model fit, it is described as

$$(3.5) \qquad f_1(\boldsymbol{\theta}) = -\log(L(\boldsymbol{\theta}|\boldsymbol{y})).$$

Although the deviance (3.5) is the most common objective function for model fit, others have been used. For example in Mallow's C_p, $f_1(\boldsymbol{\theta}) = \frac{\sum_{i=1}^{n}(y_i - \hat{\mu}_{sub})^2}{\sum_{i=1}^{n}(y_i - \hat{\mu}_{full})^2} - n$, where $\hat{\mu}_{sub}$ equals the estimated mean of a sub-model in consideration, $\hat{\mu}_{full}$ equals the estimated mean of the full model in consideration, and n equals the sample size (Mallows 1973). Hooten and Hobbs (2015) reviewed several

objective functions for model complexity using a function proportional to

$$(3.6) \qquad f_2(\boldsymbol{\theta}) = \sum_{j=1}^{p} |\theta_j - \mu_j|^{\gamma},$$

known as the *regulator*, or *regularizer*, or *penalty*. In (3.6), p is the number of parameters in the model and γ is the degree of the norm; a user-defined parameter that controls the relative penalty of the distance between θ_j and mean μ_j. Substituting the choices of $f_1(\boldsymbol{\theta})$ and $f_2(\boldsymbol{\theta})$ from (3.5) and (3.6) into (3.4), we obtain the following multi-objective optimization problem, with the objective of $\min_{\boldsymbol{\theta}}(f(\boldsymbol{\theta}))$:

$$
\begin{aligned}
f(\boldsymbol{\theta}) =& w_1 f_1(\boldsymbol{\theta}) + w_2 f_2(\boldsymbol{\theta}), \\
(3.7) \qquad =& w_1(-\log(L(\boldsymbol{\theta}|\mathbf{y}))) + w_2 \sum_{j=1}^{p} |\theta_j - \mu_j|^{\gamma}.
\end{aligned}
$$

Equation (3.7) is the general function used in many model selection methods including Akaike's information criterion (AIC), AIC for small samples (AIC$_c$), quasi-AIC (QAIC), QAIC for small samples (QAIC$_c$), Schwartz's information criterion (BIC), ridge regression, LASSO (least absolute shrinkage and selection operator), natural Bayesian shrinkage, and some forms of posterior predictive loss (Table 3.1; Gelfand and Ghosh 1998, Hooten and Hobbs 2015). Each of the listed model selection methods result from specific choices of $\boldsymbol{w}$ and γ, which I report in Table 3.1. For example, let the weights be: $w_1 = 2$, $w_2 = 2$, and set γ to zero. With these weights, (3.7) simplifies to $-2\log(L(\boldsymbol{\theta}|\boldsymbol{y})) + 2p$, or AIC (Table 3.1).

Expressing model-selection methods in terms of (3.3) has an important result that links model selection to Pareto optimality. For positive weights $\boldsymbol{w}$, any solution to (3.3) is a Pareto optimal solution (Marler and Arora 2010). Thus, any model-selection method that can be expressed in terms of (3.7) (i.e., the methods in Table 3.1) results in a solution that is Pareto optimal with respect to the objectives of maximizing model fit and minimizing model complexity.

Solving a MOO problem with competing objectives using post-optimization specification of weights requires first identifying as many Pareto optimal solutions as possible, then choosing among the Pareto optimal solutions (Chapter 2). Pareto optimal solutions for the objective functions (3.5) and (3.6) are models in which increasing the value of (3.5) requires a decrease in the values in (3.6), and *vice versa*. One method for identifying Pareto optimal solutions with two objective functions, each depending on $\boldsymbol{\theta}$ is to plot the values of (3.5) and (3.6) for each candidate model on opposing axes to identify the Pareto frontier (e.g., Fig. 3.1). After the Pareto frontier is identified, the decision maker can select the model based on the trade-offs observed in the Pareto frontier. Thus, the selection of the final model is made without explicitly choosing weights $\boldsymbol{w}$ associated with the model selection criteria listed in Table 3.1. However, if a choice from the Pareto frontier is also optimal with respect to specific model-selection criterion, the weights of that selection criterion are implied.

3.5. AN EXAMPLE: VARIABLE SELECTION IN MULTIPLE LINEAR REGRESSION

Model selection is regularly used to select variables for multiple linear regression. To examine the variable selection problem within a MOO framework, I first generated a data set consisting of 300 observations and corresponding covariates. I then chose a top model using values for (3.7) specific to AIC. I also considered models with ΔAIC< 2 (i.e., competing models), and calculated a model using model averaging with AIC weights (*sensu* Burnham and Anderson 2002, p. 150). Finally, I identified all Pareto optimal solutions and assessed the AIC top model, candidate models, and the model calculated using model averaging relative to the Pareto optimal solutions.

3.5.1. DATA GENERATION. I generated response data from the model $y_i \sim \mathrm{N}(\mathbf{x}'\boldsymbol{\beta}, 10)$, for $i = 1, ..., 300$. I selected values of $\boldsymbol{\beta}$ that had a range of both negative and positive effects on the mean; $\boldsymbol{\beta} = [5, -10, -5, -1, -0.1, 0.1, 1, 5, 10]'$. I generated the design matrix $\mathbf{X}$ using: $x_{i,1} \equiv 1, \forall i$ and $x_{i,j} \sim \mathrm{Binom}(1, 0.5), j = 2, ..., 9$. I developed a scenario reflecting an investigator who collected data on four covariates that were related to the response variable y by randomly selecting four columns of $\mathbf{X}$ that were used to generate the data (the β values associated with the random

columns of $\mathbf{X}$ that were selected were -10, -5, 1, and 5). I also assumed the investigator collected four covariates that were independent of y. I accounted for this independence by simulating data using $x_{i,j} \sim \text{Binom}(1, 0.5)$, $j = 6, ..., 9$. Thus, the data included a response variable with 300 observations, and a design matrix with an intercept, four covariates that were used to generate the response variable, and four covariates that were independent of the response variable. Four of the nine covariates used to generate the data not randomly selected were omitted from analysis (i.e., the columns of $\mathbf{X}$ associated with β equal to $-1, -0.1, 0.1$, and 10). The data to be analyzed reflected the scenario of an investigator with imperfect information in the form of missing covariates and superfluous covariates on a process of interest.

3.5.2. MODEL SELECTION USING AIC. I analyzed the data using linear regression models of the form: $y_i \sim \text{N}(\boldsymbol{x}_i'\boldsymbol{\beta}_i, \sigma^2)$. The models were comprised of covariates (x) and parameters ($\boldsymbol{\beta}, \sigma^2$). I developed models representing all possible subsets of the eight available covariates representing 256 possible models. I assumed the objective function in (3.7) with values of $\boldsymbol{w} \equiv 2$, and $\gamma = 0$ (i.e., AIC). That is, for multiple linear regression with normal, independent, and identically distributed errors:

$$(3.8) \quad f(\boldsymbol{\beta}_m) = 2\left(\frac{n}{2}\log(2\pi) + \frac{n}{2}\log(\hat{\sigma}^2_{mle,m}) + \frac{1}{2\hat{\sigma}^2_{mle,m}}\sum_{i=1}^{n}(y_i - \boldsymbol{x}'\boldsymbol{\beta}_m)^2\right) + 2\sum_{j=1}^{p_m}|\beta_{j,m}|^0,$$

where $\boldsymbol{\beta}_m$ are the subset of parameters for model $m = 1, ..., 256$, n is the sample size, and $\hat{\sigma}^2_{mle} = \frac{\sum_{i=1}^{n}(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}_m)^2}{n}$. The model that minimized (3.8) (i.e., the AIC top model) included the intercept and 4 of 8 covariates. Three of the four covariates that were included in the top model were used to generate the data (associated with $\beta = -10, -5$, and 5) and one covariate that was independent of the data. The nine competing models are shown in Table 3.2.

Models fit using multi-model averaging are special cases of (3.7) or (3.8), where instead of minimizing over $\boldsymbol{\theta}$, $\boldsymbol{\theta}$ are selected using $\hat{\boldsymbol{\theta}} = \sum_{m=1}^{M} \omega_m \hat{\boldsymbol{\theta}}_{mle,m}$, where $\omega_m = \frac{\exp(-1/2\Delta_m)}{\sum_{m=1}^{M}\exp(-1/2\Delta_m)}$ are model weights, $\hat{\boldsymbol{\theta}}_{mle,m}$ is a vector containing each of the maximum likelihood parameter estimates from each of the different M candidate models, and Δ_m is the difference between an information criterion (e.g., AIC, BIC) of the optimal model and the mth model (Burnham and Anderson 2002). Therefore, models fit using model averaging are not an optimal solution to $\min_{\boldsymbol{\theta}}(f(\boldsymbol{\theta}))$ unless

$\hat{\boldsymbol{\theta}} = \min_m(\hat{\boldsymbol{\theta}}_{mle,m})$, which will occur if $\omega_m = 1$ for any model m. I calculated the model-averaged model for comparison among the other candidate models with respect to (3.5) and (3.6) (Table 3.2).

Finally, a common practice for pre-specification of weights in MOO problems in other applications includes examining the sensitivity of the optimal choice relative to the selected weights (Barron and Schmidt 1988, Rios Insua 1990). An analogous procedure in the model-selection framework is to examine the optimal solutions relative to different information criteria because different criteria represent different objective weights (Table 3.1). The AIC, AIC$_c$, and BIC all resulted in the same top model suggesting the optimal solution for these data was robust to several different choices of weights.

3.5.3. MODEL SELECTION BY EXAMINING PARETO OPTIMAL SOLUTIONS. Using the simulated data described above, I examined model selection *via* specification of preferences post-optimization (Fig. 3.1). That is, I identified the Pareto optimal solutions among the 256 models, and then considered potential methods for selecting a model. To identify the Pareto optimal solutions, I used a graphical approach and plotted the values f_1 and f_2 described in (3.8) for each model on opposing axes to identify the Pareto Frontier (Fig. 3.1). Note that identifying the Pareto optimal solutions did not require specifying w_1 or w_2, and therefore did not require adhering to an information criterion. The Pareto optimal set included 9 models; one model for each number of parameters $2, ..., 10$. Each Pareto solution represented the model that minimized (3.8) among all models with the same number of parameters. There were 247 dominated models (i.e., models that were not Pareto optimal; Fig. 3.1). The AIC top model was a Pareto optimal solution; this was expected because AIC (and other information criteria) is a specific formulation of the weighted-sum method and is therefore *sufficient* for Pareto optimality (Marler and Arora 2010, see Chapter 2 for definition of sufficient). Each of the Pareto solutions correspond to a specific set of weights in (3.7). The dominated solutions included six models that had ΔAIC < 2 (i.e., competing models). Finally, the model-averaged model was a dominated solution; this was expected because maximum likelihood estimates will result in smaller values of $-\log(L(\boldsymbol{\theta}|\boldsymbol{y}))$ than any other estimates, by definition.

Given the information on the Pareto optimal solutions in Fig. 3.1, selecting a final model for inference can proceed in many ways, depending on the application and the nature of the parameters under consideration. A decision maker can use the information on the Pareto optimal solutions to view trade-offs of fit gained by adding (or subtracting) additional parameters from the model, and choose a Pareto optimal solution with trade-offs acceptable to the decision maker. For example, many management decisions use model-based predictions for covariate data collected annually (e.g., Johnson et al. 1997). Some parameters might be associated with covariates in which annual data are difficult, expensive, or impossible to collect. The trade-offs in terms of model fit can be assessed relative to the expense of collecting additional data for these parameters to make future predictions. If the increase in model fit from the Pareto optimal solution that requires the additional (expensive) covariate data does not justify the additional expense, another Pareto optimal solution might be preferred. Another approach is to examine the curvature of the Pareto Frontier. Shapes such as elbows (e.g, Fig. 3.1: $p = 5$) can be identified, in which increasing the number of parameters has diminishing marginal returns in terms of f_1, and decreasing the number of parameters has a large effect on f_1. Another approach is to compare the trade-offs to ecological significance of the parameters involved and the need to make inference on those parameters. For example, if a parameter is required to inform a management decision, such as survival rates for harvest decisions, a decision maker would prefer to choose a Pareto optimal solution that included survival rates. Another approach might be to choose a Pareto optimal solution such that the maximum number of parameters is constrained by the amount of data. For example, if an investigator would like to constrain the number of parameters in the model such that $p < \frac{n}{30}$, the investigator could select the Pareto optimal solution that maximized model fit within the constrained set. Constraining the number of parameters could also be conducted during the model building phase. However, this might limit an investigator from observing the potential benefits of including additional parameters. If the benefit was large, it might provide sufficient motivation for the investigator to collect more data. When model averaging, the choice of models over which to average is not always apparent. The Pareto set could be used to select models for model averaging; the final model would be a composition of each of the Pareto optimal solutions. This approach would not guarantee the

final model is Pareto optimal in terms of the original objective functions, as was shown with model averaging over all models. Finally, models that are optimal in terms of model selection criteria could be highlighted as reference points on the Pareto frontier to guide decisions. Ultimately, the use of the Pareto frontier is that it provides visual information on the trade-offs of the objectives of the decision maker; in this case maximizing model fit and minimizing complexity.

3.6. DISCUSSION

The explicit application of multi-objective optimization to model selection using the objective functions defined in (3.5) and (3.6) ties several important properties of MOO to common methods for model selection. First, many different model selection methods are special cases of the weighted-sum method; each method representing different objective weights. This provides a unifying framework to quantitatively and visually compare model-selection methods based on different theoretical foundations. Practitioners of multi-objective optimization in operations research or other decision theoretic fields usually recommend sensitivity analyses of the resulting decisions given the choice of objective weights (Keeney and Raiffa 1976, Gregory et al. 2012). A sensitivity analysis for the model selection problem consists of evaluating multiple model selection criteria (representing different objective weights) to examine the robustness of the solution to the choice of criterion. Many practitioners argue against this approach, suggesting that a criterion should be selected based on its theoretical motivation (e.g., AIC is asymptotically efficient; BIC is consistent Aho et al. 2014). Others view a specific information criterion as one line of evidence to assist in a decision and report different criteria side-by-side (e.g., Araújo and Luoto 2007, Parviainen et al. 2008). The former appears to be the dominant paradigm in ecological research, whereas the latter is common in other fields. Second, many model selection methods result in Pareto optimal solutions because they are specific formulations of (3.2), which is sufficient for Pareto optimality. Thus, there is a decision-theoretic basis for model selection methods that can be expressed in the form of (3.7) in terms of optimality criteria (i.e., Pareto optimality).

Two other points require further qualification. First, competing models (i.e., ΔAIC< 2) are not necessarily Pareto optimal in terms of (3.5) and (3.6). Second, models obtained using multi-model averaging are not necessarily Pareto optimal in terms of (3.5) and (3.6). Pareto optimality is determined with respect to the objective functions. Competing models and multi-model averaging extend beyond the objectives described in (3.5) and (3.6). They concern the concept of model selection uncertainty due to the random nature of data generated from a process (Burnham and Anderson 2002). Minimizing (or accounting for) model selection uncertainty is not fully captured by (3.5) and (3.6). Model selection uncertainty represents an additional objective. To appropriately assess methods that account for model uncertainty in the multi-objective optimization framework, an additional objective function that quantifies model robustness to random variation of the data is needed. For example, the mean-squared prediction error of each model could be used as an objective for Monte Carlo simulated data sets. Burnham and Anderson (2002) found that model averaged estimates performed better than maximum likelihood estimates of top models when comparing mean-squared predicted error. Although multi-model averaging is not optimal with respect to the traditional objectives of fit and complexity, it might be optimal when considering the additional objective of model selection uncertainty or model robustness. Appropriately accounting for a decision maker's objectives is important for multi-objective optimization to be useful.

Model selection by examining trade-offs of fit and complexity is not new. Users of Mallow's C_p often conduct similar investigations (Mallows 1973). Freitas (2004) examined Pareto optimality in the related question comparing prediction and simplicity for data mining. However, I am unaware of any wildlife or ecological studies that have formally examined model selection in terms of Pareto optimality of fit and complexity and made a choice based on the resulting trade-offs observed. Viewing each model's trade-offs, in terms of objectives, provides a visual assessment of the model selection problem, a potentially useful tool for ultimately choosing a model to base inference. As is the case with any multi-objective optimization problem, the additional flexibility in model choice based on post-optimization specification of preferences could be viewed as either a positive or negative trait, depending on how an investigator values the order in which preferences are specified. Specifying preferences pre-optimization for the model selection problem benefits

from being objective in the sense that a decision maker chooses how to weigh their specific objective functions without being influenced by how weights will alter the outcome of optimization. Specifying preferences post-optimization has the added flexibility of choosing a Pareto optimal solution that provides the best trade-offs for context dependent decision problems.

Although I described the model selection problem heuristically in terms of maximizing model fit and minimizing model complexity, I could have replaced model fit with predictive ability as one objective of interest, and asymptotic bias correction as another. Predictive ability is the most commonly sought model characteristic for model selection, and many information criteria and other model selection methods were developed to optimize predictive ability (Akaike 1973, Stone 1977, Gelfand and Ghosh 1998, Hoeting et al. 1999, Burnham and Anderson 2002, Hooten and Hobbs 2015). Many information criteria methods have weights and penalties that serve as bias corrections for optimization in terms of predictive ability (Konishi and Kitagawa 1996). That is, many information criteria are based on bias-corrected log likelihoods, in which the model complexity is a correction factor to remove asymptotic bias of the log likelihood of an estimated model (Konishi and Kitagawa 1996). The MOO problem in terms of maximizing predictive ability and accounting for model bias is similar in spirit to the MOO problem of maximizing model fit while minimizing model complexity.

Table 3.1: Values of weights (w_i) and γ for the multi-objective optimization problem of model selection described in (3.7) for various model selection methods. The objective function for model fit is $-\log(L(\boldsymbol{\theta}|\boldsymbol{y}))$; the objective function for model complexity is $\sum_{j=1}^{p} |\beta_j - \mu_j|^\gamma, j = 1, ..., p$. AIC = Akaike's information criterion; AIC_c = Second-order information criterion; QAIC = quasi-AIC; BIC = Schwartz information criterion; n = sample size; p = no. parameters in model. (*) indicates objective function for model fit defined by: $\sum_{i=1}^{n} (y_i - \beta_0 - \boldsymbol{x}'\boldsymbol{\beta})^2$. See Burnham and Anderson (2002) and Hooten and Hobbs (2015) for additional details.

Model selection method	w_1	w_2	γ	Note
AIC	2	2	0	
AIC_c	2	$2(\frac{n}{n-p-1})$	0	
QAIC	$\frac{2}{\hat{c}}$	2	0	$\hat{c} = \chi^2/df$
$QAIC_c$	$\frac{2}{\hat{c}}$	$2(\frac{n}{n-p-1})$	0	$\hat{c} = \chi^2/df$
BIC	2	$\log(n)$	0	
*Ridge regression	1	User defined or estimated	2	Larger values of w_2 shrink $\boldsymbol{\beta}$ to 0.
*LASSO	1	User defined or estimated	1	Larger values of w_2 shrink $\boldsymbol{\beta}$ to 0.

Table 3.2: Top nine models with smallest values of $f(\boldsymbol{\beta})$ from example problem considering variable selection of a multiple linear regression model used to model simulated data. AIC = Akaike's information criterion, ΔAIC = difference in AIC between model m and model with smallest AIC. Asterisk (*) indicate the parameter was used to simulate the response variable (i.e., $\beta_0 - \beta_4$). $f_1(\boldsymbol{\beta})$ and $f_2(\boldsymbol{\beta})$ are described in Eq. (3.8) and Eq. (3.6), respectively. The last model in the table, weighted average of all $\beta_0 - \beta_8$, was a model in which $f(\boldsymbol{\beta})$ was calculated using parameters ($\boldsymbol{\beta}$) that were fit using model averaging of each of the 256 candidate models instead of maximum likelihood estimates.

Model variables	$f(\boldsymbol{\beta})$ (i.e., AIC)	ΔAIC	$f_1(\boldsymbol{\beta})$	$f_2(\boldsymbol{\beta})$	Pareto optimal
$\beta_0^* + \beta_1^* + \beta_2^* + \beta_4^* + \beta_7$	1919.97	0	953.99	6	Yes
$\beta_0^* + \beta_1^* + \beta_2^* + \beta_4^*$	1920.55	0.57	955.27	5	Yes
$\beta_0^* + \beta_1^* + \beta_2^* + \beta_3^* + \beta_4^* + \beta_7$	1920.94	0.97	953.47	7	Yes
$\beta_0^* + \beta_1^* + \beta_2^* + \beta_4^* + \beta_6 + \beta_7$	1920.97	1.00	953.49	7	No
$\beta_0^* + \beta_1^* + \beta_2^* + \beta_4^* + \beta_6$	1921.29	1.31	954.64	6	No
$\beta_0^* + \beta_1^* + \beta_2^* + \beta_4^* + \beta_5 + \beta_7$	1921.62	1.65	953.81	7	No
$\beta_0^* + \beta_1^* + \beta_2^* + \beta_4^* + \beta_7 + \beta_8$	1921.72	1.75	953.86	7	No
$\beta_0^* + \beta_1^* + \beta_2^* + \beta_3^* + \beta_4^*$	1921.84	1.86	954.92	6	No
$\beta_0^* + \beta_1^* + \beta_2^* + \beta_4^* + \beta_5$	1921.88	1.90	954.94	6	No
Weighted average of all $\beta_0 - \beta_8$	2223.35	303.37	2203.35	10	No
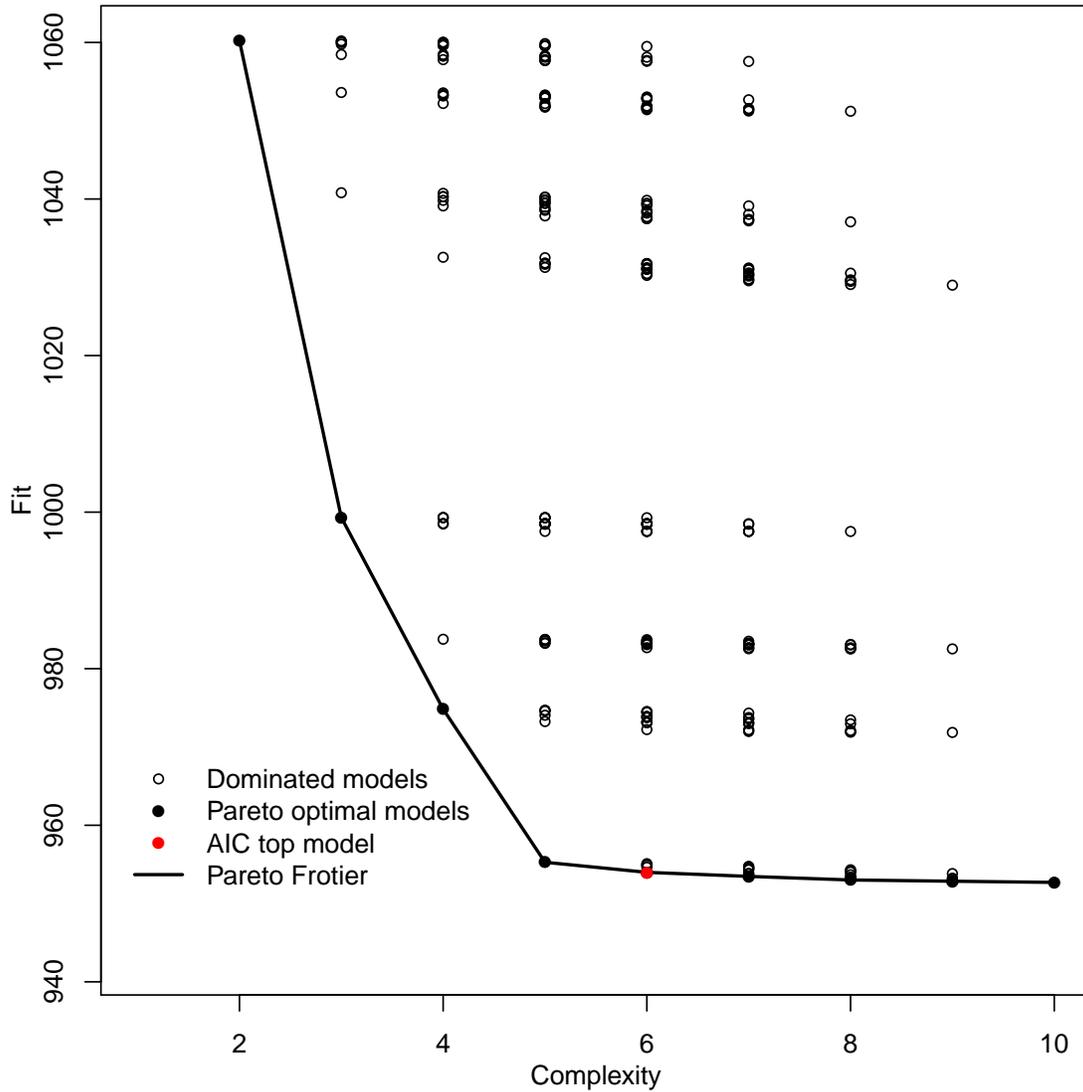
FIGURE 3.1. Plot of model fit ($f_1(\boldsymbol{\theta}) = -\log(L(\boldsymbol{\theta}|\boldsymbol{y})$; see Eq. (3.8)) versus model complexity ($f_2(\boldsymbol{\theta}) =$no. parameters; see Eq. (3.6)) for each model described in the multiple linear regression variable selection example. Model parameters include the estimated $\beta_i, i = 0, ..., 8$ values, and estimated σ^2. Optimal solutions minimize fit (moving towards bottom of figure) and complexity (moving to the left of figure). The top model using $f(\boldsymbol{\theta}) = $ AIC was a Pareto optimal solution. Six candidate solutions (i.e., ΔAIC < 2) were not Pareto optimal.

DERIVING OPTIMAL MANAGEMENT STRATEGIES FOR CACKLING GEESE FROM BAYESIAN INTEGRATED POPULATION MODELS

4.1. SUMMARY

Markov decision processes (MDPs) provide a framework for optimizing management policies for recurrent decisions. Necessary to MDPs are specifying transition probabilities describing the probability of moving from the current ecological state to future states, given the choice of management action. The Bayesian framework provides a natural method for calculating transition probabilities based on empirical data using posterior predictive distributions. I developed a Bayesian hierarchical integrated population model to estimate population dynamics of cackling geese. I estimated the effect of daily bag limits on cackling goose survival during periods of recreational hunting, and the effect of subsistence harvest on spring survival estimates. I then used management-specific survival estimates to derive posterior predictive distributions describing the probability of transitioning among population sizes, based on the management action selected. I incorporated the transition probabilities into an MDP and identified two optimal policies for selecting state-specific daily bag limits for cackling geese using stochastic dynamic programming. The first optimal policy corresponded to reward functions associated with the current management strategy for cackling geese (i.e., a population objective of 250,000 geese, and hunting closure when abundance estimates are < 80,000 geese). The second optimal policy corresponded to a reward function I developed using multi-objective optimization of three competing objectives; maximizing abundance to improve subsistence harvest, minimizing abundance to reduce agricultural depredation, and balancing stakeholder satisfaction (Chapter 2). Survival rates between periods of closed recreational hunting and periods of restricted daily bag limits (i.e., one or two birds per day) were similar among age classes. Liberal daily bag limits (three or four birds per day) were correlated with decreased survival rates among all age classes, with juvenile birds appearing to be affected

the most. Closing subsistence hunting appeared to have a positive effect on juvenile survival, but a negative effect on sub-adult and adult survival indicating that closing subsistence hunting might not be an effective management policy for increasing cackling goose abundance. The two optimal management policies were: 1) no hunting between 0–80,000 birds, restrictive daily bag limits between 80,000–235,000 birds, and liberal hunting for >235,000 birds for the current management framework, and 2) no hunting between 0–80,000 birds, restrictive daily bag limits between 80,000–223,000 birds, and liberal daily bag limits for >223,000 birds for the alternative management framework.

Key words Bayesian inference, *Branta hutchinsii minima*, cackling geese, decision theory, harvest management, hierarchical integrated population model, Markov decision process, posterior predictive distribution, stochastic dynamic programming.

4.2. INTRODUCTION

Ecological systems are dynamic, changing stochastically through time. Methods for modeling stochastic dynamic ecological processes using empirical data have led to the proliferation of advanced statistical methods. Formal methods for linking these models to natural resource management are critical for using scientific investigation to evaluate and select among alternative management strategies (Williams 2015). Markov decision processes provide a framework for linking decision making to models of ecosystem processes based on empirical data (Anderson 1975, Kendall 2001, Puterman 2014).

Markov decision processes (MDPs) are used to represent stochastic dynamic processes that can be partially controlled through the actions of a decision maker through time (Puterman 2014). MDPs have been used for many ecological applications including: modeling animal foraging behavior (Mangel and Clark 1986), describing mate desertion in birds (Kelly and Kennedy 1993), management of mallard harvest (*Anas platyrhynchos*; Anderson 1975, Williams 1996, Johnson et al. 1997), and endangered species management (Johnson et al. 2011). Models of ecosystem processes are incorporated into MDPs by describing the probability of transitioning from one system

state to any other system state, given the choice of action. Models for describing transition probabilities range from expert elicitation (e.g., Lau and Leong 1999), to fitting models using empirical evidence (e.g., Johnson et al. 1997; 2011).

Bayesian methods are becoming increasingly popular for modeling ecological processes (Hooten and Hobbs 2015). Bayesian methods provide a natural method for modeling transition probabilities in Markov decision processes because the Bayesian framework treats all unobserved quantities as random variables. A future state is an unobserved random quantity from the system that is being modeled and therefore, we can use Bayesian models of system dynamics to estimate the probability of future states, conditional on observed data (Gelman et al. 2014, Hobbs and Hooten 2015). Specifically, the probability of a future state, given data y, can be estimated using posterior predictive distributions:

$$
(4.1) \qquad [\tilde{y}|y] = \int_\theta [\tilde{y}|\theta][\theta|y]d\theta,
$$

where $[\tilde{y}|y]$ is the probability of unobserved data (or future state) given the observed data y, $[\tilde{y}|\theta]$ is the likelihood of $\tilde{y}$ given θ, and $[\theta|y]$ is the posterior distribution of θ, given y. Consider a specific application of (4.1) for the problem of predicting future population size n_{t+1}. Assume n_{t+1} is described by a function of the current population size n_t, random parameters describing the population dynamics ($\boldsymbol{\theta}$), and potential management actions a:

$$
(4.2) \qquad n_{t+1} = f(n_t, \boldsymbol{\theta}, a).
$$

For example, the function f might be a function describing the life cycle of an animal, in which case the vector $\boldsymbol{\theta}$ might consist of parameters representing the vital rates of the population, and a is the action to be implemented that, potentially, affects the vital rates. Given data $\boldsymbol{y}$, we can estimate the posterior predictive distribution for n_{t+1}, resulting in the probability of moving from the current state n_t to future states n_{t+1} for a specific management action a. That is, the predictive distribution of n_{t+1} can be written as

$$
(4.3) \qquad [n_{t+1}|y,a] = \int_\theta [n_{t+1}|\boldsymbol{\theta},a][\boldsymbol{\theta}|y,a]d\theta.
$$

Equation (4.3) can be discretized to finite states and incorporated into a MDP to describe transition probabilities among states. Combining transition probabilities with actions and reward functions, we can find an optimal solution to the MDP using solution algorithms such as stochastic dynamic programming (Puterman 2014). The optimal solution incorporates the objectives of the decision makers and statistical inference based on empirical data.

To demonstrate the application of posterior predictive distributions for modeling transition probabilities in an MDP, I consider the management problem of selecting annual daily bag limits for a harvested species. The cackling goose (*Branta hutchinsii minima*) is a migratory bird species that nests on the coastal plain of the Yukon-Kuskokwim (YK) Delta, Alaska. They winter primarily in the Willamette Valley and Lower Columbia River Valley in Oregon and Washington (Sedinger and Bollinger 1987, Pacific Flyway Council 1999). Cackling geese are hunted recreationally across their range. Recreational hunting begins in mid-September in Alaska, and often continues through a late goose season in mid-March in Oregon and Washington. Subsistence hunting is prominent to the economy and social welfare of many rural Alaskan communities (Wolfe and Walker 1987). Migratory bird subsistence hunting in Alaska has historically occurred from when birds arrive in the spring (Apr–May) through the end of August (Naves 2010; 2011; 2012). Most subsistence harvest of cackling geese occurs in the spring (Naves 2010; 2011; 2012). Cackling geese are regularly the second-most harvested avian species during subsistence hunts (behind greater white-fronted geese (*Anser albifrons*); Naves 2010; 2011; 2012).

Autumn counts of cackling geese conducted in northern California declined from $> 350,000$ birds in the mid-1960s to $< 30,000$ birds in 1983 (O'Neill 1979, Raveling 1984, King and Derksen 1986). In response to this estimated population decline, the Pacific Flyway Council, in collaboration with the U.S. Fish and Wildlife Service, state wildlife agencies, and Native Alaskans agreed to close recreational and subsistence harvest of cackling geese during the 1984–1985 hunting season until abundance estimates increased to over 110,000 birds (Pamplin Jr 1986, Pacific Flyway Council 1999). Abundance estimates increased rapidly after the 1984 hunting closure to approximately 220,000 birds in 1994 (Fig. 1.2; Pacific Flyway Council 1999). After the population recovery, subsistence and recreational hunting were re-instituted with a 1-bird daily bag limit in 1994–1995 for

recreational hunting (Table 4.1). Subsistence hunting had no bag limit, but had a 30-day cessation period after nest initiation. The Willamette Valley falls mainly in Oregon's Northwest Permit Zone (http://www.dfw.state.or.us/). Daily bag limits in the Northwest Permit Zone have ranged from 1–4 birds per day since 1994, with season length ranging from 23 to 100 days (mean = 91 days; Table 4.1).

The Pacific Flyway Council selects cackling goose harvest regulations annually based on abundance estimates from data collected when geese are nesting in Alaska (see Chapter 1 for description of aerial survey methods and estimation procedures; Pacific Flyway Council 1999). The current population objective is 250,000 geese; harvest regulations are selected to achieve this objective (Pacific Flyway Council 1999). The Pacific Flyway Council is currently revising its cackling goose management plan. A consideration for any management plan of harvested waterfowl species is the annual choice of daily bag limits. I address choosing daily bag limits by incorporating statistical inference from data collected on the population dynamics of cackling geese, into a decision theoretic framework. My objectives were to 1) reparameterize the model developed in Chapter 1 to represent the population dynamics of cackling geese relative to different harvest regulations, 2) use the population model to estimate posterior predictive distributions for transition probabilities among population sizes, and 3) incorporate the transition probabilities into an MDP framework to calculate optimal state-specific harvest regulations for cackling goose management.

4.3. METHODS

4.3.1. MARKOV DECISION PROCESS. The general MDP consists of 5 components including: 1) the potential states of a system of interest, 2) potential actions a decision maker can choose, 3) a function, or functions, describing the rewards of an action, given the current state of the system, 4) transition probabilities, and 5) a discount factor. A set of system states, denoted $\mathscr{S}$, is a finite set that contains all the potential states, s, of a system that could occur. For example, if a decision maker is interested in abundance, the set of states includes values between 0 and the carrying capacity. MDPs assume the state s_t at time t is observed without error, but moving from state s_t to future state s_{t+1} is random and described by a probability distribution, $[s_{t+1}|s_t, a]$,

that depends at least partly on the action, a, that is implemented. The distribution $[s_{t+1}|s_t, a]$ is a model (either conceptual or statistical) that describes the probability of future state s_{t+1}, given the current state s_t, and action a. A reward function $R(s_t, a)$, describes the reward for choosing action a in state s_t. Reward functions quantify the aim or interests of decision makers and are therefore subjective. Finally, a discount factor δ is a value between zero and one that describes the relative value of rewards in future time periods. It is commonly used in economic applications to model depreciation of money in future time periods. However, it is often set to one (or close to one) in ecological applications, indicating the value of a resource in the future is the same as the value of the resource now (e.g., Milner-Gulland 1997, Shea and Possingham 2000, Westphal et al. 2003).

For the cackling goose management problem, I assumed potential states included in the set $\{0, 1000, 2000, ..., 400000\}$ birds. I assumed a discount factor close to one (i.e., 0.99999) for computational reasons (i.e., to achieve a solution for an infinite time horizon problem). I describe the remaining components of the MDP for the cackling goose management problem in the following sections.

4.3.1.1. *Management actions.* The set of management actions I considered were the daily bag limits used for cackling geese in the Willamette Valley (Northwest Permit Zone) during the 1982–1983 to 2013–2014 hunting seasons (Table 4.1). Counties in the Northwest Permit Zone currently include Benton, Clackamas, Clatstop, Columbia, Lane, Lincoln, Linn, Marion, Multnomah, Polk, Tillamook, Washington, and Yamhill county (http://www.dfw.state.or.us/). These counties comprised 79% (559 of 712) of reported harvested cackling geese in the Pacific Flyway banded from 1982–2014 (https://www.pwrc.usgs.gov/bbl). The harvest regulations were restrictive hunting (one or two bird daily bag limit; 1982–1983 to 1983–1984, 1994–1985 to 1997–1998, 2005–2006 to 2009–2010), no hunting, including no subsistence hunting (zero bird daily bag limit; 1984–1985 to 1993–1994), and liberal hunting (three or four daily bag limit; 1998–1999 to 2004–2005, 2012–2013 to 2013–2014). I grouped daily bag limits into no hunting, restrictive hunting, and liberal hunting so that each management action had at least eight years of survival information from the mark-resight data to estimate survival rates associated with the management action.

4.3.1.2. *Transition probabilities.* To estimate the probability of moving from state n_t to any other state n_{t+1}, given an action (i.e., $[n_{t+1}|n_t, a]$), I derived posterior predictive distributions from a Bayesian hierarchical integrated population model fit with 26 years of cackling geese mark-resight data, and 32 years of relative abundance data. I used the data described in *abundance data* and *mark-resight data* in Chapter 1.

Model development—I used a model similar to the model used in Chapter 1, with different parameterization to account for the different potential management actions available. As in Chapter 1, I assumed aerial abundance data from $t = 1, ..., 33$ (corresponding to 1982–2014) was normally distributed, with mean $\mu = n_t$, and variance σ^2. The latent variable n_t represented pre-breeding abundance and was defined by the process:

$$n_t = n_{sub,t} + n_{ad,t},$$

$$n_{sub,t} = \frac{n_{ad,t-1}}{2} f_{t-1} \phi_{juv,t-1},$$

(4.4)
$$n_{ad,t} = n_{sub,t-1} \phi_{sub,t-1} + n_{ad,t-1} \phi_{ad,t-1},$$

$$f_{t-1} = \frac{n_{sub,t-1} \rho_{sub,t-1} f_{sub,t-1}}{n_{ad,t-1}} + \rho_{ad,t-1} f_{ad,t-1},$$

where $n_{sub,t}$ were the number of sub-adult birds in year $t = 1, ..., 33$, $n_{ad,t}$ were the number of adult birds, f_{t-1} is a composite reproductive statistic consisting of breeding propensities for each age, $\rho_{sub,t}$ and $\rho_{ad,t}$, and fecundity estimates for each age $f_{sub,t}$ and $f_{ad,t}$ (i.e., defined as the mean number of birds per pair that survived until summer banding). Thus, $f_{t-1} \times \frac{n_{ad,t-1}}{2}$ is interpretable as the number of juveniles produced in year $t - 1$ by sub-adults and adults that survived to summer banding. The parameters $\phi_{juv,t}$, $\phi_{sub,t}$, and $\phi_{ad,t}$ were annual survival probabilities. I assumed sub-adults were 1-year-old birds, and adults were $\geq$ 2-year-old birds. I set $n_{ad,1} = y_1 - n_{sub,1} = y_1 - 0.2y_1$ for year 1. $n_{ad,1} = 0.2$ was used for consistency with the model fit in Chapter 1. I assumed a uniform prior for fecundity, $f_t \sim \text{Unif}(0, 15)$. I assumed an inverse-gamma prior for σ^2, $\sigma^2 \sim \text{IG}(r, q)$, where r and q were selected such that the mean and variance of the inverse-gamma distribution both equaled $5,000^2$, respectively (using moment matching; Hobbs and Hooten 2015).

I chose $5,000^2$ as the mean for σ^2 because the maximum annual standard error for y_t calculated by the Pacific Flyway Council was 5,202 (mean = 2,967; T. Sanders, personal communication). Thus, the variance of our prior for σ^2 was conservative. I assumed the demographic stochasticity of the process was accounted for in the uncertainty in each parameter and therefore did not assume an additional probability model for the process in (4.4).

To model the mark-resight data $m_{i,j}$, for $i = 1, ..., 10,428$ individuals, $t = 1, ..., 33$, years, $s = 1, ..., 6$ seasons, and $j = 1, ..., 192$ resight occasions ($j = s + (t - 1) \times S$), I used the state-space formulation of the Cormack-Jolly-Seber model. That is,

$$
(4.5) \qquad m_{i,j} \sim \begin{cases} 0 & , z_{i,j} = 0 \\ \text{Bern}(p_j) & , z_{i,j} = 1 \end{cases},
$$

where $z_{i,j}$ was the latent, true state (alive = 1, dead = 0) of cackling goose i during resight occasion j. The parameter p_j was the resight probability of a marked goose during resight event j. I estimated p_j using a generalized linear model with a logit link: $\text{logit}(\boldsymbol{p}) = \boldsymbol{W}\boldsymbol{\alpha}$, where $\boldsymbol{W}$ is a design matrix for the detection probability parameters $\boldsymbol{\alpha}$. The design matrix $\boldsymbol{W}$ included indicator variables for season and year of resight events. I assumed vague priors for $\boldsymbol{\alpha}$ defined by $\text{N}(\boldsymbol{0}, 1.5^2\boldsymbol{I})$, where $\boldsymbol{I}$ is the identity matrix. I modeled the true latent state of each individual using

$$
(4.6) \qquad z_{i,j} \sim \begin{cases} 0 & , z_{i,j-1} = 0 \\ \text{Bern}(\phi_{i,j}) & , z_{i,j-1} = 1 \end{cases},
$$

where the first case ensures that if an individual was dead during the previous resight event it remained dead during the present resight event. The second case assumes that if an individual was alive during the previous resight event, it was alive with probability equal to the season- and time-specific apparent survival probability $\phi_{i,j}$. In Chapter 1, I modeled apparent survival for each season, age, sex, and year to develop an IPM framework that permitted inference on survival at two temporal scales. In this chapter, I consider an alternative parameterization to assess the effect of different management actions on survival. That is, I allowed survival to vary for the different time frames in which different harvest regulations were instituted. I also modeled age specific survival,

and an interaction with harvest regulations and age. The relationships between survival and the covariates were

(4.7) $\text{logit}(\phi_{i,j}) = \beta_0 + \boldsymbol{\beta}_1'(\mathbf{season}_j) + \boldsymbol{\beta}_2'(\mathbf{age}_{i,j}) + \boldsymbol{\beta}_3'(\mathbf{mgmt}_j) + \boldsymbol{\beta}_4'(\mathbf{age}_{i,j} \times \mathbf{mgmt}_j),$

for $i = 1, ..., 10,428$ individuals and the season-year combination $j = s + (t-1) \times S$, corresponding to the different management actions in seasons $1, ..., S = 6$, and years $1, ..., 32$ (i.e., 1982,...,2013). The six seasons included late summer (Aug–Sep) on the YK Delta, fall (Oct–Nov) in Oregon, Washington, and California, winter (Dec–Feb) in Oregon, Washington, and California, spring (Mar–Apr) in Oregon, Washington, and California, early summer (May–Jun) on the YK Delta, and during subsequent capture occasions on the YK Delta (late-Jul–Aug). Age included juvenile, sub-adult, and adult birds. The management covariate (mgmt) identified years in which no hunting, restrictive hunting, or liberal hunting regulations were implemented. The management covariate was only applied to seasons in which the action was most likely to occur. The seasons included Oct–Nov, Dec–Feb, and Mar–Apr for recreational hunting, and May–Jun for subsistence hunting (Naves 2010; 2011; 2012). I assumed vague priors for $\boldsymbol{\beta}$ defined by $\text{N}(\mathbf{0}, 1.5^2 \boldsymbol{I})$. The full model statement is given in Appendix C.

I calculated posterior distributions using a Markov Chain Monte-Carlo algorithm I wrote using a combination of R version 3.0.2 and C++ (R Core Team 2013, supplementary material). The full algorithm is given in Appendix E. I calculated marginal posterior distributions for the parameters $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, $\boldsymbol{f}$ and σ^2 and the derived parameters $\phi_{s,a,mgmt}$, $p_{s,t}$ and n_{t+1,a,n_t} (n_{t+1,a,n_t} is described in the next section). I obtained 3 parallel chains of 200,000 iterations using the algorithm. I removed the first 50,000 burn-in iterations and I examined convergence using trace plots and Gelman-Rubin diagnostics (Gelman et al. 2014).

Posterior predictive distribution for $\boldsymbol{n_{t+1}}$—To estimate the probability of transitioning from one state (i.e., population size n_t) to another state (n_{t+1}), given the choice of management action, I derived the parameter n_{t+1} from the estimated posterior distributions described above. I calculated

future population size, for any n_t, and any management action using

$$(4.8) \qquad n_{t+1,n_t,a} = n_t \gamma \prod_{s=1}^{S} \phi_{ad,s,a} + n_t(1-\gamma) \prod_{s=1}^{S} \phi_{sub,s,a} + \frac{n_t \gamma}{2} f \prod_{s=2}^{S} \phi_{juv,s,a},$$

where $\gamma = \frac{n_{ad}}{n_{ad}+n_{sub}}$, was the proportion of adults in the population. Using each value of $\gamma^{(k_i)}$, $\phi_{ad,s,a}^{(k_i)}$, $\phi_{sub,s,a}^{(k_i)}$, $f^{(k_i)}$, and $\phi_{juv,s,a}^{(k_i)}$ from the $k_i = 50,001,...,200000$ iterations of the $i = 1,2,3$ MCMC chains, I calculated the values of n_{t+1,a,n_t}, for all $n_t \in \{1000, 2000, ..., 400,000\}$ birds, and all $a \in \mathscr{A}$. These calculations resulted in a distribution for each combination of action and population size. I discretized the continuous distribution into a set of 400 finite states by rounding values of n_{t+1} to the nearest thousand birds. I chose 400 states because I assumed 400,000 birds was the carrying capacity of cackling geese (Chapter 1). I standardized each of these distributions to sum to one. Thus, for any value of n_t, and any action a, I derived a probability mass function describing the probability of moving from n_t to n_{t+1} (i.e., $[n_{t+1}|n_t, a]$; Fig. 4.1). Because γ and f were calculated for all years $t = 1,...,32$ in each iteration of the MCMC chain, I used the mean among all years of each iteration to use in (4.8).

4.3.1.3. *Reward functions.* When a decision maker chooses an action a given a state n, at time t, the decision maker receives a reward $R(a,n)$. I considered two scenarios for developing reward functions for cackling goose management. The first scenario represented the current regulatory framework for cackling geese reported in Pacific Flyway Council (1999). That is, a population objective of 250,000 birds, and a management trigger point of 80,000 birds for closing recreational and subsistence harvest. The second, alternative scenario, considers the objective function that I developed using the weighted-sum method to identify a population objective for cackling geese described in Chapter 2 (Eq. (2.5), Eq. (2.6), and Fig. 2.5**D**).

Reward functions for the current regulatory framework—To develop reward functions for each potential management action for the current regulatory framework, I first developed several assumptions. I made the following four assumptions for liberal and restrictive harvest. First, the maximum reward occurred when subsequent abundance equaled 250,000 birds (Pacific Flyway Council 1999). Second, the reward function was symmetric on both sides of the population objective. A symmetric reward function implies that values of n greater than the population objective

have the same decrease in reward as values below the population objective (Williams 2015). Cackling goose management balances the competing objectives of providing a population size adequate for subsistence harvest, and minimizing crop depredation in their wintering area (Chapter 2). A symmetric reward function represents these competing objectives. Third, population sizes close to the management goal had relatively smaller losses in reward per bird compared to population sizes far from the management goal. Thus, the reward functions were concave (Williams 2015). Fourth, if the population was below 80,000 birds, implementing restrictive or liberal hunting would have no reward. A standard reward function meeting these axioms is the negative squared error loss function ($R^*(n) = -(n - 250,000)^2$). I scaled $R^*(n)$ to be between 0 and 1 using

$$(4.9) \qquad R(n) = \begin{cases} \frac{R^*(n) - R^*(80,000)}{\max(R^*(n)) - R^*(80,000)} & 80,000 < n < 420,000 \\ 0 \text{ otherwise} \end{cases}$$

(Fig. 4.2**A**). The reward function in (4.9) is not currently a function of any action a. I assumed the reward accrued by a decision maker depends not only on state n_t, but the future state n_{t+1}, which depend on the action implemented. Thus, I calculated the reward function conditioned on actions by calculating the expected reward, with expectation taken over the conditional transition probability $[n_{t+1}|n_t, a]$:

$$(4.10) \qquad R(n, a) = \sum_{n_{t+1} \in \mathcal{N}} R(n)[n_{t+1}|n_t, a],$$

(Fig. 4.2**B**). For no hunting, I assumed a reward function that equaled zero for values of $n_t > 80,000$ birds. For values of $n_t < 80,000$ birds, I assumed the reward function equaled $R(80,000, a)$, for $a = $ restrictive hunting (Fig. 4.2**B**). The no-hunting reward function implies that no hunting is preferred to other actions when the population is below 80,000 birds, but when the population reaches 80,000 birds, no hunting has zero reward.

Reward functions for an alternative regulatory framework—In addition to the negative squared-error loss function described for $R(n)$ in (4.9), I considered an additional reward function that I described in Chapter 2 (Eq. (5), Eq. (6), and Fig. 2.5D), for comparison. Briefly, this

reward function considered three objectives for cackling goose management: 1) minimizing cackling goose depredation on agriculture fields, 2) maximizing subsistence harvest, and 3) balancing competing stakeholder objectives. These functions were described in (2.5) as:

$$f_s(n) = \begin{cases} 0 \, , n \leq \chi \\ \log(b + cn) \, , n > \chi \end{cases},$$

$$f_a(n) = 1/2 f_{a,1}(n) + 1/2 f_{a,2}(n),$$

$$f_{a,1}(n) = 1 - \frac{n}{\max(n)},$$

$$f_{a,2}(n) = \begin{cases} 1 \, , n \leq K_{pl} \\ 1 - \frac{n - K_{pl}}{\max(n)} \, , n > K_{pl} \end{cases},$$

$$f_b(n) = - \left(f_s(n) - f_a(n) \right)^2.$$

The function $f_s(n)$ is an objective function associated with subsistence hunting, in which $b = 1$, $c = 1 \times 10^{-6}$, and χ is a variable for the trigger point at which subsistence and recreational hunting are closed. Values of b and c were chosen to scale $f_s(n)$ to the support of cackling goose population size. The function $f_a(n)$ is an objective function associated with agriculture and was a weighted average of two hypotheses, $f_{a,1}(n)$ and $f_{a,2}(n)$. The first hypothesis, $f_{a,1}(n)$, assumes cackling geese fed exclusively on private fields. The second hypothesis, $f_{a,2}(n)$, assumes cackling geese first fed on public land until the public land carrying capacity, K_{pl}, is obtained. After K_{pl} is obtained, remaining birds feed on private land. The function $f_a(n)$ is a mixture of these two extreme hypotheses. The final function, $f_b(n)$ returns small values when $f_s(n)$ and $f_a(n)$ are different, and large values when $f_s(n)$ and $f_a(n)$ are similar, representing the objective of maintaining similar satisfaction levels between stakeholders. I assumed $\chi = 80,000$ birds to be consistent with current regulations. I also assumed $K_{pl} = 54,000$ birds, the estimated carrying capacity for cackling geese on public U.S. Fish and Wildlife Service land in the Willamette Valley (A. Mini, unpublished technical report). Finally, I assumed the carrying capacity of the population ($\max(n)$) was 400,000 birds, reflecting the largest abundance estimates calculated for cackling geese (Chapter 1). I scaled

the functions $f_s(n)$, $f_a(n)$, and $f_b(n)$ between 0 and 1 using the equation $\frac{f(n)-\min(f(n))}{\max(f(n))-\min(f(n))}$. I then took the weighted average of each function using $f(n) = 1/3f_s(n) + 1/3f_a(n) + 1/3f_b(n)$ (Fig. 4.2**C**). Similar to the methods used for developing action-specific reward functions for the current regulatory framework (described in the previous section), I calculated the expected reward for both restrictive and liberal daily bag limits using (4.10), and I used the same method for calculating the reward function for closing subsistence and recreational hunting (Fig. 4.2**D**).

4.3.2. OPTIMAL MANAGEMENT POLICY. I calculated optimal management policies for each of the management scenarios described above. I used stochastic dynamic programming to calculate expected cumulative reward for each management policy (Puterman 2014, Williams et al. 2002). That is, for each potential policy π, consisting of state-dependent management actions of either closing hunting, restrictive daily bag limits, or liberal daily bag limits, I calculated the expected total reward:

$$(4.11) \qquad \nu_\pi(N) = \lim_{T\to\infty} E\left\{ \sum_{t=1}^{T} \lambda^{t-1} R(n, a) \right\},$$

in which $0 \le \lambda < 0$ is the discounted reward which I set close to one, ensuring the limit in (4.11) existed (Puterman 2014). The optimal policy was the policy that maximized (4.11). I used policy iteration (Puterman 2014) to solve (4.11) for each potential policy using the MDPToolbox package in R version 3.0.2 (R Core Team 2013, Chades et al. 2014).

4.4. RESULTS

4.4.1. MANAGEMENT SPECIFIC SURVIVAL RATES. Apparent survival estimates during periods of no hunting (neither recreational hunting nor subsistence hunting) and restrictive hunting (open subsistence hunting and a one or two bird daily bag limit for recreational hunting) were similar for juveniles, sub-adults, and adults, with a three notable differences (Figs. 4.3, 4.4). First, juvenile survival was lower during hunting seasons (Oct-Nov–Mar-Apr) when restrictive recreational hunting occurred than when no hunting occurred (Fig. 4.3). Second, juvenile survival during May-Jun was higher during closed subsistence hunting, compared to when subsistence hunting

was open (Fig. 4.4). Third, sub-adult and adult survival was lower during May-Jun when subsistence hunting was closed compared to when it was open (Fig. 4.4). Liberal hunting decreased survival in all age classes during Oct-Nov–Mar-Apr, and had the largest effect on juvenile birds (Fig. 4.3). The mean effect of liberal hunting during Oct-Nov–Mar-Apr was similar for sub-adults and adults, but more variable for sub-adults than adults (Fig. 4.3).

4.4.2. OPTIMAL MANAGEMENT STRATEGIES. The action-specific transition probabilities are shown in Fig. 4.1. No hunting resulted in the highest probability of increasing population size in future years, followed by restrictive hunting, then liberal hunting. Transition probabilities were more variable for no hunting than restrictive hunting. The optimal management strategy for the negative squared-error-loss reward function (used to represent the current regulatory framework), and the transition probabilities shown in Fig. 4.1, was no hunting when the population size was less than 80,000 birds, restrictive hunting when the population size was between 80,000 and 235,000 birds, and liberal hunting when the population size was greater than 235,000 birds (Fig. 4.5). The optimal management strategy for the reward functions described using the weighted-sum approach of Chapter 2 (the alternative regulatory framework) was no hunting when the population size was less than 80,000 birds, restrictive hunting when the population size was between 80,000 and 223,000 birds, and liberal hunting when the population size was greater than 223,000 birds (Fig. 4.5).

4.5. DISCUSSION

Bayesian population-dynamics models that incorporate action-specific survival rates provide a natural method for calculating action-specific transition probabilities among population states using posterior predictive distributions. The transition probabilities can be incorporated in an MDP, which can be solved using stochastic dynamic programming to identify optimal management policies. I applied these methods to the problem of selecting harvest regulations for the management of cackling geese.

A requirement for incorporating Bayesian posterior predictive distributions in an MDP is discretization of the posterior predictive distribution into finite states. It is possible that such a discretization does not ensure the posterior uncertainty in transition probabilities are adequately specified (Dorazio and Johnson 2003). Here, I considered 400 discrete states. In earlier simulations I considered 400,000 states. Although the computational requirements were greater, the optimal policy was the same (after rounding). Dorazio and Johnson (2003) discussed the computational limitation of stochastic dynamic programming when the number of state and control variables are large. In the application to cackling geese which only included 400 potential state values, 3 management actions, and a 400×400 transition probability matrix, the computational time to find an optimal solution took less than a second. Thus, combining the Bayesian approach described by Dorazio and Johnson (2003) for objectives that concern cumulative rewards obtained from a future sequence of management actions using stochastic dynamic programming appears to be tractable for problems similar in size to the cackling goose management problem.

MDPs assume that the state s_t in time t can be observed without error. Many ecological applications that use MDPs make this assumption, even when it is clearly violated (e.g., Johnson et al. 1997, this study). Partially observable MDPs (POMDPs) relax the assumption of no observation error (Williams 2009; 2011). POMDPs represent a form of statistical decision theory (*sensu* Williams 2015) that account for recurrent decisions. POMDPs are theoretically attractive because they relax the assumption of no observation error, but they pose the greatest challenge in both analysis and computation. POMDPs can easily exceed computational capacity for many applications due to the added uncertainty incorporated in the observation process (Williams 2009). However, for problems similar in size to the cackling goose management problem, they are likely tractable. Because uncertainty in the observation process is incorporated in Bayesian integrated populations models, developing POMDP is a natural extension for future research for using Bayesian models in a decision theoretic framework.

4.5.1. MANAGEMENT-SPECIFIC SURVIVAL RATES. In Chapter 1, I found that seasonal survival estimates were lowest from Oct–Nov in California, Oregon, and Washington and highest during May–Jun in Alaska. Modeling the effects of management actions and how they interact

with age classes reveals new insight into the population dynamics of cackling geese. In the cackling goose wintering area in the Willamette Valley, survival rates appeared to be affected by the choice of daily bag limits (Fig. 4.3). The difference between no hunting and restrictive hunting (i.e., a one- or two-bird daily bag limit) was small. However, liberal hunting (a three or four bird daily bag limit) appeared to have a large influence on survival rates. The influence of liberal hunting on juveniles was disproportionately large. This result supports other studies investigating the effects of hunting on juvenile birds (Francis et al. 1992*a*, Schmutz et al. 1994, Frederiksen et al. 2004).

I found that survival rates were the highest during the season with the highest subsistence hunting pressure (May–Jun). The closure of the subsistence harvest only appeared to have a positive influence on juvenile survival. It appeared to have a negative influence on sub-adult and adult survival (Fig. 4.4). There are at least four possible explanations for this result. The first explanation is these results were due to random or unexplained variation and any correlation with harvest regulation is coincidental. Mean spring survival on the YK Delta could have been lower from 1985–1994 (years with closed subsistence hunting) than from 1995–2002 (years with open subsistence hunting) due to factors unrelated to harvest regulations. Despite having 18 years of data with 2,605 birds resighted in this space-time combination (Chapter 1, Appendix B), there is no way to rule out this possibility in observational studies. The second explanation is that when recreational hunting is closed, an increased number of birds survive the winter and die due to other factors (e.g., migration) suggesting compensatory hunting mortality realized during spring migration (Anderson and Burnham 1976). The third explanation (related to the second explanation) is that the closure of subsistence hunting during 1984–1993 was ineffective, and subsistence harvest continued. In this scenario, during years of closed recreational hunting, cackling geese did not experience hunting pressure until reaching their nesting area. After they reached their nesting area, they might have been more susceptible to harvest, given they did not experience hunting pressure in their wintering area. An ineffective closure of subsistence hunting would explain similar survival rates between periods of subsistence harvest and an agreed upon closure to subsistence harvest. However, this reasoning does not explain why juvenile survival was higher in years during the agreed closure to

subsistence harvest, as juveniles were the most susceptible to harvest in California, Oregon, and Washington. The fourth explanation (related to the second explanation) is that subsistence harvest has a limited influence on survival rates of cackling geese. Regardless of the explanation, closing subsistence harvest appeared to only have a small impact on spring survival rates of cackling geese in Alaska during the years of these data. Schmutz et al. (1994) and Petersen (1992) found similar survival rates for emperor geese (*Chen canagica*) in periods of subsistence hunting vs. years during the closed subsistence hunting seasons, indicating similar survival rates during both open and closed subsistence hunting are more general than cackling geese.

4.5.2. Optimal management trigger points. The optimal policy for selecting daily bag limits was similar among the two scenarios I considered. The optimal policies only differed by 12,000 birds for when to select restrictive vs. liberal daily bag limits (235,000 vs. 223,000; Fig. 4.5). The reward functions I considered did not account for the value to recreational hunters of potentially increased harvests due to differing regulations, and only considered achieving the population objective. However, I indirectly incorporated hunter satisfaction into the management action of no harvest by specifying reward functions for this action based on the current regulatory framework and independent of transition probabilities. That is, for both management scenarios, I did not use the expected reward calculated in (4.10) for closing hunting and used a piece-wise constant function with breaks at 80,000 birds. The piece-wise constant function ensured no hunting would occur below 80,000 birds and some hunting would occur above 80,000 birds. Extensions to this framework could include objective functions for maximizing the cumulative harvest, while maintaining a population size within a desired limit (e.g., Johnson et al. 1997). These extensions require additional assumptions on regulation-specific harvest rates and the value of each additional harvested bird.

Finally, additional management actions might be considered by incorporating additional data or assumptions. For example, egg collection is a common practice among Native Alaskans. Incorporating the effects of egg-collecting on transition probabilities might be examined by assuming egg collecting has a negative effect on cackling goose fecundity. In future analyses, prior distributions could incorporate knowledge on the impact of egg collecting on fecundity and be propagated

into posterior predictive distributions for transition probabilities and assessed in the MDP framework.

Table 4.1: Management regulations for cackling geese. Daily bag limits and season lengths were obtained from the Federal Register (https://www.federalregister.gov) for the Northwest Special Permit Goose Zone in Oregon. The Northwest Special Permit Zone covers the Willamette Valley, Oregon where 79% of harvested banded geese were shot and reported between 1982–2014. Data on harvest regulations from 1982–1984 were not available in the federal register, but consisted of restrictive harvest regulations (B. Reischus, personal communication). [1]A collaborative effort between federal and state wildlife agencies and Alaska Natives resulted in an agreed closure of cackling goose subsistence hunting between January 1984 and 1993 (Pamplin Jr 1986). [2]Subsistence hunting formally recognized with ammendments to the Migratory Bird Treaty Act in 1997.

Hunting year	Daily bag limit	Season length (days)	Subsistence hunting	Action designation
1982–1983	-	-	Not recognized	Restrictive
1983–1984	-	-	Not recognized	Restrictive
1984–1985	0	0	Closed[1]	Closed
1985–1986	0	0	Closed	Closed
1986–1987	0	0	Closed	Closed
1987–1988	0	0	Closed	Closed
1988–1989	0	0	Closed	Closed
1989–1990	0	0	Closed	Closed
1990–1991	0	0	Closed	Closed
1991–1992	0	0	Closed	Closed
1992–1993	0	0	Closed	Closed
1993–1994	0	0	Closed	Closed
1994–1995	1	23	Not recognized	Restrictive
1995–1996	1	100	Not recognized	Restrictive
1996–1997	2	89	Open[2]	Restrictive
1997–1998	2	89	Open	Restrictive
1998–1999	4	88	Open	Liberal
1999–2000	4	96	Open	Liberal
2000–2001	4	93	Open	Liberal
2001–2002	4	93	Open	Liberal
2002–2003	4	90	Open	Liberal
2003–2004	4	90	Open	Liberal
2004–2005	1	96	Open	Restrictive
2005–2006	2	96	Open	Restrictive
2006–2007	2	96	Open	Restrictive
2007–2008	2	97	Open	Restrictive
2008–2009	2	97	Open	Restrictive
2009–2010	2	97	Open	Restrictive
2010–2011	2	97	Open	Restrictive
2011–2012	2	96	Open	Restrictive
2012–2013	3	97	Open	Liberal
2013–2014	4	98	Open	Liberal
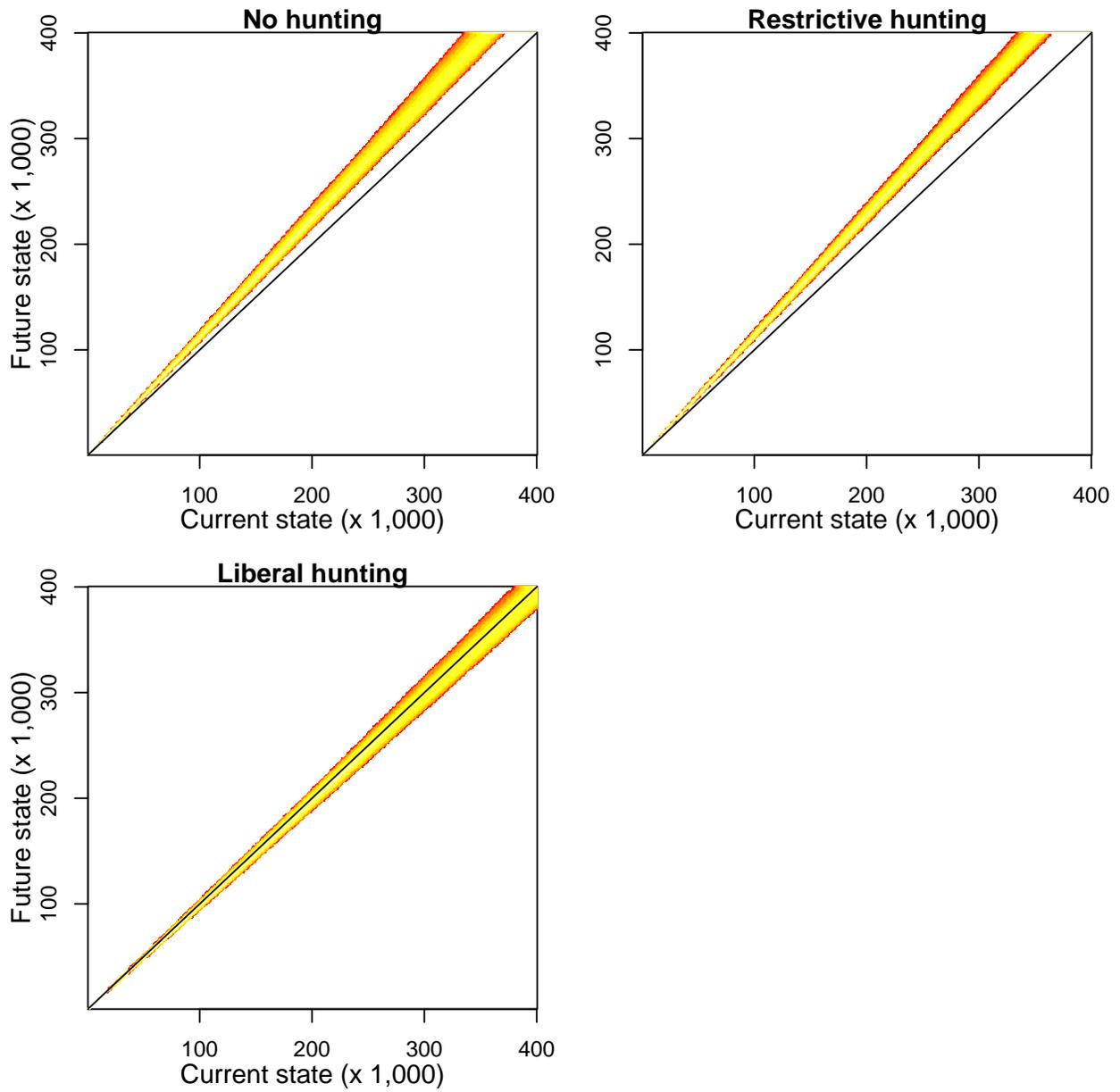
FIGURE 4.1. Transition probabilities representing the probability of moving from the current state (x-axis) to a future state (y-axis) for three different management actions. Lighter colors indicate higher probabilities. The diagonal black line indicates remaining in the current state.
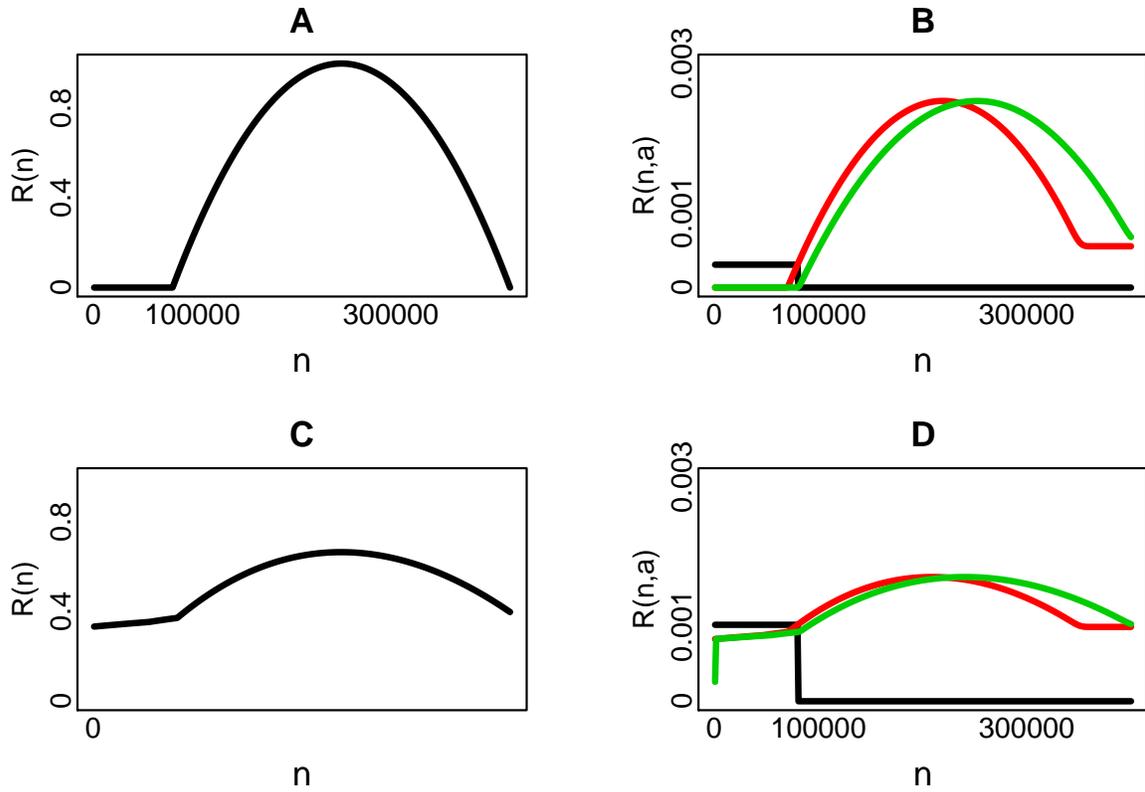
FIGURE 4.2. **A:** Standardized (between zero and one) negative squared error loss function representing the rewards associated with the current management framework with maximum reward occurring at the population objective of $n = 250,000$ birds and no reward below the trigger point at which hunting is closed (i.e., $n = 80,000$ birds). **B:** Expected rewards given the reward function in **A** and transition probabilities in Fig. 4.1. **C:** Reward function representing an alternative management framework derived using the weighted combination of stakeholder utility functions described in Chapter 2. **D:** Expected rewards given the reward function in **C** and action specific transition probabilities in Fig. 4.1. Black indicates closing hunting, red indicates restrictive bag limits (one or two birds per day), and green indicates liberal bag limits (three or four birds per day).
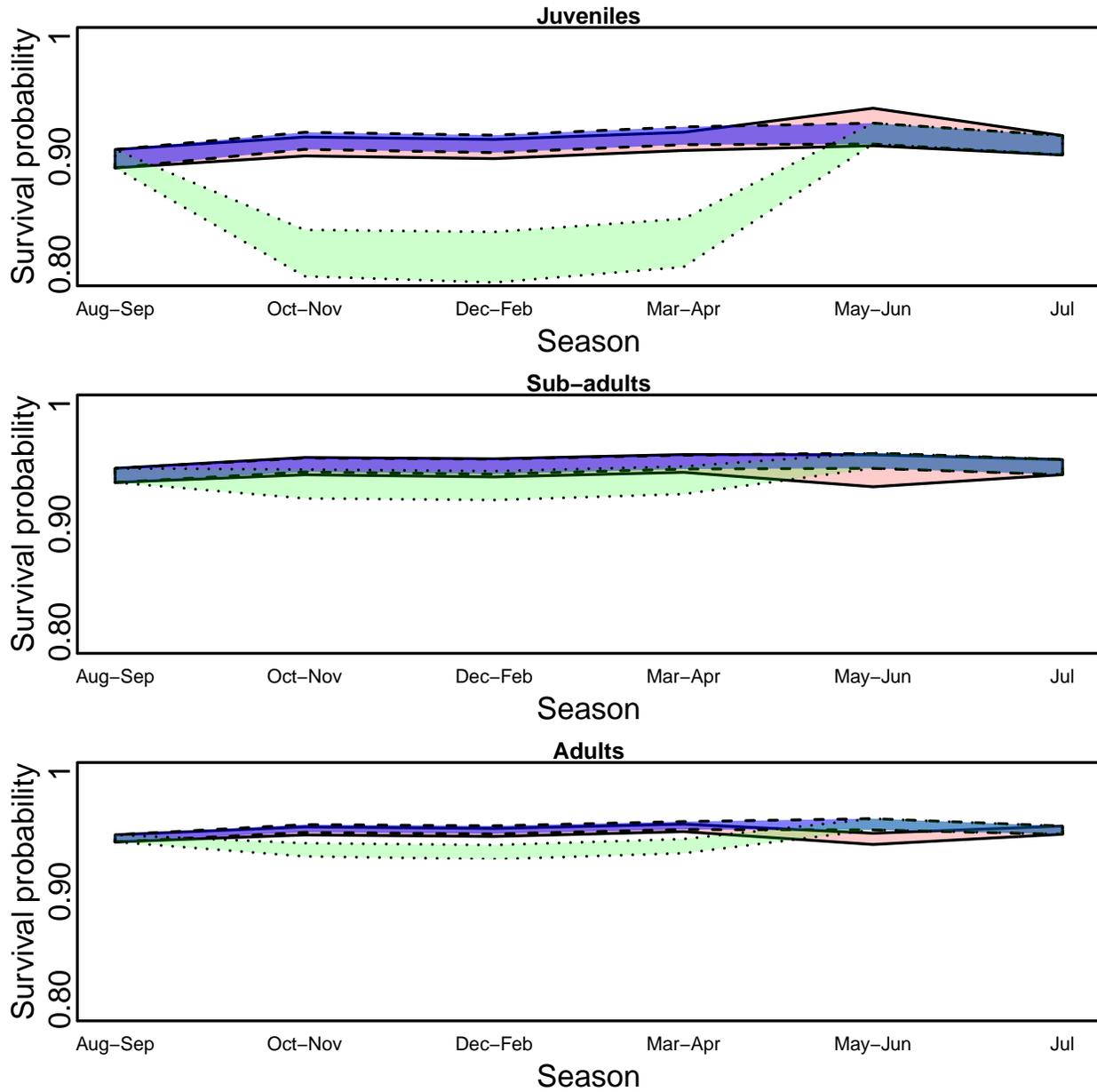
FIGURE 4.3. Age and management specific 95% credible intervals of seasonal survival probability for cackling geese. Red credible intervals indicate no hunting; blue indicates restrictive harvest regulations, and green indicates liberal harvest regulations.
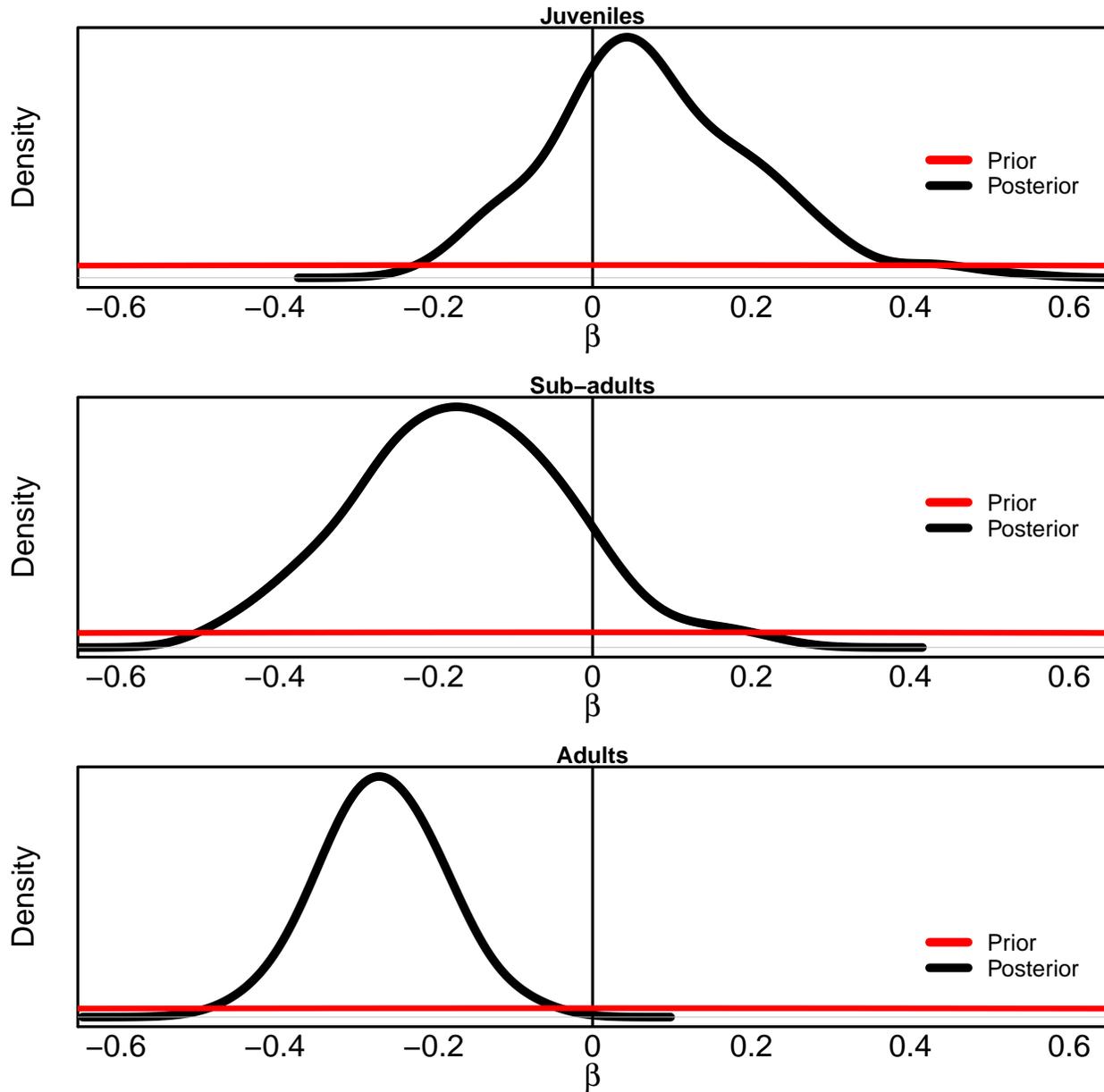
FIGURE 4.4. Prior and posterior distributions for the parameters associated with the effect of closing subsistence harvest on spring survival rates of three age classes of cackling geese in Alaska. Closing subsistence harvest had a slight positive effect on juvenile survival, but a negative effect on sub-adult and adult survival. The closure of subsistence harvest always succeeded closure of cackling goose recreational hunting.
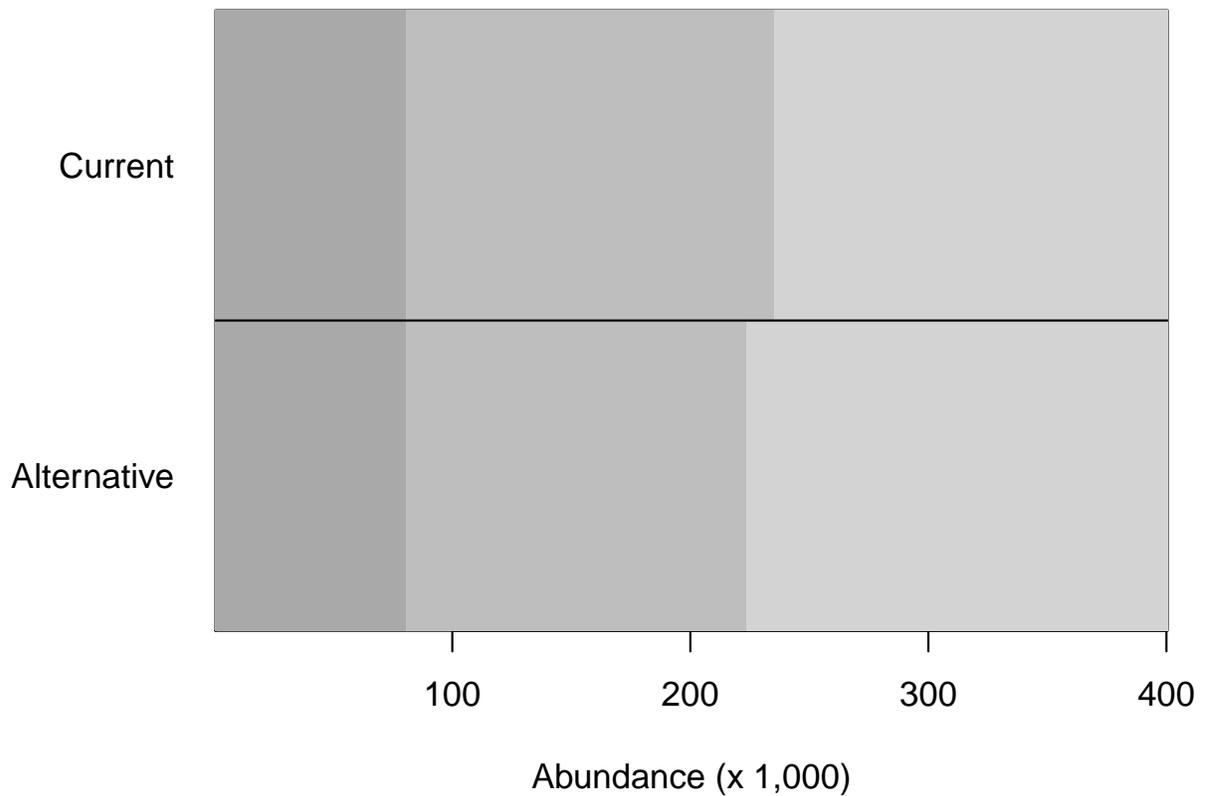
FIGURE 4.5. Optimal management strategies for two management frameworks; the current regulatory framework, and an alternative regulatory management framework that was suggested in Chapter 2. Dark gray represents closing subsistence and recreational hunting and occurs when population sizes fall below 80,000 birds. Medium gray indicates restrictive harvest regulations and occurs between 80,000 and 235,000 birds for the current management framework and 80,000 and 223,000 birds for the alternative management framework. Light gray indicates liberal harvest regulations and occurs above 235,000 for the current management framework and above 223,000 birds for the alternative management framework.

Bibliography

Abadi, F., O. Gimenez, B. Ullrich, R. Arlettaz, and M. Schaub. 2010. Estimation of immigration rate using integrated population models. Journal of Applied Ecology **47**:393–400.

Aho, K., D. Derryberry, and T. Peterson. 2014. Model selection for ecologists: the worldviews of AIC and BIC. Ecology **95**:631–636.

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. Pages 267–281 *in* B. N. Petrov and F. Csáki, editors. 2nd International Symposium on Information Theory.

Alisauskas, R., and M. Lindberg. 2002. Effects of neckbands on survival and fidelity of white-fronted and Canada geese captured as non-breeding adults. Journal of Applied Statistics **29**:521–537.

Ananda, J., and G. Herath. 2003. The use of Analytic Hierarchy Process to incorporate stakeholder preferences into regional forest planning. Forest Policy and Economics **5**:13–26.

Anderson, D. R. 1975. Optimal exploitation strategies for an animal population in a Markovian environment: a theory and an example. Ecology **56**:1281–1297.

Anderson, D. R., and K. P. Burnham, 1976. Population ecology of the mallard: VI. The effect of exploitation on survival. Technical report, US Fish and Wildlife Service.

Ankney, C. D. 1996. An embarrassment of riches: too many geese. The Journal of Wildlife Management **60**:217–223.

Araújo, M. B., and M. Luoto. 2007. The importance of biotic interactions for modelling species distributions under climate change. Global Ecology and Biogeography **16**:743–753.

Athan, T. W., and P. Y. Papalambros. 1996. A note on weighted criteria methods for compromise solutions in multi-objective optimization. Engineering Optimization **27**:155–176.

Balachandran, M., and J. Gero. 1984. A comparison of three methods for generating the Pareto optimal set. Engineering Optimization **7**:319–336.

Barron, H., and C. P. Schmidt. 1988. Sensitivity analysis of additive multiattribute value models. Operations Research **36**:122–127.

Belton, V., and T. Stewart. 2002. Multiple criteria decision analysis: an integrated approach. Springer Science & Business Media.

Berger, J. O. 2013. Statistical decision theory and Bayesian analysis. Springer Science & Business Media.

Bernoulli, D. 1954. Exposition of a new theory on the measurement of risk. Econometrica **22**:23–36.

Besbeas, P., S. N. Freeman, B. J. Morgan, and E. A. Catchpole. 2002. Integrating mark–recapture–recovery and census data to estimate animal abundance and demographic parameters. Biometrics **58**:540–547.

Bridgman, P. W. 1922. Dimensional analysis. Yale University Press.

Brooks, S., R. King, and B. Morgan. 2004. A Bayesian approach to combining animal abundance and demographic data. Animal Biodiversity and Conservation **27**:515–529.

Brown, K., W. N. Adger, E. Tompkins, P. Bacon, D. Shim, and K. Young. 2001. Trade-off analysis for marine protected area management. Ecological Economics **37**:417–434.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer Science & Business Media.

Caswell, H. 2001. Matrix population models. Wiley Online Library.

Caswell, H., and M. Fujiwara. 2004. Beyond survival estimation: mark-recapture, matrix population models, and population dynamics. Animal Biodiversity and Conservation **27**:471–488.

Caughley, G. 1977. Analysis of vertebrate populations. The Blackburn Press.

Cave, V. M., R. King, and S. N. Freeman. 2010. An integrated population model from constant effort bird-ringing data. Journal of Agricultural, Biological, and Environmental Statistics **15**:119–137.

Chades, I., G. Chapron, M.-J. Cros, F. Garcia, and R. Sabbadin, 2014. MDPtoolbox: Markov Decision Processes toolbox. URL `http://CRAN.R-project.org/package=MDPtoolbox`.

Chankong, V., and Y. Haimes. 1983. Multiobjective Decision Making Theory and methodology. Elsevier Science Publishing.

Charnes, A., and W. Cooper. 1977. Goal programming and multiple objective optimizations. Part I European Journal Operational Research **1**:39–54.

Coello, C. C., G. B. Lamont, and D. A. Van Veldhuizen. 2007. Evolutionary algorithms for solving multi-objective problems. Springer Science & Business Media.

Cohon, J. L. 2013. Multiobjective programming and planning. Courier Dover Publications.

Converse, S. J., C. T. Moore, M. J. Folk, and M. C. Runge. 2013. A matter of tradeoffs: reintroduction as a multiple objective decision. The Journal of Wildlife Management **77**:1145–1156.

Cressie, N. A. 1996. Change of support and the modifiable areal unit problem. Geographical Systems **3**:159–180.

Das, I., and J. Dennis, 1999. An improved technique for choosing parameters for Pareto surface generation using normal-boundary intersection. Pages 411–413 *in* Short Paper Proceedings of the Third World Congress of Structural and Multidisciplinary Optimization, volume 2.

Das, I., and J. E. Dennis. 1997. A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. Structural Optimization **14**:63–69.

Davis, A. J., M. B. Hooten, M. L. Phillips, and P. F. Doherty. 2014. An integrated modeling approach to estimating Gunnison sage-grouse population dynamics: combining index and demographic data. Ecology and Evolution **4**:4247–4257.

Deb, K. 2001. Multi-objective optimization using evolutionary algorithms. John Wiley & Sons.

Diaz-Balteiro, L., and C. Romero. 2008. Making forestry decisions with multiple criteria: a review and an assessment. Forest Ecology and Management **255**:3222–3241.

Dorazio, R. M., and F. A. Johnson. 2003. Bayesian inference and decision theory-a framework for decision making in natural resource management. Ecological Applications **13**:556–563.

Doubleday, W. G. 1975. Harvesting in matrix population models. Biometrics **31**:189–200.

Duke, J. M., and R. Aull-Hyde. 2002. Identifying public preferences for land preservation using the analytic hierarchy process. Ecological Economics **42**:131–145.

Edwards, W. 1977. How to use multiattribute utility measurement for social decisionmaking. IEEE Transactions on Systems, Man and Cybernetics **7**:326–340.

Flockhart, D., J.-B. Pichancourt, D. R. Norris, and T. G. Martin. 2015. Unravelling the annual cycle in a migratory animal: breeding-season habitat loss drives population declines of monarch butterflies. Journal of Animal Ecology **84**:155–165.

Folk, T. H., R. R. Holmes, and J. B. Grand. 2007. Variation in northern bobwhite demography along two temporal scales. Population Ecology **49**:211–219.

Francis, C. M., M. H. Richards, F. Cooke, and R. F. Rockwell. 1992*a*. Changes in survival rates of lesser snow geese with age and breeding status. The Auk **109**:731–747.

Francis, C. M., M. H. Richards, F. Cooke, and R. F. Rockwell. 1992*b*. Long-term changes in survival rates of lesser snow geese. Ecology **73**:1346–1362.

Frederiksen, M., R. D. Hearn, C. Mitchell, A. Sigfússon, R. L. Swann, and A. D. Fox. 2004. The dynamics of hunted Icelandic goose populations: a reassessment of the evidence. Journal of Applied Ecology **41**:315–334.

Freitas, A. A. 2004. A critical review of multi-objective optimization in data mining: a position paper. ACM SIGKDD Explorations Newsletter **6**:77–86.

Fretwell, S. D. 1972. Populations in a seasonal environment. Princeton University Press.

Gauthier, G., R. Pradel, S. Menu, and J.-D. Lebreton. 2001. Seasonal survival of greater snow geese and effect of hunting under dependence in sighting probability. Ecology **82**:3105–3119.

Gelfand, A. E., and S. K. Ghosh. 1998. Model choice: a minimum posterior predictive loss approach. Biometrika **85**:1–11.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2014. Bayesian data analysis. Taylor & Francis.

Goodwin, P., and G. Wright. 2004. Decision analysis for management judgement. John Wiley and sons.

Gotway, C. A., and L. J. Young. 2002. Combining incompatible spatial data. Journal of the American Statistical Association **97**:632–648.

Greenberg, P. P., Russelland Marra. 2005. Birds of two worlds: the ecology and evolution of migration. Johns Hopkins University Press.

Gregory, R., L. Failing, M. Harstone, G. Long, T. McDaniels, and D. Ohlson. 2012. Structured decision making: a practical guide to environmental management choices. John Wiley & Sons.

Groot, J. C., and W. A. Rossing. 2011. Model-aided learning for adaptive management of natural resources: an evolutionary design perspective. Methods in Ecology and Evolution **2**:643–650.

Hanks, E. M., M. B. Hooten, and F. A. Baker. 2011. Reconciling multiple data sources to improve accuracy of large-scale prediction of forest disease incidence. Ecological Applications **21**:1173–1188.

Havmøller, R. G., J. Payne, W. Ramono, S. Ellis, K. Yoganand, B. Long, E. Dinerstein, A. C. Williams, R. H. Putra, J. Gawi, et al. 2015. Will current conservation responses save the Critically Endangered Sumatran rhinoceros *Dicerorhinus sumatrensis*? Oryx pages 1–5.

Hennig, C., and M. Kutlukaya. 2007. Some thoughts about the design of loss functions. REVSTAT–Statistical Journal **5**:19–39.

Herath, G. 2004. Incorporating community objectives in improved wetland management: the use of the analytic hierarchy process. Journal of Environmental Management **70**:263–273.

Hestbeck, J. B. 1994. Survival of Canada geese banded in winter in the Atlantic Flyway. The Journal of Wildlife Management **58**:748–756.

Hobbs, N. T., and M. B. Hooten. 2015. Bayesian analysis: A Statistical Primer for Ecologists. Princeton University Press.

Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky. 1999. Bayesian model averaging: a tutorial. Statistical Science **14**:382–401.

Holland-Bartels, L., and B. Pierce, 2011. An evaluation of the science needs to inform decisions on Outer Continental Shelf energy development in the Chukchi and Beaufort Seas, Alaska. Page 278 *in* U.S. Geological Survey Circular 1370. U.S. Geological Survey Circular.

Hooten, M. B., and N. T. Hobbs. 2015. A guide to Bayesian model selection for ecologists. Ecological Monographs **85**:3–28.

Hostetler, J. A., T. S. Sillett, and P. P. Marra. 2015. Full-annual-cycle population models for migratory birds. The Auk **132**:433–449.

Huang, I. B., J. Keisler, and I. Linkov. 2011. Multi-criteria decision analysis in environmental sciences: ten years of applications and trends. Science of the Total Environment **409**:3578–3594.

Hupp, J. W., J. A. Schmutz, and C. R. Ely. 2008. Seasonal survival of radiomarked emperor geese in western Alaska. The Journal of Wildlife Management **72**:1584–1595.

Hwang, C.-L., and A. S. M. Masud. 1979. Multiple objective decision making methods and applications. Springer.

Ims, R. A., and H. P. Andreassen. 2000. Spatial synchronization of vole population dynamics by predatory birds. Nature **408**:194–196.

Jelinski, D. E., and J. Wu. 1996. The modifiable areal unit problem and implications for landscape ecology. Landscape Ecology **11**:129–140.

Johnson, F. A., D. R. Breininger, B. W. Duncan, J. D. Nichols, M. C. Runge, and B. K. Williams. 2011. A Markov decision process for managing habitat for Florida scrub-jays. Journal of Fish and Wildlife Management **2**:234–246.

Johnson, F. A., C. T. Moore, W. L. Kendall, J. A. Dubovsky, D. F. Caithamer, J. R. Kelley Jr, and B. K. Williams. 1997. Uncertainty and the management of mallard harvests. The Journal of Wildlife Management **61**:202–216.

Kaliszewski, I. 1987. A modified weighted Tchebycheff metric for multiple objective programming. Computers & Operations Research **14**:315–323.

Keeney, R. L. 1982. Decision analysis: an overview. Operations Research **30**:803–838.

Keeney, R. L., and H. Raiffa. 1976. Decisions with multiple objectives: preferences and value trade-offs. Cambridge University Press.

Kelly, E. J., and P. L. Kennedy. 1993. A dynamic state variable model of mate desertion in Cooper's hawks. Ecology **74**:351–366.

Kendall, W. L. 2001. Using models to facilitate complex decisions. Modeling in Natural Resource Management, Island Press .

Kéry, M., and M. Schaub. 2012. Bayesian population analysis using WinBUGS: a hierarchical perspective. Access Online via Elsevier.

King, G. 1997. A solution to the ecological inference problem. Princeton University Press.

King, J. G., and D. V. Derksen. 1986. Alaska goose populations: past, present, and future. Transactions of North American Wildlife and Natural Resource Conference **51**:464–479.

Konishi, S., and G. Kitagawa. 1996. Generalised information criteria in model selection. Biometrika **83**:875–890.

Lack, D. 1947. The significance of clutch-size. Ibis **89**:302–352.

Lake, B. C., J. A. Schmutz, M. S. Lindberg, C. R. Ely, W. D. Eldridge, and F. J. Broerman. 2008. Body mass of prefledging Emperor Geese Chen canagica: large-scale effects of interspecific densities and food availability. Ibis **150**:527–540.

Lande, R., S. Engen, and B.-E. Saether. 2003. Stochastic population dynamics in ecology and conservation. Oxford University Press.

Lau, A.-H., and T.-Y. Leong, 1999. PROBES: a framework for probability elicitation from experts. Page 301 *in* Proceedings of the AMIA Symposium. American Medical Informatics Association.

Lebreton, J., and J. Clobert. 1991. Bird population dynamics, management, and conservation: the role of mathematical modelling. Bird population studies: Relevance to conservation and management pages 105–125.

Lepage, D., G. Gauthier, and A. Reed. 1998. Seasonal variation in growth of greater snow goose goslings: the role of food supply. Oecologia **114**:226–235.

Link, W. A., and R. J. Barker. 2006. Model weights and the foundations of multimodel inference. Ecology **87**:2626–2635.

Maguire, L. A., U. S. Seal, and P. F. Brussard, 1987. Managing critically endangered species: the Sumatran rhino as a case study. Pages 141–158 *in* M. E. Soulé, editor. Viable populations for conservation. Cambridge University Press.

Mallows, C. L. 1973. Some comments on C_p. Technometrics **15**:661–675.

Mangel, M., and C. W. Clark. 1986. Towards a unified foraging theory. Ecology **67**:1127–1138.

Marler, R. T., and J. S. Arora. 2004. Survey of multi-objective optimization methods for engineering. Structural and Multidisciplinary Optimization **26**:369–395.

Marler, R. T., and J. S. Arora. 2010. The weighted sum method for multi-objective optimization: new insights. Structural and Multidisciplinary Optimization **41**:853–862.

McWilliams, S., and D. Raveling, 2004. Energetics and time allocation of cackling Canada geese during spring. Pages 99–110 *in* Biology and Management of Canada geese. Proceedings of the 2nd International Canada Goose Symposium, Milwaukee, Wisconsin, USA.

Mendoza, G., and H. Martins. 2006. Multi-criteria decision analysis in natural resource management: a critical review of methods and new modelling paradigms. Forest Ecology and Management **230**:1–22.

Menu, S., G. Gauthier, and A. Reed. 2002. Changes in survival rates and population dynamics of greater snow geese over a 30-year period: implications for hunting regulations. Journal of Applied Ecology **39**:91–102.

Menu, S., J. B. Hestbeck, G. Gauthier, and A. Reed. 2000. Effects of neck bands on survival of greater snow geese. The Journal of Wildlife Management **64**:544–552.

Messac, A., A. Ismail-Yahaya, and C. A. Mattson. 2003. The normalized normal constraint method for generating the Pareto frontier. Structural and Multidisciplinary Optimization **25**:86–98.

Messac, A., and C. A. Mattson. 2002. Generating well-distributed sets of Pareto points for engineering design using physical programming. Optimization and Engineering **3**:431–450.

Messac, A., C. Sukam, and E. Melachrinoudis. 2001. Mathematical and pragmatic perspectives of physical programming. AIAA journal **39**:885–893.

Milner-Gulland, E. 1997. A stochastic dynamic programming model for the management of the saiga antelope. Ecological Applications **7**:130–142.

Mustajoki, J., R. P. Hämäläinen, and A. Salo. 2005. Decision support by interval SMART/SWINGincorporating imprecision in the SMART and SWING methods. Decision Sciences **36**:317–339.

Naves, L. C., 2010. Alaska Migratory Bird Subsistence Harvest Estimates. Technical report, Alaska Department of Fish and Game Division of Subsistence and Alaska Migratory Bird Co-Management Council.

Naves, L. C., 2011. Alaska migratory bird subsistence harvest estimates. Technical report, Alaska Department of Fish and Game Division of Subsistence and Alaska Migratory Bird Co-Management Council.

Naves, L. C., 2012. Alaska Migratory Bird Subsistence Harvest Estimates. Technical report, Alaska Department of Fish and Game Division of Subsistence and Alaska Migratory Bird Co-Management Council.

Nelson, U. C., and H. A. Hansen. 1959. The cackling goose, its migration and management. US Fish and Wildlife Service, Bureau of Sport Fisheries and Wildlife.

O'Neill, E. 1979. Fourteen years of goose populations and trends at Klamath Basin refuges. Management and biology of Pacific Flyway geese. 0regon State Univ. Bookstores, Inc. .

Osyczka, A. 1984. Multicriterion optimization in engineering with fortran programs. John Wiley and Sons.

Pacific Flyway Council. 1999. Pacific Flyway management plan for the cackling Canada goose. Cackling Canada Goose Subcomm., Pacific Flyway Study Comm. [c/o USFWS], Portland, OR. Unpubl. Rept. 36 pp. + appendices.

Pacific Flyway Council. 2015. Pacific Flyway Council Recommendations, Informational Notes, and Subcommittee Reports. Pacific Flyway Council, Omaha, NE. Unpubl. Rept. 145 pp.

Pamplin Jr, W. L., 1986. Cooperative efforts to halt population declines of geese nesting on Alaska's Yukon-Kuskokwim Delta. Pages 487–506 *in* Transactions of the North American Wildlife and Natural Resources Conference, volume 51.

Parviainen, M., M. Luoto, T. Ryttäri, and R. K. Heikkinen. 2008. Modelling the occurrence of threatened plant species in taiga landscapes: methodological and ecological perspectives. Journal of Biogeography **35**:1888–1905.

Petersen, M. R. 1992. Reproductive ecology of emperor geese: survival of adult females. Condor **94**:389–406.

Puterman, M. L. 2014. Markov decision processes: discrete stochastic dynamic programming. John Wiley & Sons.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org/`.

Rao, J., and N. Roy. 1989. Fuzzy set theretic approach of assigning weights to objectives in multicriteria decision making. International Journal of Systems Science **20**:1381–1386.

Raveling, D. G. 1978. The timing of egg laying by northern geese. The Auk **95**:294–303.

Raveling, D. G., 1984. Geese and hunters of Alaska's Yukon Delta: management problems and political dilemmas. Pages 555–575 *in* Transactions of the North American Wildlife and Natural Resources Conference, volume 49.

Raveling, D. G., J. D. Nichols, J. E. Hines, D. S. Zezulak, J. G. Silveira, J. C. Johnson, T. W. Aldrich, and J. A. Weldon. 1992. Survival of cackling Canada geese, 1982-1988. The Journal of Wildlife Management **56**:63–73.

Redpath, S., B. Arroyo, F. Leckie, P. Bacon, N. Bayfield, R. Gutierrez, and S. Thirgood. 2004. Using decision modeling with stakeholders to reduce human–wildlife conflict: a raptor–grouse case study. Conservation Biology **18**:350–359.

Rexstad, E. A. 1992. Effect of hunting on annual survival of Canada geese in Utah. The Journal of Wildlife Management **56**:297–305.

Reynolds, K. M., 2001. Prioritizing salmon habitat restoration with the AHP, SMART, and uncertain data. Pages 199–217 *in* The Analytic Hierarchy Process in Natural Resource and Environmental Decision Making. Springer.

Reynolds, K. M., and P. F. Hessburg. 2005. Decision support for integrated landscape evaluation and restoration planning. Forest ecology and management **207**:263–278.

Rios Insua, D. 1990. Sensitivity analysis in multi-objective decision making. Springer-Verlag.

Rizzolo, D. J., J. A. Schmutz, S. E. McCloskey, and T. F. Fondell. 2014. Factors influencing nest survival and productivity of red-throated loons (*Gavia stellata*) in Alaska. The Condor **116**:574–587.

Saaty, T. L. 1988. What is the analytic hierarchy process? Springer.

Saaty, T. L. 1990. How to make a decision: the analytic hierarchy process. European Journal of Operational Research **48**:9–26.

Samuel, M. D., D. H. Rusch, and S. Craven. 1990. Influence of neck bands on recovery and survival rates of Canada geese. The Journal of Wildlife Management **54**:45–54.

Schaub, M., and F. Abadi. 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. Journal of Ornithology **152**:227–237.

Schaub, M., O. Gimenez, A. Sierro, and R. Arlettaz. 2007. Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. Conservation Biology **21**:945–955.

Schmutz, J. A., S. E. Cantor, and M. R. Petersen. 1994. Seasonal and annual survival of emperor geese. The Journal of Wildlife Management **58**:525–535.

Schmutz, J. A., and C. R. Ely. 1999. Survival of greater white-fronted geese: effects of year, season, sex, and body condition. The Journal of Wildlife Management **63**:1239–1249.

Schmutz, J. A., R. F. Rockwell, and M. R. Petersen. 1997. Relative effects of survival and reproduction on the population dynamics of emperor geese. The Journal of Wildlife Management **61**:191–201.

Schwarz, G., et al. 1978. Estimating the dimension of a model. The Annals of Statistics **6**:461–464.

Sedinger, J. S., and K. S. Bollinger. 1987. Autumn staging of cackling Canada geese on the Alaska Peninsula. Wildfowl **38**:13–18.

Shan, S., and G. G. Wang. 2005. An efficient Pareto set identification approach for multiobjective optimization on black-box functions. Journal of Mechanical Design **127**:866–874.

Shea, K., and H. P. Possingham. 2000. Optimal release strategies for biological control agents: an application of stochastic dynamic programming to population management. Journal of Applied Ecology **37**:77–86.

Sillett, T. S., and R. T. Holmes. 2002. Variation in survivorship of a migratory songbird throughout its annual cycle. Journal of Animal Ecology **71**:296–308.

Skinner, W., R. Jefferies, T. Carleton, and R. Abraham. 1998. Prediction of reproductive success and failure in lesser snow geese based on early season climatic variables. Global Change Biology **4**:3–16.

Snyder, S., C. ReVelle, and R. Haight. 2004. One-and two-objective approaches to an area-constrained habitat reserve site selection problem. Biological Conservation **119**:565–574.

Souchay, G., G. Gauthier, and R. Pradel. 2013. Temporal variation of juvenile survival in a long-lived species: the role of parasites and body condition. Oecologia **173**:151–160.

Steuer, R. E., and E.-U. Choo. 1983. An interactive weighted Tchebycheff procedure for multiple objective programming. Mathematical Programming **26**:326–344.

Stone, M. 1977. An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. Journal of the Royal Statistical Society. Series B (Methodological) **39**:44–47.

U.S. Fish and Wildlife Service. 2014. 2014 Yukon-Kuskokwim Delta Coastal Zone Survey of Geese, Swans, and Sandhill Cranes. U.S. Fish and Wildlife Service, Migratory Bird Management - Region 7.

van Strien, N., B. Manullang, W. Isnan, M. Khan, E. Sumardja, S. Ellis, K. Han, P. Boeadi, and M. Bradley, 2008. *Dicerorhinus sumatrensis*.

Von Neumann, J., and O. Morgenstern. 2007. Theory of games and economic behavior. Princeton University Press.

Ward, D. H., E. A. Rexstad, J. S. Sedinger, M. S. Lindberg, and N. K. Dawe. 1997. Seasonal and annual survival of adult Pacific brant. The Journal of Wildlife Management **61**:773–781.

Westphal, M. I., M. Pickett, W. M. Getz, and H. P. Possingham. 2003. The use of stochastic dynamic programming in optimal landscape reconstruction for metapopulations. Ecological Applications **13**:543–555.

Wierzbicki, A. P., 1986. A methodological approach to comparing parametric characterizations of efficient solutions. Pages 27–45 *in* Large-scale modeling and interactive decision analysis. Springer.

Williams, B. K. 1996. Adaptive optimization and the harvest of biological populations. Mathematical Biosciences **136**:1–20.

Williams, B. K. 2009. Markov decision processes in natural resources management: Observability and uncertainty. Ecological Modelling **220**:830–840.

Williams, B. K. 2011. Resolving structural uncertainty in natural resources management using POMDP approaches. Ecological Modelling **222**:1092–1102.

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press.

Williams, J. C. 1998. Delineating protected wildlife corridors with multi-objective programming. Environmental Modeling & Assessment **3**:77–86.

Williams, P. J., 2015. Combining statistical inference and decisions in ecology. Master's thesis, (Plan B), Colorado State University.

Williams, P. J., and M. B. Hooten. In Review. Combining statistical inference and decisions in ecology. Ecological Applications .

Wolfe, R. J., and R. J. Walker. 1987. Subsistence economies in Alaska: productivity, geography, and development impacts. Arctic Anthropology **24**:56–81.

Yu, P.-L., and G. Leitmann. 1974. Compromise solutions, domination structures, and Salukvadze's solution. Journal of Optimization Theory and Applications **13**:362–378.

Zeleny, M. 1982. Multiple criteria decision making. McGraw Hill.

Zitzler, E., J. Knowles, and L. Thiele, 2008. Quality assessment of Pareto set approximations. Pages 373–404 *in* J. Branke, K. Deb, K. Miettinen, and R. Slowinski, editors. Multiobjective Optimization. Springer.

APPENDIX A

FULL-CONDITIONAL DISTRIBUTIONS

$$[\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{Z}, \boldsymbol{f}, \sigma^2 | \boldsymbol{y}, \boldsymbol{M}] \propto$$

$$\prod_{t=1}^{T} [y_t | \boldsymbol{\beta}, f_t, \sigma^2] \prod_{i=1}^{n} \prod_{j=2}^{J} [m_{i,j} | \boldsymbol{\alpha}]^{z_{i,j}} I_{m_{i,j}=0}^{1-z_{i,j}} [z_{i,j} | \boldsymbol{\beta}]^{z_{i,j-1}} I_{z_{i,j}=0}^{1-z_{i,j-1}} [\boldsymbol{\beta}][\boldsymbol{\alpha}][\boldsymbol{f}][\sigma^2]$$

$$[\boldsymbol{\beta} | \cdot] \propto \prod_{t=1}^{T} [y_t | \boldsymbol{\beta}, f_t, \sigma^2] \prod_{i=1}^{n} \prod_{j=2}^{J} [m_{i,j} | \boldsymbol{\alpha}]^{z_{i,j}} I_{m_{i,j}=0}^{1-z_{i,j}} [z_{i,j} | \boldsymbol{\beta}]^{z_{i,j-1}} [\boldsymbol{\beta}]$$

$$[\boldsymbol{\alpha} | \cdot] \propto \prod_{i=1}^{n} \prod_{j=2}^{J} [m_{i,j} | \boldsymbol{\alpha}]^{z_{i,j}} [\boldsymbol{\alpha}]$$

$$[z_{i,j} | \cdot] \propto [m_{i,j} | \boldsymbol{\alpha}]^{z_{i,j}} I_{m_{i,j}=0}^{1-z_{i,j}} [z_{i,j} | \boldsymbol{\beta}]^{z_{i,j-1}} I_{z_{i,j}=0}^{1-z_{i,j-1}} [z_{i,j+1} | \boldsymbol{\beta}]^{z_{i,j}} I_{z_{i,j+1}=0}^{1-z_{i,j}}$$

$$[\boldsymbol{f} | \cdot] \propto \prod_{t=1}^{T} [y_t | \boldsymbol{\beta}, f_t, \sigma^2][\boldsymbol{f}]$$

$$[\sigma^2 | \cdot] \propto \prod_{t=1}^{T} [y_t | \boldsymbol{\beta}, f_t, \sigma^2][\sigma^2]$$

MARKING AND RESIGHT DATA

Table B.1: Marking and resight numbers of neck-collared cackling geese by time and location.

Year	New marked juve- niles	New marked adults	Jul–Aug (AK)	Oct– Nov (CA, OR, WA)	Dec– Feb (CA, OR, WA)	Mar– Apr (CA, OR, WA)	Apr– May (AK)	Recap (AK)
1982	150	73	0	85	74	32	0	0
1983	71	39	0	92	58	34	0	0
1984	72	47	0	0	0	0	0	0
1985	168	418	0	509	436	325	11	0
1986	245	168	100	572	598	374	36	1
1987	363	170	165	565	501	386	89	12
1988	205	140	97	657	574	297	153	7
1989	205	318	225	745	679	548	149	17
1990	187	369	263	902	869	491	183	22
1991	211	335	283	849	943	610	214	19
1992	297	303	367	959	746	601	252	17
1993	221	394	262	996	1059	839	214	16
1994	288	310	81	911	1050	878	218	24
1995	365	206	434	897	845	735	189	18
1996	258	113	96	731	698	755	207	54
1997	397	152	194	684	868	761	273	72
1998	392	198	77	469	717	654	99	23
1999	285	69	1	493	831	542	185	18

Table B.1: Marking and resight numbers of neck-collared cackling geese by time and location.

Year	New marked juve-niles	New marked adults	Jul–Aug (AK)	Oct–Nov (CA, OR, WA)	Dec–Feb (CA, OR, WA)	Mar–Apr (CA, OR, WA)	Apr–May (AK)	Recap (AK)
2000	244	97	0	457	645	362	56	13
2001	202	79	0	91	407	310	47	35
2002	148	93	0	73	189	155	30	1
2003	11	10	0	135	260	203	0	32
2004	59	8	0	0	0	0	0	0
2005	0	0	0	0	0	0	0	0
2006	0	0	0	0	0	0	0	0
2007	0	0	0	0	0	0	0	0
2008	0	0	0	0	0	0	0	0
2009	0	0	0	0	0	0	0	0
2010	0	0	0	0	0	0	0	0
2011	50	536	0	333	187	240	0	0
2012	0	210	0	148	87	98	0	0
2013	0	547	0	384	318	329	0	0

COMPLETE MODEL STATEMENT FOR CHAPTER 1 AND

CHAPTER 4

$$y_t \sim \text{N}(n_t, \sigma^2) \qquad\qquad t = 1982, ..., 2014$$

$$m_{i,j} \sim \begin{cases} 0 & , z_{i,j} = 0 \\ \text{Bern}(p_{s,t}) & , z_{i,j} = 1 \end{cases} \qquad\qquad j = 1, ..., 192$$

$$z_{i,j} \sim \begin{cases} 0 & , z_{i,j-1} = 0 \\ \text{Bern}(\phi_{s,a,sex,t}) & , z_{i,j-1} = 1 \end{cases}$$

$$n_t = n_{sub,t} + n_{ad,t}$$

$$n_{sub,t} = \frac{n_{ad,t-1}}{2} f_{t-1} \phi_{juv,t-1} \qquad\qquad t = 1983, ..., 2014$$

$$n_{ad,t} = n_{ad,t-1}\phi_{ad,t-1} + n_{sub,t-1}\phi_{sub,t-1}$$

$$\phi_{a,t} = E_{sex}[\phi_{a,sex,t}] \qquad\qquad t = 1982, ..., 2013$$

$$\phi_{a,sex,t} = \prod_{s=1}^{6} \phi_{s,a,sex,t},$$

$$\text{logit}(\boldsymbol{\phi}) = \boldsymbol{X}\boldsymbol{\beta}$$

$$\boldsymbol{\beta} \sim \text{N}(\boldsymbol{0}, 1.5^2\boldsymbol{I})$$

$$\text{logit}(\boldsymbol{p}) = \boldsymbol{W}\boldsymbol{\alpha}$$

$$\boldsymbol{\alpha} \sim \mathrm{N}(\boldsymbol{0}, 1.5^2 \boldsymbol{I})$$

$$f_t \sim \mathrm{Unif}(0, 10)$$

$$N_{ad,1} = 0.8 y_1$$

$$N_{sub,1} = y_1 - N_{ad,1}$$

$$\sigma^2 = \mathrm{IG}(\mathrm{mean} = 5,000^2, \mathrm{var} = 5,000^2)$$

APPENDIX D

CHAPTER 1 MCMC ALGORITHM

```
run.mcmc <- function(y,Y,X,W,I,beta.tune.1,beta.tune.2,alpha.tune,
                     fec.tune,starting.values,n.iter,name){

    #
    # (08June2014)  Perry Williams  [last modified:  09May2015]
    # Integrated Population Model With state-space formulation of
    # Cormack-Jolly-Seber model for mark-resight data (CJS)
    # and population counts.Y
    #
    # Updates:
    # 20 Sept 2014 added c++ code for Z sampler
    # 22 Sept 2014 changed posterior distribution of nsub,x,w
    # 29 Sept 2014 fixed c++ code for Z sampler
    # 05 October added interactions for survival
    # 04 Dec 2014 changed year to a random variable
    # 09 Jan 2015 added adult data
    # 12 Jan 2015 got rid of uncertainty in nsub, x,q
    # 16 Jan 2015 added 2011-2013 data.
    # 09 May 2015 Simplified for chapter 1



    ###
    ### Subroutines and packages
    ###

    library(mvtnorm)    # multivariate normal
    library(Rcpp)
    library(RcppArmadillo)
    library(truncnorm)
    library(compiler)

    sourceCpp('~/Dropbox/PhD/chapter1/analysis/
    supporting.material/cackler.binom.cpp')

    expit=function(x){
      1/(1+exp(-x))
    }

    logit=function(x){
      log(x/(1-x))
```

```
}

calc.annual.survival=function(beta,I,seasons){
  years=32
  annual.survival.juv=numeric(years*2)
  annual.survival.sub=numeric(years*2)
  annual.survival.ad=numeric(years*2)
  beta=matrix(beta,nrow=dim(I)[1],ncol=dim(I)[2],byrow=T)
  survival=1/(1+exp(-rowSums(beta*I)))
  for(i in 1:(years*2)){
    annual.survival.juv[i]=
    prod(survival[(i*18-17):(i*18-17+5)])
  }
  for(i in 1:(years*2)){
    annual.survival.sub[i]=
    prod(survival[(i*18-11):(i*18-11+5)])
  }
  for(i in 1:(years*2)){
    annual.survival.ad[i]=
    prod(survival[(i*18-5):(i*18)])
  }
  male.juv.survival=annual.survival.juv[seq(1,
    length(annual.survival.juv), 2)]
  male.sub.survival=annual.survival.sub[seq(1,
    length(annual.survival.sub), 2)]
  male.ad.survival=annual.survival.ad[seq(1,
    length(annual.survival.ad), 2)]
  female.juv.survival=annual.survival.juv[seq(2,
    length(annual.survival.juv), 2)]
  female.sub.survival=annual.survival.sub[seq(2,
    length(annual.survival.sub), 2)]
  female.ad.survival=annual.survival.ad[seq(2,
    length(annual.survival.ad), 2)]
  juv.survival=rowMeans(cbind(male.juv.survival,
    female.juv.survival))
  sub.survival=rowMeans(cbind(male.sub.survival,
    female.sub.survival))
  ad.survival=rowMeans(cbind(male.ad.survival,
    female.ad.survival))
  list(male.juv.survival=male.juv.survival,
    male.sub.survival=male.sub.survival,
      male.ad.survival=male.ad.survival,
      female.juv.survival=female.juv.survival,
      female.sub.survival=female.sub.survival,
      female.ad.survival=female.ad.survival,
      juv.survival=juv.survival,sub.survival=
```

```
        sub.survival,ad.survival=ad.survival
    )
  }

calc.pop=function(N.ad1,N.sub1,N.ad,N.sub,fec,
  phi.juv,phi.sub,phi.ad){
  N.sub[1]=N.ad1/2*fec[1]*phi.juv[1]
  N.ad[1]=N.sub1*phi.sub[1]+N.ad1*phi.ad[1]
  for(i in 2:years){
    N.sub[i]=(N.ad[i-1]/2)*fec[i]*phi.juv[i]
    N.ad[i]=N.sub[i-1]*phi.sub[i]+N.ad[i-1]*phi.ad[i]
  }
  N=c((N.ad1+N.sub1),(N.sub+N.ad))
  mu=N
  list(mu=mu,N.ad=N.ad,N.sub=N.sub,N=N)
}

invgammastrt <- function(igmn,igvar){
  q <- 2+(igmn^2)/igvar
  r <- 1/(igmn*(q-1))
  list(r=r,q=q)
}

###
### Dimensions
###

years=length(y)-1
n=dim(Y)[1]
J=dim(Y)[2]
seasons=(J-1)/years
ages=3
no.betas=length(X)
no.alphas=length(W)

###
### Hyperparameters
###

s2.hyper=invgammastrt(5000^2,5000^2)
sigma2.r=s2.hyper[1]
sigma2.q=s2.hyper[2]
sigma2.b=1.5^2
mu.b.p=rep(0,seasons)
mu.b.phi=rep(0,no.betas)
N.ad1=y[1]*.8
```

```
N.sub1=y[1]*.2

###
### Empty Matrices, and acceptance parameters
###

accept.N.ad1=0
accept.N.sub1=0
accept.fec=0
fec.tune.save=numeric(n.iter)
accept.alpha=0
accept.beta.1=0
accept.beta.2=0
beta.tune.save.1=numeric(n.iter)
beta.tune.save.2=numeric(n.iter)
N.sub=numeric(years)
N.ad=numeric(years)
sigma2.y.save=numeric(n.iter)
fec.save=matrix(NA,nrow=n.iter,ncol=years)
alpha.save=matrix(NA,nrow=n.iter,ncol=no.alphas)
beta.save=matrix(NA,nrow=n.iter,ncol=no.betas)
mu.save=matrix(NA,nrow=n.iter,ncol=years+1)
N.ad.save=matrix(NA,nrow=n.iter,ncol=years)
N.sub.save=matrix(NA,nrow=n.iter,ncol=years)
N.save=matrix(NA,nrow=n.iter,ncol=years+1)

###
### Starting values
###

beta.star=rnorm(40,0,1)
Y_tmp=Y
Y_tmp[is.na(Y_tmp)]=0
alpha=starting.values[[2]]
alpha[c(8,28:34)]=-10
logit.p=W[[1]]*alpha[1]
for(i in 2:no.alphas){
  logit.p=logit.p+W[[i]]*alpha[i]
}
p=expit(logit.p)
beta=starting.values[[3]]
beta[c(12,31:37)]=0
logit.phi=X[[1]]*beta[1]
for(i in 2:no.betas){
  logit.phi=logit.phi+X[[i]]*beta[i]
}
```

```
phi=expit(logit.phi)
N.ad=starting.values[[4]]
N.sub=starting.values[[5]]
N=starting.values[[6]]
phi.juv=rep(0.6,32)
phi.sub=rep(0.8,32)
phi.ad=rep(0.8,32)
Z=matrix(0,nrow=n,ncol=J)
Z[,1]=1
for(t in 2:(J-1)){
   Y.tmp=apply(Y[,t:J],1,sum,na.rm=TRUE)==0
   #if bird not seen in any future years Y.tmp=TRUE
   Z.tmp=Z[,t-1]==1 & Z[,t+1]==0
   Z[,t]=ifelse(Y.tmp,
                #if bird not seen this year (Y.tmp==TRUE):
                ifelse(Z.tmp,
                       #if bird alive last
                       #year and dead next year Z.tmp=TRUE:
                       rbinom(n,1,(phi[,t-1]*(1-phi[,t-1])*
                         (1-p[,t-1]))/
                         ((phi[,t-1]*(1-phi[,t-1])*
                         (1-p[,t-1]))+1-phi[,t-1])),
                       #survival probability is binom dist.
                       #with conjugate parameters for phi
                       ifelse(Z[,t-1]==0,0,
                              #if bird dead last year
                              #its dead this year
                              1)),
                #if bird alive last year
                #and alive next year its alive this year
                1)
   #if birds seen sometime in the future Y.tmp==1,
   #its still alive.
}
ZY.tmp=Z[,J-1]==1 & Y[,J]==0
#If alive penultimate year but not seen in
#ultimate year
ZY2.tmp=Z[,J-1]==0 & Y[,J]==0
#If dead in penultimate year and not seen in ultimate year
Z[,J]=ifelse(ZY.tmp,rbinom(n,1,(phi[,J-1]*(1-p[,J-1]))/
                                ((phi[,J-1]*(1-p[,J-1]))+
                                1-phi[,J-1])),ifelse(ZY2.tmp,0,1))
fec=pmin(starting.values[[1]],4.5)
pop=calc.pop(N.ad1,N.sub1,N.ad,N.sub,fec,phi.juv,phi.sub,phi.ad)
survive=calc.annual.survival(beta,I,seasons)
Z_new=matrix(0,nrow=n,ncol=J)
```

```
   Z_new[,1]=1

   ###
   ### MCMC Loop
   ###

   for(k in 1:n.iter){
     if(k%%100==0) cat(k," ")

     ###
     ###  Sample sigma2.y
     ###

     tmp.q=length(y)/2+sigma2.q$q
     tmp.r=(1/sigma2.r$r+.5*sum((y-pop$mu)^2))^(-1)
     sigma2.y=1/rgamma(1,tmp.q,,tmp.r)

     ###
     ### Sample fec
     ###

#      if(accept.fec/k < 0.3){   # used during burn in to tune
#        fec.tune=max(fec.tune -.001, 0.0001)
#      }
#      if(accept.fec/k >0.5){
#        fec.tune=fec.tune+0.001
#      }
     fec.star=rtruncnorm(years,a=0,b=15,fec,fec.tune)
     pop.star=calc.pop(N.ad1,N.sub1,pop$N.ad,pop$N.sub,
       fec.star,survive$juv.survival,survive$sub.survival,
       survive$ad.survival)
     pop=calc.pop(N.ad1,N.sub1,pop$N.ad,pop$N.sub,fec,
       survive$juv.survival,survive$sub.survival,
       survive$ad.survival)

     #Metropolis-Hastings
     mh1 = sum(dnorm(y,pop.star$mu,sqrt(sigma2.y),log=TRUE)) +
       sum(log(dtruncnorm(fec,a=0,b=Inf,fec.star,fec.tune)))
     mh2 = sum(dnorm(y,pop$mu,sqrt(sigma2.y),log=TRUE)) +
       sum(log(dtruncnorm(fec.star,a=0,b=Inf,fec,fec.tune)))
     mh=min(exp(mh1-mh2),1)
     if(mh>runif(1)){
       pop=pop.star
       fec=fec.star
       accept.fec=accept.fec+1
     }
```

```
####
####  Sample Z
####

Z=sample_Z(J,as.matrix(Y_tmp),Z,Z_new,phi,p)
ZY.tmp=Z[,J-1]==1 & Y_tmp[,J]==0
ZY2.tmp=Z[,J-1]==0 & Y_tmp[,J]==0
Z[,J]=ifelse(ZY.tmp,rbinom(n,1,(phi[,J-1]*(1-p[,J-1]))/
                                ((phi[,J-1]*(1-p[,J-1]))+1-
                                phi[,J-1])),ifelse(ZY2.tmp,0,1))


###
###  Sample alphas
###

#Proposal
alpha.star=rnorm(no.alphas,alpha,alpha.tune)
alpha.star[c(8,28:34)]=-10
logit.p.star=W[[1]]*alpha.star[1]
logit.p=W[[1]]*alpha[1]
for(i in 2:no.alphas){
  logit.p.star=logit.p.star+W[[i]]*alpha.star[i]
  logit.p=logit.p+W[[i]]*alpha[i]
}
p.star=expit(logit.p.star)
p.star.na=p.star
p=expit(logit.p)
p.na=p
p.star.na[is.na(Y[,2:(dim(Y)[2])])]=NA
p.na[is.na(Y[,2:(dim(Y)[2])])]=NA

#Metropolis-Hastings
mh1 = sum(dbinom(Y[,-1][Z[,-1]==1],1,
  p.star.na[Z[,-1]==1],log=TRUE),na.rm=TRUE) +
  dnorm(alpha.star,mu.b.p,sqrt(sigma2.b),log=TRUE)
mh2 = sum(dbinom(Y[,-1][Z[,-1]==1],1,p.na[Z[,-1]==1],
  log=TRUE),na.rm=TRUE) +
  dnorm(alpha,mu.b.p,sqrt(sigma2.b),log=TRUE)
mh=min(exp(mh1-mh2),1)
if(mh>runif(1)){
  alpha=alpha.star
  p=p.star
  accept.alpha=accept.alpha+1
}
```

```
###
###   Sample betas 1
###

#Proposal
#    if(accept.beta.1/k < 0.3){beta.tune.1= #used for tuning only
#       max(beta.tune.1 -.001, 0.0001)}
#    if(accept.beta.1/k >0.5){beta.tune.1=
#       beta.tune.1+0.001}
beta.star[1:9]=rnorm(9,beta[1:9],beta.tune.1)
beta.star[10:40]=beta[10:40]
beta.star[c(12,31:37)]=0
logit.phi.star=X[[1]]*beta.star[1]
logit.phi=X[[1]]*beta[1]
for(i in 2:40){
  logit.phi.star=logit.phi.star+X[[i]]*beta.star[i]
  logit.phi=logit.phi+X[[i]]*beta[i]
}
phi.star=expit(logit.phi.star)
phi.star.na=phi.star
phi=expit(logit.phi)
phi.na=phi
phi.star.na[is.na(Y[,2:(dim(Y)[2])])]=NA
phi.na[is.na(Y[,2:(dim(Y)[2])])]=NA
survive.star=calc.annual.survival(beta.star,I,seasons)
survive=calc.annual.survival(beta,I,seasons)
pop.star=calc.pop(N.ad1,N.sub1,pop$N.ad,pop$N.sub,fec,
  survive.star$juv.survival,
  survive.star$sub.survival,
  survive.star$ad.survival)
pop=calc.pop(N.ad1,N.sub1,pop$N.ad,pop$N.sub,fec,
  survive$juv.survival,
  survive$sub.survival,
  survive$ad.survival)

#Metropolis-Hastings
mh1 = sum(dbinom(Z[,-1],Z[,-J],phi.star.na,log=TRUE),
  na.rm=TRUE) + sum(dnorm(y,pop.star$mu,sqrt(sigma2.y),
  log=TRUE)) + sum(dnorm(beta.star,mu.b.phi,sqrt(sigma2.b),
  log=TRUE))
mh2 = sum(dbinom(Z[,-1],Z[,-J],phi.na,log=TRUE),na.rm=TRUE) +
  sum(dnorm(y,pop$mu,sqrt(sigma2.y),log=TRUE)) +
  sum(dnorm(beta,mu.b.phi,sqrt(sigma2.b),log=TRUE))
mh=min(exp(mh1-mh2),1)
if(mh>runif(1)){
  beta=beta.star
```

```
    pop=pop.star
    survive=survive.star
    phi=phi.star
    accept.beta.1=accept.beta.1+1
}


###
###   Sample betas 2
###

#Proposal
#    if(accept.beta.2/k < 0.3){beta.tune.2= #used for tuning only
#      max(beta.tune.2 -.001, 0.0001)}
#    if(accept.beta.2/k >0.5){beta.tune.2=
#     beta.tune.2+0.001}
    beta.star[10:40]=rnorm(31,beta[10:40],beta.tune.2)
    beta.star[1:9]=beta[1:9]
    beta.star[c(12,31:37)]=0
    logit.phi.star=X[[1]]*beta.star[1]
    logit.phi=X[[1]]*beta[1]
    for(i in 2:40){
      logit.phi.star=logit.phi.star+X[[i]]*beta.star[i]
      logit.phi=logit.phi+X[[i]]*beta[i]
    }
    phi.star=expit(logit.phi.star)
    phi.star.na=phi.star
    phi=expit(logit.phi)
    phi.na=phi
    phi.star.na[is.na(Y[,2:(dim(Y)[2])])]=NA
    phi.na[is.na(Y[,2:(dim(Y)[2])])]=NA
    survive.star=calc.annual.survival(beta.star,I,seasons)
    survive=calc.annual.survival(beta,I,seasons)
    pop.star=calc.pop(N.ad1,N.sub1,pop$N.ad,pop$N.sub,fec,
      survive.star$juv.survival,
      survive.star$sub.survival,
      survive.star$ad.survival)
    pop=calc.pop(N.ad1,N.sub1,pop$N.ad,pop$N.sub,fec,
                survive$juv.survival,
                survive$sub.survival,
                survive$ad.survival)

    #Metropolis-Hastings
    mh1 = sum(dbinom(Z[,-1],Z[,-J],phi.star.na,log=TRUE),
      na.rm=TRUE) + sum(dnorm(y,pop.star$mu,sqrt(sigma2.y),
      log=TRUE)) + sum(dnorm(beta.star,mu.b.phi,sqrt(sigma2.b),
      log=TRUE))
```

144

```
     mh2 = sum(dbinom(Z[,-1],Z[,-J],phi.na,log=TRUE),na.rm=TRUE) +
       sum(dnorm(y,pop$mu,sqrt(sigma2.y),log=TRUE)) +
       sum(dnorm(beta,mu.b.phi,sqrt(sigma2.b),log=TRUE))
     mh=min(exp(mh1-mh2),1)
     if(mh>runif(1)){
       beta=beta.star
       pop=pop.star
       survive=survive.star
       phi=phi.star
       accept.beta.2=accept.beta.2+1
     }


     ###
     ### Save Samples
     ###

     sigma2.y.save[k]=sigma2.y
     fec.save[k,]=fec
     alpha.save[k,]=alpha
     beta.save[k,]=beta
     N.ad.save[k,]=pop$N.ad
     N.sub.save[k,]=pop$N.sub
     N.save[k,]=pop$N
#     beta.tune.save.1[k]=beta.tune.1
#     beta.tune.save.2[k]=beta.tune.2
#     fec.tune.save[k]=fec.tune
     if(k%%100==0){
       out=list(sigma2.y.save=sigma2.y.save,
                fec.save=fec.save,
                alpha.save=alpha.save,
                beta.save=beta.save,
                N.ad.save=N.ad.save,
                N.sub.save=N.sub.save,
                N.save=N.save,
                accept.beta.1=accept.beta.1,
                accept.beta.2=accept.beta.2,
                accept.alpha=accept.alpha,
                accept.fec=accept.fec,
#                beta.tune.save.1=beta.tune.save.1,
#                beta.tune.save.2=beta.tune.save.2,
#                fec.tune.save=fec.tune.save
       )
       save(out,file=name)
     }
     cat("\n")
   }
```

}

CHAPTER 4 MCMC ALGORITHM

```
run.mcmc <- function(y,Y,X,W,I,beta.tune,alpha.tune,
                     fec.tune,starting.values,n.iter,name){

  #
  # (08June2014)  Perry Williams  [last modified:  5Dec2014]
  # Integrated Population Model With state-space formulation of
  # Cormack-Jolly-Seber model for mark-resight data (CJS)
  # and population counts.Y
  #
  # Updates:
  # 20 Sept 2014 added c++ code for Z sampler
  # 22 Sept 2014 changed posterior distribution of nsub,x,w
  # 29 Sept 2014 fixed c++ code for Z sampler
  # 05 October added interactions for survival
  # 04 Dec 2014 changed year to a random variable
  # 09 Jan 2015 added adult data
  # 12 Jan 2015 got rid of uncertainty in nsub, x,q
  # 16 Jan 2015 added 2011-2013 data.
  # 09 May 2015 Simplified for chapter 1
  # 03 Aug 2015 changed for chapter 3


  ###
  ### Subroutines and packages
  ###

  library(mvtnorm)    # multivariate normal
  library(Rcpp)
  library(RcppArmadillo)
  library(truncnorm)
  library(compiler)

  sourceCpp('~/Dropbox/PhD/chapter1/analysis/
           supporting.material/cackler.binom.cpp')

  expit=function(x){
    1/(1+exp(-x))
  }

  logit=function(x){
```

```r
  log(x/(1-x))
}

invgammastrt <- function(igmn,igvar){
  q <- 2+(igmn^2)/igvar
  r <- 1/(igmn*(q-1))
  list(r=r,q=q)
}

I1=c(0,0,1,1,1,1,1,1,1,1,1,1,0,0,0,0,0,
  0,0,0,0,0,0,0,0,0,0,0,0,0,0)  # No hunting
I2=c(1,1,0,0,0,0,0,0,0,0,0,0,1,1,1,1,0,
  0,0,0,0,0,1,1,1,1,1,1,1,1,0,0)  # Restrictive
I3=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,
  1,1,1,1,1,0,0,0,0,0,0,0,0,0,0)  # Liberal
I.=cbind(I1,I2,I3)

calc.annual.survival=function(beta,I.){
  # closed, restrictive, liberal
  juv.surv=c(
    expit(beta[1])*           #d closed
      expit(beta[1]+beta[2])* #r
      expit(beta[1]+beta[3])* #a
      expit(beta[1]+beta[4])* #c
      expit(beta[1]+beta[5])* #e
      expit(beta[1]+beta[6]), #f
    expit(beta[1]+beta[9])* #d restrictive
      expit(beta[1]+beta[2])*
      expit(beta[1]+beta[3])*
      expit(beta[1]+beta[4]+beta[9])*
      expit(beta[1]+beta[5]+beta[9])*
      expit(beta[1]+beta[6]),
    expit(beta[1]+beta[10])* #d liberal
      expit(beta[1]+beta[2])*
      expit(beta[1]+beta[3])*
      expit(beta[1]+beta[4]+beta[10])*
      expit(beta[1]+beta[5]+beta[10])*
      expit(beta[1]+beta[6]))
  sub.surv=c(
    expit(beta[1]+beta[7])*           #d closed
      expit(beta[1]+beta[2]+beta[7])* #r
      expit(beta[1]+beta[3]+beta[7])* #a
      expit(beta[1]+beta[4]+beta[7])* #c
      expit(beta[1]+beta[5]+beta[7])* #e
      expit(beta[1]+beta[6]+beta[7]+beta[12]+beta[16]), #f
    expit(beta[1]+beta[7]+beta[9])* #d restrictive
```

```
      expit(beta[1]+beta[2]+beta[7])*
      expit(beta[1]+beta[3]+beta[7])*
      expit(beta[1]+beta[4]+beta[7]+beta[9]+beta[13])*
      expit(beta[1]+beta[5]+beta[7]+beta[9]+beta[13])*
      expit(beta[1]+beta[6]+beta[7]),
    expit(beta[1]+beta[7]+beta[10]+beta[14])* #a liberal
      expit(beta[1]+beta[2]+beta[7])*
      expit(beta[1]+beta[3]+beta[7])*
      expit(beta[1]+beta[4]+beta[7]+beta[10]+beta[14])*
      expit(beta[1]+beta[5]+beta[7]+beta[10]+beta[14])*
      expit(beta[1]+beta[6]+beta[7]))
  ad.surv=c(
    expit(beta[1]+beta[8])*          #d Closed
      expit(beta[1]+beta[2]+beta[8])* #r
      expit(beta[1]+beta[3]+beta[8])* #a
      expit(beta[1]+beta[4]+beta[8])* #c
      expit(beta[1]+beta[5]+beta[8])* #e
      expit(beta[1]+beta[6]+beta[8]+beta[12]+beta[20]), #f
    expit(beta[1]+beta[8]+beta[9]+beta[17])* #r restrictive
      expit(beta[1]+beta[2]+beta[8])*
      expit(beta[1]+beta[3]+beta[8])*
      expit(beta[1]+beta[4]+beta[8]+beta[9]+beta[17])*
      expit(beta[1]+beta[5]+beta[8]+beta[9]+beta[17])*
      expit(beta[1]+beta[6]+beta[8]),
    expit(beta[1]+beta[8]+beta[10]+beta[18])* #a liberal
      expit(beta[1]+beta[2]+beta[8])*
      expit(beta[1]+beta[3]+beta[8])*
      expit(beta[1]+beta[4]+beta[8]+beta[10]+beta[18])*
      expit(beta[1]+beta[5]+beta[8]+beta[10]+beta[18])*
      expit(beta[1]+beta[6]+beta[8]))
  juv.annual=juv.surv[1]*I.[,1]+juv.surv[2]*I.[,2]+
    juv.surv[3]*I.[,3]
  sub.annual=sub.surv[1]*I.[,1]+sub.surv[2]*I.[,2]+
    sub.surv[3]*I.[,3]
  ad.annual=ad.surv[1]*I.[,1]+ad.surv[2]*I.[,2]+
    ad.surv[3]*I.[,3]
  cbind(juv.annual,sub.annual,ad.annual)
}

calc.pop=function(N.ad1,N.sub1,N.ad,N.sub,
  fec,phi.juv,phi.sub,phi.ad){
    N.sub[1]=N.ad1/2*fec[1]*phi.juv[1]
    N.ad[1]=N.sub1*phi.sub[1]+N.ad1*phi.ad[1]
    for(i in 2:years){
      N.sub[i]=(N.ad[i-1]/2)*fec[i]*phi.juv[i]
      N.ad[i]=N.sub[i-1]*phi.sub[i]+N.ad[i-1]*phi.ad[i]
```

```
    }
    N=c((N.ad1+N.sub1),(N.sub+N.ad))
    mu=N
    list(mu=mu,N.ad=N.ad,N.sub=N.sub,N=N)
}


###
### Dimensions
###

years=length(y)-1
n=dim(Y)[1]
J=dim(Y)[2]
seasons=(J-1)/years
ages=3
no.betas=length(X)
no.alphas=length(W)


###
### Hyperparameters
###

hp.sigma2=invgammastrt(5000^2,5000^2)
sigma2.r=hp.sigma2$r
sigma2.q=hp.sigma2$q
sigma2.b=1.5^2
mu.alpha=rep(0,seasons)
mu.beta=rep(0,no.betas)


###
### Empty Matrices, and acceptance parameters
###

accept.N.ad1=0
accept.N.sub1=0
accept.fec=0
accept.alpha=0
accept.beta=0

sigma2.y.save=numeric(n.iter)
N.ad1.save=numeric(n.iter)
N.sub1.save=numeric(n.iter)
N.sub=numeric(years)
N.ad=numeric(years)
fec.save=matrix(NA,nrow=n.iter,ncol=years)
alpha.save=matrix(NA,nrow=n.iter,ncol=no.alphas)
```

```
beta.save=matrix(NA,nrow=n.iter,ncol=no.betas)
mu.save=matrix(NA,nrow=n.iter,ncol=years+1)
N.ad.save=matrix(NA,nrow=n.iter,ncol=years)
N.sub.save=matrix(NA,nrow=n.iter,ncol=years)
N.save=matrix(NA,nrow=n.iter,ncol=years+1)
beta.tune.save=numeric(n.iter)
fec.tune.save=numeric(n.iter)

###
### Starting Values
###

Y_tmp=Y
Y_tmp[is.na(Y_tmp)]=0
sigma2.y=5000^2
N.ad1=y[1]*.8
N.sub1=y[1]*.2
alpha=starting.values[[2]]
alpha[c(8,28:34)]=-10
logit.p=W[[1]]*alpha[1]
for(i in 2:no.alphas){
  logit.p=logit.p+W[[i]]*alpha[i]
}
p=expit(logit.p)
beta=starting.values[[3]]
beta[c(11,15)]=0
logit.phi=X[[1]]*beta[1]
for(i in 2:no.betas){
  logit.phi=logit.phi+X[[i]]*beta[i]
}
phi=expit(logit.phi)
N.ad=starting.values[[3]]
N.sub=starting.values[[5]]
N=starting.values[[6]]
phi.juv=starting.values[[7]]
phi.sub=starting.values[[8]]
phi.ad=starting.values[[9]]
Z=matrix(0,nrow=n,ncol=J)
Z[,1]=1
for(t in 2:(J-1)){
  Y.tmp=apply(Y[,t:J],1,sum,na.rm=TRUE)==0
  #if bird not seen in any future years Y.tmp=TRUE
  Z.tmp=Z[,t-1]==1 & Z[,t+1]==0
  Z[,t]=ifelse(Y.tmp,
               #if bird not seen this year (Y.tmp==TRUE):
               ifelse(Z.tmp,
```

```
                              #if bird alive last year
                              #and dead next year Z.tmp=TRUE:
                              rbinom(n,1,(phi[,t-1]*(1-phi[,t-1])*
                              (1-p[,t-1]))/
                                       ((phi[,t-1]*(1-phi[,t-1])*
                                       (1-p[,t-1]))+1-phi[,t-1])),
                              #survival probability is
                              #binom dist. with conjugate
                              #parameters for phi
                              ifelse(Z[,t-1]==0,0,
                              #if bird dead last year
                              #its dead this year
                              1)),
                              #if bird alive last year
                              #and alive next year its alive this year
                    1)
                    #if birds seen sometime in the
                    #future Y.tmp==1, its still alive.
}
ZY.tmp=Z[,J-1]==1 & Y[,J]==0
#If alive penultimate year but not seen in final year
ZY2.tmp=Z[,J-1]==0 & Y[,J]==0
#If dead in penultimate year and not seen in final year
Z[,J]=ifelse(ZY.tmp,rbinom(n,1,(phi[,J-1]*(1-p[,J-1]))/
                                ((phi[,J-1]*(1-p[,J-1]))+
                                1-phi[,J-1])),
                                ifelse(ZY2.tmp,0,1))
fec=starting.values[[1]]
pop=calc.pop(N.ad1,N.sub1,N.ad,N.sub,fec,
  phi.juv,phi.sub,phi.ad)
survive=calc.annual.survival(beta,I.)
Z_new=matrix(0,nrow=n,ncol=J)
Z_new[,1]=1

###
### MCMC Loop
###

for(k in 1:n.iter){
  if(k%%100==0) cat(k," ")

  ###
  ###  Sample sigma2.y
  ###

   tmp.q=length(y)/2+sigma2.q
```

```
    tmp.r=(1/sigma2.r+.5*sum((y-pop$mu)^2))^(-1)
     sigma2.y=1/rgamma(1,tmp.q,,tmp.r)


     ###
     ### Sample fec
     ###

#       if(accept.fec/k < 0.3){ #used for tuning only
#          fec.tune=max(fec.tune -.001, 0.0001)
#       }
#       if(accept.fec/k >0.5){
#          fec.tune=fec.tune+0.001
#       }
     fec.star=rtruncnorm(years,a=0,b=15,
       fec,fec.tune)
     pop.star=calc.pop(N.ad1,N.sub1,pop$N.ad,
       pop$N.sub,fec.star,
       survive[,1],survive[,2],survive[,3])
     pop=calc.pop(N.ad1,N.sub1,pop$N.ad,pop$N.sub,fec,
                   survive[,1],survive[,2], survive[,3])

     #Metropolis-Hastings
     mh1 = sum(dnorm(y,pop.star$mu,sqrt(sigma2.y),
       log=TRUE)) +sum(log(dtruncnorm(fec,
       a=0,b=Inf,fec.star,fec.tune)))
     mh2 = sum(dnorm(y,pop$mu,sqrt(sigma2.y),
       log=TRUE)) + sum(log(dtruncnorm(fec.star,
       a=0,b=Inf,fec,fec.tune)))
     mh=min(exp(mh1-mh2),1)
     if(mh>runif(1)){
       pop=pop.star
       fec=fec.star
       accept.fec=accept.fec+1
     }

     ####
     ####  Sample Z
     ####

     Z=sample_Z(J,as.matrix(Y_tmp),Z,Z_new,phi,p)
     ZY.tmp=Z[,J-1]==1 & Y_tmp[,J]==0
     ZY2.tmp=Z[,J-1]==0 & Y_tmp[,J]==0
     Z[,J]=ifelse(ZY.tmp,rbinom(n,1,(phi[,J-1]*(1-p[,J-1]))/
       ((phi[,J-1]*(1-p[,J-1]))+1-phi[,J-1])),
       ifelse(ZY2.tmp,0,1))
```

```
###
###   Sample alphas
###

#Proposal
alpha.star=rnorm(no.alphas,alpha,alpha.tune)
alpha.star[c(8,28:34)]=-10
logit.p.star=W[[1]]*alpha.star[1]
logit.p=W[[1]]*alpha[1]
for(i in 2:no.alphas){
   logit.p.star=logit.p.star+W[[i]]*alpha.star[i]
   logit.p=logit.p+W[[i]]*alpha[i]
}
p.star=expit(logit.p.star)
p.star.na=p.star
p=expit(logit.p)
p.na=p
p.star.na[is.na(Y[,2:(dim(Y)[2])])]=NA
p.na[is.na(Y[,2:(dim(Y)[2])])]=NA

#Metropolis-Hastings
mh1 = sum(dbinom(Y[,-1][Z[,-1]==1],
   1,p.star.na[Z[,-1]==1],log=TRUE),na.rm=TRUE) +
   dnorm(alpha.star,mu.alpha,sqrt(sigma2.b),log=TRUE)
mh2 = sum(dbinom(Y[,-1][Z[,-1]==1],1,
   p.na[Z[,-1]==1],log=TRUE),na.rm=TRUE) +
   dnorm(alpha,mu.alpha,sqrt(sigma2.b),log=TRUE)
mh=min(exp(mh1-mh2),1)
if(mh>runif(1)){
   alpha=alpha.star
   p=p.star
   accept.alpha=accept.alpha+1
}

###
###   Sample betas
###

#Proposal
#    if(accept.beta/k < 0.3){beta.tune=max( #used for tuning only
#       beta.tune -.001, 0.0001)}
#    if(accept.beta/k >0.5){beta.tune=beta.tune+0.001}
beta.star=rnorm(no.betas,beta,beta.tune)
beta.star[c(11,15)]=0
logit.phi.star=X[[1]]*beta.star[1]
logit.phi=X[[1]]*beta[1]
```

```
for(i in 2:no.betas){
  logit.phi.star=logit.phi.star+X[[i]]*beta.star[i]
  logit.phi=logit.phi+X[[i]]*beta[i]
}
phi.star=expit(logit.phi.star)
phi.star.na=phi.star
phi=expit(logit.phi)
phi.na=phi
phi.star.na[is.na(Y[,2:(dim(Y)[2])])]=NA
phi.na[is.na(Y[,2:(dim(Y)[2])])]=NA
survive.star=calc.annual.survival(beta.star,I.)
survive=calc.annual.survival(beta,I.)
pop.star=calc.pop(N.ad1,N.sub1,pop$N.ad,pop$N.sub,fec,
  survive.star[,1],survive.star[,2],survive.star[,3])
pop=calc.pop(N.ad1,N.sub1,pop$N.ad,pop$N.sub,fec,
                survive[,1],survive[,2],survive[,3])

#Metropolis-Hastings
mh1 = sum(dbinom(Z[,-1],Z[,-J],
  phi.star.na,log=TRUE),na.rm=TRUE) +
  sum(dnorm(y,pop.star$mu,sqrt(sigma2.y),log=TRUE)) +
  sum(dnorm(beta.star,mu.beta,sqrt(sigma2.b),log=TRUE))
mh2 = sum(dbinom(Z[,-1],Z[,-J],phi.na,log=TRUE),
na.rm=TRUE) + sum(dnorm(y,pop$mu,sqrt(sigma2.y),
  log=TRUE)) + sum(dnorm(beta,mu.beta,sqrt(sigma2.b),
  log=TRUE))
mh=min(exp(mh1-mh2),1)
if(mh>runif(1)){
  beta=beta.star
  pop=pop.star
  survive=survive.star
  phi=phi.star
  accept.beta=accept.beta+1
}

###
### Save Samples
###

sigma2.y.save[k]=sigma2.y
fec.save[k,]=fec
alpha.save[k,]=alpha
beta.save[k,]=beta
N.ad.save[k,]=pop$N.ad
N.sub.save[k,]=pop$N.sub
N.save[k,]=pop$N
```

```
      beta.tune.save[k]=beta.tune
      fec.tune.save[k]=fec.tune
      if(k%%100==0){
        out=list(fec.save=fec.save,
                  alpha.save=alpha.save,
                  beta.save=beta.save,
                  N.ad.save=N.ad.save,
                  N.sub.save=N.sub.save,
                  N.save=N.save,
                  accept.beta=accept.beta,
                  accept.alpha=accept.alpha,
                  accept.fec=accept.fec,
#                  fec.tune.save=fec.tune.save,
#                  beta.tune.save=beta.tune.save
                  )
        save(out,file=name)
      }
      cat("\n")
    }
}
```