From Mark-Resight to Management: Bayesian Hierarchical Models for Endangered Bird Populations

Abby Bratt

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Sarah Converse, Chair

Beth Gardner

Scott Pearson

Program Authorized to Offer Degree:

Quantitative Ecology and Resource Management, College of the Environment

University of Washington

Abstract

From Mark-Resight to Management: Bayesian Hierarchical Models for Endangered Bird Populations

Abby Bratt

Chair of the Supervisory Committee:
Sarah Converse
School of Aquatic and Fishery Sciences, School of Environmental and Forest Sciences

Producing reliable estimates of demographic rates is critical to our understanding of wildlife population dynamics and can provide valuable information for prioritizing conservation and management efforts. Precise and unbiased estimates are challenging to obtain when monitoring data are sparse, knowledge gaps are pervasive, or model assumptions are violated. This is often the case for species of conservation concern, which may be poorly understood and difficult to monitor. Bayesian hierarchical models are particularly useful for estimating demographic rates because they separate imperfect observation processes from the underlying biological processes, especially when combined in an integrated framework that leverages multiple data sources for increased precision and parameter identifiability.

Here I present three case studies using Bayesian hierarchical models to better understand the demography of threatened birds, with particular contributions to mark-resight and integrated

population modeling. In Chapter 2, I addressed a common but poorly understood problem in mark-resight studies of open populations: partial mark loss and degradation. I present a novel approach to sampling latent states in a Markov Chain Monte Carlo framework using a backtracking algorithm, and I apply this approach in the context of a multi-event model to the Oregon Vesper Sparrow (*Pooecetes gramineus affinis*) in South Puget Sound, Washington, USA. The results from this model constitute some of the first estimates of age-specific survival and dispersal rates for this species of conservation concern. In Chapter 3, I developed a novel multi-site integrated population model (IPM) to better understand the population dynamics of Streaked Horned Larks (*Eremophila alpestris strigata*) in South Puget Sound, Washington. These estimates will inform future habitat management and a planned reintroduction effort, and the multi-site framework addresses a critical gap in modeling small populations monitored over fragmented landscapes. In Chapter 4, I developed an IPM to examine the impact of a cryptic threat, bycatch in commercial fisheries, on the population dynamics of Atlantic Yellow-nosed Albatross (*Thalassarche chlororhynchos*). Results from this model will motivate ongoing monitoring of Atlantic Yellow-nosed Albatross and seabird bycatch in the South Atlantic and inform fisheries regulation decisions. Broadly, the work I present here makes contributions to the development of complex demographic models with the goal of supporting conservation and management decisions by quantifying and reducing key uncertainties in the population dynamics of threatened species.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Writing this dissertation has been hard, but writing the acknowledgments section has been easy. In fact, throughout the last six months of my dissertation, I came back to it many times. Whenever I had writer's block, I would look to the acknowledgements I had drafted, and the words would start to flow. This is a small, but I hope meaningful, example of how the people listed below have supported me both directly and indirectly during my education. I could not have done it without them.

I have received financial support over the years from the University of Washington's Quantitative Ecology & Resource Management Program, the College of the Environment, the Washington Cooperative Fish & Wildlife Research Unit, the Northwest Climate Adaptation Science Center, and the Royal Society for the Protection of Birds.

I am so grateful to my coauthors of these chapters. In particular, I'd like to thank Steffen Oppel and Nathan Hostetter. To Steffen Oppel, who has been pushing Chapter 4 forward for over a decade. Thank you for trusting that I could contribute meaningfully to this project, and for your continued support and expertise. And to Nathan Hostetter, you're my hero. Thank you for your generosity, your brilliance, and your graciousness.

I'd also like to thank my colleagues in the Quantitative Conservation Lab and the co-op unit. To Staci Amburgey, Lisanne Petracca, Matt Farr, Hannah Sipe, Mark Sorel, Amanda Warlick, Amelia DuVall, Brielle Thompson, Liam Pendleton, Eve Hallock, Nate Redon, Kelly Mistry, Verna Blackhurst, and Sarah Romero. I have learned so much from each of you. Thanks for looking at code and drafts, and for being the sounding boards for all my best (and worst) ideas.

To my fellow QERMies, thanks for all the soup! To my cohort in particular, Martin Endress, Maria Kuruvilla, Yian Lin, and Megan Ferguson. You guys helped shoulder the challenges of that first year and even made it fun. To Tim Essington, Beth Gardner, Trevor Branch, and Erica Owens. Thanks for all you do to keep QERM the supportive professional and personal community that it is.

Thank you to my committee members, Sarah Converse, Beth Gardner, Gary Slater, Scott Pearson, and Ryan Kelly, for their mentorship. Ryan, thank you for going above and beyond as a member of my committee, for asking great questions, and for broadening my perspective on the world. Scott, thank you for your ever-thoughtful feedback and for pointing me in the right directions when I got lost. Gary, thank you for getting me out from behind the desk and trusting that a young statistician could turn into a young ecologist. Beth, you've worn several hats as a member of this committee and as the director of QERM and your contributions as both have been invaluable. Thank you for challenging me and reminding me to laugh at the same time. And to Sarah. I cannot begin to describe the impact you've had on my life as a scientist and a human being. Working with you has been an honor and a privilege.

I am grateful also for the unwavering support and understanding I have received from friends and family as near as the bedroom across the hall and as far as Norway. To my parents, Beth and Nick. Thank you for giving me space to dream. Thank you for teaching me to find joy in hard things. Thank you for making sure I know I always have a safe place to land. To Jesse, Miriam, Jens, and Maud. Time zones are generally a curse, but for me, they have been a blessing. What a joy it's been to wake up to messages from you nearly every day. I was also fortunate enough to gain some new family members over the course of this ride. To Halinka,

Dalton, Izzy, and George. Thank you for keeping me company during the darkest and lightest times of the past few years. Becoming a member of your family has been the honor of a lifetime.

And to Henry. 99% of people (real statistic) think statistics is the most boring subject on the planet. I feel so fortunate to have found a partner who is so curious about the world and who asks the most insightful questions. Thank you for seeing me. Thank you for knowing when I need a hug, a laugh, a break, or a meal. You are my best distraction. I love you so much.

DEDICATION

To my grandfathers:

Charles Christopher Bratt and Lloyd William Eichhorn

Champions of conservation, lovers of poetry, advocates for a good breakfast.

Chapter 1. INTRODUCTION

1.1 BACKGROUND

Understanding population dynamics and drivers of population trends can support the conservation of threatened species. However, estimating demographic rates and their variation over space and time is challenging for small, declining, fragmented, or otherwise difficult-to-monitor populations. Knowledge gaps may impede decision making when data are sparse. Population modelers aim to leverage available data to close these knowledge gaps and identify key uncertainties for future study.

One approach to population modeling is integrated population modeling, where multiple data sources with shared underlying parameters are combined in a joint analysis. Integrated population models (IPMs; Besbeas et al., 2002; Brooks et al., 2004; Schaub & Abadi, 2011; Zipkin & Saunders, 2018) have become popular in part because they leverage all available data to provide information about both demographic states (i.e., abundance) and rates (e.g., survival, productivity). Crucially, IPMs can improve precision (Abadi et al., 2010a; Schaub et al., 2007) and produce estimates of demographic rates that are unobservable (e.g., Oppel et al., 2022) or may otherwise be unidentifiable (Abadi et al., 2010b). Though much of the preliminary work demonstrating the utility of IPMs focused on taxa with simple life histories (e.g., herons; Besbeas et al., 2002), as available computing power increases, IPMs for species with complex life histories (e.g., apex predators; Regehr et al., 2018) are becoming more common. Like other hierarchical model types, IPMs can facilitate estimation of the effects of environmental conditions or anthropogenic stressors on demographic rates (e.g., Oppel et al., 2014), which is of useful when trying to identify causes of decline for threatened or indicator species.

To be a true integrated model, IPMs need to include a dataset relevant to the estimation of abundance and at least one additional dataset that allows for the estimation of one or more demographic parameters; commonly this is a mark-recapture or mark-resight dataset. Studies of marked populations can lend insights about a number of demographic parameters, including survival (e.g., Lebreton et al., 1992), recruitment (e.g., Tucker et al., 2023), movement (e.g., Sollmann et al., 2013), and productivity (e.g., Lahoz-Monfort et al., 2013). Often these parameters are strongly correlated with population trend and therefore it is valuable to estimate these parameters precisely and to identify their drivers to inform conservation decision-making. Mark-resight models vary widely in complexity, from Cormack-Jolly-Seber models (Cormack, 1964; Jolly, 1965; Seber, 1965), to multi-state models (Nichols & Kendall, 1995), to multi-event models (Pradel, 2005) and can therefore accommodate many sampling situations, provided that model assumptions are met.

While the development of complex models and integrated models can help resolve some uncertainty in our understanding of demography, there is no substitute for a well-designed monitoring program. Underpinning all models are assumptions about the underlying observation and biological processes and when those assumptions are violated it may render results invalid. Some assumptions may matter relatively little (e.g., independence of datasets in integrated population models; Abadi et al., 2010a), where others can matter more (e.g., mark loss in mark-resight models; Chapter 2). Given that we rely on long-term monitoring programs to inform conservation decisions for endangered wildlife (Nichols & Williams, 2006), great care should be taken at the outset to design a monitoring program that has the power to produce accurate and precise estimates of demographic rates, and that are more likely to detect changes in population

trends. Similarly, monitoring programs should be regularly reevaluated to ensure they are being implemented correctly and functioning as intended (Lindenmayer & Likens, 2009).

1.2 Research Objectives

The objectives of my research were twofold, including (1) advancing Bayesian hierarchical modeling of mark-resight data in the context of integrated population models, and (2) estimating vital rates for three poorly understood species of conservation concern. To this end, I developed a novel multi-state model for mark-resight data in the presence of mark loss for Oregon Vesper Sparrow (*Pooecetes gramineus affinis*) in Washington State, and IPMs for Streaked Horned Larks (*Eremophila alpestris strigata*) and Atlantic Yellow-nosed Albatross (*Thalassarche chlororhynchos*) in Washington State and on Gough Island in the South Atlantic.

The methodological advancements made here were motivated by challenges presented in the available data. For Oregon Vesper Sparrow and Streaked Horned Larks, partial mark loss is pervasive and has hindered our understanding of survival and dispersal rates for these species. Thus, in Chapter 2 I developed a novel model and approach for sampling latent states for multi-state models of mark-resight data in the presence of partial mark loss and degradation. Streaked Horned Larks are intensively monitored at numerous sites, but the region-wide population dynamics are not well understood. Thus, in Chapter 3 I built a multi-site IPM, using mark-resight data to inform movement and survival across a fragmented landscape. Atlantic Yellow-nosed Albatross are vulnerable to cryptic threats but are difficult to monitor and only observable during some life history stages. Therefore, in Chapter 4 I built an IPM around a multi-event model of mark-resight data with several unobservable states. Each challenge that I have addressed here is not unique to the case-study species. Consequently, the methodology I present is applicable to

many endangered species which are monitored with through marking and resighting, over fragmented landscapes, or only during portions of their life-history.

1.3 Broader Impacts

Collectively, the developed in these studies contribute new approaches for developing models of complex ecological processes and the specific case studies make contributions that will inform species conservation decisions. Ecologically, the Oregon Vesper Sparrow is a species of great conservation interest throughout the Pacific Northwest, including in Washington State, where it is listed as endangered. I present some of the first robust estimates of age-specific survival and dispersal probabilities for this subspecies, which will inform future population modeling efforts and influence conservation action. The Streaked Horned Lark is state and federally listed and is intensively monitored throughout South Puget Sound, where the model and demographic estimates I produced will be used to inform a reintroduction effort. Atlantic Yellow-nosed Albatross is endangered per the IUCN and is vulnerable to cryptic threats such as environmental change and anthropogenic stressors in the South Atlantic. Bycatch in commercial fisheries is a known threat to seabirds but the degree to which it impacts population dynamics of this species is not well understood. My work on this species identifies knowledge gaps for future study. Methodologically, I present novel model frameworks that facilitate robust estimation of vital rates in the face of common challenges with the integration of mark-resight data: mark loss or degradation, dispersal over fragmented landscapes, and multiple unobservable states.

1.4 REFERENCES

Abadi, F., Gimenez, O., Arlettaz, R., & Schaub, M. (2010). An assessment of integrated population models: Bias, accuracy, and violation of the assumption of independence. *Ecology*, *91*(1), 7–14. JSTOR.

Abadi, F., Gimenez, O., Ullrich, B., Arlettaz, R., & Schaub, M. (2010). Estimation of immigration rate using integrated population models. *Journal of Applied Ecology*, *47*(2), 393–400.

Besbeas, P., Freeman, S. N., Morgan, B. J. T., & Catchpole, E. A. (2002). Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. *Biometrics*, *58*(3), 540–547. https://doi.org/10.1111/j.0006-341X.2002.00540.x

Brooks, S. P., King, R., & Morgan, B. J. T. (2004). A Bayesian approach to combining animal abundance and demographic data. *Animal Biodiversity and Conservation*, 16.

Cormack, R. M. (1964). Estimates of survival from the sighting of marked animals. *Biometrika*, *51*(3/4), 429–438. https://doi.org/10.2307/2334149

Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika*, *52*(1/2), 225–247. https://doi.org/10.2307/2333826

Lahoz-Monfort, J. J., Harris, M. P., Morgan, B. J. T., Freeman, S. N., & Wanless, S. (2013). Exploring the consequences of reducing survey effort for detecting individual and temporal variability in survival. *Journal of Applied Ecology*, *51*(2), 534–543. https://doi.org/10.1111/1365-2664.12214

Lebreton, J.-D., Burnham, K. P., Clobert, J., & Anderson, D. R. (1992). Modeling survival and testing biological hypotheses using marked animals: A unified approach with case studies. *Ecological Monographs*, *62*(1), 67–118. https://doi.org/10.2307/2937171

Lindenmayer, D. B., & Likens, G. E. (2009). Adaptive monitoring: A new paradigm for long-term research and monitoring. *Trends in Ecology & Evolution*, *24*(9), 482–486. https://doi.org/10.1016/j.tree.2009.03.005

Nichols, J. D., & Kendall, W. L. (1995). The use of multi-state capture-recapture models to address questions in evolutionary ecology. *Journal of Applied Statistics*, *22*(5–6), 835–846. https://doi.org/10.1080/02664769524658

Nichols, J., & Williams, B. (2006). Monitoring for conservation. *Trends in Ecology & Evolution*, *21*(12), 668–673. https://doi.org/10.1016/j.tree.2006.08.007

Oppel, S., Clark, B. L., Risi, M. M., Horswill, C., Converse, S. J., Jones, C. W., Osborne, A. M., Stevens, K., Perold, V., Bond, A. L., Wanless, R. M., Cuthbert, R., Cooper, J., & Ryan, P. G. (2022). Cryptic population decrease due to invasive species predation in a long-lived seabird supports need for eradication. *Journal of Applied Ecology*. https://doi.org/10.1111/1365-2664.14218

Oppel, S., Hilton, G., Ratcliffe, N., Fenton, C., Daley, J., Gray, G., Vickery, J., & Gibbons, D. (2014). Assessing population viability while accounting for demographic and environmental uncertainty. *Ecology*, *95*(7), 1809–1818.

Pradel, R. (2005). Multievent: An extension of multistate capture-recapture models to uncertain states. *Biometrics*, *61*(2), 442–447.

Regehr, E. V., Hostetter, N. J., Wilson, R. R., Rode, K. D., Martin, M. St., & Converse, S. J. (2018). Integrated population modeling provides the first empirical estimates of vital rates

and abundance for polar bears in the chukchi sea. *Scientific Reports*, *8*(1), 16780. https://doi.org/10.1038/s41598-018-34824-7

Schaub, M., & Abadi, F. (2011). Integrated population models: A novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology*, *152*(S1), 227–237. https://doi.org/10.1007/s10336-010-0632-7

Schaub, M., Gimenez, O., Sierro, A., & Arlettaz, R. (2007). Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. *Conservation Biology*, *21*(4), 945–955. https://doi.org/10.1111/j.1523-1739.2007.00743.x

Seber, G. A. F. (1965). A note on the multiple-recapture census. *Biometrika*, *52*(1/2), 249–259. https://doi.org/10.2307/2333827

Sollmann, R., Gardner, B., Parsons, A. W., Stocking, J. J., McClintock, B. T., Simons, T. R., Pollock, K. H., & O'Connell, A. F. (2013). A spatial mark–resight model augmented with telemetry data. *Ecology*, *94*(3), 553–559. https://doi.org/10.1890/12-1256.1

Tucker, A. M., McGowan, C. P., Nuse, B. L., Lyons, J. E., Moore, C. T., Smith, D. R., Sweka, J. A., Anstead, K. A., DeRose-Wilson, A., & Clark, N. A. (2023). Estimating recruitment rate and population dynamics at a migratory stopover site using an integrated population model. *Ecosphere*, *14*(2), e4439. https://doi.org/10.1002/ecs2.4439

Zipkin, E. F., & Saunders, S. P. (2018). Synthesizing multiple data types for biological conservation using integrated population models. *Biological Conservation*, *217*, 240–250. https://doi.org/10.1016/j.biocon.2017.10.017

Chapter 2. AN UNBIASED SURVIVAL ESTIMATOR BASED ON MARK-RESIGHT DATA IN THE PRESENCE OF MARK DEGRADATION

Publication history: This study was co-authored with Nathan J. Hostetter, Gary L. Slater, and Sarah J. Converse. At the time this dissertation was published, this chapter was not in review with a journal.

Abstract: Mark-resight methods are a popular alternative to mark-recapture methods for estimation of demographic parameters in animal populations because individuals do not need to be physically recaptured. Investigators using mark-resight methods typically deploy uniquely identifiable marks to a subset of a population of interest, and subsequently collect resighting data non-invasively. However, it is common for artificial marks to degrade over time (e.g., individual colored bands may fall off), resulting in observations that may not be attributable to unique individuals. Existing approaches to mitigate this problem include omitting these data or assigning uncertain observations to individuals in a post-hoc manner, potentially introducing bias in estimates of survival and leading to underestimation of uncertainty. To address the challenge of mark degradation, we present a novel multi-event model linking observed marks to individuals. For the individuals with missing bands, individual encounter histories, individual mark status, and individual state are unknown. We use a backtracking algorithm to construct all individual potential matches of band-loss individuals with known individuals across sampling occasions, which forms the basis of a custom-designed Markov Chain Monte Carlo sampler. We compare our approach with the data omission approach using a simulation and we also present a case study from a population of Oregon Vesper Sparrow (*Pooecetes gramineus affinis*) in

Washington, USA. We found our approach produced unbiased results, while the data omission approach resulted in negatively biased survival estimates. However, high computational costs are a barrier to the application of our model to marked populations that are large or have a high rate of mark degradation. Improved algorithms for identifying valid matches offer promise for expanding the applicability of our approach. Partial mark loss or degradation is a frequent and likely underreported source of bias in mark-resight studies, and we provide recommendations for researchers employing study designs to minimize both the rate of mark degradation and uncertainty in identity when mark degradation does occur.
Keywords: Backtracking, capture-recapture, mark degradation, mark loss, mark-resight, multi-event model, partial identification, Oregon Vesper Sparrow, *Pooecetes gramineus affinis*

2.1 INTRODUCTION

The estimation of demographic rates is frequently critical to the conservation and management of threatened species (e.g., Williams et al., 2002). Survival, recruitment, and dispersal rates are commonly estimated from studies of marked populations (Baillie & Schaub, 2009). In mark-resight studies, both artificial marks (e.g., dye marks, color bands, leg flags, patagial tags) and natural marks (e.g., stripe patterns, permanent scars, genetic markers) can be used to identify re-encountered individuals via sight, camera trap, or DNA barcoding in the absence of physical recapture (Silvy et al., 2012).

However, without physical recaptures, there is an increased risk of being unable to accurately identify individuals due to either imperfect observation or the loss or degradation of marks, thus violating key assumptions of traditional mark-resight models. Traditional mark-recapture and mark-resight models assume that no marks are lost, that individuals are correctly classified as marked, and that marked individuals are correctly and completely identified

(Lindberg, 2012; Williams et al., 2002). Violations of some of these marking assumptions have been investigated and solutions have been proposed. For example, in the case of complete mark loss, a common solution is to double-mark a subset of the marked population to estimate a rate of complete mark loss and correct for bias in survival estimates (e.g., Cowen & Schwarz, 2006; Diefenbach & Alt, 1998).

In cases when individuals are misidentified, two types of error are possible: 1) the misidentification may confuse one individual within the dataset for another (e.g., Tucker et al., 2019); or 2) the misidentification may produce a new "ghost" individual in the dataset (i.e., a false individual; Link et al., 2010; Lukacs & Burnham, 2005; Wright et al., 2009; Yoshizaki et al., 2009). In contrast, when partial identifications occur, observations will allow only for the identification of a set of previously marked individuals that may correspond to the observed partial identity, e.g., a bird observed with only a blue color band on the left leg could be any of the set of birds that received a blue color band on the left leg.

Partial identifications can result from either observation error (i.e., only part of the mark was visible to an observer) or process error (i.e., only part of the mark remains or is identifiable). We consider the latter, which we term mark degradation. Mark degradation may occur from researcher-deployed marks degrading over time, where examples include the loss of color bands on passerines (e.g., Lovell, 1948), degradation of alphanumeric leg flags on shorebirds (e.g., Roche et al., 2014), movement of fluorescent elastomer tags within the body cavity of amphibians (e.g., Brannelly et al., 2013), the fading of lip tattoos on polar bears (e.g., Larsen, 1971), or the stretching of brands on sea lions (e.g., Schwarz & Stobo, 1999).

The result of mark degradation/partial loss is that some observations are not attributable to unique individuals. By reducing the observer's ability to uniquely identify individuals, mark

degradation complicates the way in which the resight process must be modeled. Existing approaches include 1) discarding observations, which artificially deflates both detection and, crucially, survival probabilities because the mark degradation becomes indistinguishable from an apparent mortality; or 2) assignment of marks to individuals post hoc without accounting for uncertainty in the assignment process, which results in potentially serious biases if marks are attributed to the wrong individuals. This phenomenon has been well-studied in closed populations (e.g., Augustine et al., 2019, 2020; Jiménez et al., 2019, 2021; McClintock et al., 2014) but mark degradation has received little attention in the literature in open populations, where the solutions used for closed populations are not applicable due to additional uncertainty about the true state (e.g., alive or dead) of each individual on each occasion when they are not perfectly detected.

Here we propose a solution for unbiased estimation of survival and dispersal rates in the case of uncertain individual identities due to mark degradation. We developed a novel hierarchical model linking observed marks to a discrete set of true individuals. This link could be used to develop a relatively simple extension to a Cormack-Jolly-Seber model (i.e., a mark-resight model where the only states are live and dead; Cormack, 1964; Jolly, 1965; Seber, 1965) in which mark uncertainty is accounted for. Here, we present a more complex application to multistate mark-resight models, where the true states include residence in different spatial units in addition to a dead state. In doing so, we leverage the information contained in the spatial location of observations to inform individual identity. In our model representation, true individual encounter histories, individual mark status, and individual state are latent variables. This latent hierarchy allows for process error as marks degrade. We use a backtracking algorithm (e.g., Rossi et al., 2006) to construct all valid matchings of marks to individuals. This allows us

to reconstruct all possible multi-state encounter histories. In contrast to conventional mark-resight models, this allows for marks to be only partially retained while assuming that the retained mark components are completely and correctly observed. We then develop a Markov-Chain Monte Carlo (MCMC) approach to sampling these possible histories, in addition to the underlying states.

We describe the model and develop a simulation study with a territorial passerine in mind, though the model is generalizable to alternative mark types and taxa. In mark-resight studies, passerines are commonly given a permanent mark that is not field-readable (e.g., a USGS metal band), along with field-readable auxiliary marks (e.g., a combination of colored leg bands; Calvo & Furness, 1992; Marion & Shamis, 1977). In these situations, individuals are recognizable as marked even if they lose all auxiliary marks, but they may not be individually identifiable.

We explore the properties of the estimates from this model relative to estimates obtained by omitting degraded marks. We demonstrate that this model produces unbiased estimates of survival while omitting observations of degraded marks leads to underestimation of survival. We then apply the model to a small, threatened population of Oregon Vesper Sparrows (OVSP; *Pooecetes gramineus affinis*) in Washington, USA. Our proposed approach is well suited to studies of threatened populations because it leverages all available data, which are often expensive to obtain, and without which estimates of demographic parameters may be biased or highly uncertain. We make several recommendations for practitioners to consider when designing mark-resight studies with the goals of minimizing the rate of mark degradation and minimizing uncertainty in identity when mark degradation does occur.

2.2 Materials and Methods

2.2.1 *Overview*

We present a multi-state model for estimating apparent survival and movement in a population with mark degradation. Our proposed model, while generalizable, functions best for species with some level of site fidelity, because location is used to inform individual identity. We developed the model for application to a passerine, though it is applicable to many other taxa that receive degradable marks (e.g., colored bands, leg flags, fluorescent elastomer tags, tattoos, brands, paint). We assume the population has been monitored at several sites across its breeding range. A subset of this population is marked at each site through the application of a unique combination of colored bands and one metal band on the legs. In subsequent years, researchers perform resight surveys at the same study locations to construct multi-state encounter histories for each individual (i) over time (t). We begin by describing the model using a state-space multi-state mark-resight model (Kery & Schaub, 2011; Arnason, 1973), before describing how we expand it to handle mark degradation.

2.2.2 *Example*

Consider three marked birds with the following color band combinations: Blue Green (BG), Red Green (RG), and Blue Purple (BP). Actual band combinations may be more complex than this, with multiple color bands per leg, but this example is illustrative. Each individual is also given a metal band that has a unique number imprinted on the tag but requires a physical recapture to read. Birds BG, RG, and BP were marked at sites 1, 2, and 3, respectively and were then monitored over $T = 4$ occasions at $S = 3$ sites. The latent state of each bird on each occasion

($Z_{i,t}$) is then the site (s) at which each bird occurs, where $s \in \{1, \dots, S+1\}$ and ($S\ +\ 1$) represents the dead state. In this example, the true states are

$$\boldsymbol{Z} = \begin{matrix} BG \\ RG \\ BP \end{matrix} \begin{pmatrix} 1 & 1 & 1 & 2 \\ 2 & 1 & 2 & 4 \\ 3 & 3 & 3 & 3 \end{pmatrix}.$$

That is, bird BG remained at site 1 until moving to site 2 before the last occasion, bird RG moved from site 2 to site 1 and then back to site 2 before dying, and bird BP remained at site 3 over the course of the study. This true but latent $\boldsymbol{Z}$ matrix is modeled as arising from survival and movement processes, where the state of individual i at time t, $Z_{i,t}$, is a function of the state of the individual at the previous time period, i.e.,

$$Z_{i,t}|Z_{i,t-1} \sim Categorical\left(\boldsymbol{\Psi}_{Z_{i,t-1}}\right),$$

where $\boldsymbol{\Psi}_{Z_{i,t-1}}$ is a vector composed of apparent survival (ϕ) and site-specific movement probabilities ($\boldsymbol{\psi}$) for an individual in state $Z_{i,t-1}$. An individual's transition probability is equal to the product of apparent survival, ϕ, and the probability of moving from site s to site r, $\psi_{s,r}$ where, for $s\ =\ r$, an individual remains at the same site. Individuals transition to the absorbing dead state with probability $1 - \phi$.

Given the true $\boldsymbol{Z}$ shown above, one possible encounter history matrix is

$$\boldsymbol{Y} = \begin{matrix} BG \\ RG \\ BP \end{matrix} \begin{pmatrix} 1 & 1 & 1 & 4 \\ 2 & 1 & 2 & 4 \\ 3 & 3 & 3 & 4 \end{pmatrix},$$

where $\boldsymbol{Y}_i$ represents the multi-state encounter history for individual i. Similarly, $Y_{i,t} = s$, where $s \in \{1, \dots, S+1\}$ and $(S+1)$ indicates that the individual was unobserved. This is just one of the many possible encounter history matrices. We model the observations conditional on the true state,

$$Y_{i,t}|Z_{i,t} \sim Categorical\left(\mathbf{\Pi}_{Z_{i,t}}\right),$$

where $\mathbf{\Pi}_{Z_{i,t}}$ represents the multi-state detection probability. Individuals are detected at their location with probability p and not detected with probability $1 - p$. That is, living individuals cannot be detected at any site besides the one at which they are located, and dead individuals cannot be detected. Assuming no mark loss or degradation, we observe the encounter history matrix $\boldsymbol{Y}$ directly and can use these data to estimate the parameters using a standard multi-state mark-resight model.

When auxiliary marks degrade, however, the true but latent mark state ($M_{i,t}$) of each bird on each occasion may only be partially observed. Consider an example mark loss process,

$$\boldsymbol{M} = \begin{matrix} BG \\ RG \\ BP \end{matrix} \begin{pmatrix} BG & B & B & B \\ RG & RG & G & G \\ BP & B & B & B \end{pmatrix},$$

where $M_{i,t} = k$ is an individual's mark at time *t*, where $k \in \boldsymbol{K}$, the set of all observable marks. Bird BG lost its green band after occasion 1, bird RG lost its red band after occasion 2, and bird BP lost its yellow band after occasion 1. Here, $\boldsymbol{K} = \{BG, RG, BP, G, P, B, R, A\}$ where A corresponds to total loss of auxiliary, but not permanent, bands. We assume that the metal band is never lost such that marked individuals may always be distinguished from unmarked individuals.

When mark loss occurs, the individual encounter history matrix, $\boldsymbol{Y}$, is latent because observed marks may not always be attributable to unique individuals. Since multiple individuals may now have the same mark, we instead record $Y^{obs}_{k,s,t}$, counts of mark k at site s at time t. $\boldsymbol{Y}^{obs}$ is then

$$\boldsymbol{Y}^{obs}_{\boldsymbol{,1,}} = \begin{matrix} BG \\ RG \\ BP \\ G \\ P \\ B \\ R \\ A \end{matrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \boldsymbol{Y}^{obs}_{\boldsymbol{,2,}} = \begin{matrix} BG \\ RG \\ BP \\ G \\ P \\ B \\ R \\ A \end{matrix}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \boldsymbol{Y}^{obs}_{\boldsymbol{,3,}} = \begin{matrix} BG \\ RG \\ BP \\ G \\ P \\ B \\ R \\ A \end{matrix}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

That is, at $s = 1$, mark BG was observed on occasion 1, mark RG on occasion 2, and mark B on occasions 2 and 3. At $s = 2$, mark RG was observed on occasion 1, and mark G on occasion 3. Finally, at $s = 3$, mark BP was observed on occasion 1, and mark B on occasions 2 and 3. This is represented visually in Figure 2.1. Though it does not occur in this simple example, it is possible to observe more than one count of mark k at site s on occasion t.

2.2.3 *Model Description*

Because of the uncertainty in matching marks to individuals, each hierarchical process indexed by individual identity is partially latent. As with most standard multi-state models, the true states of individuals, represented by matrix ***Z,*** are partially latent. Here we have two additional latent processes, the first being the individual encounter histories, represented by matrix $\boldsymbol{Y}$, and the second being the mark status, $\boldsymbol{M}$, where $M_{i,t}$ is conditional on the mark with which individual i was observed at $t-1$. That is,

$$M_{i,t}|M_{i,t-1} \sim Categorical\left(\boldsymbol{\theta}_{M_{i,t-1}}\right),$$

where the likelihood of individual i having mark d at time t, conditional on having mark c at time $t-1$, assuming $c, d \in \boldsymbol{K}$, follows a categorical distribution where the probability of transitioning between marks c and d is described by $\theta_{c,d}$ (see Generating the Mark Transition Matrix).

The observed counts of mark k at site s at time t ($Y_{k,s,t}^{obs}$) are related to the latent encounter histories through the introduction of latent array, $\boldsymbol{Y}^{multi}$, which links counts of marks to individual encounter histories. Specifically,

$$Y_{i,k,s,t}^{multi} = I\big(Y_{i,t} = s\big)I\big(M_{i,t} = k\big),$$

where I represents the indicator function. That is, $Y_{i,k,s,t}^{multi} = 1$ when individual i had mark k at site s at time t and 0 otherwise. The observed data are then

$$Y_{k,s,t}^{obs} = \sum_{i=1}^{N} Y_{i,k,s,t}^{multi},$$

i.e., our observations are a sum over individuals of this latent array. The model parameters, constants, and latent states are given in Table 2.1.

This novel model links observations of partially degraded marks, $\boldsymbol{Y}^{obs}$, to the latent matrices $\boldsymbol{M}$ and $\boldsymbol{Y}$ through the introduction of $\boldsymbol{Y}^{multi}$, where encounter history matrix $\boldsymbol{Y}$ is conditional on state matrix $\boldsymbol{Z}$ just as it is in a standard multi-state model. We can use MCMC to sample each latent state and estimate the underlying survival, movement, and detection parameters. We can also estimate a new parameter, β, to model the mark degradation process. In our case, this parameter represents the probability of auxiliary mark degradation, as described below.

2.2.4 *Generating the Mark Transition Matrix*

Given the set of marks deployed, we can generate the mark transition matrix, $\boldsymbol{\theta}$. We first generate all possible subsets of the original marks. Using the example above, given the original marks $\{BG, RG, BY\}$, the set of all observable marks is $\boldsymbol{K} = \{BG, RG, BY, G, Y, B, R, A\}$. The transition matrix is thus of dimension $|\boldsymbol{K}|$x$|\boldsymbol{K}|$. We then generate a matrix ($\boldsymbol{M}$) representing the

number of ways to transition from mark c to mark d, where $c, d \in \boldsymbol{K}$, and mark d is at least as degraded as mark c. For example, there is one way to transition from mark $c = BG$ to mark $d = B$ (i.e., by losing the green G band; thus $M_{c,d} = 1$). In a more complex example, an individual with mark $c = BB$ has two ways to transition to mark $d = B$ (i.e., by losing either of its blue B bands; thus $M_{c,d} = 2$). Let β be the probability of losing a single band, b_c be the number of bands associated with mark c, and b_d be the number of bands associated with mark d. If $c \neq d$ and there is at least one way to transition from mark c to mark d (i.e., $M_{c,d} \geq 1$) or $c = d$ (i.e., $M_{c,d} = 1$), the probability of transitioning from mark c to mark d is

$$\theta_{c,d} = M_{c,d} \cdot \beta^{b_c - b_d} \cdot (1-\beta)^{b_d},$$

and zero otherwise. For example, using the same example population as above, the transition matrix is

$$\boldsymbol{\theta} = \begin{array}{c} BG \\ RG \\ BY \\ G \\ Y \\ B \\ R \\ A \end{array} \begin{pmatrix} (1-\beta)^2 & 0 & 0 & \beta(1-\beta) & 0 & \beta(1-\beta) & 0 & \beta^2 \\ 0 & (1-\beta)^2 & 0 & \beta(1-\beta) & 0 & 0 & \beta(1-\beta) & \beta^2 \\ 0 & 0 & (1-\beta)^2 & 0 & \beta(1-\beta) & \beta(1-\beta) & 0 & \beta^2 \\ 0 & 0 & 0 & 1-\beta & 0 & 0 & 0 & \beta \\ 0 & 0 & 0 & 0 & 1-\beta & 0 & 0 & \beta \\ 0 & 0 & 0 & 0 & 0 & 1-\beta & 0 & \beta \\ 0 & 0 & 0 & 0 & 0 & 0 & 1-\beta & \beta \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

2.2.5 *Backtracking Algorithm*

In general, backtracking algorithms are used to enumerate all solutions to a constraint satisfaction problem. Here, we use a backtracking algorithm to enumerate all possible encounter histories $\boldsymbol{Y}$, conditional on the marks observed, $\boldsymbol{Y}^{obs}$. We represent the problem as a graph that maps marks observed at a given time to the individuals whose original marks could have degraded to produce the observed marks (i.e., $M_{c,d} \geq 1$ for original mark c and observed mark d).

The algorithm begins by creating an empty search tree which represents the set of all observed marks prior to their assignment to individuals. We then iterate over any unassigned observations in ascending order of the number of individuals to which an observation could be assigned (e.g., if we observe the complete mark BG, that uniquely identifies individual BG, but if we observe *B*, any bird that received a blue band on the correct leg could be eligible, so we would assign BG first, and B second). We make assignments of individuals to marks as the algorithm proceeds, incrementing one assignment at a time and assuming that individuals can be detected at most once per occasion. Each assignment creates a new node on the search tree, which forms the root of a unique branch. Each branch is explored recursively until it ultimately produces a terminal node that either represents a valid solution (i.e., all observed marks are mapped successfully to individuals), or fails to produce a solution (i.e., at least one observed mark has no valid individual to which it can be assigned, conditional on the other assignments of marks to individuals). Making the more constrained assignments first minimizes the computational time spent exploring branches of the tree that ultimately produce invalid solutions. For example, running the backtracking algorithm on the simple example described above for each occasion $t \in \{1,2,3\}$ produces the trees in Figure 2.2, respectively, where the terminal nodes indicate valid solutions. No individuals were observed at $t = 4$ so the backtracking algorithm is not run on this occasion.

The results of the backtracking algorithm run on each occasion $t \in 1, \dots, T$ are used to reconstruct all possible encounter history matrices, $\boldsymbol{Y_1}, \dots, \boldsymbol{Y_n}$. We do this by creating a graph ($\boldsymbol{G}$), where each node is a valid terminal node produced by the backtracking algorithm run on a single occasion t for all $t \in 1, \dots, T$. That is, each node of graph $\boldsymbol{G}$ represents a valid matching of all observed marks on a given occasion to marked individuals. Each valid matching from

occasion t was connected to each valid matching from occasion $t+1$, for all $t \in 1, \ldots, T-1$. For example, the results of the backtracking algorithm run on the example population (Figure 2.2), produces the graph shown in Figure 2.3. Each unique path through $\boldsymbol{G}$ therefore represents a unique encounter history $\boldsymbol{Y}$. However, it is possible at this point that a subset of these encounter histories $\boldsymbol{Y}$ are internally inconsistent. An internal inconsistency would occur when an incomplete mark was mapped to an individual on an occasion, but that individual was resighted with its complete mark on a subsequent occasion. This violates the assumption that mark degradation is degenerative (i.e., once marks are lost they cannot be regained), and we therefore check all paths through $\boldsymbol{G}$ for internal consistency. The valid encounter history matrices, $\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}$, were the n unique paths through $\boldsymbol{G}$ that produced internally consistent matches of observed marks to individuals.

2.2.6 *Initialization and Sampling*

We fit this model in a Bayesian framework using a custom MCMC sampler. A custom MCMC sampler was necessary to increase sampling efficiency for the latent encounter history matrix, $\boldsymbol{Y}$, because the mark transition matrix $\boldsymbol{\theta}$ imposed constraints on which individuals can be associated with which marks. Thus, generating valid proposals for $\boldsymbol{Y}$ given $\boldsymbol{Y}^{obs}$ was technically challenging. We used the backtracking algorithm described above to generate all valid encounter histories Y. The MCMC sampler took the set of all possible encounter history matrices, $\{\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}\}$, and performed a block update of the latent $\boldsymbol{Y}$, conditional on $\boldsymbol{Z}$, using a Metropolis-Hastings step.

2.2.7 *Confusion Index*

We introduce the concept of the "confusion index" for mark degradation. We define the confusion index as the average number of individuals to which each observed mark could be assigned. In a marked population with no mark degradation, the confusion index is 1 (i.e., every observation uniquely identifies an individual). If batch marks (i.e., multiple individuals receiving the same mark) are used, the confusion index is equal to the number of individuals assigned a batch mark, averaged across all batch marks used (i.e., every observation maps to a discrete set of individuals). In uniquely marked populations that experience mark degradation, the confusion index, C, is between 1 and N, where N is the number of individuals that have been marked (i.e., every observation could uniquely identify an individual, or in the case of complete auxiliary mark loss every observation could map to every individual, or somewhere in between). The confusion index is therefore a relative measure of the complexity of a mark-resight dataset with respect to mark degradation.

2.2.8 *Simulation Study*

We simulated two approaches for using mark-resight data to estimate survival in an open population with mark degradation: 1) omitting observations of degraded marks; and 2) the novel model presented here. We simulated mark-resight data under a range of demographic parameter values and observation scenarios and estimated apparent survival (ϕ), site-specific dispersal (ψ), and detection (p) using a standard multistate mark-resight model in case (1) and the model presented here in case (2). In case (2) we were also able to estimate the annual rate of mark loss (β).

For N = 50 individuals monitored at $S = 3$ sites for $T = 5$ years, we tested $\phi \in \{0.4, 0.8\}$, $p \in \{0.3, 0.7\}$ to represent low and high survival and detection, respectively. We

varied $\beta \in \{0, 0.01, 0.02 \ldots, 0.1\}$. Movement probabilities were simulated and modeled as a function of site-specific attractiveness and distance ($\boldsymbol{D}$) between sites. Simpler movement models could be substituted here, but the model presented here is identical to the model used in our case study, below, in which we were interested in dispersal rates relative to site spacing and quality. If s represents an individual's site at time $t-1$ and r is an individual's site at time t, then

$$\psi_{s,r} = \begin{cases} \left(logit^{-1}\left(\mu^{\psi} + \epsilon_s^{\psi}\right)\right) \text{if } s = r \\ \left(1 - logit^{-1}\left(\mu^{\psi} + \epsilon_s^{\psi}\right)\right) \cdot \dfrac{\exp\left(\alpha^{D} \cdot D_{s,r} + \epsilon_r^{\psi}\right)}{\sum_{s \neq r} \exp\left(\alpha^{D} \cdot D_{s,r} + \epsilon_r^{\psi}\right)} \text{if } s \neq r \end{cases}.$$

That is, an individual remained at their current site, s, with probability modeled as a function of mean site fidelity, μ^{ψ}, and random site-specific attractiveness, ϵ_s^{ψ}. An individual moved to a different site with probability equal to the product of the complement of the probability of staying at site s, and conditional on leaving site s, the probability of moving to site r modeled as a function of the distance between the two sites, $D_{s,r}$, and the attractiveness of site r. To simulate the data, we assumed $\mu^{\psi} = 2$, which corresponds to a site fidelity of approximately 0.8 on the probability scale. The random effects, $\boldsymbol{\epsilon^{\psi}}$, were normally distributed with mean 0 and standard deviation, $\sigma^{\psi} = 1$. We assumed that the distance, D, between sites 1 and 2 was 1 unit, between sites 1 and 3 was 2 units, and between sites 2 and 3 was 3 units. Finally, we assumed that $\alpha^{D} = -1$, which represented a moderately strong reduction in dispersal relative to increasing distance between sites.

In addition, we compared two mark-deployment schemes: 1) random mark deployment; and 2) sequential mark deployment. In random mark deployment, the marks deployed are randomly generated. In sequential mark deployment, researchers deploy similar marks throughout a site and/or season (e.g., all nestlings marked in the same site-year share at least 1

band). With sequential mark deployment, individuals that lose auxiliary marks are more likely to be confusable (e.g., both individuals BG and BP lose their second band, and therefore mark B does not uniquely identify an individual). In contrast, using random mark deployment, individuals could lose auxiliary marks and remain uniquely identifiable with higher probability. That is, data generated from a random mark deployment scheme is more likely to have a lower confusion index than data generated from a sequential mark deployment scheme, under conditions when fewer than the total possible marks are deployed.

For each combination of these parameters and deployment schemes, we simulated 100 datasets. We evaluated model performance based on relative bias (RB; calculated as $\frac{\hat{\theta}-\theta}{\theta}$, where $\hat{\theta}$ is the parameter value as estimated by the model, and θ is the true parameter value as used to simulate that data) and root mean square error (RMSE; calculated as $\sqrt{E((\hat{\theta}-\theta)^2)}$, where E represents the expected value).

2.2.9 *Case Study*

The OVSP is an endangered subspecies of Vesper Sparrow in Washington, USA (Altman, 2011). Like many small passerines, OVSP have high reproductive potential, early maturity, and short lifespans. OVSP are territorial and seasonally monogamous. Studies of other subspecies in the midwestern United States have shown that breeding site fidelity is high for adults (Best & Rodenhouse, 1984). However, there have been few recorded returns of marked nestlings to their natal site or elsewhere (Altman et al., 2020) which could be indicative of low site fidelity and/or low survival of hatch-year birds (Weatherhead & Forbes, 1994). Robust survival estimates for both adult (AHY) and hatch-year (HY) birds are lacking.

We analyzed mark-resight data of an intensively monitored population of OVSP on Joint Base Lewis-McChord, Washington, a US Department of Defense facility. We ran analyses both with and without observations of marks that did not uniquely identify individuals. Between 2016 and 2020, n = 134 individuals were marked across 5 sites. Of these, 78 were nestlings (L) hand-captured in the nest, while 14 were hatch-year (HY) and 42 were after-hatch year (AHY) individuals that were captured with mist-nets. Individuals were marked with between 0 and 3 colored acrylic bands (Avinet Research Supplies, 2.8 mm interior diameter) in one of 10 distinct colors as well as an aluminum USGS band. Individuals marked with 0 colored bands would be in mark state A; that is, lacking any uniquely identifying auxiliary marks but identifiable as marked. Two individuals received 0 colored bands, 1 individual received 1 colored band, and the remaining 131 individuals received 3 colored bands. Marks were deployed in an order akin to sequential deployment, as opposed to random deployment.

In the same period, there were 62 resights of individuals, including three observations of individuals that had experienced mark degradation and two sets of two individuals that received the same band combination and thus were not uniquely identifiable. Resights were obtained opportunistically during territory mapping or nest monitoring surveys at five sites monitored throughout the breeding season Figure 2.4. Resighting effort was not constant across sites and years, with two sites – Lower Weir prairie and Upper Weir prairie – being the most consistently and intensively monitored; other sites were usually visited at least three times during each breeding season.

We fit the model structure described for the simulation study to this dataset, with two small modifications. We modeled survival probabilities, $\boldsymbol{\phi}$, and the parameters controlling site-specific dispersal probabilities, $\boldsymbol{\mu^{\psi}}$, $\boldsymbol{\epsilon^{\psi}}$, and $\boldsymbol{\alpha^{D}}$, as age-specific, with ages L, HY, and AHY. We

also estimated three detection probabilities to correct for the variable monitoring effort. Annual site-specific effort levels represented were 1) no effort ($p = 0$); 2) low effort (1-3 visits in a breeding season); and 3) high effort (consistent and intensive monitoring).

2.2.10 *Implementation*

Models were written using NIMBLE (de Valpine et al., 2017), accessed through R (R Core Team, 2022). Uninformative priors were used for all parameters in both the simulation and case studies. For the simulation study, models were fit with 3 chains, 10,000 burn-in, and 10,000 iterations per chain. Samples were thinned at a rate of 10 to reduce data storage requirements. For the case study, the model was fit with 3 chains, 10,000 burn-in, and 90,000 iterations per chain with no thinning. Convergence was tested for each model using the Gelman-Rubin convergence diagnostic, where a Gelman-Rubin statistic of $\hat{R} < 1.1$ indicated convergence.

2.3 RESULTS

2.3.1 *Simulation Study*

We present relative bias (RB) and root-mean-square error (RMSE) of survival estimates in which degraded marks were either omitted or were modeled using our mark degradation model. Parameter estimates are reported as medians with 95% credible intervals (CI).

2.3.1.1 Novel Model: Mark Degradation Data Included

When marks were deployed sequentially, a relatively small subset of simulations ran to completion (Figure 2.5). Of the simulations that failed, all failed due to the high RAM requirements of the backtracking algorithm used to generate candidate encounter history matrices, $\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}$. If simulations failed at this step, we could not run the model. As the mark degradation rate increases, the average confusion index also increases, and the number of

candidate encounter history matrices grows nonlinearly. Therefore, as the confusion index increases, the frequency at which the backtracking algorithm runs to completion decreases. For simulated datasets with probability of auxiliary mark degradation of $\beta = 0.01$, the median confusion index was 1 (95% CI: 1, 1.7) across both levels of survival and detection. We ran analyses on a computer with 512 GB of available RAM, and when $\beta = 0.01$, the percentage of time the simulation failed was 17.8%. For simulated datasets with $\beta = 0.10$, the median confusion index across both levels of survival and detection was 1.933 (1, 11.050) and 73.5% of simulations failed.

Though not all simulations successfully ran, the results of those that did were unbiased on average (Figure 2.5). Because patterns are largely invariant to levels of survival, detection, and partial mark degradation probability, we present results marginalized over these parameters. The median RB in survival estimates was -0.016 (-0.09, 0.022) and the median RMSE was 0.093 (0.053, 0.157) (Figure 2.5). The median RB in detection rate estimates was 0.019 (-0.041, 0.304) and the median RMSE was 0.158 (0.072, 0.22). Because only models with relatively low confusion indexes were able to run given available computing power, and lower confusion indexes would indicate lower rates of mark degradation than may be expected given the true β, estimates of the mark degradation probability, β, showed the highest rates of RB, with a median of -0.54 (-0.749, 0.278). The median RMSE was 0.031 (0.007, 0.062).

In contrast, when marks were deployed randomly, a larger subset of simulations ran to completion (Figure 2.5). For simulated datasets with probability of auxiliary mark degradation of $\beta = 0.01$, the median confusion index was 1 (1, 1.144) across both levels of survival and detection, and 12.5% of simulations failed. For simulated datasets with $\beta = 0.10$, the median confusion index across both levels of survival and detection was 1.121 (1, 3.551) and 67.2% of

simulations failed, though note that the trend in the rate of simulation failure was nonlinear with respect to the probability of partial mark degradation, β.

Consistent with sequential deployment models, random deployment models resulted in unbiased estimates of survival. Marginalizing over survival probability, detection probability, and mark degradation probability, the median RB in survival estimates was -0.01 (-0.1, 0.03) and the median RMSE was 0.093 (0.052, 0.156) (Figure 2.5). The median RB in detection estimates was 0.021 (-0.032, 0.333) and median RMSE was 0.127 (0.071, 0.224) Estimates of the mark degradation probability showed the highest rates of RB, with a median of -0.52 (-0.668, 0.231), and RMSE of 0.031 (0.007, 0.062).

2.3.1.2 Traditional Model; Mark Degradation Data Omitted

Given that only a subset of simulations using the novel model ran to completion, we only report results from the traditional model for those same simulations to eliminate any potential biases.

When marks were deployed sequentially (i.e., individuals marked closely in time and space received very similar marks) and observations of degraded marks were omitted from the survival analysis, survival was underestimated, with the effect becoming more pronounced as the probability of mark degradation, β, increased (Figure 2.6). The median RB in survival across all levels of auxiliary mark degradation, $\beta \in \{0, 0.01, \ldots, 0.10\}$, both levels of survival, $\phi \in \{0.4, 0.8\}$, and both levels of detection, $p \in \{0.3, 0.7\}$, was -0.066 (-0.163, -0.001). When $\beta = 0.01$, the median RB across both levels of survival and detection was -0.018 (-0.077, 0.007). The median RB in survival across both levels of survival and detection when $\beta = 0.10$ was -0.119 (-0.143, -0.077). This increase in RB was more pronounced when both survival rate and detection were high (Figure 2.6). RMSE was invariant to β. Across all levels of β, RMSE was lowest when both survival rate and detection were high (i.e., 0.8 and 0.7, respectively; 0.066 (0.055,

0.105)) and highest when survival rate and detection were low (i.e., 0.4 and 0.3, respectively; 0.158 (0.148, 0.166)) (Figure 2.6). The median RB in detection rate estimates was 0.016 (-0.043, 0.348) and the median RMSE was 0.135 (0.076, 0.234).

These same patterns were present when marks were deployed randomly (i.e., individuals received a mark at random from all marks deployed) and observations of degraded marks were omitted from the survival analysis. That is, survival also tended to be underestimated, with the effect becoming more pronounced as the probability of mark degradation, β, increased (Figure 2.6). The median RB in survival across all levels of auxiliary mark degradation, both levels of survival, and both levels of detection, was -0.066 (-0.173, 0.007). In this case, the median RB across both levels of survival and detection was -0.016 (-0.072, -0.002) when $\beta = 0.01$ and -0.121 (-0.178, -0.091) when $\beta = 0.10$. This increase in RB was once again more pronounced when both survival rate and detection were high (Figure 2.6). As before, RMSE was invariant to β. Across all levels of β, RMSE was lowest when both survival rate and detection were high (i.e., 0.8 and 0.7, respectively; 0.065 (0.054, 0.107)) and highest when survival rate and detection were low (i.e., 0.4 and 0.3, respectively; 0.156 (0.15, 0.16)) (Figure 2.6). The median RB in detection rate estimates was 0.019 (-0.042, 0.313) and the median RMSE was 0.132 (0.073, 0.228).

2.3.2 *Case Study*

Using the model presented here, we generated $n = 891$ candidate encounter history matrices, $\boldsymbol{Y}$, given the data. The confusion index for this dataset was 1.177. We found that the median probability of survival was 0.184 (0.081, 0.338) for L birds, 0.539 (0.284, 0.761) for HY birds and 0.72 (0.578, 0.863) for AHY birds (Figure 2.7). We found that AHY exhibited high site fidelity across all sites, with the median dispersal probability being 0.025 (0.003, 0.103) (Figure

2.8). In first-year birds (both L and HY), the dispersal probability was 0.43 (0.238, 0.497). Conditional on dispersing, we did not find strong evidence that any sites were particularly attractive or unattractive as credible intervals for the site-level random effects, ϵ^{ψ}, all overlapped zero. All age-classes showed similar declines in dispersal probability as the distance between sites increased. The median estimate for α^{D} was -0.341 (-1.103, -0.02) for AHY birds and -0.477 (-1.182, -0.09) for L and HY birds. The estimated rate of auxiliary mark degradation, β, was 0.02 (0.007, 0.046). When monitoring effort was low, the probability of detection was 0.4 (0.02, 0.837), and 0.82 (0.619, 0.953) when monitoring effort was high.

When observations of degraded marks were excluded, we found that the probability of survival was lower for all age classes at 0.157 (0.064, 0.304) for L birds, 0.439 (0.202, 0.69) for HY birds, and 0.681 (0.538, 0.822) for AHY birds (Figure 2.7). Estimates of dispersal were slightly lower for first year birds (0.405 (0.185, 0.496)), and approximately the same for AHY birds (0.026 (0.003, 0.11)) (Figure 2.8). As above, we did not find strong evidence that any sites were particularly attractive or unattractive. Declines in dispersal relative to distance between sites were similar (-0.341 (-1.109, -0.02) for AHY birds and -0.549 (-1.332, -0.081) for L and HY birds). Detection probabilities were also similar, with 0.405 (0.02, 0.843) when monitoring effort was low, and 0.832 (0.635, 0.956) when monitoring effort was high.

2.4 DISCUSSION

We developed a novel hierarchical model that leverages all available mark-resight data, including observations of individuals with partially degraded marks, to produce unbiased estimates of survival while accounting for uncertainty in individual identity. Our approach resolves the common and often ignored challenge of mark degradation in mark-resight studies. We demonstrate both the impacts of omitting mark degradation data from survival analyses and

showcase benefits of using a model allowing for the inclusion of these data. Moreover, we present evidence that a random deployment of artificial marks reduces the impact of partial mark degradation on survival estimates. For small or threatened populations of high conservation concern, it is critical to obtain unbiased estimates of population parameters. For instance, in a typical passerine marking scheme, as presented here (i.e., 3 colored bands and an aluminum band), assuming an average lifespan of 3 years, the probability of an individual losing at least one band over the course of their life is approximately 9% at $\beta = 0.01$. At $\beta = 0.1$, this probability jumps to over 60%. Therefore, we suspect that mark degradation is likely a greater problem in mark-resight studies than has previously been acknowledged and that these biases could result in biases in estimates of population abundance and underestimates of viability in population-viability analyses.

In addition to the model presented here, we developed an MCMC sampler for the latent encounter histories, Y, that assigns observations to individuals conditional on both mark degradation and the latent state, Z, of each individual. This model is generalizable to study-specific needs (e.g., Poisson observation process, individual covariates on movement and/or detection, additional spatial information, etc.). We were primarily concerned with degradation of artificially deployed or duplicate marks, for which it is relatively simple to generate the mark transition matrix and the set of possible encounter histories. However, our approach could be extended to any marked population where partial identifications occur (i.e., partial mark observation or mark degradation observed via capture or resight) so long as researchers can enumerate all possible individuals to which uncertain observations could be assigned and specify a model for the mark degradation process.

We assumed that individuals cannot experience total mark loss. While prior work (e.g., Conn et al., 2004; Nichols & Hines, 1993) has shown that total mark loss can cause substantial bias, our study included field-readable auxiliary marks combined with a permanent metal tag, which is unlikely to be lost. In practice, it is possible that the probability of losing the permanent mark is non-zero and this may be more likely for long-lived, migratory, or marine species whose marks experience substantial abrasion. In all cases, we recommend that researchers stress-test both their permanent and auxiliary marks prior to deployment to ensure that the true probabilities of auxiliary and permanent mark degradation are as near to zero as possible. Double marking with a relatively hardy mark (e.g., two metal bands) is a design solution if complete mark loss remains a risk (Arnason & Mills, 1981; Juillet et al., 2011; Schreven & Voslamber, 2022).

An unresolved challenge with both our approach and double-marking approaches in general is the assumption that the probability of losing one mark component (e.g., leg band) is independent of losing the other(s). Schwarz et al. (2012) demonstrated that the assumption of independent mark degradation is often not valid; it is typically more likely for multiple mark components to degrade at once, in which case the assumption of independent mark degradation leads to underestimates of survival rates. For instance, some individuals may be more likely to remove their colored leg bands than others. Alternatively, some studies use both alphanumeric leg flags and colored leg bands and often these are made of different materials. As such, it may be more likely for an individual to lose their colored leg bands than their alphanumeric leg flag. The model presented here could be adapted to include heterogenous loss probabilities, β, for either specific marks, individuals, or both. However, this heterogeneity would greatly increase the complexity in generating the mark transition matrix and may not be feasible for large populations

of marked individuals, or in cases where each individual has a large number of mark components.

We also emphasize that, as presented, our model does not address the problem of incorrectly identified individuals (i.e., partial identifications as a result of observation error as opposed to process error). Imperfect identification would likely introduce serious biases if the problem was pervasive, but has been addressed by others (Morrison et al., 2011; Rakhimberdiev et al., 2022; Roche et al., 2014; Zhang et al., 2017). We recommend training observers to be conservative when identifying marked individuals. When in doubt, observers should record a sighting as incomplete, rather than risking an incorrectly read mark or assuming an individual's identity. If imperfect identification is unavoidable, our model could be extended to a multi-event framework, presuming it was feasible to generate a valid observation matrix.

The principal constraint on the wider applicability of our model is computing power. For large, marked populations and/or populations with high rates of mark degradation, it is very memory-intensive to run the backtracking algorithm and thus generate all valid encounter histories conditional on marks observed. In some cases, the number of possible encounter history matrices is in the millions. The backtracking algorithm is one relatively efficient, but certainly not the only, algorithm suited to solving complex constraint satisfaction problems. Other algorithms may provide less memory intensive ways to generate the encounter histories we feed into the MCMC sampler. However, it is worth noting that the more valid encounter histories there are, the longer MCMC chains will need to be run in order to adequately explore the parameter space.

We presented a case study demonstrating our model on a small but intensively monitored population of Oregon Vesper Sparrows in Washington, USA. In small or threatened populations,

data are often expensive or otherwise challenging to obtain. Therefore, in small datasets such as this one, leveraging all available data can improve estimates both in terms of bias and precision, as our simulation results demonstrate. Existing analyses of breeding site fidelity in other Vesper Sparrow subspecies generally conflate site fidelity with survival, in part due to few recorded returns of marked nestlings either at their natal site or elsewhere (Altman et al., 2020; Best & Rodenhouse, 1984). These analyses find that about 65% of adult (i.e., AHY) birds are resighted at the same site in subsequent years, while few estimates of natal site fidelity currently exist, which could indicate either low natal site fidelity and/or low first-year survival. In this analysis, we found that first year (i.e., L and HY) site fidelity was quite high. However, we found that while adult survival rates were high, survival was relatively low for birds banded in the nest when compared to older hatch-year birds. This suggests that post-fledging is a particularly vulnerable time for OVSP, and that this stage along with low nest success may be contributing to overall population declines.

Data from the Breeding Bird Survey (BBS) indicate a significant range-wide decline of Vesper Sparrows of 2.97% per year (Sauer et al., 2014). In Washington state, management of OVSP primary habitat, prairie-oak savannah, is a priority. However, a lack of information on OVSP vital rates has prevented an analysis of whether these actions have been effective in conserving OVSP. Because we were able to harness all available data, our results will contribute to a more robust understanding of OVSP demography. Presumably, partial mark loss will continue to be a problem in this population, and this model will be a useful tool for estimating changes in demography in response to habitat management.

Our work suggests design approaches for improving results from mark-resight studies in which mark degradation is a concern. First, when choosing marking methods, careful

consideration of mark longevity, including field tests, can help to determine which marks last longest and have the lowest risk of degrading in a way that interferes with individual identification. A random band deployment scheme can reduce the confusion index and reduce the chance of observation errors in the field, because individuals would be more likely to be uniquely identifiable even in the event of partial mark degradation. While this assumes that many mark combinations are unused, that assumption is likely to be met in many marking schemes. For instance, if 11 different band colors are used in a passerine study, with two colored bands on the left leg and 1 metal band and 1 colored band on the right leg, there are 2,662 unique band combinations. More combinations could be generated if the metal band could also go on the left leg, while combinations might be reduced to eliminate those that are difficult to distinguish in the field (e.g., two bands of the same color on the same leg). Moreover, band combinations can be reused if an individual previously assigned that combination is confirmed dead. If mark degradation remains a problem, our proposed model works well for relatively small populations with low rates of mark degradation. If the dataset is sufficiently large and multiple datasets are available, one option would be to perform an integrated demographic analysis and estimate a true survival probability, which is the survival as estimated from a mark-resight model omitting mark degradation data multiplied by a latent inflation factor (e.g., Abadi et al., 2010, Chapter 3). If neither of those are feasible, we recommend reporting the confusion index for the dataset in addition to a description and justification of what was done with the observations of degraded marks in order to address any potential biases.

In conclusion, we present a model for estimating apparent survival, detection, and mark degradation probabilities in marked populations that undergo mark loss/degradation. For multi-state models where states represent discrete sites, we also leverage the information contained in

the spatial information of mark detections. The inclusion of a mark-degradation submodel reduced biases that arise from standard models that ignore the possibility of mark loss. We emphasize that while both the proposed model and the described MCMC sampler work well, a current limitation to our approach is the computational power required to produce the candidate encounter history matrices, $\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}$, that the MCMC sampler uses to update the latent individual encounter histories. Future work could continue to improve the efficiency of the algorithm used to map observed marks to individuals such that we may tackle larger datasets with higher rates of mark degradation in the future.

2.5 ACKNOWLEDGEMENTS

We thank H. Davis for designing the schematic figures and we thank K. Bostrom and P. de Valpine for helpful discussions. We also thank numerous avian field technicians for collecting field data, most notably T. Leque and K. Kelly. Handling and marking of OVSP was approved by Washington Department of Fish and Wildlife through annual scientific collecting permits and USGS Bird Banding Permit 22932. This research was funded through support of A. Bratt by the Washington Cooperative Fish and Wildlife Research Unit and the Center for Natural Lands Management. Fieldwork was supported by Joint Base Lewis McChord Military Base (DABJ25-03-H-4002), U.S. Fish and Wildlife Service (F17AP00813), and American Bird Conservancy (#19062). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

2.6 FIGURES & TABLES

Table 2.1: Model parameters, latent states, and data for described multi-state model including observations of degraded marks.

Parameter	Definition
ϕ	Apparent survival probability
μ^{ψ}	Mean probability of site fidelity, conditional on survival
σ^{ψ}	Scale parameter for site-specific random effect on movement probability
ϵ_s^{ψ}	Site-specific random effect on movement probability
α^{D}	Effect of distance on movement probability
β	Probability of losing a single auxiliary mark
p	Probability of detection
$\theta_{c,d}$	Probability of transitioning from mark c to mark d
$\boldsymbol{Z}_{i,}$	State history for individual i
$\boldsymbol{M}_{i,}$	Mark history for individual i
$\boldsymbol{Y}_{\boldsymbol{i},}$	Multi-state encounter history for individual i
$\boldsymbol{Y}_{\boldsymbol{i},\boldsymbol{k},\boldsymbol{s},}^{multi}$	Individual i's encounter history at site s in mark k
$\boldsymbol{Y}_{k,s,}^{obs}$	Observed detection history of mark k at site s
N	Number of marked individuals, $i \in \{1, \dots, N\}$
T	Number of occasions, $t \in \{1, \dots, T\}$
S	Number of sites, $s \in \{1, \dots, S\}$
$\boldsymbol{K}$	Set of observable marks, where $\lvert\boldsymbol{K}\rvert$ is the number of observable marks
$\boldsymbol{D}$	Distance matrix
$M_{c,d}$	Matrix delineating number of possible transitions from mark c to mark d, $c, d \in \boldsymbol{K}$

Figure 2.1: Example spatial representation of marks observed in each year for three individual passerines, each marked with two colored auxiliary marks and one metal permanent band. Individuals were marked with auxiliary marks in colors Blue and Green (BG), Red and Green (RG), and Blue and Purple (BP). Marks were observed without error in Year 1 when individuals were captured at one of three sites. The same three sites were monitored for the remaining three years of study during which resight surveys were conducted. During this period, individuals could move throughout the study area and undergo partial loss of auxiliary marks, as demonstrated in years two and three. No marked individuals were observed in the study area in year four.

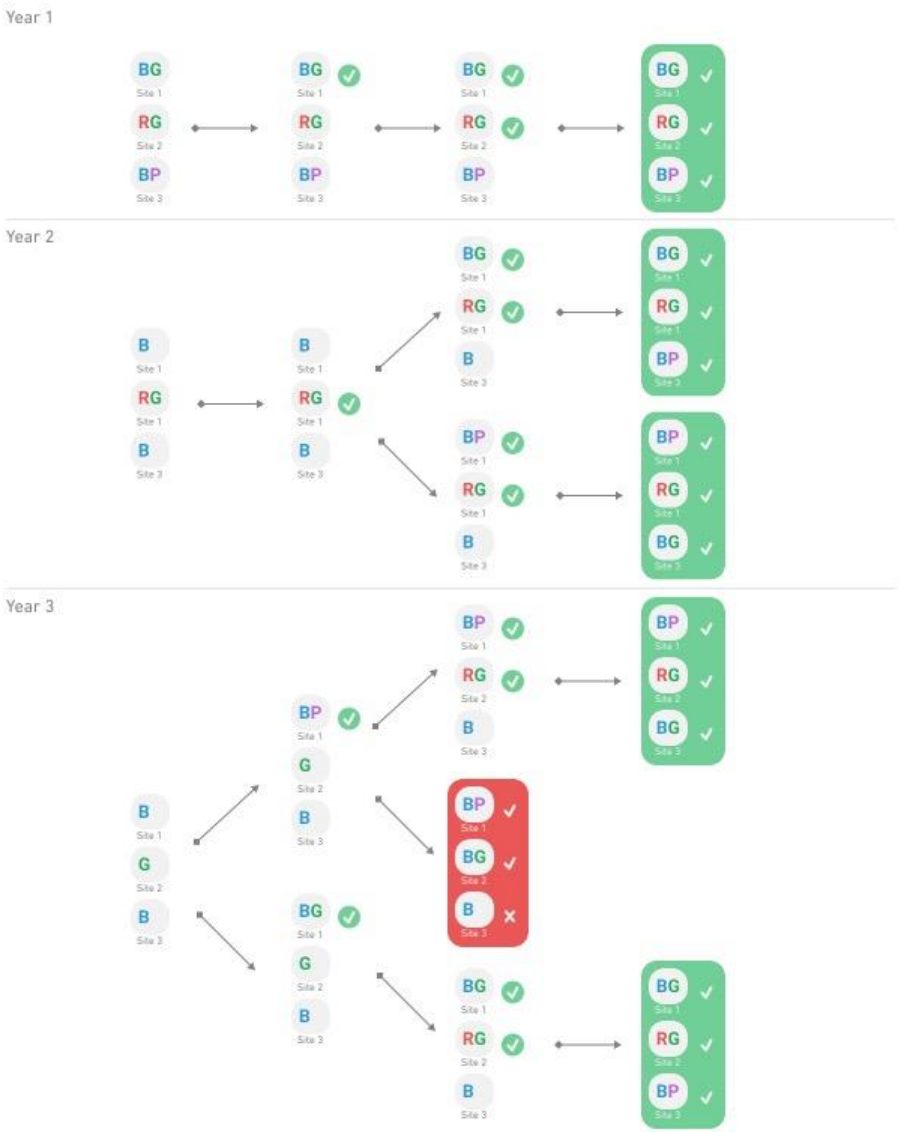

Figure 2.2: Visualization of how the backtracking algorithm assigns individuals to observed marks, using the described example. The algorithm is run separately on years 1-3 (top-bottom). Within each year, the algorithm begins with the set of observed marks, where observed marks are not yet assigned to any individuals. Each node of the search tree increments the previous partial solution by assigning one more observed mark to an individual (green check marks). Green boxes around nodes indicate that each observed mark has been successfully matched to an individual. Red boxes around nodes indicate that not all observed marks can be successfully matched to individuals and therefore this branch of the search tree has not resulted in a valid solution.

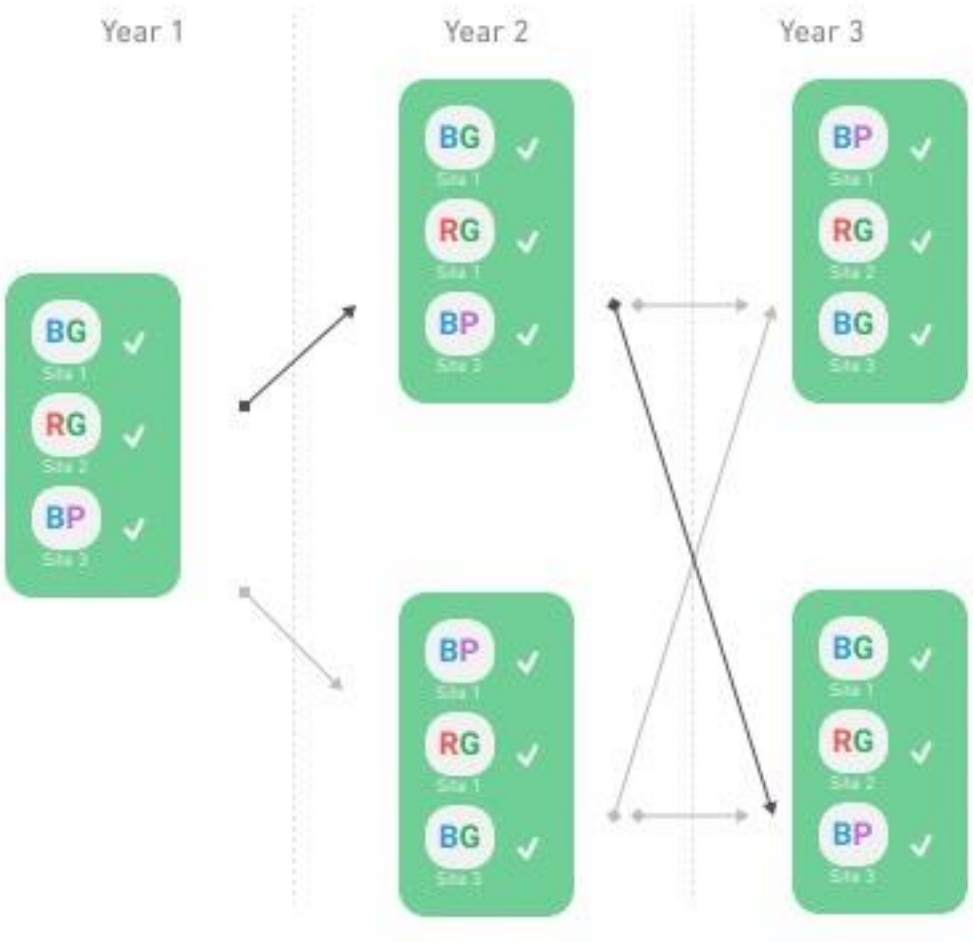

Figure 2.3: Graph constructed using the valid terminal nodes constructed by the backtracking algorithm for each of Years 1-3 (Figure 2.2). Each terminal node from Year 1 is connected to each terminal node from Year 2, and each terminal node from Year 2 is connected to each terminal node from Year 3. No marks were observed in Year 4 so it is omitted from the graph. The path through this graph indicated by the darker arrows results in the true encounter history described within the in-text example.

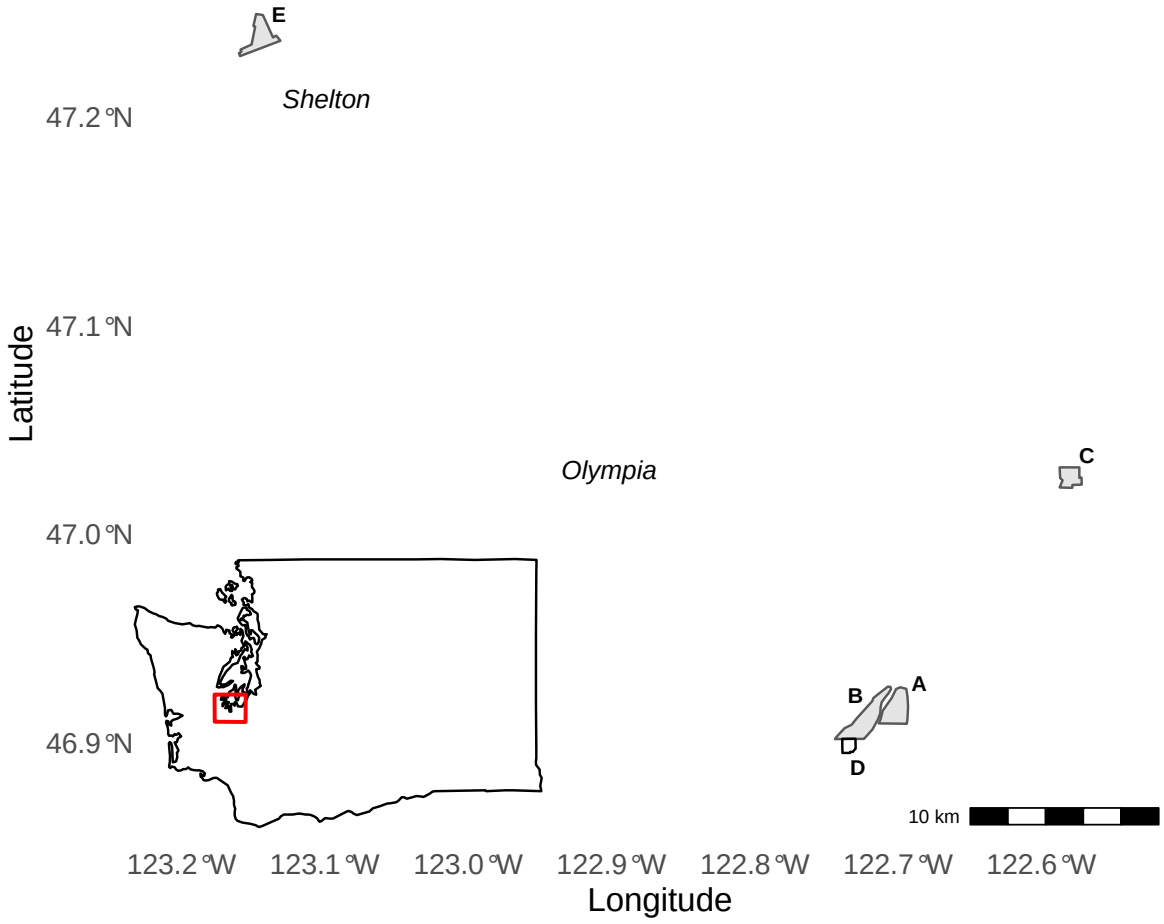

Figure 2.4: Map of Oregon Vesper Sparrow monitoring sites in western Washington, USA, where the shaded blue area represents Joint Base Lewis-McChord. Sites A-D are native prairie sites, while site E is a municipal airport. Site A is Lower Weir prairie, B is Upper Weir prairie, C is Range 76, D is Tenalquot prairie, and E is Sanderson Airport.

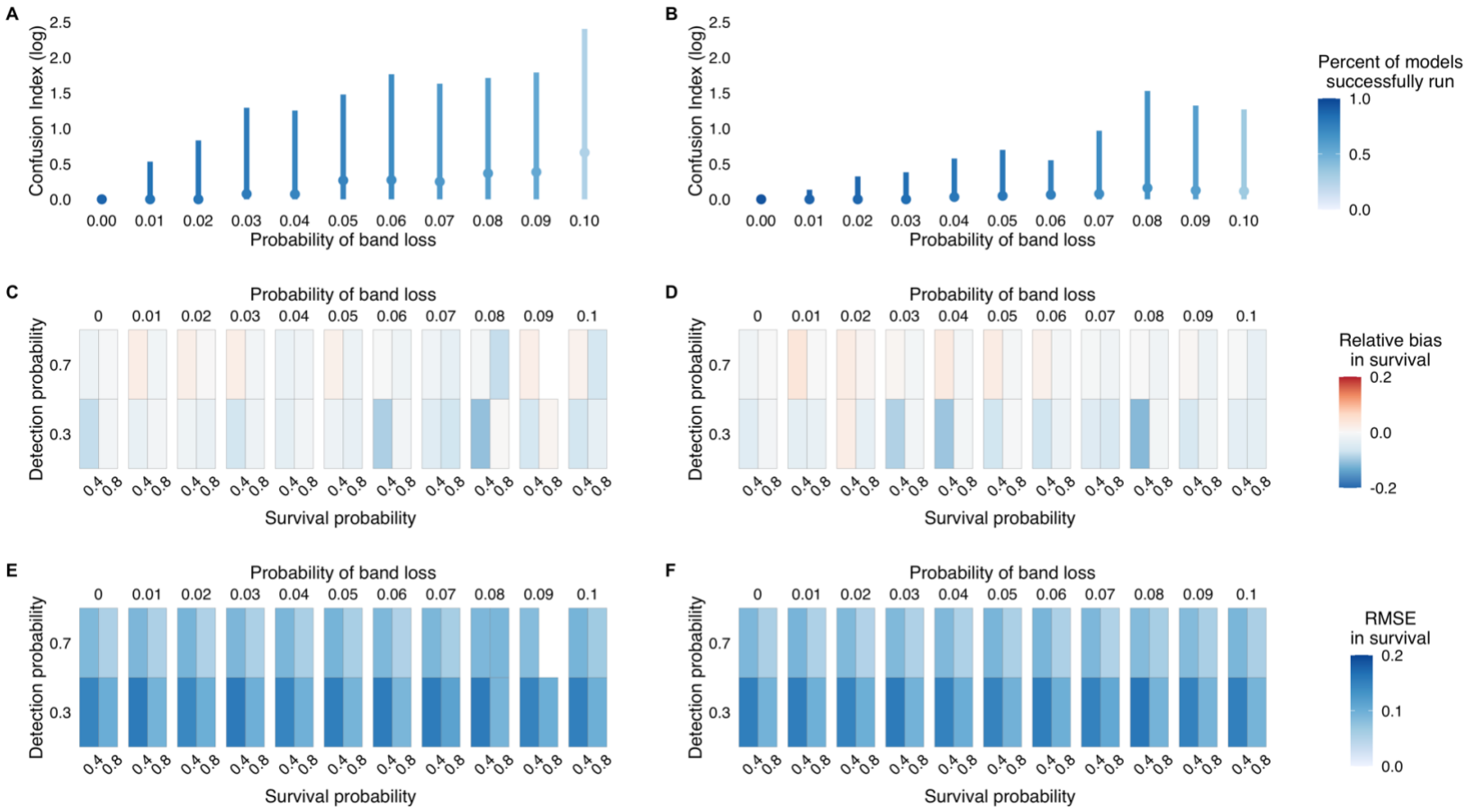

Figure 2.5: Median (points) and 95% confidence intervals (line ranges) of log-scale confusion indices across all 100 simulated datasets relative to the probability of partial mark degradation, assuming sequential mark deployment (A), or random mark deployment (B). The color of the points and line ranges are scaled to the percent of models successfully run, with darker blues indicating more models run. Also shown are median RB (C, D) and RMSE (E, F) in survival estimates when observations of degraded marks are included in the analysis, using the model described here. RB and RMSE are presented with respect to survival probability, detection probability, and the probability of partial mark degradation when marks deployed either sequentially (A, C, E), or randomly (B, D, F).

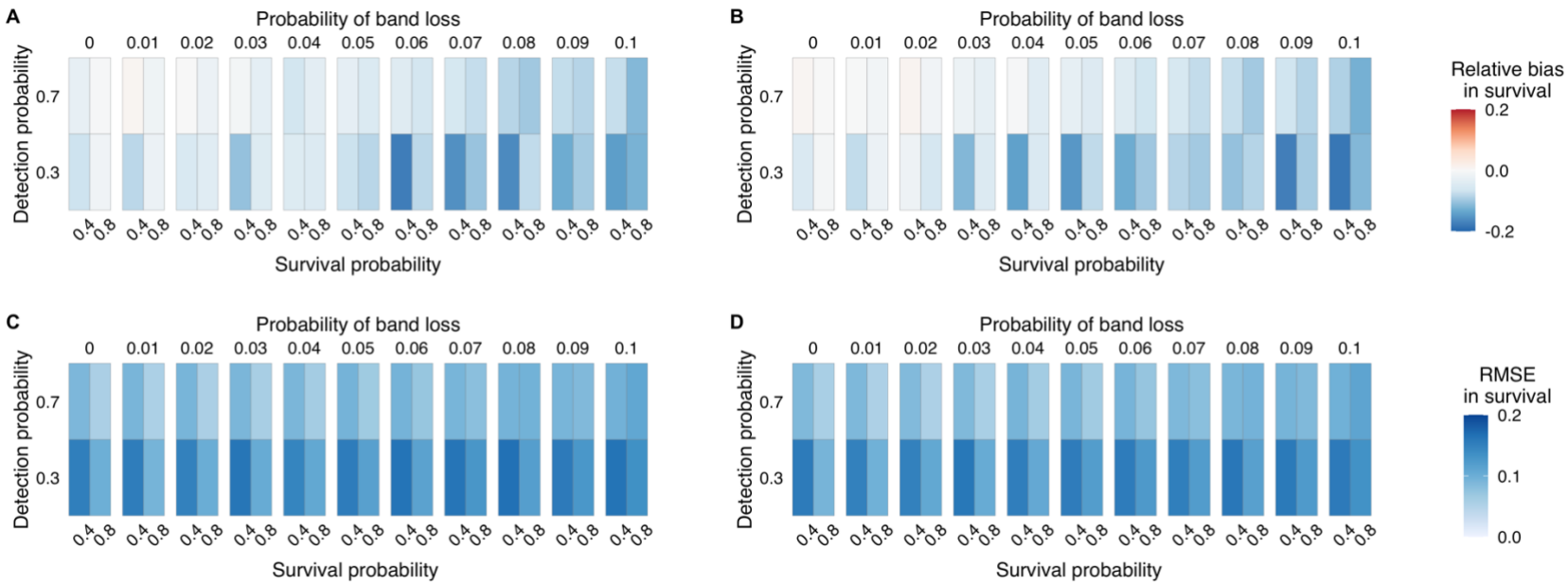

Figure 2.6: Median relative bias (RB; A, B) and root mean square error (RMSE; C, D) in survival estimates when observations of degraded marks are omitted from the analysis, and when marks are deployed either sequentially (A, C), or randomly (B, D). RB and RMSE are presented with respect to survival probability, detection probability, and the probability of partial mark degradation.

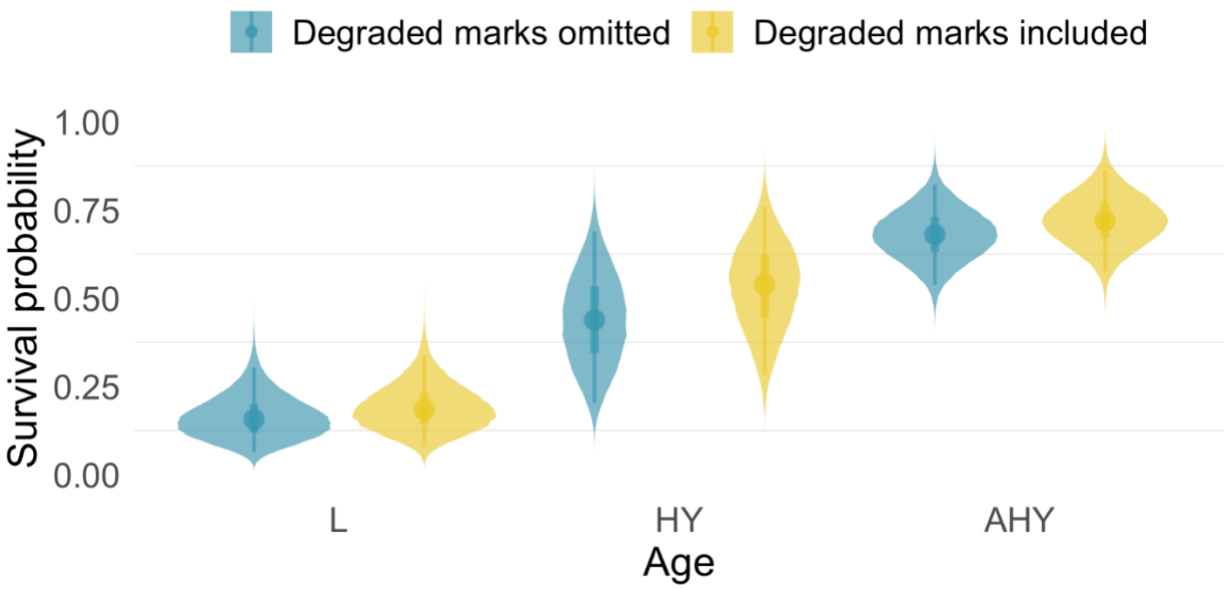

Figure 2.7: Age-specific survival estimates for western Washington Oregon Vesper Sparrows from model omitting observations of partially degraded marks (blue) compared to estimates from model including degraded marks (yellow). Posterior distributions are shown, as well as medians (points), and 95% credible intervals (line ranges). Inclusion of degraded marks leads to slightly higher and more precise estimates of survival across all age-classes.

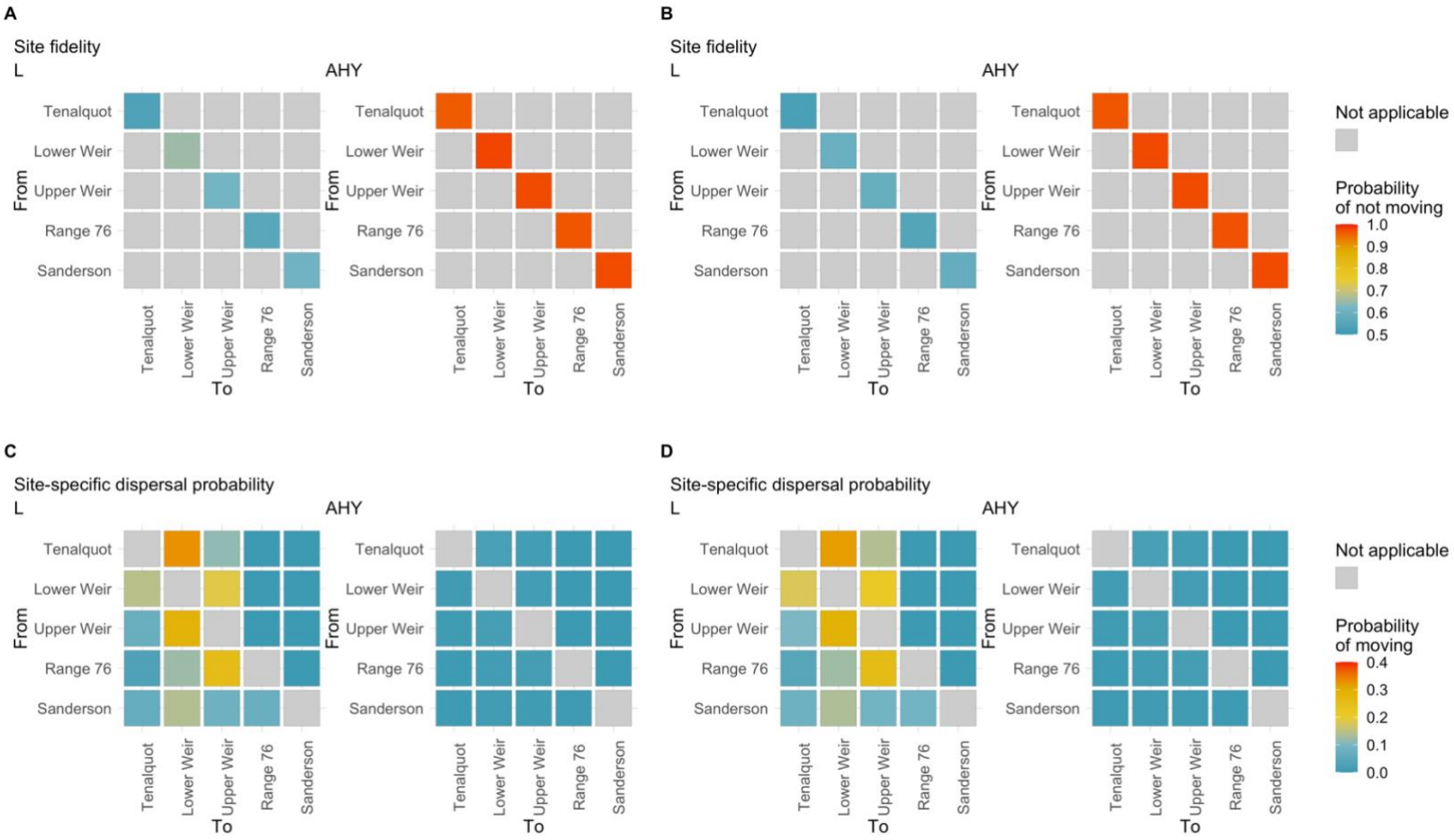

Figure 2.8: Median site- and age-specific probabilities of site fidelity for Oregon Vesper Sparrow in western Washington, USA (A, B) from model omitting observations of degraded marks (A) compared to model including observations of degraded marks (B). Also shown are site- and age-specific dispersal probabilities (C, D) from model omitting observations of degraded marks (C) compared to model including observations of degraded marks (D). L and HY birds are shown on the left, compared to AHY birds on the right. Omitting degraded marks results in slightly higher estimates of site fidelity and lower estimates of dispersal.

2.7 REFERENCES

Abadi, F., Gimenez, O., Ullrich, B., Arlettaz, R., & Schaub, M. (2010). Estimation of immigration rate using integrated population models. *Journal of Applied Ecology*, *47*(2), 393–400.

Altman, B. (2011). Historical and current distribution and populations of bird species in prairie-oak habitats in the pacific northwest. *Northwest Science*, *85*(2), 194–222. https://doi.org/10.3955/046.085.0210

Altman, B., Stinson, D., & Hayes, G. (2020). Revised draft status report for the Oregon Vesper Sparrow in Washington. *Washington Department of Fish and Wildlife*. https://wdfw.wa.gov/sites/default/files/2021-01/12oregon_vesper_sparrow_status_report.pdf

Arnason, A. N., & Mills, K. H. (1981). Bias and loss of precision due to tag loss in jolly–seber estimates for mark–recapture experiments. *Canadian Journal of Fisheries and Aquatic Sciences*, *38*(9), 1077–1095. https://doi.org/10.1139/f81-148

Augustine, B. C., Royle, J. A., Linden, D. W., & Fuller, A. K. (2020). Spatial proximity moderates genotype uncertainty in genetic tagging studies. *Proceedings of the National Academy of Sciences*, *117*(30), 17903–17912. https://doi.org/10.1073/pnas.2000247117

Augustine, B. C., Royle, J. A., Murphy, S. M., Chandler, R. B., Cox, J. J., & Kelly, M. J. (2019). Spatial capture–recapture for categorically marked populations with an application to genetic capture–recapture. *Ecosphere*, *10*(4), e02627. https://doi.org/10.1002/ecs2.2627

Baillie, S. R., & Schaub, M. (2009). Understanding changes in bird populations – the role of bird marking. *Ringing & Migration*, *24*(3), 189–198. https://doi.org/10.1080/03078698.2009.9674391

Best, L. B., & Rodenhouse, N. L. (1984). Territory preference of vesper sparrows in cropland. *The Wilson Bulletin*, *96*(1), 72–82.

Brannelly, L. A., Chatfield, M. W. H., & Richards-Zawacki, C. L. (2013). Visual implant elastomer (VIE) tags are an unreliable method of identification in adult anurans. *Herpetological Journal*, *23*, 125–129.

Calvo, B., & Furness, R. W. (1992). A review of the use and the effects of marks and devices on birds. *Ringing & Migration*, *13*(3), 129–151. https://doi.org/10.1080/03078698.1992.9674036

Conn, P. B., Kendall, W. L., & Samuel, M. D. (2004). A general model for the analysis of mark-resight, mark-recapture, and band-recovery data under tag loss. *Biometrics*, *60*(4), 900–909. https://doi.org/10.1111/j.0006-341X.2004.00245.x

Cormack, R. M. (1964). Estimates of survival from the sighting of marked animals. *Biometrika*, *51*(3/4), 429–438. https://doi.org/10.2307/2334149

Cowen, L., & Schwarz, C. J. (2006). The Jolly-Seber Model with Tag Loss. *Biometrics*, *62*(3), 699–705.

de Valpine, P., Turek, D., Paciorek, C. J., Anderson-Bergman, C., Lang, D. T., & Bodik, R. (2017). Programming with models: Writing statistical algorithms for general model structures with nimble. *Journal of Computational and Graphical Statistics*, *26*(2), 403–413. https://doi.org/10.1080/10618600.2016.1172487

Diefenbach, D. R., & Alt, G. L. (1998). Modeling and evaluation of ear tag loss in black bears. *The Journal of Wildlife Management*, *62*(4), 1292–1300. https://doi.org/10.2307/3801993

Jiménez, J., Augustine, B. C., Linden, D. W., Chandler, R. B., & Royle, J. A. (2021). Spatial capture–recapture with random thinning for unidentified encounters. *Ecology and Evolution*, *11*(3), 1187–1198. https://doi.org/10.1002/ece3.7091

Jiménez, J., Chandler, R., Tobajas, J., Descalzo, E., Mateo, R., & Ferreras, P. (2019). Generalized spatial mark–resight models with incomplete identification: An application to red fox density estimates. *Ecology and Evolution*, *9*(8), 4739–4748. https://doi.org/10.1002/ece3.5077

Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika*, *52*(1/2), 225–247. https://doi.org/10.2307/2333826

Juillet, C., Choquet, R., Gauthier, G., & Pradel, R. (2011). A capture–recapture model with double-marking, live and dead encounters, and heterogeneity of reporting due to auxiliary mark loss. *Journal of Agricultural, Biological, and Environmental Statistics*, *16*(1), 88–104. https://doi.org/10.1007/s13253-010-0035-5

Kery, M., & Schaub, M. (2011). *Bayesian Population Analysis using WinBUGS: A Hierarchical Perspective*. Academic Press.

Larsen, T. (1971). Capturing, handling, and marking polar bears in svalbard. *The Journal of Wildlife Management*, *35*(1), 27–36. https://doi.org/10.2307/3799868

Lindberg, M. S. (2012). A review of designs for capture–mark–recapture studies in discrete time. *Journal of Ornithology*, *152*(S2), 355–370. https://doi.org/10.1007/s10336-010-0533-9

Link, W. A., Yoshizaki, J., Bailey, L. L., & Pollock, K. H. (2010). Uncovering a latent multinomial: Analysis of mark–recapture data with misidentification. *Biometrics*, *66*(1), 178–185. https://doi.org/10.1111/j.1541-0420.2009.01244.x

Lovell, H. B. (1948). The removal of bands by cardinals. *Bird-Banding*, 2.

Lukacs, P. M., & Burnham, K. P. (2005). Research notes: Estimating population size from DNA-based closed capture–recapture data incorporating genotyping error. *Journal of Wildlife Management*, *69*(1), 396–403.

Marion, W. R., & Shamis, J. D. (1977). An annotated bibliography of bird marking techniques. *Bird-Banding*, *48*(1), 42–61. https://doi.org/10.2307/4512291

McClintock, B. T., Hill, J. M., Fritz, L., Chumbley, K., Luxa, K., & Diefenbach, D. R. (2014). Mark-resight abundance estimation under incomplete identification of marked individuals. *Methods in Ecology and Evolution*, *5*(12), 1294–1304. https://doi.org/10.1111/2041-210X.12140

Morrison, T. A., Yoshizaki, J., Nichols, J. D., & Bolger, D. T. (2011). Estimating survival in photographic capture–recapture studies: Overcoming misidentification error. *Methods in Ecology and Evolution*, *2*(5), 454–463. https://doi.org/10.1111/j.2041-210X.2011.00106.x

Neil Arnason, A. (1973). The estimation of population size, migration rates and survival in a stratified population. *Researches on Population Ecology*, *15*(2), 1–8. https://doi.org/10.1007/BF02510705

Nichols, J. D., & Hines, J. E. (1993). Survival rate estimation in the presence of tag loss using joint analysis of capture-recapture and resighting data. 229–243.

R Core Team. (2022). *R: A language and environment for statistical computing* [R Foundation for Statistical Computing].

Rakhimberdiev, E., Karagicheva, J., Saveliev, A., Loonstra, A. H. J., Verhoeven, M. A., Hooijmeijer, J. C. E. W., Schaub, M., & Piersma, T. (2022). Misidentification errors in reencounters result in biased estimates of survival probability from CJS models: Evidence

and a solution using the robust design. *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-210X.13825

Roche, E. A., Dovichin, C. M., & Arnold, T. W. (2014). Field-readable alphanumeric flags are valuable markers for shorebirds: Use of double-marking to identify cases of misidentification: Alphanumeric Flags as Markers for Shorebirds. *Journal of Field Ornithology*, *85*(3), 329–338. https://doi.org/10.1111/jofo.12072

Rossi, F., Beek, P. van, & Walsh, T. (2006). *Handbook of Constraint Programming*. Elsevier.

Sauer, J., Hines, J., Fallon, J., Pardieck, K., Ziolkowski Jr, D., & Link, W. (2014). *The North American Breeding Bird Survey, results and analysis 1966—2012, Version 02.19.2014*.

Schreven, K., & Voslamber, B. (2022). Neckband loss and its effect on apparent survival estimates in Greylag Geese (Anser anser): Variation with season, sex and age. *Journal of Ornithology*, *163*, 1–12. https://doi.org/10.1007/s10336-022-01993-2

Schwarz, C. J., & Stobo, W. T. (1999). *Estimation and effects of tag-misread rates in capture–recapture studies*. *56*, 9.

Schwarz, L. K., Hindell, M. A., McMahon, C. R., & Costa, D. P. (2012). The implications of assuming independent tag loss in southern elephant seals. *Ecosphere*, *3*(9), art81. https://doi.org/10.1890/ES12-00132.1

Seber, G. A. F. (1965). A note on the multiple-recapture census. *Biometrika*, *52*(1/2), 249–259. https://doi.org/10.2307/2333827

Silvy, N. J. (2012). *The Wildlife Techniques Manual: Volume 1: Research. Volume 2: Management 2-vol. Set*. JHU Press.

Tucker, A. M., McGowan, C. P., Robinson, R. A., Clark, J. A., Lyons, J. E., DeRose-Wilson, A., Du Feu, R., Austin, G. E., Atkinson, P. W., & Clark, N. A. (2019). Effects of individual

misidentification on estimates of survival in long-term mark–resight studies. *The Condor*, *121*(1). https://doi.org/10.1093/condor/duy017

Weatherhead, P., & Forbes, M. (1994). Natal philopatry in passerine birds: Genetic or ecological influences? *Behavioral Ecology*, *5*. https://doi.org/10.1093/beheco/5.4.426

Williams, B. K., Nichols, J. D., & Conroy, M. J. (2002). *Analysis and Management of Animal Populations: Modeling, Estimation and Decision Making*. https://pubs.er.usgs.gov/publication/5200256

Wright, J. A., Barker, R. J., Schofield, M. R., Frantz, A. C., Byrom, A. E., & Gleeson, D. M. (2009). Incorporating genotype uncertainty into mark—Recapture-type models for estimating abundance using dna samples. *Biometrics*, *65*(3), 833–840. JSTOR.

Yoshizaki, J., Brownie, C., Pollock, K., & Link, W. (2011). Modeling misidentification errors that result from use of genetic tags in capture–recapture studies. *Environmental and Ecological Statistics*, *18*, 27–55. https://doi.org/10.1007/s10651-009-0116-1

Zhang, J., Dennis, T. E., Landers, T. J., Bell, E., & Perry, G. L. W. (2017). Linking individual-based and statistical inferential models in movement ecology: A case study with black petrels (*Procellaria parkinsoni*). *Ecological Modelling*, *360*, 425–436. https://doi.org/10.1016/j.ecolmodel.2017.07.017

Chapter 3. POPULATION DYNAMICS AND VIABILITY OF AN ENDANGERED GRASSLAND BIRD ON A FRAGMENTED LANDSCAPE

Publication history: This study was co-authored with Gary Slater, Ilai Keren, Scott Pearson, and Sarah Converse. At the time this dissertation was published, this chapter was not in review with a journal.

Abstract: Understanding the dynamics of small, declining, or threatened populations in fragmented landscapes is critical for understanding population viability and identifying optimal management actions. On fragmented landscapes, population trends can vary widely between sites and estimating movement between sites can reveal source-sink dynamics that may be missed at broader spatial scales but have large impacts on site-specific extirpation risk and population viability. Streaked Horned Larks (*Eremophila alpestris strigata;* SHLA) are native to lowland prairies in the Cascadia bioregion and have since been extirpated from much of their historical range due to habitat loss and degradation and human activities. In Washington State's South Puget Sound region, SHLA now exist on a highly fragmented landscape. SHLA have been monitored to varying degrees at nine occupied sites in this region since 2010. We developed a multi-site integrated population model integrating nest monitoring, mark-resight, and count data for SHLA. Our results show that range-wide abundance decreased by approximately 1.2% per year since 2010, though this rate was sensitive to assumptions about initial population size and was highly variable across occupied sites, with the sub-populations at some sites showing significant growth during this period, and the sub-populations at other sites declining to near 0. Overall population growth rate was strongly correlated with survival for all age-classes and nest

survival at all nest stages, providing two key avenues for potential conservation action. Results from this work will inform ongoing habitat monitoring and management, as well as a planned reintroduction of SHLA to currently unoccupied sites in western Washington.

Keywords: Integrated population model, Bayesian hierarchical model, endangered species, population viability, Streaked Horned Lark, *Eremophila alpestris strigata*

3.1 INTRODUCTION

The widespread decline of grassland birds in North America over the last six decades or more has been called a conservation crisis (Brennan & Kuvlesky Jr., 2005). The causes of these declines are complex, including forest encroachment (Coppedge et al., 2001), habitat loss and fragmentation (Stanton et al., 2018), the introduction of invasive plants (Hamman et al., 2011), and agricultural development (Mahony et al., 2022), among others. Among grassland birds, Horned Larks (*Eremophila alpestris*) are widespread throughout North America. While Horned Larks generally have exhibited population declines, Streaked Horned Larks (*Eremophila alpestris strigata*; SHLA), are a subspecies of particular conservation concern. SHLA are the only subspecies of Horned Lark to be recognized as threatened in the US and are endangered in Washington State. SHLA historic range spanned prairie habitats throughout western Oregon, Washington, and British Columbia (Altman, 2011), although they have since been extirpated from southern Oregon and British Columbia. Like other grassland birds, SHLA have experienced habitat loss, fragmentation, and degradation in the face of the spread of invasive plants and rapid urban development in the Pacific Northwest (Dunwiddie & Bakker, 2011). SHLA are considered a sentinel species for prairie systems in western Washington, as they exhibit high site fidelity and have limited ability to withstand perturbations such as habitat loss (Camfield et al., 2011). But

despite their ecological importance, little is known about the factors driving SHLA population dynamics.

Information about drivers of population dynamics across large spatial and temporal scales can inform conservation measures for endangered species. However, developing such an understanding is generally challenging due to limited data availability. While individual data streams, such as count data, can give insights about population trends, they cannot give insights into the mechanisms driving population change. In recent years, integrated population models (IPMs) have become a valuable tool in wildlife management because they combine data from several sources to estimate both demographic rates and states (Schaub & Abadi, 2011). Because multiple data sources provide information on the same demographic rates (e.g., survival affects both mark-resight data as well as population counts), these rates can be estimated with increased precision in an integrated analysis (Schaub et al., 2007). Moreover, IPMs can estimate parameters that would otherwise be unidentifiable, such as recruitment or immigration (Abadi et al., 2010). These advances have practical benefits for resource managers who must make conservation management decisions based on limited information, particularly for populations that are small, declining, or otherwise difficult to monitor.

Despite the proliferation of IPMs in wildlife ecology, there are relatively few examples of multi-site IPMs (though see McCrea et al., 2010). Developing multi-site IPMs at any spatial scale is often made challenging by spatial and temporal mismatches in datasets, as well as variable sampling effort and observation error between sites and observers (Ahrestani et al., 2017). Some IPMs, while technically multi-site, model local populations independently, without accounting for potential between-site interactions (Weegman et al., 2017). There are notable exceptions, and recent work has demonstrated the utility of multi-site IPMs for understanding

spatial dynamics (Chandler & Clark, 2014), detecting source-sink dynamics (Weegman et al., 2016), and describing the dynamics of abundant and widely distributed species monitored across large-scale monitoring programs (Saracco & Rubenstein, 2020). However, multi-site IPMs for rare species in fragmented landscapes are underexplored.

We developed a multi-site IPM for SHLA in western Washington using 10 years of population monitoring data to estimate demographic rates and site- and region-wide abundance. We integrated nest-monitoring, mark-resight, and population count data from 9 sites in western Washington collected between 2010 and 2020. Our model provides a tool for exploring the population dynamics of SHLA on a fragmented landscape, an improved understanding of which will facilitate future management directions and priorities, including providing important information for reintroduction planning (Sipe et al. *in prep*). The model also provides a tool for estimating the effects of environmental and management covariates on vital rates when additional data become available in the future. Our modeling framework extends existing methods for the estimation of daily nest survival, annual survival in populations that undergo mark degradation, and abundance for species with unequal detectability across demographic groups. These characteristics make our modeling framework broadly applicable to other avian species, especially those inhabiting fragmented landscapes.

3.2 METHODS

3.2.1 *Study System and Species*

While SHLA are endemic to lowland habitats west of the Pacific Northwest's Cascade Mountains (Altman, 2011), their current range in Washington is largely restricted to a discrete set of sites in the South Puget Lowlands. SHLA are also found on Washington's outer coast, on Columbia River islands, and in Oregon's Willamette Valley (Altman, 2011). Within SPL there

are nine currently or recently occupied sites that are monitored to varying degrees during the breeding season. Of these sites, five are native prairies and four are airfields (Figure 3.1). Six of the sites are located on the U.S. military installation Joint Base Lewis-McChord (JBLM). These six sites include two airfields and all known occupied prairie sites. While vegetation composition is different between native prairie and airfield sites, all sites are located on former or current prairies and characterized by glacial outwash soils that can support grassland vegetation (e.g., short grasses and perennial forbs) and a high percentage of bare ground, a habitat characteristic preferred by SHLA.

SHLA are partially migratory with nearly all birds leaving the Puget Lowland sites while birds in Oregon overwinter in similar habitat on or near their breeding locations. They over-winter in large, mixed flocks in southwestern Washington State and in Oregon's Willamette Valley between late August and early April. Males typically arrive at the breeding sites in early April and establish territories in mid- to late April when pairing and nest site selection occurs (Altman, 2011). SHLA are territorial and generally monogamous throughout the May – August breeding season, during which pairs can renest 2-3 times. SHLA typically exhibit high site fidelity, though dispersal between sites has been observed and, when it does occur, is more common in first-year individuals (Wolf et al., 2020). Like many passerines, individuals reach sexual maturity in their first year of life.

3.2.2 *Population Monitoring Data*

Our model combined three demographic data sources collected over the 2010 to 2020 breeding seasons. Nest-monitoring data were collected at three intensively monitored sites, capture-mark-resight data were collected at all nine of the occupied SPL sites (though all marking and the

majority of resighting were done at the three nest-monitoring sites), and count data were collected during repeated annual surveys at the nine occupied sites.

3.2.2.1 Nest Monitoring Data

Nest-monitoring data were obtained primarily from three sites on JBLM where observers had regular access (D, G, and E; Figure 3.1), with additional data from one site (H; Figure 3.1) where observer access was more limited due to military operations. Nests were located using systematic search methods and behavioral cues as described by Martin and Geupel (1993). Once located, nests were revisited every few days to assess their status. For successful nests, the number of fledglings was recorded. Because a large proportion of SHLA at these sites were individually marked, there was a subset of nests for which observers could uniquely identify parentage. For nests with known parentage, the number of renesting attempts per year was approximated by the number of nests found within a pair's territory. However, these data are likely biased given that all nesting attempts per pair may not be located.

3.2.2.2 Capture-mark-resight Data

Most (99%) marked individuals were marked as fledglings in the nest (hereafter, age class L) just prior to their expected fledge date. Individuals were marked with a unique combination of three colored bands (Avinet Research Supplies, 2.8mm interior diameter) which are field-readable at a distance, and a unique aluminum USGS band, which is not field-readable. A smaller percentage of birds were marked as post-fledglings (HY) or birds of age 1 or greater (AHY; 0.5 and 0.1%, respectively) following capture in mist-nests deployed at the end of the breeding season at nest-monitoring sites. All individuals were marked at the same sites at which nest monitoring occurred. Resight data were collected opportunistically by observers throughout the breeding season. Most (98%) of resights were obtained at the three sites where intensive nest monitoring

occurred (D, G, and E; Figure 3.1). These opportunistic resights were processed into annual detection histories.

3.2.2.3 Count Data

Finally, counts were conducted multiple (typically 3-4) times per site per year following the protocol of Pearson et al. (2016). Per the protocol, multiple trained observers walked line transects and recorded all birds detected either aurally or visually up to 75m on either side of the transects. When possible, birds were aged and sexed upon detection. Surveys were conducted after initial pairing but early in the breeding season to maintain an assumption of population closure. However, not all transects were surveyed in all years. We account for variable effort in the model described below by including a correction factor for the proportion of suitable habitat surveyed in each site-year.

3.2.3 *Statistical Modeling*

To estimate key demographic rates (i.e., productivity and survival) and states (i.e., age- and site-specific abundance) we developed a two-sex IPM that integrated the nest monitoring, count, and mark-resight data from 2010-2020. We describe models for each of these datasets below.

3.2.3.1 Productivity Model

We developed a multi-part model to estimate overall productivity (i.e., number of fledglings per female per year). The first component of this model is a multi-state daily nest survival model for estimating probability of nest success (i.e., probability that a nest produces at least one fledgling) while accounting for uncertainty in nest age and state. The nest states were laying ($G = 1$), incubating ($G = 2$), nestling ($G = 3$), fledged ($G = 4$), and failed ($G = 5$). This model is a simplification of the multi-event Jolly-Seber model described by Warlick et al. (2022), which

conditions on first observation and assumes that nest state can be observed perfectly when checked. Critically, the model allows transitions between nest states to be described as a function of latent nest age. For each nest n, its state $\boldsymbol{G}_{n,\cdot}$ was modeled on a daily scale between the day it was first observed, $first_n$, and the day it was first observed in a fledged or failed state, $last_n$. Nests began in the state in which they were first observed and on subsequent days, nest state was categorically distributed based on the state in the previous day, where state transition probabilities are described by the matrix $\boldsymbol{\Pi}$. Specifically,

$$G_{n,d}|G_{n,d-1} \sim Categorical\left(\boldsymbol{\Pi}_{G_{n,d-1},x_{n,d-1}}\right)$$

where

$$\boldsymbol{\Pi} = \begin{bmatrix} Laying & Incubating & Nestling & Fledged & Failed \\ S_{n,d}^{lay}(\eta_{x_{n,d}}^{lay}) & S_{n,d}^{lay}(1-\eta_{x_{n,d}}^{lay}) & 0 & 0 & 1-S_{n,d}^{lay} \\ 0 & S(\eta_{x_{n,d}}^{inc}) & S_{n,d}^{inc}(1-\eta_{x_{n,d}}^{inc}) & 0 & 1-S_{n,d}^{inc} \\ 0 & 0 & S_{n,d}^{nes}(\eta_{x_{n,d}}^{nes}) & S_{n,d}^{nes}(1-\eta_{x_{n,d}}^{nes}) & 1-S_{n,d}^{nes} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The matrix Π formulates state transitions based on the underlying state-specific survival ($S_{n,d}^{state}$) and transition probabilities ($\boldsymbol{\eta}^{state}$) that depend on age-in-state $x_{n,d}$ on day d. In the laying state, nests can survive ($S_{n,d}^{lay}$) and remain in the laying state ($\eta_{x_{n,d}}^{lay}$) or transition to the incubating state ($1-\eta_{x_{n,d}}^{lay}$). In the incubating state, nests can survive ($S_{n,d}^{lay}$) and remain in the incubating state ($\eta_{x_{n,d}}^{inc}$) or transition to the nestling state ($1-\eta_{x_{n,d}}^{inc}$). In the nestling state, nests can survive ($S_{n,d}^{lay}$) and remain in the nestling state ($\eta_{x_{n,d}}^{nes}$) or fledge ($1-\eta_{x_{n,d}}^{nes}$). In any state prior to fledging, nests can also fail with probability $1-S_{n,d}^{state}$.

Transitions between states were modeled as categorical processes that were non-zero during biologically plausible ages-in-state and zero otherwise. That is, nests could be in the

laying state from 1-6 days, in the incubating state from 10-13 days, and in the nestling state from 4-14 days. Between these ages-in-state, transition probabilities, η_{θ}^{state}, were non-zero and represented the probability of transitioning at age-in-state $x_{n,d}$ conditional on having not yet transitioned, per Royle & Dorazio (2008).

Finally, daily nest survival is modeled as

$$logit(S_{n,d,s,t}^{state}) = \mu_{state}^{S} + \epsilon_{s}^{S} + \epsilon_{t}^{S}$$

where μ_{state}^{S} is a state-specific intercept, and ϵ_{s}^{S} and ϵ_{t}^{S} are site- and year-specific random effects, respectively, where nest site and year are known given n. All random effects were drawn from normal distributions with mean 0 and unique standard deviations $\boldsymbol{\sigma}$. Shrinkage priors (Simpson et al., 2017) were used for $\boldsymbol{\sigma}$. Overall, the multi-state daily nest survival model provides estimates of the probability of nest success while accounting for nests that failed before they were observed, per Mayfield (1961).

The second component of overall productivity was nest productivity, defined as the number of fledglings produced conditional on nest success (i.e., conditional on fledging at least one chick). We model observations of the number of fledged chicks per successful nest, $\boldsymbol{fl}$, as

$$fl_{n,s,t} \sim Poisson(\lambda_{n,s,t}^{fl})$$

where

$$\log(\lambda_{n,s,t}^{fl}) = \mu^{fl} + \epsilon_{s}^{fl} + \epsilon_{t}^{fl}$$

where μ^{fl} is an intercept, and random effects are analogous to above.

The third component of overall productivity was renesting rate. For nests in territories where parentage was observed and where those territories were intensively monitored such that all nest attempts were likely located, the number of renesting attempts per pair per season ($\boldsymbol{r}$) was modeled as

$$r_t \sim Poisson(\lambda^r_{s,y})$$

where t indexes unique territories and where

$$\log(\lambda^r_{s,y}) = \mu^r$$

where again μ^r is an intercept.

Overall productivity of females is then the product of these three components,

$$\widehat{f_{s,t}} = \frac{1}{2} \cdot \left(\widehat{S^{lay}_{s,t}} \right)^{\theta^{lay}_{mid}} \left(\widehat{S^{inc}_{s,t}} \right)^{\theta^{inc}_{mid}} \left(\widehat{S^{nes}_{s,t}} \right)^{\theta^{nes}_{mid}} \cdot \widehat{\lambda^{fl}_{s,t}} \cdot \widehat{\lambda^{r}_{s,t}}$$

where the first three terms represent the overall probability of nest success at site s in year t, given an estimated average age-in-state, θ^{state}_{mid}, at all nest state transitions. The overall probability of nest success is multiplied by site- and year-specific nest productivity and site- and year-specific renesting rate. The quantity is also multiplied by one half, assuming an equal sex ratio upon fledging.

3.2.3.2 Survival and Dispersal Model

We developed a hierarchical multi-state model of survival and dispersal that facilitates the estimation of annual apparent survival, $\boldsymbol{\phi}$, and site-specific movement probabilities, $\boldsymbol{\psi}$, for both age classes. The state process model is

$$Z_{i,t} | Z_{i,t-1} \sim Categorical(\mathbf{\Psi}_{Z_{i,t-1}})$$

where the states $\boldsymbol{Z}$ correspond to sites $s \in \{1, \ldots S\}$, plus an additional dead state. The vector $\mathbf{\Psi}_{Z_{i,t-1}}$ describes the state-transition probabilities for an individual in state $Z_{i,t-1}$, which are a product of apparent survival and a vector of site-specific movement probabilities, $\boldsymbol{\psi}$, with elements equal to the probability of moving from site s to site r, $\psi_{s,r}$ where, for $s = r$, an individual remains at the same site.

Apparent survival ($\phi_{a,s,t}$) was modeled as

$$logit(\phi_{a,s,t}) = \mu_a^{\phi} + \epsilon_s^{\phi} + \epsilon_t^{\phi}$$

where μ_a^{ϕ} is an age specific intercept, and ϵ_s^{ϕ} and ϵ_t^{ϕ} are site- and year-specific random effects, respectively. Ages represented were fledglings, post-fledglings, and adults (L, HY, and AHY, respectively). Shrinkage priors were used for the variances of all random effects (Table 3.1).

Movement probabilities were modeled as a function of site-specific attractiveness and distance ($\boldsymbol{D}$) between sites. If s represents an individual's site at time $t-1$ and r is an individual's site at time t, then

$$\Psi_{a,s,r} = \begin{cases} \left(logit^{-1}\left(\mu_a^{\psi} + \epsilon_s^{\phi}\right)\right) & \text{if } s = r \\ \left(1 - logit^{-1}\left(\mu_a^{\psi} + \epsilon_s^{\phi}\right)\right) \cdot \dfrac{\exp\left(\alpha^{D} \cdot D_{s,r} + \epsilon_r^{\phi}\right)}{\sum_{s \neq r} \exp\left(\alpha^{D} \cdot D_{s,r} + \epsilon_r^{\phi}\right)} & \text{if } s \neq r \end{cases}$$

That is, an individual remained at their current site, *s*, with probability modeled as a function of age-specific mean site fidelity, μ^{ψ}, and random site-specific attractiveness, ϵ_s^{ϕ}, assuming that more attractive sites are those with higher annual survival. An individual moved to a different site with probability equal to the product of the complement of the probability of staying at site s, and a conditional probability of moving from site s to site r modeled as a function of the distance between the two sites, $D_{s,r}$, and the site-specific attractiveness of site r. Finally, individuals transition to the dead state with probability $1-\phi$, and remain in the dead state with probability 1.

The observation model is that of a standard multi-state model,

$$Y_{i,t}|Z_{i,t} \sim Categorical\left(\mathbf{P}_{Z_{i,t}}\right),$$

where $\mathbf{P}_{Z_{i,t}}$, represents the multi-state detection probability. Individuals are detected at their location with site- and year-specific probability $p_{s,t}$ and not detected with probability $1 - p_{s,t}$. That is, living individuals cannot be detected at any site besides the one at which they are located, and dead individuals cannot be detected. Detection probability was modeled as

$$logit(p_{s,t}) = \mu^{p}_{effort_{s,t}} + \epsilon^{p}_{type_s}$$

where μ^p is an effort-specific intercept and ϵ^p is a site-type-specific random effect. Site-types represented were airfields and native prairies. Site- and year-specific effort levels were categorized as high (intensive nest monitoring occurred, with more than 20 individuals banded), medium (some nest monitoring occurred, with 0-20 individuals banded), low (only count surveys occurred), and zero (site not monitored). Abundance Model

We fit a hierarchical *N*-mixture model (e.g., Royle, 2004) to estimate sex- and site-specific abundance trends from spatially replicated counts. A challenge in estimating sex-specific abundance is the inclusion of data on individuals that are not sexed or aged in the field. To account for these individuals, we follow the method of Keren & Pearson (2019), where

$$Y_{g,s,t,v} \sim Binomial\left(\rho_{s,t}\left(N_{1_{g,s,t}} + N_{AD_{g,s,t}}\right) - U_{g,s,t,v},\ \theta_{g,s,t,v},\right).$$

In this model, $Y_{g,s,t,v}$ represents the count of individuals of sex g at site s in year t at visit v; $\rho_{s,t}\left(N_{1_{g,s,t}} + N_{AD_{g,s,t}}\right)$ represents the number of individuals available to be detected given the proportion of suitable habitat surveyed in a site-year, $\rho_{s,t}$, and age-specific abundances $N_{1_{g,s,t}}$ and $N_{AD_{g,s,t}}$ for first-year birds and adults, respectively; $U_{g,s,t,v}$ represents the unclassified birds that are probabilistically assigned to be adults of sex g in a particular site-year-visit combination, and $\theta_{g,s,t,v}$ represents the detection probability for a particular sex at a particular survey.

Unclassified birds are probabilistically assigned to an age and sex depending on how they were detected, with the consideration that different sexes and ages are not equally likely to be detected visually versus aurally. For instance, males tend to be both more vocal and more visually detectable. We refer readers to Keren & Pearson (2019) for details.

Sex and survey-specific detection rates are modeled as

$$\text{logit}(\theta) = \mu^c$$

where μ^c is an intercept.

3.2.3.3 Population Process

A challenge in producing reliable estimates of SHLA abundance in survival is pervasive partial mark loss. Because of the population size, length of the time series, and probability of mark loss, the scale of the problem is greater than could be addressed by the approach developed in Chapter 2. Here, we leverage all available data in the context of an IPM to estimate an additional parameter, γ, which represents an inflation factor relating apparent survival, $\boldsymbol{\phi}$, to true survival, $\boldsymbol{\phi^{true}}$.

The demographic parameters modeled above are linked to the age-, sex-, and site-specific population size via the stochastic relationship,

$$N'_{1_{g,s,t}} \sim Poisson\left(f_{s,t-1} \cdot \left(N_{1_{2,s,t-1}} + N_{AD_{2,s,t-1}}\right) \cdot \phi_L^{true}\right)$$

$$N'_{AD_{g,s,t}} \sim Binomial\left(N_{1_{g,s,t-1}} + N_{AD_{g,s,t-1}}, \phi_{AHY}^{true}\right)$$

where $N'_{1_{g,s,t}}$ and $N'_{AD_{g,s,t}}$ represents the sex-, site-, and year-level abundance of 1-year-olds and adults, respectively, prior to movement; $f_{s,t-1}$ represents overall productivity as estimated by the productivity model, assuming that productivity is female-limited; and ϕ_1^{true} represents the true

survival probability. Note that $\phi_a^{true} > \phi_a$, for both age-classes, given that marked individuals can lose their field-readable bands and that mark degradation leads to an underestimation of apparent survival (Chapter 2). The relationship between these two parameters is modeled as

$$logit(\phi_a^{true}) = logit^{-1}(\phi_a) + \gamma$$

where γ is a positive correction factor.

Movement is modeled according to

$$\boldsymbol{N}^{\boldsymbol{move}}_{\mathbf{1}_{g,s,t,\cdot}} \sim Multinomial(N'_{1_{g,s,t}}, \boldsymbol{\psi}_{\mathbf{1}_{S,\cdot}})$$

$$\boldsymbol{N}^{\boldsymbol{move}}_{\mathbf{AD}_{g,s,t,\cdot}} \sim Multinomial(N'_{AD_{g,s,t}}, \boldsymbol{\psi}_{\mathbf{AD}_{S,\cdot}})$$

where $\boldsymbol{\psi}_{\mathbf{1}_{S,\cdot}}$ is estimated by the dispersal model. True abundance at site s is then computed by summing all individuals that either remain at or move to site s according to

$$N_{1_{g,s,t}} = \sum_{k=1}^{S} N^{move}_{1_{g,k,t,s}}$$

$$N_{AD_{g,s,t}} = \sum_{k=1}^{S} N^{move}_{AD_{g,k,t,s}}$$

where $N_{1_{g,s,t}}$ and $N_{AD_{g,s,t}}$ represent the number of sex $g \in \{1,2\}$ individuals at site s at time t, for 1-year-olds and adults, respectively, and sexes correspond to males and females, respectively.

3.2.3.4 Model Fitting

The model was fitted in NIMBLE (de Valpine et al., 2017) accessed from program R (R Core Team, 2022). Within NIMBLE, we used a custom distribution modified from the nimbleEcology package (Turek et al., 2016) to avoid directly sampling the latent state histories $\boldsymbol{Z}$. We used discrete filtering of the multi-event likelihood to remove the latent states, and we reduced the

dataset to only include unique detection histories, weighting the likelihood by the frequency with which each detection history appears in the full dataset. In combination, these changes decreased a prohibitively long runtime by over an order of magnitude.

Uninformative priors were used for all parameters. Uninformative priors were also used to initialize the initial population size by age, sex, and site, where upper bounds were drawn from Keren & Pearson (2019). In particular, initial site-specific population sizes were drawn from a uniform categorical distribution with lower bound of 0 individuals per sex and age class and upper bound equal to 100, which is approximately equal to the upper bound of the 95% credible interval for males at the region's most populous site (site A) as reported by Keren & Pearson (2019). Because not all sites were surveyed in the first years of the time series, we also explored the sensitivity of estimated population trend to an alternative initialization structure. In particular, sites C, F, and I were infrequently surveyed and were not surveyed until the latter half of the time series, in large part because abundance was presumed to be low at these sites. Therefore, we also explored the effects of using somewhat restricted priors for initial population size, where upper bounds were again drawn from Keren & Pearson, (2019), but where upper bounds were site-, age, and sex-specific. That is, the population size at each site was initialized between 0 and the maximum upper bound of the site- and sex-specific 95% credible intervals across years where the site was monitored. This approach led to much lower initial abundances for sites C, F, and I.

We also explored the effects of numerous environmental covariates on demographic rates (Appendix A). However, we ultimately did not detect any meaningful effects and therefore did not include any covariates in our IPM.

Models were fit with 3 chains of 100,000 iterations with an adaptation interval of 200. The first 50,000 samples were discarded as burn-in, and the remaining samples were thinned at a

rate of 10 to reduce storage requirements. We evaluated model convergence using visual inspection of traceplots, and we ensured that the Gelman-Rubin statistic ($\hat{R}$; Gelman & Rubin, 1992) was less than 1.1 for all parameters.

3.2.4 *Population Sensitivity*

We examined the sensitivity of site-specific population growth to fecundity, age-specific survival, and state-specific nest survival by computing the correlation coefficient (*r*) between the posteriors for annual growth rate, $\lambda_t = \frac{N_{t+1}^{tot}}{N_t^{tot}}$, and the posteriors for the key demographic rates with annual variation.

3.3 RESULTS

3.3.1 *Demographic Rates*

Results reported here correspond to the population model using uninformative initial site-specific population abundance, unless otherwise noted. Estimates of demographic rates were largely invariant to initialization.

3.3.1.1 Productivity

Mean overall nest success ($\left(\widehat{S_{s,t}^{lay}}\right)^{\theta_{mid}^{lay}} \left(\widehat{S_{s,t}^{inc}}\right)^{\theta_{mid}^{inc}} \left(\widehat{S_{s,t}^{nes}}\right)^{\theta_{mid}^{nes}}$) was 0.394 (95% credible interval: 0.203, 0.573). Overall nest success ranged from 0.357 (0.174-0.502) to 0.416 (0.252-0.604), though variance between sites was generally low (0.102; 0.019, 0.377) (Figure 3.3). Variance between years was higher (ϵ_t^S = 0.198; 0.064, 0.387) than between sites, with overall nest success highest in 2014 (0.472; 0.332, 0.636) and lowest in 2011 (0.363; 0.143, 0.528) (Figure 3.4). Daily nest survival probabilities (ϕ_{state}^S) were lowest during the incubation stage,

0.965 (0.955, 0.973), and highest during the laying stage, 0.969 (0.946, 0.984). The mean number of fledglings per successful nest (λ^{fl}) was 2.846 (2.43, 3.187), and estimates were largely invariant with respect to site and year. The mean number of renesting attempts per pair per year (λ^{r}) was 1.927 (1.802, 2.059). These three components (i.e., probability of nest success, nest productivity, and number of renesting attempts) make up annual fecundity (female fledglings per female per year) which averaged 1.115 (0.556, 1.734) fledglings per female per year. Variation in annual fecundity between sites and years was primarily driven by variation in overall nest success.

3.3.1.2 Survival and Dispersal

Mean apparent survival from fledging to one year of age (ϕ_L) was 0.226 (0.182, 0.294) versus 0.401 (0.236, 0.546) from post-fledging to one year of age (ϕ_{HY}). Mean apparent survival for adults (ϕ_{AHY}) was 0.498 (0.420, 0.596) (Figure 3.5). After correcting for partial mark-loss with inflation factor, γ, true survival rates (ϕ_a^{true}) were between 20 and 40% higher for all age classes. There was moderate variance between sites (σ_s^{ϕ} = 0.334; 0.197, 0.554), but no obvious differences between site types (i.e., airfields vs. prairies) or locations (i.e., on- or off-JBLM; Figure 3.6). Variance between years was low (σ_t^{ϕ} = 0.271; 0.161, 0.442), though 2018 was a particularly low-survival year followed by a particularly high-survival year in 2019 (Figure 3.7).

Adults displayed high breeding site fidelity ($\Psi_{AHY,s,s} = 0.921$; 0.884, 0.959), whereas site fidelity for first-year birds was lower ($\Psi_{L,s,s} = 0.746$; 0.732, 0.791) (Figure 3.8). When dispersal did occur, the probability of moving between sites was somewhat influenced by the distance between sites, with posterior median of α^{D} = -0.017 (-0.045, -0.001).

When resight monitoring effort was high, the mean annual detection probability of marked birds, p, was 0.977 (0.959, 0.988) and was lower at prairie sites than airfield sites. When monitoring effort was low, annual detection probability tended to be under 0.5.

3.3.2 *Abundance*

Abundance trends differed depending on the population initialization strategy. When non-informative priors were used for initial abundance, results suggest that the region-wide population was in decline over the study period. Specifically, the abundance across all occupied sites decreased from 600 (410, 784) in 2010 to 539 (409, 696) in 2020, which represents an average annual growth rate (λ) of 0.989 (0.988, 0.999). However, abundance trends varied between sites (Figure 3.9).

In contrast, when informative priors were used for initial abundance, the population appeared more stable over the time series. Specifically, the abundance across all occupied sites decreased from 437 (330, 533) in 2010 to 539 (389, 725) in 2020, which represents an average annual growth rate (λ) of 1.021 (0.988, 1.052), with the 2019 season being particularly successful. The difference in estimated trend between initialization strategies was driven primarily by the three least-frequently monitored sites (C, F, and I; Figure 3.10), where initial population size was most uncertain.

3.3.3 *Population Sensitivity*

Of the demographic rates, apparent survival (ϕ) was most strongly correlated with λ, with r = 0.94 (0.75, 0.99) for both fledglings (L) and adults (AHY) and the correlation with respect to post-fledgling (HY) survival was not meaningfully different (Figure 3.11). Daily nest survival

probability S^{state} was also moderately correlated with λ, with r = 0.77 (0.28, 0.94) for all nest states.

3.4 DISCUSSION

We combined nest monitoring, mark-resight, and count data from 2010 to 2020 in a multi-site IPM to better understand the population dynamics of an endangered grassland bird in the South Puget Sound region of Washington State. Our analysis suggests that region-wide abundance has decreased slightly over the study period, though this result is sensitive to initial population structure. This sensitivity is driven by uncertainty in the initial population size at three infrequently monitored sites, and therefore our results highlight the importance of continued and regular monitoring at occupied sites throughout the region. Regardless of initialization strategy, we found that site-specific trends are asynchronous and highly variable year over year. We did not observe any systematic differences in population trends between site type or site location (i.e., airfields vs. prairie sites or on- or off-JBLM). Population growth rates were sensitive to both survival and nest survival for all ages and states, respectively, which is consistent with other short-lived passerines (Cox et al., 2014; Sæther et al., 2016). The sensitivity of SHLA growth rates to underlying vital rates suggests that these vital rates could be effective targets for conservation action. These results will inform future research and management priorities, including a plan to reintroduce birds to currently unoccupied sites in South Puget Sound.

Relative to previous estimates of SHLA vital rates, our results are broadly similar, but we find the population to be more stable today than in the early 2000s, when rapid declines were observed (Camfield et al., 2011). Camfield et al. (2011) reported adult site fidelity to be roughly 90%, similar to our estimate. They estimated apparent survival to be approximately 55% for adult males and 47% for females, which is similar to our estimate of apparent survival, though

we did not account for differences in sex. However, given high rates of partial mark loss, our estimate for true survival (i.e., after correcting for partial mark loss) is higher, approximately 60%. We consider this to be true survival because we account for emigration to other sites within the study region, and because observed emigration outside of the study region is near zero (Wolf et al., 2020). A weakness of our model is the fact that we did not account for any differences in survival between sexes, as biased sex-ratios have been observed in other grassland passerines. However, female survival can be challenging to estimate given their lower detection rates compared to males. Our estimate of mean apparent juvenile survival was similar to previous reports (Camfield et al., 2011), though again our estimate of true survival was higher. Fecundity was consistent with previous estimates (Camfield et al., 2011; Pearson & Knapp, 2016). Given population sensitivity to survival rates, it therefore follows that our estimates of population growth rates are greater than previously reported, regardless of initialization strategy. They are also consistent with recent population count data, which show generally stable populations region-wide, though site-specific trends are variable. Taken together, these results support the hypothesis that both demography and population trend have changed (i.e., become more stable) since the last population assessment for this species. However, the future viability of the species remains uncertain, and the isolated subpopulations remain a concern due to their vulnerability to demographic and environmental stochasticity.

A limiting factor for population growth that we have not addressed here is the effect of density-dependence. As population sizes increase, territory size, dispersal, or survival probabilities could change as site-specific carrying capacities are approached. Therefore, the upper bounds of our population projections should be interpreted with caution. Each individual site in the region is small, with sizes ranging from 56 to 309 ha. Over the data collection period,

population sizes were small enough that density-dependent effects were likely negligible (i.e., not all viable habitat was occupied and territories were larger than might otherwise be expected, suggesting that site-specific carrying capacities had not been reached), and these effects, when present, are uncertain. Future work could address this uncertainty, which would be particularly useful in informing an ongoing reintroduction effort. Given the fragmentation of suitable habitat in the region, there are currently few places to establish new breeding sites, though reintroductions are being considered at several currently unoccupied sites that could present more opportunities for region-wide population growth, and it will be important to understand both the carrying capacities of these new sites as well as the potential for dispersal between sites as carrying capacities are reached.

Results from our model suggest several opportunities for population and environmental monitoring that could reduce key uncertainties in drivers of SHLA population dynamics. First, the number of renesting attempts per female per year is not well informed by data. It is difficult to locate all nests per pair per year, so this parameter is likely underestimated in our model. Using telemetry methods on a subset of females over the course of an entire breeding season would provide richer information about this parameter. Second, the overwinter dynamics of SHLA are not well understood (though see Pearson et al., 2005). Many overwintering sites are unknown, as is what constitutes quality wintering habitat and the drivers of mortality during the nonbreeding season. One way to identify overwintering sites would be to use telemetry, and once sites were identified, on-the-ground surveys could better inform risks to SHLA over the entire annual cycle. Finally, habitat management and monitoring throughout the breeding range is very sparse, mainly occurring at one site. Recording habitat characteristics (e.g., % vegetation cover) and other types of management actions (e.g., predator control) more systematically across

monitoring sites and years could facilitate analyses of these as potential predictors of SHLA demographic rates. Information about links between habitat characteristics, management, and demographic rates has tremendous potential to inform future management, the effectiveness of which has been previously demonstrated for SHLA (Pearson & Knapp, 2016).

The widespread decline of grassland birds has been a major cause of conservation concern (Brennan & Kuvlesky Jr., 2005; Pavlacky Jr. et al., 2022), with reductions to habitat availability and quality resulting in small populations on fragmented landscapes. Populations on fragmented landscapes face a unique set of threats, including heightened vulnerability to environmental perturbation and low genetic variation, especially when birds are highly philopatric, as we observed in this study. We introduced a multi-site IPM that facilitates understanding of the dynamics of populations existing on fragmented landscapes. Several additional characteristics, including the ability to deal with complex reproduction data, render this framework broadly useful for a variety of avian species existing in fragmented landscapes.

3.5 ACKNOWLEDGEMENTS

We thank numerous avian field technicians for collecting field data, most notably T. Leque, J. Treadwell, and A. Wolf. Handling and marking of SHLA was approved by Washington Department of Fish and Wildlife through annual scientific collecting permits and USGS Bird Banding Permit 22932. This research was funded through support of A Bratt by the University of Washington College of the Environment, the Washington Cooperative Fish and Wildlife Research Unit, Ecostudies Institute, and by a U.S. Geological Survey Northwest Climate Adaptation Science Center award G17AC000218 to A Bratt. Fieldwork was supported by Joint Base Lewis McChord Military Base (DABJ25-03-H-4002), U.S. Fish and Wildlife Service (F17AP00813), and American Bird Conservancy (#19062); none of the funders required the

approval of this manuscript before submission or publication. The findings and conclusions of the WDFW and USGS authors in the paper are their own and do not necessarily represent the views of the WDFW or United States Geological Survey. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

3.6 FIGURES & TABLES

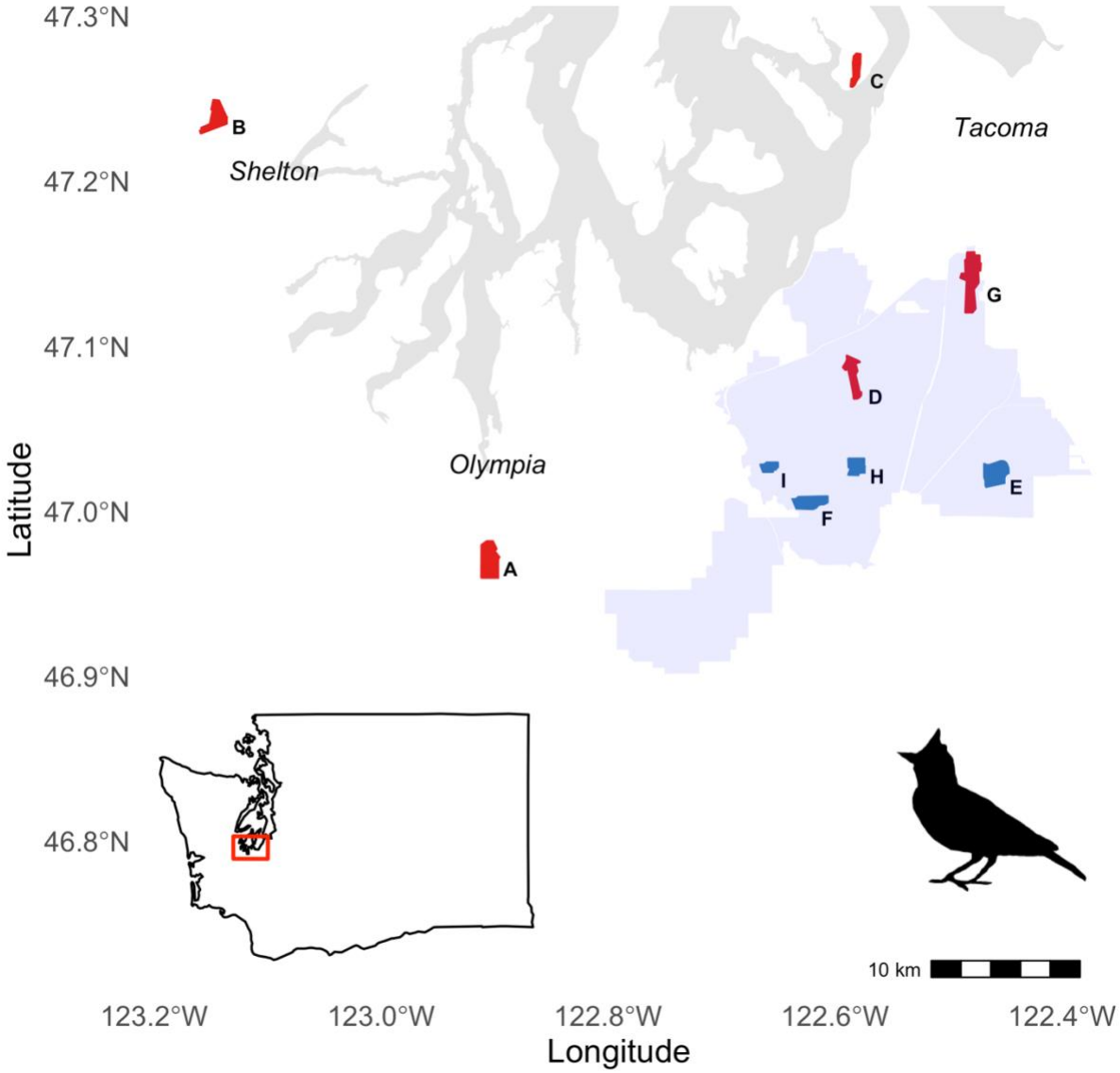

Figure 3.1: Map of study sites in the South Puget Sound region of Washington State, USA. Blue sites are native prairies and red sites are airfields. The purple shaded area represents Joint Base Lewis-McChord.

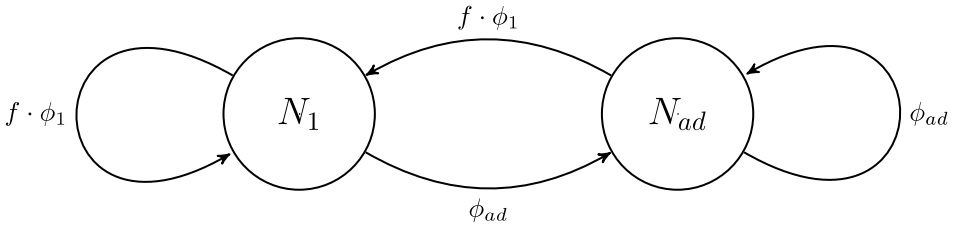

Figure 3.2: Life-cycle diagram for Streaked Horned Larks. Age-classes represented are near one-year-olds (N_1) and older adults (N_{ad}). Both age-classes survive with rate ϕ_{ad}. Both age-classes produce chicks with rate f, who then survive to become near one-year-olds with rate ϕ_1.

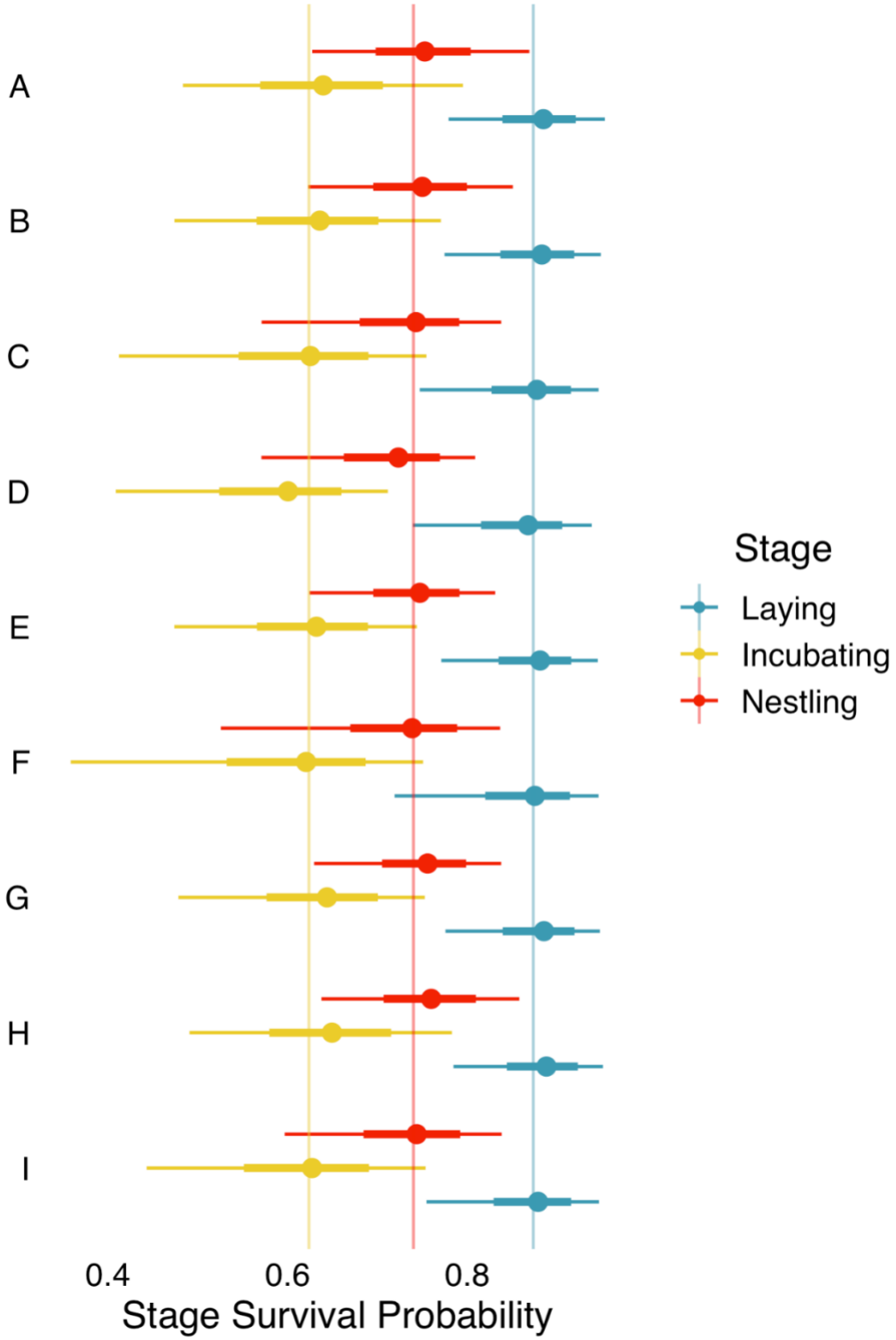

Figure 3.3: Site-specific nest survival probabilities by nest state. Site labels correspond to the site labels in Figure 1. Shown are age-specific medians (points), and 95% credible intervals (line ranges). The vertical lines represent the stage-specific means.

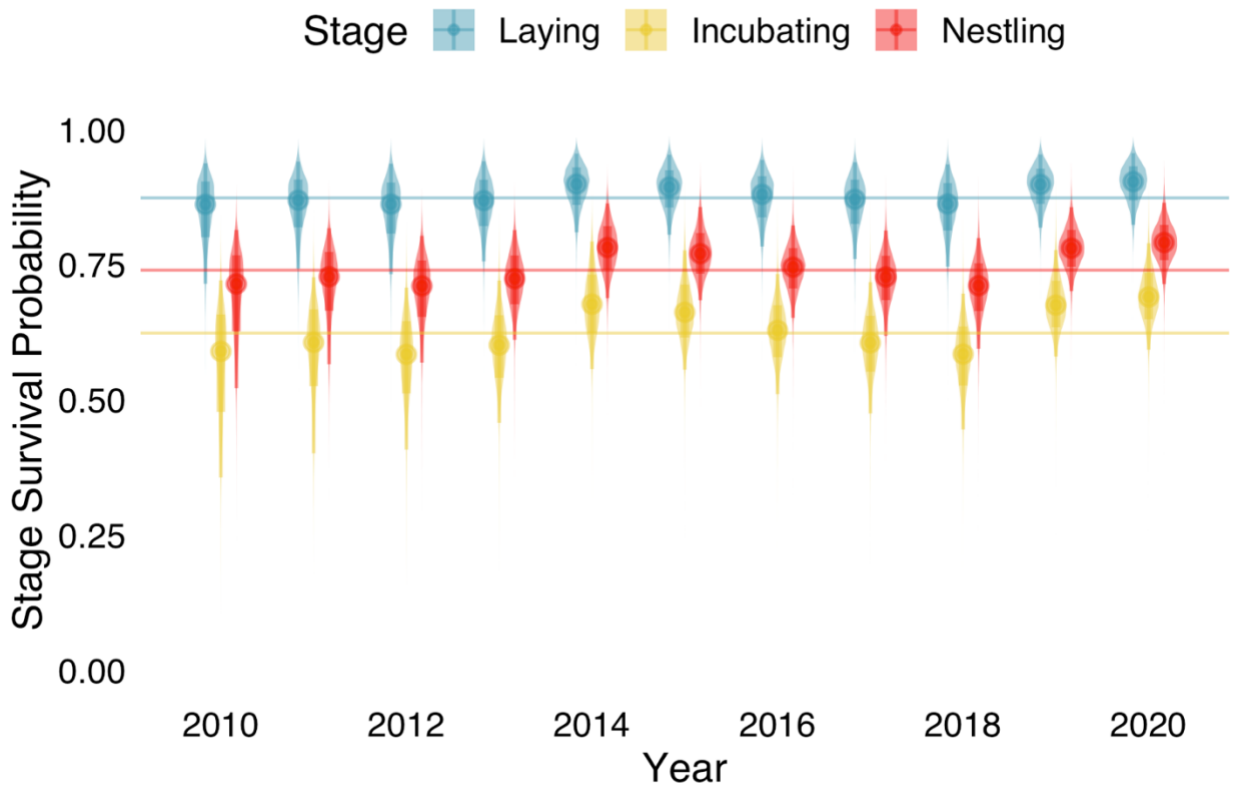

Figure 3.4: Annual nest survival probabilities by nest stage. Posterior distributions are shown along with medians (points), and 95% credible intervals (line ranges). The horizontal lines represent the stage-specific means.

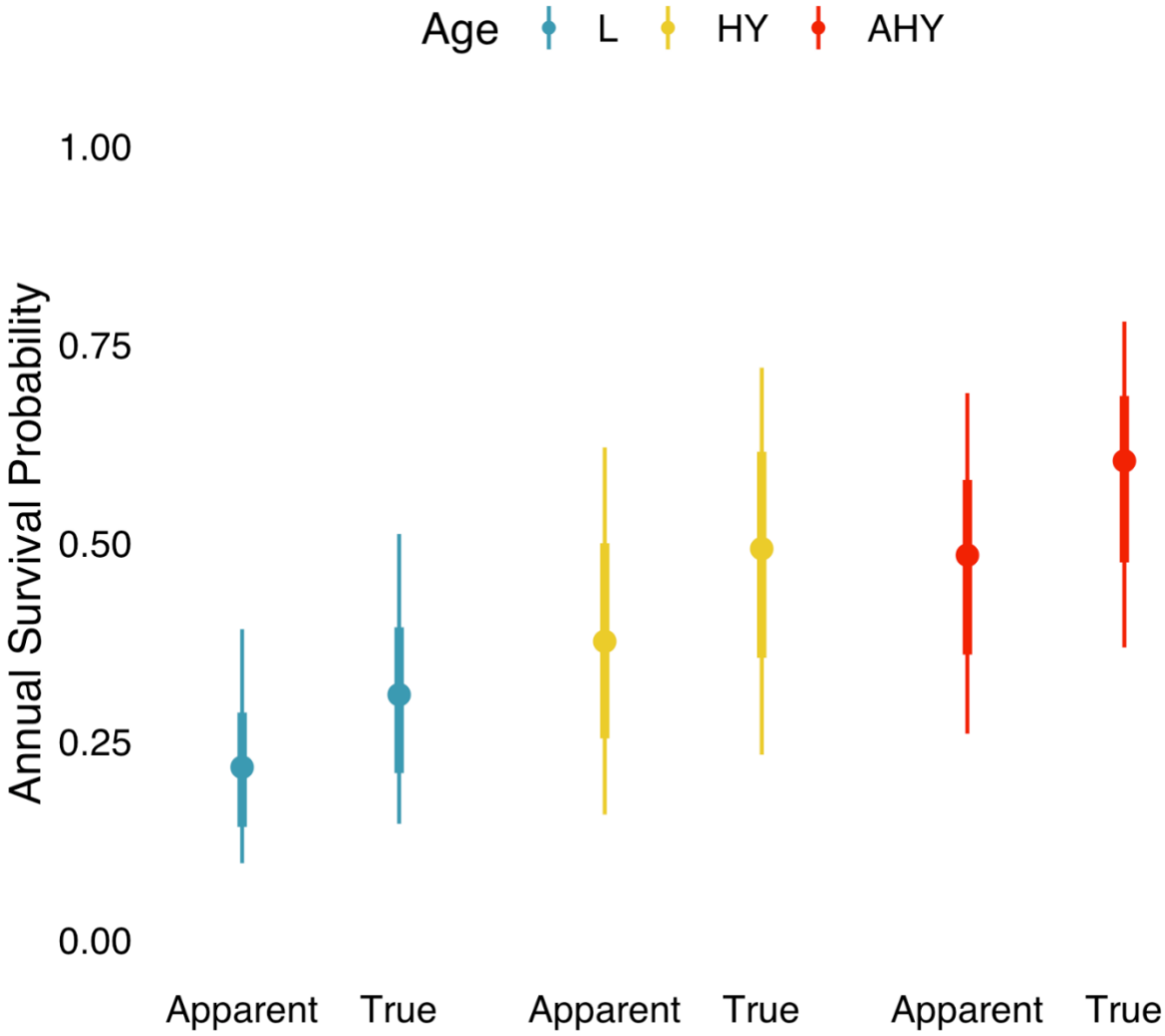

Figure 3.5: Apparent and true survival probabilities by age. Shown are age-specific medians (points), and 95% credible intervals (line ranges).

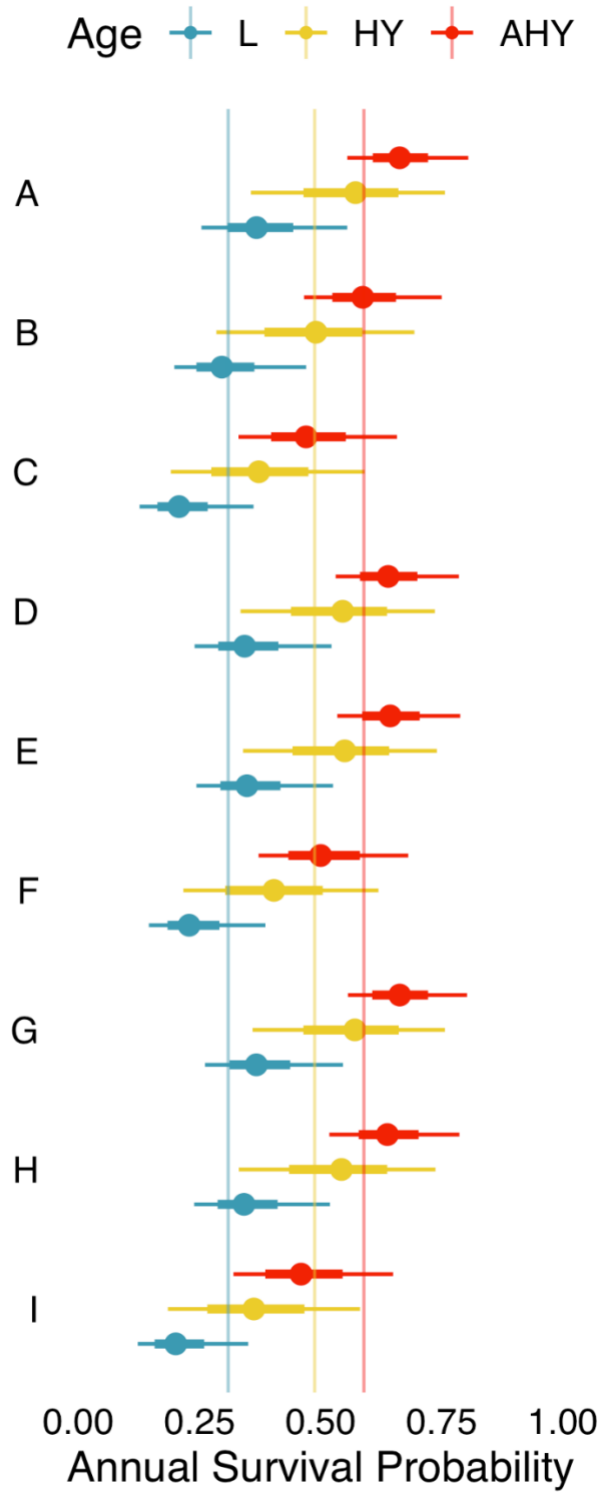

Figure 3.6: Site-specific survival probabilities by age. Site labels correspond to the site labels in Figure 1. Shown are age-specific medians (points), and 95% credible intervals (line ranges). The vertical lines represent the age-specific means.

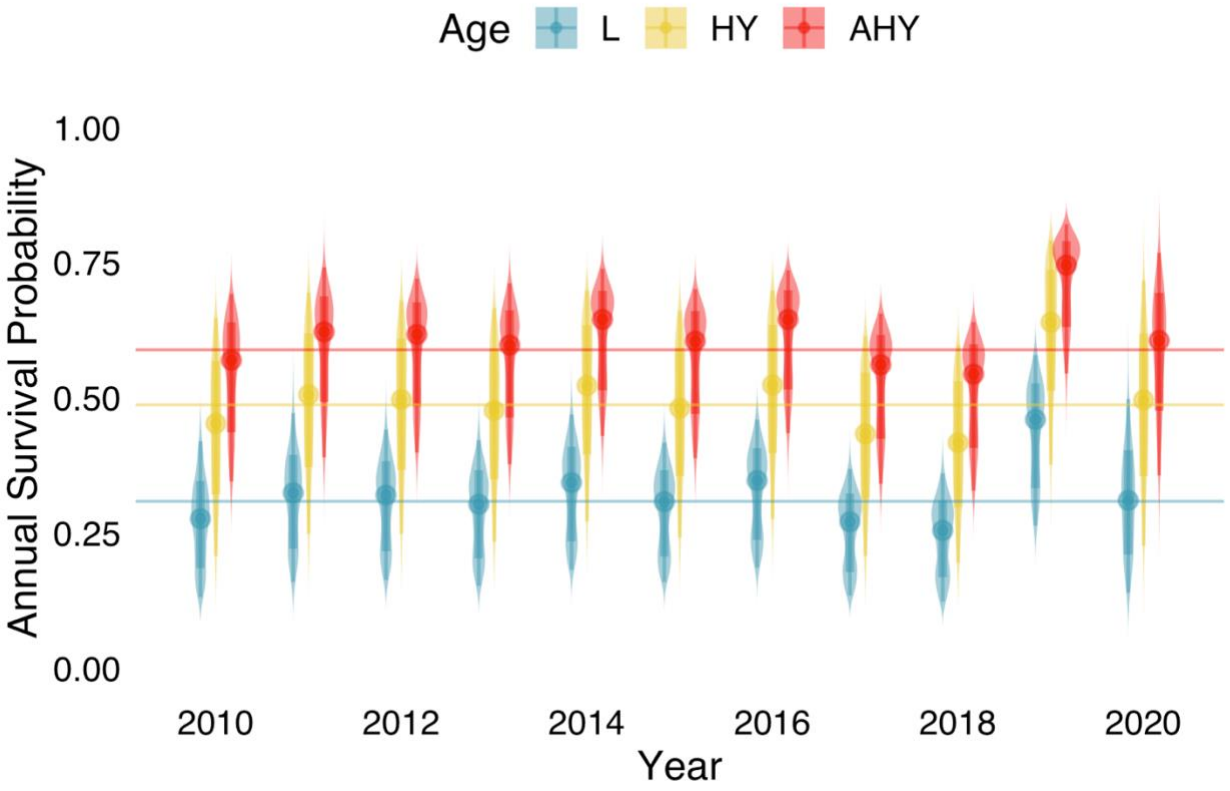

Figure 3.7: Annual nest survival probabilities by age. Posterior distributions are shown along with medians (points), and 95% credible intervals (line ranges). The horizontal lines represent the age-specific means.

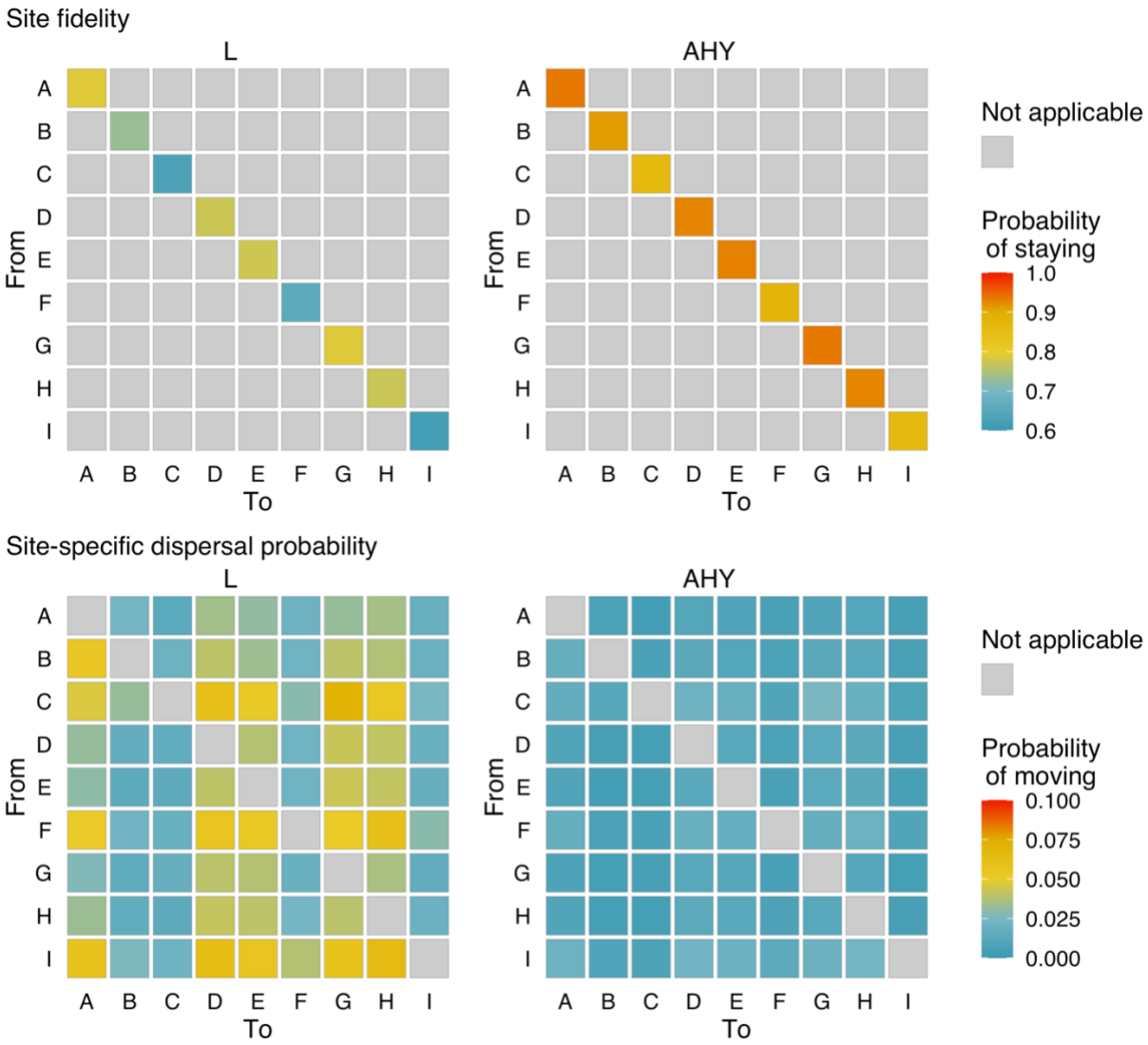

Figure 3.8: Median site- and age-specific probabilities of site fidelity (top) and site- and age specific dispersal probabilities (right) for Streaked Horned Larks in South Puget Sound. Fledglings (L) are shown on the left, compared to adult (AHY) birds on the right. Site labels correspond to the site labels in Figure 1. Post-fledglings (HY) are assumed to disperse at the same rate as fledglings and are therefore omitted from this figure.

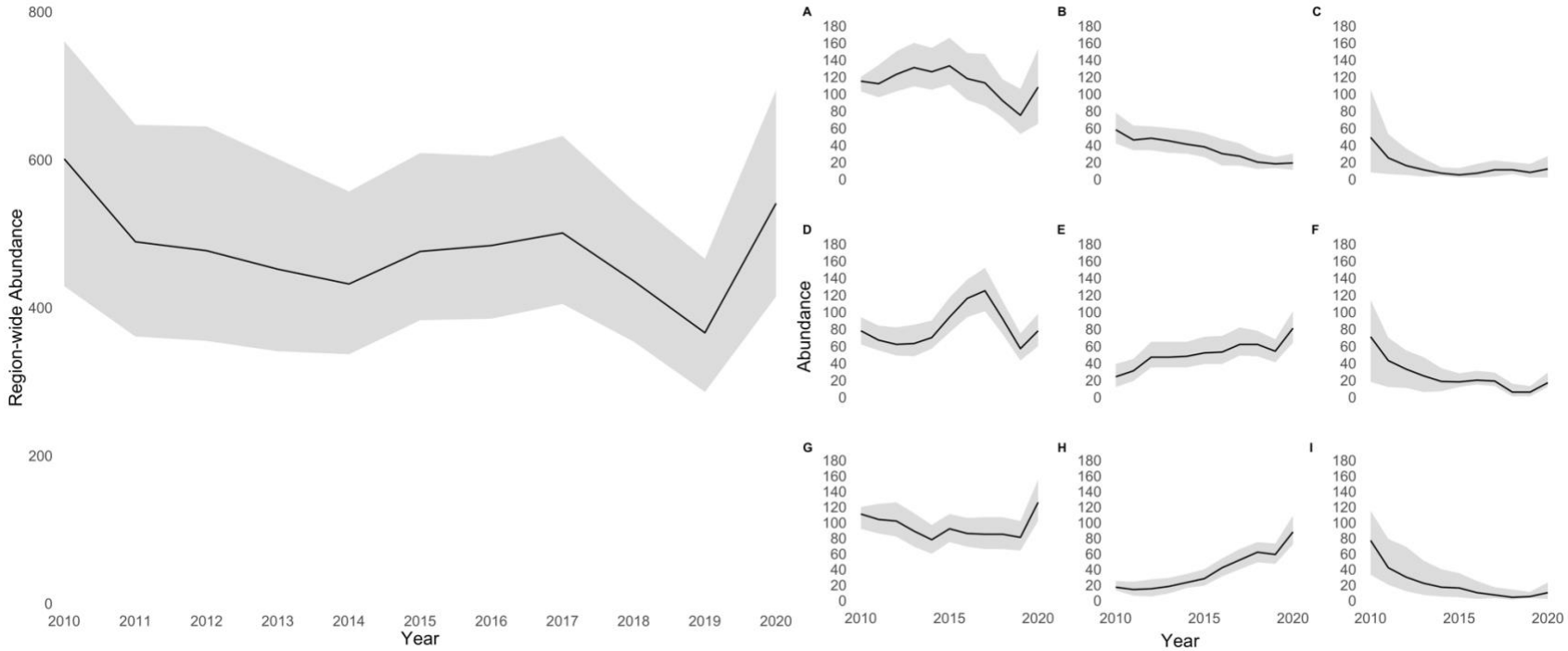

Figure 3.9: Estimated region-wide (left) and site-specific (right) abundance of Streaked Horned Larks at occupied sites in South Puget Sound over the data period, using non-informative priors for site-specific initial abundances. Medians are represented by the bold lines, while 95% credible intervals are represented by the shaded areas. Site labels correspond to the site labels in Figure 1. There is substantial variation in trend between years, and limited synchrony between sites. The resulting trend is substantially different than in Figure 3.10, revealing model sensitivity to initial population size at infrequently monitored sites (i.e., C, F, I).

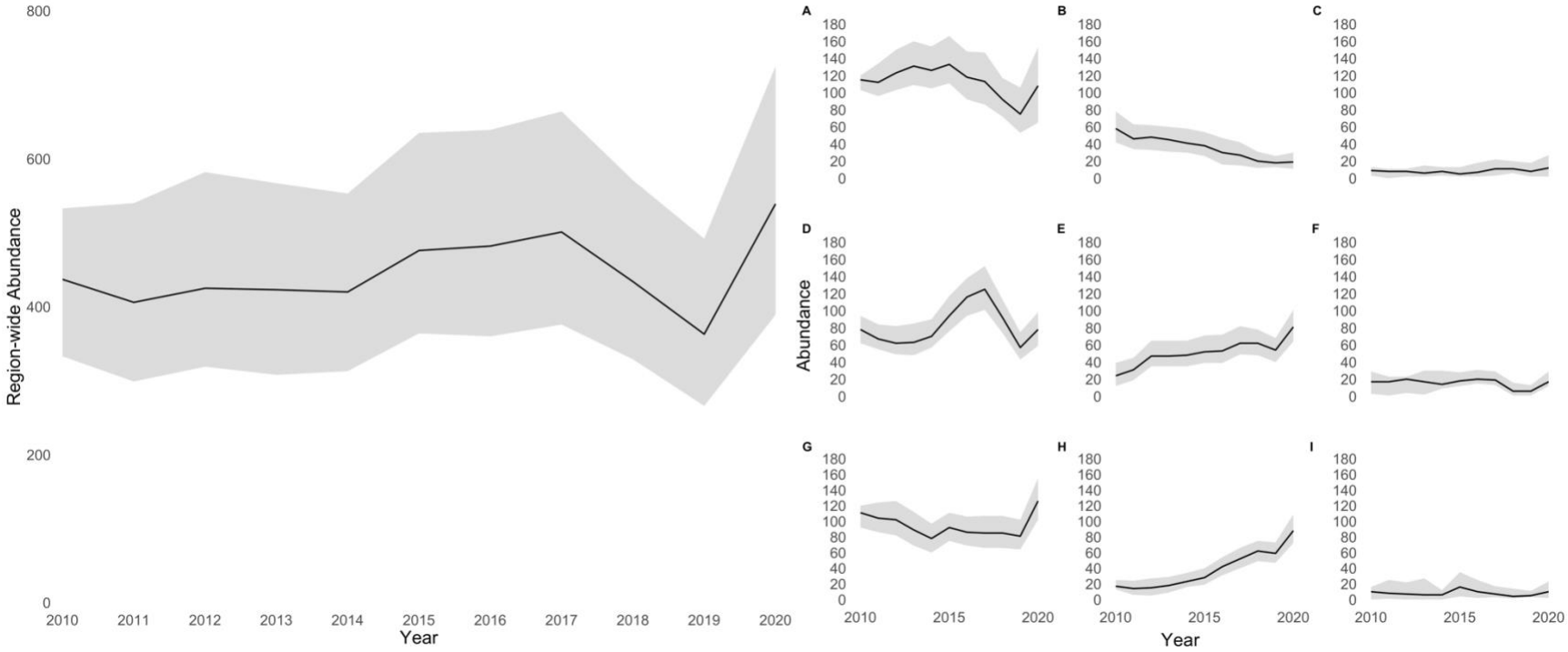

Figure 3.10: Estimated region-wide (left) and site-specific (right) abundance of Streaked Horned Larks at occupied sites in South Puget Sound over the data period, using informative priors for site-specific initial abundances. Medians are represented by the bold lines, while 95% credible intervals are represented by the shaded areas. Site labels correspond to the site labels in Figure 1. There is substantial variation in trend between years, and limited synchrony between sites. The resulting trend is substantially different than in Figure 3.9, revealing model sensitivity to initial population size at infrequently monitored sites (i.e., C, F, I).

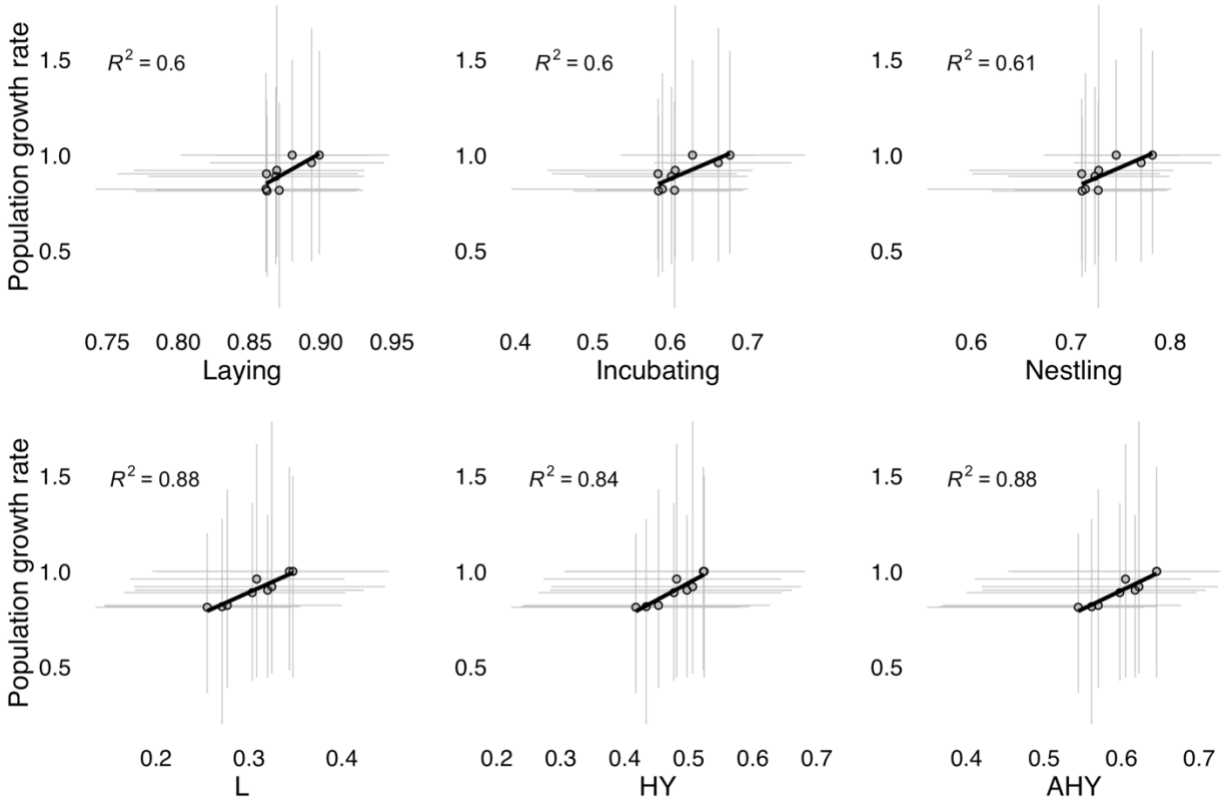

Figure 3.11: Estimated annual population growth rates relative to estimated demographic rates over the data period. Posterior medians are shown (points), with 95% credible intervals (line ranges) for stage-specific nest survival probabilities (top row) and age-specific survival probabilities (bottom row). All parameters show strong correlation with population growth rate, though the correlation is stronger for survival probabilities. There is no substantial difference in correlation between classes for stage-specific or age-specific survival probabilities.

Table 3.1: Model parameters and their priors, with references when informative priors were used.

Parameter	Description	Prior
μ_{state}^{S}	Mean state-specific daily nest survival	N(3.5, 0.5)
σ_s^S	Scale of site effect on daily nest survival	Exp(1)
σ_t^S	Scale of annual effect on daily nest survival	Exp(1)
λ^{fl}	Mean number of fledglings per successful nests	N(1, 0.5)
σ_s^{fl}	Scale of site effect on fledgling rate	Exp(1)
σ_t^{fl}	Scale of annual effect on fledgling rate	Exp(1)
λ^r	Mean number of renesting attempts per female	N(0, 0.5)
μ_a^{ϕ}	Age-specific mean annual survival	N(0, 1.5)
σ_s^{ϕ}	Scale of site effect on annual survival	Exp(1)
σ_t^{ϕ}	Scale of annual effect on annual survival	Exp(1)
α^D	Rate of decay in dispersal probability with distance	N(0, 0.5)
μ_a^{ψ}	Age-specific mean dispersal probability	N(0, 2)
μ_{ef}^{p}	Effort-specific mean mark-resight detection probability	N(0, 1.5)
γ	Annual survival inflation factor	Exp(10)
θ_g	Sex-specific detection probability on count surveys	N(0, 1.5)

3.7 REFERENCES

Abadi, F., Gimenez, O., Ullrich, B., Arlettaz, R., & Schaub, M. (2010). Estimation of immigration rate using integrated population models. *Journal of Applied Ecology*, *47*(2), 393–400.

Ahrestani, F. S., Saracco, J. F., Sauer, J. R., Pardieck, K. L., & Royle, J. A. (2017). An integrated population model for bird monitoring in North America. *Ecological Applications*, *27*(3), 916–924. https://doi.org/10.1002/eap.1493

Altman, B. (2011). Historical and current distribution and populations of bird species in prairie-oak habitats in the Pacific Northwest. *Northwest Science*, *85*(2), 194–222. https://doi.org/10.3955/046.085.0210

Brennan, L. A., & Kuvlesky Jr., W. P. (2005). North American grassland birds: An unfolding conservation crisis? *The Journal of Wildlife Management*, *69*(1), 1–13. https://doi.org/10.2193/0022-541X(2005)069<0001:NAGBAU>2.0.CO;2

Camfield, A. F., Pearson, S. F., & Martin, K. (2011). A demographic model to evaluate population declines in the endangered Streaked Horned Lark. *Avian Conservation and Ecology*, *6*(2). https://doi.org/10.5751/ACE-00467-060204

Chandler, R. B., & Clark, J. D. (2014). Spatially explicit integrated population models. *Methods in Ecology and Evolution*, *5*(12), 1351–1360. https://doi.org/10.1111/2041-210X.12153

Coppedge, B. R., Engle, D. M., Masters, R. E., & Gregory, M. S. (2001). Avian response to landscape change in fragmented southern great plains grasslands. *Ecological Applications*, *11*(1), 47–59. https://doi.org/10.1890/1051-0761(2001)011[0047:ARTLCI]2.0.CO;2

Cox, W. A., Thompson, F. R. I., Cox, A. S., & Faaborg, J. (2014). Post-fledging survival in passerine birds and the value of post-fledging studies to conservation. *The Journal of Wildlife Management. 78(2): 183-193.*, *78*, 183–193. https://doi.org/10.1002/jwmg.670

de Valpine, P., Turek, D., Paciorek, C. J., Anderson-Bergman, C., Lang, D. T., & Bodik, R. (2017). Programming with models: Writing statistical algorithms for general model structures with nimble. *Journal of Computational and Graphical Statistics*, *26*(2), 403–413. https://doi.org/10.1080/10618600.2016.1172487

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

Hamman, S. T., Dunwiddie, P. W., Nuckols, J. L., & McKinley, M. (2011). Fire as a restoration tool in Pacific Northwest prairies and oak woodlands: Challenges, successes, and future directions. *Northwest Science*, *85*(2), 317–328. https://doi.org/10.3955/046.085.0218

Keren, I.N., and S.F. Pearson. 2019. Streaked horned lark abundance and trends for the Puget lowlands and the lower Columbia River/Washington Coast, 2010-2018: Research Progress Report. Washington Department of Fish and Wildlife, Wildlife Science Division, Olympia, Washington.

Mahony, N. A., Dale, B. C., & Miller, D. A. W. (2022). Grassland bird population declines at three Breeding Bird Survey spatial scales in contrast to a large native prairie. *Ecosphere*, *13*(12), e4309. https://doi.org/10.1002/ecs2.4309

Martin, T., & Geupel, G. (1992). Nest-monitoring plots—Methods for locating nests and monitoring success. *J. Field Ornithol*, *1993*.

Mayfield, H. (1961). Nesting success calculated from exposure. *THE WILSON BULLETIN*, *73*(3).

McCrea, R. S., Morgan, B. J. T., Gimenez, O., Besbeas, P., Lebreton, J.-D., & Bregnballe, T. (2010). Multi-Site Integrated Population Modelling. *Journal of Agricultural, Biological and Environmental Statistics*, *15*(4), 539–561. https://doi.org/10.1007/s13253-010-0027-5

Pavlacky Jr., D. C., Green, A. W., George, T. L., Iovanna, R., Bartuszevige, A. M., Correll, M. D., Panjabi, A. O., & Ryder, T. B. (2022). Landscape-scale conservation mitigates the biodiversity loss of grassland birds. *Ecological Applications*, *32*(3), e2548. https://doi.org/10.1002/eap.2548

Mahony, N. A., Dale, B. C., & Miller, D. A. W. (2022). Grassland bird population declines at three Breeding Bird Survey spatial scales in contrast to a large native prairie. *Ecosphere*, *13*(12), e4309. https://doi.org/10.1002/ecs2.4309

Martin, T., & Geupel, G. (1993). Nest-monitoring plots—Methods for locating nests and monitoring success. *J. Field Ornithol*.

Mayfield, H. (1961). Nesting success calculated from exposure. *The Wilson Bulletin*, *73*(3).

Pearson, S.F., M. Hopey, W. D. Robinson, R. Moore. 2005. Range, Abundance and Movement Patterns of Wintering Streaked Horned Larks (*Eremophila alpestris strigata*) in Oregon and Washington. Natural Areas Program Report 2005-2. Washington Dept. of Natural Resources. Olympia, WA.

Pearson, S. F., & Knapp, S. M. (2016). Considering Spatial Scale and Reproductive Consequences of Habitat Selection when Managing Grasslands for a Threatened Species. *PLOS ONE*, *11*(6), e0156330. https://doi.org/10.1371/journal.pone.0156330

Pearson, S.F., M. Linders, I. Keren, H. Anderson, R. Moore, G. Slater, and A. Kreager. 2016. Survey protocols and strategies for assessing streaked horned lark site occupancy status,

population abundance, and trends. Wildlife Science Division, Washington Department of Fish and Wildlife, Olympia, Washington.

Pearson, S. F., Moore, R., & Knapp, S. M. (2012). Nest exclosures do not improve Streaked Horned Lark nest success: Nest Exclosures and Nest Success. *Journal of Field Ornithology*, *83*(3), 315–322. https://doi.org/10.1111/j.1557-9263.2012.00381.x

R Core Team. (2022). *R: A language and environment for statistical computing* [R Foundation for Statistical Computing].

Royle, J. A., & Dorazio, R. M. (2008). *Hierarchical modeling and inference in ecology: The analysis of data from populations, metapopulations and communities*. https://doi.org/10.1016/B978-0-12-374097-7.50001-5

Sæther, B.-E., Grøtan, V., Engen, S., Coulson, T., Grant, P. R., Visser, M. E., Brommer, J. E., Rosemary Grant, B., Gustafsson, L., Hatchwell, B. J., Jerstad, K., Karell, P., Pietiäinen, H., Roulin, A., Røstad, O. W., & Weimerskirch, H. (2016). Demographic routes to variability and regulation in bird populations. *Nature Communications*, *7*, 12001. https://doi.org/10.1038/ncomms12001

Saracco, J. F., & Rubenstein, M. (2020). Integrating broad-scale data to assess demographic and climatic contributions to population change in a declining songbird. *Ecology and Evolution*, *10*(4), 1804–1816. https://doi.org/10.1002/ece3.5975

Schaub, M., & Abadi, F. (2011). Integrated population models: A novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology*, *152*(S1), 227–237. https://doi.org/10.1007/s10336-010-0632-7

Schaub, M., Gimenez, O., Sierro, A., & Arlettaz, R. (2007). Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. *Conservation Biology*, *21*(4), 945–955. https://doi.org/10.1111/j.1523-1739.2007.00743.x

Sipe, H.A., Slater, G.L., & S.J. Converse. *In prep*. Decision analysis for Streaked Horned Lark (*Eremophila alpestris strigata*) reintroduction.

Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, *32*(1), 1–28. https://doi.org/10.1214/16-STS576

Stanton, R. L., Morrissey, C. A., & Clark, R. G. (2018). Analysis of trends and agricultural drivers of farmland bird declines in North America: A review. *Agriculture, Ecosystems & Environment*, *254*, 244–254. https://doi.org/10.1016/j.agee.2017.11.028

Turek, D., de Valpine, P., & Paciorek, C. J. (2016). Efficient Markov chain Monte Carlo sampling for hierarchical hidden Markov models. *Environmental and Ecological Statistics*, *23*(4), 549–564. https://doi.org/10.1007/s10651-016-0353-z

Warlick, A. J. (2022). *Understanding the effects of environmental variability on demography in species with complex life histories through integrated population modeling.* https://www.proquest.com/openview/53058a6573d69b1146e276690115c81f/1.pdf?pq-origsite=gscholar&cbl=18750&diss=y

Weegman, M. D., Arnold, T. W., Dawson, R. D., Winkler, D. W., & Clark, R. G. (2017). Integrated population models reveal local weather conditions are the key drivers of population dynamics in an aerial insectivore. *Oecologia*, *185*(1), 119–130. https://doi.org/10.1007/s00442-017-3890-8

Weegman, M. D., Bearhop, S., Fox, A. D., Hilton, G. M., Walsh, A. J., McDonald, J. L., & Hodgson, D. J. (2016). Integrated population modelling reveals a perceived source to be a cryptic sink. *Journal of Animal Ecology*, *85*(2), 467–475. https://doi.org/10.1111/1365-2656.12481

Wolf, A. L., Slater, G. L., Pearson, S. F., Anderson, H. E., & Moore, R. (2020). Range-wide patterns of natal and breeding dispersal in the streaked horned lark. *Northwest Science*, *94*(1), 31–43. https://doi.org/10.3955/046.094.0103

Chapter 4. QUANTIFYING THE EFFECT OF BYCATCH MITIGATION EFFORTS ON THE POPULATION DYNAMICS OF A LONG-LIVED SEABIRD

Publication history: This study was co-authored with Steffen Oppel, Sarah Converse, Oliver Yates, Catharine Horswill, Alexander Bond, Richard Cuthbert, John Cooper, and Peter Ryan. At the time this dissertation was published, this chapter was not in review with a journal.

Abstract: Bycatch in commercial fisheries is widely considered to be one of the greatest threats to seabird populations globally and considerable efforts have been made to reduce bycatch through implementation of mitigation measures. However, given the challenges of monitoring seabirds, the effect of bycatch on survival probabilities is unknown for many species, as is the extent to which mitigation actions mediate this effect. Atlantic Yellow-nosed Albatross (*Thalassarche chlororhynchos*), which are listed as Endangered on the IUCN Red List, are frequently bycaught in longline and demersal trawl fisheries in the South Atlantic. A portion of the species' second-largest breeding colony, located on Gough Island in the Tristan da Cunha archipelago, has been monitored since 1982. In this analysis, we integrated mark-resight data from 1985-2020 with nest and population count data from 2008-2020. We used a Bayesian integrated population model to improve our understanding of demography and estimate the effects of bycatch mitigation rates in four key fisheries on adult and juvenile survival. Our results suggest that mean juvenile and adult survival are both high (0.89 and 0.92, respectively) and that population growth rate was not as strongly correlated with survival as it was with fecundity. We found that after a period of decline, the total population has been increasing since 2014 while the breeding population has remained relatively stable over the time series. We did not detect an

effect of bycatch mitigation in any fishery on the survival of either age class. We caution against interpreting this result as suggesting that there is no effect of bycatch or mitigation on population dynamics, but rather that the data are currently insufficient to detect any effect or the effect is small. We also conducted a population viability analysis under a status-quo scenario, and we predicted that the population will gradually increase. The predicted extinction risk in 2070 was near 0. The results from this analysis can inform future improvements to population monitoring, bycatch tracking, and mitigation regulations. More generally, the model we introduced can be used in the future to reevaluate recovery for this species, and this framework can be applied more broadly for difficult-to-monitor species vulnerable to cryptic threats.

Keywords: Integrated population model, Bayesian hierarchical model, hidden Markov model, population viability, seabird bycatch, bycatch mitigation, Atlantic Yellow-nosed Albatross, *Thalassarche chlororhynchos*

4.1 INTRODUCTION

Understanding the effects of environmental and anthropogenic stressors on population dynamics can inform the development of effective wildlife conservation efforts. However, knowledge of population dynamics is often difficult to obtain, particularly for populations that are declining, distributed over large areas, or are otherwise difficult to survey (Zipkin & Saunders, 2018). Procellariforms (albatrosses, petrels, and shearwaters) exemplify each of these difficulties. Of 147 procellariform species, 86 are listed as near threatened, vulnerable, endangered, or critically endangered on the IUCN Red List (IUCN 2022). These species spend most of their lives at sea and generally can be monitored only on their often-remote breeding islands.

Albatrosses, the largest of the procellariforms, are also the most threatened. All but 1 of the 22 albatross species are listed as near threatened, vulnerable, endangered, or critically endangered (IUCN 2022). Because of their life-history, albatrosses are vulnerable to cryptic threats such as broad-scale environmental change and anthropogenic stressors. Changes to prey distribution and phenology (e.g., Croxall et al., 2012), ocean warming (Ventura et al., 2021), plastic ingestion (Roman et al., 2021), introduced mammalian predators (Davies et al., 2015), and bycatch in commercial fisheries (Véran et al., 2007) have all been demonstrated to have negative effects on albatross vital rates. However, the effects of these threats can be challenging to estimate because of uncertainty in the underlying population dynamics and because the threats themselves are poorly understood.

Atlantic Yellow-nosed Albatross (AYNA; *Thalassarche chlororhynchos*) breed only on the Tristan da Cunha archipelago, including on Gough Island, which is centrally located in the South Atlantic Ocean. On Gough Island, AYNA have been intensively monitored since 1982 (Figure 4.1; Cuthbert & Sommer 2004). Though the current population abundance and trend are uncertain, previous population modeling efforts found that the Gough Island population has declined at about 1% per year since 1982 (Cuthbert et al. 2003). The same study found that observed decreases were most likely caused by low adult and juvenile survival, one possible driver of which is bycatch in commercial fisheries.

AYNA face a number of threats, some of which are specific to Gough and some of which occur range-wide. On Gough, introduced house mice (*Mus musculus*) are a significant source of chick and even adult mortality for many of the island's seabirds (Cuthbert & Hilton, 2004). Fortunately, AYNA appear to be less affected by mouse predation than other congeners (e.g., Tristan Albatross), possibly because they breed in the austral summer when other food sources

are more plentiful (Cuthbert et al., 2013). Of greater concern is the effect of AYNA bycatch in commercial fisheries (Cuthbert et al., 2013). In longline fisheries, birds are attracted to baited hooks, where they may be hooked and then drown (Bugoni et al., 2008; Tuck et al., 2003; Yeh et al., 2013; Zhou et al., 2019). In trawl fisheries, birds may strike the cables supporting nets, which can lead to injury and death (da Rocha et al., 2021; Watkins et al., 2008). Observer data from the South Atlantic suggest that AYNA, and particularly juvenile AYNA, are frequently bycaught (Yeh et al., 2013). However, the degree to which bycatch has impacted overall AYNA survival probabilities or population trend remains unknown.

Outside of seasonal and spatial fishing restrictions, vessel-level bycatch mitigation measures include night-setting, using bird-scaring lines, and weighting or shielding hooks (Gilman et al., 2022; Kroodsma et al., 2023; Paterson et al., 2019). These measures can decrease seabird bycatch by up to 98% (da Rocha et al., 2021). Recently, there has been considerable effort from a group of international seabird experts, the Albatross Task Force (BirdLife International, Royal Society for the Protection of Birds), to substantially reduce bycatch by working with fishing fleets to encourage use of mitigation measures and with governing bodies to change gear regulations. These efforts have resulted in large reductions in albatross bycatch in some fleets, including one fleet of significance to AYNA, the Namibian fleet (da Rocha et al., 2021), though it is not well understood how reductions in individual mortality have translated to the population-level.

Quantifying the effects of seabird bycatch as well as bycatch mitigation on AYNA populations is challenging for several reasons. First, our knowledge of AYNA demography is relatively poor given the logistical difficulties associated with accessing and monitoring AYNA in the Tristan da Cunha archipelago, including on Gough (R. Cuthbert et al., 2003). Second, the

at-sea distributions of AYNA, and therefore the fisheries with which they interact, are not well understood (though see Dias et al., 2017). Third, fishery-specific estimates of bycatch are generally not available. Observer coverage on vessels in the South Atlantic is sparse, and even when observers are present, bycaught seabirds may not be identified to species (Zhou et al., 2019b). In the absence of observer data, bycatch is sometimes estimated as a function of fishing effort (e.g., birds/millions of hooks; da Rocha et al., 2021; Paterson et al., 2019; Robertson et al., 2018), but metrics of fishing effort have not been consistently reported for all fisheries in the South Atlantic (Tuck et al., 2003). Finally, no data are available with respect to the use of mitigation measures in fisheries where use is opt-in (i.e., international waters) or where compliance rates are unknown.

Integrated population modeling is a useful framework for getting the most out of all available demographic data. Integrated population models (IPMs) are a class of hierarchical population models that leverage multiple data sources to produce estimates of demographic rates and states (Besbeas et al., 2002; Brooks et al., 2004; Schaub & Abadi, 2011; Zipkin & Saunders, 2018). IPMs allow key demographic rates and states to be jointly estimated, offering numerous benefits. First, IPMs can produce more precise and less biased estimates of demographic rates than possible from disparate analyses of the same data (Schaub et al., 2007). Second, it is possible to estimate parameters that would otherwise be unidentifiable (e.g., recruitment or immigration; Abadi et al., 2010). And third, Bayesian IPMs allow for the seamless propagation of uncertainty. IPMs are particularly useful for understanding the population dynamics of species with complex life-histories that are otherwise challenging to monitor (Lawson et al., 2022). IPMs often form the bases of population viability analyses (PVAs; Oppel et al., 2022; Rosenblatt et al., 2021; Saunders et al., 2018), with the goal of quantifying the degree of threat faced by a

population, given the current state of knowledge of that population. Given that they make the most of all available information and appropriately represent uncertainty, IPM-based PVAs can be useful for managers who need to make robust decisions to manage populations of at-risk species.

Here, we evaluate the impacts of fishing bycatch on AYNA at Gough Island. We hypothesized that bycatch mitigation has a positive effect on AYNA survival and population trends. Of the four fisheries considered (i.e., International Commission for the Conservation of Atlantic Tunas (ICCAT), Namibia, South Africa, Uruguay), we hypothesized that mitigation in the ICCAT and Namibian fleets have the greatest effect, based on the relative size of these fisheries and the degree of spatial overlap with AYNA foraging ranges (Dias et al., 2017). We also hypothesized that bycatch avoidance is a learned skill and that the benefits of mitigation are therefore larger for juveniles than for adults (Gianuca et al., 2017). We evaluated our hypotheses by modeling the effect of estimated rates of mitigation implementation within key fisheries on AYNA survival within an IPM. We then built an IPM-based PVA to project the AYNA population into the future. Our results suggest that the AYNA population is stable to slightly increasing. Overall, our analysis introduces a model that can be used to reevaluate population trends in the future, and a framework that can be broadly applied to species with complex life-histories that are difficult to monitor and vulnerable to cryptic threats.

4.2 METHODS

4.2.1 *Study System and Species*

Gough Island is part of the UK Overseas Territory of Tristan da Cunha and is located in the south-central Atlantic Ocean (Figure 4.1). It is an approximately 65 km^2 volcanic island with

steep, mountainous terrain. Across its varied habitats, Gough supports approximately 20 breeding seabird species and is considered one of world's most important islands for seabirds (Caravaggi et al., 2019; R. J. Cuthbert & Sommer, 2004; Holmes et al., 2019; Wanless et al., 2007) and it hosts the second largest breeding population of AYNA. Because of its remoteness and difficult terrain, seabird population monitoring on Gough is challenging. The study area on the southeastern portion of the island is relatively flat and open, a preferred habitat for AYNA. This area constitutes approximately 10% of the suitable breeding habitat for AYNA on Gough and has been monitored since 1982, with more consistent effort applied since 2008 (Cuthbert & Sommer, 2004).

AYNA are colonial breeders, nesting on islands in the mid-Atlantic; the largest breeding island is Tristan da Cunha, which hosts approximately 50% of the global breeding population, while Gough is the second largest, with 10% of the global breeding population (Cuthbert et al., 2003). Unlike many albatrosses, particularly the larger species, AYNA can breed annually, with pairs producing a single egg at each breeding attempt. While breeding, AYNA generally forage close to the colony within international waters, or off the southwestern coast of Africa. Nonbreeding birds may forage further from the colony, as far as southeastern South America, though nonbreeding distributions demonstrate significant overlap with estimated breeding distributions (da Rocha et al., 2021).

4.2.2 *Population Monitoring Data*

AYNA monitoring on Gough consists of counting and marking individuals within the study area over the course of the breeding season and, when possible, recording breeding status. These monitoring data are summarized in three different datasets, described below.

4.2.2.1 Nest Monitoring Data

Breeding success has been recorded for nests within the Gough Island study area for several decades. We model data from 2008 onward, which represent a period with greater survey effort and data quality. Monitoring was conducted per the recommendations of Cuthbert and Sommer (2004). In this protocol, surveys of AYNA consisted of weekly visits to a subset of the study area during the early breeding season (mid-September thru mid-October) to mark and identify active nests. After egg-laying was completed, nests were visited roughly biweekly to determine status (i.e., egg, chick, or empty), and visits continued until all nest fates (i.e., fledged or failed) could be determined.

Data were summarized at the season level to nest and fledgling counts within the study area. Because breeding females are monogamous and produce one egg in one nest per year, we define fecundity as the average number of chicks produced per female per year. Previous analyses (e.g., Cuthbert et al., 2003) have modeled overall fecundity as the product of hatching success and, conditional on hatching, fledging success, where these rates are estimated separately. Here, we favor a simpler model due to interannual variability in nest-check frequency and inconsistent assignment of nestlings and fledglings to nests.

4.2.2.2 Mark-resight Data

AYNA have been marked within the study area and resighted opportunistically on nest-monitoring and abundance surveys since 1982. During the egg-laying season (i.e., Sept-Oct), the study area was visited approximately once a week to mark or identify both individuals in each breeding pair. After eggs were laid, the study area was visited approximately biweekly until nests either fledged or failed. In March, chicks within the study area were marked just prior to fledging. Individuals were banded with metal Monel bands (Mechaniska, Bankyerd, Sweden and

Lambournes, Leominster, UK), and, prior to 2001, unique combinations of field-readable colored bands. After 2001, the colored bands were replaced by field-readable alpha-numeric Darvic bands (ProTouch, Saskatoon, Saskatchewan, Canada). In the rare event of band loss, individuals could be identified by the unique number on their metal bands and these individuals were targeted for recapture to replace their field-readable bands.

Ideally, records of resighted birds would also include their annual breeding state, where observable states of interest are 1) immature (i.e., pre-breeder); 2) breeding; and 3) loafing. In practice, it is difficult to distinguish between these states on a single visit unless the resighted individual is actively incubating or tending a chick, and inconsistency in data recording practices over the time series make breeding state difficult to resolve even at the season-level. For the majority of the time-series, resights of individuals are summarized to detected and non-detected events at the season-level, though breeding state has been recorded with greater accuracy since 2019, with the considered states listed above. Given the frequency of visits to the study area throughout the breeding season, we assume that annual detection of marked individuals is near 1, given that movements outside of the study area were rare (Cuthbert et al., 2003).

4.2.2.3 Population Count Data

Count surveys have been conducted within 11 sub-areas of the Gough Island study area since 1982. However, not all sub-areas of the study area have been surveyed in all years, nor were the boundaries of the sub-areas consistently defined. Therefore, we restricted the count time series considered here to 2008 onward, when count surveys have been conducted over the entire study area. During this period, count surveys were conducted multiple times per breeding season, each capturing different components of the population. In mid-October breeding birds were counted, in early January chicks were counted, and in mid-March fledglings were counted. Here, we focus

on the mid-October counts of breeding birds aggregated over the entire study area (i.e., all 11 sub-areas).

4.2.3 *Bycatch Mitigation Data*

To evaluate the effects of bycatch mitigation efforts, we focused on four key longline and demersal trawl fisheries in the South Atlantic that target areas also used by foraging AYNA (Dias et al., 2017): 1) the Namibian fleet; 2) the South African fleet; 3) the Uruguayan fleet; and 4) the international South Atlantic tuna fleet (ICCAT). Most metrics related to bycatch mitigation are at the fishery level (e.g., birds killed per unit effort, across all bycaught species) but we were interested in determining the effect of region-wide bycatch mitigation on AYNA survival probabilities. Unfortunately, metrics of bycatch mitigation effort are not currently available from regional fisheries themselves, nor are fishing effort metrics available for all fisheries of interest over the time series. Instead, we elicited annual estimates of the proportion of the four key fleets that have implemented bycatch mitigation measures from subject-matter experts at the Albatross Task Force (Figure 4.3). For each fishery of interest, experts reported annual estimates of the mean proportion of the fleet that had used bycatch mitigation measures (e.g., night fishing, bird-scaring lines, line weighting) that year.

4.2.4 *Statistical Modeling*

We developed a female-based IPM that integrated the nest monitoring data (i.e., nest and fledgling counts) and population counts from 2008 to 2020 with mark-resight data from 1985-2020. Using the longer time series for the mark-resight data allowed for more precise estimates of annual survival probabilities and the age-at-recruitment curve. The other two datasets were

restricted to 2008-2020 to ensure a greater degree of reliability in the data, given historically variable monitoring effort and data recording practices.

4.2.4.1 Productivity Model

To estimate AYNA productivity, annual fledgling counts, F_t, were modeled as

$$F_t \sim Binomial(B_t, f_t)$$

where B_t represents annual nest counts, and f_t represents the annual fecundity probability. Annual fecundity was modeled as

$$logit(f_t) = \mu^f + \epsilon_t^f$$

where μ^f is the average fecundity rate and ϵ_t^f is an annual random effect drawn from a normal distribution with mean 0 and standard deviation σ^f. A shrinkage prior (Simpson et al., 2017) was used for σ^f (Table 4.1), both to prevent overfitting and to shrink insignificant effects to near zero while maintaining true large effects.

4.2.4.2 Mark-resight Model

To estimate age-specific survival, and transitions between breeding states, we implemented a multi-event model for the mark-resight data. Given inconsistencies in the accuracy and specificity with which breeding state was recorded over time, the multi-event framework is preferred because the observed events are not always equivalent to the true state. The true state process was modeled as

$$Z_{i,t} \sim Categorical(\mathbf{\Psi}_{Z_{i,t-1},i,t})$$

where the state of individual i in year t was conditional on the individual's state in the previous year and was categorically distributed according to the transition matrix $\mathbf{\Psi}$. States were immature

(i.e., pre-breeding), at-sea non-breeder, non-breeding loafer (i.e., on the breeding colony but not breeding), breeding, and dead. For individual i in year t, the matrix $\boldsymbol{\Psi}$ was

$$\boldsymbol{\Psi}_{i,t} = \begin{array}{c} \\ Immature \\ At\ sea \\ Loafer \\ Breeder \\ Dead \end{array} \begin{pmatrix} \phi_{i,t}(1-\rho_{i,t}) & 0 & 0 & \phi_{i,t}\rho_{i,t} & 1-\phi_{i,t} \\ 0 & \phi_{i,t}\alpha & \phi_{i,t}(1-\alpha)(1-\gamma) & \phi_{i,t}(1-\alpha)\gamma & 1-\phi_{i,t} \\ 0 & \phi_{i,t}\alpha & \phi_{i,t}(1-\alpha)(1-\gamma) & \phi_{i,t}(1-\alpha)\gamma & 1-\phi_{i,t} \\ 0 & \phi_{i,t}\alpha & \phi_{i,t}(1-\alpha)(1-\gamma) & \phi_{i,t}(1-\alpha)\gamma & 1-\phi_{i,t} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

where $\phi_{i,t}$ represents the apparent survival probability, $\rho_{i,t}$ represents the recruitment probability, α represents the probability of being at sea, and γ represents the probability of breeding, conditional on being in the colony. Age-specific apparent annual survival probabilities were modeled as:

$$logit(\phi_{i,t}) = \mu^{\phi}_{a_{i,t}} + \epsilon^{\phi}_{t} + \beta^{\phi}_{a'_{i,t}} X_t$$

where μ^{ϕ} represents the mean survival probability and is equal to μ^{ϕ}_{juv} when age class $a'_{i,t} = 1$ and μ^{ϕ}_{ad} otherwise, ϵ^{ϕ}_{t} is an annual random effect drawn from a normal distribution with mean 0 and standard deviation σ^{ϕ}, and β^{ϕ} is the effect of bycatch mitigation index (X_t) that is equal to β^{ϕ}_{juv} when $a'_{i,t} = 1$ and β^{ϕ}_{ad} otherwise. Given the presence of multicollinearity amongst the fleet-specific bycatch mitigation rates (Figure 4.3), we modeled survival as a function of a linear combination of them, X_t. However, the relative influence of individual fisheries on survival was also of interest, so we modeled X_t as

$$X_t = \sum_{m=1}^{4} w_m x_{m,t}$$

where m indexes the fisheries and the vector $\boldsymbol{w}$ represents the relative influence of mitigation rates in each fishery ($\boldsymbol{x}$) on survival and was sampled via a Dirichlet process (Table 4.1). The covariate effects $\boldsymbol{\beta^{\phi}}$ were drawn from a normal distribution with mean 0 and standard deviation σ^{β} and shrinkage priors were used for both σ^{ϕ} and σ^{β}. We emphasize that $\boldsymbol{\phi}$ represents

apparent survival, given that neither permanent nor temporary breeding emigration (i.e., breeding outside of the study area) are estimable under the current monitoring program. The sigmoidal age-at-recruitment curve ($\boldsymbol{\rho}$) was modeled as

$$logit(\rho_{i,t}) = \mu^{\rho} + \beta^{\rho}(a_{i,t} - 1)$$

where $expit(\mu^{\rho})$ represents the probability of recruiting into the breeding population at age $a_{i,t} = 1$, and β^{ρ} represents the slope of recruitment probability with age, on the logit scale.

In practice, breeding states were not always recorded completely or accurately, so we modeled the observation process as

$$Y_{i,t} \sim Categorical(\mathbf{\Pi}_{Z_{i,t},t})$$

where the observed event Y for individual i at time t was conditional on the true state $Z_{i,t}$ and was categorically distributed according to observation matrix $\mathbf{\Pi}$. The events were: 1) observed breeding; 2) observed loafing, had bred before; 3) observed with uncertain breeding state; and 4) not observed. Events 1 and 2 only occurred in the last 3 years of the time series when data recording consistency was adequate; including these data is useful for estimating γ, which separates breeders from loafers. For the majority of the time-series, the only events in the detection histories are 3 and 4. Moreover, some states were unobservable (i.e., immature, at sea, dead). Thus, individuals were only observed when their true states were either loafing or breeding, but whether or not these events were recorded varied over the time series. We assume that annual detection probabilities are 1 provided that an individual is in the study area, that at-sea adults and dead individuals are not observable, and that immature birds are first observable in the year they recruit.

4.2.4.3 Count Model

Annual counts of breeders, C_t, were modeled as

$$C_t \sim Poisson(2 \cdot N_t^{breed})$$

where N_t^{breed} represents the number of breeding females within the study area, which are multiplied by 2 as counts include both males and females. Because individuals are not always identified to sex during population counts, we assume that there is an equal sex ratio within the population, for simplicity. We also assume that both under- and over-counts are possible. A Poisson distribution is used under the assumption that annual variation in counts was large, though it is likely that data are underdispersed. We explored an alternative model using a Normal distribution (i.e., $C_t \sim Normal(2 \cdot N_t^{breed}, \sigma^{obs})$ with observation error σ^{obs}estimated as an additional parameter. Preliminary results yielded similar estimates of abundance, but MCMC mixing was poor. Therefore, we chose the Poisson model.

4.2.4.4 Population Process

Age- and state-specific abundances were estimated for immature and adult birds, respectively, as stochastic functions of estimated prior abundance and the demographic rates described above. As this is a female-only model, all estimates of abundance refer to females.

The annual abundance of juveniles (i.e., fledglings) was modeled as

$$N_t^{juv} \sim Poisson(\frac{1}{2} f \cdot N_t^{rebreed} + \frac{1}{2} f \cdot N_t^{recruit})$$

where $N_t^{rebreed}$ represents adult females who have bred at least once, and $N_t^{recruit}$ represents females who are breeding for the first time. Assuming an equal sex ratio at fledging, the productivity of females is $\frac{1}{2} f$. For years 2009 (i.e., one year after the model start) to the end of the time series, the number of immature (i.e., pre-breeding) females that are one year old was modeled as

$$IM_{1,t+1} \sim Binomial(N_t^{juv}, \phi_{juv,t})$$

and number of surviving immature females that are ages 1 and greater was modeled as

$$IM'_{a+1,t+1} \sim Binomial\left(IM_{a,t}, \phi_{ad,t}\right)$$

while the number of recruits of age a was

$$N_{a,t}^{recruit} \sim Binomial\left(IM'_{a,t}, \rho_a\right)$$

and the individuals that do not recruit was thus

$$IM_{a,t} = IM'_{a,t} - N_{a,t}^{recruit}.$$

For years 2009 to the end of the time series, the number of surviving adult (i.e., previously recruited) females was modeled as

$$N_{t+1}^{adult} \sim Binomial(N_t^{breed} + N_t^{sea} + N_t^{loaf}, \phi_{ad,t})$$

where the adults are partitioned into three states: 1) breeders; 2) individuals at sea; and 3) loafers. The number of individuals at sea was modeled as

$$N_t^{sea} \sim Binomial\left(N_t^{adult}, \alpha\right).$$

Of the surviving individuals who return to the colony and have bred before, the number of those who make a breeding attempt was modeled as

$$N_t^{rebreed} \sim Binomial(N_t^{adult} - N_t^{sea}, \gamma)$$

and all remaining adults are presumed to be loafers. That is,

$$N_t^{loaf} = N_t^{adult} - N_t^{sea} - N_t^{rebreed}.$$

In any year, the total number of breeders is the number of established breeders, plus recruits of all ages. That is,

$$N_t^{breed} = N_t^{rebreed} + \sum_a N_{a,t}^{recruit}.$$

We consider the total population size in any given year to be the total number of individuals that are at least one year of age (i.e., omitting N_t^{juv}) such that

$$N_t^{tot} = N_t^{breed} + N_t^{sea} + N_t^{loaf} + \sum_a IM_{a,t}.$$

4.2.5 *Model Fitting*

The model was fitted in NIMBLE (de Valpine et al., 2017) accessed from program R (R Core Team 2022). Within NIMBLE, we used a custom distribution modified from the nimbleEcology package (Turek et al., 2016) to avoid directly sampling the latent state histories $\boldsymbol{Z}$. We used discrete filtering of the multi-event likelihood to remove the latent states, and we reduced the dataset to only include unique detection histories, weighting the likelihood by the frequency with which each detection history appears in the full dataset. In combination, these changes decreased a prohibitively long runtime by over an order of magnitude.

Informative priors were used for some parameters, drawn from the literature on AYNA where available (Table 4.1). Informative priors were also used for the initial population size, where the initial age- and stage-specific abundances were based both on previous literature and monitoring data prior to the model start. In particular, age- and state-specific initial population sizes were drawn from Poisson distributions with rates equal to the age- and state-specific counts from years prior to 2008 (i.e., the start of the count surveys) multiplied by mean prior probabilities for age-class-specific survival, colony return, and breeding propensity probabilities.

Models were fit with 3 chains of 100,000 iterations with an adaptation interval of 200. The first 50,000 samples were discarded as burn-in, and the remaining samples were thinned at a rate of 10 to reduce storage requirements. We evaluated model convergence based on visual inspection of traceplots and ensuring that the Gelman-Rubin statistic ($\hat{R}$; Gelman & Rubin, 1992) was less than 1.1 for all parameters.

4.2.6 *Population Sensitivity and Viability*

We examined the sensitivity of population growth to annual fecundity and juvenile and adult survival by computing the correlation coefficient (r) between the posteriors for annual growth rate, $\lambda_t = \frac{N_{t+1}^{tot}}{N_t^{tot}}$, and the posteriors for the demographic rates with annual variation.

To examine future viability, we projected the population 50 years into the future under current levels of bycatch mitigation (i.e., status quo). We simulated 10,000 trajectories using the model structure described above along with the estimated posterior distributions for key demographic rates and abundances. We summarized predictions to the average population growth rates, probability of extinction, and predicted future abundance.

4.3 RESULTS

4.3.1 *Demographic Rates*

Posterior mean fecundity ($\widehat{\mu^f}$) was 0.501 (95% credible interval: 0.452, 0.551) (Figure 4.5). There was some interannual variability, with the posterior mean variance of the annual random effect ($\widehat{\sigma^f}$) equal to 0.335 (0.228, 0.503). The sigmoidal age-at-recruitment curve had intercept $\widehat{\mu^\rho}$ = -4.357 (-4.511, -4.206) and slope $\widehat{\beta^\rho}$ = 1.033 (0.941, 1.136) (Figure 4.6). Once individuals recruited into the breeding population, the probability of returning to the colony, and conditional on returning, the probability of making a breeding attempt, are both high. That is, $\hat{\alpha}$ = 0.259 (0.248, 0.271) and $\hat{\gamma}$= 0.804 (0.77, 0.836). Posterior mean apparent juvenile survival ($\widehat{\mu_{juv}^{\phi}}$) was 0.881 (0.847, 0.909). Adult survival ($\widehat{\mu_{ad}^{\phi}}$) was estimated with less uncertainty, with posterior mean 0.923 (0.908, 0.935). Survival was markedly lower in some years, with recent examples being 2010, 2014, and 2018 (Figure 4.4). The standard deviation of the annual random effect

($\widehat{\sigma^{\phi}}$) equal to 0.662 (0.493, 0.907), with annual deviates being larger since 2009, relative to the rest of the time series (i.e., 1985-2008).

We did not detect an effect of region-wide bycatch mitigation on annual survival probability for either age class, nor did we detect a meaningful difference in the effect of bycatch mitigation in any individual fishery. That is, the posterior mean for the effect of bycatch mitigation ($\widehat{\beta^{\phi}}$) was 0.001 (-0.25, 0.296) for juveniles and 0 (-0.237, 0.229) for adults. The posterior probability that the effect on survival was positive was 0.513 for juveniles and 0.494 for adults. The posterior weights for the relative impact of each fishery ($\widehat{\boldsymbol{w}}$) were all roughly equal, with mean 0.25 (0.074, 0.466).

4.3.2 *Abundance and Population Viability*

During the data period (i.e., 2008-2020), the average population growth rate ($\hat{\lambda}$) was 0.99 (0.983, 0.998). The decline was more severe between 2008 and 2014, with the population declining at a rate of 0.957 (0.944, 0.97) (Figure 4.7). Since 2014, the population has been increasing at a rate of 1.025 (1.009, 1.041), but as of 2020 it had not reached the estimated abundance in 2008. Over the same period the breeding population was estimated to be stable, with an average growth rate of 1.001 (0.994, 1.008). In 2020, the posterior means for the total number of females associated with the study area (i.e., loafers, breeders, and females at sea), and the number of females observed breeding in the study area were 700 (646, 759) and 309 (289, 329), respectively.

In 50-year population projections under the status-quo scenario (i.e., no further increases to bycatch mitigation measures in the region past 2020) the population increased on average, with mean growth rate 1.007 (0.957, 1.044) (Figure 4.9). The probability of extinction under this scenario was 0. However, the uncertainty on estimates is large, with credible intervals for abundance at the end of the projection period of 969 (76, 6017). In particular, the upper bounds

should be interpreted with caution, as density-dependent effects would likely mediate the true carrying capacity within the study area.

Of the demographic rates, fecundity had the strongest correlation with $\hat{\lambda}$, with $r = 0.17$ (-0.179, 0.785). Neither survival probability had a meaningful correlation with population growth rate with $r = 0.021$, (-0.54, 0.57) and $r = 0.02$ (-0.54, 0.56) for adults and juveniles, respectively (Figure 4.8).

4.4 DISCUSSION

We developed an IPM to improve understanding of AYNA population dynamics and estimate the effects of bycatch mitigation rates in four key South Atlantic fisheries on survival rates. This analysis represents an updated population assessment and is the first to show the population is stable to recovering, with a low probability of extinction in the next 50 years. However, the degree of uncertainty in our predictions are largely driven by variation in data quality prior to 2008, underscoring the need for continued study of this species. This analysis also constitutes one of the first attempts to demonstrate a population-level effect of bycatch mitigation on seabird survival. While we did not detect a strong effect of bycatch mitigation on AYNA survival, our ability to detect an effect was sharply limited by the quality of data on bycatch and bycatch mitigation. Our results therefore underscore the importance of continued research into AYNA population dynamics and drivers, including additional study of bycatch and bycatch mitigation in the South Atlantic.

We emphasize that our results do not rule out an effect of bycatch or bycatch mitigation on AYNA survival. Rather, population and fisheries data are currently insufficient to detect an effect. In practice, mitigation rates are unknown, but likely are quite low on the high seas where most mortalities likely occur (Dias et al., 2017; Michael et al., 2021; Paterson et al., 2019).

National fisheries tend to have stronger regulations around use of mitigation measures (da Rocha et al., 2021), though compliance rates are unknown. Moreover, current observer coverage is sparse in the South Atlantic and when observers are present, bycaught seabirds are generally not identified to species (Zhou et al., 2019a, 2019b). Instead, most existing estimates of bycatch rates are reported as number of birds per unit effort (e.g., number of hooks in a longline fishery, or set hours in a trawl fishery). Because rates are not species-specific and historic effort data are not available for all fisheries in the South Atlantic, it is difficult to estimate the total number of AYNA that are bycaught annually. Despite this, there is substantial evidence at the vessel-level that AYNA are frequently bycaught in the absence of mitigation, and that mitigation can nearly eliminate bycatch (da Rocha et al., 2021). Presumably, there should be a negative effect of fishing effort on AYNA survival rates and this effect should be mediated by introducing mitigation measures throughout the fleet. In this analysis, we used expert-elicited estimates of fleet-specific bycatch mitigation rates for key fisheries in the South Atlantic and these estimates may not accurately represent trends in bycatch mitigation, and they do not capture the relative influence of each fishery.

It is possible that some aspects of AYNA life history and fishing trends are masking the effects of bycatch on survival. One possibility is that the fishing effort is lower on vessels that use mitigation measures relative to vessels that do not, in which case the proportion of fleets using mitigation measures would not be an appropriate metric for estimating the effect of mitigation on AYNA survival. Alternatively, true fishing effort could be higher than expected, and if effort increases were outpacing mitigation increases, any effect could be masked. Ideally, fishery-specific effort data would be available, but in this case they are not available for all fisheries of interest over the time series. A potential weakness of our model is that we assume

that the dynamics of both sexes are the same and that females are representative of the entire population. This is not always the case, and previous work has demonstrated the long-term impacts of sex-biased mortality (Mills & Ryan, 2005). These impacts can be difficult to detect before the biased sex-ratio is extreme. We also assumed that individuals survive at the adult rate from age 2 onward. However, it is likely that survival gradually increases during immaturity, and if younger birds are more susceptible to bycatch, the changes to the underlying age-structure can be difficult to detect but yield large impacts on long-term population trend (Oppel et al., 2022). Finally, it is possible that the Gough Island colony is not representative of AYNA population and bycatch trends globally. Gough is the second largest breeding colony and has been most consistently monitored, but is much smaller than the colony on Tristan da Cunha, hosting approximately 10% of the global AYNA population. Previous estimates of population trends have varied throughout the region, with steeper population declines observed on Tristan da Cunha and Inaccessible Island than on Gough (Cuthbert et al., 2003).

An additional challenge in estimating the effects of bycatch on survival is uncertainty surrounding the survival rates themselves. Though AYNA have been monitored on Gough since 1982, the reliability of some aspects of the data (e.g., breeding state of marked birds), has improved dramatically since 2019. Prior to then, we modeled only where individuals were or were not detected, and because transitions between states were not directly observed, the parameters governing transitions (i.e., age-specific survival, age-specific recruitment, probability of remaining at sea, and breeding probability) are more challenging to estimate. In this analysis, we used a simplified model of the true biological process with relatively few parameters to ensure parameter identifiability. Simplifying assumptions included perfect detection of individuals within the study area, and that temporary breeding emigration was negligible for both

age classes. Both assumptions are likely satisfied for adults, where breeding counts are considered near censuses and fidelity is near 1 (Cuthbert & Sommer, 2004). These assumptions may be less appropriate for prebreeders as they exhibit lower rates of fidelity (approximately 80%; Cuthbert et al., 2003). Moreover, due to limitations with the data, we were not able to subcategorize loafers as prebreeders or experienced breeders. We therefore make the simplifying assumption that individuals are breeding in the first year they are resighted. Consequently, survival may be slightly underestimated for both age classes, and the age-at-recruitment curve may skew young.

In 2003, an analysis was conducted to predict the power of the current monitoring protocol to detect different rates of population decline and this study found that a 1% annual decline would take over 30 years to detect with 80% power (Cuthbert et al. 2003). Here, population trend estimates were based on only 12 years of data. Provided that improvements to the data (i.e., consistently recording breeding state) and monitoring protocol (i.e., searching for marked birds in a "buffer zone" surrounding the primary study area to account for movement out of the study area; Kendall et al., 2009) continue, future analyses may produce more precise estimates of parameters, bolstering our ability to identify population trends.

Our results deviate from previous analyses of AYNA demography for a few key parameters. The most recent population assessment was in 2003, at which time the population was reportedly trending downwards by approximately 1.2% per year, with breeders declining more rapidly at 2.3% per year (Cuthbert et al. 2003). In contrast, we estimated an initial population decline, with the negative trend reversing after 2014 indicating a relatively stable breeding population. Our estimate of mean fecundity was substantially lower (~ 50% vs. ~ 67%), though our estimate of breeding probability was higher (~80% vs ~66%), which is more

consistent with breeding rates of the other small albatrosses known as mollymawks (*Thalassarche spp.*; Converse et al., 2009; Kendall et al., 2009; Ponchon et al., 2019; Ryan et al., 2007). Our estimates of mean juvenile and adult survival were nearly identical (0.88 and 0.92, respectively) to previous estimates. Cuthbert et al. (2003) speculated, based on comparisons with congener species, that population decreases were most likely caused by low adult and juvenile survival, but we found that population growth rates were more strongly correlated with annual fecundity than survival, which may suggest a need for additional conservation action at the breeding colony in addition to bycatch mitigation at-sea.

We did not address the potential impact of another cryptic threat to AYNA: nest predation by introduced mice. In response to research demonstrating large effects of mouse predation on Gough's nesting seabirds (Caravaggi et al., 2019; Cuthbert & Hilton, 2004; Jones et al., 2021; Oppel et al., 2022; Wanless et al., 2007, 2009), an eradication was attempted in 2021 but failed (Samaniego et al., 2022). While a smaller proportion of AYNA nests have been predated than other species on the island (e.g., Tristan Albatross; Oppel et al., 2022), it is possible that mouse predation is having a greater impact on AYNA than previously thought. This is a key uncertainty that should be resolved as the mouse population approaches pre-eradication levels.

The results from this analysis will support future improvements to AYNA population monitoring as well as bycatch tracking and mitigation regulations in the South Atlantic. Though uncertainty remains, these results fill a need for updated estimates of AYNA vital rates, abundance, and population trend. In particular, our results identify key areas of remaining uncertainty and demonstrate a need for continued and consistent population monitoring and better data on both fishing effort, seabird bycatch, and mitigation rates in the South Atlantic. The

model we introduce can be used in the future to reevaluate recovery for this species, and this framework can be applied more broadly for difficult-to-monitor species vulnerable to cryptic threats.

4.5 ACKNOWLEDGEMENTS

We thank numerous avian field technicians for collecting field data and staff at the Royal Society for the Protection of Birds for developing a database. This research was funded through support of A. Bratt by the Washington Cooperative Fish and Wildlife Research Unit and the Royal Society for the Protection of Birds. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

4.6 FIGURES & TABLES

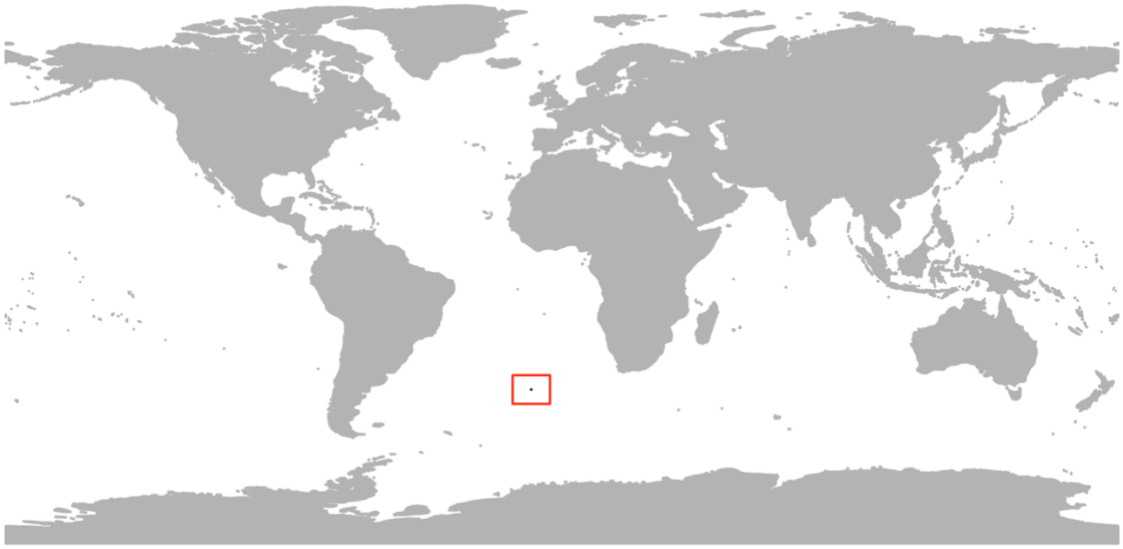

Figure 4.1: Location of Gough Island which is centrally located in the South Atlantic Ocean, surrounded by the red box. Atlantic Yellow-nosed Albatross (AYNA) foraging ranges span the Southern Atlantic between South America and Africa. On Gough, the AYNA study area is located on the southern tip of the island, which is a lowland plateau.

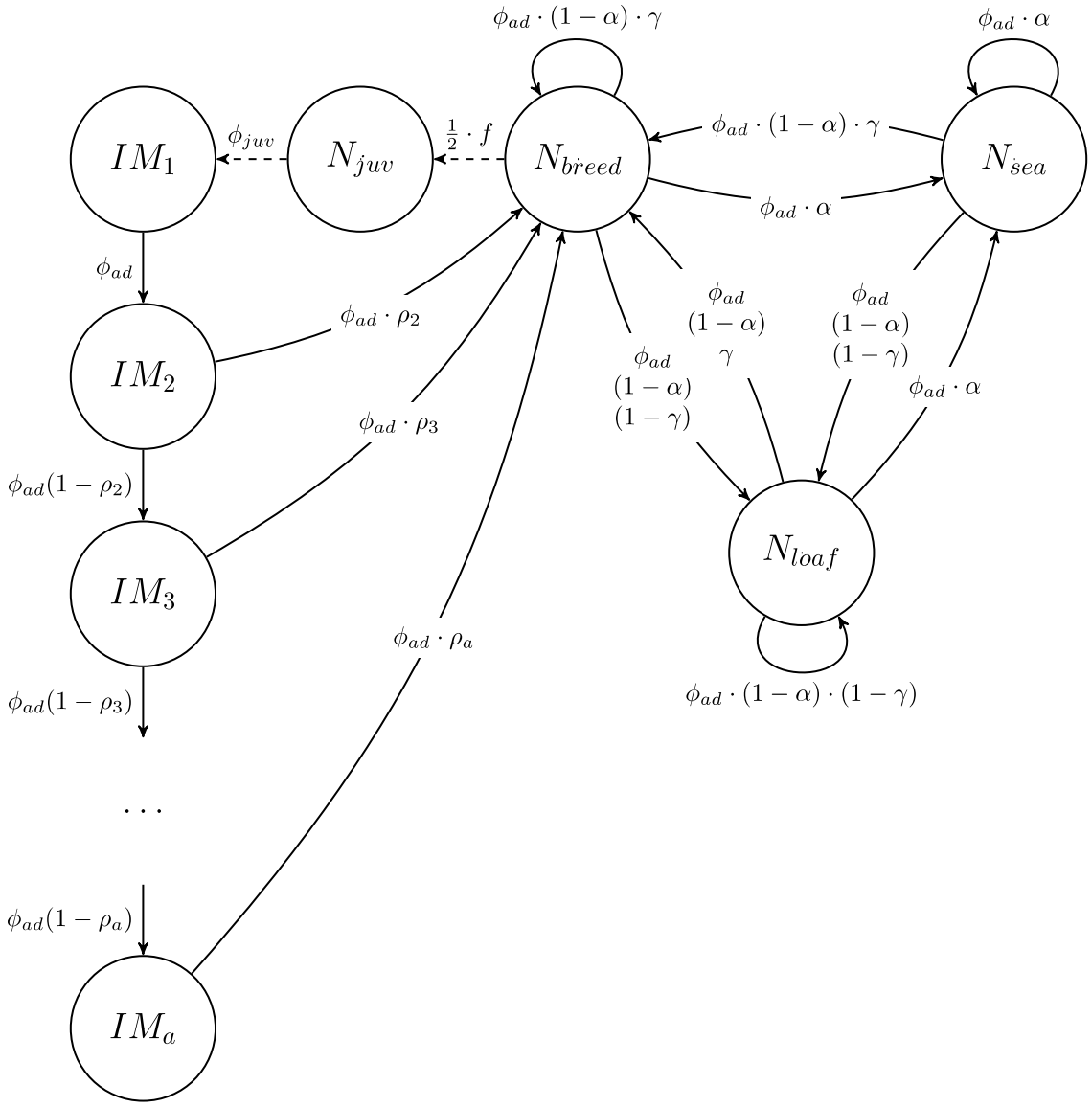

Figure 4.2: Life-cycle diagram for Atlantic Yellow-nosed Albatross as described by the population model. For adults, represented states are 1) breeding; 2) loafing within the breeding colony; and 3) at sea. Conditional on surviving, adults may cycle between these three states according to the transition probabilities described in the text. If breeding, adults produce juveniles, which may then survive to become immature individuals of age 1. Conditional on surviving, immature birds may recruit into the breeding population with an age-specific transition probability. Note that this is a female-based model and thus fecundity is halved, assuming an equal sex-ratio at fledging.

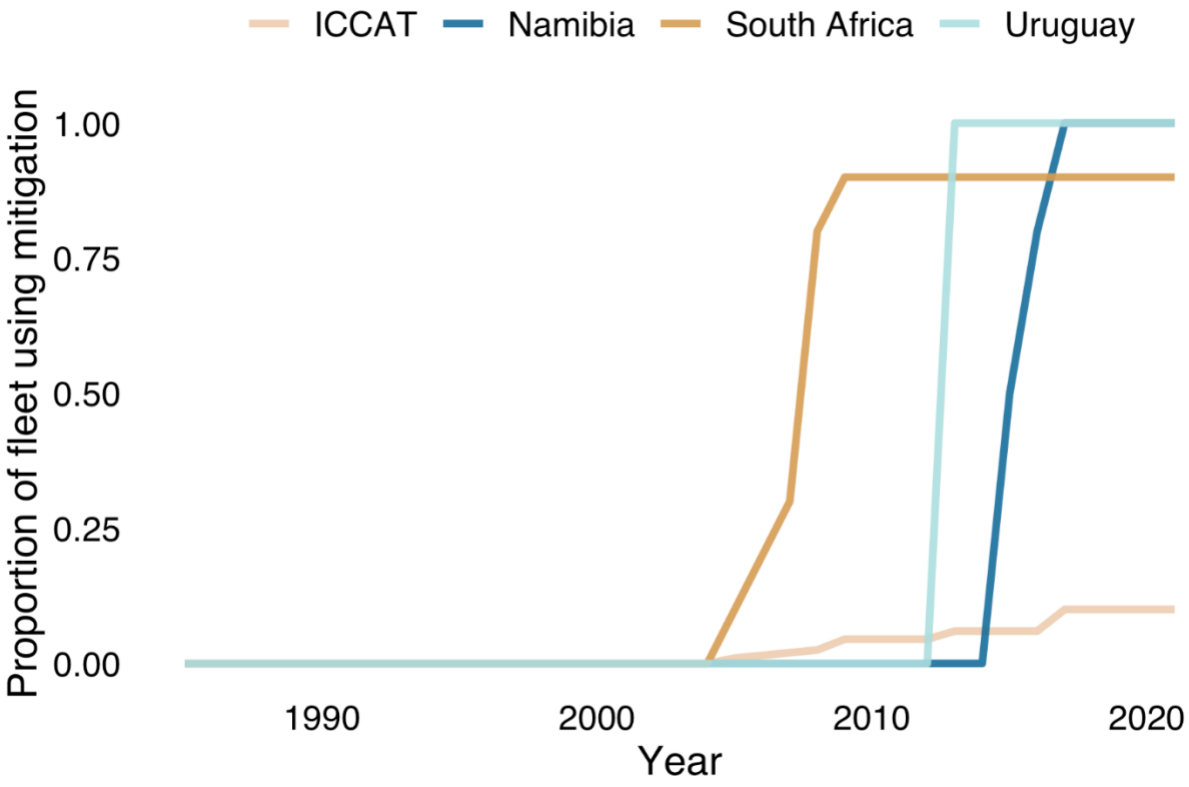

Figure 4.3: Proportion of international (International Commission for the Conservation of Atlantic Tuna; ICCAT) and national (Namibia, South Africa, and Uruguay) longline and demersal trawl fishing fleets that have implemented seabird bycatch mitigation measures by year, as elicited from experts from the Albatross Task Force. The large discontinuities amongst the national fleets reflect changes to mitigation mandates, whereas mitigation is still "opt-in" for the ICCAT fleet.

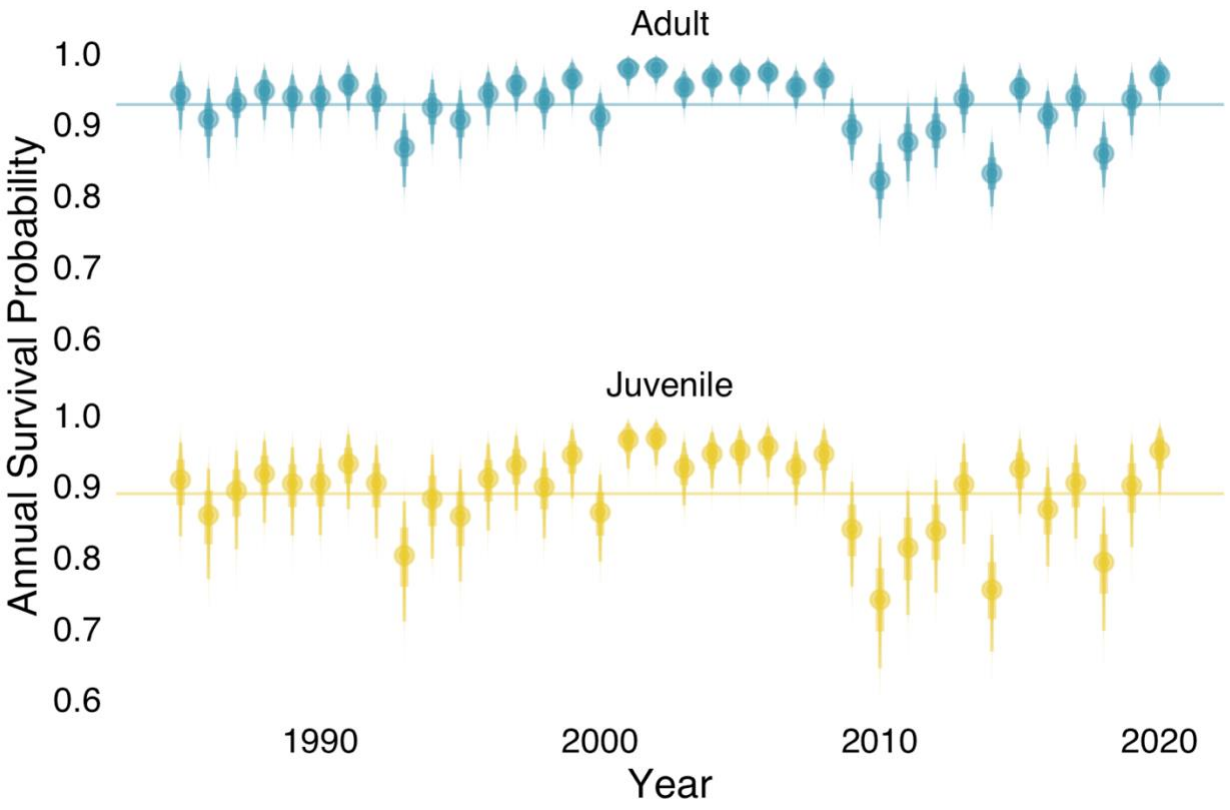

Figure 4.4: Annual survival probability for juvenile (yellow) and adult (blue) Atlantic Yellow-nosed Albatross, where any bird over 1 year old is assumed to survive at the adult rate. Posterior distributions are shown along with medians (points), and 95% credible intervals (line ranges). Horizontal lines represent age-specific means, noting that y-axis begins at 0.6.

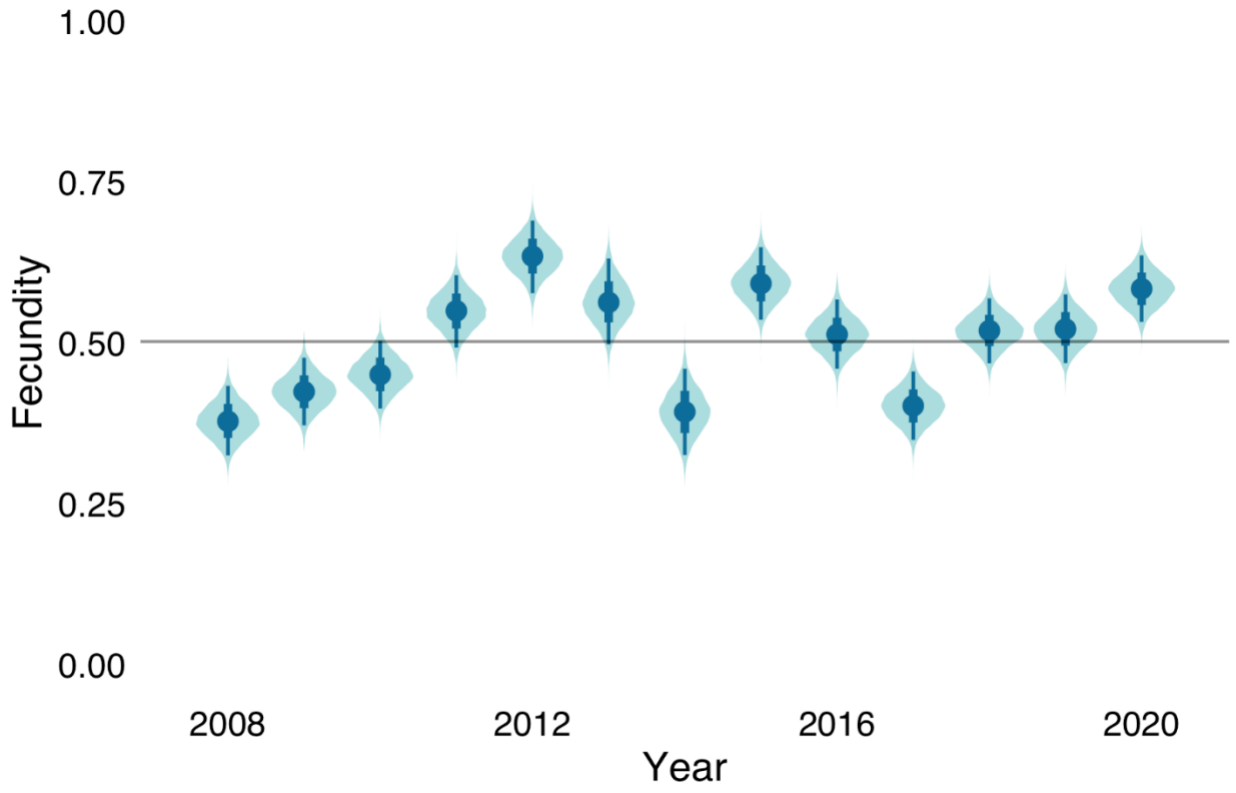

Figure 4.5: Annual fecundity estimates. Fecundity represents the probability of successfully fledging a chick of either sex. Posterior distributions are shown along with medians (points), and 95% credible intervals (line ranges). The horizontal line represents the inter-annual mean.

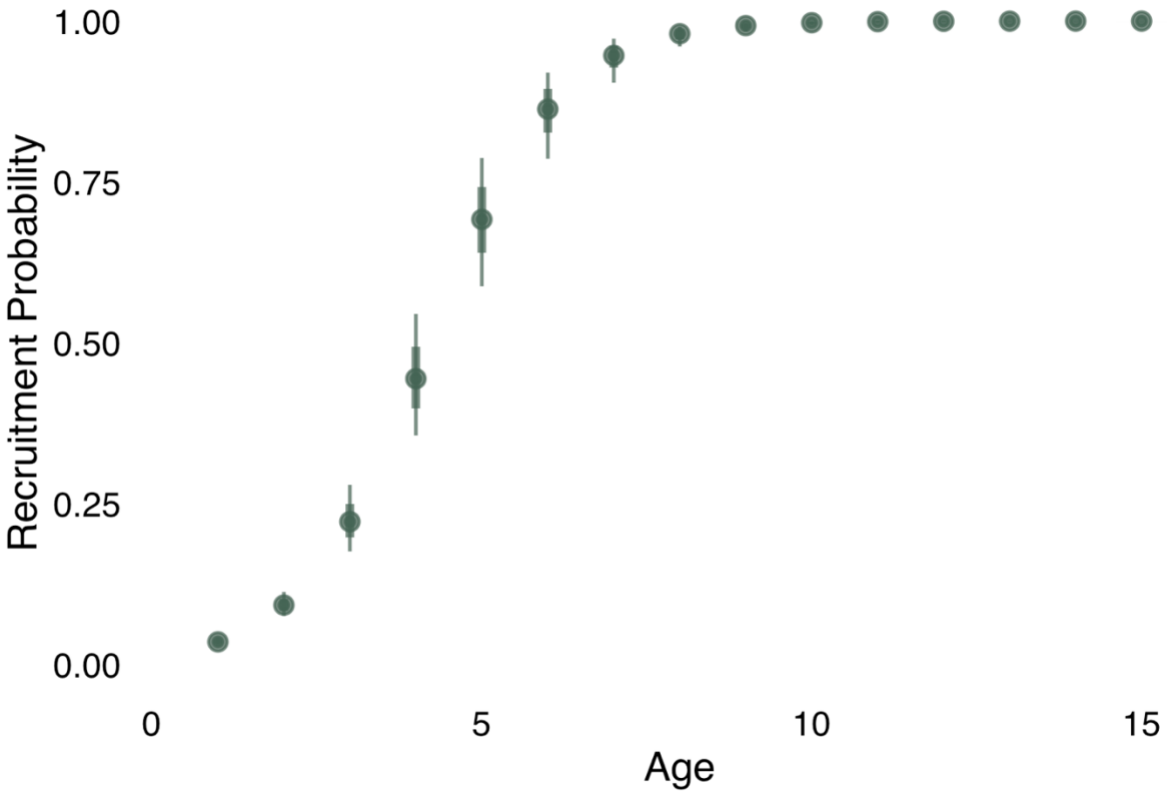

Figure 4.6: Estimated age-at-recruitment curve. The curve takes a sigmoidal shape, where average age at recruitment is around 9. Recruitment probability is low prior to this, and increases steeply as individuals age, with an asymptote of 1. Shown are age-specific medians (points), and 95% credible intervals (line ranges).

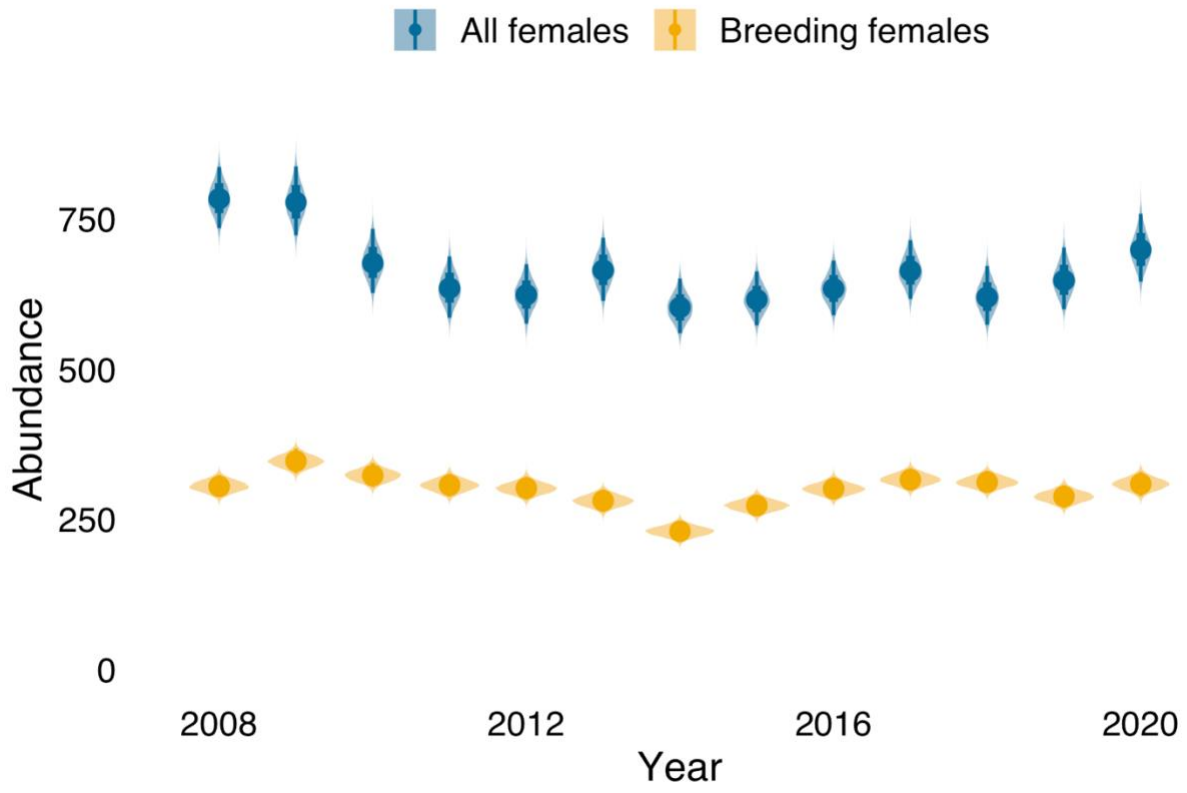

Figure 4.7: Estimated abundance of breeding (yellow) and all (blue) females associated with the Gough Island study area over the data period. Posterior distributions are shown along with medians (points), and 95% credible intervals (line ranges). Despite some interannual variation, both groups appear relatively stable over this period. There is more variability present amongst all females, which includes loafers, adults at sea, and immature birds in addition to breeding individuals. There is some asynchrony between breeding and total abundance.

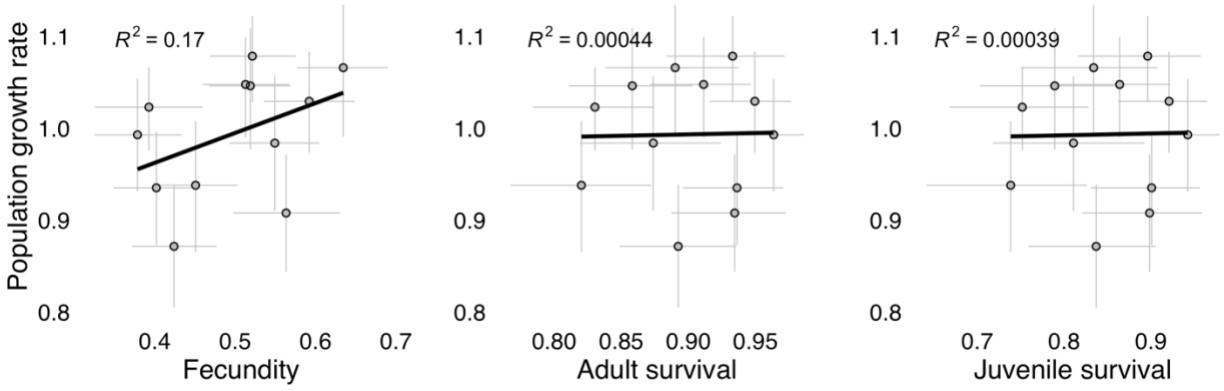

Figure 4.8: Estimated annual population growth rates relative to estimated demographic rates over the data period. Posterior medians are shown (points), with 95% credible intervals (line ranges) for fecundity, adult survival, and juvenile survival.

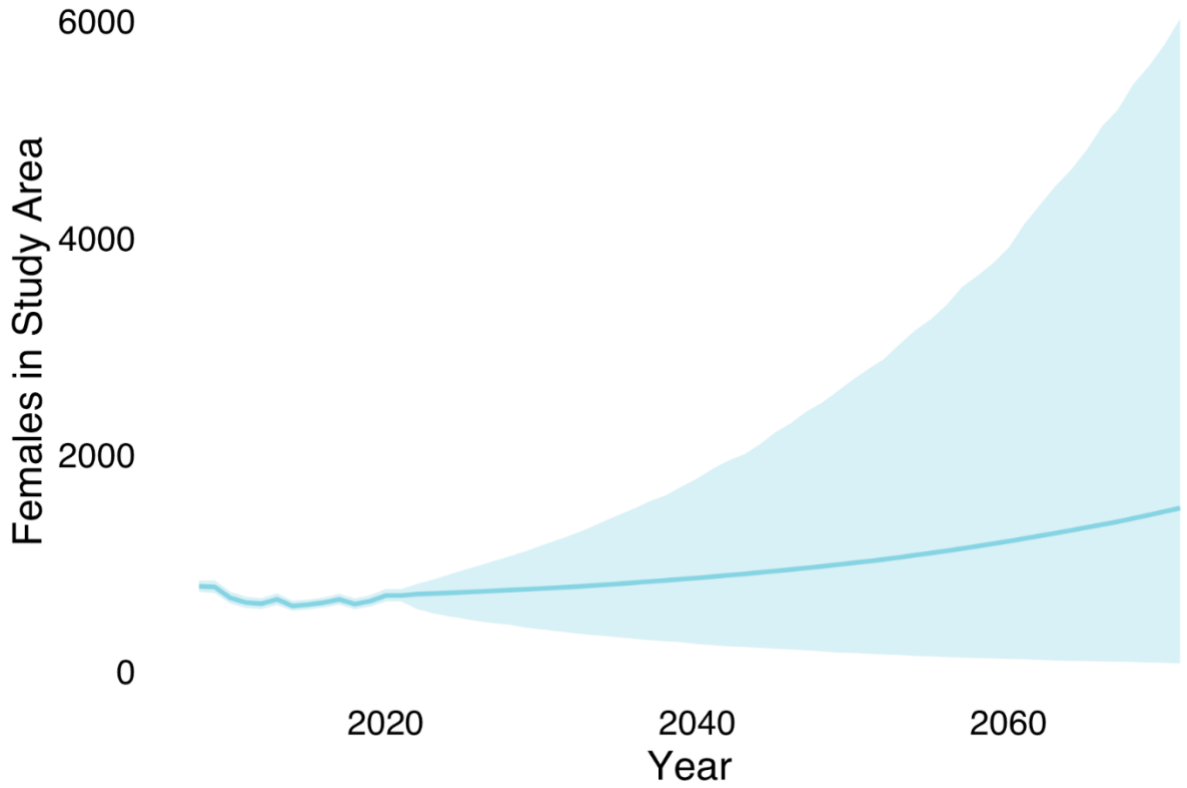

Figure 4.9: Projected total abundance of females within the Gough Island study area. Medians are represented by the bold lines, while 95% credible intervals are represented by the shaded areas. Projections are 50 years long, with blue representing the data period prior to 2020 and the "status quo" projection.

Table 4.1: Model parameters and their priors, with references when informative priors were used.

Parameter	Description	Prior	Prior reference
μ^f	Mean fecundity	Beta(1,1)	--
σ^f	Scale of annual fecundity effect	Exp(1)	Simpson et al. 2017
α	Probability of breeder remaining at sea	Beta(1,1)	--
γ	Probability of breeding, conditional on returning to colony	Beta(1,1)	--
μ^ρ	Probability of recruitment at age 1 (logit scale)	Normal(-6, 1)	Cuthbert et al. 2003
β^ρ	Effect of age on recruitment probability	Normal(0.75, 0.25)	Cuthbert et al. 2003
μ^ϕ_{juv}	Mean juvenile survival probability	Beta(20,5)	Cuthbert et al. 2003
μ^ϕ_{ad}	Mean adult survival probability	Beta(60,5)	Cuthbert et al. 2003
σ^ϕ	Scale of annual survival effect	Exp(1)	--
$\boldsymbol{\delta}$	Scale of variance for relative weights	Poisson(10)	--
$\boldsymbol{w}$	Relative weight of effective of mitigation in each fishery	Dirichlet($\boldsymbol{\delta}$)	--
σ^β	Shrinkage for mitigation effect	Exp(1)	Simpson et al. 2017
p	Probability of prebreeder returning to colony	Normal(0,2)	--

4.7 REFERENCES

Abadi, F., Gimenez, O., Ullrich, B., Arlettaz, R., & Schaub, M. (2010). Estimation of immigration rate using integrated population models. *Journal of Applied Ecology*, *47*(2), 393–400.

ACAP. Agreement on the Conservation of Albatrosses and Petrels. http://www.acap.aq.

Besbeas, P., Freeman, S. N., Morgan, B. J. T., & Catchpole, E. A. (2002). Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. *Biometrics*, *58*(3), 540–547. https://doi.org/10.1111/j.0006-341X.2002.00540.x

Brooks, S. P., King, R., & Morgan, B. J. T. (2004). A Bayesian approach to combining animal abundance and demographic data. *Animal Biodiversity and Conservation*, 16.

Bugoni, L., Mancini, P. L., Monteiro, D. S., Nascimento, L., & Neves, T. S. (2008). Seabird bycatch in the Brazilian pelagic longline fishery and a review of capture rates in the southwestern Atlantic Ocean | Bycatch Management Information System (BMIS). *Http://Www.Int-Res.Com/Articles/Suppl/N005p137_app.Pdf*. https://doi.org/10.3354/esr00115

Caravaggi, A., Cuthbert, R. J., Ryan, P. G., Cooper, J., & Bond, A. L. (2019). The impacts of introduced House Mice on the breeding success of nesting seabirds on Gough Island. *Ibis*, *161*(3), 648–661. https://doi.org/10.1111/ibi.12664

Converse, S. J., Kendall, W. L., Doherty, P. F., & Ryan, P. G. (2009). Multistate Models for Estimation of Survival and Reproduction in the Grey-headed Albatross (*Thalassarche chrysostoma*). *The Auk*, *126*(1), 77–88. https://doi.org/10.1525/auk.2009.07189

Croxall, J. P., Butchart, S. H. M., Lascelles, B., Stattersfield, A. J., Sullivan, B., Symes, A., & Taylor, P. (2012). Seabird conservation status, threats and priority actions: A global

assessment. *Bird Conservation International*, *22*(01), 1–34. https://doi.org/10.1017/S0959270912000020

Cuthbert, R., & Hilton, G. (2004). Introduced house mice Mus musculus: A significant predator of threatened and endemic birds on Gough Island, South Atlantic Ocean? *Biological Conservation*, *117*(5), 483–489. https://doi.org/10.1016/j.biocon.2003.08.007

Cuthbert, R. J., Louw, H., Parker, G., Rexer-Huber, K., & Visser, P. (2013). Observations of mice predation on dark-mantled sooty albatross and Atlantic yellow-nosed albatross chicks at Gough Island. *Antarctic Science*, *25*(6), 763–766. https://doi.org/10.1017/S0954102013000126

Cuthbert, R., Ryan, P. G., Cooper, J., & Hilton, G. (2003). Demography and Population Trends of the Atlantic Yellow-Nosed Albatross. *The Condor*, *105*(3), 439–452. https://doi.org/10.1093/condor/105.3.439

Cuthbert, R., & Sommer, E. (2004). *Gough Island Bird Monitoring Manual Research Report*. 54.

Cuthbert, R., Sommer, E., Ryan, P., Cooper, J., & Hilton, G. (2004). Demography and conservation of the Tristan albatross Diomedea [exulans] dabbenena. *Biological Conservation*, *117*(5), 471–481. https://doi.org/10.1016/j.biocon.2003.08.006

da Rocha, N., Oppel, S., Prince, S., Matjila, S., Shaanika, T. M., Naomab, C., Yates, O., Paterson, J. R. B., Shimooshili, K., Frans, E., Kashava, S., & Crawford, R. (2021). Reduction in seabird mortality in Namibian fisheries following the introduction of bycatch regulation. *Biological Conservation*, *253*, 108915. https://doi.org/10.1016/j.biocon.2020.108915

Davies, D., Dilley, B., Bond, A., Cuthbert, R., & Ryan, P. (2015). Trends and tactics of mouse predation on Tristan Albatross Diomedea dabbenena chicks at Gough Island,

South Atlantic Ocean. *Avian Conservation and Ecology*, *10*(1). https://doi.org/10.5751/ACE-00738-100105

de Valpine, P., Turek, D., Paciorek, C. J., Anderson-Bergman, C., Lang, D. T., & Bodik, R. (2017). Programming with models: Writing statistical algorithms for general model structures with nimble. *Journal of Computational and Graphical Statistics*, *26*(2), 403–413. https://doi.org/10.1080/10618600.2016.1172487

Dias, M. P., Martin, R., Pearmain, E. J., Burfield, I. J., Small, C., Phillips, R. A., Yates, O., Lascelles, B., Borboroglu, P. G., & Croxall, J. P. (2019). Threats to seabirds: A global assessment. *Biological Conservation*, *237*, 525–537. https://doi.org/10.1016/j.biocon.2019.06.033

Dias, M. P., Oppel, S., Bond, A. L., Carneiro, A. P. B., Cuthbert, R. J., González-Solís, J., Wanless, R. M., Glass, T., Lascelles, B., Small, C., Phillips, R. A., & Ryan, P. G. (2017). Using globally threatened pelagic birds to identify priority sites for marine conservation in the South Atlantic Ocean. *Biological Conservation*, *211*, 76–84. https://doi.org/10.1016/j.biocon.2017.05.009

Gilman, E., Musyl, M., Wild, M., Rong, H., & Chaloupka, M. (2022). Investigating weighted fishing hooks for seabird bycatch mitigation. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-06875-4

Holmes, N. D., Spatz, D. R., Oppel, S., Tershy, B., Croll, D. A., Keitt, B., Genovesi, P., Burfield, I. J., Will, D. J., Bond, A. L., Wegmann, A., Aguirre-Muñoz, A., Raine, A. F., Knapp, C. R., Hung, C.-H., Wingate, D., Hagen, E., Méndez-Sánchez, F., Rocamora, G., … Butchart, S. H. M. (2019). Globally important islands where eradicating invasive mammals will benefit

highly threatened vertebrates. *PLoS ONE*, *14*(3), e0212128. https://doi.org/10.1371/journal.pone.0212128

IUCN. 2022. *The IUCN Red List of Threatened Species. Version 2022-2*. https://www.iucnredlist.org.

Jones, C. W., Risi, M. M., Cleeland, J., & Ryan, P. G. (2019). First evidence of mouse attacks on adult albatrosses and petrels breeding on sub-Antarctic Marion and Gough Islands. *Polar Biology*, *42*(3), 619–623. https://doi.org/10.1007/s00300-018-02444-6

Jones, C. W., Risi, M. M., Osborne, A. M., Ryan, P. G., & Oppel, S. (2021). Mouse eradication is required to prevent local extinction of an endangered seabird on an oceanic island. *Animal Conservation*, *24*(4), 637–645. https://doi.org/10.1111/acv.12670

Kendall, W. L., Converse, S. J., Doherty, P. F., Naughton, M. B., Anders, A., Hines, J. E., & Flint, E. (2009). Sampling design considerations for demographic studies: A case of colonial seabirds. *Ecological Applications*, *19*(1), 55–68. https://doi.org/10.1890/07-1072.1

Kroodsma, D., Turner, J., Luck, C., Hochberg, T., Miller, N., Augustyn, P., & Prince, S. (2023). Global prevalence of setting longlines at dawn highlights bycatch risk for threatened albatross. *Biological Conservation*, *283*, 110026. https://doi.org/10.1016/j.biocon.2023.110026

Lawson, A. J., Jodice, P. G. R., Rainwater, T. R., Dunham, K. D., Hart, M., Butfiloski, J. W., Wilkinson, P. M., McFadden, K. W., & Moore, C. T. (2022). Hidden in plain sight: Integrated population models to resolve partially observable latent population structure. *Ecosphere*, *13*(12), e4321. https://doi.org/10.1002/ecs2.4321

Michael, P. E., Wilcox, C., Barbraud, C., Delord, K., Sumner, M., & Weimerskirch, H. (2021). Dynamic enforcement of bycatch via reproductive value can increase theoretical efficiency. *Marine Policy*, *132*, 104684. https://doi.org/10.1016/j.marpol.2021.104684

Mills, M. S. L., & Ryan, P. G. (2005). Modelling impacts of long-line fishing: What are the effects of pair-bond disruption and sex-biased mortality on albatross fecundity? *Animal Conservation*, *8*(4), 359–367. https://doi.org/10.1017/S1367943005002386

Oppel, S., Clark, B. L., Risi, M. M., Horswill, C., Converse, S. J., Jones, C. W., Osborne, A. M., Stevens, K., Perold, V., Bond, A. L., Wanless, R. M., Cuthbert, R., Cooper, J., & Ryan, P. G. (2022). Cryptic population decrease due to invasive species predation in a long-lived seabird supports need for eradication. *Journal of Applied Ecology*, *n/a*(n/a). https://doi.org/10.1111/1365-2664.14218

Oppel, S., Hilton, G., Ratcliffe, N., Fenton, C., Daley, J., Gray, G., Vickery, J., & Gibbons, D. (2014). Assessing population viability while accounting for demographic and environmental uncertainty. *Ecology*, *95*(7), 1809–1818.

Paterson, J. R. B., Yates, O., Holtzhausen, H., Reid, T., Shimooshili, K., Yates, S., Sullivan, B. J., & Wanless, R. M. (2019). Seabird mortality in the Namibian demersal longline fishery and recommendations for best practice mitigation measures. *Oryx*, *53*(2), 300–309. https://doi.org/10.1017/S0030605317000230

Ponchon, A., Cornulier, T., Hedd, A., Granadeiro, J. P., & Catry, P. (2019). Effect of breeding performance on the distribution and activity budgets of a predominantly resident population of black-browed albatrosses. *Ecology and Evolution*, *9*(15), 8702–8713. https://doi.org/10.1002/ece3.5416

R Core Team. (2022). *R: A language and environment for statistical computing* [R Foundation for Statistical Computing].

Robertson, G., Ashworth, P., Ashworth, P., Carlyle, I., Jiménez, S., Forselledo, R., Domingo, A., & Candy, S. G. (2018). Setting baited hooks by stealth (underwater) can prevent the incidental mortality of albatrosses and petrels in pelagic longline fisheries. *Biological Conservation*, *225*, 134–143. https://doi.org/10.1016/j.biocon.2018.06.026

Roman, L., Butcher, R. G., Stewart, D., Hunter, S., Jolly, M., Kowalski, P., Hardesty, B. D., & Lenting, B. (2021). Plastic ingestion is an underestimated cause of death for southern hemisphere albatrosses. *Conservation Letters*, *14*(3), e12785. https://doi.org/10.1111/conl.12785

Rosenblatt, C. J., Gates, R. J., Matthews, S. N., Peterman, W. E., & Stricker, N. J. (2021). An integrated population model to project viability of a northern bobwhite population in Ohio. *Ecosphere*, *12*(9), e03731. https://doi.org/10.1002/ecs2.3731

Ryan, P. G., Phillips, R. A., Nel, D. C., & Wood, A. G. (2007). Breeding frequency in Grey-headed Albatrosses Thalassarche chrysostoma. *Ibis*, *149*(1), 45–52. https://doi.org/10.1111/j.1474-919X.2006.00594.x

Samaniego, A., Jolley, W., McClelland, P., Samaniego, A., Jolley, W., & McClelland, P. (2022). A lesson for planning rodent eradications: Interference of invasive slugs during the Gough Island mouse eradication attempt in 2021. *Wildlife Research*, *50*(5), 344–355. https://doi.org/10.1071/WR22024

Saunders, S. P., Cuthbert, F. J., & Zipkin, E. F. (2018). Evaluating Population Viability and Efficacy of Conservation Management Using Integrated Population Models. *Journal of Applied Ecology*, *55*(3), 1380–1392. https://doi.org/10.1111/1365-2664.13080

Schaub, M., & Abadi, F. (2011). Integrated population models: A novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology*, *152*(S1), 227–237. https://doi.org/10.1007/s10336-010-0632-7

Schaub, M., Gimenez, O., Sierro, A., & Arlettaz, R. (2007). Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. *Conservation Biology*, *21*(4), 945–955. https://doi.org/10.1111/j.1523-1739.2007.00743.x

Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, *32*(1), 1–28. https://doi.org/10.1214/16-STS576

Thomson, R. B., Alderman, R. L., Tuck, G. N., & Hobday, A. J. (2015). Effects of Climate Change and Fisheries Bycatch on Shy Albatross (Thalassarche cauta) in Southern Australia. *PLOS ONE*, *10*(6), e0127006. https://doi.org/10.1371/journal.pone.0127006

Tuck, G. N., Polacheck, T., & Bulman, C. M. (2003). Spatio-temporal trends of longline fishing effort in the Southern Ocean and implications for seabird bycatch. *Biological Conservation*, *114*(1), 1–27. https://doi.org/10.1016/S0006-3207(02)00378-6

Turek, D., de Valpine, P., & Paciorek, C. J. (2016). Efficient Markov chain Monte Carlo sampling for hierarchical hidden Markov models. *Environmental and Ecological Statistics*, *23*(4), 549–564. https://doi.org/10.1007/s10651-016-0353-z

Ventura, F., Granadeiro, J. P., Lukacs, P. M., Kuepfer, A., & Catry, P. (2021). Environmental variability directly affects the prevalence of divorce in monogamous albatrosses. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1963), 20212112. https://doi.org/10.1098/rspb.2021.2112

Véran, S., Gimenez, O., Flint, E., Kendall, W. L., Jr, P. F. D., & Lebreton, J.-D. (2007). Quantifying the impact of longline fisheries on adult survival in the black-footed albatross. *Journal of Applied Ecology*, *44*(5), 942–952. https://doi.org/10.1111/j.1365-2664.2007.01346.x

Wanless, R. M., Angel, A., Cuthbert, R. J., Hilton, G. M., & Ryan, P. G. (2007). Can predation by invasive mice drive seabird extinctions? *Biology Letters*, *3*(3), 241–244. https://doi.org/10.1098/rsbl.2007.0120

Wanless, R. M., Ryan, P. G., Altwegg, R., Angel, A., Cooper, J., Cuthbert, R., & Hilton, G. M. (2009). From both sides: Dire demographic consequences of carnivorous mice and longlining for the Critically Endangered Tristan albatrosses on Gough Island. *Biological Conservation*, *142*(8), 1710–1718. https://doi.org/10.1016/j.biocon.2009.03.008

Watkins, B. P., Petersen, S. L., & Ryan, P. G. (2008). Interactions between seabirds and deep-water hake trawl gear: An assessment of impacts in South African waters. *Animal Conservation*, *11*(4), 247–254. https://doi.org/10.1111/j.1469-1795.2008.00192.x

Yeh, Y.-M., Huang, H.-W., Dietrich, K. S., & Melvin, E. (2013). Estimates of seabird incidental catch by pelagic longline fisheries in the South Atlantic Ocean. *Animal Conservation*, *16*(2), 141–152. https://doi.org/10.1111/j.1469-1795.2012.00588.x

Zhou, C., Jiao, Y., & Browder, J. (2019). Seabird bycatch vulnerability to pelagic longline fisheries: Ecological traits matter. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *29*(8), 1324–1335. https://doi.org/10.1002/aqc.3066

Zhou, C., Jiao, Y., & Browder, J. (2019). How much do we know about seabird bycatch in pelagic longline fisheries? A simulation study on the potential bias caused by the usually

unobserved portion of seabird bycatch. *PLoS ONE*, *14*(8), e0220797. https://doi.org/10.1371/journal.pone.0220797

Zipkin, E. F., & Saunders, S. P. (2018). Synthesizing multiple data types for biological conservation using integrated population models. *Biological Conservation*, *217*, 240–250. https://doi.org/10.1016/j.biocon.2017.10.017

APPENDIX A

Model selection and evaluation (Chapter 2)

Climate-related covariates

We examined the effect of climatic conditions on SHLA at two spatial and temporal scales. First, we examined the effect of localized climate conditions at breeding sites during the breeding season on fecundity rates. Second, we examined the effect of region-wide overwinter climate conditions on annual survival rates. We assumed these climate-related covariates affected survival and productivity both directly and indirectly through several ecological mechanisms.

Localized climate conditions included the average daily temperature across the breeding season, the number of days in which the average temperature was under 10 degrees Celsius, the number of days in which the average temperature was under 5 degrees Celsius, the number of days in which the average temperature was over 30 degrees Celsius, the number of precipitation days, and the number of storm days, where a precipitation day was defined as a day with > 1mm of rainfall and a storm day was defined as a day with > 10mm of rainfall. Each of these metrics has been negatively associated with nest development and success in other Horned Larks (de Zwaan et al., 2019, 2022). We therefore hypothesized that for sites and years in which conditions are extreme (i.e., particularly hot, cold, or stormy), overall productivity is lower.

Region-wide climate conditions included the average daily temperature across the nonbreeding season, the number of days in which temperatures dipped below freezing, the number of precipitation days, and the number of storm days. Based on studies of other horned larks (e.g., de Zwaan et al., 2019, 2022), we hypothesized that colder, stormier conditions would be associated with lower annual survival.

All climate-related covariates were obtained from PRISM (PRISM 2014), accessed from the Google Earth Engine (GEE) using R package rgee (Aybar 2023). All localized climate variables were obtained at daily levels for each site, aggregated to the level of breeding season, and converted to Z-scores (by subtracting the mean and dividing by the standard deviation) for the respective time series. Because little is known about where SHLA overwinter, region-wide climate variables were obtained at daily levels for areas where we assumed larks were likely to be. In winter surveys of SHLA, Pearson et al. (2005), detected SHLA in 10 counties in Oregon and Washington state (WA: Pierce, Thurston, Grays Harbor, Pacific, Wahkiakum; OR: Clatsop, Multnomah, Polk, Linn, Benton). Approximately 70% of SHLA detected on these surveys were in Willamette Valley (i.e., Polk, Linn, and Benton counties), 20% on Columbia River Islands (i.e., Wahkiakum, Clatsop, and Multnomah counties), and 10% elsewhere in western Washington state (Pierce, Thurston, Grays Harbor, and Pacific counties). Given that a small percentage of land cover in each county represents suitable habitat for SHLA, we obtained daily climate variables only for areas in which we assumed larks could be overwintering according to land-cover type. Land-cover types within each county were obtained from MODIS, also accessed from GEE using rgee. Categories assumed to represent suitable habitat for SHLA were Croplands, Grasslands, Savannas, Open Shrublands, and Barren land cover types. Daily- and county-level climate metrics were first aggregated across counties using a weighted average, where weights were equal to the proportion of SHLA detected in each county. Then, daily region-wide metrics were aggregated to the nonbreeding season and converted to Z-scores.

Habitat-related covariates

We examined the effect of habitat characteristics on SHLA nest survival during the breeding season. In particular, we included site-level covariates, including site size and interior-to-edge

ratio, and a nest-level covariate, distance to site edge. These covariates are related to predation risk, where nests in smaller sites, sites with smaller interior to edge ratios, and nests nearer to site edge tend to have greater predation risk.

Seasonality-related covariates

We also examined the effect of seasonality on SHLA nest survival during the breeding season. In particular, we included an effect of day of season and day of season squared to account for seasonality in daily nest survival, which we hypothesized is lower at the beginning and end of the season.

Management-related covariates

Given variable monitoring and management effort throughout the region, management-related covariates of interest (i.e., buffered areas around active nests, predator control, and habitat quality) were not available for all sites or all years and therefore any potential effects were not estimable in our model.

Variable and Model Selection

In order to reduce the number of covariates used in the full IPM, where penalized complexity priors were used, we first eliminated covariates that showed limited support. To identify which covariates were well supported, we used Reversible Jump Markov Chain Monte Carlo (RJMCMC; Link & Barker, 2006) on each component of the IPM (i.e., the productivity and mark-resight models alone). Each variable had a prior probability of inclusion of 0.5 in the RJMCMC. The posterior probability distributions for each covariate's inclusion in the model are computed from the proportion of posterior samples in which the regression coefficient was equal to zero. When the 95% credible interval for the posterior probability of covariate inclusion excluded zero, the covariate was presumed to have strong support. Covariates with strong support were generally

included in the IPM, although we eliminated some covariates that were strongly correlated (e.g., number of storm days was highly correlated with number of precipitation days) and any covariates measured at a temporal resolution finer then breeding season (i.e., day of year).

Selecting covariates in this manner resulted in a full model that included no covariates. That is, no covariates showed strong support.

References

Aybar C (2023). *rgee: R Bindings for Calling the 'Earth Engine' API*. https://github.com/r-spatial/rgee/, https://r-spatial.github.io/rgee/, https://github.com/google/earthengine-api/.

de Zwaan, D. R., Camfield, A. F., MacDonald, E. C., & Martin, K. (2019). Variation in offspring development is driven more by weather and maternal condition than predation risk. *Functional Ecology*, *33*(3), 447–456. https://doi.org/10.1111/1365-2435.13273

de Zwaan, D. R., Drake, A., Camfield, A. F., MacDonald, E. C., & Martin, K. (2022). The relative influence of cross-seasonal and local weather effects on the breeding success of a migratory songbird. *Journal of Animal Ecology*, *91*(7), 1458–1470. https://doi.org/10.1111/1365-2656.13705

Link, W. A., & Barker, R. J. (2006). Model weights and the foundations of multimodel inference. *Ecology*, *87*(10), 2626–2635. https://doi.org/10.1890/0012-9658(2006)87[2626:MWATFO]2.0.CO;2

Pearson, S.F., M. Hopey, W. D. Robinson, R. Moore. 2005. Range, Abundance and Movement Patterns of Wintering Streaked Horned Larks (*Eremophila alpestris strigata*) in Oregon and Washington. Natural Areas Program Report 2005-2. Washington Dept. of Natural Resources. Olympia, WA.

PRISM Climate Group, Oregon State University, https://prism.oregonstate.edu, data created 4 Feb 2014, accessed 22 Oct 2022.

VITA

Ms. Abby Bratt is a graduate researcher in Quantitative Ecology & Resource Management Program at the University of Washington. She has conducted her graduate research in the Quantitative Conservation Lab, in the School of Environmental and Forest Sciences and the School of Aquatic and Fishery Sciences. Ms. Bratt's research focuses on innovations in and applications of Bayesian hierarchical and integrated modeling for improving our understanding of wildlife demography and population dynamics. She is particularly interested in the dynamics of small and declining populations and producing actionable science to improve their management. Prior to her graduate studies, Ms. Bratt achieved a Bachelor of Science in Statistics from the University of Washington. Born and raised in the Pacific Northwest, Ms. Bratt is committed to science that is in the public interest and is rooted in a sense of place. She is also passionate about increasing statistical literacy amongst ecologists and resource managers. When she's not at a computer or a whiteboard, she is keen to learn some new fiber art or get out in the field, whether it's for work or play.